SIM
AUD
2020

ONLINE

2020 Proceedings of the
**Symposium on Simulation for
Architecture and Urban Design**

Edited by
Angelos Chronis
Gabriel Wurzer
Wolfgang E. Lorenz
Christiane M. Herr
Ulrich Pont
Dana Cupkova
Gabriel Wainer

2020 Proceedings of the

Symposium on Simulation for Architecture and Urban Design

Edited by

Angelos Chronis
Gabriel Wurzer
Wolfgang E. Lorenz
Christiane M. Herr
Ulrich Pont
Dana Cupkova
Gabriel Wainer

Cover by
John Yee

Preface

The COVID-19 crisis hit us on 26th February 2020. Before that day, we had planned to hold SimAUD 2020 as a physical conference at TU Wien, loosely themed around the *Fridays for Future* movement and its implications for the AEC world. Accordingly, we wanted to avoid handing out conference bags and other plastic giveaways; we also intended to offer only vegan food, not only because it fit the subject but also as a homage to Azam Khan—one of SimAUD's founders whom we invited as a keynote speaker—who is an outspoken proponent of going vegan.

As more and more news of lockdowns in Europe came in, our chances for holding the conference physically became slim. When WHO finally declared COVID-19 a global pandemic on March 11th, 2020, we had to choose whether to cancel, postpone, or move the conference online. Canceling was no option since the authors would lose their chance of presenting their work, possibly leading to problems with research funding. This would also apply had we postponed the conference to some later date, e.g. after summer or even one year later; now that we know that COVID-19 is going to be around for quite some time, it seems wise that we did not settle on this option as well.

On March 18th, 2020, we decided to make SimAUD2020 into an online conference and fully embrace the advantages of this medium: Availability 24/7 through the use of pre-recorded video presentations, to be discussed in live webinars held during the conference; the ability to sift through papers using data analytics, and—since we speak about data— a new session for discussing datasets handed in together with the papers; a new session for education, which is what most of us use the Internet for nowadays.

Strangely, the shift to the online format also made us aware of existing physical borders: Some countries restrict the use of video-sharing platforms and videoconferencing tools, which has made participation impossible for some of our authors. Differences in time zones and daylight savings time regulations also made it necessary to hold each session *twice* so that authors could participate *during their daytime*. Twice the amount of sessions would have meant extending the duration of the conference—a no-go if one considers that some sessions are only sparsely filled. Instead, we opted to merge such sessions in order to densify our online schedule (figure 1). The proceedings, however, remained unaffected by this decision, as each of our sessions has its own chapter.
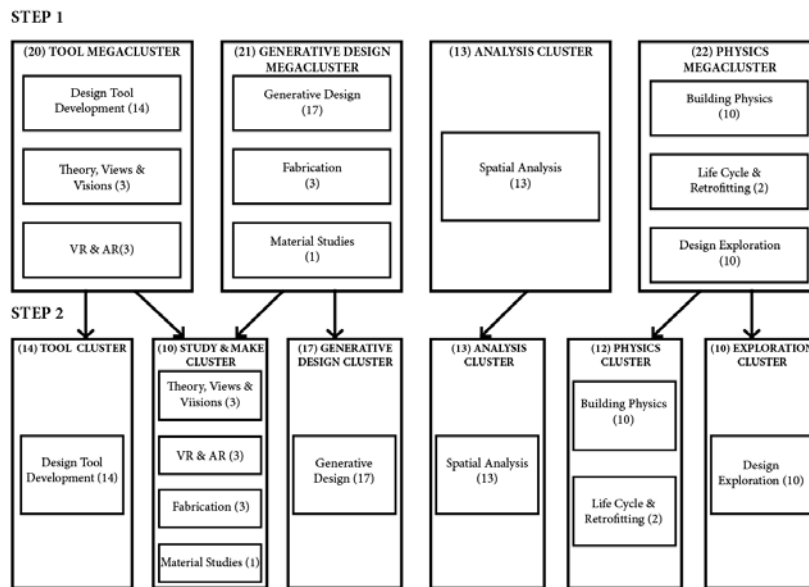

Figure 1. *Merging sessions. (Step 1) Clustering related sessions, (Step 2) splitting/merging these clusters again if they are too big. Numbers in (brackets) denote the amount of papers. Image by Myrto Karnoutsou.*

Coming to the contents of this *Proceedings of the 11ᵗʰ Annual Symposium on Simulation for Architecture and Urban Design*, we see a total of 72 regular papers grouped into 10 chapters (=sessions). We additionally have 4 keynote papers by Christina Hopfe, Ardeshir Mahdavi, Bige Tuncer and Azam Khan. All papers were subject to a double-blind peer review-process (headed by our scientific chairs Christiane M. Herr, Ulrich Pont, Dana Cupkova and Gabriel Wainer). Reviews were conducted by our scientific committee which consisted of 93 reviewers, who ensured the high quality of the conference was kept and proper feedback was provided to all of the authors who handed in (111 submissions, 80 accepted). Individual papers are indexed in the ACM Digital Library after the conference.

The actual conference takes place 25-27 May (*when measured in UTC*; however, in Alaska it starts on May 24 and in New Zealand it ends on 28 May). It features 6 paper discussion sessions that happen twice a day. We warmly thank all authors and chairs for participating in these discussions, as these now form the core of the online conference. These discussion sessions are summarized for the first time in *papers panel* sessions by each of the respective chairs (the papers panel itself is moderated by our keynote speakers). Another novelty of this year's conference lies in the appointment of a workshops coordinator (Wolfgang E. Lorenz) to manage our 5 workshop sessions happening in parallel (the rest of the conference is sequential, i.e. *single-track*). Help was provided by our 5 student volunteers (Jakub Klas, Judith Hofer, Dominik Fill, Androniki Pappa and Myrto Karnoutsou) who are also attending/assisting the live discussions and giving people a sense of *presence* (actually being at the conference even when physically sitting at home) by using social media to report "who says what" and "what happens where". Lastly, we offer the "SimAUD Educational Community" and "SimAUD Dataset Repository" sessions (both for the first time; the latter is prepared by Anastasia Gassia and Myrto Karnoutsou) in order to further extend the scope of SimAUD.

All of this would have not been possible without institutional support. We are incredibly thankful to SCS (Oletha Darensburg and Carmen Ramirez), ACM, IBPSA, TU Wien's Center for Computational Complex Systems (Niki Popper) and TU Wien's research unit of Digital Architecture & Planning (Bianca Braun, Peter Kompatscher), and to our exclusive sponsor, Autodesk, for their continuing support of the SimAUD conference.

Gabriel Wurzer

SimAUD 2020 Program Chair

Technische Universität Wien, Vienna, Austria

Angelos Chronis

SimAUD 2020 General Chair

Austrian Institute of Technology, Vienna, Austria

Contents

CONFERENCE PROCEEDINGS

Design Tool Development 395

Material Studies

Theory, Views & Visions

Fabrication

Keynote Speakers

Azam Khan

Trax.GD & University of Toronto

Lecture Title: **SimAUD Grand Challenge: A Decade of Action**

Azam Khan is Co-founder & CEO of http://Trax.GD developing an urban simulation and visualization platform. He founded SimAUD and is Adjunct Professor at University of Toronto. He has held positions at Autodesk Microsoft and Adobe.

Bige Tunçer

Singapore University of Technology and Design

Lecture Title: **Data Driven Design to Design Driven Data**

Bige Tunçer is professor at Singapore University of Technology and Design leading the Informed Design Lab focused on data driven architectural & urban design. She leads multidisciplinary projects in evidence informed design, IoT & big data.

Christina Hopfe

TU Graz

Lecture Title: **The dilemmas of decision making**

Christina Hopfe is full Professor in Building Physics at TU Graz. She is a Fellow of the HEA, CIBSE, and IBPSA. She is a Director-at-Large of IBPSA and the chair of the communication committee.

Ardeshir Mahdavi

TU Wien

Lecture Title: **Bringing HIM closer to HER**

Ardeshir Mahdavi is an internationally acclaimed expert in Building Physics, Computational Modelling, and Human Ecology. He has authored over 700 scientific publications and supervised over 60 doctoral students.

Keynote Papers

SimAUD Grand Challenge: A Decade of Action

Azam Khan

Trax.GD
Toronto, Canada
azam.khan@trax.gd

Department of Computer Science
University of Toronto
Toronto, Canada

ABSTRACT

Zoonosis, passing disease from animals to humans, may be the central topic of this decade due to the unprecedented global impact of the SARS-CoV-2 virus. With *land-use change* identified as the primary driver of emerging infectious disease (EID) events, we have the opportunity within the SimAUD community to study and positively impact urban expansion, deforestation, and pandemic preparedness through design. The first decade of the Symposium on Simulation for Architecture and Urban Design has built a strong community. In the second decade we can leverage our community to be greater than the sum of its parts. We can focus our sustainability science efforts around the United Nations Sustainable Development Goals which include land-use change, good health and well-being, clean energy, sustainable cities and communities, and more. As a high-impact empirical target, our own campuses and institutions each support tens of thousands of people, all impacted by the pandemic, and consume large amounts of water, food, energy, and materials.

We propose to the community a SimAUD Grand Challenge to adopt a mission-based research agenda to demonstrate leadership, working together with each of your institutions, to plan, design policies, and implement a net-zero campus transformation by 2030 that also minimizes encroachment and increases resilience to vector-borne and zoonotic diseases. This would result in examples of successful science-policy interface efforts that can inspire regional and national groups to raise their understanding and expectations for progress on the global United Nations 2030 Agenda for Sustainable Development.

Author Keywords

Sustainable Development Goals; land-use; human encroachment; zoonosis; net-zero campus.

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING (e.g. Model Development).

1 INTRODUCTION

Zoonosis, passing disease from animals to humans, may be the central topic of this decade. Land-use change is the primary driver of emerging infectious disease (EID) events at 31% of all causes of zoonosis [1,2] (see Figure 1). As such, we have the opportunity in the SimAUD community

to study and have a positive impact on issues like urban expansion, deforestation, and pandemic preparedness through design and action research. Still, land-use change is part of a complex system of interacting resource and societal flows [3] and system-level effects must be considered.
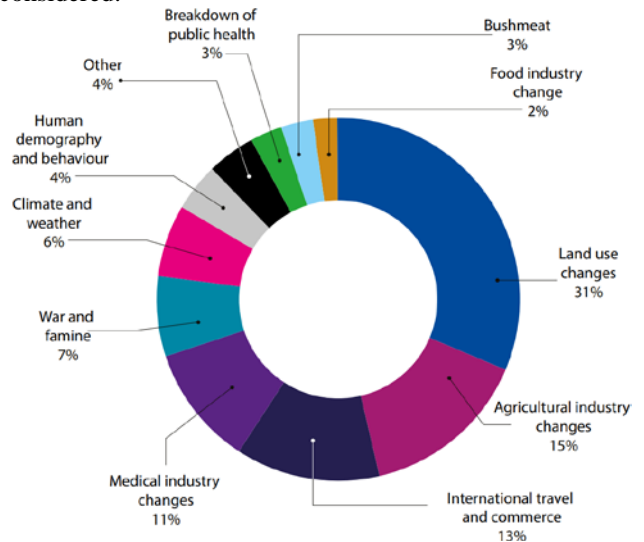
Figure 1. Primary drivers of past disease emergence.

A high-level analysis [4] of land-use indicates that 50% of all habitable land is used for agriculture and 77% of that is used for livestock (meat and dairy), see Figure 2.
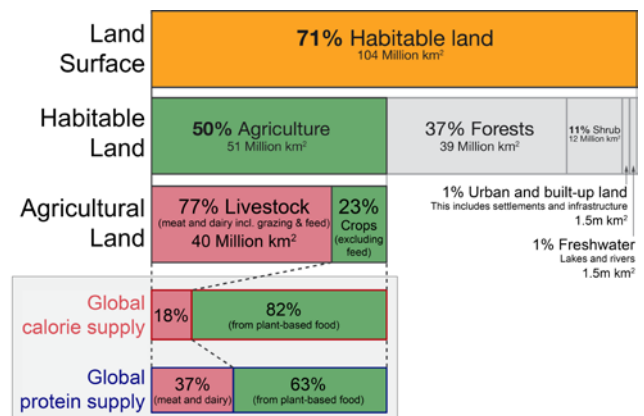
Figure 2. Global land-use for food production [4].

A considered approach to the use of land will be needed to support the growing human population while using less

land overall. Ideally, large amounts of farmland could even be restored back to forests and natural habitats to address other aspects of climate change, if populations could significantly increase plant consumption and reduce meat consumption. In planning the sustainable transformation of communities at the scale of a university campus, local district-level food, water, and energy production must be maximized.

2   ZOONOSIS AND ECOSYSTEM DESTRUCTION

Zoonosis can happen in industrial animal facilities, as with the H1N1 Swine Flu virus pandemic. Wild animal hunting can result in disease transmission as with the Human Immunodeficiency Virus (HIV) global epidemic. Illegal and poorly regulated wildlife trade, as with the current coronavirus pandemic, is also a major cause of zoonosis, as was the first SARS pandemic. Ebola, West Nile virus, avian influenza, Lyme disease, Zika virus, and many neglected zoonotic diseases continue to have a significant impact on society [1,2].

The unprecedented global action in 2020 against the spread of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) demonstrates a new sense of urgency in government to act relatively quickly, broadly, and co-operatively to control the international spread of the virus. The United Nations Environment Program [1] states that the "emergence of zoonotic diseases is often associated with environmental changes or ecological disturbances, such as agricultural intensification and human settlement, or encroachments into forests and other habitats." Instead, preserving ecosystem integrity buffers humans from pathogens both physically and through species diversity.

3   SCIENCE FOR SUSTAINABLE DEVELOPMENT

At the same time, sustainable development has been an ongoing concern in our community and was in fact a central consideration in the creation of SimAUD in 2010. Recent EID events further highlight the systems nature of our chosen problem domains. As a scientific community, we must now expand our thinking (and impact) to topics that were previously considered to be outside of our purview. As our point of departure, we take the recommendations of the Global Sustainable Development Report 2019: The Future is Now – Science for Achieving Sustainable Development [3]. The report asks, "what do we know about transforming the world?" and proposes a matrix of six domains as entry points and four sectors as levers, each combination forming a pathway. This approach for organizing effective action toward the Sustainable Development Goals (SDGs) [5] can also be used for planning campus transformation projects at a systems-level. At the architecture-level, Mossin et al. [7] provide a guide of sample projects interpreting the SDGs. Finally, the SDGs represent a societal lens to bridge between architecture and urban design while also providing a common reporting structure for comparability, peer-review, collaboration, and education, while also fulfilling our namesake.

4   GRAND CHALLENGE

Inspired by both the United Nations global call for a Decade of Action [6] and the Global Sustainable Development Report [3] call for strong institutions (Goal 16) as being an enabler for implementing all of the goals, we propose a SimAUD Grand Challenge for members of our community to work with their institutions to develop a ten-year campus plan for sustainable development. Specifically, our campuses should be net-zero energy, water, food, and waste by 2030. Moreover, our solutions must increase the resilience of our campuses to vector-borne and zoonotic diseases and minimize human encroachment. This would result in examples in many contexts of successful science-policy interface efforts that can inspire regional and national groups to raise their understanding and expectations for progress on the global United Nations 2030 Agenda for Sustainable Development.

REFERENCES

1. UNEP (2016). UNEP Frontiers 2016 Report: Emerging Issues of Environmental Concern. *United Nations Environment Programme*, Nairobi. pp. 73.

2. Loh, E.H., Zambrana-Torrelio, C., Olival, K.J., Bogich, T.L., Johnson, C.K., Mazet, J.A.K., Karesh, W., and Daszak, P. (2015). Targeting Transmission Pathways for Emerging Zoonotic Disease Surveillance and Control. *Vector-Borne and Zoonotic Diseases*. 15(7). pp. 432-437.

3. Independent Group of Scientists appointed by the Secretary-General. (2019). Global Sustainable Development Report 2019: The Future is Now – Science for Achieving Sustainable Development. United Nations, New York, 2019. pp. 216.

4. Ritchie, H. and Roser, M. (2020)."Land Use". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/land-use'.

5. UN General Assembly (2015). Transforming our world: the 2030 Agenda for Sustainable Development. United Nations, New York. pp. 41.

6. Guterres, A., Secretary-General of the United Nations (2019). Retrieved from: 'https://www.un.org/sustainabledevelopment/decade-of-action/'.

7. Mossin, N, Stilling, S, Bøjstrup, T, Larsen, VG, Blegvad, A (ed.), Lotz, M (ed.) & Rose, L. (2018). An Architecture Guide to the UN 17 Sustainable Development Goals. The Royal Danish Academy of Fine Arts, Schools of Architecture, Design and Conservation. pp. 158.

Bringing HIM closer to HER

Ardeshir Mahdavi

Department of Building Physics and Building Ecology
TU Wien
Vienna, Austria
amahdavi@tuwien.ac.at

ABSTRACT

The approach to planning and maintaining the built environment should be arguably people-centric ideologically and evidence-based methodologically. Whereas the former requires concerted Human-Environment-Research (HER), the latter could be implemented in terms of Human-Information-Modeling (HIM). HER is complex, and HIM is challenging. But to properly approach HIM, we need to rely on HER. As things are, there is still a gap between HIM and HER. The present contribution addresses this gap and the prospects of reducing it.

Author Keywords
Human-Environment-Research; Human-Information-Modeling

1   OF EFFECTIVENESS AND EFFICIENCY

Information technology has the potential to support various aspects and phases of building design, construction, and operation. Ongoing research and development efforts can further strengthen this potential. It is thus important to regularly reflect on the orientation and quality of such efforts. The present contribution focuses on a specific aspect of recent research regarding building design and operation support, characterized by increased recognition of the importance of the human dimension of the building delivery process. From the computational modeling perspective, this underlines the need for more detailed and realistic models of people, their presence and movements in buildings, as well as their interactions with buildings' features and systems. Toward this end, agent-based modeling (ABM) [7] ranks high amongst approaches that have been deemed to be promising.

Note that the systematic inclusion of high-resolution representations of people in buildings, for instance, via ABM, is relevant to both effectiveness and efficiency of buildings' performance. Whereas provision of conditions conducive to people's health, comfort, satisfaction, and productivity are associated with building performance's effectiveness, such conditions must be provided in an efficient manner from the point of energy and environmental impact. As such buildings are expected to provide a high degree of habitability (associated with the effectiveness requirement) in an energy and resource saving manner (associated with the efficiency requirement).

Deployment scenarios of computational design support methods typically address both of these dimensions. The adequate (effective) attributes of designs (mainly in view of indoor-environmental quality) are to be realized in an efficient manner (e.g., in terms of energy and resource use). Thus, computational tools help us to figure out if

i)   a design proposal is effective, i.e., if it leads to a more habitable environment, and

ii)   the provision of habitability is accomplished in an efficient manner.

Particularly the evaluation of buildings' effectiveness necessitates a deep understanding of habitability requirements. As such, this is not a predominantly computational issue, as the representation of the human dimension in building modeling, or Human-Information-Modeling (HIM) requires domain knowledge, and must rely on Human-Environment-Research (HER). The strength of computational methods and associated formalisms lies in identifying how matters can be done efficiently, once the effectiveness criteria are settled. But sometimes, it seems as if it is believed that formalisms on their own (or based on simplistic and superficial content-related input assumptions) can lead to optimal solutions. This contribution argues that, while efficient formalisms are necessary and essential, formalism-centric thinking may divert from the appreciation of the indispensable significance of dependable domain knowledge.

2   OF HAMMERS AND NAILS

Prior to taking a closer look at the scope and effectiveness of occupancy-related ABM applications in building modeling, it may be instructive to briefly reflect on the relationship between formalisms on the one side and the circumstances (or phenomena) they aim to frame, capture, describe, or explain on the other side. At a general level, we may roughly think of the formalism (involving, for instance, a tool, an instrument, or a technique) as a kind of abstract structure projected onto observed circumstances one tries to understand.

In this context, occasionally, the following problem arises: Fixation on the formalism can not only lead to inaccurate or mistaken projections, but in a sense to supplant the essence of the circumstances or phenomena that were meant to be understood and represented. Application-seeking formalisms, occasionally referred to metaphorically as hammers seeking for nails, may follow motivationally a well-known trait of human cognition system, namely the tendency to seek for meaningful patterns in the observational field. Arguably, such a trait is, from the evolutionary standpoint, of high survival value. The consequences of falsely recognizing a non-existent threat may be inconvenient, but failing to recognize the pattern associated with a real threat may be fatal.

Note that, abstract formalisms, such as theories in mathematics (e.g., topology) can be subject of queries and probing independent of physical realities they may be deployed to describe. Such formalisms may exist long before they are found to be relevant for developing physically relevant models. For instance, 17th century progress in description of the planetary movements could refer back to prior – much older (i.e., ancient Greece) and independently established – investigations of conic sections [17]. Likewise, formulation of Einstein's general relativity benefited from Grossmann's input concerning the previously established understanding in non-Euclidean geometry, differential geometry, and tensor calculus [12, 13, 32].

On the other hand, explorations in physics might trigger new developments in the domain of abstract mathematical formalisms. For instance, Newton's formulation of his "fluxional" was arguably motivated by his prior investigations into the behavior of physical systems (for instance, the increase in the speed of falling objects). Likewise, Paul Dirac introduced, in his book "The Principles of Quantum Mechanics", the delta function as a continuous analogue of the Kronecker delta (the discrete version of the delta function) [10]. As such, this mathematical structure for the representation of point objects (e.g., mass, charge) is widely applied in quantum physics. Later, Laurent-Moïse Schwartz developed the theory of distributions [33], which provided a solid interpretational basis for objects such as the Dirac delta function.

The dialectic of mathematical formalism and descriptions of physical phenomena has been as such highly fruitful. But overtly formalism-centric approaches can – in specific circumstances – end up diverting or even impairing research progress. The following brief references to a few instances of such formalism-centric attitudes do not all fit along the abstract versus concrete axis. But they illustrate instances of what can be expected when a discourse starts with a cherished formalism (or a pattern, tool, technique, algorithm, etc.) searching for factual (or observational, empirical, etc.) circumstances onto which they could be projected:

- There is a vast body of treatments (some scholarly, some amateurish) dealing with the illustration of real and putative (geometric) patterns and proportions behind objects in nature, arts, and architecture. Common instances of such patterns involve the golden section, Fibonacci series, and harmonic intervals [30]. Even while discussing the one and the same subject (for instance, the geometric features of a cathedral), there are disputes as to which of such patterns are the ones truly underlying the compositions [16]. This reinforces the impression, that once a certain interest – if not passion – is aroused concerning a specific pattern, scheme, or proportional system or theory, there is tendency to look for – and claim to have shown – its prevalence. It is of course quite possible, particularly in case of artifacts of art and architecture, that designers and creators indeed consciously drew on such patterns. But in many other cases, the multitude of both candidate patterns and the geometric affordance of the observed material as well as the fuzziness of mapping operations confers a notable degree of arbitrariness to the alleged discoveries.

- Fascination with formalisms and the urge to widely project them onto all kinds of objects, entities, and phenomena appears at times to follow certain cycles and fashion-like tendencies. One such formalism – highly popular in the last decades of the previous century emerged from an extension of the classical geometry, namely fractal geometry. Here again, the claim to interpretative potency goes beyond the initially postulated "Geometry of Nature" [28] to cover "… carpets, bricks, and much else besides" [5], not to forget about applications in architecture [18].

- Another instance of the proverbial hammer looking for nails may be recognized in the popularity of the so-called shape grammars [35] in a number of academic institutions in the eighties and nineties of the last century and associated meetings and symposia in the design computing community. Aside from purported applicability toward compositional design, it was also suggested to provide means of analyzing various design languages [9]. It appears as if a sense of exclusivity and entitlement in matters of design interpretation – on occasions, not far from a kind of universality claim – accompanied shape grammar discourse and theory developments. However, it has arguably fallen short of providing convincing instances of practical applications.

- In multiple instances in the past, an entire scientific system (discipline, theory) has been adopted in terms of formal explanatory vehicles in architectural domain. Notable instances of such systems may be found, for instance, in semiotics, cybernetics, and information

theory [22]. Semiotics is a theory of signs, how they function, and how they acquire meaning [8, 11]. Information theory concerns the measurement and transmission of information [34]. Cybernetics is not only about regulatory processes in biological and technological systems, but has been applied also in cognitive, and social systems [3, 37]. Thereby, notwithstanding the actual value and utility of these disciplines, once again aspects of aforementioned nail-seeking hammer symptom are recognizable, including an initially enthusiastic adoption (particularly in architectural discourse), a tendency to claim wide applicability despite insufficient evidence, and a slow fading process, typically followed up by another fashionable paradigm.

- The potential and value of parametric simulation – particularly in design development and optimization scenarios is well established. However, in the context of research in relevant areas of building science, parametric simulation is of very limited use, if it is carried out solely based on an empirically untested computational model [31]. Carrying simulation-aided virtual experiments using an unavoidably reduced model of reality may be useful in certain situations, but does not substitute the need for independent empirical examination and validation of the respective findings. A simulation model may provide only a limited proxy of the reality, not a legitimate replacement of empirical information. The tendency to throw parametric simulation runs at fundamental research questions may be thus interpreted as yet another instance of tools and techniques aiming to demonstrate their applicability independent of availability of empirically-based evidence of the validity of the underlying models.

These instances may also help better understand some of the contemporary tendencies in view of formalism-centric approaches in the present contribution's topical focus, namely the inclusion of the occupancy-related factors in computational building performance assessment. Factors that motivate some of these approaches may be discussed in terms of the following – roughly sketched – categories:

- A number of efforts have attempted to develop predictive models of occupants' behavior based on rather limited sets of data. Specifically, efforts have been made to represent the dynamics and probabilistic appearance of occupancy-related information (for instance, patterns of occupants' arrival in an office building, or the occupants patterns of window operation in a residential building) using stochastic methods (involving, for example, generalized linear mixed models) [24]. There is as such nothing wrong with the motivational background of such efforts. A problem arises though, when the deployed formalism is claimed to have explanatory or predictive power beyond the scope of data underlying its development. A number of

studies have falsified such claims using independent empirical information [23, 24, 25, 26]. To be clear, there is no reason why we should not be able to introduce data-driven stochastic formalisms that can capture occupants' behavioral traits with more fidelity. But it is not the specific type of a formalism as such that establishes its credibility. Rather, as always in science, conformance to observational data remains the ultimate validity criterion.

- A second group of ongoing efforts [20] are propelled by the perceived success of the so-called "big data" approaches in other fields [29]. An implied assumption is that application of data mining techniques to data streams from multiple sources (e.g., monitoring systems, building automation, occupants' phones and other personal devices) can effectively address building-related challenges, including operational optimization pertaining to energy use and indoor-environmental conditions. Data mining, pattern recognition, and further AI-based techniques have been indeed demonstrated to provide efficient computational solutions in a number of domains (such as those requiring powerful search algorithms) without relying on white-box models (e.g., explicit, causal or first-principles based). But the applicability and effectiveness of such solutions in the behavioral modeling field has not been clearly established. A relevant question is thereby the depth of contribution of inherently black-box type methods when – particularly in design support applications – uses are expected to obtain a deeper understanding concerning the causal and motivational factors influencing behavioral manifestations.

- A further class of tools and techniques are motivated by the – principally reasonable – premise that no model can be ever considered to be fully correct. Rather, there will always be some level of uncertainty associated with both model input assumptions and the resulting computational output. Sensitivity analysis, uncertainty analysis, and Bayesian reasoning [15, 36] are geared toward operationalization of uncertainties, be those the result of insufficient or low quality input data or a consequence of knowledge gaps in the models' underlying representations of real-world phenomena and processes. Applications of such techniques and methods can be undoubtedly useful – for instance when evaluating the reliability of computationally derived building performance indicator values. However, the usefulness level of such application scenarios is a direct function of the empirical validity level of assumed uncertainty distributions of pertinent model variables. In the absence of such empirical grounding, formally produced ranges of uncertainty would be of little use. Likewise, the utility of Bayesian reasoning remains limited, as long as assumptions concerning the pertinent priors are entirely arbitrary. In such instances, we end up with yet another instance of models pretending to

7

generate behavioral data (e.g., patterns of building occupants' interactions with buildings' environmental control systems) when there is little or no empirical evidence for their validity.

To venture into a bit of meta-critical reflection, even search and finding instances of confirmation bias itself may be an instance of confirmation bias at work. Nonetheless, the preceding critical remarks do exemplify manifestations of a general tendency that could be viewed as a specific case of the confirmation bias: Having developed or adopted certain tools, methods, and formalisms, people (including researchers) are invested in demonstrating their broad utility and effectiveness. Whereas developing and tweaking formalisms is something we can pursue and control, relevance and applicability can only be established based on – frequently cumbersome – empirical data collection and analysis. It is thus not surprising that in the domain of building design and operation support too, formalisms have been at times granted credibility prematurely – that is without a thorough evidence-based examination [23, 24, 25].

All this is supposed to underline a central concern: Reliability and fruitful application of formalisms meant to capture people's behavioral patterns (in other terms, Human Information Modeling or HIM) depends on the depth and credibility of the underlying empirical understanding of people's perceptual, evaluative, and behavioral processes (resulting from properly conducted Human-Environment Research or HER). This applies also when we look at the potential of formalisms associated with agent-based modeling (ABM) techniques, tools, and platforms toward productive implementation of HIM in the building design and operation domain. Here too, we need to examine if and to which extent HIM is supported by findings from HER.

3 OF AGENTS AND ATTITUDES

Agent-based modeling (ABM) has been applied in a large and diverse set of domains [7, 21, 38], including also those pertaining to the built environment [1, 2, 4]. ABM appears to provide an adequate vehicle for the detailed and dynamic representation of building occupants in performance simulation applications. Thereby, behavioral tendencies of occupants – for instance, their decision making processes with regard to actions relevant to buildings' operation – can be coded in terms of a set of rules. As such, ABM has the potential to address the complexity and dynamics of such processes. It can also facilitate the representation of inter-occupant behavioral influences and learning phenomena.

As a simulation technique, the previously raised general issues concerning the applicability and utility of formalisms are also relevant in the case of ABM. Accordingly, we can argue that, aside from coding integrity and usability issues, ABM's usefulness in building design and operation scenarios directly depends on the knowledge base underlying assumptions regarding agents' behavioral repertoire.

To explore the sources of knowledge that have been used to define agents' behavioral patterns in building-related ABM applications, we recently reviewed a number of related scientific publications [6]. Specifically, we identified 23 publications with a focus on the simulation of buildings' energy and indoor-environmental performance.

The result of this review suggest that only a minority of the research efforts (about 30%) rely on an explicit underlying behavioral theory. Some 39% of the contributions rely exclusively on basic general material included in codes and standards, meaning they neither refer to a theoretical foundation nor include any original empirically-based information on occupants' behavior.

Some form of empirical data (interviews, observations) can be found in roughly 39% of the efforts. However, such data cannot be assumed to be generally representative or applicable, as none of the contributions include empirical data obtained from multiple buildings or locations. A simultaneous inclusion of theoretical underpinnings, standard-based information, and – at least a minimum of – empirical data could be found only in 13% of the reported ABM development efforts.

The above brief reference to the conducted review exercise concerning ABM applications in building performance simulation yields a couple of essential insights. From a purely technical point of view, it can be concluded that the ABM-based emulation of occupants' presence and behavior in buildings has become increasingly feasible. However, the actual utility of ABM-based building design and operation instruments remains limited due to both *i)* insufficiently developed explanatory theories of human behavior as individuals and as members of socially relevant groups; *ii)* the paucity of sufficiently detailed and representative empirical data.

4 OF PSYCHE AND PHYSICS

The main point of the arguments so far has been to underline the indispensable role of domain knowledge (empirical observation and explanatory theories) for the development and validation of formalisms in general and ABM in particular.

Relevant for the present discussion are explanatory theories of building occupants' perception of and behavior in indoor environments. Such theories are expected to provide the foundation for an evidence-based approach (together with related codes, standards, guidelines, and tools) to building design and operation. However, despite much research in the past decades (including recent studies in psychology and neuroscience that deal with human perception and cognition), one needs to go back to intellectual currents in the twentieth century, if not the late nineteenth century, to locate high-level explanatory theories of human perception and behavior (as proposed, for instance, in Psychophysics, Human Ecology, Ecological Psychology, and Cybernetics).

The examination of these theories reveals a few recurrent trends [22].

One trend (mainly relevant to perception models) aims at operationalization of purported correlations between the strength of physical stimuli (e.g., temperature, luminance, sound level) and the intensity of people's subjective perception (e.g., thermal sensation, perception of brightness and loudness). Such correlations may be the reason for prevalence (in codes, standards, and guidelines) of scales that map measurable environmental stimuli onto occupants' physiological and psychological responses. However, the utility of psychophysical scales – as well as other similar or derivative approaches – is limited due to multiple circumstances, including:

i)   the non-linear nature of perceptual responses to multi-sensory (i.e., thermal, visual, olfactory, auditory) exposure situations;
ii)   the difference between immediate perceptual correlates of physical stimuli and the higher-level assignment of values to such perceptions;
iii) people's attitude toward and prior experience with the sources of stimuli as well as their true and perceived level of control over such sources; and
iv)   people' diversity, as evidenced in inter-individual differences in age, gender, cultural background, physical and mental health, attitudes, expectations, preferences, etc.

Another trend more relevant to behavioral models, starts with survival as the most fundamental priority (value) of an organism and the requisite of the organism to maintain its life-critical internal state in a proper range. In case of organisms with high-level environmental mapping capability such as human beings, both non-conscious homeostatic processes and conscious decisions to act triggered by feelings of pain and pleasure facilitate the move toward the homeostatic range beneficial to survival and well-being. Amongst other things, this insight has been fundamental to development of comfort models pertaining to indoor-environmental conditions. For instance, conventional thermal comfort models typically highlight the human body's need to maintain heat balance over the long run [14]. Protracted departure from body states associated with heat balance is assumed to trigger thermal discomfort and encourage compensatory (homeostatic) processes.

Nonetheless, extensive research and standardization in the thermal comfort domain has not resulted in definitive instances of models and codes. Alterations suggested by the adaptive thermal comfort theory may be seen as progress, but they too do not conclusively resolve challenges encountered by thermal comfort models. Multiple reasons can be cited for this circumstance. One problem is the number of suspected independent variables of comfort and associated behavioral models, as well as the available information on the possible values of such variables. Even if we could construct a multi-disciplinary multi-layered

(physical, physiological, psychological, social) causal model of thermally relevant comfort or behavior, we would still need to obtain reliable values for the multitude of variables involved.

Additional complexity results from the fact that both indoor-environmental quality judgments and adaptive behavior are influenced by occupants' views and attitudes. For instance, in office buildings, not only the scope of real and perceived control may matter, but also occupants' views on the commitment and professionalism of building management personnel. Such a circumstance is exemplified by occupants' tolerance (the "so-called forgiveness factor") of occasional departure from comfort conditions due to inhabitants' social and environmental attitudes [19].

Moreover, comfort and behavioral models face difficulties in accounting for the sizable diversity amongst the population of buildings' occupants. Besides, in building design scenarios, the detailed composition of future occupants is rarely known, making it difficult – if not impossible – to predict occupants' perception and behavior patterns. As such, uncertainties resulting from inter-individual diversity of building occupants can outpace the differences between predictions of rival models.

We should also bear in mind that, as alluded to before, occupants' perceive (and act in) a multi-dimensional exposure field, including – amongst other things – thermal, visual, auditory, and olfactory factors. Perception of comfort and disposition to control-oriented behavior with regard to one environmental factor is not independent of exposure parameters pertaining to other factors. But existing models do not capture such interdependency. As suggested by the brief treatment of recent publications on this subject in the following section of this paper, much more research is needed to arrive at a deeper understanding and explanatory models of occupants perceptual and behavioral tendencies in multi-domain exposure situations.

5   OF EYES AND EARS

Development of robust theories in all fields of inquiry requires a sufficient density of available facts. This of course applies also to the field of behavioral research relevant to building occupants. As mentioned before, despite a long tradition of studies, health standards and comfort guidelines are still mostly single-domain (i.e., they frequently address thermal, visual, olfactory, and auditory factors in isolation). As such, they fall short of accounting for real indoor-environmental exposure situations involving multiple kinds of stimuli. This assertion was confirmed by the results of a recent international review effort concerning this topic [27].

Past multi-physical perceptual research has undeniably generated much valuable knowledge. However, the results are frequently inconclusive, and in part even contradictory. The majority of the studies were short-term, and the human participants not necessarily representative of the general

9

population. Moreover, most studies were conducted in rather artificial settings and could not truly succeed in addressing the implications of the Hawthorne effect. The studies also do not sufficiently build on past research: As such, there is little evidence of actual carryover from past studies' findings into more recent ones. Consequently, studies do not apply standard research designs, data collection strategies, metrics, and statistical analysis techniques. This makes it thus very challenging – if not infeasible – to conduct meta-analyses of the subject.

These observations underline the importance of future international, collaborative, and interdisciplinary experimental studies. Thereby, the critical importance of conceptual and methodological integration of engineering and human science methods must be addressed. Moreover, fundamental theories must be developed that effectively accommodate the processes of hypothesis formulation and testing. To increase the credibility of occupant-centric computational formalisms for building design and operation support, we need to work toward an operationalized general theory of the nature of the perceptual and behavioral processes involved in multi-domain exposure situations.

6 OF PAST AND PROSPECT
Speaking in code, if BIM (Building Information Modeling) is to support occupant-centric building design and operation, it must incorporate HIM (Human Information Modeling). But HIM cannot be approached solely in terms of a series of alternative formalisms (e.g., stochastic methods, agent-based modeling), but must be based on solid domain knowledge derived from HER (Human Environment Research).

This contribution reflected on the understandable yet problematic tendency to approach HIM predominantly from a formalism-centric angle. Specially, given the pressures of funding acquisition in the academic and research institutions, there is a tendency to exaggerate the potential and practical applicability of proposed tools and models. Formalisms and abstract models can be played around with extensively (resulting, perhaps, in academically exploitable output) without the need to engage deeply in the unforgivingly laborious empirical research into the nature of human cognition, perception, and behavior. But, short of a rich body of findings and insights in HER, HIM developments remain superficial, and so too their incorporation in BIM.

REFERENCES
1. Alfakara, A., and Croxford, B. Using agent-based modelling to simulate occupants' behaviours in response to summer overheating. Simulation Series (2014), 46(7), 90–97.

2. Andrews, C. J.; Yi, D.; Krogmann, U.; Senick, J. A.; Wener, R. E. Designing Buildings for Real Occupants: An Agent-Based Approach. IEEE Transactions on systems, man, and cybernetics (2011), 1–15.

3. Ashby, W. R. An Introduction to Cybernetics, Chapman & Hall. (1956).

4. Azar, E., and Menassa, C. A conceptual framework to energy estimation in buildings using agent based modeling. Proceedings of the 10th Winter Simulation Conference (2010), 3145–3156.

5. Barnsley, M.F. Fractals Everywhere. Morgan Kaufmann Pub; Subsequent edition (1993).

6. Berger, C. and Mahdavi, A. Review of current trends in agent-based modeling of building occupants for energy and indoor-environmental performance analysis. Building and Environment (2020). https://doi.org/10.1016/j.buildenv.2020.106726.

7. Bonabeau, E. Agent-based modeling: methods and techniques for simulating human systems. Proceedings of the National Academy of Sciences of the United States of America, 99 Suppl 3, 7280–7287 (2002). https://doi.org/10.1073/pnas.082080899.

8. Buchler, J. Philosophical writings of Peirce. New York: Dover Publications (1955).

9. Chiou, S. and R. Krishnamurti, R. 1995. The grammar of Taiwanese traditional vernacular dwellings. Environment and Planning B: Planning and Design (1995). 22: 689-720.

10. Dirac, P. The Principles of Quantum Mechanics. Oxford University Press 1930.

11. Eco, U. A Theory of Semiotics. Indiana University Press; First Edition (1978).

12. Einstein, A.; Grossmann, M. Entwurf einer verallgemeinerten Relativitätstheorie und einer Theorie der Gravitation. Zeitschrift für Mathematik und Physik (1913). 62: 225–261.

13. Einstein, A.; Grossmann, M. Kovarianzeigenschaften der Feldgleichungen der auf die verallgemeinerte Relativitätstheorie gegründeten Gravitationstheorie. Zeitschrift für Mathematik und Physik (1914). 63: 215–225.

14. Fanger, P.O. Thermal Comfort. McGraw-Hill Book Company (1972). ISBN 0-07-019915-9.

15. Gelman, A. Bayesian Data Analysis. Chapman and Hall/CRC; 3. Edition (2013).

16. Haase, R. 20 Jahre Hans-Kayser-Institut f. harmonikale Grundlagenforschung. Schriften über Harmonik Band Nr. 17. Kreis der Freunde um Hans Kayser (1988) ISBN: 9783906643021.

17. Howard. E. An introduction to the history of mathematics. Cengage Learning; 6 edition (1990).

18. Iasef Md Rian, I.M.; Park, J.; Ahn, H.U.; Chang, D. Fractal geometry as the synthesis of Hindu cosmology in Kandariya Mahadev temple, Khajuraho. Building and Environment, Volume 42, Issue 12 (2007), pp. 4093-4107.

19. Leaman, A. and Bordass, B. Are users more tolerant of 'green' buildings? Building Research & Information. (2007) 35(6), 662 –673.

20. Linder, L.; Vionnet, D.; Bacher, J.; Hennebert, J. Big Building Data - a Big Data Platform for Smart Buildings. Energy Procedia, Volume 122 (2017), pp. 589-594.

21. Macal, C. M., and North, M. J. Agent-based modeling and simulation: ABMS Examples. Proceedings of the 2008 Winter Simulation Conference. Miami, Florida. (2008).

22. Mahdavi, A. Explanatory stories of human perception and behavior in buildings. Building Environment 168, 106498 (2020).

23. Mahdavi, A.: Common fallacies in representation of occupants in building performance simulation. Keynote: BSA2015 - Building Simulation Applications 2015 - 2nd IBPSA-Italy Conference (2015), Bolzano, Italy.

24. Mahdavi, A., Tahmasebi, F. People in building performance simulation. Building Performance Simulation for Design and Operation - Expanded Second Edition. Hensen, J., Lamberts, R. (Ed.); Routledge, New York (2019) pp. 117 - 145.

25. Mahdavi, A. and Tahmasebi, F. On the quality evaluation of behavioural models for building performance applications. Journal of Building Performance Simulation, 10, 5-6 (2017), pp. 554 - 564.

26. Mahdavi, A. and Tahmasebi, F. The deployment-dependence of occupancy-related models in building performance simulation. Energy and Buildings, 117 (2016), pp. 313 - 320.

27. Mahdavi, A.; Berger, C.; Jamrozik, A.; Chinazzo, G.; Prabha, L.; Schweiker, M. Understanding multi-aspect indoor-environmental exposure situations: Past insights and future needs. Indoor Air (2020). To be published.

28. Mandelbrot, B. The Fractal Geometry of Nature. Times Books; 2nd prt. edition (1982).

29. Mayer-Schönberger and Cukier, K. Big Data. Eamon Dolan/Mariner Books (2014).

30. Naredi-Rainer, P.v. Architektur und Harmonie. DuMont Köln (1982) ISBN 3-7701-1196-6.

31. Samuelsona, H.; Claussnitzer; S. Goyal, A.; Chen, Y.; Romo-Castilloa, A. Parametric energy simulation in early design: High-rise residential buildings in urban contexts. Building and Environment, Volume 101, 15 (2016) pp 19-31.

32. Sauer, T. Marcel Grossmann's contribution to the general theory of relativity. Proceedings of the 13th Marcel Grossmann meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories. World Scientific (2015), pp. 456–503.

33. Schwartz, L. Théorie des distributions. Hermann, 2 vols. (1966).

34. Shannon, C.E. The Mathematical Theory of Communication. The University of Illinois Press; 16th edition (1971).

35. Stiny, G. and Gips, J. Shape Grammars and the Generative Specification of Painting and Sculpture. Information Processing 71 (1972) 1460-1465.

36. Tian, W.; Heo, Y.; De Wilde, P.; Li, Z.; Yan, D.; Park, CS.; Feng, X.; Augenbroe, G. A review of uncertainty analysis in building energy assessment. Renewable and Sustainable Energy Reviews 93 (2018), 285-301.

37. Wiener, N. Cybernetics, or Control and Communication in the Animal and the Machine. Cambridge: MIT Press. (1948).

38. Wilensky, U. and Rand, W. An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with NetLogo. (2015).

The dilemmas of decision making

Christina J Hopfe

Graz University of Technology
Graz, Austria
C.J.Hopfe@tugraz.at

In times of crisis (such as climate change, heat waves, and more recently viral pandemics), it becomes apparent that the decisions we have made in the design of our built environment are crucial and can significantly impact our quality of life. Decision making in building performance simulation can add significant support to the design process from conceptual to detailed design and beyond. In support of this process building simulation is a tool which can provide quantitative feedback on the likely performance of alternative design options [1]. Building performance evaluators have to respond to a plethora of different decisions throughout the design process from conceptual ones such as the geometry and layout of the building to more complex ones in terms of code compliance, system design, material selection and control algorithms.

These decisions become more difficult and onerous to make if they are final and the user is unable to subsequently change them throughout the buildings' lifespan. The decision-making process becomes even more complex when there is more than one decision maker and conflicting priorities involved. The architect might aim for an exceptional design, the building physicist for the best performance, the user for the highest comfort, the client for the lowest costs. It is widely documented that resolving such complex and messy situations constitutes a significant part of the daily workload in most design practices [4,6,7].

To make the process even more cumbersome uncertainties compound this situation, including: uncertainties in the design, in physical parameters (that will be entered into building simulation), and in scenario conditions (such as the particular climate, extreme weather scenarios or the stochastic nature of user behaviour etc.) [3]. This not being enough, the uncertainty the decision maker contributes to the process is exacerbated by the choice and prioritization of more than one objective. What is more important, for example, comfort or costs - and if both then how should they be weighted?

Hazelrigg states that "good decision making is very often quite far from intuitive" [2], which is why designers tend to apply mathematical models to support their decision process. In contrast to this assertion interaction design researchers such as Stolterman [6] and Rogers [5] argue that theoretical approaches involving the use of

computational methods to resolve complex design decisions have a poor track record of success in design practice. Rogers [5] presents evidence showing that such approaches are often too time consuming, too difficult to learn and too abstract to be properly used and understood by the majority of design practitioners.

This talk focuses on decision making by categorizing the diverse approaches into four different schools of thoughts: (1) deterministic decision making; (2) decision making with Pareto; (3) decision making through the help of expert judgment; and (4) ethical decision making.

Examples are given where decisions have been made successfully (and unsuccessfully) using one or other approach. These examples will partially make use of building performance simulation; however, to allow a broader insight into the process of decision-making examples beyond the field of architecture are considered.

It is shown that the quality of decision is not necessarily proportional to the time and effort that goes into the decision-making process. Truly successful decision making appears to rely on reducing the decision space to its simplest dimensions and then finding the right balance between deliberate and intrinsic thinking. Thus, good decision making is both a science and an art.

REFERENCES

1. Clarke, J.A. and Hensen, J.L.M. 2015, Integrated building performance simulation: Progress, prospects and requirements, Building and Environment 91, 294-306

2. Hazelrigg, G.A., 2012. Fundamentals of Decision Making for Engineering Design and Systems Engineering.

3. Hopfe, C.J., Augenbroe, G.L.M. & Hensen, J.L.M. 2013, "Multi-criteria decision making under uncertainty in building performance assessment", Building and Environment, vol. 69, pp. 81-90.

4. Lawson, B., 2005. How Designers Think: The Design Process Demystified, 4 edition. ed. Routledge, Oxford; Burlington, MA.

5.Rogers, Y., 2004. New theoretical approaches for HCI. Annu. Rev. Inf. Sci. Technol. 38, 87–143.

6.Stolterman, E., 2008. The Nature of Design Practice and Implications for Interaction Design Research. Int. J. Dsign 2, 55–65.

7.Thackara, J., 2006. In the Bubble: Designing in a Complex World, New Ed edition. ed. MIT Press, Cambridge, Mass.

Data Driven Design to Design Driven Data

Bige Tunçer

Singapore University of Technology and Design
bige_tuncer@sutd.edu.sg

Cities are hubs for interaction among their inhabitants, and these interactions drive liveliness, innovation, culture, and economic welfare (Batty, 2013; Bertaud, 2018; Christiaanse et al., 2016; Storper and Venables, 2004; West, 2017). As complexity of big and developing cities grow with increase in density patterns, challenges arising from climate change, and socio-economic divisions in societies, urban planning and design practices have been adopting increasingly more data driven approaches with the aim of improving all forms of sustainability (economic, environmental, social) in cities.

The use of big data as an enabler of the smart cities vision has been investigated and put into practice in the recent decades (Albino et al., 2015; Kitchin, 2014). The transition from smart cities to responsive cities encompasses the placement of the inhabitants in the focus of city development, rather than the technological emphasis that utilizes, among others, sensor data streaming from all types of urban infrastructures (Schmitt et al., 2017). Smart cities target good performance and effective use of resources, however, technology itself can't automatically transform and improve cities and the lives of their inhabitants.

Evidence based design support in the context of urban design and planning, recently, makes use of big urban data and data analytics in order to replace some of the assumptions made during design and planning by grounded substantiation. However, having access to evidence related to aspects pertaining to various aspects of design and planning contexts and parameters does not lead to a linear translation of evidence to design and planning solutions.

At the Informed Design Laboratory of Architecture and Sustainable Design, Singapore University of Technology and Design, and Future Cities Laboratory, SEC, Singapore, we adopt an approach to evidence based design and planning that involves the collection, analysis, visualization and learning using multi-source, multi-scale, and multi-time data. This approach involves a combination of big and small data and qualitative and quantitative approaches, centered around a set of questions related to a design or planning issue, context, stakeholders and governance structure (Benita and Tunçer, 2019; Benita et al., 2019; Hertogs et al., 2018; Richards and Tunçer, 2018; Subramanian et al., 2019; Tunçer, 2020; Tunçer and You, 2017; Zhou et al., 2018). Differently from a data-driven design approach, we advocate a design-driven data approach. In this context, the set of questions that motivate

the investigation and the user/behavioral and physical/spatial parameters drive which data sets ought to be involved, how the user interaction of the computational design/planning tools with the stakeholders ought to be designed, and how the design/planning evidence ought to be translated into action, hand in hand with the domain expertise of the designers/planners.

The current COVID-19 pandemic that affects all of our urban environments around the world presents many challenges for our cities. Although it is not yet clear how the health related consequences will unfold in the coming year, we already have rather concrete indications for how city dwelling will be affected in the future, related to this pandemic or related possible future scenarios. The ongoing discourse (Bassett, 2020; Cairns, 2020) about the 'new normal' in urban life considers a fine balance between the social and economic benefits of high urban density and access of inhabitants to resources such as health care, green, energy, knowledge, education, etc. This is an exciting time to be a researcher.

The keynote talk discusses the above, giving the various thoughts context through completed and ongoing research of the Informed Design Laboratory (Figures).

Figure 1. A visualization of the frequency of runs, jogs, and fitness walks in the street network of Singapore (right) juxtaposed with green index, land use, street network, and eye level view classification data (left). Runners prefer recreational areas as destinations; however, street level mixed-use functions do not seem to affect running behavior as opposed to utilitarian walking. Image by Özgün Balaban (Balaban 2019).
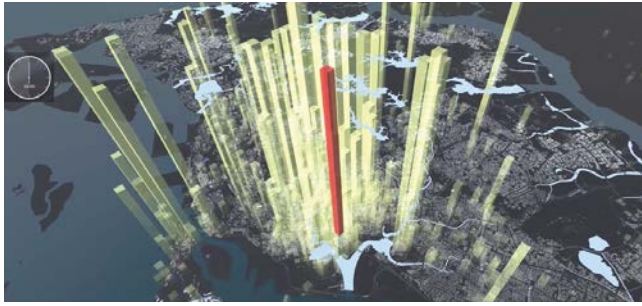
Figure 2. Cultural planning in cities is changing because of the effect that social media has on the creation and consumption of art A collection of social media feeds over two years was filtered in order to select the feeds related to art. These were mapped to locations and districts that receive public subsidies in Singapore. The result presents 'confirmations', negations', and 'discoveries' of art spaces in Singapore. Image by Ludovica Tomarchio (Tomarchio et al., 2016).

REFERENCES

Albino, V., Berardi, U., and Dangelico, R.M. (2015). 'Smart cities: definitions, dimensions, performance, and initiatives'. Journal of Urban Technology, 22(1): 3–21

Balaban, Ö. (2019). 'Understanding Urban Leisure Walking Behaviour: Correlations between Neighbourhood Features and Fitness Tracking Application Data ', Ph.D. Dissertation, Singapore University of Technology and Design.

Bassett, M.T. (2020). 'Just Because You Can Afford to Leave the City Doesn't Mean You Should', Opinion Piece, The New York Times, May 15, 2020, https://www.nytimes.com/2020/05/15/opinion/sunday/coronavirus-cities-density.html

Batty, M. (2013). The new science of cities. Cambridge, Massachusetts: MIT Press.

Benita F. and Tunçer B. (2019). "Exploring the effect of urban features and immediate environment on body responses", Urban Forestry & Urban Greening 43.

Benita F., Bansal G. and Tunçer B. (2019). "Public spaces and happiness: Evidence from a large-scale field experiment", Health & Place 56, pp. 9-18.

Bertaud, A. (2018). Order without design: How markets shape cities. Cambridge, Massachusetts: MIT Press.

Cairns, S. (2020). 'Is density doomed? Cities in a post-Covid-19 world', Staits Times, May 14, 2020, https://www.straitstimes.com/opinion/is-density-doomed-cities-in-a-post-covid-19-world

Christiaanse, K., Hetey, A., Schmidt, V. and Zou, Y. (2016). 'Urban breeding grounds', Future Cities Laboratory Magazine Special Issue. Singapore: Singapore-ETH Centre.

Hertogs P., Tunçer B., Schläpfer M. and He P. (2018). "A weighted graph model to estimate people's presence in public space - The Visit Potential Model", eCAADe2018, pp. 611-620.

Kitchin, R. (2014). 'The real-time city? Big data and smart urbanism'. GeoJournal; 79(1): 1-4.

Richards D. and Tunçer B. (2018). "Using image recognition to automate assessment of cultural ecosystem services from social media photographs". Ecosystem Services 31(C), pp.318-325.

Schmitt, G., Klein, B., König, R., Schläpfer, M., Tunçer, B. and Buš, P. (2017). "Big Data-Informed Urban Design", Future Cities Laboratory: Indicia 01, Lars Müller Publishers, pp. 103-113.

Storper, M. and Venables, A.J. (2004). 'Buzz: face-to-face contact and the urban economy', Journal of Economic Geography 4(4): 351–70.

Subramanian R., Tunçer B. and Binder A. (2019). "Thermal comfort based performance appraisal of covered walkways in Singapore", CAADRIA 2019, pp. 805-814.

Tomarchio, L., Tunçer B., You, L. and Klein, B. (2016). "Mapping Planned and Emerging Art Places in Singapore through Social Media Feeds", 34th Annual Conference on Education and research in Computer Aided Architectural Design in Europe, eCAADe 2016, Oulu, Finland, pp. 437-446.

Tunçer, B. (2020). 'Augmenting Reality: (Big-)Data-informed Urban Design and Planning'. Architectural Design; 90(3): 52-59.

Tunçer B. and You L. (2017). "Informed Design Platform: Multi-modal Data to Support Urban Design Decision Making", 35th Annual Conference on Education and research in Computer Aided Architectural Design in Europe, eCAADe 2017, Rome, Italy, pp. 545-552.

West, G. (2017). Scale: the universal laws of growth, innovation, sustainability, and the pace of life in organisms, cities, economies, and companies. New York: Penguin Press.

Zhou Y., Lau. P.L.B., Yuen C., Tunçer B and Wilhelm E. (2018). "Understanding Urban Human Mobility through Crowdsensed Data", IEEE Communications Magazine 56(11), pp. 52-59.

Spatial Analysis

SpaceAnalysis: A Tool for Pathfinding, Visibility, and Acoustics Analyses in Generative Design Workflows

Rhys Goldstein[1], Simon Breslav[1], Kean Walmsley[1] and Azam Khan[2]

[1]Autodesk Research, Toronto, Canada, {firstname.lastname}@autodesk.com
[2]University of Toronto, Toronto, Canada, azam.khan.ca@acm.org

ABSTRACT

A growing number of architectural design efforts are making use of spatial metrics that characterize the experience of people in built environments. Metrics can make qualitative experience-related factors quantitative, and thereby assist in the exploration of a parametric or generative design space. To facilitate the adoption and development of generative design workflows, we introduce a tool called SpaceAnalysis that performs pathfinding, visibility, and acoustics analyses from which a variety of metrics can be computed. A theoretical contribution arising from this work is a new discretization method that converts 2D building geometry into a grid-based data structure supporting all three types of analyses. Experimental results show that the new method accommodates narrow corridors and small doorways with an efficient grid resolution of about 25 cm. We apply SpaceAnalysis to recreate and make publicly available a generative design workflow that was previously used to lay out a 250-person office.

Author Keywords

Grid-based methods, discretization; Dijkstra's algorithm; isovists; sound simulation, transmission-line matrix method.

ACM Classification Keywords

J.6 COMPUTER-AIDED ENGINEERING (Computer-aided design); G.2.2 DISCRETE MATHEMATICS: Graph Theory

1 INTRODUCTION

Human comfort, productivity, and wellness are greatly influenced by the design of the buildings and public spaces in which we spend the vast majority of our lives [11]. Recent investigations into the visual [16, 17] and acoustic [1, 15] implications of indoor space geometry are among many research efforts aimed at improving peoples' day-to-day experiences in built environments. Several spatial metrics characterizing a human experience—including adjacency, distraction, views to outside, and daylight, as shown in Figure 1—were incorporated by Nagy et al. [13] into a generative design workflow used to lay out a two-story office in Toronto for 250 knowledge workers. The idea behind the approach was to explore a diverse set of design options that would likely perform well from an office worker's perspective.
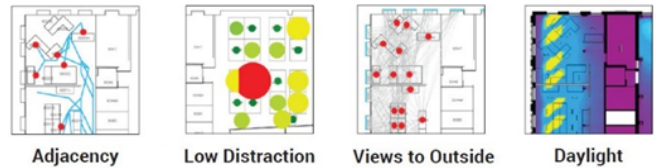
Figure 1: Four of the metrics characterizing human experience in the generative design work of Nagy et al. [13].

Our goal is to facilitate those who wish to adopt and further develop generative design workflows similar to the one pioneered for the Toronto office. To this end, we developed a tool called SpaceAnalysis that performs pathfinding, visibility, and acoustics analyses from which other experience-related metrics can be computed. Pathfinding is vital for adjacency calculations. It can also be used to identify congestion areas, a source of inconvenience and distraction yet also a potentially desirable element of a productive social environment [14]. Visibility is useful for quantifying both unwanted visual distraction and desirable views to the outside. Acoustics results enable privacy and noise-related distraction metrics.

The development of the SpaceAnalysis tool led to a number of theoretical contributions including a new discretization method and novel pathfinding and visibility algorithms inspired by Pascal's triangle. The discretization method converts line segments representing walls or other building elements into a space lattice data structure on which all three types of analyses are based. The use of a single grid-based data structure simplifies both the implementation and use of the tool, as a single geometry conversion function is all that is needed for pathfinding, visibility, and acoustics. Some precision is lost in converting building geometry into a grid, but our priority is computational speed. It is understood that for generative design workflows, the results need only be "good enough" to compare different options. Separate, precision-oriented tools can be used to fine-tune the final design.

SpaceAnalysis takes the form of a package for Dynamo [3], a visual programming environment. After presenting the discretization method and describing the three types of analyses, we present an open source recreation of the generative design project of Nagy et al. [13] that makes use of the new tool.

2 RELATED WORK

The features of SpaceAnalysis were guided by the Toronto office generative design project [13], which was itself an application of research in space layout generation, pathfinding, space syntax, and multi-objective optimization.

Space layout generation is a part of many generative design efforts in the architectural domain. Numerous layouts were generated for the Toronto office using Voronoi patterns controlled by a set of seed points. A variety of other geometry generation approaches are reviewed by Du et al. [6]. SpaceAnalysis focuses on the process that comes after space layout generation: the analysis of the generated design options.

Pathfinding is the computational process of finding shortest paths in environments such as buildings or urban neighborhoods. It enables a variety of metrics related to travel: the time and energy to walk from point A to point B, the distraction caused by people moving, unwanted crowding, but also the health and social benefits of passing through pleasant and engaging spaces. The simplest pathfinding methods are Dijkstra's algorithm [5] and the A* ("A-star") algorithm [7] applied on a regular grid of points. The method used in the Toronto project belongs to this category. Pathfinding is a heavily researched field, however, and as described in a review by Algfoor et al. [2] there are methods that use navigation meshes instead of grids, there are "any-angle" grid-based methods where paths are not constrained to grids, and there are hierarchical methods designed to efficiently handle large environments. Despite extensive research in the field, relatively little attention has been given to the discretization method by which travel barriers are converted to a grid representation. As will be shown in Section 3, the discretization method affects the spatial resolution of the grid, which in turn affects the speed of the pathfinding process.

Space Syntax [8] refers to a research community and its large collection of spatial analyses that have been pioneered and applied to architectural and urban design projects. Their contributions include isovists [18]: spatial geometries representing the area that is visible from a single point. Isovists were used in the Toronto project to evaluate workers' views to the outside and the extent to which they might be distracted by other occupants. While most visibility analyses of this nature rely on vector-based calculations, the visibility computation in SpaceAnalysis is unusual in that it employs a discrete, grid-based approach with no rays and no intersection tests.

Multi-objective optimization, the exploration of design options and their trade-offs, is the next step in a generative design workflow. An example of the technique is provided by Keough and Benjamin [12], who optimize both structural and material objectives in a context where the aesthetics of the design is of paramount concern. Fundamental to this approach is the idea that no single design performs best on every measure. It is therefore neither practical nor desirable to have the computer choose the final design. The designer still makes the decisions, but does so with access to numerous relatively well-performing generated options. Metrics computed using SpaceAnalysis can be optimized in this fashion using a separate Dynamo extension called Project Refinery [4].

Acoustics analyses were not included as part of the Toronto office project, but promise to enhance future generative design efforts by allowing noise and privacy considerations to be part of the process. One way to simulate acoustics is the Transmission-Line Matrix (TLM) method [9], a simple algorithm that propagates impulses on a grid. If cell A transmits an impulse of 1 to cell B, then in the next time step cell B radiates impulses of $\frac{1}{2}$ in three directions and reflects an impulse of $-\frac{1}{2}$ back to cell A. Huang et al. [10] provide an example of indoor sound propagation using the TLM method.

3 SPACE LATTICES

The design, implementation, and use of the SpaceAnalysis package was greatly simplified by the decision to employ a single data structure for all types of supported analyses. The downside of this approach is that no single data structure is optimal for every analysis, and as a consequence we had to adapt some of our algorithms. Overall, the benefits of using a single data structure, which we refer to as a *space lattice*, outweighed the costs. Here we describe the structure of the space lattice, the discretization method used to construct it, and an experiment evaluating the discretization method.

3.1 Structure

The space lattice data structure is a 2D rectangular 8-neighbor grid of points. An 8-neighbor grid means that each point may be connected with its 8 closest neighbors on the axes and diagonals, in contrast to a 4-neighbor grid which excludes diagonal connections. Figure 2 illustrates both types of grids.

(a) 4-neighbor grid   (b) 8-neighbor grid
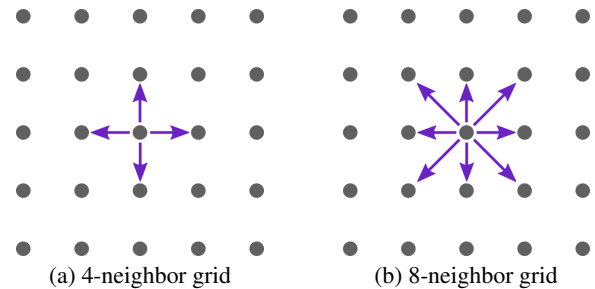
Figure 2: Illustration of (a) 4-neighbor and (b) 8-neighbor grids. The space lattice data structure of the SpaceAnalysis package is based on an 8-neighbor grid.

In the absence of barriers, all pairs of neighboring points are considered connected. In the case of pathfinding, this means a person at any point can travel in any of the 8 possible directions. Connections are bidirectional; if one can travel directly from one point A to a neighboring point B, then one can travel directly from B to A. Barriers such as walls are represented by severing connections between neighboring points.

Figure 3 shows a Dynamo node supplied by the SpaceAnalysis package to create a space lattice from an outer bounding box, a list of line segments representing barriers, and the grid resolution. The grid resolution is the spacing between neighboring grid points on the same axis—the shortest distance between points. Connections between points, shown in purple, are broken by a line at the bottom representing a barrier.
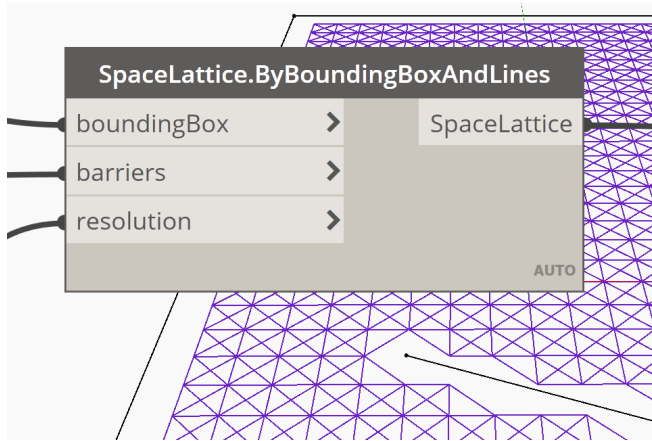
Figure 3: A Dynamo node that creates a SpaceLattice object from an outer bounding box, a list of line segments representing barriers, and the spacing between grid points.

3.2 Discretization

The line segments that represent barriers are primitive Dynamo objects and are not constrained to points on the space lattice grid. When these line segments are used to sever connections between grid points, they are effectively being converted from a continuous vector-based representation to a discrete grid-based representation. There are multiple ways to perform this discretization, and the method chosen may affect both the results of the various analyses and the appropriate choice of grid resolution. We implemented two discretization methods: a basic method and an enhanced method.

The **basic discretization method** in Figure 4 is the more obvious approach. The initial layout shows a 4x5 grid of points that may or may not be accessible depending on which connections get severed by six barriers represented by line segments (Figure 4a). The first step is to establish a cellular grid for processing the barriers (Figure 4b). In this case it is essentially the same as the original grid. The next step is to use Bresenham's standard line drawing algorithm to process the barriers by filling in grid cells as if they were pixels in an image (Figure 4c). The final step is to ensure the original points are connected to their neighbors wherever the path is not cut off by blocked cells (Figure 4d). Observe that the path from A1 to B2 is cut off by a pair of blocked cells that meet at a corner. The path from C1 to B2 touches the corner of a blocked cell, but is not cut off since one could skirt around.

The **enhanced discretization method** shown in Figure 5 differs from the basic method in two ways. First, the grid used to process the barriers is twice the resolution of the initial grid of points (Figure 5b). Second, after the barriers are processed using Bresenham's algorithm, blocked cells are expanded to fill their 4 nearest neighbors (Figure 5d), and then contracted by expanding the unblocked cells in the same manner (Figure 5e). These expansion and contraction steps fill in small gaps. The original points are then connected where possible. Notice that the path from B1 to C2 is cut off by a pair of blocked cells, but the path from C1 to B2 is traversable.
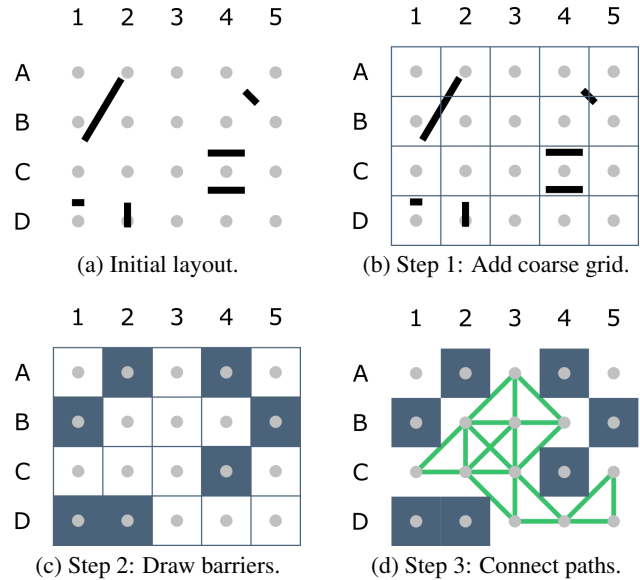
Figure 4: Basic discretization method with one resolution.
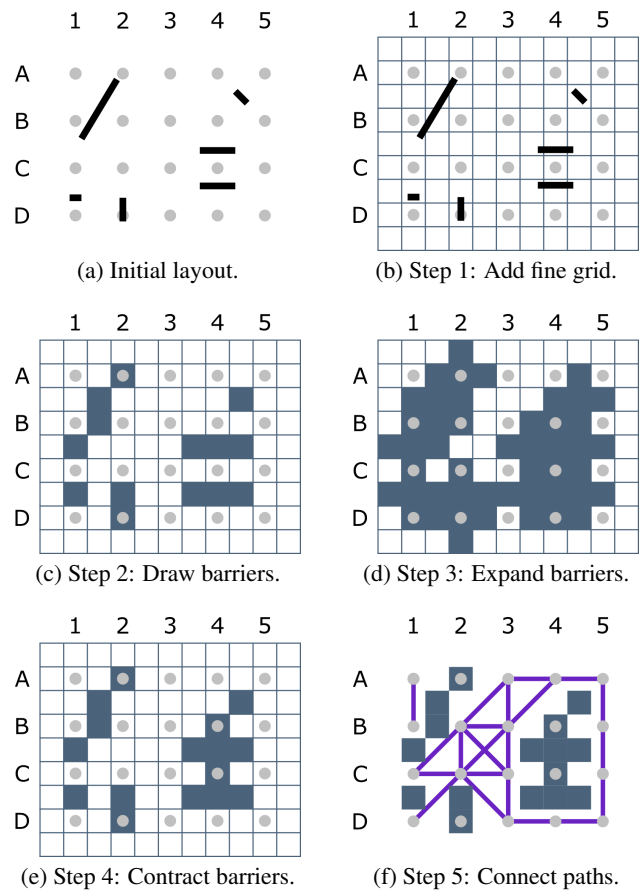
Figure 5: Enhanced discretization method where barriers are draw at a finer resolution, then expanded, then contracted.

If we compare the outcome of the basic method in Figure 4d with that of the enhanced method in Figure 5f, we find that the same initial set of barriers can generate very different space lattices. For example, point B4 is accessible only when using the basic method, and point D1 is only accessible using the enhanced method. The enhanced method produces a route that encircles the barriers on the right-hand side of the environment, whereas the basic method eliminates this route.

The SpaceAnalysis tool uses only the enhanced method. This decision was based on the intuition that the enhanced method would more faithfully capture users' intentions when applied to real-world buildings and urban environments. It is worth noting that as long as any intended gaps between barriers are large compared with the resolution of the space lattice grid, then the choice of discretization method may not matter. The initial layouts in Figures 4 and 5 are somewhat contrived in that the gaps between the barriers are inadvisably small relative to the grid spacing. Nevertheless, it is helpful to investigate how the two methods handle small openings and narrow corridors where the grid resolution becomes important.

3.3 Experiment

Here we systematically compare the basic and enhanced discretization methods using the three scenarios shown in Figure 6. In each scenario, we are testing whether one can travel from point A to point B, which involves traversing a certain type of narrow passage. In the Wall Opening Scenario, the passage is a gap or doorway in the middle of a straight wall. In the Wall Junction Scenario there is also a gap or doorway, but the passage is situated at the corner of a T-junction formed by two walls. In the Corridor Scenario, the passage is the entire region between two parallel walls, a region that contains the two points.
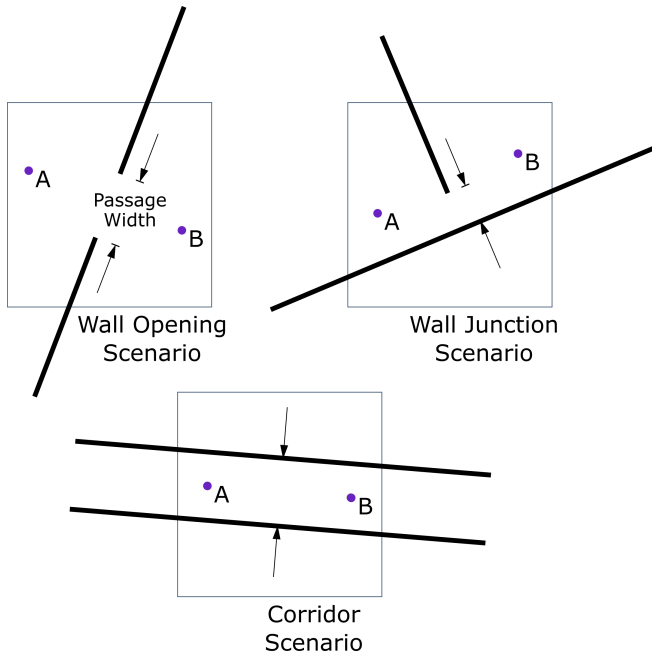
Figure 6: Scenarios used to compare discretization methods.

For each scenario, we varied the passage width from 0 to 4 at increments of 0.1, where the unit of measure is the resolution of the space lattice grid (i.e the grids shown in Figures 4a and 5a). To clarify, suppose the grid resolution is 25 cm. In that case, a passage width of 1 means the passage is 25 cm wide, and a passage width of 2 means the passage is 50 cm wide.

For each scenario and passage width, we ran 10000 tests to see whether one could travel from point A to point B. For each test, we rotated all of the geometry by a random angle, and translated it by a random displacement. Once the 10000 tests were run, we recorded the traversal rate: the fraction of tests where travel was successful. For extremely small passage widths, all routes from A to B were certain to be cut off regardless of how the scenario was rotated or translated. In those cases, all tests failed and the traversal rate was 0. For large passage widths, it was certain there would be a path from A to B through the passage. In those cases, all tests succeeded and the traversal rate was 1. But for certain passage widths in between, the orientation and placement of the geometry relative to the grid determined whether a route could be established or whether it was cut off. In those cases, the traversal rate was between 0 and 1.

The entire set of trials (3 scenarios $\times$ 41 passage widths $\times$ 10000 tests) was repeated for both the basic and enhanced methods. The results are plotted in Figure 7. For each scenario and discretization method, the curve has two notable passage widths: the minimum width where at least one test was successful; and the maximum width where all tests were successful. These width are listed in Table 1.

Scenario	Minimum Width		Maximum Width	
	Basic	**Enhanced**	**Basic**	**Enhanced**
Wall Opening	1.0	0.6	2.8	1.6
Wall Junction	0.6	0.8	3.0	2.3
Corridor	1.1	0.9	3.3	2.3

Table 1: Passage widths at which traversal is found to be possible (minimum width) and guaranteed (maximum width).

There are two criteria we should use to evaluate the methods. First, the incline of the curves in Figure 7 should be as steep as possible. Ideally the curves would be step functions, meaning that travel from A to B would depend only on the passage width and not on the rotation or translation of the geometry relative to the grid. The second criterion is that the incline should be as far to the left as possible. The further the incline is to the left, the coarser the resolution one can choose while still ensuring passage through small doorways and narrow corridors. Coarser resolutions mean fewer grid points, faster analyses, and potentially better experiences for users interactively exploring numerous designs options.

Based on these criteria, the enhanced method clearly outperforms the basic method in the Wall Junction and Corridor scenarios. The inclines are narrower and further to the left. With the enhanced method, a resolution of 25 cm essentially guarantees that travel paths will connect through passages of at least 60 cm. The basic method would require a finer resolution of 18 cm to ensure all 60 cm gaps are traversable.
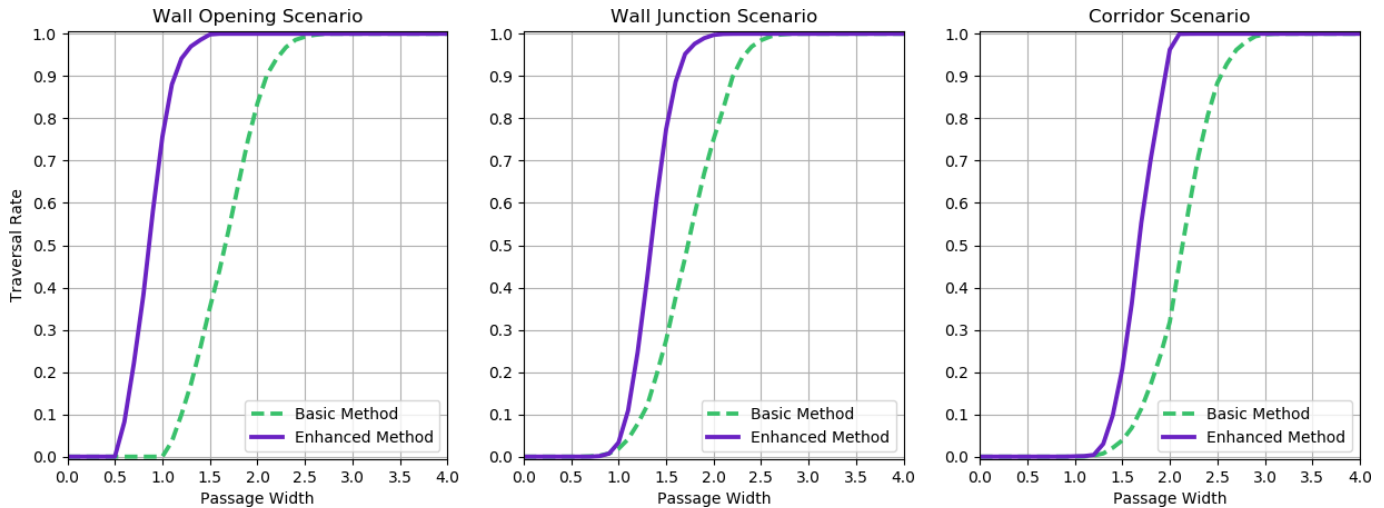
Figure 7: Traversal rates as a function of passage width for the three scenarios and two discretization methods.

The enhanced method does have a weakness, which is revealed by the Wall Opening Scenario. With a resolution of 25 cm, it is possible for a route to pass through a gap of only 15 cm. For applications involving human travel, this is too narrow. The steps of expanding and then contracting blocked cells in Figures 5d and 5e are intended to cut off these small gaps by causing a few accessible cells to become filled in. But this is only effective in situations like the Wall Junction and Corridor scenarios where the passage is bounded by the face of a wall. In the Wall Opening situation, those steps have no effect. Nevertheless, our perspective is that designers rarely create narrow vertical gaps with the intent of restricting travel. Furthermore, if the basic method were used with the necessary resolution of 18 cm or finer, then the Wall Junction tests show it is possible for a route to pass through a gap of only 11 cm. Thus we find that the experimental results validate our decision to use the enhanced method, and we can recommend a space lattice resolution of 25 cm for analyzing human travel in environments with doorways of at least 60 cm in width.

4 ANALYSES

Three areas of focus—(1) how people travel through a space and what interactions arise as a result, (2) what people can see from various locations, and (3) what people can hear—account for much of the impact that a layout option can have on people's experiences. These are the areas addressed by the by SpaceAnalysis tool's analyses: (1) pathfinding, (2) visibility, and (3) acoustics. The analysis methods are chosen with the aim of providing rapid feedback at a level of quality sufficient for comparing a diverse set of design options. Here we describe the implementation and use of each analysis.

4.1 Pathfinding

Pathfinding is a principal feature of the SpaceAnalysis package. It can be used for visualizing travel routes and computing metrics that capture the adjacency of locations, desirable and undesirable congestion, and worker distraction. In addition, the code that supports pathfinding serves also as a foundation for the visibility calculations described in Section 4.2.

Paths in SpaceAnalysis are computed by applying Dijkstra's algorithm to the 8-neighbor space lattice grid. With this approach, one starts with a single point, then computes paths inward from the closest adjacent neighbors, then computes paths inward from the next closest set of neighbors, and so on until the processed region either fills the entire lattice or encompasses some other specified point. For simplicity, steps to nearest and diagonal neighbors have distances of 5 and 7.

To take advantage of Dijkstra's algorithm, the PathField node precomputes all paths to or from a single point. Figure 8 shows an example of this node in use. It is also possible to bypass the PathField node and compute a single route directly from a SpaceLattice object and a start and end point.
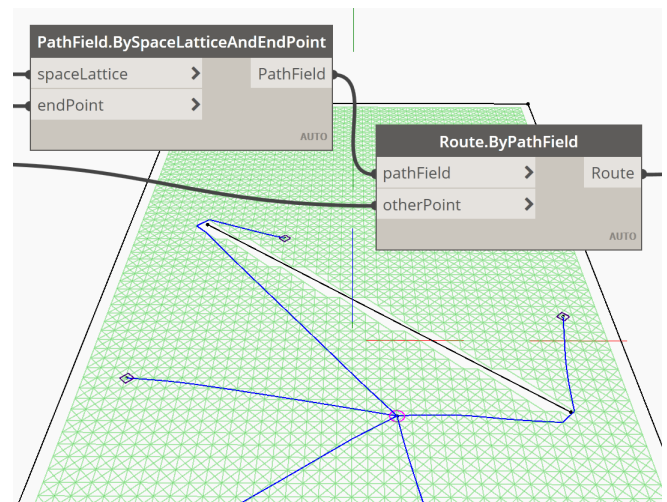
Figure 8: An example of pathfinding in SpaceAnalysis.

When finding paths on a grid, there is generally more than one shortest path between any two points. SpaceAnalysis aims to select the path that crosses empty regions with minimal deviation from the direct line-of-sight route. This is accomplished

by assigning what we call a "Pascal number" to each point on any shortest path. Inspired by Pascal's triangle, a Pascal number is the number of shortest paths that go through a point. Repeatedly selecting the highest Pascal number results in paths that tend to appear jagged, but can easily be smoothed into relatively straight curves. A "tidyIterations" parameter is provided to iteratively smooth jagged paths.

4.2 Visibility

Maximizing compelling views, minimizing visual distraction, and ensuring certain elements can be seen from certain locations are among the objectives that require visibility analysis. This feature of SpaceAnalysis takes advantage of the decision to use a single data structure for all analyses. Whereas most methods for computing visibility involve geometric projections or ray-object intersection tests, SpaceAnalysis uses a grid-based approach that re-purposes the implementation of Dijkstra's algorithm developed for pathfinding.

Starting from a single view point, shortest paths are computed outward. By "outward", we mean that the connection from any point A to a neighboring point B exists only if the trajectory of this step differs by at most 45 degrees from the vector originating at the view point and terminating at point A. In other words, all paths head as straight as possible outwards. Every visited point is assigned a Pascal number, the number of shortest paths to that point. To compute the point's visibility, its Pascal number is divided by the Pascal number that would be obtained if there were no barriers. Hence the visibility from the view point to any other point is approximated as the fraction of shortest paths that are not cut off by barriers.

A Dynamo user performs the analysis by supplying parameters to a ViewField node, and then using a VisibilityGrid node to access the resulting array of fractions approximating visibility. An example with two barriers is shown in Figure 9.
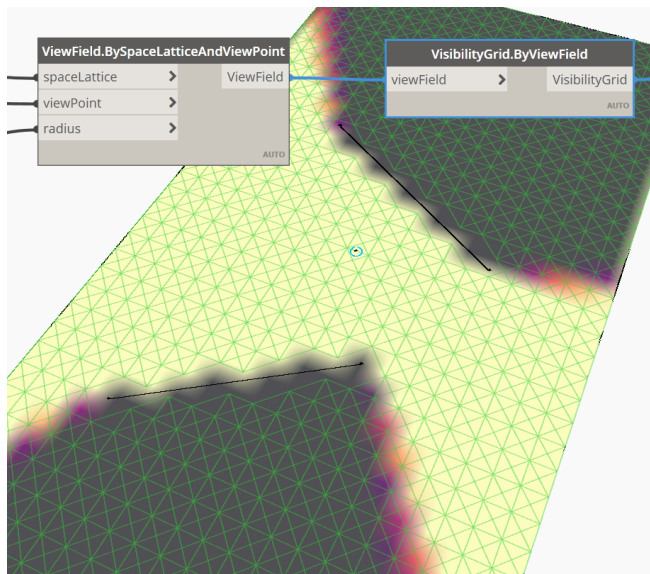
Figure 9: An example of visibility analysis. Surface colors indicate the region visible from a view point in the middle.

Two additional nodes assist with visibility analyses involving multiple points. One node creates a VisibilityGrid by taking the union of the results of multiple ViewField nodes. Technically the node reports the maximum visibility value of each grid point for all ViewFields. In essence, however, this can be interpreted as the union of a set of view fields: the region visible from any of the view points. The other node creates a VisibilityGrid from the intersection—or technically the minimum values—of multiple ViewField nodes: the region visible from all of the view points.

4.3 Acoustics

It can be desirable for a person at one location to hear sound from another, or it can be a privacy concern or source of distraction. The acoustics analysis supported by SpaceAnalysis provides a convenient way to evaluate these types of concerns. The tool simulates sound using the transmission line matrix (TLM) method described in Section 2. The basic TLM method assumes a 4-neighbor grid, which posed a challenge for us because the space lattice data structure is based on an 8-neighbor grid (see Figure 2). We therefore adapted the method to propagate sound impulses from a point to a diagonal neighbor, but only if barriers would prevent the impulse from reaching that neighbor via two axis-aligned steps (i.e. one step forward, one step sideways). The user interface for acoustic analysis is similar to that of the visibility analysis, with SoundField, AudibilityGrid, and associated union and intersection nodes. The node setup is illustrated in Figure 10.
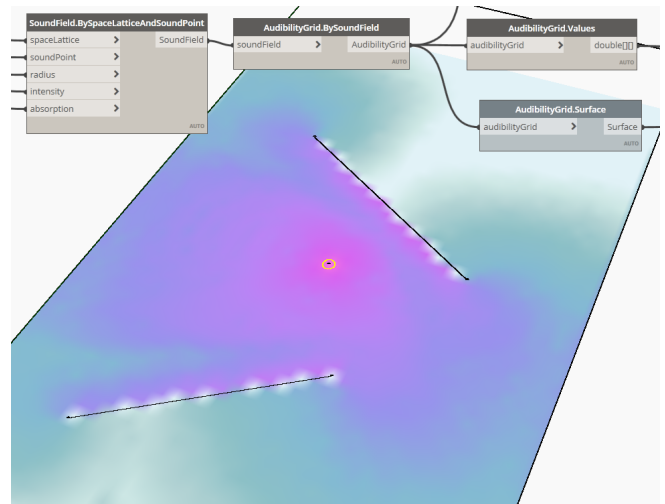
Figure 10: An example of acoustics analysis. Surface colors indicate the propagation of sound from a point source.

The sound simulation assumes single-frequency waves. This has the noticeable effect of causing interference patterns, particularly when the SoundSystem node is used to model multi-speaker arrangements. The wavelength of the sound can be obtained using the Wavelength node. In the current version of the tool, this wavelength is always $8\sqrt{2}$ times the space lattice grid resolution, which works out to 120 Hz for a resolution of 25 cm. The reason the wavelength-resolution ratio is fixed is to prevent degenerate results (small ratios) or long computation times (large ratios).
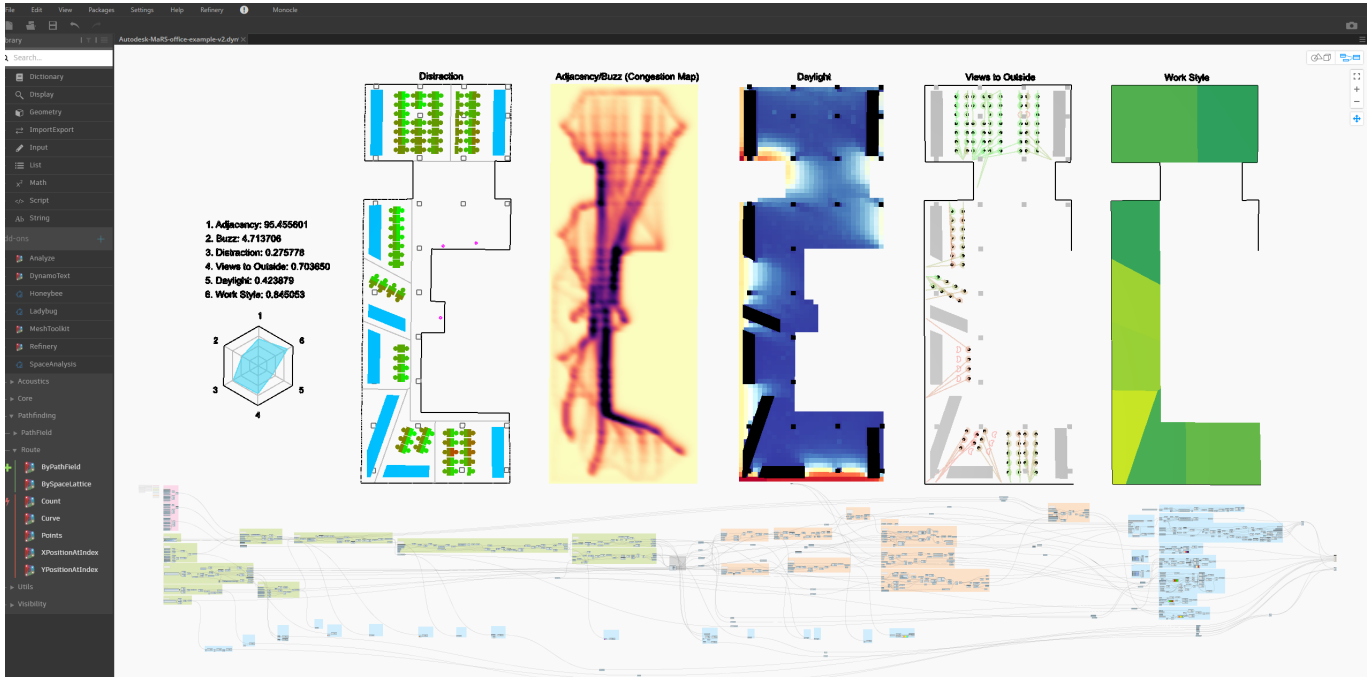
Figure 11: The recreated Toronto office generative design project. Distraction, Buzz, Daylight, Views to Outside, and Work Style metrics are visualized at the top; the Dynamo graph is shown beneath. Most of the metrics were computed using SpaceAnalysis.

5 CASE STUDY

To demonstrate features of the SpaceAnalysis package, we have performed a recreation of Autodesk's generatively designed office in Toronto's MaRS Discovery District using Dynamo. The structure of the graph can be seen in Figure 11, where nodes can roughly be grouped into four categories: input parameters, geometry system, analysis, and visualization nodes. Input parameters are located on the left side of the graph inside pink groups and are defined by slider nodes. Geometry system nodes are inside green groups, analysis or metrics related nodes are inside orange groups, and visualization related nodes are inside blue groups.

The geometry system closely follows Nagy et al. [13], where we define a number of neighborhoods using Voronoi partitioning. Input parameters control the shape and size of neighborhoods as well as amenity locations within the neighborhoods. By manipulating input parameters, the geometry system is able to generate different layouts for the office, defining the solution space for the optimization search.

Each metric is aggregated into a single number to be used by the optimizer; however for visualization purposes, metrics are calculated at the individual desk or neighborhood level as seen in Figure 11. While Daylight and Work Style metrics did not require SpaceAnalysis, the package was used for Adjacency, Buzz, Views to the Outside, and Visual Distraction.

Adjacency characterizes the distance from desks to amenities such as meeting rooms, washrooms, or egress points. Once SpaceAnalysis finds the shortest path from each desk to each amenity, we average the distances to obtain an overall score. Lower scores indicate shorter, more desirable travel paths.

The *Buzz* metric used in our graph is slightly different from the original equation in Nagy et al. [14]. We define it as

$$Buzz = \frac{\sum_i d_i H_i}{\sum_i d_i h_i} \qquad (1)$$

where i represents a given path in a set of all adjacency paths, d_i is the length of path i, $H_i = \sum_j |c_{i,j+1} - c_{i,j}|$ with $c_{i,j}$ being the level of congestion at point j along path i, and $h_i = \max_j(c_{i,j}) - \min_j(c_{i,j})$. Congestion values $c_{i,j}$ are calculated by aggregating and smoothing all travel paths. Whereas the Buzz metric defined by Nagy et al. [14] characterizes the extent to which congestion areas are dispersed throughout a space, ours attempts to quantify the degree to which travel routes intersect one another. Designs achieving higher scores are seen as promoting interaction among people who frequent different areas of a productive social environment. We leave a more extensive validation and comparison of the metric for future work.

Views to Outside assesses visibility from each desk to the outside windows, giving seats closer to the windows a higher score and the ones further away a lower one.

Visual Distraction measures the amount of visual distraction people have at their desks. Distractions can come either from people at other desks or parts of the office that have high levels of congestion.

To cut down on the size of the graph, part of the implementation of the metrics is done using Python scripts embedded in Dynamo nodes. The Python code in these nodes can invoke functions corresponding with the nodes of the SpaceAnalysis package. Please refer to the source code of the case study

for more details on how to use SpaceAnalysis from within Python scripts in Dynamo. The code can be downloaded from https://autode.sk/mars-graph. The SpaceAnalysis package itself can be installed from within Dynamo, by going to the Packages → Search for a Package menu item and searching for "SpaceAnalysis". As there have been over 2,000 downloads, the tool appears to be of interest and value to the computational design community.

In order to find the best designs according to the defined metrics, Project Refinery was used to perform multi-objective optimization. Besides optimization, Refinery supports exhaustive and randomized searching of a parameter space. After a set of design options are evaluated, the user can explore the results of the optimization using various plotting techniques such as scatter and parallel coordinate plots.

6 CONCLUSION

The types of spatial analyses, simulations, and metrics considered in this paper share a common goal of making qualitative experience-related factors quantitative, thereby assisting designers in the pursuit of more functional, productive, safe, healthy, visually compelling, and socially stimulating built environments. The presented tool contributes to that goal on both a theoretical and practical level. On the theoretical side, we have proposed a new discretization method for 2D building geometry, and shown that it accommodates narrow corridors and small doorways with an efficient grid resolution of about 25 cm. On the practical side, SpaceAnalysis supports those seeking to adopt and further pioneer generative design workflows for architecture and urban design.

The main strengths and weaknesses of the tool relate to the fact its underlying data structure is limited to two dimensions and a single scale. Future research will explore how a similar representation can support three dimensions and multiple scales, enabling movement on slopes and staircases, daylight analysis, acoustic simulations that account for room height, and projects with both building- and urban-scale elements.

REFERENCES

1. Alambeigi, P., Burry, J., and Cheng, E. Investigating the effects of the geometry on speech privacy of semi-enclosed meeting spaces. In *Proceedings of the Symposium on Simulation for Architecture and Urban Design*, SimAUD (2017).

2. Algfoor, Z. A., Sunar, M. S., and Kolivand, H. A comprehensive study on pathfinding techniques for robotics and video games. *Int. J. Comput. Games Technol. 2015* (2015).

3. Autodesk, Inc. Dynamo (**https://dynamobim.org**).

4. Autodesk, Inc. Project Refinery Beta (**https://www.autodesk.com/campaigns/refinery-beta**).

5. Dijkstra, E. W. A note on two problems in connexion with graphs. *Numerische Mathematik 1*, 1 (1959), 269–271.

6. Du, T., Turrin, M., Jansen, S. C., van den Dobbelsteen, A., and Biloria, N. A review on automatic generation of architectural space layouts with energy performance optimization. In *Proceedings of the International Conference On Building Energy & Environment*, COBEE (2018).

7. Hart, P. E., Nilsson, N. J., and Raphael, B. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics 4*, 2 (1968), 100–107.

8. Hillier, B., Leaman, A., Stansall, P., and Bedford, M. Space syntax. *Environment and Planning B: Planning and Design 3*, 2 (1976), 147–185.

9. Hoefer, W. J. R. The transmission-line matrix method—theory and applications. *IEEE Transactions on Microwave Theory and Techniques 33*, 10 (1985), 882–893.

10. Huang, P., Kapadia, M., and Badler, N. I. SPREAD: Sound propagation and perception for autonomous agents in dynamic environments. In *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA (2013), 135–144.

11. International WELL Building Institute (IWBI). WELL v2 pilot (Q3 2019), Accessed October 7, 2019 at **https://v2.wellcertified.com/v/en/overview**.

12. Keough, I., and Benjamin, D. Multi-objective optimization in architectural design. In *Proceedings of the Symposium on Simulation for Architecture and Urban Design*, SimAUD (2010).

13. Nagy, D., Lau, D., Locke, J., Stoddart, J., Villaggi, L., Wang, R., Zhao, D., and Benjamin, D. Project Discover: An application of generative design for architectural space planning. In *Proceedings of the Symposium on Simulation for Architecture and Urban Design*, SimAUD (2017).

14. Nagy, D., Villaggi, L., Stoddart, J., and Benjamin, D. The buzz metric: A graph-based method for quantifying productive congestion in generative space planning for architecture. *Technology—Architecture + Design 1*, 2 (2017), 186–195.

15. Peters, B. Integrating acoustic simulation in architectural design workflows: the FabPod meeting room prototype. *SIMULATION 91*, 9 (2015), 787–808.

16. Rockcastle, S., Ámundadóttir, M. L., and Andersen, M. A simulation-based workflow to assess human-centric daylight performance. In *Proceedings of the Symposium on Simulation for Architecture and Urban Design*, SimAUD (2017).

17. Turan, I., Reinhart, C., and Kocher, M. Evaluating spatially-distributed views in open plan work spaces. In *Proceedings of the IBPSA International Building Simulation Conference*, BS (2019).

18. Turner, A., Doxa, M., O'Sullivan, D., and Penn, A. From isovists to visibility graphs: A methodology for the analysis of architectural space. *Environment and Planning B: Planning and Design 28*, 1 (2001), 103–121.

cogARCH: Simulating Wayfinding by Architecture in Multilevel Buildings

Michal Gath-Morad[1]**, Leonel Aguilar**[1,2] **, Ruth Conroy Dalton**[4] **and Christoph Hölscher**[1,3]

[1]Chair of Cognitive Science, ETH Zurich, Switzerland, michal.gath@gess.ethz.ch
[2]Data Science, Systems and Services Laboratory, ETH Zurich, Switzerland, leonel.aguilar@gess.ethz.ch
[3]Future Cities Laboratory, ETH-Singapore Research Center, Singapore, choelsch@ethz.ch
[4]Lancaster School of Architecture, UK, r.dalton1@lancaster.ac.uk

ABSTRACT

Findings from cognitive science link the architectural complexity of multilevel buildings with occupants' difficulty in orienting and finding their way. Nevertheless, current approaches to modelling occupants' wayfinding reduce the representation of 3D multilevel buildings to isolated 2D graphs of each floor. These graphs do not take account of the interplay between agents' 3D field of view and buildings' 3D geometry, topology, or semantics, yet these are necessary to inform occupants' path differentiation during wayfinding. Instead, agents are often modeled as unbounded and rational, able to calculate complete paths towards goals that are not immediately visible using direct routing algorithms. In turn, simulated behavior in most cases is unrealistically optimal (e.g. shortest or fastest route). This gap may hinder architects' ability to foresee how their design decisions may result in suboptimal wayfinding behavior, whether intended or not. To bridge this gap, the paper presents cogARCH, a computational, agent-based simulation framework. cogARCH is grounded in research on spatial cognition and heuristic decision making to support pre-occupancy evaluation of wayfinding in multilevel buildings. To demonstrate the relevance of cogARCH to architectural design, we apply it to assess wayfinding performance across three architectural variations of a multilevel education building. Preliminary results showcase significant variability in cognitive agents' wayfinding performance between building scenarios. In contrast, behavior of shortest-path agents sampled across respective conditions displayed significantly less variance and thus failed to reflect potential effects of architectural changes applied to 3D building configuration on wayfinding behavior.

Author Keywords

Pre-occupancy simulation; Cognitive agents; Wayfinding; BIM; Agent-based simulations;

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING: I.6.5Model Development; J.6 COMPUTER-AIDED ENGINEERING; I.2 ARTIFICIAL INTELLIGENCE

1 INTRODUCTION

Within the building simulation community, research in the area of pre-occupancy simulations has mainly focused on

modelling egress evacuation[2], crowd behavior[21], and most recently building narratives[23, 17]. These studies focus on modeling collaborative behaviors to analyze macroscale patterns. As a result, less emphasis is placed on modeling low-level, microscale behaviors such as wayfinding. Instead, direct routing algorithms (e.g. A*) are used to simulate obstacle and crowd avoidance along a shortest path between origin and destination pairs. This approach overlooks agents' bounded 3D field of view and cognitive limitations observed during wayfinding in unfamiliar environments, and especially in multilevel buildings[9, 11].

Cognitive science research on wayfinding in multilevel buildings has linked architectural complexity with occupants' difficulty in orienting and finding their way[19, 9, 11]. In particular, discontinuous and discretized horizontal and vertical sightlines from decision points to such key building elements as vertical circulation, entrances, and exits is associated with occupants' having to rely on fragments of perceived or stored information to support wayfinding towards goals that are not directly visible[9, 11]. The incompleteness of this information may result in erroneous decisions, loss of orientation, confusion, and frustration[16, 9, 11].

Nevertheless, current approaches to simulating movement in buildings decompose the spatial complexity of 3D multilevel buildings into isolated 2D graph representations of each floor[14]. The main shortcoming of this approach is its inability to support 3D field of view calculations in a way that captures agents' visual perception of multilevel spaces. Representing both the navigation space and agents' vision in 2D does not allow dynamic simulation of occupants' field of view through a vertical void in the building, such as an atrium or staircase and towards potential goals at different floors. Whereas this limitation plays a minor role in wayfinding simulations set in single-level environments, it poses a barrier to wayfinding simulations set in multilevel buildings.

Moreover, architectural, semantic, and topological features linked with 3D building configurations are often not encoded in the navigation space, although these have been reported to influence occupants' ability to differentiate between path choices[5, 3]. Instead, agents are modeled as essentially unbounded and rational and given complete knowledge of the navigation environment[14]. In turn, simulated behavior in most cases is optimal (i.e. shortest or fastest route) and fails

to account for reported findings regarding loss of orientation as a function of vertical travel[19], specific wayfinding strategies observed in multilevel buildings[9], and expectations, that might be unmet, concerning the association of external cues with typical building destinations[5]. In the absence of an integrated simulation model that reflects both the cognitive and the architectural complexity of wayfinding in multilevel buildings, architects are not necessarily able to foresee the ways in which their design decisions may result in suboptimal wayfinding behavior, whether intended or not.

This paper aims to bridge this gap and make a direct contribution to the building simulation community and in particular to occupant-centered simulations. We thus present the development of cogARCH: a computational, agent-based simulation framework to simulate wayfinding in multilevel buildings. cogARCH draws on research from spatial cognition and heuristic decision making to support pre-occupancy evaluation of wayfinding during the design process. The paper's main contributions are threefold: (1) a parametric cognitive agent model grounded in theories of spatial decision making and spatial analysis, and incorporating reported wayfinding strategies in multilevel buildings; (2) a combined approach to simulating agents' visual perception in multilevel building space that integrates 2D isovist fields and a 3D field of view; and (3) an automated preprocessing pipeline to generate a 3D, hierarchical and semantically rich navigation space from a 3D Building Information Model (BIM).

To demonstrate the relevance of cogARCH to architectural design, we perform a sensitivity analysis and simulate wayfinding across three architectural variations of a multilevel education building. To highlight the need to model the interplay between 3D field-of-view and complex geometry of multilevel space, we consider architectural variations related to vertical visibility, including variation in building materials (e.g. glass vs. concrete) and floor permeability (e.g. addition of atria voids). Preliminary results showcase significant variability in cognitive agents' wayfinding performance between building scenarios. In contrast, the behavior of shortest-path agents sampled across respective settings displayed significantly lower variance. Finally, we briefly discuss planned calibration and validation experiments in virtual reality and critically assess the relevance of the proposed method to inform architectural design decisions.

2 RELEVANT STUDIES

2.1 The Process of Wayfinding in Buildings
Visual information is considered a primary input to wayfinding decisions[3, 9]. The building design may facilitate or hinder visual access in both the horizontal and vertical axes with walls, doors, choice of materials, atria, and shafts. Information perceived by occupants is used to construct internal representations both of the local choice set and of global search structures, depending on the building's configuration[22].

Various attempts have been made to quantify visibility based on 2D floorplans using space syntax methods[3, 1]. Both global and local measures of visibility can be described using these methods. For instance, axial map analysis focuses on global measures of visibility[7]. In contrast, isovists[1, 3]describe the local spatial properties of a visible area from a given observation point using viewshed polygons. Amongst 18 isovist measures[1, 3], building locations at which the area of

each isovist polygon is larger have been correlated with the locations of decision points[3].

A further link between visible information and background expectations in directed wayfinding in unfamiliar buildings is made by Frankenstein et al[5]. They conclude that occupants' background expectations regarding the association between perceived environmental cues and locations of typical building destinations predict local wayfinding choices[5]. Environmental cues are broadly defined and include configuration, materials, objects, people, and activities. Results demonstrate that building destinations such as auditorium, main exits and restrooms were associated with more central and public locations. In contrast, destinations such as rear exit, entrance to the cellar, and broom closet were associated with peripheral locations. Central locations were also rated as more public than peripheral ones. Specific environmental cues were significantly associated with some goal destinations; for instance, a chair was rated as indicative of a waiting area.

The ability to differentiate between local choices is crucial and depends on both the level of architectural differentiation between path choices[22] and occupants' reasoning on the basis of background expectation, local heuristics, and search strategies. Hölscher et al[9] provide an overview of the types of strategies that occupants employ in multilevel buildings. Novice occupants are most likely to follow a central-point strategy of finding one's way by sticking as much as possible to central and public parts of the building, even if this requires considerable detours. More complex strategies include the direction strategy: choosing routes towards the horizontal position of the goal as directly as possible[3], irrespective of level changes. In contrast, the floor strategy is applied when occupants aim to reach the floor of the destination first, irrespective of the horizontal position of the goal. Occupants familiar with the building tend to rely on either of the complex strategies. cogARCH uses these findings to provide a formal model in which expectations, strategies, and heuristics are used to simulate unaided and directed wayfinding in multilevel buildings.

2.2 Wayfinding Simulation Approaches
Various models of human navigation exist, most of which are grounded in the field of pedestrian dynamics. In this field, the use of microscopic models is becoming especially prevalent to decompose the complexity of collective emergent behavior by modeling a single individual agent. The majority of microscopic models applied to simulate navigation include three interconnected layers; a strategic layer in which agents choose between possible destinations to form an activity schedule[4], a tactical layer that focuses on route choice and path planning[8], and an operational layer that describes occupants' local 'steering behavior,' such as obstacle avoidance, speed, and acceleration[6].

The tactical layer models path planning towards a destination, essentially modeling the process of wayfinding. However, the destination is usually given to agents explicitly without them having to search for it on the basis of perceived information. In such cases, the simulation model is focused on path-planning, assumes global knowledge, and uses graph-based routing methods to calculate a walking path from an origin to a destination in accordance with some optimization criteria (e.g. distance, speed, density, etc.).

Approaches to generating navigation graphs from either 2D or 3D building geometries are various[8] and could be broadly classified into (1) semantic approaches that identify objects and relations between them; (2) topological methods that exploit visual connectivity and accessibility between semantic or geometric structures in space, such as rooms; (3) metric approaches such as occupancy grids that describe the distance and angle between locations in space; and (4) hierarchical models[12] that propose a multilevel hierarchical representation combining two or more of these approaches to model indoor navigation space.

Once a navigation graph is generated, a distinction is made between egocentric and allocentric routing algorithms. Direct routing solves the routing problem by providing an optimal path to the destination based on shortest or fastest path solutions. In contrast, iterative routing algorithms provide the next node to visit stepwise based on local optima. These algorithms are often used in egress simulations where agents are used to simulate evacuation to one or more exits. Given that agents have full knowledge of the navigation environment, the process of uninformed search using perceived environmental information is eliminated. Instead, agents calculate a shortest path towards a goal destination, regardless of its visibility. The main considerations that govern agents' movement are distance minimization, obstacle avoidance, and crowd evasion.

Cognitive approaches to simulating navigation and wayfinding have mostly focused on wayfinding aided by externalized symbolic information, such as signage, or on undirected wayfinding (i.e. exploration). Penn and Turner[15] developed an agent that perceives a space and infers spatial relations from a lookup table that encodes global measures of a visibility graph. This agent has access to cells captured in its 2D field of view and moves towards the cells with the highest visual connectivity value. Although this approach takes account of agents' bounded visibility, information provided relates to global measures describing the buildings' configuration beyond agents' natural perception. Furthermore, this model does not consider 3D visibility and overlooks the role of background expectations during directed wayfinding.

With regards to 3D visual perception, multilevel building complexity is decomposed into isolated 2D graph representations of each floor, although agents' movement is often visualized on a 3D building model. The main shortcoming of this approach is its inability to support 3D field of view calculations in a way that captures agents' visual perception in a multilevel space. A few notable exceptions simulate wayfinding using 3D visibility[18, 13], but these studies focus on wayfinding aided by signage perception using a 3D field of view.

More recently, Kielar et al developed Spice, a framework for cognitive-pedestrian modeling[10]. Spice's architecture combines the three-layered approach to pedestrian modeling with concepts from cognitive and social science. Agents' decision-making process is hierarchical and is informed by a 2D field of view and a 2D representation of an environment. Although Kielar's work is very much in-line with our aim of combining pedestrian modeling with cognitive principles, it does not take account of wayfinding towards non-visible goals and for 3D visual perception of multilevel spaces.

3 THE COGARCH SIMULATION FRAMEWORK

3.1 Overview

The development of cogARCH aims to overcome the limitations discussed above and simulate the process of unaided and directed wayfinding in unfamiliar, multilevel buildings. Achieving this requires both a cognitive agent model to represent human spatial search process and a navigation space to represent the building search space, and these must be both sufficient and complementary. To this end, cogARCH provides a computational framework in which a cognitive and parametric agent model interacts with a 3D, hierarchical and semantically rich navigation space generated from a BIM representation of a multilevel building.

cogARCH consists of three core modules: (1) spaces (2) agents, and (3) tasks. Simulation parameters are specified and controlled by (4) a batch simulator module or by a (5) GUI module that enables manual assignment of wayfinding scenarios and supports expert-based visual inspection of the simulation in real-time. Data collected during the simulation is stored in (6) a data collector module from which it is further analyzed and visualized using (7) analysis and visualization engine. The chosen development environment is Unity3D, a video-game engine that is able to support the physics and processing of complex geometries. Figure 1 provides an overview of cogARCH's framework structure.

3.2 Spaces

The navigation space in cogARCH is hierarchical. Each layer describes geometric, topological, semantic or metric properties of the building that are necessary to support agents' wayfinding decisions. Three interconnected layers are generated from the building description, imported as an IFC file from a BIM environment. The first layer consists of 3D convex zone partitions that discretize the building space into high-level path choices. Zones are created manually during the modeling of the building in a BIM environment. Zone boundaries are determined by two criteria: visibility and semantics. In contrast to purely continuous representations of navigation spaces such as those that discretize space uniformly into equal-sized cells, each of which is a unit of choice, the syntactical nature of zones allows similar and adjacent cells to be grouped semantically into meaningful units of spatial choice.

We distinguish two types of zone: (a) room-like zones, which are 'naturally' bounded in both horizontal and vertical dimensions by walls, floors, ceiling, doors, and other surfaces; and (b) open-area zones, which are spaces that are not directly bounded by a surface, but rather indirectly, by applying surface extension[16] in the horizontal and vertical dimensions. For example, a long corridor along a row of offices is subdivided into zones to reflect the variation in information continuity that results from the ability to look through office doors and towards office spaces. Accordingly, while moving from one zone to the next, new visual information is available. Transition points are thus located on the boundaries of zones. We refer to these locations as decision thresholds in which agents make local decisions and choose amongst alternative zones. Each zone stores features related to its semantics (e.g. office, public), objects (e.g. desk, chair), and connected thresholds.

The second layer of the navigation space builds upon the first to derive a threshold graph. This graph represents the topological relation between zones and thresholds. Whereas the first
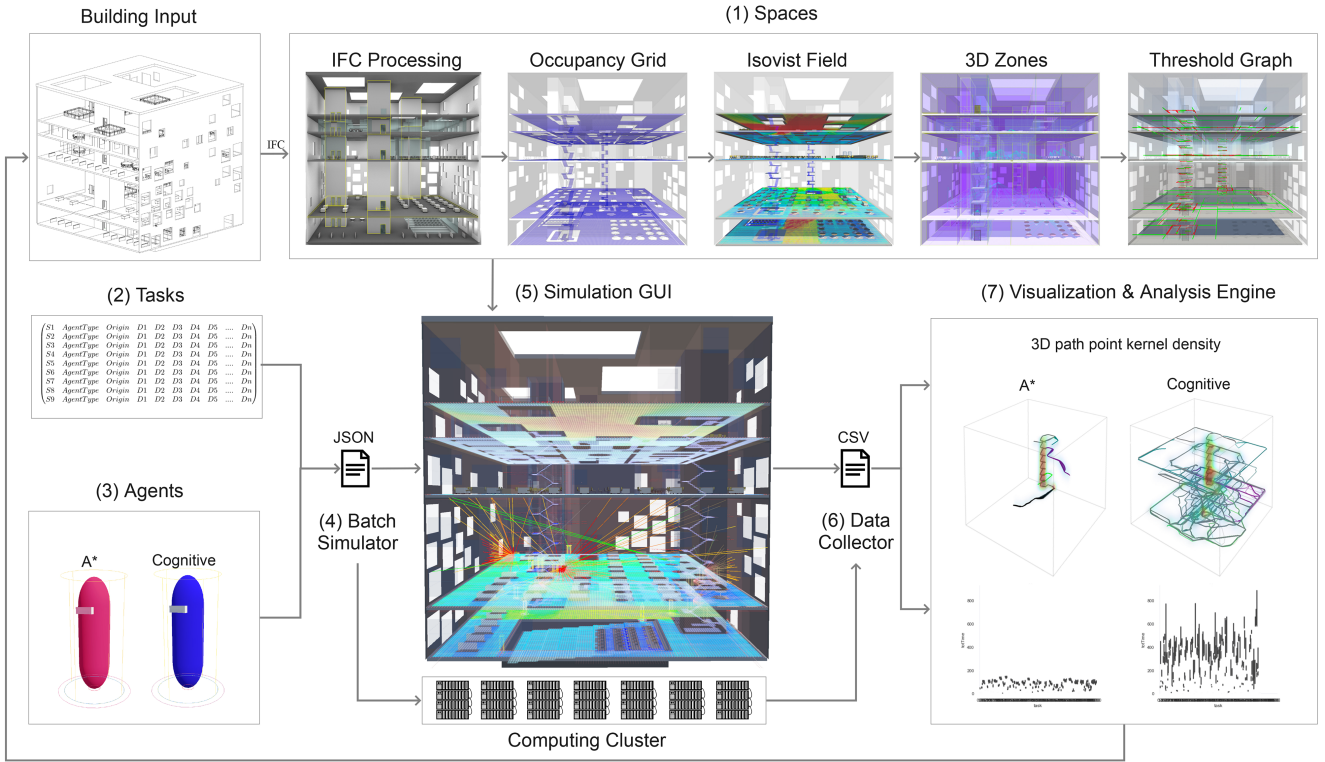
Figure 1: The cogARCH framework: A BIM model input is automatically processed into the layers of the hierarchical navigation space. Tasks and agent type are set using a batch simulator to execute simulations directly inside the GUI or remotely using a computing cluster. The simulation output is exported to a database for further analysis and visualization.

two layers represent high-level movement choices, the third layer provides low-level metric information to support operational movement (e.g. path calculation, obstacle avoidance). This layer consists of a multi-floor occupancy grid that provides a metric representation of walkable space. The occupancy grid graph is continuous along staircases without any need to link separate floors manually. The occupancy grid graph is also used to cast 2D isovists from each grid cell to form an isovist field. An area measure per isovist is calculated and mean isovist area for all cells within each zone is computed. The generation of the hierarchical navigation space is automated using a preprocessing pipeline. The output of preprocessing steps using this pipeline is visualized in Figure 1.

3.3 Agents

We assume that agents have no prior knowledge of the navigation space and that they are searching for a semantically defined destination, such as an office or entrance. Agents' template consists of the following four components: (1) a goal-specific expectation function (2) a 3D field of view (3) working memory, and (4) a state machine that represents the agent's decision-making process.

Similarly to the hierarchical navigation space described in the previous section, agents' wayfinding decisions follow a hierarchical structure from coarse to fine [9]. Three decision-making layers guide agents' wayfinding: (1) a strategic layer supports agents' ability to approximate the location of non-visible destinations; (2) a tactical layer allows agents to make local turn choices to move in the direction of the approximated destination; and (3) an operational layer allows agents to calculate an

obstacle-free shortest path towards a visible target. The three layers composing the navigation space correspond to each of these decision-making layers. The strategic layer is informed by 3D zones, the tactical layer is supported by both 3D zones and the threshold graph, and the operational layer is supported by the occupancy grid.

An agent's access to each layer of the navigation space is bounded by its 3D visual field of view. At each decision threshold, an agent casts a 3D field of view and constructs a choice set of adjacent and visible zones and a set of zones that are only visible but not reachable in one time-step. This is critical for multilevel buildings in which potential goals could be visible, for instance. through atria, but are not directly accessible (e.g. because they are at another level).

To determine the approximate location of the destination, agents' strategic decision-making layer is in charge of ranking zone alternatives for expected cues according to its expectation function [5] or chosen wayfinding strategy [9, 3]. The wayfinding literature leads us to assume that agents would strive to maximize the correspondence between features of perceived zones and their goal-specific expectations or strategies. We thus formulate both expectations and strategies as a multi-objective optimization function per goal.

Instead of evaluating all zone alternatives for all possible features, a process that would require significant cognitive resources, agents apply the notion of heuristics to differentiate between alternative zones. Heuristics are used to compare zone alternatives in a more human-like manner. Accordingly, agents prune the number of alternative zones to only those sat-

| (a) T=1 | (b) T=2 | (c) T=6 | (d) T=12 | (e) T=23 |
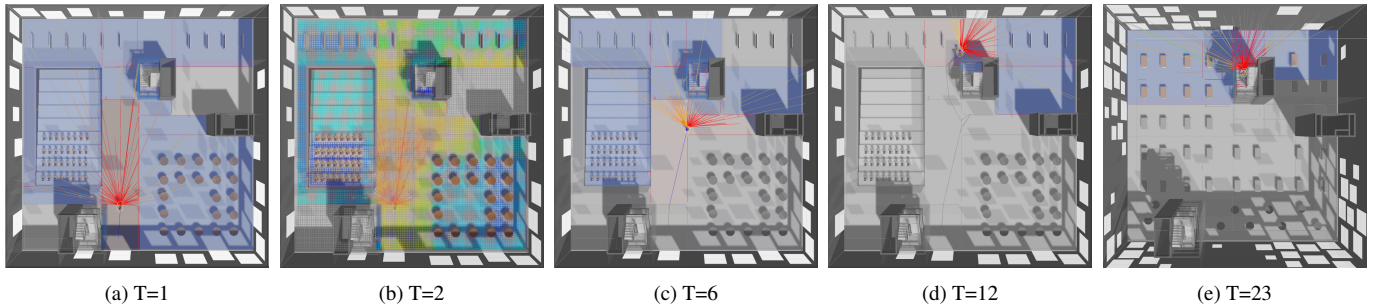
Figure 2: Selected time steps showcasing a sample task performed by a cognitive agent in building scenario 2 (Glass): At T=1, the agent casts a 3D field of view. At T=2 it evaluates zone alternatives based on destination expectations. The global goal is not visible so it ranks zones based on central point strategy using mean isovist area measure per zone. At T=6 the agent moves towards the threshold of the zone with the highest mean of isovist area. At T=12 the agent doesn't gain new visibility information and the floor strategy is triggered. The agent moves to a visible zone with vertical circulation. At T=23 the agent moves up the staircase and during so perceives its global goal.

isfying distinct criteria. This heuristic follows the elimination-by-aspects theory of decision making. This theory states that alternatives will be ignored if they do not fit within specific definitions of acceptability [20].

Algorithm 1: Agents' Decision-Making Process

Input: Agent's 2D and 3D perception of the environment
Output: Time history of agent's position and decision making

1 **while** *Global Goal Is Not Visited* **do**
2   Cast **3D field of view**;
3   Log to **Memory**;
4   Rank zones by **expectation**;
5   **if** *Global Goal is Visible* **then**
6   **if** *Path is Visible* **then** **Follow Visible Path** ;
7   **else if** *Path in Memory* **then**
8   **Follow Path** from **Memory**;
9   **else if** *Goal in same level* **then**
10   **if** *NewRandomNumber* < 0.3 **then**
11   Rank zones by **Direction Strategy**;
12   **else** Rank zones by **Floor Strategy** ;
13   **else** Rank zones by **Floor Strategy** ;
14   **else** Rank Zones by **Central Point Strategy** ;
15   Set **Local Goal Zone**;
16   Chose **threshold**;
17   Find **Shortest Path**;
18   Move to **threshold**;
19 **end**

Upon pruning the set of alternatives, agents aim to identify the highest ranked zone alternatives (i.e. that best match agents' expectations or strategy). This process is performed by ranking each zone for each relevant cue. If a zone matches all cues, the agent sets it as global goal, meaning the location of the final destination has been identified. If none of the zones meet agents' expectations, the agents rank the zones again for cues associated with the central point strategy. In this case, agents constrain their choices to public zones and aim to maximize mean isovist area per zone. Isovist area provides agents with a local measure of integration and is thus preferred over global measures such as visual graph connectivity.

Once a goal has been set, agents rely on their tactical decision-making layer to choose between visible zones that are also ad-

jacent. If an agent is at the same floor as the goal, it may choose either the central point strategy or the direction strategy to rank adjacent zones. In the direction strategy, the agent strives to minimize the horizontal angle deviation towards the goal, thus choosing the zone that best meets this condition. If the agent is at a different floor from the goal, it can either perform the direction strategy or execute the floor strategy, in which it searches for an escalator, staircase, or elevator based on its expectation of vertical circulation cues. Once an adjacent zone has been chosen, the agent relies on its operational layer to calculate a shortest path using the A* algorithm and reach a threshold at a boundary with the adjacent zone.

This decision-making process repeats at every threshold until the agent finds its destination. Visited zones and thresholds are logged to construct a graph as part of the agents' working memory and allow the agent to follow a path from memory. The current stage of the framework's development does not incorporate angular distortion of stored locations or memorability on the basis of visual and semantic saliency. Additionally, a shortest-path agent using A* algorithm is implemented to simulate a benchmark shortest-path and compare it against cognitive agents' paths.

3.4 Wayfinding Tasks
In cogARCH, wayfinding tasks consist of a single origin and either one or multiple destinations. Both origin and destination are defined by a zone rather than a point. Zones can be selected through direct interaction with the building model using the GUI or indirectly by specifying the semantics of destination zones: entrance, exit, staircase, elevator, or sitting area.

3.5 Simulation setup
Several parameters are used to define each simulation: (1) building scenario, (2) tasks, (3) type of agents performing each task, (4) the number of agents of each type to perform each task, and (5) the number of samples to be collected per scenario. Parameters can be set either manually through the GUI, or with a text file in JSON format to support headless batch execution.

4 CASE STUDY RESULTS
To demonstrate the potential of cogARCH to architectural design, we apply it to evaluate variability in wayfinding performance across three multilevel building scenarios with systematically varied architectural design. Figure 3 shows 3D models

(a) Scenario 1: Base case   (b) Scenario 2: Glass shafts   (c) Scenario 3: Atria
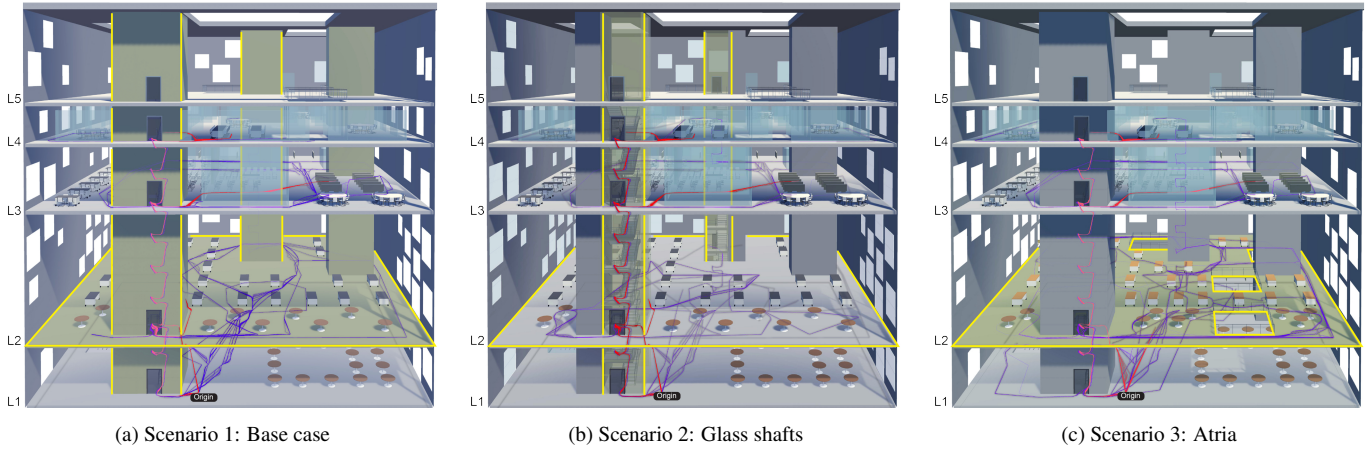
Figure 3: Three building scenarios were evaluated. Scenario 1 is the base case building and the highlighted elements are those subject to variations. In Scenario 2, we apply material variation to the circulation shaft. Instead of a concrete enclosure, as in the base case, a glass facade is set. Scenario 3 presents a variation in floor permeability in the form of atria added to the second floor. Paths of both agent type for a subset of 5 tasks are included per scenario. (—, Shortest-path agent), (—, Cognitive agent)

(a) Shortest-path agents, across scenarios   (b) Cognitive agents, Scenario 1.(Base Case)   (c) Cognitive agents, Scenario 2. (Glass shafts)   (d) Cognitive agents, Scenario 3. (Atria)
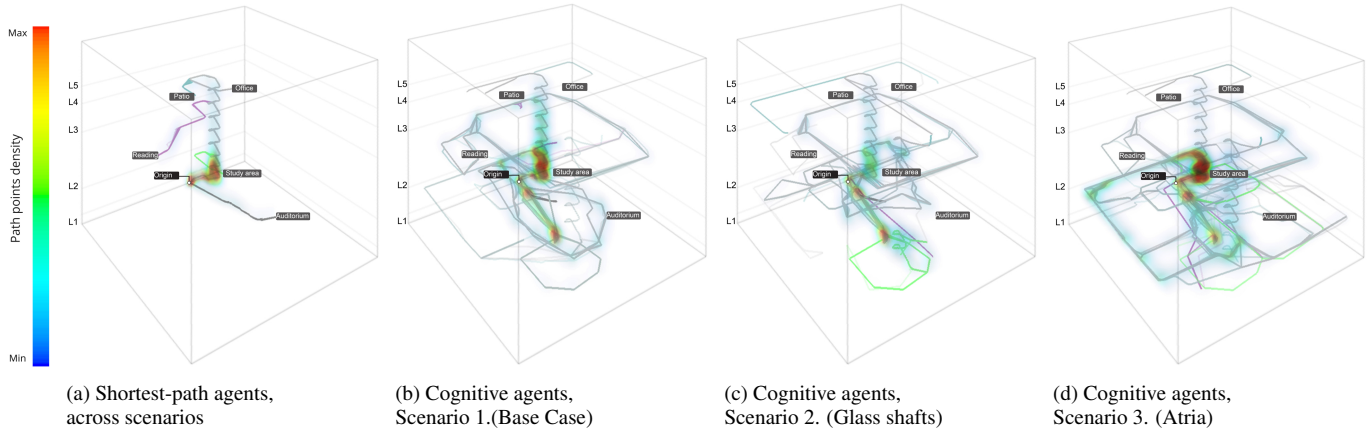
Figure 4: Overlay of paths with 3D path point kernel density across 5 tasks for both agent types for each building scenario

Building Scenario			
Agent Type	**1 (Base case)**	**2 (Glass Shaft)**	**3 (Atria)**
Cognitive	10500	10500	10500
Shortest-Path	10500	10500	10500

Table 1: The experimental setup used to conduct simulation experiments. 21 initial zones × 5 destinations × 100 samples =10500 samples in total per agent case pair. A total of 63000 samples

Model	Effect	Coef.	Std.Err.	z
1 $dist \sim buil. + type$	buil.	2.701	1.768	1.528
	type	96.682	2.887	33.485
2 $dist \sim buil.$ S(B-G)	buil.	2.737	0.131	20.926
3 $dist \sim buil.$ C(B-G)	buil.	-26.323	1.487	-17.697
4 $dist \sim buil.$ S(B-A)	buil.	-3.781	0.385	-9.825
5 $dist \sim buil.$ C(B-A)	buil.	14.343	2.190	6.549

Table 2: Results of a Mixed Linear Model regression for preliminary data

of all three building scenarios. The base case building (Scenario 1) spans 5 building floors. The building program consists of a cafeteria, auditorium, exhibition space, open-space study areas, classrooms, office spaces, meeting rooms, indoor patios, and a roof terrace. Vertical circulation between floors is enabled by two staircases and two elevators located in enclosed concrete shafts. The building has two main entrances on the ground floor, located on opposing facades. Two variations to the base case building are introduced. The first one, Scenario 2, replaces the concrete enclosure of both circulation shafts with a glass facade. The second variation, Scenario 3, introduces a series of small-scale atria on the second floor similar to those on the fifth floor.

The proposed variations aim to make wayfinding in the base case building more efficient by increasing vertical visibility between floors. Architectural changes are specifically applied to the three-dimensional configuration of the building. We do so to demonstrate the need for a cognitive agent with a 3D field of view in order to capture potential effects these variations may have on wayfinding performance and could not be observed using a shortest-path agent. Accordingly, we simulate both shortest-path and cognitive agents' wayfinding across typical tasks typical of novice occupants. These tasks consider 21 initial origin zones combined with one of 5 semantically defined destinations. A total of 63,000 samples is generated, see the experimental setup in Table 1. Monte Carlo type sim-
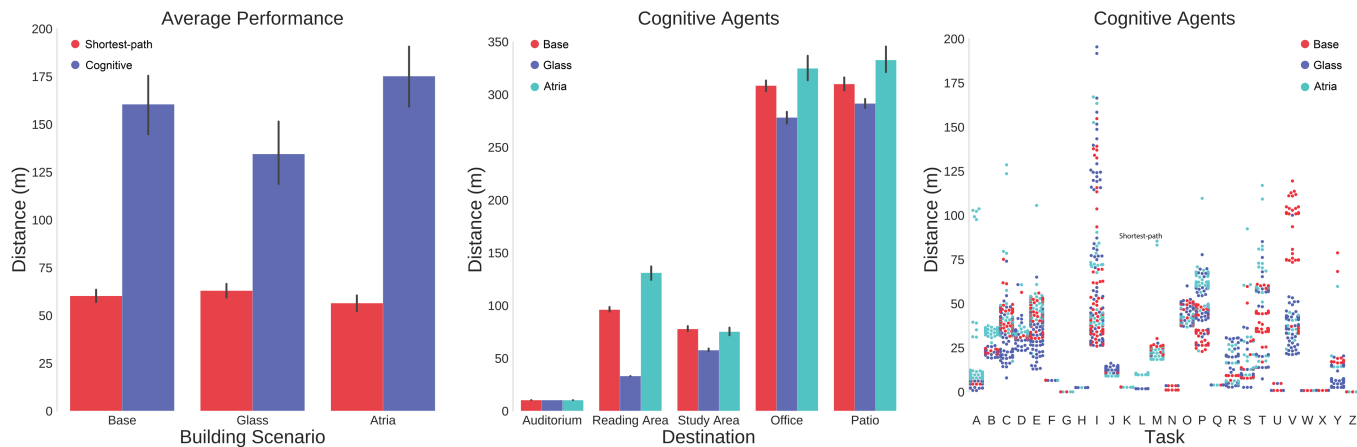
(a) Comparison of both agent types performance across building scenarios

(b) Cognitive agents performance across 5 typical wayfinding tasks across building scenarios

(c) Cognitive agents performance across 26 tasks across building scenarios

Figure 5: Comparison of simulated wayfinding performance across tasks, agents and building scenarios

ulations were performed to evaluate the distribution of the results. Randomization has been introduced to agents' initial headings. The simulations were executed using a computing cluster through singularity-based containerization.

Preliminary results regard a subset of 5 tasks that origin from the buildings' main entrance. Figure 4 shows the difference in path dispersion using kernel density estimation. It can be observed that path density in shortest-path agents (see Figure 4a), is highly localized and shows minimal variability across scenarios in comparison to cognitive agents, (see Figures 4b,4c,4d). This results is further reflected when comparing average distance performance between cognitive and shortest-path agents across those 5 tasks Figure 5a. Whereas shortest-path agents showcase minimal performance variability, cognitive agents' performance varies largely between scenarios.

Given that distance performance varies across the 5 selected tasks (see Figure 5b), we provide further evidence through 5 Mixed Effects Model Regressions (MEMR) considering the wayfinding tasks as random effects, summarized in Table2. MEMR 1. models the building and agent type as fixed effects, and is evaluated across the subset of 5 wayfinding tasks. The model results show that the major effect across the entire sample is due to the agent type (an increase of 96.68m in the path lengths). All subsequent models evaluate individually either the shortest-path agent (S) or the cognitive agent (C) against a pair of scenarios (i.e. Base case vs Glass (B-G) or Base vs Atria considering only the building as the fixed effect. MEMR 2. shows that the architectural variation (Glass) has a marginal variation in shortest-path agents' performance and a minimal increase in the path length (2.737m); In contrast, MEMR 3. shows that this effect is substantial in the case of cognitive agents with an improvement of performance of (26.323m) from the base case scenario. Although MEMR 4 leads to the same observations as MEMR 2, it is interesting to notice that the effect of the atria (scenario 3) is actually a decrease of performance (14.343m) for the cognitive agents, meaning that agents covered more distance to find their way across tasks. Although this result could be counter-intuitive, it highlights the need for simulation tools such as cogARCH to inform architectural intuition with quantifiable metrics.

The presented sub-sample intends to show the potential of cogARCH to quantify the effects of architectural variations on wayfinding performance for a small-scale scenario that directly translate to a specific use-case (i.e. novice occupants). The rest of the results showcase cogARCH's ability to process larger data sets (e.g. from the 105 possible task we selected 26 named A through Z in Figure 5c). Figure 5c showcases mean distance covered by cognitive agents across 26 tasks for all three building scenarios. As can be seen, mean distance across tasks within samples from the same building scenario, as well as between scenarios, varies. In this case the analysis is unequivocal in favour of one building scenario with respect to our design objective (i.e. more efficient wayfinding). This finding reflects the complex nature of architectural design and the need for simulation tools to reveal potential performance tradeoffs that are not necessarily intuitive to predict otherwise.

5 CONCLUSION

The ability to foresee how architectural design decisions impact occupants' wayfinding can provide ample inspiration to inform design decisions and actively facilitate *wayfinding by architecture*. Multilevel buildings, with their multiple usages, complex three-dimensional configurations, and many origin and destination pairs, provide particular challenges that current simulation tools have not yet addressed.

To bridge this gap, the paper presented cogARCH, a computational simulation framework to evaluate wayfinding performance in multilevel buildings. Preliminary results have demonstrated cogARCH's capacity to capture significant variability in wayfinding behavior given architectural design variations applied to a 3D multilevel building. This finding stands in contrast to the significantly less varied wayfinding performance of shortest-path agents sampled under the same scenarios. Extending our work and validating our findings requires a validation protocol. To that end, a series of virtual-reality and real-world experiments is underway to replicate the behaviors simulated in cogARCH. By comparing observed behavior against simulated, the predictive power of our model will be assessed and the model adapted as necessary.

Future applications of cogARCH aim to extend its cognitive agent to deal with prior knowledge of the building and pro-

33

cesses of knowledge acquisition. We further aspire to integrate other models of cognitive agents and to establish benchmark cases for assessing variability between models against real-world and lab-based data. Concurrently with extending cogA-RCH, we aspire to integrate it in simulation-based generative design workflows. Such integration would incorporate evaluation of wayfinding performance alongside other performance criteria to prune the solution space. Such an approach has the potential to introduce a much-needed occupant-centered perspective to future applications of architectural optimization and design automation.

REFERENCES

1. Benedikt, M. L. To take hold of space: isovists and isovist fields. *Environment and Planning B: Planning and design 6*, 1 (1979), 47–65.

2. Chu, M. L., Parigi, P., Law, K., and Latombe, J.-C. Modeling social behaviors in an evacuation simulator. *Computer Animation and Virtual Worlds 25*, 3-4 (2014), 373–382.

3. Dalton, R. C. The secret is to follow your nose: Route path selection and angularity. *Environment and Behavior 35*, 1 (2003), 107–131.

4. Dijkstra, J., Timmermans, H., and Jessurun, J. Modeling planned and unplanned store visits within a framework for pedestrian movement simulation. *Transportation Research Procedia 2* (2014), 559–566.

5. Frankenstein, J., Brüssow, S., Ruzzoli, F., and Hölscher, C. The language of landmarks: the role of background knowledge in indoor wayfinding. *Cognitive processing 13*, 1 (2012), 165–170.

6. Helbing, D. Agent-based modeling. In *Social self-organization*. Springer, 2012, 25–70.

7. Hillier, B., Penn, A., Hanson, J., Grajewski, T., and Xu, J. Natural movement: or, configuration and attraction in urban pedestrian movement. *Environment and Planning B: planning and design 20*, 1 (1993), 29–66.

8. Höcker, M., Berkhahn, V., Kneidl, A., Borrmann, A., and Klein, W. Graph-based approaches for simulating pedestrian dynamics in building models. *eWork and eBusiness in Architecture, Engineering and Construction* (2010), 389–394.

9. Hölscher, C., Brösamle, M., and Vrachliotis, G. Challenges in multilevel wayfinding: A case study with the space syntax technique. *Environment and Planning B: Planning and Design 39*, 1 (2012), 63–82.

10. Kielar, P. M., and Borrmann, A. Spice: a cognitive agent framework for computational crowd simulations in complex environments. *Autonomous Agents and Multi-Agent Systems 32*, 3 (2018), 387–416.

11. Kuliga, S. F., Nelligan, B., Dalton, R. C., Marchette, S., Shelton, A. L., Carlson, L., and Hölscher, C. Exploring individual differences and building complexity in wayfinding: The case of the seattle central library. *Environment and Behavior 51*, 5 (2019), 622–665.

12. Melgar, L. E. A., Lalith, M., Hori, M., Ichimura, T., and Tanaka, S. A scalable workbench for large urban area simulations, comprised of resources for behavioural models, interactions and dynamic environments. In *International Conference on Principles and Practice of Multi-Agent Systems*, Springer (2014), 166–181.

13. Motamedi, A., Wang, Z., Yabuki, N., Fukuda, T., and Michikawa, T. Signage visibility analysis and optimization system using bim-enabled virtual reality (vr) environments. *Advanced Engineering Informatics 32* (2017), 248–262.

14. Pelechano, N., and Malkawi, A. Evacuation simulation models: Challenges in modeling high rise building evacuation with cellular automata approaches. *Automation in construction 17*, 4 (2008), 377–385.

15. Penn, A., and Turner, A. Space syntax based agent simulation. Springer-Verlag, 2002.

16. Peponis, J., Wineman, J., Rashid, M., Kim, S. H., and Bafna, S. On the description of shape and spatial configuration inside buildings: convex partitions and their local properties. *Environment and Planning B: planning and design 24*, 5 (1997), 761–781.

17. Schaumann, D., Gath-Morad, M., Zinger, E., Pilosof, N. P., Sopher, H., Brodeschi, M., Date, K., and Kalay, Y. E. A computational framework to simulate human spatial behavior in built environments. In *Proceedings of the Symposium on Simulation for Architecture & Urban Design* (2016).

18. Schrom-Feiertag, H., Stubenschrott, M., Regal, G., Schrammel, J., and Settgast, V. Using cognitive agent-based simulation for the evaluation of indoor wayfinding systems. *arXiv preprint arXiv:1611.02459* (2016).

19. Soeda, M., Kushiyama, N., and Ohno, R. Wayfinding in cases with vertical motion. *Proceedings of MERA 97* (1997), 559–564.

20. Tversky, A. Elimination by aspects: A theory of choice. *Psychological review 79*, 4 (1972), 281.

21. Van Schyndel, M., Hesham, O., Wainer, G., and Malleck, B. Crowd modeling in the sun life building. In *Proceedings of the Symposium on Simulation for Architecture & Urban Design (SimAUD'16). SCS* (2016).

22. Weisman, J. Evaluating architectural legibility: Way-finding in the built environment. *Environment and behavior 13*, 2 (1981), 189–204.

23. Wurzer, G. In-process agent simulation for early stages of hospital planning. *Mathematical and Computer Modelling of Dynamical Systems 19*, 4 (2013), 331–343.

Data Modeling of Cities, a Machine Learning Application

Arman Najari[1], Diego Pajarito[2], Areti Markopoulou[3]

[1]IAAC
Barcelona, Spain
arman.najari@iaac.net

[2]IAAC
Barcelona, Spain
diego.pajarito@iaac.net

[3]IAAC
Barcelona, Spain
areti@iaac.net

ABSTRACT

Today cities are the main humankind settlement relying on intricate systems. The technological growth in the context of urban produces a vast amount of data. Analyzing and visualizing this data have provided insights into this complex environment. However, these mere approaches are inert and, due to technical constraints, hard to be integrated into urban planning.

. Consequently, we adopt a research hypothesis in which the adaptability and complexity of urban systems can be replicated, or partially replicated, by the machine-learning algorithms.

The current study aims to define a process for evaluating the capabilities for analyzing and predicting urban data with machine-learning (ML) algorithms. This process starts with constructing a data structure for inputting into the ML algorithm. Following this, different tests are applied to identify valid combinations of data and models that allow understanding of urban patterns. The bicycle-sharing system is used as a case study. The process ends discussing the options to replicate the experiment in different urban areas as well as to adapt it to different problems.

Different datasets from different cities have been explored and considered for this experiment. Across many cities' open dataset platforms, the NYD platform offered the most reliable data. From the sub-systems of the city, the mobility network was selected as a case study for exploration. More specifically, data on shared-bicycle mobility and use were selected as a result of its exciting raise as travel choices reported by New York City's department of transportation. The urban data analysis and prediction process of this research paper identifies the neighborhood as the unit of the model. Additionally, to illustrate and analyze the relationship between the selected mobility sub-system and other urban systems, contextual indicators such as land use indexes were added in the modeling process.

Despite the prediction modeling machine for the bike-sharing system coming out of this study, the main achievement is the introduction of a collection of analysis and prediction processes for urban data beyond mobility. For further advancement, implementing this approach in a different urban system and context is crucial. By this means, the replicability of the process could be evaluated and tested.

Author Keywords

Predictive urban model, Urban analytics, Big-data, Machine learning, Bicycle Mobility data

ACM Classification Keywords

I.6.1 •Computing methodologies~Machine learning~Machine learning approaches~Neural networks•Information systems~Information systems applications~Spatial-temporal systems~Location based services•Computing methodologies~Machine learning~Machine learning approaches~Classification and regression trees•Applied computing~Operations research~Transportation

1 INTRODUCTION

Advancements of Information and Communication technologies have changed the availability of data in an urban environment. Mobile networks, Internet of Things, and many other technologies are providing vast amounts of data surrounding the cities that could be analyzed besides their primary service, [1]. The accumulation of data reveals a novel perspective for the urban complexity through a new digital layer in addition to the physical one. One of the significant features of big data is the capacity to analyze the urban environment in a dynamic way through the accurate cartography of continuously changing flows [2]. However, despite these advancements, the processes of urban analysis and planning still follow the traditional model of approximate data and static cartographies [3].

Planning has been applied in contexts of limited amounts of data while relying on the analysis of the current situation and a series of future assumptions. The limited data led to an understanding of the cities constrained to human brain capacity and speed. However, nowadays, the available data in urban areas are significantly large and out of human capability to be handled. For this reason, the necessity of building a set of tools to organize and understand such a vast amount of information is crucial [4]. New tools for analysis as well as prediction-based planning are needed to understand and define the interconnections more accurately within urban environments and cities.

This research aims to discuss and evaluate the role of machine learning (ML) algorithms in the analysis and prediction of urban issues. In the context of massive

datasets available, this paper looks for an evaluation of ML capabilities oriented to urban planners that focuses on the opportunities of analysis as well as the technical advantages of the method. The contributions are a general methodology as well as a technical discussion of the analysis of mobility conditions in the city of New York. Although technological advancements of capturing and processing data can potentially capture more dimensions of the complex and dynamic nature of urban development. These advancements still lack accessibility and connections with real-case scenarios that validate the advances of recent years [5]. Urban scholars, along with other experts and professionals, just start to operate this new gear to get a better understanding and prediction capabilities for the urban environment.

2 BACKGROUND

In recent years the application of machine learning in the field of urban planning has increased exponentially [6]. Some of the urban issues that are being addressed by this method are mobility, air quality, food supplement, energy, and resources. The evolution of that research shows a promising outcome and a bright future for ML application in this scope [7]. The ML algorithms that are mainly applied in contemporary research are artificial neural networks (ANN), support vector machines, decision trees, neuro-fuzzy, and deep learning. These algorithms are usually applied for classification of information and the creation of predictive models [8]. The reviewed papers present applications of ML algorithms to understand issues commonly faced by cities and support for the discussion on the role of these algorithms in urban analysis.

Papadopoulos, for instance, introduces a machine learning approach for both modeling and proposing new policies for energy consumption in the city. In this research, the k-means algorithm was utilized to cluster the consumers based on EUI (energy use intensity). Additionally, a regression model has been applied to estimate the reduction of energy usage after implementing new policies. The author eventually redefines a new model for energy grading the buildings, which can be flexible and deployable for different contexts. This model is a clustering approach based on the gradient tree-boosting algorithm. [9].

In another study, Shafizadeh-Moghadam used a synthesis approach by building a LULC (land use and land cover) model with various algorithms to simulate urban growth. For this modeling SVG (support vector regression), ANN (Artificial neural network), RF (Random forest), CART (classification and regression trees),LR(Logistic regression), MARS (multivariate adaptive regression splines) was integrated. Eventually, by taking into consideration the neighborhood status, a bottom-up CA (Cellular automaton) algorithm is applied to filter the results of the previous mentioned six algorithms. [10].

In the public health framework, Wang explores the resident's mental health by the perception of neighborhood safety through the street view image analysis. In this research, a set of images have been selected to train a convolutional neural network algorithm to classify the objects seen in the images. The participants from the neighborhood, then, score each image. Later, an RF model was used to classify a different set of images based on the detected objects and the residents' scores [11].

Although energy and health are the fields that got most of the scholar's attention, the research on applications of these approaches in the field of mobility has increased in recent years. ities' reports focused on this subject also increased, mainly targeting urban traffic, network optimization, prediction model, and public transportation. Pan introduces a differential evolution back propagation neural network model for general traffic trend prediction [12]. Moreover, Badii, Nesi and Paoli built a prediction model for parking slots. They used Bayesian regularized ANN as well as SVR, ARIMA (Autoregressive integrated moving average), RNN (Recurrent neural network) algorithms in this study. The traffic prediction research is one of the prevailing norms in the scope of ML [13].

Furthermore, the selection of mobility type is one of the vital components of the current research. The bike-sharing system, as one of these types, starts to appear in numerous research recently. The scholars of numerous emerging research argue that the bicycle could form a sustainable future for addressing current mobility issues. Additionally, the growth of these services in many cities around the world resulted in resource savings of the mobility network [14]. While it has been much investment to optimize this system through urban data analytics, machine learning plays an essential role in this analysis.

One of the main problems, for instance, in the deck base system, is managing the demand of bikes to facilitate the service for residents. Gao and Lee tried to address this issue by developing a model to predict the future demands for each station. They combined an FCM (Fuzzy C-means) algorithm and a genetic algorithm with building an unsupervised classifier to pre-classify the data. This process is followed by an NN (neural network) to make the prediction model. In their model, variables of the seasons, day of the week and hour of the day were considered due to people's different lifestyles depending on time [15].

The other issue that raised the attention of the scholars is a comparison of this system with other available transportation means, in order to expand the service more efficiently. In a study, Zhou, Wang and Li compared the bike-sharing system and taxi in the context of Chicago. As a result, they developed a model for the estimation of travel choice tendency between the bike-sharing system and taxi. For building this model, they investigate various algorithms such as SGD (stochastic gradient descent), KNN (k-nearest neighbors), SVM (support vector machine), GNB (Gaussian naive Bayes), and DT (decision tree) [16].

Some research tries to analyse the system by considering the context and other interference variables. Gao and Lee studied the character of the land use of San Francisco with its sharing bike system. In this research, they have integrated a deep learning approach for zone clustering. They experimented with a range of two to ten hidden layers with different activation levels, to find the best fit for the data model. It is worth to mention that the traffic analysis zones applied for this study included the physical boundaries such as railway, roads, and rivers.

To summarize, Machine learning applications have grown in recent years as a new paradigm for the new challenges in urban studies. The availability of the data and accessibility of the computational tools have enhanced this growth. The topic where scholars had most of their attention to was energy, health, and mobility. The studies of transportation have raised during the last years, mostly because of the increase of data available in this scope. The bike-sharing system, as one of the components of this framework, showed a promising future to the urban mobility system. Although this system could address the future issues of transportation, it presents certain limitations in relation to processing complex types of data. The use of machine-learning algorithms not only showed promising results in understanding mobility issues but also framed mobility as a representative use case when it comes to urban analysis. Out of the extensive literature and based on the best of authors knowledge, this section only includes those with stronger links to urban design according to the scope of the paper. Moreover, state of the art research on urban issues grouped methods into clustering and prediction approaches. The general trend for ML algorithms in clustering studies are K-means, Ant colony, KNN, SVM while for the prediction investigation are ANN, RNN, Random forest, GCNN.

3 METHODOLOGY

The current research initiates with a process of vast data collection. In this stage, choosing a valid and reliable source is crucial. The availability of data is a key aspect, which is why the open data platforms became the initial data source for this research.

Open data platforms are the result of efforts to open up city data and make them accessible. In most of the cases, governments or municipalities support and maintain these platforms. For this reason, they are vigorous and concrete sources of clean and regularly updated data sets.

Urban data can be additionally obtained through the API (Application programming interface) of specific mobile applications or social media. The data from these sources are not unified, and usually require to be preprocessed for cleaning and verification [17].

In this research, geospatial data was the main interest, which applied as a filter to narrow down the field to reach the required dataset. Due to its impact on urban systems,

and as one of the spotted subjects for machine learning application [18], the transportation framework has been chosen for this case study. This type of data is fundamental in the urban study because they relate an attribute to the physicality of the city [19].

By studying the available data, the questions, which can be addressed by this method has been reviewed. This step is crucial for defining the reset of the research since the objectives will formed around this query.

After selecting the dataset, processing data is applied by exploring and validating the information through data visualization combined with statistical analysis. The next step after validation is merging transport and other urban data sets (i.e. land use, GDP) to feed a machine learning algorithm.

The unit of study chosen for correlating the geospatial datasets together was the neighborhood boundary. Thereby, every geospatial data has to have the indicator of the neighborhood in its attribute table. In this way, the geolocation of each dataset intersects with the boundary of each neighborhood while in the geospatial data with dynamic attributes, the static part requires filtering out for unitizing. Furthermore, the static part can join back to the original dataset. An example of this is the data of the weather forecast sensor systems in the city. The attribute table of this dataset contains the location, reading time, temperature, humidity, and the level of the pollution. The static part of this dataset is the location of each sensor, while the rest data are continuously altering.

The current research considers the geographic position for clustering since this helps to make a bridge for assigning the other geospatial data to the dataset. For this objective, two approaches have been explored (Fig 1). The first method is an unsupervised clustering with k-means, which works with the distance matrix of the positions. This approach has been opt because of extensive implications by other scholars in the same scope [20]. The other approach is based on using existing polygons such as traffic analysis zones, neighborhood and postal code zones to cluster the data.

To make the data understandable for the algorithm, a set of preprocessing has been achieved. This process consists of functions that translate the data in different forms. The data types are usually text (string), real numbers (float), conditional (Boolean), and natural numbers (integer). In comparison to this, the neural network requires numbers between zero and one to perform efficiently. To prepare the data for the algorithm, the text data is transformed into a set of integers. For the float data that are not in the range of zero and one, by comparing the number with the minimum and maximum, they can be mapped between the required bounds. Integer data also need to be remapped to fit the criteria of the network. The same process that transforms the float data is applicable for this case as well. In the

current research. two approaches (i.e., a node per data and separated binary data for granular data) have been tested and compared despite the impact in computing performance[21].
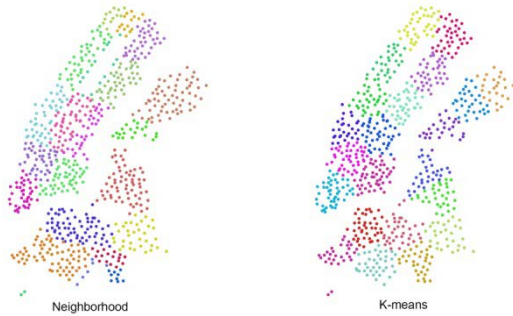
Figure 1. The clustering methods

The ANN variable and functions that have to be assigned for this research are the sampling ratio, network architecture, activation function of each layer, optimizer and epochs. Since the efficiency of this algorithm is out of the scope of this research, for most of these variables a generic choice recommended by other scholars with similar data type has been obtained [13]. However, in the case of activation function and while experimenting with gradient and binary nodes, different functions were implemented [22].

4 CASE STUDY

To examine the described methodology, a case study has been designed considering not only available geospatial open data of various urban systems but also the contextual information needed to minimize the impact of not having primary data. New York City has been selected as the case study due to its increased urban (and mobility) complexity as well as due to the availability of reliable mobility data that could be obtained to feed the developed system. By studying the mobility network in this municipality and through the detailed analysis of transportation department's reports, we can observe that the Number of travels with sharing bikes has increased in the past five years [23].

Although the bicycle is considered to be the future of urban transportation due to its small carbon footprint [25], the national household travel survey report in 2017, showcases that more than 80% of the trips in the United States were driven, and only 7% of them were provided by public transportation [26]. In the case of New York City, Faghih-Imani et al. found out that the bike-sharing system is even faster than the taxi during the weekdays [27]. Despite this result, the public usage of bike-sharing systems is lower than any other mobility choices. In the first stage of this research, the dataset of the New York bike-sharing for 2019 is downloaded from the open data provider web site [28].

As shown in the table 1, the provided dataset contains individual trips that can be divided into time, station, and user profile. Time data consists of duration, start time, and stop time. Station data contains id, name latitude, longitude for starting and ending station. On the user side, the id of the token bike, the user profile, birth year, and gender has been specified.

In the next step, summary statistics for trip duration and starting time are calculated (i.e., mean, variance, and median). This analysis is repeated for the age of the users while in the next step, a set of visualizations are developed in order to further explore the stations' data. In this stage, the stations' locations are added to the map, and the density of them is visualized. Through this visualization, the stations, which are not part of the service provider are removed. A heat map visualization depicts the trend of starting and ending stations (Fig 2). The Thousand most popular trips through the dataset are visualized and the result is overlapped with the current bike network roads (Fig 6).
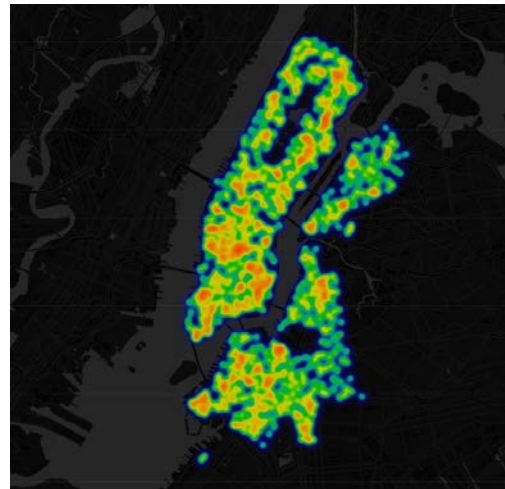
Figure 2. Density heat map of bike trips

Moreover, an analysis on the specio-temporal visualization of journeys dataset with intersection with the landuse and other terms of transportation (i.e. bus station, metro station) has been done to indicate any posible relation.

tripduration	starttime	stoptime	start station id	start station name	start station latitude	start station longitude	end station id	end station name	end station latitude	end station longitude	bikeid	usertype	birth year	gender
2287	02:45.3	40:52.9	334	W 20 St & 7 Ave	40.74239	-73.9973	3086	Graham Ave	40.71514	-73.9445	20310	Subscriber	1976	0

Table 1. The provided dataset by the bike sharing company

The next phase is dedicated to experimentation with the machine-learning algorithm. Since ML algorithm input has to be a number, the text data are transformed into numerical values. This process is followed by the application of time and boundary units to the dataset. Consequently, sets of secondary data, which had been calculated from the initial dataset, are added to the latter one. Secondary data include the number of the weekday (0-6), the hours of the day (0-23), the duration boundary (0-2), the cluster ID of the starting station, and the cluster ID of the ending station . The start time of the trip, which includes the date that trip occurred, is used for extracting the weekday variables that are used in the dataset. The days are mapped from Monday to Saturday into 0 to 6 while the duration boundary is a transformation from quantitative to qualitative indicators. For this study the trip duration stacks in three groups of low (t > 15 minute), medium (30 minute > t > 15 minute), and high (t > 30 minute). Eventually, the values are mapped through the numerical values of 0 to 2.

For clustering the starting stations based on their location, two approaches have been applied. The first method is K-mean clustering, which works with the distance between the stations. The second approach uses the neighborhood polygon as a geometrical definition for each cluster. Both instance, the k-means clustering cannot consider the approaches have their advantages and limitations. For physical barriers between stations such as the water body; however, it is a reliable unsupervised clustering method if the Number of the clusters is optimized. The clustering based on the polygon of the neighborhood does not include the barrier issue, but this is mostly related to the land and not the stations. It can be considered therefore an appropriate approach for the mediation and integration of other datasets.

Conversion of numbers to the binary is a technique that is used for increasing the number of nodes of the network. The critical aspect that should be considered here, is that this transformation changes the gradient nodes to the boolean, and such alternation in the type of data, might affect the activation function required in the application.

The next step includes the integration of a shallow neural network model in order to build a prediction model based on the provided data. As it has been discussed, there are numerous techniques available for building a prediction model. Nevertheless, this research is dedicated to the data modeling necessary for the prediction processes but not the prediction model itself.

5 RESULTS

The outcome of this study can be divided into three categories, including statistical analysis, visualization, and prediction modelling.

The figure 3 illustrates the starting time of the trip during the summer and winter months. As expected, there are two

obvious picks, which show the rushing hours at around seven in the morning, and five in the evening. The only difference between the seasons is the number of travels, which in summer are almost twice than wintertime. A sharp decline in the evening and over the night trips is observed in both seasons, although in summer this drop has smaller slop in comparison to the winter months.

The figure 4 shows the trips number in the summer and winter month period. The value is fluctuating based on the day of the week and holidays, although this pattern is much more recognizable in the winter season, compared to summer.
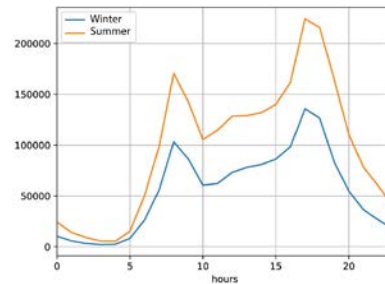
Figure 3. The number of the trips in each hour of the day in a summer and winter
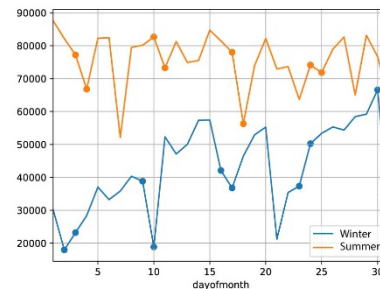
Figure 4. The number of trips in each day of month in a summer and winter
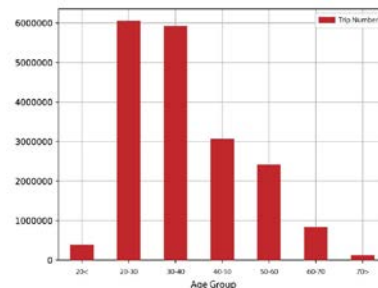
Figure 5. The age of the users and the number of the trips that they done in a summer and winter month

The figure 5 demonstrates the number of trips performed by specific ages. As it appears, most of the riders in 2019 were

in the group age of 20-30 years old. Furthermore, the frequent users belong to the group age of 20-30 and 30-40 years old.

The figure 1 illustrates the location of the station with two clustering systems, with the same amount of clusters, for comparison. We can observe the limitations of K-mean clustering in this context. Since this is an unsupervised and objective method, it cannot distinguish between urban spaces in which the bicycle trip can happen and areas in which such activity is not possible (e.g., rivers, parks or the ocean)

To this aspect, the neighborhood boundary unit definition, although being a subjective approach, solves this limitation. The figure 6 depicts that most of the trips happened in Manhattan, even though most of its roads do not have a dedicated line for a bicycle.

Figure 6. Visualization of first thousand most popular bike travels

The table 2 summarizes the results for the prediction models developed during this research. The major focus of all these models is to find a certain parameter(s), which can give a forecast for user destination. In this aspect, some of the models are built to predict the usage time of the bike based on the starting time of the trip and the station that the trip started from. Moreover, few models are trained to predict whether bike travel will go outside its star cluster or will remain in the same one. It should mention that the training is performed on a random sample with the 80% size of the available data while the validation uses 20%.

These models can be used to estimate the load of the network at a given point, which is helpful for managing the system. Furthermore, small modifications can help with estimations of the impact of, for instance, a new station into the overall network.

6 CONCLUSION

In this research, we evaluated the role of ML capabilities for analyzing and predicting urban issues. Through the exploration of prediction models for NY bike-sharing mobility sub-system, we found three main conclusions related to urban datasets, machine learning algorithms and future replication of the methodology.

Cities and urban areas have not only generated enormous amounts of data but also, an increasing number of cities, have made this data available for citizens, urban designers and researchers. Having access to such datasets enables multiple analysis like the one presented. Additionally, this variety of analysis forces experts out of information technologies to get involved in advanced analysis techniques such as ML.

This research approximates future applications of ML in which urban design can have a relevant role. In the process, researchers identified mobility, in particular bicycle sharing systems, as one of those interconnected systems in which machine learning can successfully improve urban design and urban planning activities

The particular use case involved different combinations of ANN according to the types of data inputs. Among the results obtained and presented in Table 1, relevant prediction rates (above 90%) were obtained. Considered these rates as a performance indicator, the research points to a promising field of application for ML, out of the current developments of computer vision and closer to the needs of cities.

The adopted methodology as well as the discussion after the implementation showed two main research directions. First, the replication of the experimental setup for a city with a similar dataset of a bicycle sharing system, which allows contrasting the differences with NYC. Latest work in progress has shown positive results in this direction recently. The same methodology has been successfully tested using the London bicycle trip dataset and confirmed the potential of improving the current development.

Second, an adaptation of the process to a new field of analysis (e.g., land use, land price, air pollution, or others). Either direction would provide new relevant research questions.

The results serve as an evidence for the advantages of interdisciplinary research. Especially the combination of the urban designer, serving as a frame for the selected urban issues, and the data analysts perspectives. When it comes to ML, this approach gains relevance due to the novelty and complexity of the tasks involved. The research team

Time period	Input tag	Output prediction	Training time	Accuracy	Neural Architecture
Week	Start station K-means Cluster, day of week, hours	Duration estimation	4 min	51.65%	3,3,3
Week	Start station K-means Cluster In binary, day of week, hours, duration	End station K-means Cluster in binary	5 min	53.67%	9,5,3,4
Week	Start station K-means Cluster In binary, day of week, hours, duration	End station K-means Cluster in binary	5 min	66.20%	9,5,4
Week	Start station K-means Cluster In binary, day of week, hours	End station K-means Cluster in binary	5 min	53.68%	6,5,4
Week	Start station boundary Cluster In binary ,day of week, hours,	End station boundary Cluster in binary	8 min	61.50%	9,5,7
Month	in or out of boundary, Start station boundary Cluster In binary, day of week, hours,	End station boundary Cluster in binary	12 min	55.43%	7,5,5
Month	Start station boundary Cluster In binary, day of week, hours, residential area, commercial area	in or out bool	7 min	44.74%	9,5,2
Month	Start station boundary Cluster In binary, day of week, hours, residential area, commercial area	End station boundary Cluster in binary	7 min	32.56%	9,7,5
Month	Start station boundary Cluster In binary, day of week in binary, hours in binary, residential area, commercial area	in or out of boundary	12 min	60.28%	17,10,2
Month	Start station boundary Cluster In binary, day of week in binary, hours in binary	in or out of boundary	12 min	60.47%	15,12,2
Winter	Start station boundary Cluster In binary, day of week in binary, hours in binary,staition_id in binary	End station cluster	20 min	93.43%	23,12,8,5
Summer	Start station boundary Cluster In binary, day of week in binary, hours in binary,staition_id in binary	End station cluster	20 min	92.66%	23,12,8,5
Year	Start station boundary Cluster In binary, day of week in binary, hours in binary,staition_id in binary	End station cluster	45 min	90.21%	23,12,8,5

Table 2. Predict models experimentation result.

constantly exchanged ideas and concepts while testing the different models.

Another contribution of this research is the preparation of geospatial data for machine learning applications. The clustering method has been applied in most of the scholars. Nevertheless, casting a binary node from the clustering output for the learning algorithm is a novel approach for expanding the network while reducing its complexity.

The future work coming out of this research lays on the clustering method of the bike station. For instance, the ant colony approach considering road network and trips from each station can correlate station's mutual attractiveness instead of their generic distance. Additionally, developments might come from deep learning algorithms instead of ANN since this method can observe an overall prerogative picture of the entire system.

REFERENCES

[1] V. Moustaka, A. Vakali, and L. G. Anthopoulos, "A Systematic Review for Smart City Data Analytics," vol. 51, no. 5, 2018.

[2] S. E. Bibri, *Advancing Sustainable Urbanism Processes: The Key Practical and Analytical Applications of Big Data for Urban Systems and Domains*. 2019.

[3] Y. Long and L. Liu, "Transformations of urban studies and planning in the big / open data era : a review," vol. 9832, no. August, 2016.

[4] A.-T. Umit Deniz Ulusar,Deniz Gul OzcanFadi, "Open Source Tools for Machine Learning with Big Data in Smart Cities," *Smart Cities Performability, Cogn. Secur.*, pp. 153–168, 2019.

[5] R. M. Anandakumar, HaldoraiArulmurugan, "Machine Learning and Big Data for Smart Generation," *Comput. Commun. Syst. Urban Dev.*, pp. 185–203, 2019.

[6] S. Nosratabadi, A. Mosavi, R. Keivani, S. Ardabili, and F. Aram, "State of the art survey of deep learning and machine learning models for smart cities and urban sustainability," no. August, 2019.

[7] J. Liu, T. Li, P. Xie, S. Du, F. Teng, and X. Yang, "Urban big data fusion based on deep learning : An overview," *Inf. Fusion*, vol. 53, no. June 2019, pp. 123–133, 2020.

[8] J. T. de Souza, A. C. de Francisco, C. M. Piekarski, and G. F. do Prado, "Data mining and machine learning to promote smart cities: A systematic review from 2000 to 2018," *Sustain.*, vol. 11, no. 4, 2019.

[9] Sokratis Papadopoulos, "MACHINE LEARNING

APPLICATIONS TO ADVANCE THE FIELD OF URBAN ENERGY DYNAMICS," New york university, 2019.

[10] H. Shafizadeh-Moghadam, A. Asghari, A. Tayyebi, and M. Taleai, "Coupling machine learning, tree-based and statistical models with cellular automata to simulate urban growth," *Comput. Environ. Urban Syst.*, vol. 64, pp. 297–308, 2017.

[11] R. Wang *et al.*, "Health & Place Using street view data and machine learning to assess how perception of neighborhood safety in fl uences urban residents ' mental health," vol. 59, no. July, 2019.

[12] X. Pan, W. Zhou, Y. Lu, and N. Sun, "Prediction of network traffic of smart cities based on DE-BP neural network," *IEEE Access*, vol. 7, pp. 55807–55816, 2019.

[13] C. Badii, P. Nesi, and I. Paoli, "Predicting Available Parking Slots on Critical and Regular Services by Exploiting a Range of Open Data," *IEEE Access*, vol. 6, pp. 44059–44071, 2018.

[14] D. Namiot and M. Sneps-sneppe, "On Bikes in Smart Cities 1," vol. 53, no. 1, pp. 63–71, 2019.

[15] X. Gao and G. M. Lee, "Computers & Industrial Engineering Moment-based rental prediction for bicycle-sharing transportation systems using a hybrid genetic algorithm and machine learning," *Comput. Ind. Eng.*, vol. 128, no. December 2018, pp. 60–69, 2019.

[16] X. Zhou, M. Wang, and D. Li, "Bike-sharing or taxi？ Modeling the choices of travel mode in Chicago using machine learning," *J. Transp. Geogr.*, vol. 79, no. June 2018, p. 102479, 2019.

[17] A. Psyllidis, A. Bozzon, S. Bocconi, and C. T. Bolivar, "A Platform for Urban Analytics and Semantic Data Integration in City Planning," pp. 21–36, 2015.

[18] Z. Liu *et al.*, "A tailored machine learning approach for urban transport network flow estimation," *Transp. Res. Part C*, vol. 108, no. December 2018, pp. 130–150, 2019.

[19] J. Lee, M. Kang, J. Lee, and M. Kang, "Geospatial Big Data： Challenges and Opportunities Procedia Computer Science Geospatial Big Data： Challenges and Opportunities," *Big Data Res.*, 2015.

[20] L. Caggiani, R. Camporeale, M. Marinelli, and M. Ottomanelli, "User satisfaction based model for resource allocation in bike-sharing systems," *Transp. Policy*, vol. 80, no. March, pp. 117–126, 2019.

[21] C. Shi and N. Lin, "SURVEY ON BINARY NEURAL NETWORKS A," pp. 1–17, 2019.

[22] Charu C. Aggarwal, *Machine Learning with Shallow Neural Networks*. Springer London, 2018.

[23] "New York City Mobility Report," 2019.

[24] N. Mcguckin, A. Fucci, and D. E. Jenkins, "Trends in travel behavior," pp. 1969–2017, 2018.

[25] A. Nikolaeva, P. Adey, T. Cresswell, J. Y. Lee, A. Nóvoa, and C. Temenos, "Commoning mobility: Towards a new politics of mobility transitions," *Trans. Inst. Br. Geogr.*, vol. 44, no. 2, pp. 346–360, 2019.

[26] North American Cities and Transit Aencies, "Bike Share in the U . S .: 2017 More Systems , More Cities , More Bikes , More Companies," 2017.

[27] A. Faghih-Imani, S. Anowar, E. J. Miller, and N. Eluru, "Hail a cab or ride a bike? A travel time comparison of taxi and bicycle-sharing systems in New York City," *Transp. Res. Part A Policy Pract.*, vol. 101, pp. 11–21, 2017.

[28] I. Motivate International, "Citi Bike System Data," 2013. [Online]. Available: https://www.citibikenyc.com/system-data.

From Semantic-Based Rule Checking to Simulation-Powered Emergency Egress Analytics

Muhammad Usman[1], Yi Mao[2], Davide Schaumann[3], Petros Faloutsos[15] and Mubbasir Kapadia[4]

[1]York University, Toronto, Canada, {usman; pfal}@eecs.yorku.ca
[2]University of Science and Technology of China, Hefei, China, mypeter@mail.ustc.edu.cn
[3]Jacobs Technion-Cornell Institute, New York, USA, davide.schaumann@cornell.edu
[4]Rutgers University, New Jersey, USA, mk1353@cs.rutgers.edu
[5]UHN - Toronto Rehabilitation Institute, Toronto, Canada

ABSTRACT

Complying with the *International Building Code* (IBC) is essential in architectural design. Computer-Aided Design (CAD) tools have been developed to perform BIM-based (Building Information Modeling) rule checking for fire egress scenarios. Such tools are helpful in identifying design flaws for potential egress evacuations. However, these rule checking tools consider static building design features without considering the space semantics (i.e. what the space is designed for), and more importantly, time-based dynamics to understand how the design would impact the behavior of human inhabitants. As a result, a building design may pose threats for human safety. This work aims to develop a computational tool to perform an automatic semantic-based simulation-powered rule checking for fire egress scenarios in architectural design. Beyond static geometric information, the tool considers space semantics and utilizes dynamic crowd simulations to predict the impact of a design-space with respect to potential egress evacuation on the safety of human inhabitants. A case-study showcases how simulation-powered semantic-based egress analyses can identify problematic areas for potential safety hazard and provide time-based dynamics of potential building–occupant interactions.

Author Keywords

crowd simulation; space semantics; egress planning, building code; means of egress; spatial analytics

ACM Classification Keywords

I.6.3 SIMULATION AND MODELING : Applications; I.6.5 SIMULATION AND MODELING : Model Development; I.6.8 SIMULATION AND MODELING : Types of Simulation; J.5 ARTS AND HUMANITIES: Architecture; J.6 COMPUTER-AIDED ENGINEERING.

1 INTRODUCTION

Understanding the extent to which a design solution supports the safe living and working conditions for the inhabitants is

a critical aspect in architectural design. Building Information Modeling (BIM) and Computer-Aided Design (CAD) tools enable the generation of computational building models that are amenable to evaluation of different functional specifications, including the International Building Code (IBC) for built designs [7].

An automated rule-checking process generally validates IBC by applying certain rules and constraints on the building specifications without changing the design itself, and reports the outcome as "pass" if they comply by the codes or "fail" otherwise.

Some computational tools to perform rule-checking exist as standalone software [24, 19], while others are integrated into commercial architecture design platforms [20]. A large research focus has been on rules for planning egress [1]. However, current workflows for egress planning rely mostly on static building specifications (e.g. geometry information), do not take into account space semantics and only consider a static "distance" measure while planning the egress routes.

We posit that a simulation-powered rule checking with additional space usage information (e.g. semantics) can extend current static rule-checking approaches, and thus help architects design safer environments. To this end, a computational tool is proposed to perform an automatic semantic-based simulation-guided rule checking of IBC rules for fire egress of a building. Unlike standard egress planning workflows, it also uses crowd simulations which yield a time-based representation of dynamic behaviors of occupants in a built-environment (e.g. evacuation times). We showcase the limitation of standard egress rule-checking workflows in favor of a semantic-based simulation-guided rule checker with a help of a case study.

2 RELATED WORK

In 1994, the International Code Council (ICC) was formed and ever since, ICC has been publishing International Building Code (IBC) to assure fire prevention and human safety in buildings. With the advent of CAD and BIM, re-
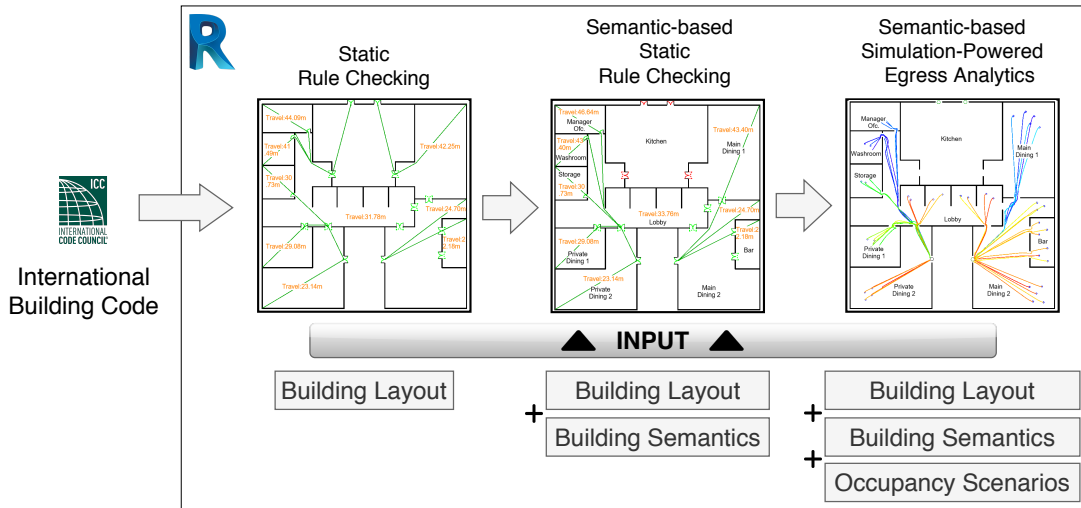
Figure 1: Overview – From static rule checking to semantic-based simulation-powered egress analytics.

searchers started developing computational approaches to validate these codes using automated workflows.

2.1 Automated IBC Rule-Checking

Some of early works in automatic IBC rule checking for egress include Solibri Model Checker [24], EDM ModelChecker [25], SMARTcodes [5], CORENET e-PlanCheck [19]. These tools have been surveyed in [6]. In the recent past, a collaborative framework is presented for all the participants to keep track of building codes during design process of a project using BIM and to assist in alternative building code-based designs [20]. A rule-based platform is presented to analyze BIM models to detect fall-related safety hazards and make preventive recommendations [30]. And more recently, a BIM-based rule checking interface is developed to logically track the influence of design adjustments against building code violation to prevent further violations [8]. Few efforts have been made to translate semantic information from IBC regulations to logic clauses used for automated rule checking [17, 29].

These tools and platforms are useful to validate IBC rules and identify design flaws for potential safety hazards, but they solely rely on static geometric components of architectural structures, ignoring space semantics i.e. how the building space is going to be used by the people, not the semantics from IBC regulations, and potential human–building interactions.

2.2 Simulation-guided Analysis of Environments

Several models have been proposed to mimic human movements in virtual environments [18, 21]. Dynamic aspects of human behavior can be analyzed using crowd simulations which provide a time-based representation of architectural structure in-use by the potential occupants. Such analysis, however, requires a user to define certain crowd specifications like behaviors, walking velocities, target positions, etc., for the given building model.

A rich literature exists on using crowd simulations in improving and analyzing architectural and building designs. The position of architectural design elements is optimized by means of simulation to improve pedestrian flow [2, 12, 11]. An optimization framework is proposed to automatically find optimal egress routes for a building evacuation using crowd simulations [3]. An interactive tool is presented to optimize obstacles in a virtual environment w.r.t. walking flows, but only supports small-scale and less-complex architectural buildings [13].

More recently, an automated parameter exploration framework is presented to jointly explore architectural and crowd behavioral parameters as templates for efficient building evacuations [27]. An event-based, model-independent and general-purpose framework is proposed for joint simulation of human–building interactions. The modeling and visualization of footstep noise simulation is shown in the case study [22]. A framework for parametric modeling is presented to jointly model a building and the bounds of its permissible layout alterations, occupants and activities the occupants engages-in, using a node-based representation widely used in main stream parametric design workflows [28]. An interactive computational tool for crowd-aware analyses of buildings and environments is proposed. Spatial visualization from the simulations are shown as walking trajectories of occupants [26].

2.3 Our Contribution

In this work we demonstrate that it is important to account for space semantics (e.g. how the architectural structures are to be used by its future inhabitants) and dynamic human-behavioral aspects in validating and planning egress plans, to ensure enhanced human safety and generate more safer building designs. We integrate our workflow into a professional industry platform, Autodesk Revit ®. A case study is presented to showcase the effectiveness of the proposed approach.

3 SYSTEM ARCHITECTURE
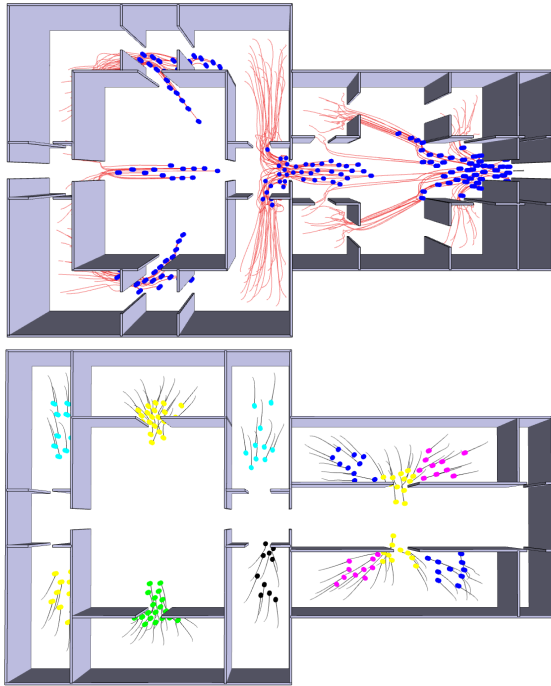
Figure 2: Sample snapshots of crowd simulations during egress at $t = 234^{th}$ (Top) and $t = 16^{th}$ (Bottom) frame respectively. The circular disks represent virtual occupants, whereas *Red* and *Black* line segments represent crowd trajectories.

We present an interactive computational tool which allows designers and architects to evaluate International Building Code for *Means of Egress*, read and/or assign semantics to the building (i.e. labeling a BIM model), compute and visualize an egress plan for emergency evacuation, analyze the travel distances and evacuation times for egress routes, and analyze dynamics of potential occupant–building interactions for different crowd occupancies. Figure 1 shows an overview of the proposed system.

3.1 International Building Code

Validation of several building design rules (e.g. geometric rules) are implemented in accordance with International Building Code (IBC) 2018 developed by *International Code Council* [15]. In principle, these rules must be adopted as baseline building standards while designing an architectural environment. One of ideas behind adopting IBC is the safety concern of potential occupants of the built-environment. Since a significant amount of IBC rules deal with fire emergency situations, in this work, we primarily focused on selected IBC rules [16] which are involved in designing the *Means of Egress* to achieve timely evacuations for the safety of potential occupants. These include rules for "ceilings" (e.g. vertical rise, headroom of protruding objects from ceilings) "doors" (e.g. width and height of a door leaf, minimum/maximum door opening angles), "ramps" (e.g. slope, vertical rise, width of a ramp), "egress paths" in emergent evacuation of a building (e.g. travel distances, permissible

and prohibited room types for egress) , and "corridors" (e.g. fire-resistance, width, capacity) We summarize these considered rules in Table 1.

3.2 Space Semantics

The word semantic is widely known as the study of "meaning". In the domain of architecture and building design, semantics are means to understand the built space, allowing designers and architects to consider for potential building's space usage and accordingly account for the foreseen behavioral properties of the built-environment. An example of building space semantics for a house would be: "bedroom area" – a space where people can sleep, "kitchen" – a space to prepare food, "laundry room" – a space to do the laundry, a "washroom", etc.

Our tool requires designers to input semantic information directly into the BIM model, and then it automatically extracts this information to validate the integrated IBC rules. If the semantics are missing in a model, the tool notifies the user about their absence, so that the designer can add the missing semantics.

Egress routes are computed using shortest paths on graph (G) of the building environment where nodes (N) and edges (E) represent rooms and doors in the graph respectively. When semantics are added, the graph is updated. If the semantic assigned to a room is prohibited to pass through during egress, all the edges connected to the node of that room becomes untraversable in the graph, and hence, they do not participate in egress routes.

3.3 Simulation-Powered Crowd Movements

We model the potential occupant–building interactions using crowd simulations which yield a time-based representation of dynamic behaviors of occupants in a built-environment. Such analysis requires a specification of the *building layout* (e.g. geometry information), the *occupants* to populate the environment (e.g. spawn regions for the crowd, one or more target destinations, walking velocities, etc.) and the *activities* they engage in (in this case, an emergency evacuation). SteerSuite [23], an open framework for crowd simulation and optimization techniques, is used as a platform to run simulations. We use a social-forces based steering technique for simulating virtual crowds [14]. The tool itself, however, is not bound to just use a single kind of crowd steering model. More steering techniques can be integrated.

For spawning virtual occupants (e.g. crowds) in a built-environment, our tool automatically iterates over all the closed-spaces (e.g. rooms) in the given BIM model. It then maps the number of virtual occupants to be spawn within these closed spaces using a standard qualitative classification, Level of Service (LoS) [10], in accordance to available area within these spaces. There are six Levels of Service. LoS has been used in traffic and crowd simulations to measure the quality of movement flow both for automotive and pedestrian applications. LoS classes are generally given a grade level (from A-F), which are summarized in Table 2. We further categorize these classes into three levels: *LoS Low* – it is an

Category	Section	Description
Ceilings	1003	*"the means of egress shall have a ceiling height of not less than 7 feet 6 inches above the finished floor"* and *"protruding objects are permitted to extend below the minimum ceiling height...where a minimum headroom of 80 inches is provided over any circulation paths"*
Doors	1010	*"a door should provide a minimum clear opening width of 32 inches"*, *"maximum width of a swinging door leaf shall be 48 inches nominal"* and *"minimum clear opening height of doors shall be not less than 80 inches"*
Ramps	1012	*"ramps used as part of a means of egress shall have a running slope not steeper than one uni t vertical in 12 units horizontal"*, *"the rise for any ramp run shall be 30 inches maximum"* and *"the clear width of a ramp between handrails, if provided, or other permissible projections shall be 36 inches minimum"*
Egress Paths	1016/1017	*"egress shall not pass through kitchens, storage rooms, closets or spaces used for similar purposes"* and *"exit access travel distance shall be measured from the most remote point of each room, area or space"*
Travel Distance	1017	for most building types *"exit access travel distance shall not exceed the value of 200 ft without sprinkler system"* and *"250 ft with sprinkler system" installed*
Corridors	1020	*"the minimum corridor width with an occupant load of less than 50 shall be not less than 44 incles"*

Table 1: Summary of rules for "Means of Egress". Extracted from Chapter 10 of *International Building Code*, 2018/19. Developed by *International Code Council*.

average of grade A & B (a sparse crowd), *LoS Medium* – average of grade C & D (a moderate crowd) and *LoS High* – average of grade E & F (a dense crowd). Figure 2 shows two sample simulation snapshots for egress.

The crowd simulations yield not just the traveled distances from the starting position of virtual occupants to the nearest exit, but also evacuation times, for individual occupants as well as average over the whole population. Therefore, unlike static egress planning which only relies on the distance information, with simulation-powered crowd movements, more efficient egress planning can be done. Hence, more safe building layouts can be designed. Using the tool, users can also compute an average *Egress Flow*, which represents the rate at which virtual occupants vacate the environment (the higher the better).

4 CASE STUDY
This case study first investigates the normative egress-planning workflow in accordance to IBC rules (e.g. Section 1016/1017, Chapter 10, IBC 2018) which illustrate the requirements for egress paths in case of an emergency evacuation of a building. It then highlights the limitation of standard workflow which does not take into account space semantics and relies on a static distance measure alone, and shows how crowd simulations can provide time-based representation of potential occupant–building interactions (e.g. provide both distance and time measures) which may be of assistance in designing more safer egress plans. Finally, it demonstrates

LoS from [9, 11]	Crowd Density	Selected Levels
A	≤ 0.27	*Low*
B	0.43 to 0.31	
C	0.72 to 0.43	*Medium*
D	1.08 to 0.72	
E	2.17 to 1.08	*High*
F	≥ 2.17	

Table 2: Level of Service (LoS) values and crowd density mapping from [9, 11]. Density is measured in occupants per square meter.

that using simulations users can account for different crowd occupancy in different areas of a building design to understand the dynamics of potential occupant–building interactions for different occupancy levels.

Environment. A variant of a real-world restaurant style environment is used in the case study, Figure 3. The environment consist of 10 room spaces and 4 external exits. Overall area of the environment is approximately 7545 meters. Potential usage of the design space is also shown as room labels (i.e. *Space Semantics*).

4.1 Static Analysis with No Building Semantics
First, we compute an emergency egress plan for the built-environment using normative workflow. In normative work-
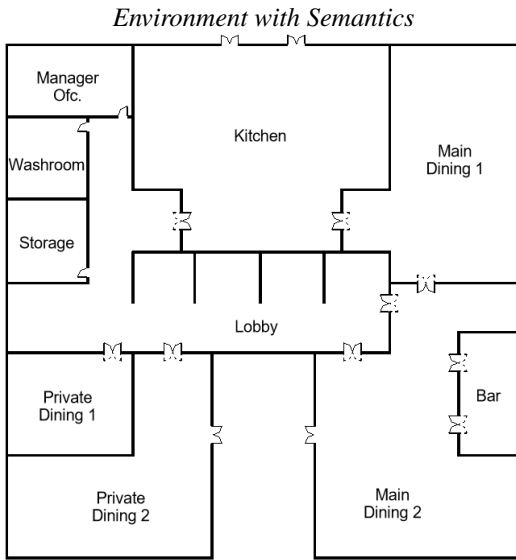
Figure 3: A variant of a real-world environment (e.g. a restaurant), used in the case study. Space semantics (e.g. room labels) are also shown.

flow, only the architectural elements like rooms, exits, corridors, ramps, etc. are considered while planning for the egress routes. Hence, for a given a BIM model, an egress plan simply consists of paths from rooms to their nearest exits, relying on the geometric information of the given model alone.

Since IBC rules state that travel distances to exits are to be measured from the most remote location of each room (Section 1), therefore, to devise an egress plan, we select the farthest point in every room. We then calculate shortest navigable routes from every room to its nearest exit.

Figure 4 (Left) shows an egress plan for static analysis in absence of space semantics which is a default workflow. Room and building exits in *green* showcase that there is absolute no restriction for these built spaces (e.g. associated rooms) and they can be part of egress routes. However, such a normative workflow might lead to an egress plan which violates certain IBC rules for *Means of Egress*. This is because one of the rooms in built-environment is to be used as "Kitchen", and therefore, it can not be a part of egress plan. This egress planning violation can be seen in Figure 4 (Middle). In addition, with standard workflow, only a static "distance" information is available to architects and designers for egress routes.

4.2 Static Analysis using Building Semantics

Next, we introduce *Space Semantic* information in static rule-checking for egress situations. The semantics for the model used in this study include *Kitchen*, *Main Dining* (2), *Private Dining* (2), *Bar*, *Washroom*, *Storage*, *Lobby* and *Manager's Office*. Such representations allow users to label the spaces (e.g. rooms) based on their potential usage. As a result, our tool is able to compute an egress plan which could not be computed in absence of semantic information. For example, it ensures that egress routes do not pass through certain

restricted building areas which are not allowed to be passed under IBC rules for *Means of Egress* (Table 1), hence, maximizing the quality assurances of built-environment for human safety.

Figure 4 (Right) shows an egress plan for static analysis in presence of space semantics. We constrained certain rooms which are to be used as "Kitchen", "Storage areas", etc., not to participate in egress routes. The exits of such rooms are shown in *red* to showcase that these are restricted areas, and as per IBC rules, they are not allowed to pass through during an emergency egress evacuation. As a result, avoiding any of IBC violations, and empowering designers and architects to further enhance the safety of potential building occupants by considering space semantics in preparing for an egress plan. The travel distances are also reported along with egress routes to further help in making informed design decisions. However, the egress decisions are still relying on static distance information alone even after incorporating space semantics in the planning.

4.3 Dynamic Analysis using Building Semantics

In previous static analyses, egress planning is conducted by relying on a static distance measure alone (i.e. travel distances from rooms to nearest exits). Since a building-design itself is static but the potential occupants of built-environment are not, therefore, we introduce dynamic crowd simulations to compute egress routes as well as to understand dynamics of potential occupant–building interactions for different crowd occupancies. As a result, designers and architects can make egress plans by not just considering travel distances but the evacuation times as well. In addition, they can also analyze exit flows of different crowd occupancies in their building designs to make informed egress design decisions, hence, empowering them to make more realistic and safe building environments.

Figure 5 shows egress analytics for different crowd occupancies. In each example, a different crowd occupancy behavior is tested (e.g. Right–Left: LoS Low, LoS Medium and LoS High). Egress trajectories of the occupants are shown as color gradient from *Red* to *Blue*. Trajectories in *Red* show shorter traveled distances and evacuation times, whereas in *Blue* show longer distances and evacuation times. Average exit flow values for whole population is also shown for every example. In the second row of the figure, crowd-density heatmaps are shown where problematic areas are highlighted in *Red* compared to *Blue* ones, which are comparatively less congested. *LoS High* exhibits multiple bottlenecks at the lobby, main dining 1 & 2 and near the exits. They are to further assist designers to understand dynamics of potential occupant–building interactions in planning egress.

As another example, selected room spaces in a building design is populated with varied count of occupants. Figure 6 shows egress analytics of such scenarios. As in Figure 5, egress trajectories and crowd-density heatmaps are shown here as well. This is to showcase that using the tool designers and architects can also analyze the dynamics of handpicked building spaces with their selected crowd occupancy behavior (e.g. LoS Low, LoS Medium and LoS High).
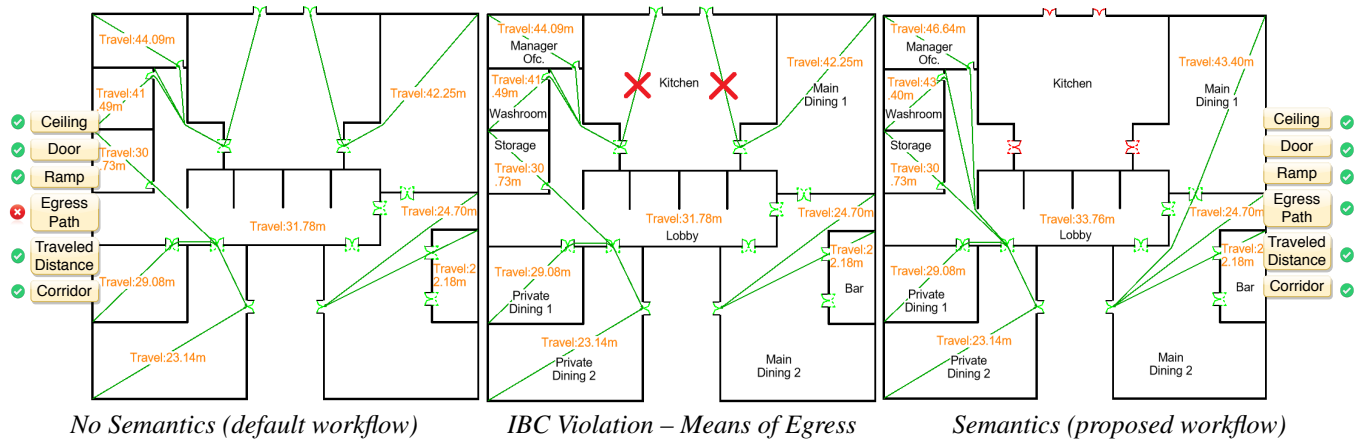
Figure 4: Emergency egress plan of the selected environment. Left: routes are computed using default egress planning workflow (i.e. in absence of space semantics). Middle: adding room labels to the egress plan shown in (Left). Right: routes are computed by taking into account the space semantics (i.e. semantics are defined). For both analyses, traveled distances are reported from the farthest point in every room to the nearest exit. Doors in red represent the entrance to areas which are not to be considered as part of egress routes under *Means of Egress* rules defined in building codes.

5 CONCLUSION

The presented work motivates the development of a semantic-based simulation-guided IBC rule checker for egress. The case study results indicate that the standard egress planning workflow does not take into account space semantics, violates certain IBC rules for *Means of Egress*, and rely on a static distance measure alone. Thus, poses threats for human safety in built-designs. On the other hand, by taking into account semantics of potential space usage and crowd simulations, a more secure egress plan can be achieved which takes into account evacuation times as well in planning for egress routes, and the dynamics from different crowd occupancy analyses, to help in making more realistic and safe egress planning decisions. Currently, the tool only validates and account for IBC rules for *Means of Egress*. Future work will involve incorporating social interactions in egress simulations [4]. Furthermore, additional IBC rules will be integrated which can be validated by means of dynamic simulations. A user evaluation and integration of the current tool into other architecture design platforms (e.g. Digital Project, Rhinoceros, Sketchup, etc.) will also be done in future work.

REFERENCES
1. Balaban, Ö., Kilimci, E. S. Y., and Cagdas, G. Automated code compliance checking model for fire egress codes. *Digital Applications in Construction 2* (2012).

2. Berseth, G., Usman, M., Haworth, B., Kapadia, M., and Faloutsos, P. Environment optimization for crowd evacuation. *CAVW 26*, 3-4 (2015), 377–386.

3. Cassol, V. J., Testa, E. S., Jung, C. R., Usman, M., Faloutsos, P., Berseth, G., Kapadia, M., Badler, N. I., and Musse, S. R. Evaluating and optimizing evacuation plans for crowd egress. *IEEE computer graphics and applications 37*, 4 (2017), 60–71.

4. Chu, M. L., Parigi, P., Law, K., and Latombe, J.-C. Safegress: a flexible platform to study the effect of human and social behaviors on egress performance. In *Proceedings of the Symposium on Simulation for Architecture & Urban Design*, Society for Computer Simulation International (2014), 4.

5. Conover, D. Development and implementation of automated code compliance checking in the us. *International Code Council* (2007).

6. Eastman, C., Lee, J.-m., Jeong, Y.-s., and Lee, J.-k. Automatic rule-based checking of building designs. *Automation in construction 18*, 8 (2009), 1011–1033.

7. Eastman, C., Teicholz, P., Sacks, R., and Liston, K. *BIM handbook: A guide to building information modeling for owners, managers, designers, engineers and contractors*. John Wiley & Sons, 2011.

8. Fan, S.-L., Chi, H.-L., and Pan, P.-Q. Rule checking interface development between building information model and end user. *Automation in Construction 105* (2019), 102842.

9. Fruin, J. J. *Designing for pedestrians: A level-of-service concept*. No. HS-011 999. 1971.

10. Fruin, J. J. Pedestrian planning and design. Tech. rep., 1971.

11. Haworth, B., Usman, M., Berseth, G., Kapadia, M., and Faloutsos, P. Evaluating and optimizing level of service for crowd evacuations. In *Proceedings of MIG*, ACM (2015), 91–96.
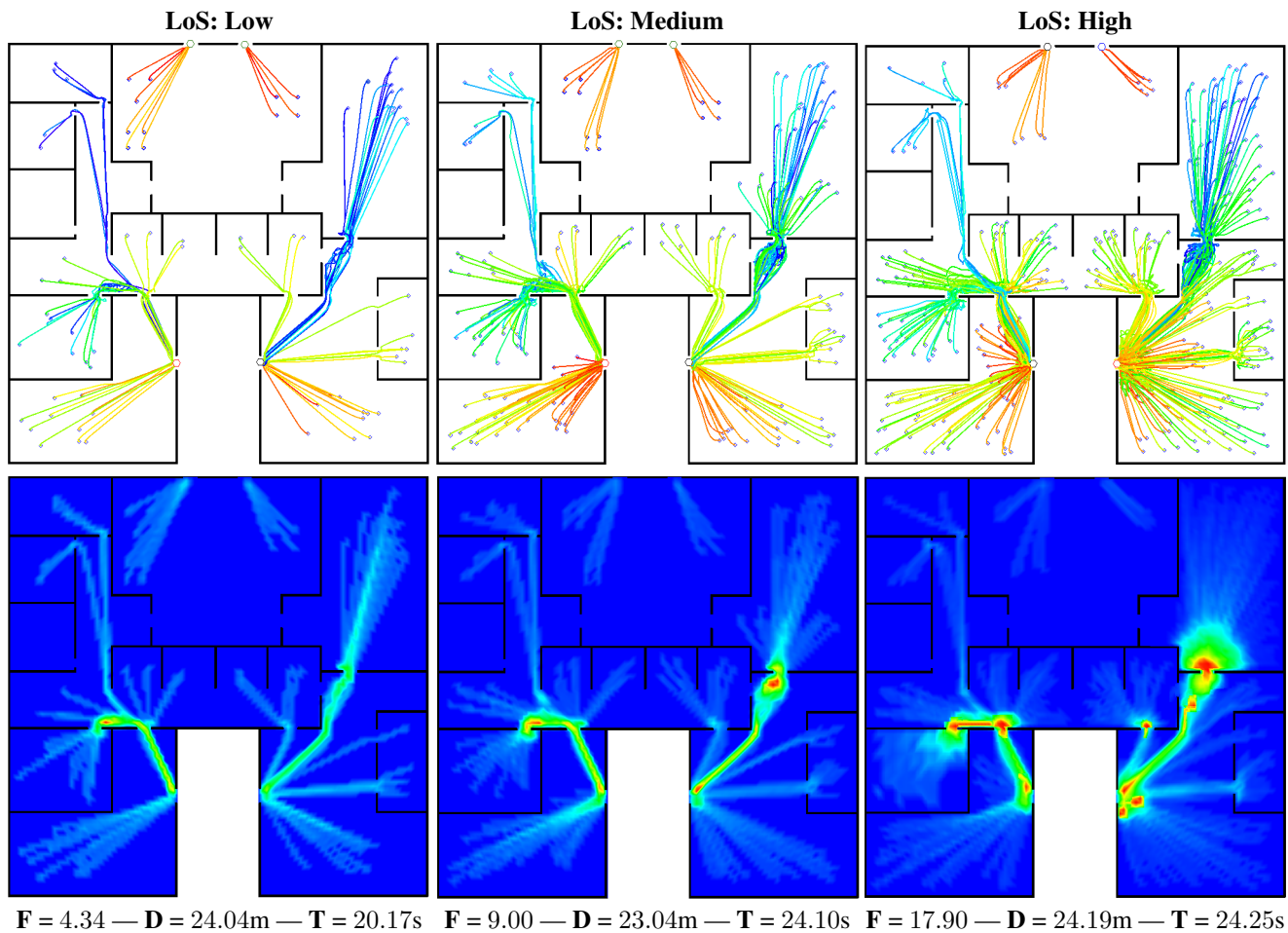
Figure 5: Spatial analytics for varied crowd occupancies. Each example maps a different crowd occupancy level (e.g. Level of Service). Top: color-coded (Red–Blue) trajectories of occupants' evacuations are shown based on average of evacuation times (**T**) and traveled distances (**D**). Trajectories in *Red* show shorter traveled distances and evacuation times, whereas in *Blue* show longer distances and evacuation times. Average exit flow values (**F**) (i.e. occupants/second) are also shown. Bottom: crowd density heatmaps with high density in red (problematic areas) and low in blue.

12. Haworth, B., Usman, M., Berseth, G., Khayatkhoei, M., Kapadia, M., and Faloutsos, P. Towards computer assisted crowd aware architectural design. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, ACM (2016), 2119–2125.

13. Haworth, B., Usman, M., Berseth, G., Khayatkhoei, M., Kapadia, M., and Faloutsos, P. Code: Crowd-optimized design of environments. *CAVW 28*, 6 (2017), e1749.

14. Helbing, D., and Molnar, P. Social force model for pedestrian dynamics. *Physical review E 51*, 5 (1995), 4282.

15. International Code Council. International building code, 2018.

16. International Code Council. Means of Egress. International Building Code, 2018.

17. Junior, J. S., Formoso, C. T., and Tzortzopoulos, P. A semantic-based framework for automated rule checking in healthcare construction projects. *Canadian Journal of Civil Engineering* (2019).

18. Kapadia, M., Pelechano, N., Allbeck, J., and Badler, N. Virtual crowds: Steps toward behavioral realism. *Synthesis lectures on visual computing: computer graphics, animation, computational photography, and imaging 7*, 4 (2015), 1–270.

19. Khemlani, L. Corenet e-plancheck: Singapore's automated code checking system. *AECbytes, October* (2005).

20. Nguyen, T.-H., and Kim, J.-L. Building code compliance checking using bim technology. In *Proceedings of the 2011 Winter Simulation Conference (WSC)*, IEEE (2011), 3395–3400.

21. Pelechano, N., Allbeck, J. M., Kapadia, M., and Badler, N. I. *Simulating heterogeneous crowds with interactive behaviors*. CRC Press, 2016.

Kitchen — Private Dining 2 — Bar
Occupants = 11 — 50 — 11

Manager Office — Lobby — Bar
Occupants = 4 — 20 — 14

Kitchen — Main Dining 1 — Main Dining 2
Occupants = 11 — 38 — 28

F = 6.82 — **D** = 12.36m — **T** = 10.42s   **F** = 2.24 — **D** = 29.24m — **T** = 24.10s   **F** = 4.02 — **D** = 23.13m — **T** = 24.57s
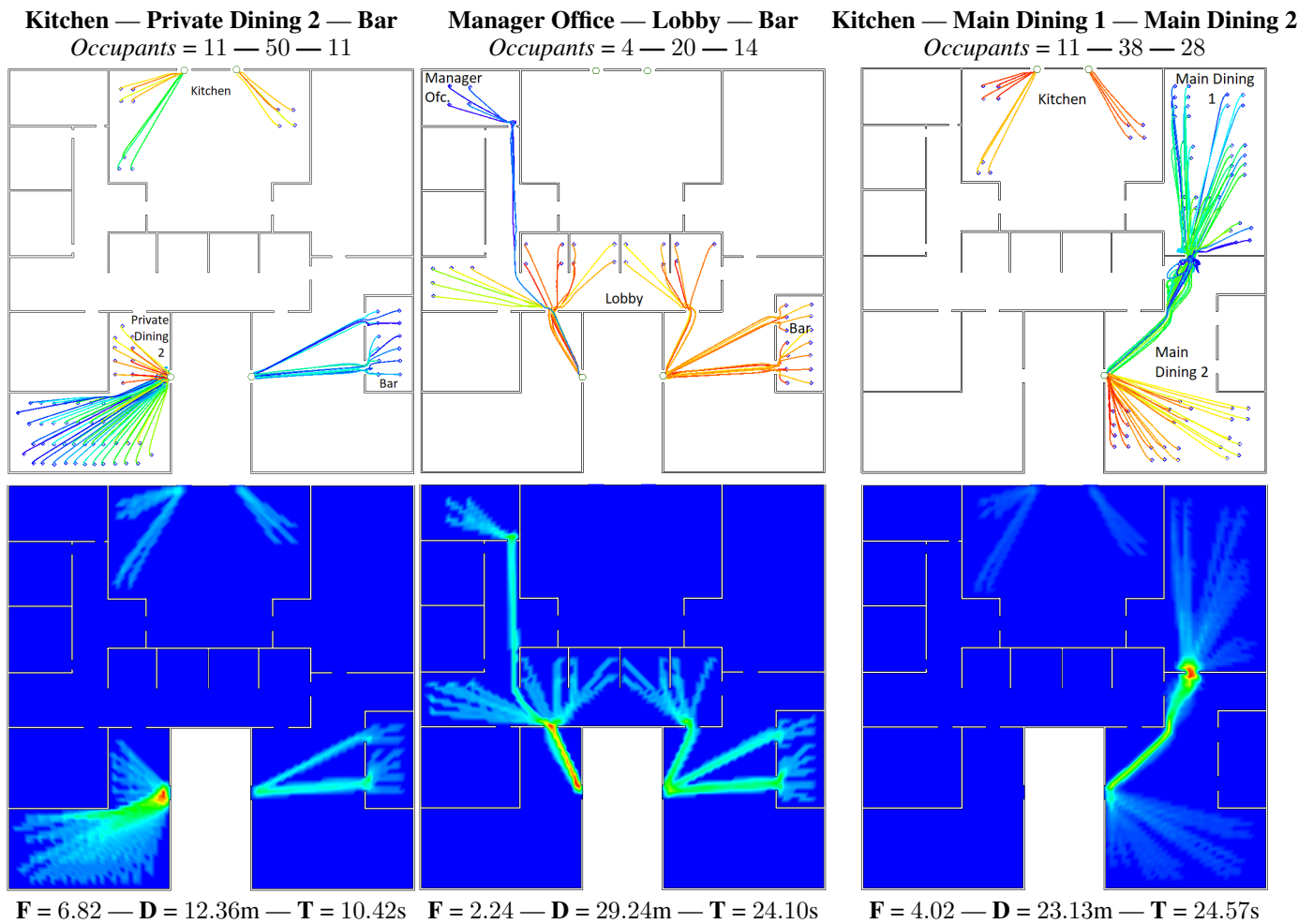
Figure 6: Crowd analytics for varied user-selected areas using *Medium* LoS. In each example, different rooms in a building space is populated with varied count of occupants. Top: color-coded (Red–Blue) trajectories of occupants' evacuations are shown based on average of evacuation times (**T**) and traveled distances (**D**). Trajectories in *Red* show shorter traveled distances and evacuation times, whereas in *Blue* show longer distances and evacuation times. Average exit flow values (**F**) (i.e. occupants/second) are also shown. Bottom: crowd density heatmaps with high density in red (problematic areas) and low in blue.

22. Schaumann, D., Moon, S., Usman, M., Goldstein, R., Breslav, S., Khan, A., Faloutsos, P., and Kapadia, M. Join: an integrated platform for joint simulation of occupant-building interactions. *Architectural Science Review* (2019), 1–12.

23. Singh, S., Kapadia, M., Faloutsos, P., and Reinman, G. An open framework for developing, evaluating, and sharing steering algorithms. In *Motion in Games*, Springer-Verlag (2009), 158–169.

24. Solibri. Solibri model checker, 2009.

25. Technology, J. E. Express data manager (edm), 2009.

26. Usman, M., Schaumann, D., Haworth, B., Berseth, G., Kapadia, M., and Faloutsos, P. Interactive spatial analytics for human-aware building design. In *Proceedings of MIG*, ACM (2018), 13.

27. Usman, M., Schaumann, D., Haworth, B., Kapadia, M., and Faloutsos, P. Joint exploration and analysis of high-dimensional design–occupancy templates. In *Proceedings of MIG*, ACM (2019), 35.

28. Usman, M., Schaumann, D., Haworth, B., Kapadia, M., and Faloutsos, P. Joint parametric modeling of buildings and crowds for human-centric simulation and analysis. In *International Conference on Computer-Aided Architectural Design Futures*, Springer (2019), 279–294.

29. Zhang, J., and El-Gohary, N. M. Semantic-based logic representation and reasoning for automated regulatory compliance checking. *Journal of Computing in Civil Engineering 31*, 1 (2017), 04016037.

30. Zhang, S., Teizer, J., Lee, J.-K., Eastman, C. M., and Venugopal, M. Building information modeling (bim) and safety: Automatic safety checking of construction models and schedules. *Automation in Construction 29* (2013), 183–195.

Urban Public Spaces as Network Configurations Through Real-Time Traffic Data

Eirini Androutsopoulou [1], Eleftherios Anagnostopoulos[2]

[1] National Technical University
of Athens
Athens, Greece
iandroutsopoulou@gmail.com

[2] WePraos
Athens, Greece
praos.corp@gmail.com

ABSTRACT

Based on the research of urban public spaces as network configurations, we present a novel methodology which takes into account real-time traffic data in order to determine and categorize the network's connections. Three means of transport produce respective variables per network connections at every instance which correspond to trip durations.

Through the application of k-means clustering, different categories of connections arise. As different time instances correspond to different traffic data, the network's connections change accordingly. The introduction of actual variations between the relationships of the network's nodes produces time sequences of network reconfigurations thanks to the network's inner dynamics.

Since the network's original structure is based on dynamic connections between urban public spaces, the resultant real-time sequences of network configurations reveal the self-adapting behavior of the network of urban public spaces.

Author Keywords

Dynamic network; real-time traffic data; urban public spaces; cluster analysis

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING

1 INTRODUCTION

The study of the urban environment as a network configuration is based on the definition of the network nodes and their interconnections. Public spaces are treated as network nodes and their attributes come from urban analysis, as well as from their network function. Through the application of network algorithms (degree, betweenness centrality, eigenvector centrality, closeness centrality and eccentricity), attributes referring to the network structure are applied. These are called network characteristics of urban spaces [1].

In contrast to the hierarchical constructions, network constructions allow for multiple connections between elements [2], therefore being closer to the complexity of the associative forces found in the structure of the urban environment. The transition from the physical space to the network structure involves the retention of nodes' proximity through network's connections only, exempting any other analogy or dependency from the objects' Cartesian's coordinates at the physical space.

This research focuses on dynamic network connections. Ten public spaces of Zografos Municipality in Attica, Greece are selected (Figure 1) and the connections between them are determined and categorized through real-time traffic data. Car, bus and pedestrian traffic data form dynamic data sets corresponding to trip durations (Figure 2).

Given the fact that the nodes' attributes depend on the overall network structure, that is, the nodes and their connections, a dynamic network structure produces transitional data of network characteristics per network node. What is more, the use of real-time data reveals in real conditions the self-adapting behavior of the network of urban public spaces.
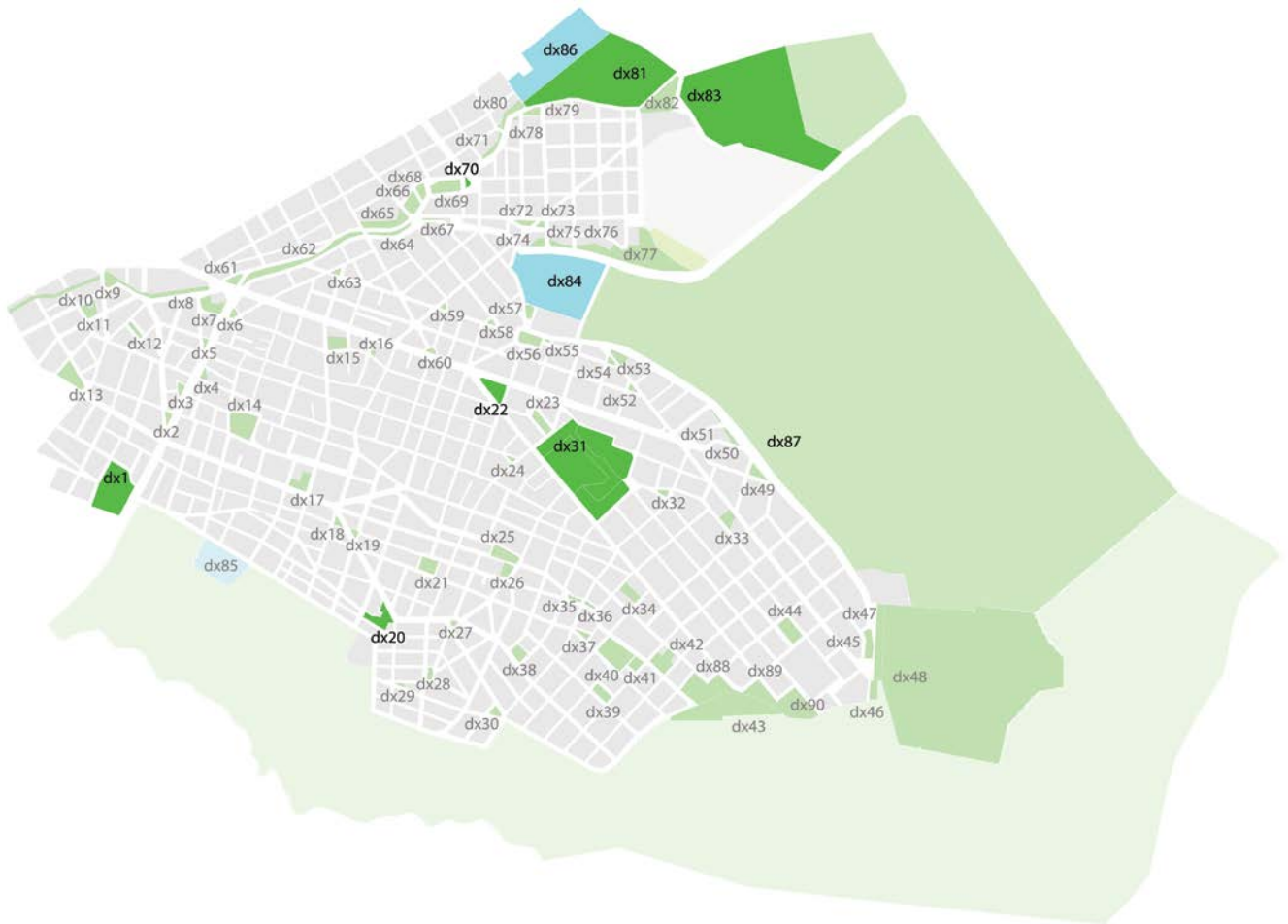
Figure 1. The selection of ten existing urban public spaces of the Municipality of Zografos. The three new metro stations will be located at Eleftherias square (dx70), Gardenia square (dx22) and Cyprus square (dx20), while Zografou park (dx1), Zografou Villa (dx31), Goudi park (dx81), Goudi grove (dx83), Zografou stadium (dx84), Zografou Sports Center (dx86) and National Technical University Campus (Zorgafou Gate, dx87) are highly visited public spaces of Zografos Municipality.

Figure 2. Car, bus and pedestrian traffic data are collected for every pair of urban public spaces.

52

2 REAL-TIME TRAFFIC DATA COLLECTION

The initial network configuration is formed through the mutual connection of every pair of public spaces. All ten public spaces are initially interconnected using bidirectional connections of equal strength (weight), each pointing to a node. This appears to be a symmetric graph, since all nodes are equally interconnected (Figure 3).
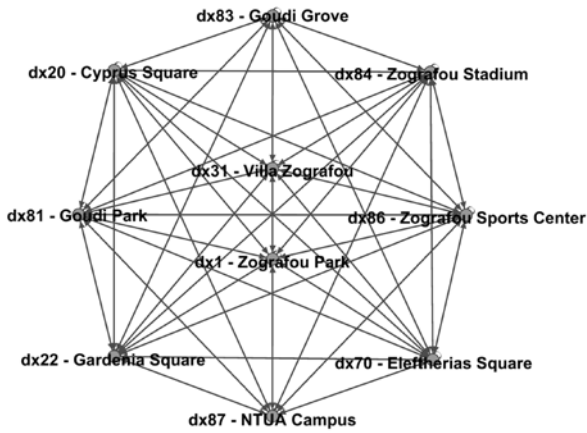
Figure 3. The symmetric graph produced if all nodes are interconnected using bidirectional connections of equal weight.

In order to construct a dynamic network whose connections will alter in strength according to real-time traffic data, car, bus and pedestrian trip durations from each node to every other node of the network are collected. For every connection between any pair of public spaces, the actual time needed to travel between them at the actual physical urban space is collected during 12 hours of a specific day (22.10.2019). Note that the time needed to reach urban space B from urban space A is not necessarily the same as the time needed to reach urban space A starting from urban space B using car or bus, as the proposed routes and the respective distances covered may differ (Figure 4). This is not the case for the pedestrian trip duration, as the shortest proposed route to follow is the same.

The collection of real-time traffic data is made possible through the use of data sets retrieved from Google Maps for the three means of transport. Trip durations from any given node (urban public space) to any other node are collected for a total time of 12 hours, starting from 08h45 to 20h15 and for a total of 24 steps, using 30 minutes intervals between them (Figure 5).

This produces 24 data sets of 100 time values in seconds (10x10 nodes) and for three means of transport, that is 7.200 values that depict real-time travel data.

Trip duration by car 8h45 - 22.10.2019 (SEC)

	dx1	dx20	dx22	dx31	dx70	dx81	dx83	dx84	dx86	dx87
dx1	0	241	374	421	449	549	565	633	569	1001
dx20	268	2	222	181	531	630	665	587	672	1017
dx22	664	521	21	484	398	498	514	259	515	900
dx31	428	272	82	6	442	545	549	295	565	954
dx70	410	368	294	335	2	168	211	552	188	710
dx81	541	503	424	468	137	0	109	683	59	844
dx83	579	548	461	506	177	102	0	725	122	879
dx84	424	263	180	223	147	253	272	13	273	660
dx86	560	524	439	485	156	63	128	702	0	708
dx87	730	557	468	511	408	440	447	728	458	0

Trip duration by bus 8h45 - 22.10.2019 (SEC)

	dx1	dx20	dx22	dx31	dx70	dx81	dx83	dx84	dx86	dx87
dx1	0	331	802	866	668	1003	1272	947	1144	1179
dx20	546	0	762	557	1471	1344	1951	910	1485	1062
dx22	763	733	0	271	666	1093	912	284	1234	463
dx31	948	525	214	0	707	1042	1150	661	1183	614
dx70	635	1076	570	673	0	397	576	318	539	986
dx81	987	1343	809	940	305	0	260	535	195	1250
dx83	1141	1427	963	1024	390	223	0	639	363	1404
dx84	862	858	264	379	278	562	670	0	703	692
dx86	1152	1847	954	1085	486	199	391	700	0	1393
dx87	977	963	337	449	792	1071	885	503	1212	0

Trip duration for pedestrians 8h45 - 22.10.2019 (SEC)

	dx1	dx20	dx22	dx31	dx70	dx81	dx83	dx84	dx86	dx87
dx1	0	815	1064	1175	1043	1432	1641	1156	1574	1798
dx20	676	0	716	561	1032	1344	1432	864	1485	1145
dx22	943	733	0	271	548	851	939	284	992	717
dx31	961	525	214	0	797	1042	1150	493	1183	649
dx70	943	1076	570	834	0	397	576	318	539	1060
dx81	1291	1343	826	1085	340	0	260	555	195	1240
dx83	1445	1427	910	1169	494	223	0	639	363	1337
dx84	1016	858	264	529	278	562	670	0	703	734
dx86	1437	1489	972	1231	486	199	391	700	0	1386
dx87	1520	1019	562	540	862	1078	1182	578	1219	0

Figure 4. The collection of real-time traffic for the three means of transport at step 1 (8h45).
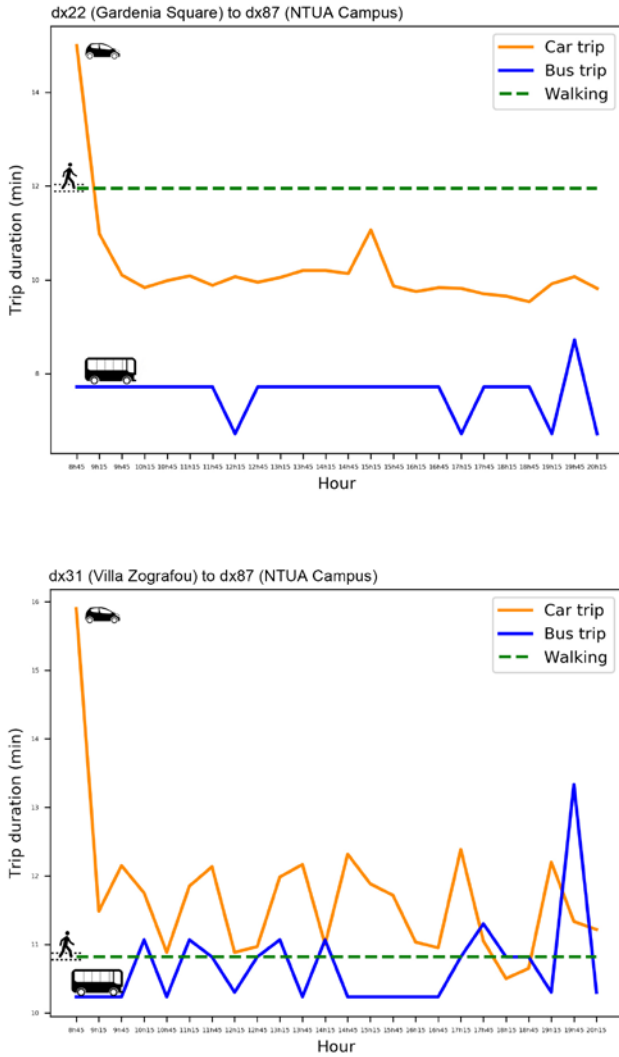
Figure 5. The collection of real-time traffic data for the three means of transport for a total of 24 steps (8h45 to 20h15)

3 CLUSTERING AND CATEGORIZATION OF NETWORK CONNECTIONS

The dynamic network construction presupposes requires a way of computing the transition from the actual connections of the urban environment to the network's connections, using real-time traffic data. These data refer to the actual time needed to travel from one public space to another. As the rules of the reconstruction of the urban environment to a network configuration may differ from one application to another, this research focuses on the constant alteration of the urban connections which produce respective alterations of the network's structure.

In order to be able to categorize the network's connections and produce variations of the network's structure which correspond to real-time traffic data, clusters of the connections between pairs of public spaces are produced, using K-means clustering algorithm. Within a labeled three-dimensional dataset (1:1, 1:2, 1:3, ..., 10:8, 10:9, 10:10 correspond to dx1:dx2, dx1:dx2, dx1:dx3, ..., dx10:dx8, dx10:dx9, dx10:dx10 connections), k-means algorithm searches for clusters of connections per instance. This way, each connection of three variables, car (sec), bus (sec) and pedestrian(sec), is grouped within a cluster whose center is closer to this connection than to other cluster centers (Figure 6).

For every instance, a different clustering result is produced. These 24 clustering results correspond to 24 different groupings of network connections, according to the k-means clustering algorithm application.

4 CONNECTIONS STRENGTH ACCORDING TO REAL-TIME TRAFFIC DATA CLUSTERING

4.1 Connections strength according to real-time traffic data clustering results

Being able to collect real-time data from the actual urban environment, a series of clustering results is produced. The translation of real-time traffic data into network connections is made possible through an interpretation of cluster results into connections weight (Table 1).

A specific weight value is assigned on connections belonging to the same cluster. Shorter trip durations are translated into stronger directed connections between network nodes, while longer trip durations are translated into weaker network connections. This way, different network configurations are constructed for every instance. Connection strength values alter continuously, following the alterations of the actual urban connections' groupings.

4.2 The dynamic network structure is based on transitional data

Since the time needed to travel from node A to node B (car (sec), bus (sec)) is not necessarily equal to the time needed to travel from node B to node A, all ten public spaces (nodes) are interconnected using pairs of bidirectional connections which differ in weight. Dynamic networks containing bidirectional edges produce transitional network characteristics per node, depending solely on the network structure. With the application of the weighted degree algorithm [5], the weighted bidirected network of public spaces acquires a ranking of public places, according to their average degree. The most powerful nodes in terms of weighted degree are those most strongly connected with the other nodes of the network. This means that the weaker nodes are these public spaces which require longer trips for the three means of transport to be reached. Weak urban spaces are exposed, appearing as less influential urban elements of the network overall structure (Figure 7).
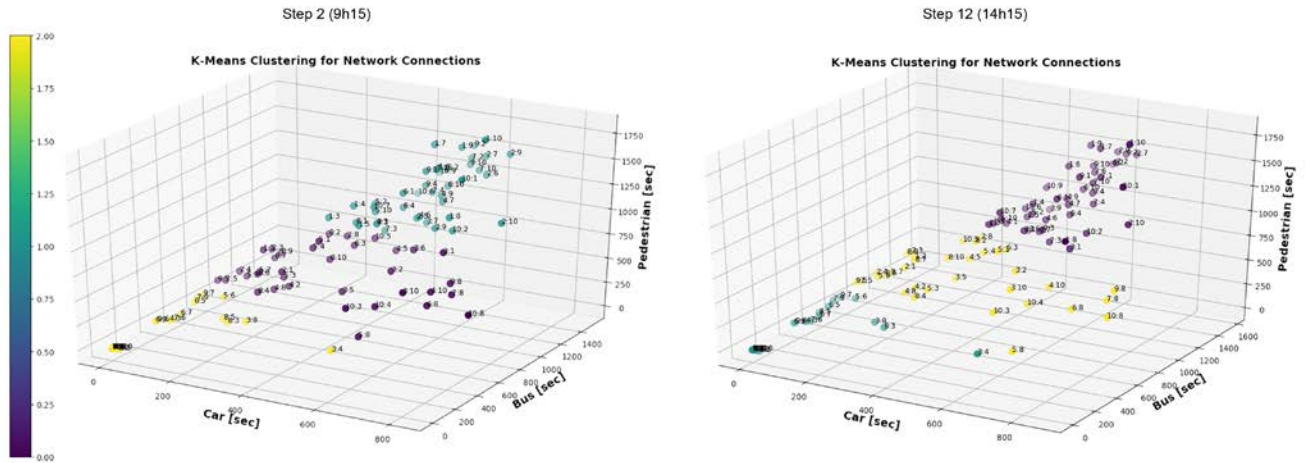
Figure 6. The scatter plots of the produced clusters using K-means clustering algorithm.
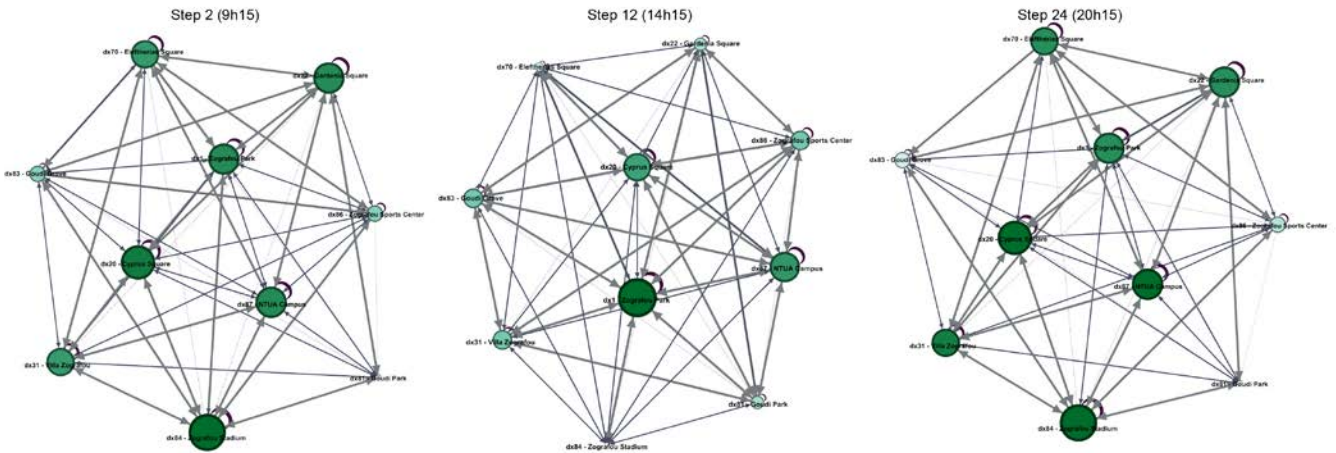
Figure 7. With the application of the weighted degree algorithm weak urban spaces are exposed, appearing as less influential urban elements of the network overall structure. The nodes' size and color correspond to the weighted degree value.

Source	Target	Cluster	Weight	Type
1	1	0	1	Directed
1	2	2	2	Directed
1	3	1	3	Directed
1	4	1	3	Directed
1	5	1	3	Directed
1	6	1	3	Directed
1	7	1	3	Directed
1	8	1	3	Directed

Table 1. Table extract of network edges for step 12 (14h15)

5 THE SELF-ADAPTING BEHAVIOR OF THE NETWORK OF URBAN PUBLIC SPACES

The transitional qualities of the urban environment produce simultaneous but not analogous changes at the network structure. The network configuration relies on nodes, and connections in order to reconstruct the urban environment. The clusters [3] and proximities between elements emerge from the topology produced by the strength of the connections and not from the Cartesian topology. Centrality studies [4], made possible through the network configuration, reveal the self-adaptation of the urban body, provoked by dynamic network connections.

Through the application of algorithms which re-evaluate the connection forces between nodes, as well as the mathematical rules witch define the cluster formations, the result-output is generated through the processing of the parameters that determine the urban structure. Real-time data determine the construction of the network itself, that is to say the relationships between component parts that describe the mutational procedure, while the alteration of the initial structure of the urban configuration produces a time-based sequence of urban mutations.

6 CONCLUSION

The multiple connections between elements, found inherent in the main body of the network construction involves the interpretation of the changing qualities of the urban environment. On the other hand, the mathematical rules describing the way the urban spaces relate to each other involves the reinterpretation of the urban change based on the forces that cause change and not on the result itself. With the combined management real-time data related to the actual urban environment and with the transition to an ever-changing network structure, it is possible to formulate a methodology of simulating the self-adapting behavior of the network of urban public spaces.

Since many of the research topics for strategies for urban interventions aim at the treatment of the public spaces of Athens as a coherent whole and not as a sum of sparse entities scattered at the urban environment, the study of the connections between public spaces contributes at the enrichment of the criteria for the prioritization of urban interventions on public spaces (nodes-public spaces and their interconnections).

REFERENCES

1. Alexander Ch., 'A City is not a Tree', 1965, *Architectural Form*, vol. 122, nr. 1, pp. 58–61.

2. Androutsopoulou, E.: 2016. 'Urban body network configurations through attributes of network elements', *Conference Proceedings of the Symposium on Simulation for Architecture and Urban Design*, Simaud 2016, 111-117.

3. Blondel,VD, Guillaume, JL, Lambiotte, R, Lefebvre, E, 'Fast unfolding of communities in large networks', *Theory and Experiment 2008 (10)*, P1000

4. Brandes,U, A Faster Algorithm for Betweenness Centrality, 2001, *Journal of Mathematical Sociology*, 25(2):163-177.

5. Newman, M. E. J. 2001. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*. 64 (1): 1–7.

An Innovative Approach to Determine Building Window-To-Wall Ratios for Urban Energy Simulation

Xing Shi[1], Chao Wang[2], Meng Wang[3], Peng Tang[4]

[1]School of Architecture,
Southeast University
Nanjing, China
shixing_seu@163.com

[2]School of Architecture,
Southeast University
Nanjing, China
chaowang_seu@163.com

[3]Schooll of Computer Science
and Engineering, Southeast
University
Nanjing, China
meng.wang@seu.edu.cn

[4]School of Architecture,
Southeast University
Nanjing, China
tangpeng@seu.edu.cn

ABSTRACT

Window-to-wall ratio (WWR) is one of the most important parameters influencing building and urban energy consumption. In the context of urban energy simulation, determining the WWRs of numerous urban buildings is a challenging task. This paper presents an artificial intelligence (AI) built on semantic segmentation and convolutional neural network. The AI was trained to recognize opaque walls and transparent windows and thus be able to calculate the WWRs from building elevation images. The AI was tested using manual calculations and building elevations with true WWRs known. The testing results show that the AI is able to determine the WWRs with satisfactory accuracy.

Author Keywords

Urban energy; window-to-wall ratio; artificial intelligence; semantic segmentation; convolutional neural network

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING

1   INTRODUCTION

Urbanization is a process that dramatically changes cities and the way people live and work there. It is also one of the major reasons for the increase of energy consumption and carbon emission. According to the UN [1], by 2050 there will be 68% of the total population worldwide living in cities. Therefore, urbanization has a profound impact on many aspects of the development for nations in the world, especially those undergoing a fast urbanization such as China and some countries in southeast Asia and Africa.

Energy shortage is a challenge faced by the entire world and exacerbated by urbanization. Many countries have issued and implemented laws or standards to regulate the energy consumption and efficiency for buildings and cities, examples being The Energy Performance of Buildings Directive in the EU [2], ANSI/ASHRAE/IES Standard 90.1 in the USA [3], and GB50189 in China [4]. In addition to implementing energy efficient codes and standards, research and development to achieve urban energy efficiency and reduce carbon emission in cities has been active. Among various research subjects, urban energy simulation is a promising one.

Urban energy simulation (or modelling), described as a "nascent field" [5], is a technique to model and simulate urban energy consumption, including its magnitude and preferably its spatial and temporal pattern. In terms of its theoretical framework, urban energy simulation can be broadly divided into two classifications, namely the top-down approach and the bottom-up approach. The top-down approach depends on statistical and historical data from industrial sectors and estimate urban energy consumption using regressional or other statistical methods. The bottom-up approach on the other hand relies on the calculation of the energy consumption of individual buildings and accumulates them to obtain urban energy consumption. Due to its methodical nature, the bottom-up approach is more suitable for architects and urban designers and planners because it reflects the impact of urban form, which is the central theme for those professionals, on urban energy performance.

To conduct a bottom-up urban energy simulation, a large number of data are required, among which the window-to-wall ratio (WWR) of urban buildings is a critical one. The WWR of a building is defined as the ratio between the area of the windows and that of the solid walls in the same elevation. Since windows with transparent glazing allow more solar heat gain into buildings in summer and more heat conductive loss in winter compared with solid walls, a large WWR indicates that the heating and cooling load of the building is likely to be high, which will lead to a higher

energy consumption. Therefore, the WWR is a critical parameter to affect building energy consumption and consequentially urban energy consumption.

To obtain the WWR of an individual building is straightforward. In many cases, the design and construction drawings of the building being investigated are available. The WWR can be determined by calculating the area of the windows and solid walls on each elevation of the building. If the drawings are not available, for one single building it is both technically feasible and economically viable to conduct some field measurements to determine its WWR. However, neither of these two approaches work well for cities.

For a typical urban energy simulation task, there are hundreds and even thousands of buildings. It is unlikely, if not impossible, to have access to the design and construction drawings for all of the buildings. On the other hand, the measurement technique to obtain the WWR for single buildings, if applied to hundreds or thousands of urban buildings, is likely to become too time-consuming and resource-demanding to be feasible. Therefore, it is necessary to develop new methods and technologies to determine the WWR of buildings in the context of urban energy simulation.

Currently, two approaches to obtain the WWR for a large number of urban buildings are commonly used, namely oblique photography and expert judgement. Oblique photography is a type of aerial photography in which the camera axis is deliberately kept tilted from the vertical by a specified angle. The photographs, thus taken, reveal details masked in some ways in vertical photographs. Oblique photography has been used in acquiring three dimensional information of cities and creating their maps and models [6]. The problem with oblique photography is the cost of the equipment and a very large number of data that need to be post-processed for any meaningful city models to be generated. Expert judgement is another approach that has been used in determining the WWRs of urban buildings. However, experts' knowledge has its limit and cannot be exclusively relied on. In some cases, the WWR is assumed to be a randomly distributed value in a certain range [7], which is clearly an over-simplification.

Hence, both oblique photography and expert judgement have some limitations and even flaws in determining the WWR for urban energy simulation. There is clearly a need to develop an innovative methodology to obtain the WWRs for many urban buildings with relatively low cost, high efficiency, and acceptable accuracy.

This paper presents an AI (Artificial Intelligence) to determine the WWRs for urban buildings in the context of urban building simulation. The AI is built on semantic segmentation and convolutional neural network. It is trained using a set of images which represents various typical building elevations. The reliability and accuracy of the AI is tested. The results are satisfactory in terms of both the efficiency and accuracy.

2 METHODOLOGY

2.1 Algorithms of the AI

The primary algorithms of the AI consist of semantic segmentation and convolutional neural network. Semantic segmentation is a machine learning algorithm that assigns a tag or category to each pixel in an image. It is often used to recognize a group of image pixels which belong to a distinct object or category. A convolutional neural network, sometimes referred to as shift invariant or space invariant artificial neural networks, is a class of deep and feed-forward artificial neural networks that are applied to analyzing visual imagery. Semantic segmentation and convolutional neural network are combined and realized through Unet framework, which was originally designed and developed for medical imaging processing [10].

Figure 1. Some of the training images of building elevations

58

2.2 Functional modules of the AI

The AI consists of three main functional modules, namely RESIZE, SPLIT, and CRF. RESIZE is the functional module to recognize the exterior wall in an image containing a building elevation. SPLIT is to distinguish the windows from the solid part of the wall. CRF is the functional module to repair the missing parts of the wall or windows due to blocking obstacles such as trees. The former two functional modules are run for every building elevation image while the latter is only run when a blocking obstacle is identified in the image.

2.3 Training of the AI

After setting up the AI in Unet framework, the next step was to train him to gain the capability of recognizing solid walls and transparent windows in building elevation images. The training set was built with more than 150 images containing a variety of building elevations (Figure 1). These images were selected by the following criteria:

- Cover a broad range of building types, wall types, window types and building ages.

- Avoid the blocking of the building elevation as much as possible. The blocking can be caused by trees, vehicles, small buildings, and other obstacles.

The images in the training set were marked with pixel-level precision, i.e., the walls and windows in the images were manually marked as shapes containing a certain number pixels. Although this operation is time-consuming and labor-intensive, it can be sped up by using image mark-up programs.

Since the Unet framework requires that input images have the same format and size, the training images were compressed or resized to be 512x512 pixels.

3 RESULTS AND ANALYSIS

3.1 Recognition of the walls and windows by the AI

After training, the AI was able to recognize the walls and windows in a building elevation image. Figure 2 illustrates a building elevation image and the shape of the wall and windows recognized by the AI.

3.2 Accuracy of the AI compared with manually calculated WWR

To test the accuracy of the trained AI, 40 images with building elevations were accumulated. The WWRs were determined using the AI and manual calculation, respectively. Figure 3 illustrates the relative errors of the WWRs determined by the AI compared with those calculated manually. The relative error is defined by Eq. 1,

$$RE = \frac{WWR_{AI} - WWR_{MANUAL}}{WWR_{MANUAL}} \times 100\% \qquad (1)$$

where RE is the relative error; WWR_{AI} is the window-to-wall ratio determined by the AI; WWR_{MANUAL} is the window-to-wall ratio calculated manually.

As shown in Figure 3, the AI determined the WWRs with fairly good accuracy since 35 out of 40 building elevations show the relatively error (the absolute value) is less than 10%. In only 5 cases, the relatively error (the absolute value) is larger than 10%.

Figure 2. A building elevation image and the wall and windows recognized by the AI
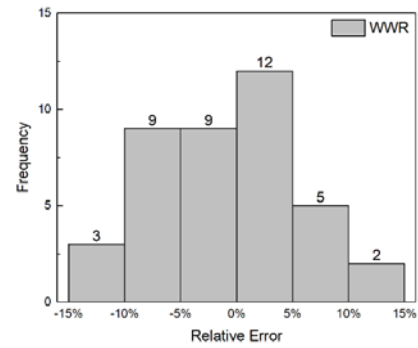
Figure 3. The relative errors of the WWRs determined by the AI compared with those calculated manually
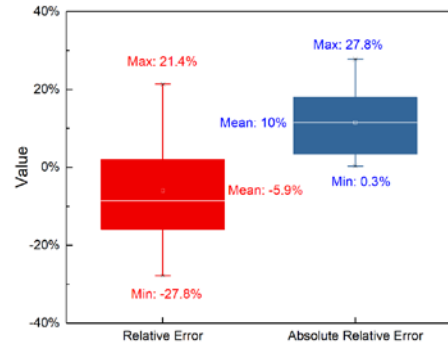
Figure 4. The box error bars of the WWRs determined by the AI compared with those calculated using the design drawings

Table 1. Relative errors between the WWRs calculated using the design drawings and those determined by the AI

D	WWR$_{actual}$	WWR$_{AI}$	Relative Error (%)
1	0.15	0.1311	-12.6
2	0.13	0.1274	-2.0
3	0.34	0.2876	-15.4
4	0.15	0.1396	-6.9

59

5	0.13	0.1011	-22.2
6	0.34	0.3044	-10.5
7	0.13	0.1320	1.5
8	0.14	0.1434	2.4
9	0.13	0.1283	-1.3
10	0.13	0.1126	-13.4
11	0.16	0.1309	-18.2
12	0.31	0.3240	4.5
13	0.16	0.1437	-10.2
14	0.31	0.2239	-27.8
15	0.16	0.1706	6.6
16	0.16	0.1315	-17.8
17	0.04	0.0486	21.3
18	0.24	0.2392	-0.3
19	0.29	0.2426	-16.3
20	0.14	0.1693	21.0

3.3 Accuracy of the AI compared with known WWRs

Twenty building elevations with design drawings were obtained to further benchmark the accuracy of the AI. Since the design drawings were available, the WWRs could be accurately calculated without taking elevation images. The true WWRs calculated using the design drawings, those determined by the AI, and the relative errors are presented in Table 1. Figure 4 shows the boxplots of the relative errors and their absolute values. Overall, the largest positive relative error is 21.4% and the largest negative relative error is -27.4%. The mean relative error is -5.9%, which is considered quite small. Note that although the mean relative error is usually calculated based on the absolute value to avoid the cancellation of the positive and negative numbers, using the relative errors with signs is valid in this case because the objective of determining the WWRs is to simulate urban energy consumption. Higher WWRs are likely to cause a higher urban energy simulation result and vice versa. Therefore, it is reasonable to expect the overall effect to be more realistic.

3.4 Limitations

The most important limitation of this study is probably the applicability of the approach to cities and urban districts with a large number of buildings. Although the speed of AI to recognize WWRs is fast, the preparation, mainly the collection of building elevation images, takes time. Based on our current experience, the approach works well on hundreds of buildings. When the number reaches a few thousands, the time required would be significantly more, perhaps still manageable for a small research team with several people. However, for cities and urban districts with

tens of thousands of buildings, the approach may require too much time and too many resources to be feasible.

4 CONCLUSION

The window-to-wall ratio is one of the most important parameters influencing the energy performance of buildings and cities. In the context of urban energy simulation, it is challenging to obtain the WWRs of hundreds and even thousands of buildings. Conventional methods such as oblique photography and expert judgement both have limitations. The AI developed by combining semantic segmentation and convolutional neural network is able to automatically recognize the solid walls and windows in a building elevation image. Its accuracy, both compared with manual calculation and known values, is satisfactory. Therefore, the technique can be applied to the acquisition and determination of the WWRs for urban buildings.

REFERENCES

1. The United Nations. *World Urbanization Prospects: The 2018 Revision*. 2018.

2. Energy performance of buildings, European Commission. https://ec.europa.eu/energy/en/topics/energy-efficiency/energy-performance-of-buildings/overview#content-heading-0. As of 3 December 2019.

3. *Energy Standard for Buildings Except Low-Rise Residential Buildings*, ANSI/ASHRAE/IES Standard 90.1, 2019

4. *Design Standard for Energy Efficiency of Public Building* (GB 50189-2019), Ministry of Housing and Urban-Rural Development, China, 2019.

5. Reinart, C., Davila, C. Urban building energy modeling – A review of a nascent field. *Building and Environment*, 97(2016), 196-202.

6. Liang, J., Gong, J., Liu, J., Zou, Y., Zhang, J., Sun, S., Chen, S. Generating Orthorectified Multi-Perspective 2.5D Maps to Facilitate Web GIS-Based Visualization and Exploitation of Massive 3D City Models, *International Journal of Geo-Information,* 5(2016), 212.

7. Davila, C., Reinhart, C., Bemis, J. Modeling Boston: A workflow for the efficient generation and maintenance of urban building energy models from existing geospatial datasets, *Energy,* 117(2016), 237-250.

Spatial Interpolation of Outdoor Illumination at Night using Geostatistical Modeling

Bess Krietemeyer, Jason Dedrick, Camila Andino, Daniela Andino

Syracuse University
Syracuse, NY, USA
{eakriete, jdedrick, candino, danino}@syr.edu

ABSTRACT

There is a growing interest in studying spatial and temporal variation of artificial nighttime light in order to design safe and ecologically-balanced lighting strategies. Geostatistical analysis tools can aid in characterizing spatial illuminance patterns and interpolating values for locations where light samples were not taken. However, much of the research on spatial interpolation of light has studied sky brightness or ground illumination, without focusing on illuminance as perceived by the human eye. This paper investigates how geostatistical modeling can be used to visualize and predict outdoor illumination at night at the pedestrian scale. Field measurements of vertical illumination are taken across a campus environment and mapped using GIS software. Geostatistical models using kriging techniques are developed and cross-validated to evaluate prediction certainty. Results suggest that the use of geostatistical analysis to map and analyze the spatial distribution of illuminance across a campus area has potential to better predict the nighttime light environment and offers a promising method for producing spatially-complete maps of illuminance ranges to understand existing conditions and identify future needs for exterior lighting design. A discussion includes opportunities and improvements to the methodology for predicting illuminance, as well as considerations for the dynamic characteristics of nighttime lighting and the variable visual qualities of the surrounding built environment.

Author Keywords

GIS; Spatial interpolation; Illumination; Night

ACM Classification Keywords

CCS—Computing Methodologies—Modeling and Simulation—Model Development and Analytics (Model Verification and Validation; Uncertainty Analysis)

1   INTRODUCTION

1.1   Outdoor lighting in sustainable design and planning

Architecture, urban design and planning, and information systems have related and ongoing roles in developing innovative, sustainable solutions to urban development challenges. Artificial lighting accounts for about 8% of electricity use in the U.S. commercial and residential sectors, with resultant environmental impacts from electricity generation [5]. Recent smart city initiatives addressing issues

of energy efficiency and safety in outdoor spaces have led to the replacement of street lighting with solid-state lighting, such as light emitting diodes, or LEDs. The transition to LED lighting across the globe is well underway. In the United States, 10% of all street lighting has been converted; New York City is changing all 250,000 of its street lights, and Milan in Italy was the first city in Europe to do so on such a scale, with results visible from space [22]. Although the transition to LEDs can increase efficiency and lower cost, some lighting researchers and designers argue that the negative impacts of this transition far outweigh the benefits for human health and the environment. For many metropolitan areas, global use of artificial lighting at night has actually increased [15]. While the energy savings may be large, the use of LED lighting can increase light pollution, ecological impacts, and environmental degradation, for instance if lower operational costs encourage installing more lighting (the so-called rebound effect) [12]. Measures to mitigate the negative consequences of wasteful lighting practices are becoming of critical importance as the general growth of lighting of all types continues around the world [14]. Municipalities are feeling increasing pressure to incorporate sustainable night-time illumination policies and design guidelines into their urban lighting master plans [22].

These issues pose significant research and design challenges. For one, there are already contradicting demands put on a city's exterior lighting for balancing energy efficiency, safety, human health, and light pollution. While exterior lighting has the potential to improve safety, it could encounter challenges associated with poorly designed lighting in dark areas, creating real and perceived security risks [18]. If luminaires (lighting units) exist but their light distribution and control systems are not matched to the demands of the location and perceptions of users, they can lead to unintentional increases in energy use and light pollution without having the intended impact on community sustainability and safety. Good lighting design can mitigate some problems through the design of luminaires that are shielded and directed downwards, or by integrating intelligent controls and wireless sensing systems for dimming, detecting, and tracking movement into lighting design strategies [22]. To understand where such strategies could be implemented, as well as how best to adapt them to

the needs of different social and physical contexts for future applications, requires knowledge on the intensity and spatial characteristics of current outdoor nighttime lighting conditions.

1.2 Measuring outdoor light at night

The intensity of light depends on the light source and the direction in which it radiates light. The amount of light falling on a surface is known as illuminance and is measured in lux. In photometry, this is used as a measure of the intensity, as perceived by the human eye, of light that hits or passes through a surface. Devices such as hand-held cameras and light meters are commercially available for the evaluation of illumination—measured in lux or klux—for both interior and exterior spaces. Instruments such as these can be used to measure ambient light at night (ALN) for local studies at a single point in time. Photometric instruments measure light in the visible spectrum and are designed so that their spectral response matches a simplified theoretical response curve for the human visual system [3,9]. A terrestrial photometric sensor and hand-held data logger were used in the work presented here and will be described in Section 2.1.

Many examples of research studies examine the impact of outdoor lighting by conducting in-situ measurements with light meters to quantify light at night. Several of these investigate the conditions on campus environments, where students, faculty, and staff are often walking outside during dark hours. One study at the University of Minnesota-Duluth measured light intensity and compared those values to crime rates in order to identify problem areas and safe routes for students [10]. Another study at Princeton University measured lux levels along pedestrian and bicycle pathways to catalog existing light fixtures, create a lamp identification system, and design strategies for reducing energy use by outdoor lighting [19]. A study at the University of Washington Seattle collected photographs and mapped luminaire types across campus to analyze their color temperatures as they related to perceived qualities of light [4]. Each of the studies measured illuminance and mapped values as single points in time and space. This provides a useful snapshot of specific locations, but does not provide a complete picture of the variation in light intensity across all areas due to the time and cost of conducting the measurements.

1.3 Geospatial visualization and interpolation

Geospatial visualization techniques have been used to map lighting conditions across a range of spatial scales, as indicated by the campus studies. However, it is impossible in practice to take measurements at every location in an area of interest to get a complete picture. Recognition of the wide-ranging effects of artificial outdoor lighting at night has led to growing interest in studying spatial and temporal variation in the nighttime light environment. This is especially important as the use of LED lamps becomes more widespread, with greater variation in color, dimming capacity and switching speed, leading to potentially ecologically significant heterogeneity in light intensity, duration and spectra [8]. Of fundamental importance for the design of new or renovated lighting systems is the ability to appropriately characterize the nighttime light environment at a given place and time [20]. This requires models or simulations to estimate light levels in between field measurements, in order to obtain more complete views of lighting conditions and their impacts on perceptions of safety, sky brightness, and light pollution.

Geostatistical analysis methods can be used to describe spatial patterns and interpolate values for locations where samples were not taken. They can also provide measures of uncertainty for those values. This is important for informed decision making, as the uncertainty value allows for further information on the possible outcomes for each location rather than just one interpolated value [6]. Several research studies have utilized geostatistical modeling for examining and estimating light levels at local, regional, and global scales, with the aim of producing spatially-complete maps of illuminance or ALN. Garstang (1986) developed a model that has become essentially the standard in the field for predicting sky brightness produced by artificial lighting [7]. Biggs et al. (2012) measured night-sky brightness in Perth, Australia with upward-pointing sky quality meters (SQMs) and spatially interpolated the results to produce a map of night-sky brightness [1]. A study by Xu et al. (2018) compared the use of field observations and high-resolution remote sensing imagery to map ALN. They developed machine-learning algorithms to generate inputs for models that predict the in-situ illuminance measurements from high-resolution remote sensing imagery data. They assessed the models' capacity to accurately portray intra-urban variations in ALN [20].

Spatial interpolation methods have been used for predicting illuminance at the building scale. Jin et al. (2017) used the inverse distance weighted interpolation (IDW) method to study the vertical illumination from residential building windows at night [13]. Yi (2016) investigated the use of kriging modeling for interpolating indoor daytime illuminance as compared to advanced physics-based light simulation and energy simulations at the scale of an interior room. The kriging model was validated with indoor illuminance measurements and showed to be a better predictor than the physics-based energy simulation daylight model [21]. The probabilistic method of kriging incorporates randomness and allows for the inclusion of variance and statistical significance of the predicted values [2]. Kriging offers numerous advantages over other linear interpolation methods, such as IDW [1]. There are many types of kriging methods with varying assumptions. Ordinary kriging, which is the most common, uses a weighted average of neighboring samples to estimate the unknown value at a given location.
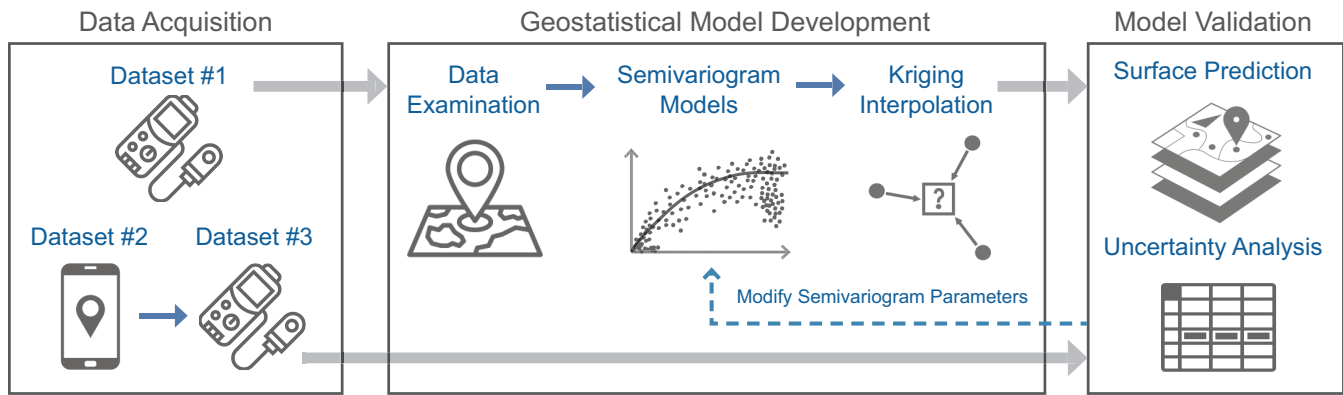
Figure 1. Diagram of the methodology for data acquisition, geostatistical model development, and model validation.

1.4 Research question and objectives

Much of the research on spatial interpolation of illuminance has studied impacts of artificial lighting on light pollution, measuring sky brightness or ground illuminance, without necessarily focusing on vertical illuminance as perceived by the human eye. As cities and campus communities continue to examine outdoor lighting strategies for energy efficiency and human circadian health and safety, there is a need to spatially interpolate vertical illuminance at the pedestrian scale to understand how existing lighting systems can be improved to increase safety and reduce energy use. Given this need, this research asks: can geostatistical analysis and interpolation methods, such as kriging, be used to predict nighttime lighting conditions across a campus area, using vertical illuminance data measured at the pedestrian scale?

This research examines the ordinary kriging interpolation technique for predicting nighttime illumination across a university campus. Using field measurements of outdoor vertical illumination, the ArcGIS geostatistical analysis software is used to generate and cross-validate experimental semivariogram models for spatial interpolation. Field-measured and interpolated values of illuminance are compared to determine which semivariogram models produce the most reliable set of predicted illuminance values, and to quantify the certainty of those models.

This work investigates the potential of geostatistical analysis tools for generating spatially-complete maps of outdoor nighttime illumination. The broader aim is to study how geospatial and geostatistical analysis tools can be used to visualize and predict nighttime lighting conditions to aid in the design of lighting strategies that meet criteria for ecological and human well-being.

2 METHODOLOGY

A university campus area located in the Northern United States was used for this study. Because of its northern latitude, the sun sets as early as 4:30pm during winter months, leaving many dark hours during which the campus operates. The size of the campus area is approximately 720 acres and includes over 160 buildings. A mixed-methods

approach is used to collect, analyze, and spatialize quantitative data within the campus study area. The methodology consists of three main stages: data acquisition, geostatistical modeling, and model validation. Figure 1 provides a diagram of the methodology.

1. Data acquisition: Collected illuminance field measurements across the campus that were used as inputs for the interpolation models and for model validation.

2. Geostatistical model development: Conducted geostatistical analysis and visualization of data, cross-comparison of semivariogram models, and kriging interpolation.

3. Model validation: Studied the accuracy of the interpolation models through cross-validation and through field measurements of illuminance data to determine model reliability and quantify uncertainties.

2.1 Data acquisition

Field measurements were conducted to measure three sets of data: The first includes field measurements of illuminance that were taken along campus pathways and parking lots (dataset #1), the second includes user perception data collected via a mobile application (dataset #2), the third includes field measurements of illuminance taken at the same locations as the user data (dataset #3). Each dataset is further described below. A portable device for measuring and logging data was assembled comprising a tripod with two LI-210R photometric sensors and an attachment for an LI-1500 data logger for data acquisition and storage. The photometric sensors have a spectral response between 400 nm and 700 nm which closely matches the CIE Standard Observer Curve and a cosine response that enables them to be sensitive to light from all directions up to an 82° angle of incidence [16]. One sensor was mounted to the tripod at 5 feet off the ground pointed outward for vertical eye-level measurements, and the other sensor mounted at 1'-0" height pointed towards the ground for horizontal measurements. For each photometric measurement (lux), the data logger captures information on date (yyyymmdd), time (hhmmss), latitude, longitude, elevation from sea level, and a measure of the geometric

quality of a GPS satellite configuration in the sky, determining the relative accuracy of a horizontal position.

The first set of illuminance data, dataset #1, included field measurements along pathways and parking lots. A total of 1150 illuminance measurements were acquired across the campus using the photometric sensor. The vast majority (80%) of this data was collected in the non-summer months (approximately October to April) of 2017 and 2018 between the hours of 10:00pm and 4:00am. Measurements were acquired along pedestrian pathways in regular increments every 10 feet, in parking lots along the central driving area every 10 feet, and at each stand-alone light fixture of the campus. Each data collection point consisted of 5 lux readings: four readings were taken at eye level 5 feet off the ground and in all cardinal directions (north, south, east, west) to represent the vertical illuminance that the human would view. These readings were averaged for a single reading at that point. This method for illumination data collection matches those of previous studies [13,17,19]. The fifth reading was taken 1 foot off the ground pointed down to collect measurements at ground level, in order to obtain information related to how illumination at the horizontal surface might impact modes of mobility, such as bicycling.

The second set of data, dataset #2, included user data collected via a mobile application. This measured user perceptions, not actual illuminance levels. The project team developed a mobile app to anonymously rate perceived levels of safety across campus. Features of the app included: GPS tracking to locate the latitude and longitude of each rating; the option to rate the perceived safety of a location on a scale of 1 to 5; a recording of the date and time of ratings; and the option to enter a short comment about why a particular rating was given. The app was publicly available and actively distributed to 300 students, and ratings were recorded across campus between Fall 2017 and Winter 2018. Users were recruited in large lecture courses and received extra credit for participation. Users had to download the app, and then open the app and record data at locations of their choice. Approximately 125 useful data points were acquired after removing data that were reported from inside a building, beyond the perimeter of the campus area, or during any hours outside of the 10pm to 4am time frame. These data represent approximately 125 students.

Following collection of the user ratings via the app, another set of illuminance field measurements (dataset#3) were taken at the user rating locations that would be used as an independent dataset to test the kriging interpolation of dataset #1. These data were collected in Fall 2019 between 10:00pm and 4:00am. A Garmin eTrex Touch 35 GPS device was used to locate the latitude and longitude of previously recorded user ratings, so that the illumination of those locations could be measured. The same equipment and method for measuring dataset #1 was used in the dataset #3 illumination measurements.

2.2 Development of geostatistical model

A geostatistical model was developed involving three steps:

Step 1: Examination of data through mapping and non-spatial analysis. First the dataset #1 lux levels were mapped using the ArcGIS Pro software to visualize spatial characteristics of the data. This allowed us to assess whether the data was taken at the appropriate scale. Each measured point was mapped and color-coded to indicate different lux ranges (0-5, 5-10, 10-15, 15-20, 20-25, 25-30). The initial geospatial mapping of dataset #1 (Figure 2) showed zones where lux levels were generally higher, such as pathways commonly used by the community on evenings for campus sporting events, or areas where there was greater commercial activity and pedestrian and car flow, and thus most likely more street and exterior building lights. It also showed zones that were generally lower, such as areas along the periphery where students often walked to off-campus housing, or places in between campus buildings that were less active during nighttime hours.

The data were examined to obtain reliable values for kriging interpolation. First, points were removed from dataset #1 that were flagged as potentially unreliable when recorded in the field, due to highly reflective surface conditions or proximity to campus construction sites. Next, illuminance measurements from dataset #1 that were nearest the locations of dataset #3 were compared. Illumination levels that were off by a factor of 4 or greater than 50 lux were removed from both datasets. 118 data points were deemed reliable from datasets #1 and #2. This sampling scheme created a quasi-random sampling that satisfies the geospatial statistics requirement that in order to minimize sampling bias in a two-dimensional map the samples should be acquired at random positions within each grid element [1]. Figure 3 shows a histogram of the reduced dataset #1.
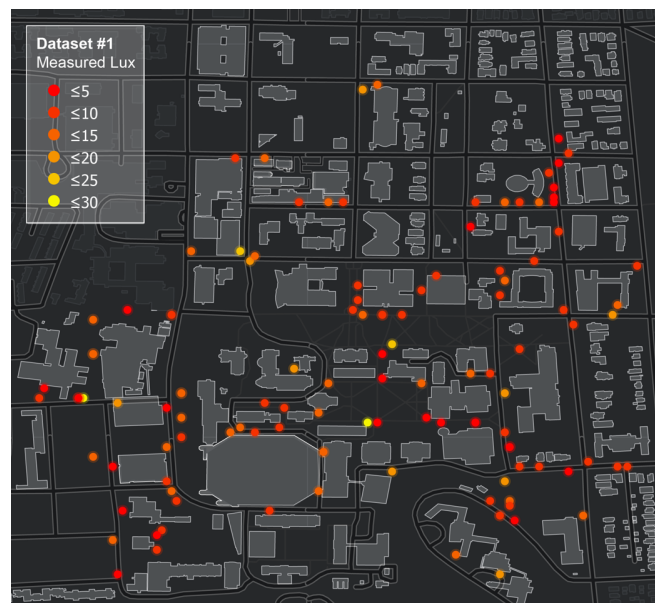
Figure 2. Map of dataset #1 using ArcGIS.
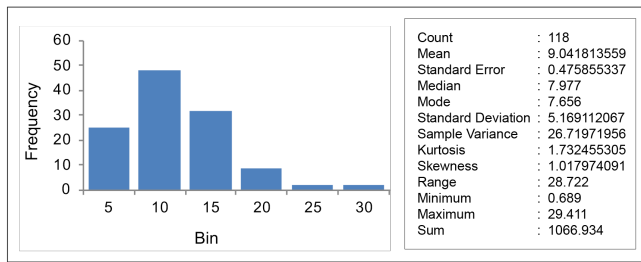
Figure 3. Histogram of dataset #1.

Step 2: Cross-comparison of semivariogram models. The semivariogram model is a key component in kriging. It is a function that describes the differences (variance) between data samples separated by varying distances. It quantifies the assumption that things nearby tend to be more similar than things that are farther apart. Semivariograms estimated from sample data are empirical semivariograms and are represented as a set of points on a graph. A function is fitted to these points and is known as a semivariogram model [6]. The ArcGIS Geostatistical Analyst tool was used to compute ten semivariogram models using dataset #1. Cross-comparison of semivariogram model results allows for comparison of the predicted value to the observed value and provides information about the quality of the different semivariogram models used for kriging. The semivariogram model outputs include a graph with a covariance curve that represents trends in the empirical data, shown as plotted points. The semivariogram curve that best fits the empirical data should pass through the center of the 'cloud' of plotted points, or 'binned' values, and pass as closely as possible to the averaged values. Figure 4 shows an example of the semivariogram curve with binned values in red and averaged values in blue.
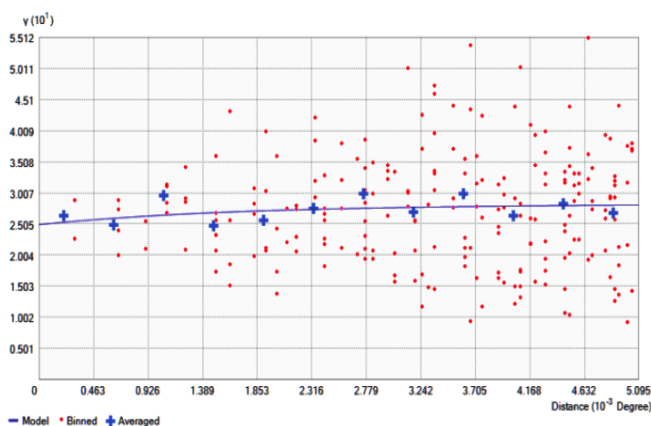
Figure 4. Graph of a semivariogram for the K-Bessel model.

Step 3: Map generation using ordinary kriging interpolation. The kriging interpolation provides an optimal prediction surface and a measure of confidence of how likely that prediction will be true. The ordinary kriging method is a flexible spatial interpolation method that can accommodate changes in the mean value of the surface. Weights are optimized using the semivariogram model, which indicates

the location of the samples and the relevant interrelationships between known and unknown values. This technique also provides a "standard error" which may be used to quantify confidence levels. Each semivariogram model was used to generate interpolation maps and to test how well they predicted the illumination values.

2.3 Model validation
Cross-validation was used to assess the semivariogram models in terms of kriging prediction accuracy. It allows for comparison of the predicted value to the observed value, and it provides information about the quality of the different kriging semivariogram models. The ArcGIS Geostatistical Analysis software includes a cross-validation tool using the same dataset that produces the semivariogram model. In this process, measured points are removed from dataset #1 and all remaining points are used to generate a prediction surface. The predicted value for the points that were removed are then compared to the measured value, and statistics are generated to determine the certainty of the prediction.

The semivariogram produces a quantification of predicted errors including: average standard error (ASE) (least average margin of error), root mean square prediction error (RMSE) (greatest prediction accuracy), difference (D), average error (AE) (least margin of error), mean standardized error (MSE) (least model bias), and root mean square standardized error (RMSSE) (least model bias). Results of interpolation models are provided in surface prediction maps and tables with predicted values. Ten models were assessed based on how well they predicted illuminance within certain lux ranges (0-5, 5-10, 10-15, 15-20, 20-25, 25-30). The predicted values were compared to the measured values from dataset #1.

The best-performing semivariogram models were further tested using dataset #3, an independent dataset that was collected at a later date under similar conditions and by a different person. Measured values from dataset #3 were compared to predicted values of the same locations. Using dataset #3 for validation allowed for the use of an independent set of field-measured illuminance values, so no points needed to be removed from the interpolated dataset. This avoids model bias as the ArcGIS cross-validation tool uses all the data to both estimate the trend and autocorrelation models. In addition to the scatterplot analyses, interpolated maps of all semivariograms were spatially compared to the field measurements from dataset #3 using overlay analysis.

3   RESULTS

3.1 Cross-comparison of model predicted errors
The semivariogram models were compared according to the semivariogram graph covariance curves and the predicted error values to assess which model would generate the closest prediction surface for illuminance values. Cross-comparison of predicted error values showed minimal difference in errors between models (Table 1). The models with the lowest margins of error (measured by ASE and AE)

were J-Bessel and Rational Quadratic. The models with the least bias, or lowest tendency to over-predict or under-predict values (measured by MSE and RMSSE), were Rational Quadratic and Exponential. The models with the lowest RMSE values were K-Bessel, Stable, Gaussian, and J-Bessel, respectively. RMSE indicates how closely the model predicts the measured values. K-Bessel was the lowest at 5.338 and Circular was the highest at 5.372, demonstrating little variation in errors across the models.

3.2 Cross-validation of kriging interpolation

The models were validated using the ArcGIS cross-validation tool, in which data samples from dataset #1 were removed for kriging, and the interpolated values were compared to the measured values. According to the geostatistical cross-validation analysis of predicted errors shown in Table 1, the K-Bessel, Stable, Gaussian, and J-Bessel semivariogram models had the closest predictions to the measured values. However, according to kriging interpolation, these models did not predict illumination values within certain lux ranges as well as other models. Prediction accuracy differed significantly depending on the number of measured data points and on the specific lux range in which measured and predicted values were compared (Table 2). Overall, predictions most closely matched those of measured values within a lux ranges of 5 to 10 for all

semivariogram models, averaging 74% accuracy out of a total of 48 measured points. The models that performed the best were J-Bessel, Hole Effect, Pentaspherical, and Tetraspherical at 75%. Predictions averaged 48% accuracy for the range of 10 to 15 lux, which included 32 measured values. Predictions were extremely low for the lux range of 0 to 5, averaging only 2% for all models despite the total of 25 measured values. For the higher lux ranges (15 to 20, 20 to 25, and 25 to 30), there was a much lower data sample (9, 2, and 2 samples), and prediction accuracy was 0% for each range above 15 lux.

The low accuracy for higher lux levels may be due to the small sample size. The poor performance at low lux levels is more likely to be a result of limitations in the models, or could reflect limits in the ability of the instrument to accurately measure illuminance at low levels.

3.3 Model validation with independent dataset #3

The predicted surface for the J-Bessel model was used to compare interpolated values to the 118 measured samples of the independent dataset #3. The prediction certainty overall was 40.7% for all lux ranges, and 71.4% accurate for the lux range of 5 to 10. These findings were similar to the kriging cross-validation results for dataset #1. A map of the predicted surface with dataset #3 is shown in Figure 5.

Semivariogram Model	Average Standard Error (ASE)	Root Mean Sq Prediction Errors (RMSE)	Difference (D)	Average Error (AE)	Mean Standardized Error (MSE)	Root Mean Sq Standardized Errors (RMSSE)
Stable	5.30331576	*5.34092633*	*0.03761057*	0.03088099	0.00590872	1.00816662
J-Bessel	*5.29655603*	5.34267447	0.04611844	0.02068223	0.00402837	1.00954201
K-Bessel	5.30245211	*5.33884856*	*0.03639645*	0.01297600	*0.00256883*	*1.00782568*
Hole Effect	*5.29812541*	5.34377043	0.04564502	0.02193561	0.00430012	1.00949327
Rational Quadratic	5.30299993	5.34592098	0.04292105	*0.01114996*	*0.00215596*	1.00891882
Gaussian	5.30331576	*5.34092633*	*0.03761057*	0.03088099	0.00590872	1.00816662
Exponential	5.30939087	5.34531581	*0.03592493*	0.01740526	0.00321788	*1.00761251*
Circular	5.30290350	5.37246749	0.06956399	0.01939569	0.00378988	1.01409469
Spherical	5.30483251	5.34420638	0.03937387	0.02743037	0.00523317	1.00862099
Pentaspherical	5.30285372	5.34283782	0.03998411	0.01998778	0.00377935	1.00833274
Tetraspherical	5.30346226	5.34737978	0.04391752	*0.01436445*	*0.00272696*	1.00906517

Table 1. Cross-comparison of predicted error values for ten semivariogram models. Shaded cells indicate lowest values.

Lux Range	0 to 5	5 to 10	10 to 15	15 to 20	20 to 25	25 to 30	All lux
Measured Data Count	25	48	32	9	2	2	118
Average Prediction Certainty (%) for All Semivariogram Models	2%	74%	48%	0%	0%	0%	43%

Table 2. Average prediction certainty for all semivariogram models within specific lux ranges.
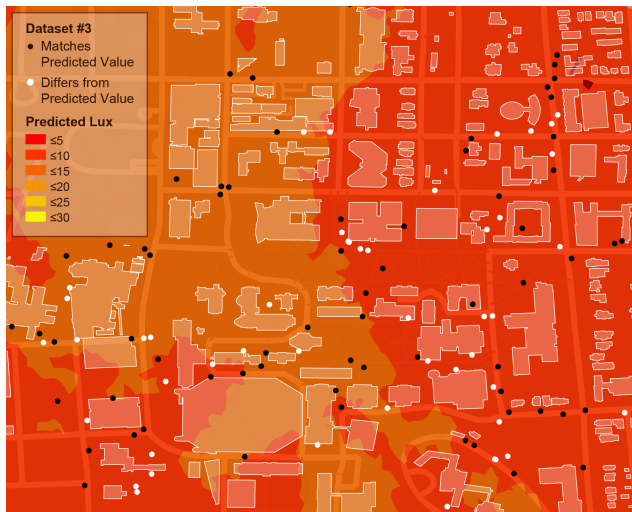
Figure 5. A surface prediction map using the J-Bessel semvariogram model with point locations of dataset #3.

4 DISCUSSION

4.1 Discussion of results and limitations

The results suggest that the use of geostatistical analysis to map and analyze the spatial distribution of illuminance across a campus area has potential to predict information on the nighttime light environment with complete areal coverage and spatial detail. Results of cross-validation using spatial and non-spatial methods demonstrated that the ordinary kriging interpolation technique was most accurate for predicting the illumination data for a range of 5 to 10 lux, which is the Illuminating Engineering Society's recommended maintained illuminance levels for pedestrian sidewalks and bikeways [11]. However, prediction certainty was very low for the 0 to 5 lux range and significantly decreased for any lux ranges above 15. Therefore, the kriging method was only useful for a relatively narrow range of vertical illumination data.

Error prediction results demonstrated that the semivariogram models were unbiased. They also showed that the models produced errors due to the low variability of the data and underestimations of the variability of the data. For example, the root mean squared standardized errors were slightly over 1, indicating that the models were underestimating the variability in their predictions. This shows that the actual differences among values could be higher than predicted; therefore, the model might predict a higher value. This would bias the results and make it appear from the model that the whole campus is brighter than it actually is. As a tool for decision-making in sustainable lighting design, this could be misleading, as it could overlook areas that appear adequately lit but actually require better lighting for safety, or lead to overestimations of light pollution.

The prediction uncertainty and underestimated variability in the predictions could be due to a number of issues, including the lower number of data samples to pull from, since the greater number within each range yielded a higher percentage of correct predictions. The low variability of the measured data did not allow us to test kriging using a small sample zone, which would have been a useful validation step prior to analyzing the complete campus study area. The prediction uncertainty is also likely related to the equipment used for data collection in combination with the nature of the phenomena being measured, impacting predictions for both extremely low and high lux values. Lower vertical illuminance values are more challenging to measure accurately in the field, and the photometric sensor used in this study might not have the sensitivity to capture the accurate measurements. Higher vertical illuminance values are also challenging to measure where there is a bright light source in stark contrast to its dark surroundings, potentially skewing the average cardinal readings for that location.

4.2 Future work

While the interpolated maps provide a description of the illuminance distribution in the study area, there are a number of improvements that can be addressed in future mapping projects. The method of measuring illuminance at eye-level and at four cardinal directions could be improved to more accurately reflect the lighting conditions for each area. While this averaged value takes into consideration the surface luminance of buildings, foliage, benches, signs, and ground surface within its field of view, it could also lead to over-simplification of nuanced conditions that are likely to vary in spectral reflectance over time, such as during wet or dry evenings, or with the fluctuating presence of light coming from building interiors. These errors could be reduced by examining the vertical illuminance data alongside the ground surface illuminance measurements. It is also recommended that data be collected only on dry evenings, and in a shorter and more consistent time periods, such as during a day or week when building interior lights follow more predictable schedules based on occupancy patterns. Improvements are also needed in the geostatistical analysis methods. In order to improve prediction certainty of the models, further model corrections are needed to modify the semivariogram parameters to better fit the data. Future work should consider model corrections that account for reflective surfaces, vegetation, and building shadows and interior lighting. The user safety ratings and HDR photos from this research will be important resources to examining illuminance relative to these social and physical conditions of the built environment.

Given that the nighttime light environment is highly dynamic at multiple temporal scales (i.e. interior building lights, car headlights, cloud coverage and moon phases) a single map of light at night can never be perfectly representative for all times of night and all environmental conditions. Development of methods to spatialize outdoor illumination for various diurnal and seasonal conditions will be an important area of future work. With improvements to data collection and spatial interpolation techniques, maps such as these can aid as tools for the design of exterior buildings and landscape lighting strategies that are adapted to the needs and conditions of the specific area. Further research opportunities

exist to study the spectrum of light sources and their impacts on safety as perceived by individuals—this would require different instruments than the one used in this study.

5 CONCLUSION

This paper described the use of geostatistical analysis methods for spatial interpolation of outdoor nighttime lighting. Field measurements of illumination across a campus environment were mapped using GIS, and a series of geostatistical models were developed and cross-validated to identify the models that had the highest prediction certainty. Results demonstrated how the geostatistical analysis and visualization tool offers a promising method for producing spatially-complete maps of illuminance within certain ranges for better understanding existing conditions and future needs for exterior lighting design.

REFERENCES

1. Biggs, J.D., Fouché, T., Bilki, F. Zadnik, M.G. Measuring and mapping the night sky brightness of Perth, Western Australia. *Monthly Notices of the Royal Astronomical Society 421, 2* (2012), 1450–1464.

2. Burrough, P.A., McDonnell, R. *Principles of Geographical Information Systems*. Oxford University Press, Oxford, 1998.

3. CIE ISO23539:2005(E)/ CIE S 010/E:2004. *Photometry – The CIE System of Physical Photometry.* Standard, CIE, Vienna, Austria, 2005.

4. Douglas, K. Campus Illumination: A Roadmap To Sustainable Exterior Lighting At The University Of Washington Seattle Campus. *A Report for the University of Washington's Integrated Design Lab* (2017).

5. EIA (Energy Information Administration). How much electricity is used for lighting in the United States. 2018. https://www.eia.gov/tools/faqs/faq.php?id=99&t=3. Accessed December 18, 2019.

6. ESRI. ArcGIS Pro: What is Geostatistics? https://pro.arcgis.com/en/pro-app/help/analysis/geostatistical-analyst/what-is-geostatistics-.htm. Accessed December 16, 2019.

7. Garstang, R. H. Model for Artificial Night-Sky Illumination. *Publications of the Astronomical Society of the Pacific*, *98, 601* (1986), 364-375.

8. Gaston, K.J., Davies, T.W., Bennie, J., Hopkins, J. REVIEW: Reducing the ecological consequences of night-time light pollution: options and developments. *The Journal of Applied Ecology 49, 6* (2012), 1256–1266.

9. Hänel, T. Posch, S.J. Ribas, M. Aubé, D. Duriscoe, A. Jechow, Z. Kollath, D.E. Lolkema, C. Moore, N. Schmidt, H. Spoelstra, G. Wuchterl, C.C.M. Kyba. Measuring night sky brightness: methods and challenges. *Journal of Quantitative Spectroscopy and Radiative Transfer 205* (2018), 278–290.

10. Hollingsworth, J.L. UMD students map Duluth's lighting levels. *Duluth News Tribune* (2009) https://www.duluthnewstribune.com/news/1446014-UMD-students-map-Duluths-lighting-levels.

11. Illuminating Engineering Society. *Lighting for Exterior Environments. IES Recommended Practice* (2014), RP-33-99, Table 6.

12. Jägerbrand, K. New framework of sustainable indicators for outdoor LED (light emitting diodes) lighting and SSL (solid state lighting). *Sustainability 7* (2015), 1028–1063.

13. Jin, X., Li, Y. Zhang, J., Zheng, J., Liu, H. An Approach to Evaluating Light Pollution in Residential Zones: A Case Study of Beijing. *Sustainability 9, 4* (2017), 652.

14. Kinsey, B.R., Miller, N.J. Light at night--A delicate balance. Pacific Northwest National Laboratory. *Report for Pacific Northwest National Laboratory* (2018). PNNL-SA-131714.

15. Kyba, C.C.M., Kuester, T., Sanchez de Miguel, A., Baugh, K., Jechow, A., Holker, F., Bennie, J., Elvidge, C.D., Gaston, K.J., Guanter, L. Artificially lit surface of Earth at night increasing in radiance and extent. *Science Advances 3* (2017), e1701528.

16. LiCOR 210R Photometric Sensor. https://www.licor.com/env/products/light/photometric. Accessed December 2019.

17. Miller N.J., Koltai, R., McGowan, T. Pedestrian Friendly Outdoor Lighting. *Pacific Northwest National Laboratory* (2013) PNNL-23085.

18. Rea, M.S., Bullough, J.D. Brons, J.A. Spectral considerations for outdoor lighting: Designing for perceived scene brightness. *Lighting Research & Technology 47* (2015), 909–919.

19. Seward, A. Light Mapping, Architectural Lighting Magazine. https://www.archlighting.com/industry/light-mapping_o. Accessed December 16, 2019.

20. Xu, Y., Knudbyb, A., Côté-Lussierc, C. Mapping ambient light at night using field observations and high-resolution remote sensing imagery for studies of urban environments. *Building and Environment 145* (2018), 104-114.

21. Yi, Y.K. Adaptation of Kriging in daylight modeling for energy simulation. *Energy and Buildings 111* (2016), 479–496.

22. Zielinska-Dabkowska, K.M. Make lighting healthier. *Nature 553* (2018), 274-276.

Vision-Based Lighting State Detection and Curtain Openness Ratio Prediction

Esat Kalfaoglu [1], **Ipek Gursel Dino** [12], **Orcun Koral Iseri** [12], **Sahin Akin** [12], **Alp Eren Sari** [1], **Bilge Erdogan** [3], **Sinan Kalkan** [1], **Aydin Alatan** [1]

[1]Center for Image Analysis (OGAM), Middle East Technical University, Ankara, Turkey
[2]Department of Architecture, Middle East Technical University, Ankara, Turkey
[3]Heriot-Watt University, Institute for Sustainable Building Design, Edinburgh, Scotland

ABSTRACT

In non-residential buildings, space lighting accounts for 17 % of the total energy consumption. Effective use of daylighting has great potential to reduce lighting energy use in buildings. The amount of daylighting through the building windows is influenced by shading devices (i.e. curtains) that limit the visible light being introduced to the space. Therefore, there is a strong relationship between artificial lighting and shading. In this work, artificial lighting state prediction and curtain openness ratio prediction from the visual data is studied. For lighting state prediction, local and global approaches have been proposed and the performances are improved with the addition of background modeling and light sensor information. For the curtain openness prediction, a method with background modeling, binarization and morphology has been proposed. The performance of the proposed methods is evaluated with video datasets captured during one week in December 2018 and one week in March 2019. It has been shown that promising results are obtained with a 92.79 % and 97.40 % correctness ratio for light state detection for the same weeks, respectively; and with a 2.55 % mean average percentage error for curtain openness ratio prediction.

Author Keywords

Human-building Interaction; Lighting State Detection; Shading Opening Ratio Detection

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING: Applications; I.4.9 IMAGE PROCESSING AND COMPUTER VISION: Applications

1 INTRODUCTION

Space lighting represents 17 % of the total building energy use in commercial buildings in the US [8]. Despite the increased use of energy-efficient light bulbs and the decrease in lighting energy use, the reduction of the total electric energy use remains a key issue. The efficient use of daylighting can reduce the need for space lighting, reduce lighting energy use

and improve indoor conditions, resulting in increased occupant productivity. However, excessive amounts of daylighting negatively impacts cooling energy consumption due to excessive solar heat gain and leads to glare-related discomfort. Therefore, the correct operation of shading devices (i.e. curtains) is critical for the right balance between energy efficiency and occupant comfort [14]. Monitoring of plug loads and the related occupant interventions can increase the precision of calculation of energy savings [1]. However, occupant behavior is difficult to represent accurately in energy models due to the stochastic nature of occupants' actions. Moreover, energy consumption and occupant density are loosely correlated in institutional buildings [1]. The differences of occupants' comfort preferences, satisfaction levels and indoor environment perceptions makes it difficult to reliably associate electric-related information with occupant presence in energy models [7]. Reliable performance predictions depend on realistic representations of the interaction between occupants and control systems [9].

To represent occupant-building interactions, the typical means is to use existing schedule libraries for different space types, such as those specified in ASHRAE [2]. However, these libraries cannot capture the temporal complexity of human behavior. The most common alternative is predictive models that attempt to model correlations between environmental variables (i.e. illumination, temperature) and the associated likelihood of occupants engaging in adaptive actions (i.e. changing setpoint, turning lights on). Regression models, as used in [4, 17], can describe the probabilities of occupant actions at various time intervals. Alternatively, agent-based approaches can model occupants as autonomous entities with unique behavioral characteristics, thereby integrating the behavioral diversity of multiple occupants. Other approaches, as identified by [6], include Bernoulli models that predict the probability of finding a building component with which occupants interact [10, 13], Discrete-time Markov models that predict the likelihood of an adaptive behavior in the next time-step [11], Discrete-event Markov models that associate adaptive actions to external events [13] and survival models that predict the lifetime of an occupant action or the state of a building component [12].

The current approaches aim to predict the lighting / shading states either as a function of one or more environmental variables that are directly associated with these components, or as based on the presence of building occupants that can change the component states. These existing approaches also need to make many assumptions to model these correlations, both spatially and temporally, which might not hold true in all cases. Therefore, these approaches can be said to be derivative, in that they are only reliable in certain pre-determined situations.

As opposed to the above-mentioned predictive models, we propose a vision-based method that detects the lighting on/off states and curtain openness ratio. Our video-content analytical approach uses an IP camera installed in a room. For lighting state detection (Section 3), two different approaches has been proposed and the performance is improved with background modeling and light sensor information of the IP camera. For curtain openness ratio prediction (Section 4), a three-stage algorithm that consists of background modeling, thresholding and morphology stages, is proposed.

The precise prediction of lighting and curtain openness in our work is important in two aspects. Firstly, the correlation between lighting and shading can help identify unnecessary use of lighting when the curtains are drawn. Secondly, the estimated values can be used in the energy modeling of existing buildings, as an alternative to the standard schedule templates. The increased precision in these schedules have the potential to increase the precision of simulation results and better-informed energy analysis of existing buildings. Finally, data captured during long-term monitoring can be used to understand the complex interactions between occupants, lighting and shading by machine learning models.

2 EXPERIMENTAL ENVIRONMENT, CAMERA AND DATASET

Since both problems (lightning state detection and curtain openness ratio prediction) use the same environment and the dataset, this section is introduced before explaining the proposed methods.

The environment chosen for both lightning state detection and curtain openness ratio prediction tasks is a classroom with area $49m^2$ (located at the Department of Architecture at Middle East Technical University). The room layout and the installed camera are shown in Figure 1a. As seen from the figure, the environment has six artificial light sources which are denoted as $L1, L2, ..., L6$. These six lights sources are connected to two switches which are denotes as $S1$ and $S2$ in Figure 1a. The light sources $L1$, $L2$ and $L4$ are connected to switch $S1$, and $L3$, $L5$ and $L6$ are connected to switch $S2$. The camera is located in the lower-right part of the layout and shown in red color. The horizontal field of view of the camera is also demonstrated in this layout in Figure 1a.

The field of view of the camera to the class can be seen in Figure 1b. From this figure, it can be understood that $L6$ is not in the field of the camera, and $L1$ cannot observed because of occlusion of projector. The four visible light source locations, the ceiling region and switch location are annotated
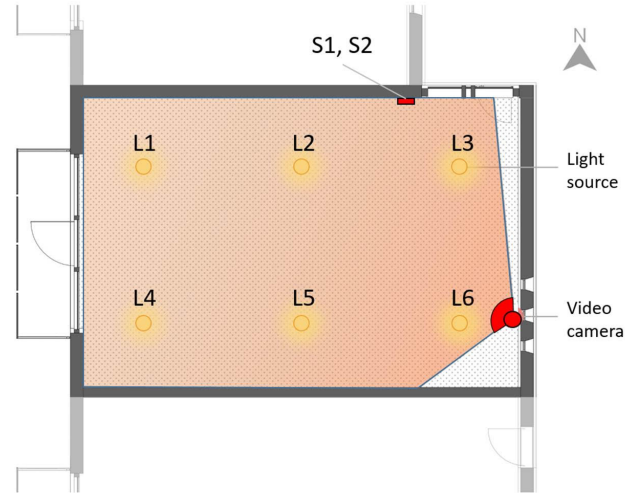
(a) The room layout for the experimental setup
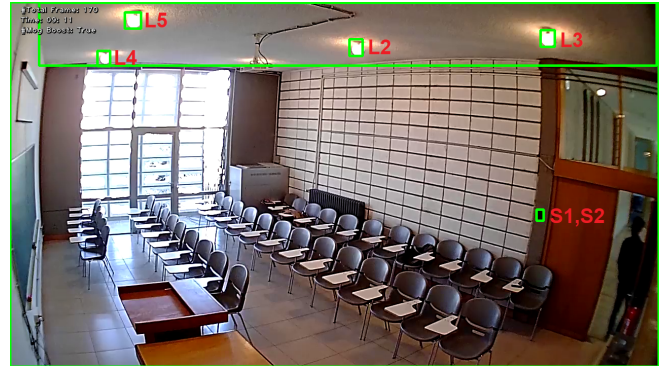
(b) The camera's view of the classroom. The regions of light sources, the ceiling and the switch location are denoted with green rectangular bounding boxes

Figure 1: The visual information about the experimental setup

using green rectangle bounding boxes as illustrated in Figure 1b.

The camera used for this experiment is a 135-degree wide angle IP camera. The horizontal and vertical field of view of the IP camera are about 112-degree and 62-degree, respectively. The videos have a frame rate of 10fps. The camera sends videos as 10-minute clips to the Amazon cloud when there is a movement in the field of view of the camera.

To test the performance of the proposed algorithms, the class environment is recorded twice for a duration of one week: in December 2018 (24.12.18 - 31.12.18) and one March 2019 (13.03.19 - 20.03.19). Additionally, to test the proposed methods of light state detection, 12 light opening closing event couples (in total 24 events) are collected in a controlled manner from three different times of a day, at 10am, 13pm and 16pm. The whole procedure for collecting these events takes at most ten minutes for each time of the day. Therefore, during this short period of time, there is not a challenging case that can be faced with in real life scenario which may

force the proposed methods to create false alarms. The possible challenges in real life scenarios will be analyzed in detail in Section 3.2. The proposed light state detection methods are tested on these cases before testing them on more challenging one-week recordings.

It should be noted that the proposed methods should not be considered as specific to this environment setup. They can be adapted and utilized in any environment – see also Section 5.

3 LIGHTING STATE DETECTION

This section presents the problem of detecting light states which are either turned on or off. In this section, the proposed methods are presented in Section 3.1. Then, challenges of the problem are explained in Section 3.2. Finally, the experiments and results are discussed in Section 3.3.

3.1 Proposed Methods

In order to handle the light state detection problem, it is preferred that catching the change in the light states, namely light events, is a possible solution. Mainly depending on this idea, two methods are proposed to detect the light events, namely detection via *local* illumination change and detection via *global* illumination change. The local method assumes that the location of every light source is known, while the global method tries to detect the changes from the full field of view or from only the ceiling. In addition, the fusion of light sensor information of IP cameras to the global approach is analyzed.

It is assumed that a video clip consists of N frames denoted as $f_1, f_2, ..., f_N$. For both methods, the average intensity of the pixels in a region of interest R_n of the frame f_n is calculated and can be denoted as i_n^R. These regions of interest are the pre-defined bounding boxes in the frame as shown in Figure 1b. It should be noted depending on the approach of the proposed methods, the union of the some of the regions in Figure 1b is also a possible solution to define a region. For example, the union of bounding boxes of all light sources in the field of view ($L2, L3, L4$, and $L5$) is possible to make a global region of all light sources. After calculating average intensity for the pre-defined regions of interest for all frames, we can analyze the signal $I^R = \left[i_1^R, i_2^R, ..., i_n^R \right]$ for our purposes, which reduces the problem to a 1D signal processing problem.

Detection via Local Illumination Change
This method assumes that the locations of the light sources (the bounding boxes in Figure 1b) are previously known and annotated. In this method, only the illumination changes in these known light regions are examined. This approach is not affected by other environmental factors (such as outdoor light) since only limited regions around light sources are considered, in contrast to the global approach in which the bounding box of ceiling or the whole frame is utilized. However, this approach is impractical in the sense that there is a need to manually annotate all light sources in advance, which may not be possible when there are too many light sources or when diffuse, wide-area light sources are used.

The local method follows these steps:

Step 1 - Noise Reduction: The signal I^R is smoothed with the Gaussian filter in order to eliminate the noisy samples:

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2} \right). \qquad (1)$$

The tap number T is defined to create Gaussian filter in discrete domain where $x = [-T, -T+1, ..., 0, ..., T-1, T]$, resulting in $2T + 1$ elements in $G(x)$ Gaussian filter.

Step 2 - Illumination Change Detection: Using a simple first-order filter like $F = [-1, 1]$ or $F = [-1, 0, 1]$, the illumination changes over time are detected. Gaussian and first order filtered signal can be denoted as $I_{filtered}^R = I^R * G * F$ where $*$ denotes the convolution operation.

Step 3 - Find Peaks and Dips in Changes: The method then finds the peaks and dips in $I_{filtered}^R$ to detect the events and a threshold is applied in order to eliminate the possible false alarms.

For the light state detection task, depending on the usage, it is considered that there are two possible versions of the algorithm. The first version is called "multiple switches" and assumes that lights might be controlled by different switches ($L2, L4$ connected to $S1$ and $L4, L5$ connected to $S2$) but all of the switches should be either on or off (Both $S1, S2$ are on or off). This assumption is to decrease the false alarms. However, the problematic aspect of this assumption is that it ignores the cases that some of switches can be on while the others are off. The second version called "single switch" assumes that there is only one global switch, to which all of the light sources are connected ($L2, L3, L4, L5$ connected to global S). The limitation of this approach is that it is more susceptible the false alarms as compared to the "multiple switch" assumption.

Detection via Global Illumination Change
This method utilizes the exposure change of the camera to the light change in the environment. The main idea behind the camera exposure change is that the camera adjusts its exposure to maintain the average intensity of the pixels to a certain extent. The IP camera used in this experiment keeps the average intensity of pixels at the level of about 110 (over 255). Therefore, if the average light intensity changes, the camera uses automatic exposure to restore the intensity level to 110. As a result, peaks or dips are created in the average illumination graph when the lighting state changes in the room. Similar observation is also made in the average intensity of the pixels that belong to the ceiling. One of the advantages of using ceiling is that the effect of light change is more noticeable due to the close proximity with the lights. Moreover, the ceiling is affected by the other environmental factors less as compared to the rest of the scene.

The global method follows these steps:

Step 1 - Mean Extraction: Because of the non-ideal usage of LoG filter in the discrete domain, the filter does not create a zero response for even a constant signal. Therefore, the mean of I^R is subtracted from I^R, resulting in a zero-mean signal I_0^R

Step 2 - Exposure Blob Detection: To capture the exposure blobs (light events) in I_0^R, the Laplacian of Gaussian (LoG) is applied to the illumination. LoG is the second-order derivative of the Gaussian filter and commonly used in blob detectors like SIFT [15] and Harris-Laplace detector [16] in order to find the scale of the blobs or interest points. The scale normalized LoG filter in 1D domain can be defined formally as:

$$LoG(x) = -\frac{x^2 - \sigma^2}{\sigma^3 \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right), \qquad (2)$$

with a tap number T, resulting in a LoG filter with $2T + 1$ elements. As a result, the LoG filtered signal is obtained as $I_{LoG}^R = I_0^R * LoG$. For exposure blob detection, it is desired that only the blobs belonging to light change are detected. For example, the curtain movement can also create exposure blobs in I_0^R. However, it would be a slower process with respect to light change. Therefore, the scale parameter (σ) of the LoG should be chosen suitably.

Step 3 - Find Peaks and Dips in Changes: The method finds the peaks and dips in I_{LoG}^R to detect the exposure blobs and a threshold is applied in order to eliminate the possible false alarms.

Step 4- False Exposure Blob Elimination by Zero Crossings: To eliminate the false blob detection in I_{LoG}^R, the zero crossings of the first derivative ($I_{filtered}^R = I^R * G * F$) at the locations of peaks and dips are also checked in order to decrease the false alarms.

Background Modeling

Optionally, for both local and global methods, the background modeling at the switch location(s) can be used to decrease the false alarms of the method from the other environmental light factors. In order to turn on/off lights, a hand should be around the light switch and should be detected as foreground by a background modeling algorithm. Background modeling is performed using Mixture of Gaussian (MOG2) algorithm [18]. The main idea behind this method is that every pixel is represented with Gaussian color clusters by using the first- and second-order statistics. The algorithm assumes that frequent clusters belong to the background and rare clusters belong to the foreground.

A visual demonstration for the background modeling utilization for light event detection is shown in Figure 2. The view that captures the instant that a person is starting to turn on the light is shown in Figure 2a. The background mask belonging to this view is demonstrated in Figure 2b. The region of the switch location in the background modeling mask is shown in Figure 2c, which includes some foreground part because of the hand reaching the switch.

Sensor Fusion

Both local and global methods have a similar limitation: They are *differential* systems. In other words, both methods try to detect the light state changes (lights on/off events) instead of predicting the light states directly. Therefore, if an event of turning on/off is missed, the method still assumes that the light is still on/off, which results in significant performance reduction. To address this problem, the light sensor installed
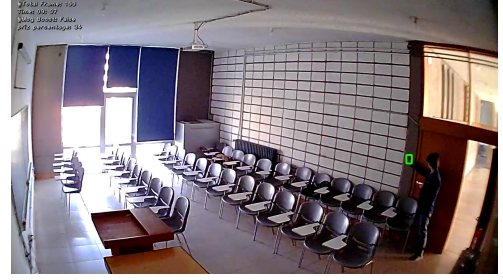
(a) The camera view capturing light turning on instance

(b) Background mask of view

(c) Background mask of switch region

Figure 2: Utilization of background modeling in light event detection

in the IP cameras is utilized. In this way, the current state of the light sensor can be fused with the change detection: if the sensor has high confidence that the artificial light is not on, then the light state is switched to turned off state.

3.2 Challenges

There can be other factors that influence the illumination levels in an environment, and potentially misinform the proposed methods: One of them is other light sources that have other purposes than illuminating the room, but nevertheless trigger a change in illumination. This can potentially affect both the local method to a certain degree, and the global method significantly. An example in our case study is the projector installed on the ceiling of the room, which can significantly change the illumination level when slides change. Another factor is a sudden change in daylighting conditions in the room (i.e. sunlight appearing behind a cloud and reaching the indoors). This event also affects the global approach significantly.

The strength of peaks and dips in changes of both algorithms may vary significantly according to the time of the day and the curtain openness ratio. According to the time of the day, the amount of daylight entering the environment changes. The curtain also controls the amount of daylight entering the environment. In a darker environment, a transition from lights-on to lights-off (and vice versa) is significantly noticeable as compared to a directly day-lit environment.

3.3 Experimental Results and Analysis

The analysis in this section can be categorized into two sections: the controlled and uncontrolled experiments. In the controlled experiments, 24 light turning on/off instances (see Section 2) collected for this task are used. This experiment

is "controlled" because a person, as part of the experiment setup, is required the turn on/off the switch. Moreover, these instances do not include the challenging factors identified in Section 3.2 because they only cover a short time period, thereby decreasing the possibility of observing unexpected cases. Therefore, with the controlled experiment, the general capability of the proposed method at detecting light events is aimed to be evaluated. In the uncontrolled experiments, video recordings of a one-week duration in December 2018 and March 2019 (see Section 2) are used. This dataset represents a more realistic scenario, as the changes in lighting conditions are induced by the actual room occupants. Moreover, covering a longer time period, this dataset has a greater possibility of observing possible challenging cases under actual human-building interaction conditions.

Implementation Details: For the parameters of the experiments, the tap number T mentioned in Section 3.1 is set to 12. The standard deviation (σ) in (Eq. 1 and 2) is set to four. To estimate the first-order derivative of the filter (denoted as F in Section 3.1), $[-0.5, \ 0, \ 0.5]$ is used. The peaks and dips in change in proposed methods are found by a function called 'find peaks' of the 'Scipy' library. The threshold for the dips and heights of the signal are determined as ten for the local approaches and six for the global approaches. These thresholds are set according to the observed performances in the uncontrolled experiment. Additionally, the derivatives are taken considering the time between the frame interval in the discrete domain, which is 0.1 second in this work (because the camera is 10 fps).

Performance Measures: To evaluate the performance of the proposed light state detection methods, four different measures are used. Three of the these measures are well-known measures which are precision, recall and F1-score. The other metric is called "correctness".

Precision and recall are calculated as:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}},$$
$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}, \quad (3)$$

where True Positive is the number of instances which are correctly found; False Positive is the number of instances which are falsely found; False Negative is the number of instances which exist but are missed by the method. F1-score is can be interpreted as the weighted balance between precision and recall, and is calculated as:

$$\text{F1-score} = 2 \frac{\text{Precision Recall}}{\text{Precision} + \text{Recall}}. \quad (4)$$

In addition, we use a correctness measure that evaluates the correct proportion of the estimated light states. To state it in mathematical terms, assume a predicted light state signal and ground truth signal in discrete domain denoted as $P(n)$ and $GT(n)$, respectively, where n is the time index of the signal and these signals obtains the value of 1 when the light is on, and 0 when the light is off. With these notations, correctness

Methods	Precision	Recall	F1 Score
Local (Single Switch)	100.0	100.0	100.0
Global (Without MOG)	91.3	87.5	89.4

Table 1: Results of controlled experiment

can be defined as:

$$\text{Correctness} = \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}(P(n) = GT(n)), \quad (5)$$

where N is the total number of time indices. Correctness deals with the performance in estimating the light states, while the other three measures evaluate the performance of detecting light events. Obtaining a high correctness measure is the final aim of the light state detection task. To make it more clear, for energy simulations, the number of turning on/off events that are detected correctly is not a final concern. Instead, estimating whether light sources are on or off correctly for every time instant is important.

Controlled Experiment: The controlled experiment consists of 24 events from the three different times of the day. The result of this experiment can be seen in Table 1. The results are given as precision, recall and F1 scores in percentage values. As seen in the table, local method (Single Switch) is better than the global method without background modeling (Without MOG).

Uncontrolled Experiment: For the uncontrolled experiment, the local (multiple switches), local (single switch), global without background modeling (without MOG) and global with background modeling (with MOG) are compared on the weekly data recorded in March 2019. During this week, there are five light on and off events (in total 10 events). In the 3rd event, only one switch out of two (so three out of six light sources) are turned on/off (See Section 2 for details of the experiment setup).

The results are provided on Table 2. For the correctness measure, the local (Multiple Switch) method is the most successful method, as shown in Table 2. However, in terms of the other three measures, the local (Multiple Switch) method is not the best. The reason for obtaining the lowest recall with the local (multiple switches) method is the requirement for two light switches turning on/off. To make it more clear, the 3rd instance contains only one switch turning on/off, which does not satisfy the requirement of the local (multiple switches) method; therefore, the 3rd instance is missed by the local (multiple switches) method. One of the reasons to obtain higher correctness values for the local method (multiple switches) is that 3rd closing instance is very challenging (Due to the time of the day and one switch turning on/off instead of two. For challenges, See Section 3.2). Because of the fact that 3rd opening instance is not detected in the local method (multiple switches), the error caused by not detecting the 3rd closing instance is not propagated to the other time instances,

73

Method	Correctness	Precision	Recall	F1 Score
Local (Multiple Switches)	**96.62**	70.00	70.00	70.00
Local (Single Switch)	83.15	64.29	**90.00**	75.00
Global (without MOG)	62.96	23.68	**90.00**	37.49
Global (with MOG)	84.34	**81.82**	**90.00**	**85.72**

Table 2: The proposed algorithm performances on March 2019

	Method	Correctness Ratio
December 2018	Without Sensor	44.66
	With Sensor	**92.79**
March 2019	Without Sensor	84.34
	With Sensor	**97.40**

Table 3: The performance improvements brought by sensor fusion

resulting in improved correctness for local method (multiple switches).

Another important conclusion that can be derived from Table 2 is that using a background model improves the performance of the global approach significantly. In the global method, without background, many false positives are observed. With background modeling, false positives are significantly eliminated, resulting in an increase in precision and therefore the F1-score.

For the final analysis, the performance improvement introduced by sensor fusion on the correctness of light states is investigated. The selected method to test the effect of sensor fusion is the global method (with MOG). The result of sensor fusion can be seen in Table 3. For this test, the weekly datasets from December-2018 and March-2019 are used. In this dataset, there are eight light on and off events (in total 16 events). As seen from the table, sensor fusion increases the correctness measure significantly especially for December-2018, where the correctness measure increases from 44.66 % to 92.79 %. Although the effect of sensor fusion is shown with global (with MOG) method, utilizing it in other methods is also possible.

A counter argument could be that if sensor fusion increases the performance significantly, it can also be used for the detection of light events without the visual data. However, it should be denoted that sensor fusion is utilized only for the missed closing events. When the room in properly day-lit, the light sensors might not be able to capture the opening events, because daylight -most of the times- results in stronger sensor output than the artificial light sources.

4 CURTAIN OPENNESS RATIO PREDICTION

The goal of the Curtain Openness Ratio Prediction algorithm is to estimate the portion of the window that is not covered by the shading devices (or most commonly the curtains). The same experimental setup as the light state detection is used (See Section 2). In this task, the four corners of the window are assumed to be already annotated by the users, so to mark the boundaries of the window region. In this section, firstly, the proposed method is explained in Section 4.1. Then, experimental results are given in Section 4.2

4.1 Proposed Method

The proposed curtain openness ratio prediction method consists of three stages. The four corners of the window area should be given as input to the proposed method in order to guide the method where it should look to. The stages of the algorithm are visualized in Figure 3. The proposed method mainly benefits from the fact that the shaded region of the window is darker as compared to the non-shaded regions of the window.

Stage 1 - Background Modeling: In the first stage, a background modeling algorithm (MOG2) [18] is implemented. The aim of background modeling here is to increase the robustness of our algorithm against instantaneous occlusions (i.e. people in the room that pass in front of the window). As seen in the Figure 3b, the two people who are occluding the window gradually became a part of the background instead of directly observing them in the scene as in Figure 3a. Background modeling was also utilized in light state detection, as explained in Section 3.1. The only difference in two implementations is that, in the lighting detection method, the background modeling output is in the format of binary mask (see Figure 2b), whereas in this method, the image is constructed as the weighted mean of the Gaussian color clusters detected as background in the background modeling algorithm (see Figure 3b). The weights are determined from the observation frequency of the clusters.

Stage 2 - Binarization: In the second stage, the binarization of the image is implemented. Binarization is the process of converting the information of every pixel into the one-bit encoding, which are black and white as seen in Figure 3c. For binarization, the output of the background modeling is converted from the color image to a gray-scale image. Then, for every pixel, a value greater than a pre-determined threshold L is set to white, and to black otherwise.

Stage 3 - Morphological Operation: After the binarization of the background-modeled image, in the third stage, a *closing* morphological operation is used. This operation is utilized in order to fill the black holes inside the white regions in the binarized image. This operation is beneficial in eliminating the negative influence of the occluding objects inside the room. For instance, the binarized output (See Figure 3c) illustrates that neither background modeling nor binarization can completely eliminate the occluding people. However, as seen in Figure 3d, not only the occluding people, but also the occluding chairs are eliminated from the binarized image. Moreover, some objects outside the room, such as buildings seen through the window, might not reflect the sunlight enough to

(a) The input

(b) Background image in background modeling

(c) Thresholding

(d) Closing Morphology

Figure 3: The steps of the curtain openness ratio estimation

surpass the pre-determined threshold of binarization. Morphological operations are beneficial also in eliminating the effects of objects outside which has a lower reflectivity, creating undesired holes inside the thresholded image.

4.2 Experimental Results

Implementation Details: For the threshold value L in binarization process, three values are tested and the best value is selected using visual inspection. These values are 140, 220 and 250, considering that the pixel values are scaled from 0 to 256. It is observed that using 140 results in a higher possibility of detecting some closed regions as opened regions, such as the curtain regions close to the edges where sunlight is diffused into. Using 250 causes missed detection of some portions of the open curtain parts. Therefore, 220, which gives the best compromise in our settings, is set as the threshold for the binarization stage. The kernel size used in morphology operation is set to 25×25 which is a relatively big kernel. This kernel size is also set by visual inspection. Using smaller kernel is not adequate in eliminating the imperfections because they have relatively bigger sizes. The further increase in kernel size may lead to destruction of correctly identified shaded regions.

Results: For the curtain openness ratio prediction test, 18 measurements are taken randomly throughout the one-week March 2019 dataset. These measurements are the proportion of non-shaded (open) regions inside the window area to the total area of the window. Non-shaded (open) regions are manually labeled by a person. For error calculation, the absolute difference between the predicted openness ratio percentage and the percentage of curtain non-shaded (open) region of the manually labeled ground truth is calculated. For example, for a given sample, if GT is 40% and the prediction is 43%, then the error is calculated as 3%. The average percentage error of all samples is calculated as **2.55** % following this rule from these 18 samples.

5 GENERALIZABILITY AND APPLICABILITY OF THE APPROACHES

The proposed solutions for both problems are generalizable to some extent. However, there are some cases where there is applicability might be challenging, or the parameters may need to be reconsidered based on the context.

For light state detection, the first problem is the characteristic of turning on/off of the artificial lights. For instance, certain light bulbs, such as fluorescent bulbs, can flicker when they are switched on, which can misinform the algorithm. Secondly, the IP camera works at 10 fps. When a camera with different fps is used, some of the parameters needs to be updated, such as the standard deviation parameter (σ) of the LoG filter (which specifies the scale of the blob) used in the global approach. Another concern is the number of switches that control lighting in the room. In the experiment environment of this work, there are only two switches, both of which are usually operated together by the users. For cases with multiple switches, the ratio of turned-on artificial light sources is crucial for a better representation in energy simulations. In such a case, the proposed global approach would completely fail, whereas the local approach would be in need of being applied separately for every switch.

For curtain openness ratio prediction, one of the possible problems is the reflectivity of the material of the shading device. The curtain in our classroom environment is dark blue, thereby absorbing a high percentage of the light (Fig. 3b). In contrast, if the color of the curtain is very light, the artificial light reflected from the curtain turn out to be very bright too. However, the brightness difference between the daylight and curtain is one of the most crucial requirements for the proposed algorithm. Therefore, implementing the proposed method in such a case may pose a disadvantage to the algorithm and reduce the performance of the estimation. Another limitation is that the method cannot make estimations during the evening and night due to the absence of daylight. However, for energy simulations, the main goal of curtain openness ratio prediction is to measure the amount of daylight or radiative heat entering to the indoor environment. Therefore, this does not pose a serious problem for energy simulations.

The calculated results can be used in energy simulation tools. Our previous work, in which we implement object detection methods to count the number of occupants in a room using IP cameras, demonstrates the feasibility of real-time data capturing for energy simulations [5]. In this previous work, we have demonstrated that (1) the actual number of occupants can be realistically estimated using video-analytical methods,

and (2) the estimated number of people, when added as 'People' schedules, can improve the precision of results.

The methods used in this study too have similar potential in several areas. First, similar to our previous work, the captured data can be converted into hourly or sub-hourly schedules for lighting and shading. However, in our work, one-week recordings were analysed. Executing annual energy simulations would require annual video recordings and the conversion of this data into annual schedules. Alternatively, the captured weekly data could be converted into weekly schedules, to be repeated during the year. The use of the lighting and shading data in energy simulation models remains as a future work.

6 CONCLUSION
In this study, local and global illumination change methods are proposed for the light state detection problem. For the verification of the proposed light state detection methods, two one-week video-recordings which are from a winter season and a spring season are utilized. For curtain openness ratio detection, a three-stage algorithm is proposed and the algorithm performance is tested on manually-labeled shading masks. For both tasks, we consider that promising results are obtained. Vision-based methods for light state detection and curtain openness ratio prediction has been an hitherto unaddressed area in the literature. As future work, we consider that the performance of the light state detection can be improved further by benefiting from existing pose estimation methods, e.g. OpenPose [3] in order to eliminate false detection. Moreover, the use of the calculated data in energy simulation tools remains a future work.

REFERENCES
1. Anand, P., Cheong, D., Sekhar, C., Santamouris, M., and Kondepudi, S. Energy saving estimation for plug and lighting load using occupancy analysis. *Renewable Energy 143* (dec 2019), 1143–1161.

2. ANSI/ASHRAE/IES. *Standard 90.1-2013, Energy Standard for Buildings Except Low-Rise Residential Buildings*. 2013.

3. Cao, Z., Hidalgo Martinez, G., Simon, T., Wei, S.-E., and Sheikh, Y. A. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (7 2019), 1–1.

4. Delzendeh, E., Wu, S., Lee, A., and Zhou, Y. The impact of occupants' behaviours on building energy analysis: A research review, dec 2017.

5. Dino, I., Kalfaoglu, E., Sari, A., Akin, S., Iseri, O., Alatan, A., Kalkan, S., and Erdogan, B. Video content

6. D'Oca, S., Gunay, H. B., Gilani, S., and O'Brien, W. Critical review and illustrative examples of office occupant modelling formalisms. *Building Services Engineering Research and Technology 40*, 6 (nov 2019), 732–757.

7. D'Oca, S., Hong, T., and Langevin, J. The human dimensions of energy use in buildings: A review, jan 2018.

8. EIA. 2012 commercial buildings energy consumption survey: Energy usage summary. **https://www.eia.gov/consumption/commercial/ reports/2012/energyusage/index.php**, 2012. Online; accessed 15 March 2020.

9. Gunay, H. B., O'Brien, W., Beausoleil-Morrison, I., D'Oca, S., and Corgnati, S. P. On modelling and simulation of occupant models (2015).

10. Haldi, F., and Robinson, D. On the behaviour and adaptation of office occupants. *Building and Environment 43*, 12 (dec 2008), 2163–2177.

11. Haldi, F., and Robinson, D. Interactions with window openings by office occupants. *Building and Environment 44*, 12 (dec 2009), 2378–2395.

12. Haldi, F., and Robinson, D. Adaptive actions on shading devices in response to local visual stimuli. *Journal of Building Performance Simulation 3*, 2 (jun 2010), 135–153.

13. Herkel, S., Knapp, U., and Pfafferott, J. Towards a model of user behaviour regarding the manual control of windows in office buildings. *Building and Environment 43*, 4 (apr 2008), 588–600.

14. Kirimtat, A., Koyunbaba, B. K., Chatzikonstantinou, I., and Sariyildiz, S. Review of simulation modeling for shading devices in buildings, 2016.

15. Lowe, D. G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision 60*, 2 (Nov 2004), 91–110.

16. Mikolajczyk, K., and Schmid, C. Scale & affine invariant interest point detectors. *Int. J. Comput. Vision 60*, 1 (Oct. 2004), 63–86.

17. Rijal, H. B., Tuohy, P., Humphreys, M. A., Nicol, J. F., and Samuel, A. An algorithm to represent occupant use of windows and fans including situation-specific motivations and constraints. *Building Simulation 4*, 2 (jun 2011), 117–134.

18. Zivkovic, Z., and van der Heijden, F. Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern recognition letters 27*, 1636734 (P (1 2006), 773–780. 10.1016/j.patrec.2005.11.005.

Urban Intervention Strategy for Networking Zografos Public Spaces

Eirini Androutsopoulou [1]

[1]National Technical University of Athens
Athens, Greece
iandroutsopoulou@gmail.com

ABSTRACT

The purpose of the research is to develop urban environment networks in order to formulate strategic urban interventions for the networking of public spaces.

In the case of Zografos Municipality (Attica, Greece), the research aims to establish a coherent network of public spaces, with the prioritization of urban interventions at nodes (public spaces) and their connections, as the existing public spaces do not constitute a network. The methodology starts with the configuration of urban networks by defining nodes and connections and proceeds to export data per network node. Applying Hierarchical Clustering as a methodology for grouping and analyzing network and other features (ownership, accessibility, total area) results in the formulation of strategic urban interventions, with the selection of nodes that meet multiple criteria.

Author Keywords
Network; Urban; Hierarchical Clustering; Multiple Criteria

ACM Classification Keywords
I.6.1 SIMULATION AND MODELING

1   INTRODUCTION
The methodology for prioritizing urban interventions in Zografos nodes (public spaces) and their interconnections was developed within the framework of the research program entitled "Intervention Strategies for Networking and Upgrading of Zografos Public Spaces" [1].

The description of the urban environment as a network configuration presupposes the definition of the network nodes and connections [2], [3]. In order for the network connections to be defined, the proximity relations of the actual geographic space are translated to network connections. With the consideration of the urban environment as a network configuration two groups of nodes' characteristics arise, those that refer to the network structure and those that describe urban characteristics, such as physical structure characteristics, ownership status and plantation cover.

2   EXISTING CONFIGURATION OF URBAN OUTDOOR SPACES OF ZOGRAFOS
In its current state, the urban outdoor spaces configuration of the Municipality of Zografos includes 24 subnets (one subnet with 11 nodes, 4 subnets with 4 nodes, 5 subnets with 3 nodes and 14 subnets with 2 nodes) and 46 isolated nodes. It includes in total 116 nodes (based on the research team's records) and 52 existing links between them. Urban outdoor spaces are interconnected when there is direct proximity to each other in the real geographical space (Figure 1).

The network dynamics (the dynamic network topology) are quantified by applying specific network algorithms (degree, betweenness centrality, eigenvector centrality, closeness centrality and eccentricity), which translate the network dynamics into variable values per node. The values of variables per node resulting from the application of the network algorithms derive solely from the network structure. The implementation of the network algorithms shows a great network component, in number of connected nodes, comprising 11 nodes (i22, i3, i4, i2, g4a, g4b, g4c, g4d, g4e, g4f, g5) and 12 interconnections (Table 1, Figures 2,3). Lacking an internal coherence, the existing configuration of urban outdoor spaces of the Municipality of Zografos does not constitute a network.

3   NETWORK OF URBAN OUTDOOR SPACES OF ZOGRAFOS
In order to select the nodes, and therefore the connections, which are more important for the gradual establishment of a coherent network of urban outdoor spaces, we start with a coherent hypothetical network where all the nodes are interconnected. This constitutes the desired final outcome of the urban interventions strategy. At the new hypothetical network, all nodes are connected using short paths, that is to say up to four building blocks of the existing urban fabric, provided that nodes - free spaces are directly connected by road axes.   This sets up a cohesive network of urban outdoor spaces with 116 nodes and 156 connections (Figure 4). At a next step, the network algorithms are implemented. For the purpose of this research we focus on the network feature of degree, betweenness centrality, eigenvector centrality, closeness centrality and eccentricity (Table 2, Figures 5, 6) [4].
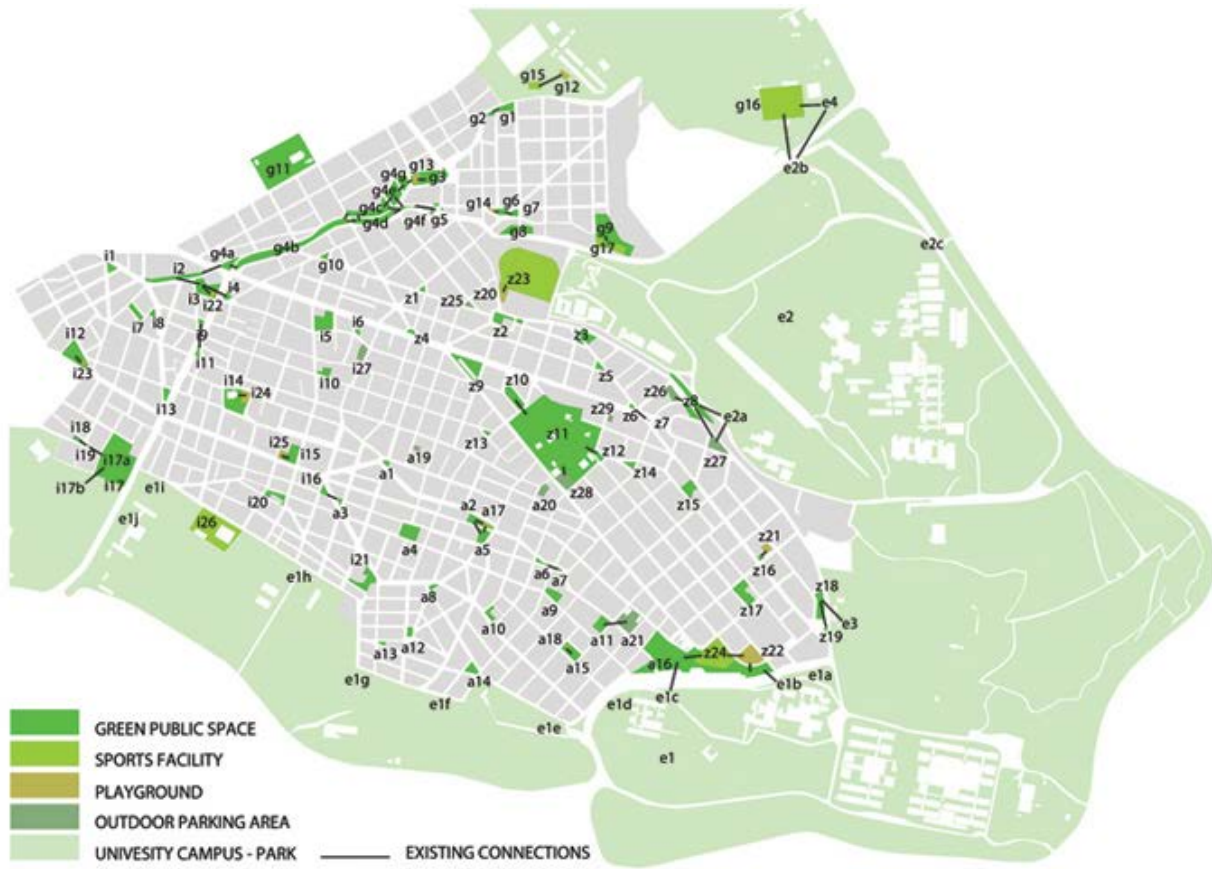
Figure 1. The existing configuration of urban outdoor spaces of the Municipality of Zografos.

Id	Degree	Betweenness Centrality	Eigenvector Centrality	Closeness Centrality	Eccentricity
G1	1	0	0.021835	1	1
G3	1	0	0.223319	1	1

Table 1. Table extract of network characteristics per node.
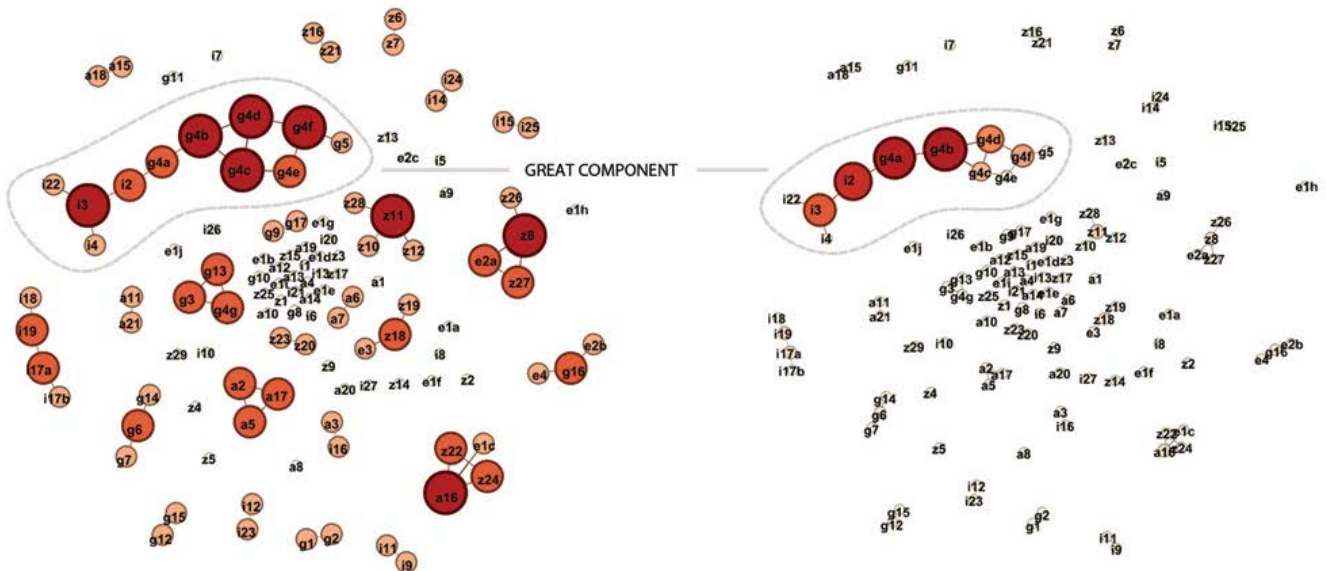
Figures 2,3 The existing configuration of urban outdoor spaces of the Municipality of Zografos. The nodes` size and color correspond to degree (left) and betweenness centrality measure (right).

Figure 4. The coherent hypothetical network of urban outdoor spaces of the Municipality of Zografos.

Id	Degree	Betweenness Centrality	Eigenvector Centrality	Closeness Centrality	Eccentricity
G1	3	1074	0.101668	10.31304	17
G3	3	226	0.088618	12.2087	18

Table 2. Table extract of network characteristics per node.
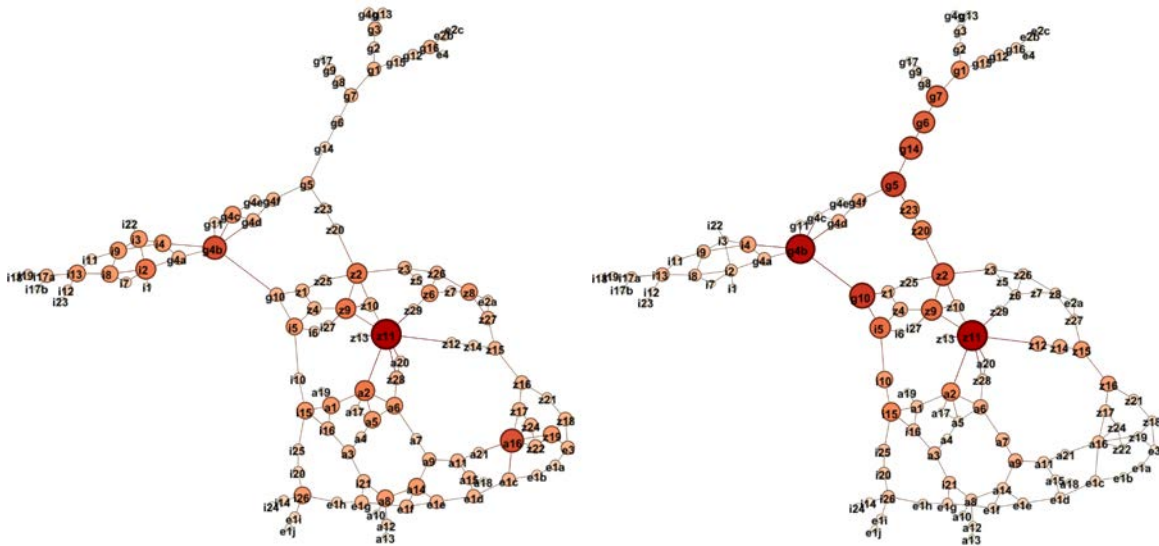
Figures 5,6 The coherent hypothetical network of urban outdoor spaces of the Municipality of Zografos. The nodes` size and color correspond to degree (left) and betweenness centrality measure (right).

79

4 HIERARCHICAL CLUSTERING

To this point, network algorithms have been implemented and variables referring to the network structured are assigned per node. In addition, characteristics referring to physical structure features (material, cover, surface materials, vegetation), ownership, status, and accessibility are assigned to each node. Each network node therefore has multiple network and other features. In order to be able to analyze and manage data in the direction of building strategic urban interventions, a methodology for finding associations between nodes and their characteristics is applied (Hierarchical Clustering) [5].

This allows for the clustering of nodes and their attributes, with the application of similarity measures (Pearson's correlation coefficient), of distance metrics (average group linkage) and the definition of the minimum similarity index [6].

4.1 Node similarity results based on network characteristics

The application of Hierarchical Clustering on datasets of nodes with their network characteristics results in groups of more similar nodes being detected solely based on the network structure.

The node clusters that emerge with the application of Hierarchical Clustering have a similar role in the Zografos Municipality's free space network in terms of network function, measured as network characteristics (degree, eigenvector centrality, betweenness centrality, closeness centrality, eccentricity). These include nodes that appear to be scattered but have similar relative positions in the network (Figure 7).

Group 06 includes nodes (z9, z11, z2, a2, g10, i5) with high degree, high betweenness centrality and eigenvector centrality and low closeness centrality and eccentricity values (mean and maximum network distances). Group 08 includes nodes (z20, z23, g5, g6, g14, g7) with high betweenness centrality and eigenvector centrality and low closeness centrality and eccentricity values, while group 07 includes nodes (g4d, i4, g4a, g4b, g4f) with average to high values of degree, betweenness centrality and eigenvector centrality and low closeness centrality and eccentricity values. Groups 02 and 03 include nodes with average degree and betweenness centrality values, low closeness centrality values, and average eccentricity values. Groups 01, 04 and 05 include nodes with low degree and low betweenness centrality values, and high values of closeness centrality and eccentricity (long mean and maximum network distances).

4.2 Node similarity results based on urban characteristics

By shifting the similarity index, different clusters of urban outdoor spaces are produced, producing groups of closer or most distant correlations, namely clusters of nodes that have more or less similar urban identity, based on other characteristics. In the urban outdoor spaces network of the Municipality of Zografos the characteristics that don't refer to network structure are quantitative and qualitative. Coverage is a quantitative characteristic, while other physical structure features (materialized, surface materials, vegetation cover), ownership status, status and accessibility are qualitative features. In order to generate clusters of free spaces using Hierarchical Clustering, the following values are applied:

- constructed urban outdoor space: yes (1), no (0)

- cover area: numeric value

- surface materials: hard surfaces (0), hard surfaces and soil (5), soil (10)

- vegetation cover: little or no cover (0), low plantation (5), high plantation (10)

- ownership status: special status (0), public (10)

- current condition: poor (0), average (5), good (10)

- accessibility: no access (0), controlled entrance (5), free access (10)

By setting the minimum similarity index to 0.95, 11 clusters of urban outdoor spaces and 3 isolated spaces are produced (Figure 8). From the generated clusters, groups 03, 07 and 08 include constructed free access public spaces which are in good condition. Of these, groups 03 and 07 include urban spaces with high plantation, while group 08 includes urban spaces with low plantation. Isolated node z11 is also a public space with free access and high plantation that does not belong to the above groups due to its large cover area. Cluster 02 also includes free access and constructed public spaces with high plantation, with the exception of two public spaces (z9 and a10) that do not have sufficient vegetation cover. Group 11 includes privately owned parking areas with controlled entrance, while groups 01, 04, 06, 09 and 10 include public spaces in moderate or poor current condition. Isolated public spaces z16 and i10 are also in moderate and poor current condition.
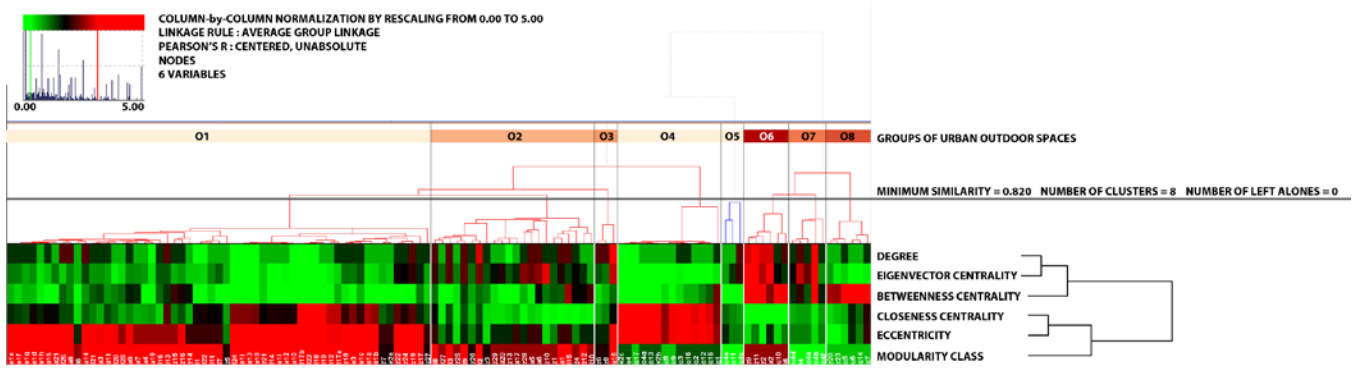
Figure 7. Groups of urban outdoor spaces based on network characteristics.
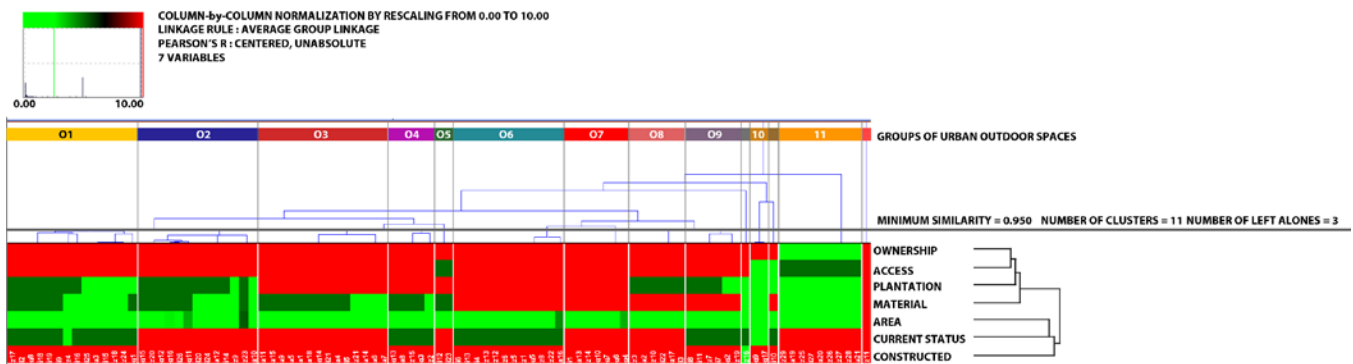
Figure 8. Groups of urban outdoor spaces based on urban characteristics.

5 URBAN INTERVENTION STRATEGY FOR NETWORKING ZOGRAFOS PUBLIC SPACES

5.1 Selection of groups of nodes important for their role in the network

In order to set up a network of public spaces, priority is given to nodes that are included in stronger groups in terms of their ability to link subnets and isolated nodes. Therefore, the nodes that are most important for their overall role in the network are selected at first.

The choice of these specific network features is related to the research question. As the purpose of urban interventions is the ultimate establishment of a cohesive network, nodes that have greater ability to connect subnets and isolated nodes are considered more important (Figure 9).

By grouping the nodes according to the values of variables resulting from the application of network algorithms, the groups that include the most important nodes in the network are selected based on the following criteria:

- high degree value (nodes that are connected to more neighboring nodes)

- high value of betweenness centrality (nodes that are more important in their ability to group nodes together)

- high value of eigenvector centrality (nodes that are important because the adjacent nodes have a high degree value)

- low closeness centrality (nodes having short average network distances from all other nodes in the network)

- low eccentricity value (nodes with a short maximum distance from the farthest node).
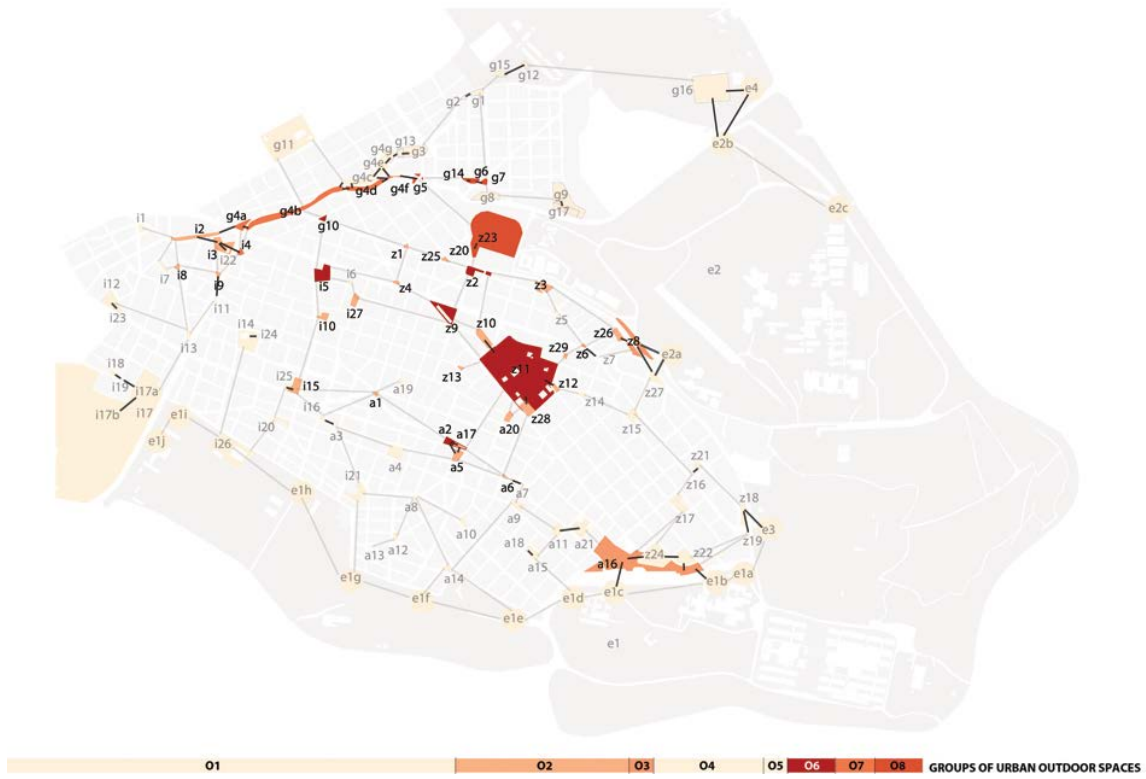
Figure 9. Selection of nodes belonging to groups 02, 03, 06, 07, 08, based on network characteristics.
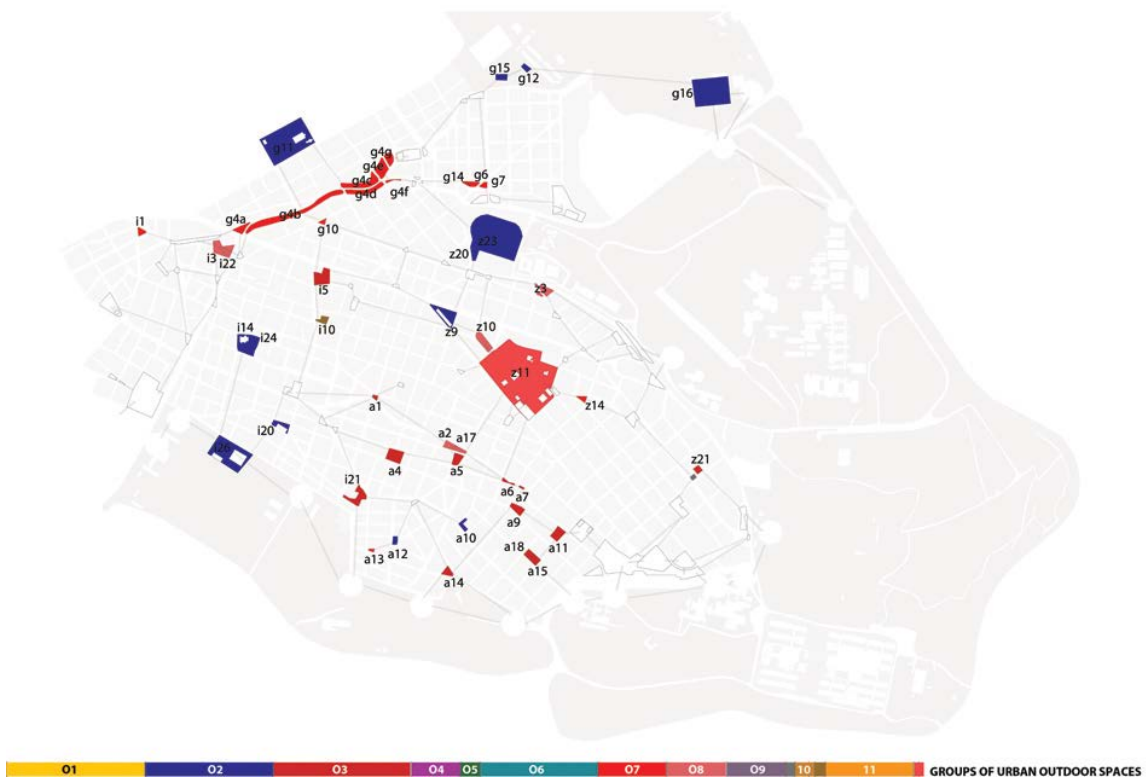
Figure 10. Selection of nodes belonging to groups 02, 03, 07, 08 and left-alone node z11, based on urban characteristics.

Figure 11. Selection of public spaces that meet multiple criteria.

5.2 Selection of groups of nodes that meet other criteria

Hierarchical clustering results based on other attributes produced some groups of nodes being public, constructed outdoor spaces with free access and sufficient vegetation cover. Since the intervention in these nodes is rather optional, giving priority to implementation in these public places, the financial cost will only concern networking with intervention in the connections between them (Figure 10).

Clustering the nodes according to their attributes yields groups of nodes that are more important based on the following criteria:

- ownership status: public

- accessibility: free access

- vegetation cover: low and mainly high plantation

- constructed urban outdoor space: yes

- current condition: good

5.3 Urban Intervention Strategy by selecting for a first stage implementation of nodes that meet multiple criteria

At this point the nodes that meet multiple criteria are selected. The aim is to set up a network by intervening only in the links between them, as these are public, constructed spaces in good current condition, with free access and sufficient vegetation cover. In addition, these are public spaces that are important for their role in the network and therefore the intervention at the connections between them will quickly give a network configuration to the public spaces in the Municipality of Zografos.

6   CONCLUSION

With the combined management of data related to network as well as urban characteristics, it is possible to formulate an strategy that prioritizes the urban interventions. The application of network algorithms allows us to extract information about the importance of some public spaces for their relative position in the network. By applying the Hierarchical Clustering to the network and urban features, we are able to select connections between selected nodes that meet multiple criteria. In a next step, nodes - public spaces that are directly linked to the original nodes are selected for implementation. In this way, with the least

possible use of financial resources, an initial network of public spaces for the Municipality of Zografos is created.

REFERENCES

1. Municipality of Zografos, N.T.U.A.

2. Newman, M.E.G., 2010. *Networks: An Introduction*, Oxford University Press, Oxford, New York.

3. Androutsopoulou, E.: 2014. 'Urban body network configurations - Attica', *Conference proceedings eCAADe 2014, (1) 191-199.*

4. Blondel,VD, Guillaume, JL, Lambiotte, R, Lefebvre, E.: 2008. 'Fast unfolding of communities in large networks', *Journal of Statistical Mechanics: Theory and Experiment (10).*

5. Androutsopoulou, E.: 2016. 'Urban body network configurations through attributes of network elements', 2016 Proceedings *of the Symposium on Simulation for Architecture and Urban Design, 111-117.*

6. Hastie, T., Tibshirani, R., Friedman, J.: 2003. *The elements of Statistical Learning*, Springer-Verlag, New York.

Predicting Space Usage by Multi-Objective Assessment of Outdoor Thermal Comfort around a University Campus

Patrick Kastner[1], Timur Dogan[1]

[1]Environmental Systems Lab, Cornell University, Ithaca, NY, USA, pk373@cornell.edu

ABSTRACT

With the impending issues regarding global warming, urban design is considered a key driver to improve the microclimate in cities. For public spaces, studies suggest that outdoor thermal comfort may be seen as a proxy for space usage in turn attractiveness to people. Although the topic has gained interest in recent years, the discussion so far has focused on computing the metrics rather than deriving interventions from them. Here, we use the tool Eddy3D to model and analyze the outdoor thermal comfort of a designated area around a university campus. Further, we demonstrate how to estimate space usage from those results. Finally, we conduct a spatial sensitivity analysis of the underlying results as a step towards decision aiding. Our work demonstrates how decision-makers may derive areas where interventions will likely have the largest impact on outdoor thermal comfort performance.

Author Keywords

Outdoor Thermal Comfort, Sensitivity Analysis, CFD, MRT, Space Usage

1 INTRODUCTION

Climate change scenarios suggest disconcerting outcomes for the future of cities, especially in urbanized and dense locations. As a result of changing urban microclimates, there is not only increased competition over the wind, sun, and daylight [6], but also a growing risk for detrimental impacts on aspects of the city life; public health, commercial activities, active mobility, and space usage among them [17]. However, the fact that urban microclimates are shaped by the way cities are built also opens new opportunities. For these reasons, the prediction of outdoor thermal comfort has become an increasingly relevant area of research.

In the early 1980s, Whyte [19] devoted a significant amount of time to studying how people use public spaces in US cities, and why some plazas show different usage patterns than others. From video recordings, they counted the number of dwellers which revealed that access to sunlight can be a major factor for plaza use. For NYC, they found that there is a strong correlation between space usage and exposure to sunlight while in other months the microclimate was less relevant. Direct exposure to sunlight is preferred through May during which

time the ambient temperature has yet reached comfortable levels. By contrast, in months from June through August, people show little preference between sun and shade and other social parameters and amenities were determining factors [18].

This study frames simulated-based microclimate assessment as a multiple-criteria evaluation problem. It aims to assess the microclimate conditions on a university campus from which outdoor space use, a proxy for the attractiveness of outdoor spaces, is derived. This is motivated by the fact that the majority of the student population is on campus during spring and fall when outdoor thermal comfort is highly dependent on sun and wind exposure. The framework of this analysis is justified with a two-at-a-time and a global sensitivity analysis. The spatial sensitivity adds a layer to the results that allow decision-makers to deduce where geometric interventions will likely have the largest impact on the performance.

2 METHODOLOGY

This study utilizes a simulation tool called Eddy3D to simulate the annual outdoor thermal comfort around an existing building arrangement on the Cornell University campus. The weather data for this study were retrieved from the EnergyPlus repository by the Department of Energy[1] and classifies as warm-summer humid continental climate (Dfb) according to the Köppen classification.

2.1 Case study: Engineering Quad at Cornell University

The CAD model used is comprised of Cornell's Engineering Quad in Ithaca, upstate NY, see Figure 1. As the elevation differences between the buildings are insignificant, they have been neglected in the CAD model.

2.2 Universal Thermal Climate Index (UTCI)

Like other outdoor comfort metrics, the UTCI was developed as an equivalent temperature measure. It is based on a multi-node thermo-physical model that was coupled with an adaptive clothing model [8]. Thus, for any combination of air temperature, wind, radiation, and humidity, the UTCI is defined as the air temperature of a particular reference condition which would cause the same thermal sensation as predicted by the model [5]. We chose the UTCI as a metric, as several studies have shown the UTCI to be a suitable metric that is sensitive and accurate for cold temperatures while also achieving good

---

[1] Syracuse-Hancock.Intl.AP.725190_TMY3
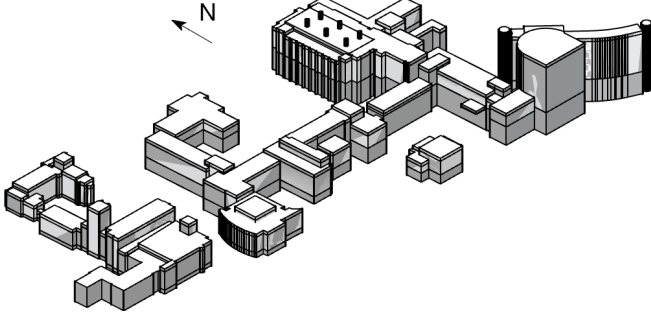
Figure 1. CAD model of the Cornell University engineering quad.

agreement with human responses in tropical climates [14]. Mathematically, the UTCI is computed with a polynomial approximation of four input variables:

$$UTCI = T_{amb} + f(T_{amb}, T_{MRT}, U_{Wind}, p_{vapour}) \quad (1)$$

where T_{amb} is the ambient temperature, T_{MRT} is the mean radiant temperature, U_{Wind} is the wind velocity, and p_{vapour} refers to the vapor pressure, which is a function of the dry-bulb temperature. The resulting UTCI temperatures may be classified and reported as thermal stress categories ranging from "extreme cold stress" to "extreme heat stress". Ideally, one would strive for maximizing the annual urban climate for the "No thermal stress" condition, which falls in the range of 9°C and 26°C.

Empirical data that has been collected by Reinhart et al. [15], confirmed studies undertaken by White. Over one year, the number of people with WiFi devices in a public courtyard on a university campus has been recorded. It has been shown that people take longer lunch breaks when in "no thermal stress" conditions. This suggests that people prefer being in outdoor spaces if the microclimatic conditions are within the "no thermal stress" category. Their findings are summarized in Figure 2.
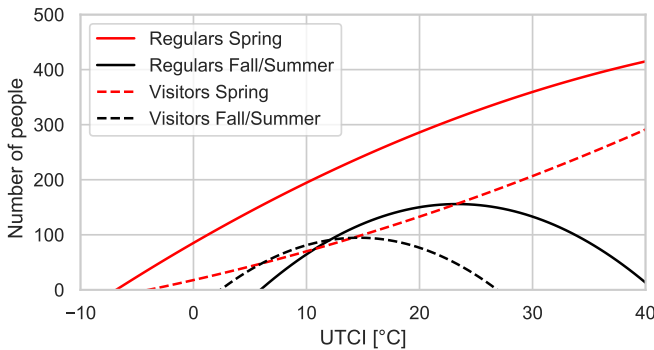
Figure 2. Collocation of empirical outdoor thermal comfort and space usage correlations collected by Reinhart et al.

2.3 Sensitivity analysis

To identify the dominant input parameters of the UTCI metric in this particular temperate climate, and to deduce actionable suggestions on how the microclimate could be improved,

sensitivity analyses were carried out in three stages. First, a two-at-a-time sensitivity analysis was carried out by varying the ambient temperature with the remaining three parameters, see Figure 3. For this, the hidden variables were fixed to 20 °C for the dry-bulb temperature, 1 $^m/_s$ for the wind velocity and 50 % for the relative humidity. The code was adapted from [13]. Further, we ran a global, variance-based Sobol sensitivity analysis. It is a method that decomposes the variance in the UTCI into fractions which can be attributed to the four input variables. Generally, it is used to derive the rank-order importance of input variables in multivariate analyses. Beyond that, it is also helpful in not only gaining an understanding of how the variables affect the solution in sets of two but allow us to quantify the total effect when varying all parameters simultaneously (which cannot be plotted in two dimensions). The Saltelli implementation in SALib [9] was used to evaluate 100'000 samples of the UTCI problem space with the upper and lower bounds from Table 1. The upper and lower bounds were derived from the Syracuse weather data, except for the MRT, which has been assumed to be within a ±5 $C°$ offset from the dry-bulb temperature.

Table 1. Variables and bounds for the global sensitivity analysis from the Syracuse weather data.

Variable	Unit	Min	Max
AmbTemp	[°C]	-19.4	37.8
MRT	[°C]	-24.4	42.8
Wind	[$^m/_s$]	0	17
RelHum	[%]	0	100

The aforementioned methods provide a good starting point to understand the dominance of the input variables in a theoretical framework. However, as cities shape their microclimates, they also shape the upper and lower bounds of the UTCI input variables in time and space. In this regard, the wind velocity and the MRT can be especially spatially-inhomogeneous due to wind sheltering effects or extreme surface temperatures. Therefore, we conducted a spatially-resolved global sensitivity analysis, for which the annual upper and lower bounds for each probing point serve as an input. For this, we probed the wind velocity and the MRT from the outdoor thermal comfort simulation results and made use of the weather data for ambient temperature and relative humidity.

2.4 Annual outdoor thermal comfort simulation framework

The simulation framework consists of two simulation engines, namely: OpenFOAM and Radiance. Both engines are centered around a toolkit called Eddy3D[2] that is implemented in Rhinoceros and Grasshopper which handles pre-, post-processing and the data handling between the simulation engines. Specifically, Eddy3D creates the simulation domain, specifies the boundary conditions, and takes care of processing the weather data based on the building geometry in Rhinoceros. Similarly, it uses Rhinoceros' meshing capabilities to export

---

[2]www.eddy3d.com

building meshes for both OpenFOAM and Radiance. Further, it uses Radiance to calculate irradiation and view factors for each sensor point.

Wind velocity
Due to the computationally-expensive nature of Computational Fluid Dynamics (CFD) simulations, it is infeasible to run a single analysis for an entire year. Considering this, we made use of the wind reduction factors method which has been implemented in Eddy3D. The tool utilizes Open-FOAM's $blockMesh$ utility for the background mesh and $snappyHexMesh$ to subsequently snap the background mesh to the building geometry. For the background mesh, we used a cylindrical simulation domain approach which allows reusing the same computational mesh for every wind direction, thus reducing the computation time and storage space [11]. Within the cylindrical mesh, we further refined the mesh within a refinement box that surrounds the buildings of interest. The simulation domain was set up according to best practices while taking into account all relevant surrounding buildings which resulted in $6.6 \cdot 10^6$ cells for the global mesh. This methodology makes use of a set of CFD simulations from several wind directions. For this study, we used 8 RANS simulations (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) in a 45° interval. Depending on the direction, we mapped the inlets on the one-half circle of the simulation domain and the outlet on the opposite side. We used an incompressible, isothermal, steady-state solver from OpenFOAM in combination with a $k - \omega - SST$ turbulence model. The half-circular domain inlet was set to an atmospheric boundary layer (ABL) profile for U, k, and ω, and a roughness length $z_0 = 1$ that corresponds to a suburban environment. At the outlet of the computational domain, constant pressure is assumed, while the other variables are imposed to be zero-gradient. The ground and the building geometry used the same boundary conditions, a no-slip condition for velocity, a zero-gradient condition for the pressure and wall functions for U, k, and ω. Going forward, the 8 RANS simulations served as a nearest neighbor lookup table of wind velocities in concert with the annual weather data. For each probing point, we probed the simulated velocity from the 8 CFD simulations. This multidimensional array is used to calculate the dimensionless wind velocity for every probing point by dividing the simulated velocity magnitude by the scaled-down inlet velocity with the logarithmic wind power profile. This yields a spatial wind reduction matrix with information for every probing point for each of the 8 wind directions. From here, we converted the spatial matrix into a temporal matrix. For every hour of the year and its corresponding wind direction, we looked up the nearest neighbor CFD simulation and multiplied the velocities from the spatial velocity matrix with the wind velocity from the weather data that has been scaled down to the probing height. This operation yields a temporal velocity matrix with wind reduction data from which the wind velocities for the UTCI calculation are retrieved [10]. For cases where the wind velocity was outside the bounds of the UTCI calculation (0.5 $^m/_s$ < applicable range < 17 $^m/_s$), we replaced the values with lower and upper bounds and lifted the velocity at a height of 10 m as advised in [4].

Mean radiant temperature
The mean radiant temperature (MRT) calculated in this study consists of three components: the sky temperature T_{sky}, the solar gain (ΔMRT_{ds}) of being exposed to direct sun, and the ground and building surface temperature T_{gb}.

To estimate the ratio between T_{sky} and T_{gb}, we used Radiance to run a view factor analysis for every probing point. We calculated T_{sky} through the sky emissivity and the horizontal infrared radiation intensity [7]. We approximated T_{gb} with the ambient temperature as it has been shown that typical differences are less than ±5 K [12]. The solar gain to the human body is calculated using the Effective Radiant Field (ERF) [1] which we adapted for an outdoor setting and from which we derived (ΔMRT_{ds}). For the irradiance that is used as an input for the ERF, we implemented a Radiance-based Two-Phase (DDS) method. The DDS approach was chosen because it provides a better spatial resolution of the direct component. The Two-Phase DDS method is a daylight-coefficient-based simulation with an all-weather dynamic sky model (Perez Sky model). Instead of approximating the position and shape of the sun with few sky patches, we used 577 sun patches for the direct and diffuse simulation and 2305 direct sun patches. The illuminance is a linear combination of: (1) an annual daylight coefficient simulation, (2) annual direct-only daylight coefficients, and (3) an annual sun-coefficients simulation [3]:

$$E = C_{dc} \cdot S - C_{dcd} \cdot S_d + Csun \cdot S_{sun} \qquad (2)$$

where C_{dc} and S denote the daylight coefficient matrix and the sky vector, C_{dcd} and S_d denote the direct-sky coefficient and the direct sky matrix, and C_{sun} and S_{sun} denote the direct-sun coefficient and the sun matrix, respectively. [16]. Finally, we computed the MRT for every hour h and every probe i according to Equation (3).

$$MRT_{h,i} = T_{sky,h} \cdot f_{sky,i} + T_{gb,h} \cdot f_{sky,i} + \Delta MRT_{ds,h} \quad (3)$$

2.5 Space usage
We used the data collected by Reinhart et al. [15] to estimate space usage around the Cornell University Engineering Quad to provide a meaningful metric for campus decision-makers. For this, we calculated the number of people for each hour and probing point as a function of the UTCI. The space usage is calculated for the three seasons spring, summer, and fall for occupancy periods between 10 am - 7 pm. All values were then normalized between 0 - 1 as the empirical data cannot be directly applied to a different university campus without further investigation.

3 RESULTS
Both non-spatial sensitivity analyses justify the assumptions that have been made when setting up the simulation framework.

Two-at-a-time sensitivity analysis
In Figure 3, each input variable to the UTCI metric shows a non-linear behavior when varying it with the ambient temperature, however, their amplitude and parameter range differ
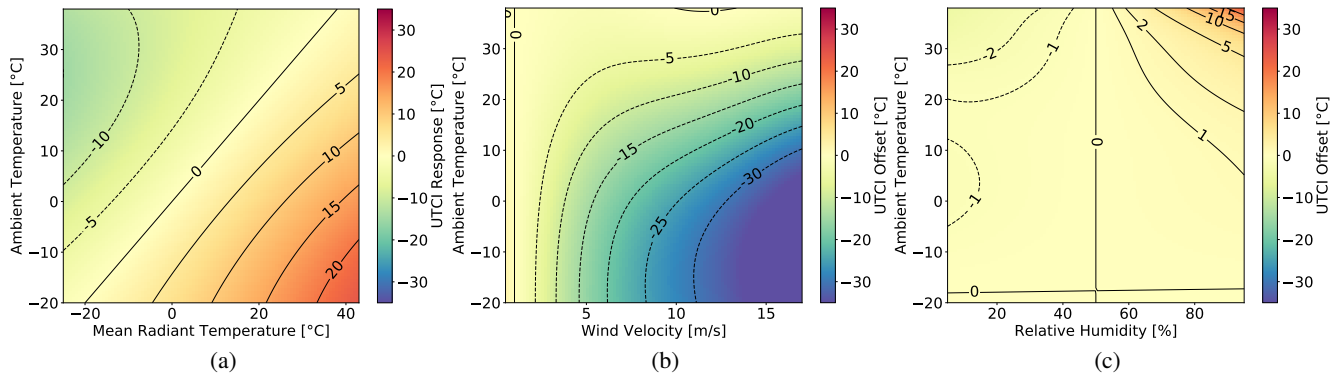
Figure 3. Two-at-a-time sensitivity analysis of the UTCI input variables.

significantly. When varying the MRT with the ambient temperature in Figure 3 (a), high and low MRTs create significant discomfort through temperature asymmetry. For the wind velocity in Figure 3 (b), we see highly non-linear behavior for both low ambient temperatures and high wind velocities. This effect is commonly known as the wind chill due to increased forced convection. When varying the ambient temperature with relative humidity, non-linear behavior is evident only for high levels of both relative humidity and ambient temperature.

Global sensitivity analysis

Figure 4 quantifies the first-, second-, and total-order sensitivity indices illustrated as black dots, gray lines, and black circles, respectively. The confidence intervals for each index and variable were found to be below $1 \cdot 10^{-2}$. We confirm that the dominant variables contributing to the UTCI in a climate in upstate NY were ambient temperature, MRT, and wind velocity. Those three variables show total-order sensitivity indices of 83 %, 5 %, 16 %, respectively. The least dominant factor in this climate is the relative humidity which only accounts for 1 % of the variance. This suggests that the humidity may be neglected for the Dfb climate zone. Note that the total sensitivity indices sum to a value greater than one, indicating that the input variables are to some extent correlated. The second-order sensitivity indices range from ~1 - 5 % with wind and ambient temperature showing the largest interaction effects. This confirms the findings from the two-at-a-time sensitivity analysis, see Figure 3 (b).

Annual outdoor thermal comfort

All results in this section have been computed and probed at a height of 2 m above ground. Figure 5 (a) shows the averaged annual wind velocity magnitudes. We see that points that are sheltered from the wind experience a lower pedestrian comfort rating and vice versa. It is also evident that venturi effects exist in two areas, channeling the wind from north to south. Figure 5 (b) shows the averaged annual mean radiant temperature. The average MRT ranges from 4.5-6.5°C where higher temperatures are seen closed to building geometries and vice versa. Figure 5 (c) shows the averaged annual UTCI. While points close to building geometries generally see favorable conditions in terms of wind velocities and MRTs, it is evident that the venturi effect sports are rendered less comfortable. Overall, the seating arrangements on the Engineering Quad are within areas that show a rather high annual averaged UTCI.
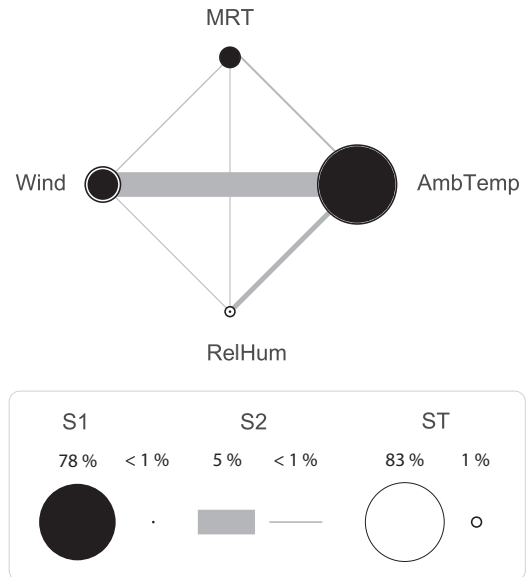
Figure 4. Global sensitivity indices where S1 is the first-order, S2 is the second-order, and ST is the total-order effect for the UTCI estimated with the Sobol method.

Space usage

In Figure 5 (d) we show the predicted space usage around the campus on June 30, 16:00 h, a hot and calm day from the weather file for which the average UTCI was estimated to 33°C. We focus on hourly data in this part of the results as an annual average would not allow to highlight the temporal variance and its effect on space usage. Areas with exposure to direct sun receive a penalty in terms of space usage due to wind not being present. By contrast, shaded areas experience increased numbers of space usage as those are the only ones with a comfortable UTCI. Here, the seating arrangements on the Engineering Quad suffer from direct exposure to sunlight. Given the already uncomfortable conditions (UTCI = 33°C), the seating arrangements and other areas exposed to direct sunlight suffer a penalty.

Spatially-resolved global sensitivity analysis

Figure 6 (a) and Figure 6 (b) show the spatially-resolved total sensitivity indices for the wind velocity and the MRT. The upper bound for ST for the wind velocity is close to 40 %,
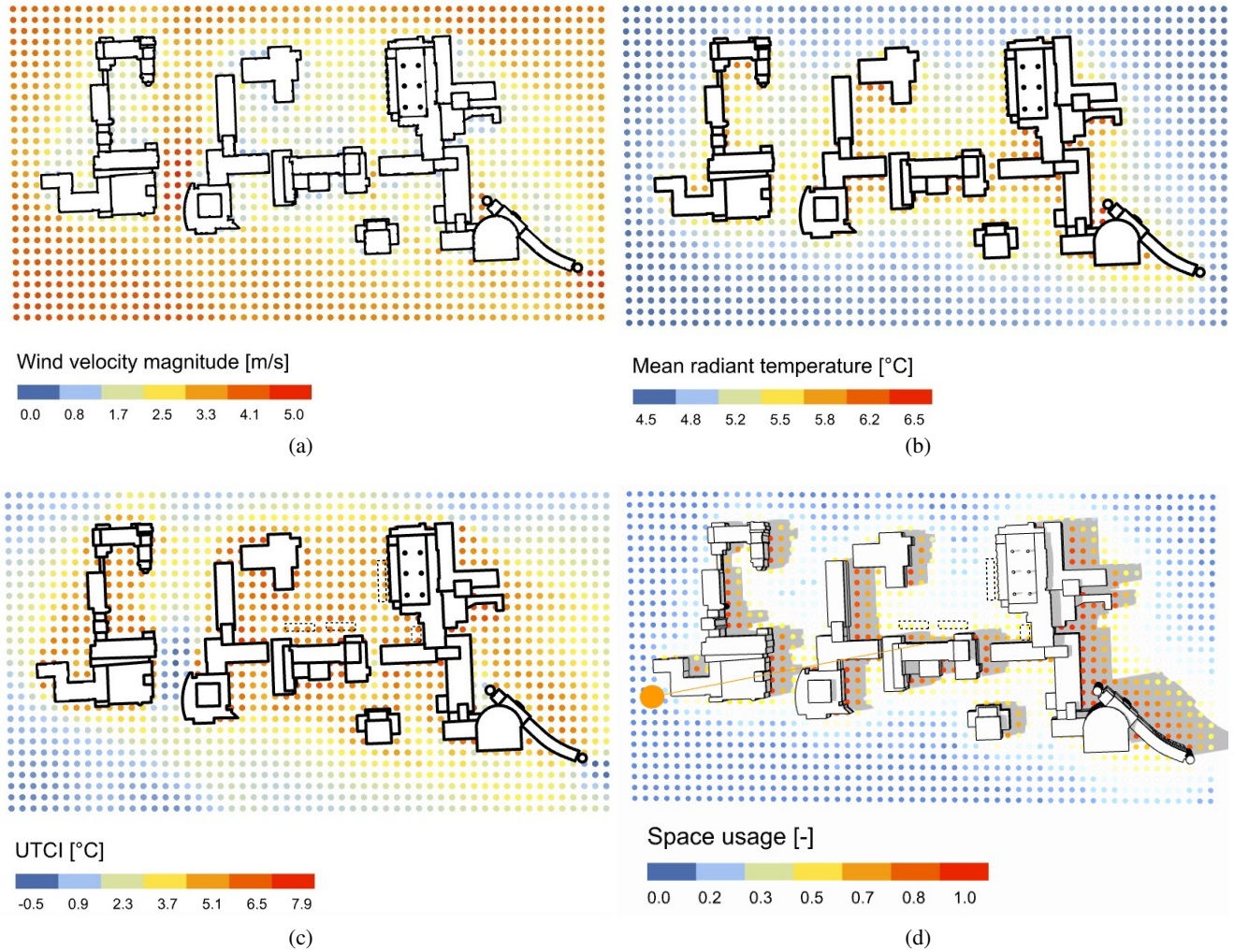
Figure 5. (a) Annual wind velocity magnitudes; (b) average annual mean radiant temperature; (c) average annual UTCI; (d) estimated space usage on June 30, 16:00 h, a hot and calm day from the weather file. The dashed markers indicate actual outdoor seating arrangements on campus.
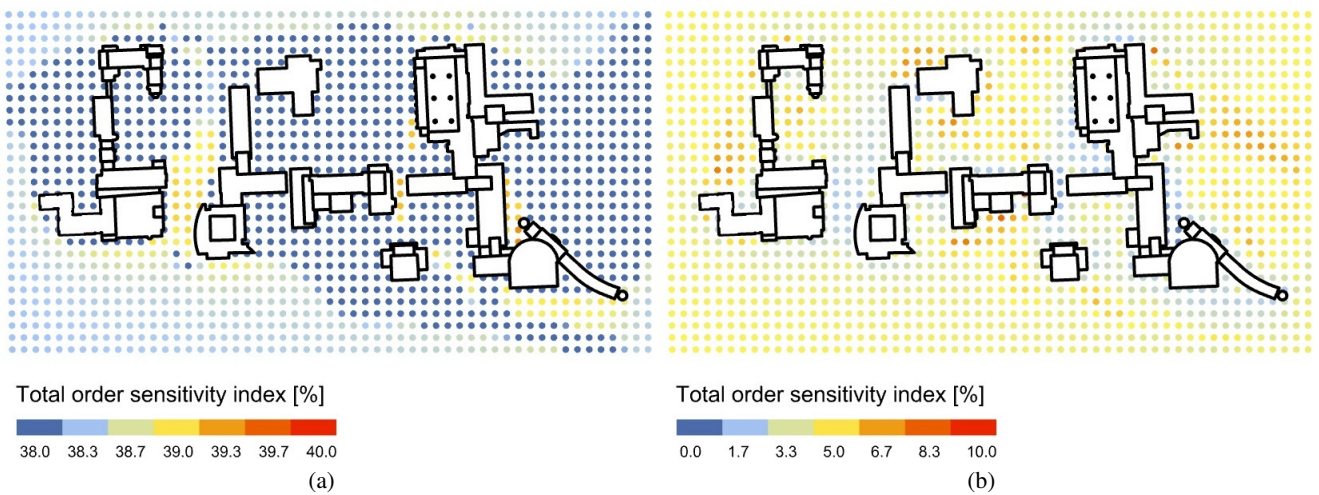
Figure 6. Spatially-resolved total sensitivity index for the (a) the wind velocity and (b) the mean radiant temperature.

whereas the upper bound for the MRT is around 10 %. For example, the wind velocity around two north-south corridors in Figure 6 (a), explains a large part (~40 %) of the variance in the UTCI output considering an entire year.

4 DISCUSSION

We used weather data from Syracuse, NY as boundary conditions for this study as there is limited data available for the site in Ithaca, NY. While ambient temperature and irradiation of both sites are comparable, it is conceivable that Ithaca, NY experiences deviating wind velocity distributions due to its overall differences in elevation and its location at the edge of a valley. Further, using empirical data from another university campus in Boston to estimate space usage has limitations. Despite the overall climate in Boston being similar, it lies in a different sub-climate zone with hotter summers than those in upstate NY. Although Reinhardt et al. [15] suggest their results hold for climate comparable to Boston's, the relationship between UTCI and space usage in upstate NY be likely different. Nonetheless, in this study, we argue for the importance of decision aiding based on outdoor thermal comfort results rather than their validation.

We have shown that local, global, and spatially-global sensitivity analyses can help understand the dominant input variables of a complex metric such as the UTCI. The results of the global sensitivity analysis in Figure 4 reveal the strong second-order interaction between the ambient temperature and the wind velocity. That is to say that, without plotting the relationship between both variables, we know that a large part of the variance in the UTCI can be attributed to varying both the ambient temperature and the wind velocity simultaneously. It is worth noting that the global analysis itself does not reveal the inverse nature of this relationship (an increase in wind velocity and a decrease in dry-bulb temperature yields a non-linear decrease in UTCI). As sensitivity analyses help assess arbitrary functions (such as the UTCI), generalizing this insight might be helpful for practitioners in the design process. In that regard, we suggest to use a global sensitivity analysis as a first step when analyzing a (potentially unknown) metric, and to derive a factor prioritization from it. Identifying second-order effects (gray lines in Figure 4) will help to determine which "slices" of the solution space to plot in two dimensions to better understand the relationship of the variables.

The 16 % total sensitivity of the wind velocity warrants a discussion of the limitation of the annual wind velocity method, see Figure 5 (a). First, the number of wind directions simulated determines the accuracy of the result, although, the increase shows diminishing returns. Although the authors are aware that a higher number of wind directions will increase the overall accuracy of the annual wind velocity magnitudes, we refrained from simulating additional wind direction in the interest of time. Further, for each of the 8 simulations, a constant inlet wind velocity of 5 $^m/_s$ at 10 m was assumed, regardless of the distribution of wind directions and their corresponding hourly wind velocities. This was justified by Becker et al. [2] that have shown that the reattachment length behind a cube does not change significantly for Reynolds numbers greater than $1 \cdot 10^5$ which applies for this study. In general, however, the

wind velocities between the buildings on campus are likely overestimated as not all surrounding buildings on campus have been included in the CFD simulation.

The range of the averaged annual MRT in Figure 5 (b) seems reasonable as the average dry-bulb temperature is 10 °C. A closer inspection of the annual MRT time series revealed that the shaded MRT is generally very close to the dry-bulb temperature, whereas the nighttime MRT lies ~10°C below. This confirms the findings by [12]. The current simulation framework, however, does not take into account radiative heat exchange and therefore does not take into account urban heat islands that are likely to occur during summer. Although Kessling et al. [12] argue that the difference between surface and dry-bulb temperatures usually below ±5 K even in very hot climates, it is not clear if those findings are generalizable for arbitrary building densities, other than Riyadh. Here, more research is required to establish a more holistic and robust workflow.

It is worth noting that the ranges for the spatially-resolved sensitivity analysis in Figure 6 (a) and Figure 6 (b) differ from those in Figure 4. That is to say that locally, the bounds for each point differ significantly from the assumptions made in Table 1. Moreover, the areas with high total-order sensitivity indices in Figure 6 (a) and Figure 6 (b) confirm what was expected by the average annual UTCI distribution. The areas colored in red may be interpreted as the areas for which the respective variable (wind velocity / MRT) is controlling most of the variance in the UTCI output. For example, in areas where the wind velocity is high on average (north-south corridors), the wind velocity is relatively dominant. Besides, Figure 6 (b) suggests that wind shelters might be beneficial to improve the annual outdoor comfort for the seating arrangements to the west of the building as the MRT variance in the output is relatively low. Insights from such visualizations will prove to be helpful in other studies when decision-makers decide about possible interventions to promote good outdoor thermal comfort.

Future studies should combine spatial and temporal sensitivity studies. A preliminary sensitivity analysis of the UTCI for different climate zones has shown that the complexity of outdoor comfort metrics typically causes a subset of parameters to be inactive at any particular time; and that the inactive input variables differ from climate zone to climate zone. This sparsity of activation may lead to needless computational complexity and inappropriate assumptions of parameters that are inactive in the first place. Combining spatial and temporal sensitivity analyses (ideally by season) might present a valuable opportunity to overcome the complexity of simulation engines. In effect, this would be achieved by restricting the variable input space to only those parameters which are actively contributing to the output at a specific time and location for the particular climate zone.

5 CONCLUSIONS

Exploring parameter activation at the spatial and temporal scales is important not only for diagnostic analyses of biometeorological indices such as the UTCI but also to provide

actionable insights when deciding about potential interventions. This study represents a novel step in this direction by visualizing the spatial sensitivity of the UTCI and exemplifies how to derive the outdoor space usage as a proxy for the attractiveness of outdoor spaces. We conclude that the spatial variability of any outdoor comfort metric can easily be visualized which provides valuable information about its behavior. As the availability of computing power continues to increase, we anticipate the community to look beyond results only; instead, we anticipate using them to derive actionable insights which to date remained largely unexplored.

REFERENCES

1. Arens, E., Hoyt, T., Zhou, X., Huang, L., Zhang, H., and Schiavon, S. Modeling the comfort effects of short-wave solar radiation indoors. *Building and Environment 88* (2015), 3–9.

2. Becker, S., Lienhart, H., and Durst, F. Flow around three-dimensional obstacles in boundary layers. *Journal of Wind Engineering and Industrial Aerodynamics 90*, 4-5 (2002), 265–279.

3. Bourgeois, D., Reinhart, C. F., and Ward, G. Standard daylight coefficient model for dynamic daylighting simulations. *Building Research & Information 36*, 1 (2008), 68–82.

4. Bröde, P., Fiala, D., Błażejczyk, K., Holmér, I., Jendritzky, G., Kampmann, B., Tinz, B., and Havenith, G. Deriving the operational procedure for the universal thermal climate index (utci). *International journal of biometeorology 56*, 3 (2012), 481–494.

5. Broede, P., Blazejczyk, K., Fiala, D., Havenith, G., Holmer, I., Jendritzky, G., Kuklane, K., and Kampmann, B. The universal thermal climate index utci compared to ergonomics standards for assessing the thermal environment. *Industrial health 51*, 1 (2013), 16–24.

6. Bui, Q., and White, J. Mapping the shadows of new york city: Every building every block. *New York Times* (2016).

7. EnergyPlus Documentation. Engineering Reference - EnergyPlus 8.9. *The Reference to EnergyPlus Calculation* (2019).

8. Fiala, D., Havenith, G., Bröde, P., Kampmann, B., and Jendritzky, G. Utci-fiala multi-node model of human heat transfer and temperature regulation. *International journal of biometeorology 56*, 3 (2012), 429–441.

9. Herman, J., and Usher, W. Salib: an open-source python library for sensitivity analysis. *Journal of Open Source Software 2*, 9 (2017), 97.

10. Kastner, P., and Dogan, T. Towards high-resolution annual outdoor thermal comfort mapping in urban design. In *Proceedings of Building Simulation 2019: 16th Conference of IBPSA* (2019).

11. Kastner, P., and Dogan, T. A cylindrical meshing methodology for annual urban computational fluid dynamics simulations. *Journal of Building Performance Simulation 13*, 1 (2020), 59–68.

12. KESSLING, W., ENGELHARDT, M., and KIEHLMANN, D. The human bio-meteorological chart. In *PLEA* (2013).

13. Kongsgaard, C. Utci parameter sensitivity, 2012.

14. Provençal, S., Bergeron, O., Leduc, R., and Barrette, N. Thermal comfort in quebec city, canada: sensitivity analysis of the utci and other popular thermal comfort indices in a mid-latitude continental city. *International journal of biometeorology 60*, 4 (2016), 591–603.

15. Reinhart, C. F., Dhariwal, J., and Gero, K. Biometeorological indices explain outside dwelling patterns based on wi-fi data in support of sustainable urban planning. *Building and Environment 126* (2017), 422–430.

16. Subramaniam, S. Daylighting simulations with radiance using matrix-based methods. *Lawrence Berke-ley National Laboratory* (2017).

17. Tumini, I., and Rubio-Bellido, C. Measuring climate change impact on urban microclimate: A case study of concepción. *Procedia engineering 161* (2016), 2290–2296.

18. Whyte, W. H. *City: Rediscovering the center*. University of Pennsylvania Press, 2012.

19. Whyte, W. H., et al. *The social life of small urban spaces*. Conservation Foundation Washington, DC, 1980.

Evaluating Jogging Routes in Mass Models

Mathew Schwartz[1]

[1]New Jersey Institute of Technology, Newark, USA, cadop@umich.edu

ABSTRACT

Designers and Urban Planners use various tools and metrics to help guide their decisions. Evaluating an urban environment based on distance to amenities (i.e., a source and destination) is only part of the movement experience of an occupant. In recreational movement such as walking, jogging, and running, people plan a path in which the start and end point is the same. The route taken and the experience along this route has implications for the satisfaction of the occupant. This paper introduces a straightforward method of generating all possible routes via two waypoints that meet a user-defined criteria for the total length of the route based on travel speed and time. Two new evaluative metrics for the paths are introduced: the percentage of the path that is in shade as well as the percentage that contains the Sun in the glare range of a person. As the latter aspect is a function of direction, a directed graph is used to factor in the directionality of the recreational route. Similarly, the truest location of the person is required for evaluating the impact of the built environment so the sidewalk, rather than road centerlines, are used.

Author Keywords

human factors, walkability, computation, recreation, evaluation, urban planning

ACM Classification Keywords

I.6 SIMULATION AND MODELING : [General]; J.5 ARTS AND HUMANITIES: [Architecture]; D.2.2 SOFTWARE ENGINEERING: Design Tools and Techniques—*Human Factors*; I.3.6 COMPUTER GRAPHICS: Methodology and Techniques—*Ergonomics*

1 INTRODUCTION

The goal of this paper is to provide an evaluation of the virtual environment to a designer that encompasses the availability of comfortable jogging routes. While existing tools evaluate human experiences based on visibility to other locations or distances to transportation, there is yet to be a focus on occupant physical and visual experience based on solar lighting. Unlike real-time path planners, the goal is not to find an optimal or most comfortable route, but rather evaluate an exhaustive set of routes formulated based on the environment (in particular, the sidewalk and crosswalk surfaces) given. Using these routes (e.g., (Fig. 1)), various metrics can be calculated and provide a score– not of the route itself, but of the environment that led to those routes being created.
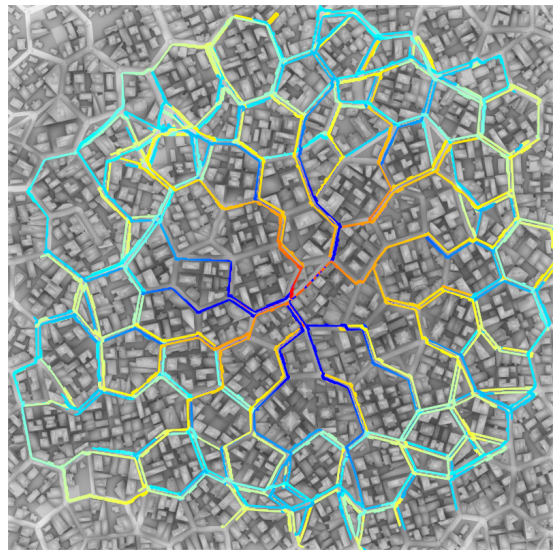
Figure 1: Overlay of 300 jogging paths generated. A color gradient is applied to each path corresponding to the start and end of the individual path color blue to red respectively.

When considering an urban environment for analysis related to the occupants, walkability is often at the forefront. The term walkability is a topic of research in numerous fields [1], yet even within urban design itself there are various meanings [9]. While specific terms and factors being analyzed vary, there is an underlying unification in goals that the built environment can influence both the behaviour and the health of the people within – albeit physical or mental.

Human subject studies are vital in the development of design process and thinking as they help to refine, and at times reject, the intuitive knowledge designers bring to the process. For example, a past study using infrared cameras and tracking of pedestrians throughout the trails in a city found and unexpected relationship that countered the urban design theory in which shorter street blocks induce higher levels of activity [12]. Research has shown the facilitation and increased recreational movement within an environment can be brought about by a variety of nature-based settings [15]. Consequently, it is feasible to consider the uncontrollable factors of jogging routes (i.e., the sun) when evaluating an environment, as additional interventions (e.g., creating a park or modifying the sidewalk surface condition) are suitable for increased activity. There is also an element of perceived vs. actual

attributes of an environment as Zuniga-Teran et al. (2019) argues; the perception of a neighborhood being pedestrian friendly is linked to visitation rates of greenspace [23].

Along with the risks (e.g., skin cancer), the benefit of sun exposure has been detailed in literature [16]. As the preferences and health benefits to every combination of sun and lighting conditions is a complex and still developing knowledge, this paper takes a computational approach to generating the possibilities and presenting the results, without determining the value of one over another. The proposal is then to develop analysis methodologies and techniques that find a large number of possible routes within a given constraint in order to analyze the environmental conditions one would experience along these routes.

With the existing literature emphasizing the importance of restoration and facilitation of the built environment for recreational mobility as a means towards a healthy community, the work presented here is positioned within the neighbourhood. The jogging routes calculated have a source and destination of the same location as the assumption placed is the person jogging, does so from their home, and returns back.

To avoid confusion with nuances in meaning, the terms; route, trail, and path, are used interchangeably to refer to an ordered list of locations a person travels along. The various terms are all used to stay consistent in terminology with the original authors of the literature review. In the computational perspective the term path is often used to describe an ordered list of nodes as it is derived from shortest-path algorithms, the underlying method used in this research. Recreational routes refers to walking, jogging, or running for leisure rather than an alternative mode of transportation. The emphasis on jogging is somewhat arbitrary for the methods in this research, but in terms of distance and speeds used in the results, it is the most similar type of activity.

1.1 Contribution

This paper focuses on route planning of closed cycles (Fig. 1) in a graph for evaluating the environment itself, rather than the path created. The metrics often used in route planning for filtering options for a user are stored as the layered costs for each route found to satisfy a given distance. Additionally, the metrics of route percentage that is in shade and in glare is demonstrated as a new addition to recreational routing evaluation. The evaluation system is written as a plugin to the Rhino 3D software in which a user can input their routing constraints and visualize aggregate or individualized data through Grasshopper.

The methods presented can benefit the urban design and planning field by evaluating and visualizing data related to the possible routes for recreational activities. This data can be used as an aid in decision making for defining areas of safer and pedestrian-centric intersections/crosswalks as well as unique (possibly more costly and with limited budgets) sidewalk materials and maintenance schedules. The methods for evaluating the paths are applicable to the evaluation of any route. However, the emphasis on jogging with route generation is due to the existing literature showing the circular path

attributes to be most applicable to this type of recreational path.

2 RELATED WORK

2.1 User Tracking

The impact of the built environment on a persons perception of their running route has been a topic of recent research, although it is still lacking. While there has been a longer history of study relating to walkability, there is still variation in what relationships exist between environmental factors and walking behaviour [8]. From the literature, there seems to be a difference on the restorative properties of a space based on the type of recreational activity.

A method for estimating trail traffic using recorded data was presented in [12] by extrapolating correlated features from a large dataset although it lacked in resolution for individuals and tracking source-destination pairs. It is important to note, however, the recent studies on evaluating the built environment related to comfort or attractiveness of a route are surveys with predefined questions (e.g.,[6, 8]) and historically the majority of analysis methods were either surveys or interviews (see the review on Landscape and Urban Planning pre-2008 [15]. The importance of this is that the predefined aspect may not be interpreted in relation to sun conditions. For example, Deelen et al. (2019) found a positive association with poor lighting and attractiveness of the environment [6]. While the researchers explain a possible reason, the question (which the authors state is based on [8]) was asking about street lighting. Likewise, these surveys lack resolution in the time of a recreational run as the reported categories are: before 9am, after 9am, afternoons, evenings. Depending on the time of year and how much *before 9am* it is, the condition of lighting can drastically change.

In Renswouw et al. (2019), the difference in GPS trails of people running showed that some areas in which people run during the day are avoided at night [21]. Hot climates (e.g., Singapore) may have significant differences in recreational running patterns than more moderate areas as suggested by Balaban and Tunçer (2017) – runners avoid the time of day the sun is strongest [2]. With this assumption, the increase in shade along a running path would be beneficial to facilitating this activity during these intense sunlight hours. Comparatively, a cooler climate may want to promote more direct sunlight exposure, while still avoiding glare. Studying attributes of parks that lead to physical activity, Brown et al. (2014) used a method of public participation geographic information system (PPGIS) technology to overcome shortcomings with survey data, but did not present an analysis related to sunlight [5].

With the known benefits of sun exposure, the impact of radiant heating and various climates, and physiological response to glare, the evaluation metrics of a path in this paper are well established for their merit, although have not been directly studied for jogging.

2.2 Path Planning

Computing recreational routes (i.e., walking, jogging, running, cycling) in a city has been demonstrated most often in
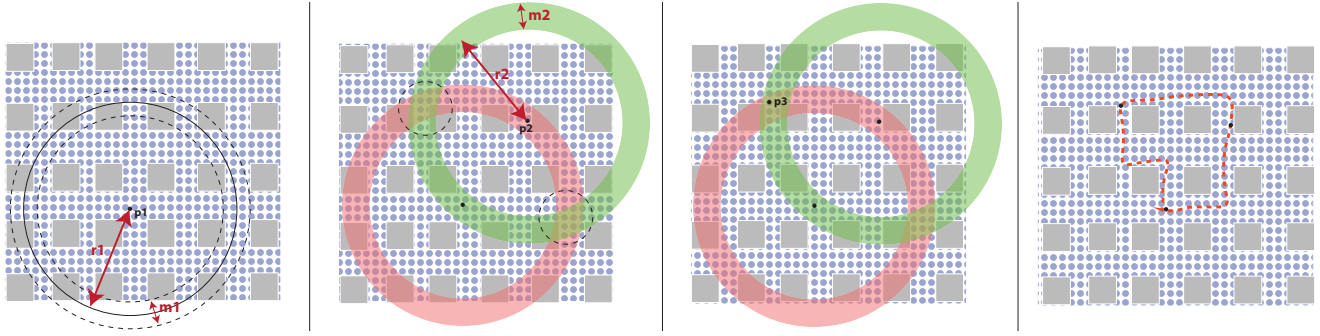
Figure 2: Visualization of waypoint selection process (from left to right) of a single given starting position. The process in steps 2,3,4 is applied for every node in the initial ring of step 1.

the field of computer graphics [10, 13, 22, 20]. However often in these works, particularly in the domain of human computer interaction, the emphasis is on the real-time generation of a single path for a user (e.g., on a phone app). The input data is largely based on maps (e.g., openstreetmap) and has little access to the three dimensional aspect of the built environment. The weights and environmental factors considered for the *goodness* of a route are often points of interest (POI) and influence the way in which the path is created, for example, in tourist routes [3, 4]. Sabine (2018) uses a method to query a source-destination shortest path with visible regions of interest without forcing the path to route through the POI directly [19]. Birsak et al. (2014) generates a tourist brochure using route planning between all pairs of POIs [3]. Their work intentionally reduces the graph iteratively as to reduce the number of overlapping routes.

A recommendation system for route planning of a closed cycle was shown in [13]. The framework also included numerous metrics to integrate in the final rankings (i.e., length, uniqueness, shape, lighting, elevation, pedestrian friendliness, turns, nature, history). In the case of lighting, the streets containing street lights were used and this metric was ignored during daytime. Similarly, Willamowski et al. (2019) focused on elevation as a driving input for the route planning of runners, emphasizing the generation of a new path using user preferences during construction (computationally, not physically) of the path compared to a final filtering step [22].

Waypoints and action descriptions (e.g., goals or tasks to complete) are often used in agent-based simulations to drive the agent in a particular direction. In optimizing sign placement for wayfinding, Huang et al. (2017) used a graph with paired source-destination points such as the bus stop to a POI, and POIs to eachother [11]. The nodes in the graph are based on road segments, a common approach of urban path planning and analysis. While much of the work in route planning for multiple POIs involves the computation of an *attractive* path, this goal does not align well with the evaluation of an environment. For example, Birsak et al. (2017) plans routes that may not be the shortest, but rather are more interesting by passing multiple POIs [4]. In the design of an urban environment however, if a large number of paths generated have

mostly uninteresting routes, it is the design itself that can be modified. Mathew et al. (2018) generates an environment to create a desired crowd simulation effect [14]. This is similar to the generative design approach when the optimization metrics are evaluative scores on the environment.

While not integrating path planning, to analyze the impact of shadows in the urban environment, Miranda et al. (2018) introduced shadow accrual maps as a fast method for visualizing overlapping shadows in a city for city planners and architects [17]. This evaluation method integrated the three dimensional aspect of the buildings as it was similarly geared towards spatial evaluation for designers and planners.

Differences in both the initial data and goal of the path planning present a gap in the current state of the art relating to jogging route generation. As the evaluation of a 3D model for design purposes is evaluating possibilities of multiple people, the distance between locations is fully computed. Similarly, evaluating the impact of a building design on the human experience is done with the intention of understanding how changing that building will also impact people. Therefor the dimensionality is critical, rather than the network by itself.

3 METHODOLOGY

To provide an overall evaluation of diversity among loop-based jogging routes in the environment, three main steps are taken. First, the traversable surfaces of the environment are discretized into a grid-based dense graph in which the nearest neighbors are connected (eight neighbors) with an edge cost equal to the euclidean distance using the method in [18]. For the mass model type used in this paper and output by CityEngine, these surfaces are the sidewalks and crosswalks. Once this graph is created, Dijkstra's All-Pairs-Shortest-Path (ASPS) algorithm is run to get the distance from all nodes to each other. The next step is to find a sequence of nodes that satisfy a user-specified distance a jog should be such that the sum of edge costs in the sequence are within some threshold of the total distance. This distance is input as a parameter composed of speed of the jog, desired length of time, and a percentage of tolerance to that length of time. The loop-like jogging routes available from a given point (e.g., a home in the environment) is calculated by finding the nodes within
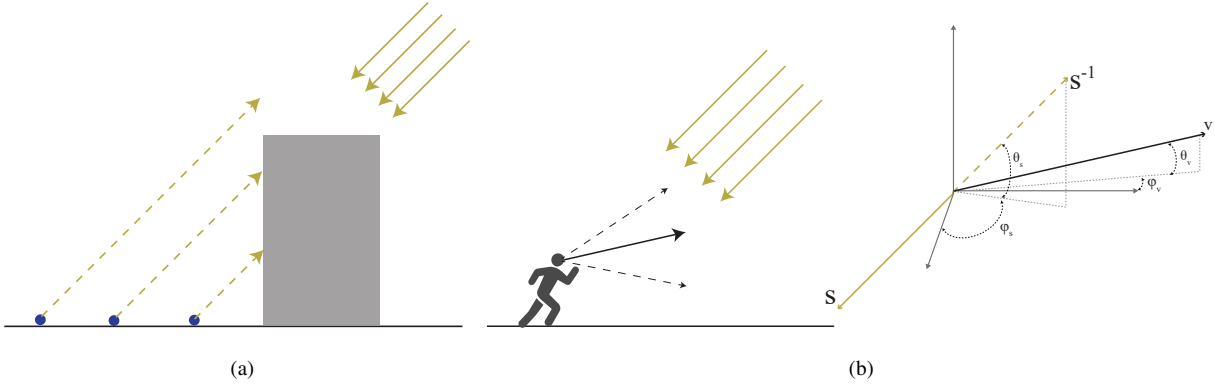
Figure 3: (3a): Illustration of the inverse sun vector being cast from a node to determine if it is within a shadow or not. (3b): Calculation method for determining if a direction a person is moving would be subject to glare from the sun.

one third the total distance and constructing a boolean vector for each of those valid nodes. This vector is then multiplied against each row of the distance matrix in which there is a corresponding true element, providing a set of tertiary nodes for the sequence.

Once the routes are created, the two evaluation metrics are calculated for a series of time periods defined by a user. First, the nodes that belong to the set of jogging routes are checked to be in a direct view of the sun or not (shadow). This is done by casting a ray in the opposite direction of the sun vector, starting at the node location with an offset to a human eye level (essentially a *shadow ray*). If the ray intersects any geometry (e.g., buildings) within the scene, it must be in shade. The next step is to take the set of edges, which are ordered node pairs, where the second node is not in shade, and compare the vector from the first to second node with the Sun vector. As there are varying amounts of discomfort for glare based on the location within the field of view, the angle threshold between these two vectors is determined by a parameter entered by the user.

3.1 Paths Computation

Fundamentally there are an infinite number of ways a path can be created in a given length. For an urban environment discretized to a 1 or 2 meter grid, while not infinite, the number of possible paths of a jogging length (e.g., 1000 meters) is impractical to calculate. More importantly, these path options are impractical to use in a meaningful way. For example, from a given location, walking between two nodes 1000 times would qualify. Techniques for the generation of paths often use a cost structure and find the minimal cost between two points. Rather than creating rules such as adding costs to edges that have been traversed, this paper uses a method of selecting sets of nodes that are inherently at appropriate distances. The general approach to finding the waypoints of the cycle is similar to the *2-via-routes* ring method in [10] with a modified ring distance and precomputed APSP distance matrix.

A graph representation of the environment provides a discretized value of distances between locations that can be considered for a path. For evaluation based on pedestrian movement, these locations are within the bounds of sidewalks and crosswalks. From this graph, an All Pairs Shortest Path (APSP) algorithm, in this case Dijkstra, to generate a distance and predecessor matrix. The distance matrix is an $n \times n$ matrix corresponding to the distances between each node such that a row i contains the distances of v_i from $v_1 \ldots v_n$ and where the distance value of v_i to v_i is 0. The predecessor matrix is structured the same as the distance matrix except that instead of the distance between the two given nodes, it is the node with the minimal cost to it.

The distance matrix can then be used to find the set of nodes in the graph that are within some distance threshold from it. In the case of jogging path calculations, a desired distance can be given directly, or as implemented in here, a combination of jogging speed and desired time. To aid in the user experience of the grasshopper plugin a pedestrian speed of kilometers per hour and the total time in minutes is entered. As the distance matrix is set in meters, the total distance calculation is $T_d = k * 16.6667$. While T_d can be used to find all nodes from the start location that are the desired distance away, it would not account for the pedestrians travel time back. Unlike a single source-destination pair, the start and end location is the same, and when considering the distance matrix that is calculated by the Dijkstra APSP, the value of the same node to itself is 0. To resolve the distance discrepancy the easiest solution is to divide T_d by half. While the directionality would be useful in this approach, the path itself would remain as a line and have little variation for the pedestrian. Instead, waypoints are created to divide the path into multiple segments.

In the case of a path created by three nodes n_1, n_2, n_3 the predecessor matrix will be used to calculate the path of $n_1 \rightarrow n_2$, $n_2 \rightarrow n_3$, and $n_3 \rightarrow n_1$. Therefore, a path can be constructed by selecting three evenly distanced nodes of the predecessor matrix. The exact distance is unlikely to exist and therefore a margin of error must be incorporated. This approach uses four variables; a radius from the start node r_1, a margin threshold for variation of that radius m_1, a secondary radius defined from the second node r_2, and likewise a second margin m_2.
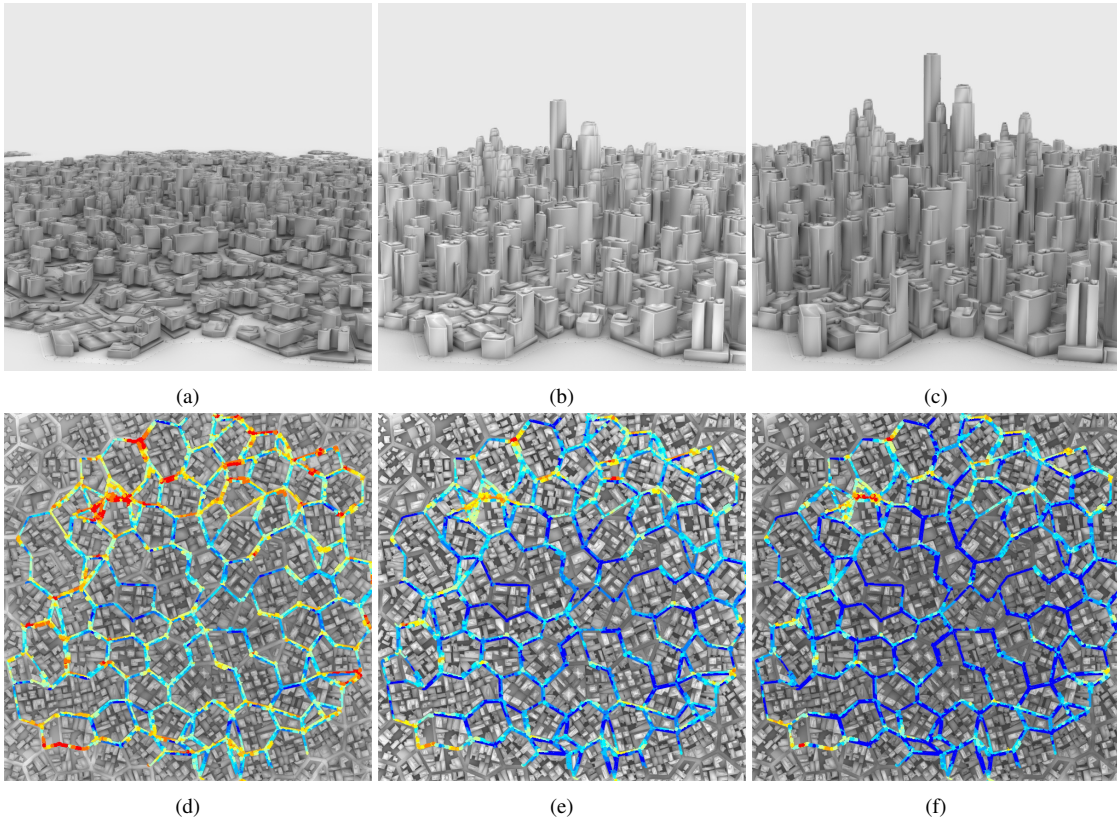
Figure 4: (4a-4c) perspective view of the generated city with increasing building heights from ESRI CityEngine. (4d-4f) A top view of the city (corresponding to the city directly above each image) with accumulative shadow scores calculated by the average number of times a node was in shadow or not over the designated sample period defined in the input parameters. The color is based on a low occurrence of times in the sun (blue) to high rate of occurrence (red).

In Figure 2, the process of finding the waypoints is demonstrated. First a node (referred to as point p_1) is selected as the start and end. The set of possible secondary nodes L are found by checking which nodes in the distance matrix are within the threshold, such that $\{v \in L : r_1 - (m_1/2) > D_{p_1}(v) > r_1 + (m_1/2)\}$, where $D_{p_1}(v)$ is the distance of a node v from p_1. The next step is for each possible secondary point, (in Fig. 2, one example is shown labeled p_2 which falls within the red zone of p_1) find the final third node within the threshold distance. To be able to accurately retrieve a three node combination of a specified distance, the set of tertiary nodes is intersected with the secondary node set (dotted circles in Fig. 2).

The final step in Figure 2 is the dotted red line representing a calculated path that avoids obstacles such as buildings (as their are no nodes there), and generates an appropriately long path. The values of r_1 and r_2 are functionally independent. However in this work they are set to be equal as this creates evenly distanced waypoints such that $D_{p_1}(p_2), D_{p_2}(p_2), D_{p_3}(p_1) \approx T_d/3$ and a mostly circular route is formed within the city.

After generating all the paths, a random sample is taken for visualization and analysis in Rhino, assuming the desired number of paths is less than the total calculated. As the values of m_1 and m_2 increase, the total number of paths increases as well. In rough terms, on a graph of $\approx 50,000$ nodes, $\approx 9,000$ paths are calculated in ≈ 10 seconds. From the sampled paths, any node used in a path is stored for analysis. The use of this path set has the additional benefit of reducing the calculations needed of the entire graph as only the nodes associated with a jogging path are used in the jogging path analysis.

3.2 Evaluation Metrics

The first evaluation metric is determining which nodes along the path are in shade. As this is independent of the path, all nodes can be calculated simultaneously without regard for order. A range of values for time of day are supplied, including the month, longitude, and latitude. For all the dates in range, the sum of the time a node is within shadow or not is stored. For determining the node is in a shadow or not for a given time, the inverse vector of the sun direction is cast as a ray from the node and checked for intersections with the models in the environment (Fig.3a).

The second evaluation metric is to determine if an edge along a path would be within a specified range such that the sun would cause discomfort from glare to a person. As illustrated

in Figure 3b, the view direction of a person is estimated by the vector of the ordered nodes making up an edge in a path. If both nodes of an edge are in shadow, the glare calculation can be skipped as the sun would not be visible even if the vectors are the same. A range determined by the user of the plugin for determining the cutoff for glare is used as the threshold t_θ, t_φ between the Sun and human view vectors. To make the calculation method understandable and transparent, the cartesian space-based vectors are converted into spherical coordinates. Similar with the shadow calculation, the comparison of view to the sun vector s is done with the inverse of the vector s^{-1}. The x,y,z coordinates are then converted to azimuth φ, elevation θ, and a radius (which is ignored for unit vectors of length 1). The sun is within the glare range if $abs(\theta_v - \theta_s) < t_\theta$ and $abs(\varphi_v - \varphi_s) < t_\varphi$.

To verify the sampling choice of 300 paths to be adequate for this preliminary analysis and demonstration, the evaluation metrics for shade and glare of two sets of separately generated paths are plotted as histograms in Figure 5. As evident in the image, the consistency between the two sets is strong.

(a) Glare (b) Shadow
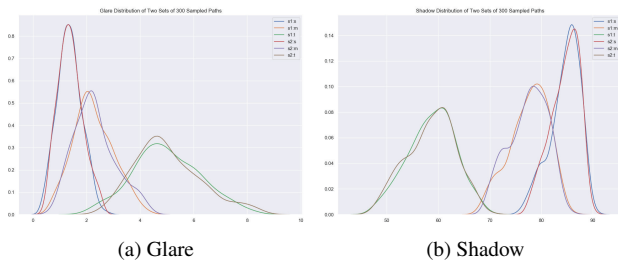
Figure 5: Histogram comparison between two sets of 300 paths sampled from the total possible with the shadow analysis.

4   RESULTS

With the aggregated jogging paths shown in Figure 1, the choice of certain sidewalks and streets over others can be seen in relation to the given starting point (where the blue and red parts of the paths meet). While it is clearly possible one would choose to use these side streets, the algorithms results are logical in that the most underutilized streets are perpendicular to the outgoing radial lines. To further illustrate the types of paths that are created, individual examples are given in Figure 6.

The example city used for this demonstration was generated in ESRI CityEngine [7]. A hexagonal grid was used with three variations of the same city in which the building height scale was modified (Fig. 4a-4c). The height is a multiplying factor from the original, rather than an overall addition of floors. This demonstrates the impact of relatively similar sized buildings (Fig. 4a) to one in which the city center has skyscrappers (Fig. 4c).

Based on the peak times and month for urban trail utilization in [12], the following demonstration of results utilize a time of day ranged from 06:00 - 18:00 during August 1st, 2019. The longitude and latitude were set as $44.725°$ and $-74.919°$

(a) (b) (c)
(d) (e) (f)
(g) (h) (i)
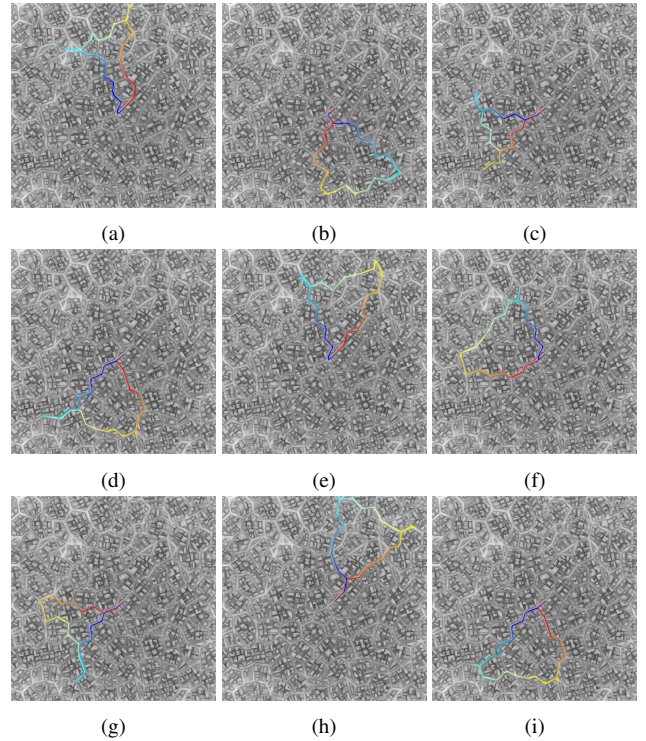
Figure 6: Sample Jogging Paths. Each path is colored by the distance from the start, where blue is the beginning and red is the end.

respectively. The impact of sun glare to a human varies by angle, however a threshold of $50°$ is used for the results here. For this example, the jogging speed was defined as 7km/hr (a light jog) and the city evaluation was performed for 15 minute jogs (7km/hr = 116.7 m/min), at 1750 meters. Each margin as described in Sec. 3 were set to 15%, resulting in a jogging path range between $\approx 1485 - 2000$ meters.

The visualization of the shadow parameters is shown in Figures 4d-4f. In this visualization it is easy for a designer to see the locations in which sun will often be on a node or not. Depending on the design intention, this could be an opportunity to add shade (naturally e.g., with trees) or emphasis for the jogging route. The shorter building heights in Fig. 4a correspond to the shadow scores in Fig. 4d. Comparing Fig. 4d to Fig. 4f, the individual normalized values (i.e., the normalization in the colors are not impacted by the design alternatives) demonstrate how the larger buildings within the center provide very little sunlight compared to more open areas near the outer limits of the jogging paths.

One method of understanding the relationship between the urban design options is through the histogram plots. As seen in Figure 7, the difference in percentage of the jogging paths in shadow are significantly different between the three building height models. For the small building heights (colored in green), the range of path percentage in shadow is from $\approx 50\%$ to $\approx 70\%$ while the tallest building variant (colored in blue) ranges from $\approx 70\%$ to $\approx 90\%$. Just as noticeable is the re-
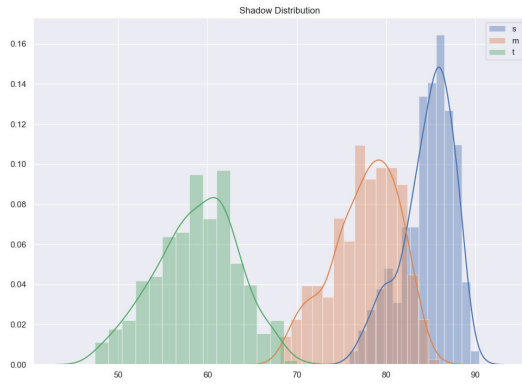
Figure 7: Histogram with kernel density estimation for the percentage of nodes of each jogging path that are in shadow for the three building heights. The colors correspond as follows: Fig. 4a-green, Fig. 4b-orange, Fig. 4c-blue.
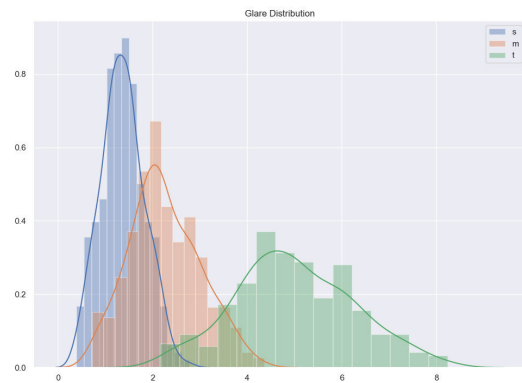
Figure 8: Histogram with kernel density estimation for the percentage of edges in each jogging path that are within the glare range for the three building heights.

duced distribution of bins between the two options. If consistency along a path is desired, the more concentrated values of option 4c may be considered better.

The evaluation scores based on the percentage of the paths that contain sun glare are similarly proportioned with slightly more overlap, seen in Figure 8. As may be expected, the tall building city variant has very little of the path in the glare range. While demonstrated in this paper is the difference in building heights, alternative design options related to glare could include the orientation of the street network. As the angle of glare is dependent on the direction of the path, an environment with hills and a surface terrain that is not flat would likely have a larger impact in the variation of glare (which the current framework supports).

To illustrate the importance of the utilization of sidewalks and crosswalks for the evaluation, a comparison between a graph that is generated within a 1 meter road and the sidewalks was done. In both cases, 300 samples were taken of the calculated jogging paths with the same approximate source-destination
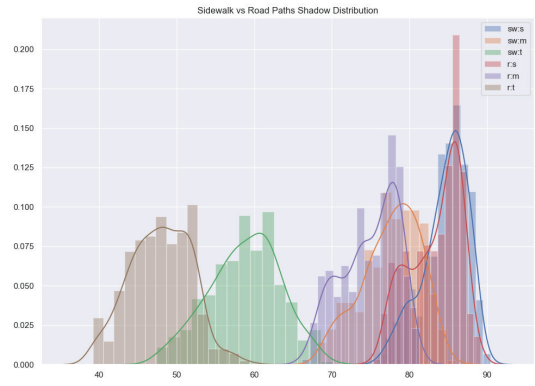
Figure 9: Comparison of histograms for the percentage of all paths that are in shadow when using the road center vs the sidewalk.

location. In Figure 9 the two sets of results are shown for the shadow percentage along paths.

It is easy to see why the large difference in road vs sidewalk evaluations exist when considering Figure 3a as the points closest to the building will have a larger angle in which the sun vector must be to provide a line of sight. When the center of the road is used, the shadow cast by a building is less likely to impact it.

5 DISCUSSION

This paper presented a method for generating paths within a user-specified distance that start and end in the same location (i.e., jogging paths). Two methods for evaluating the paths are presented; the percentage of a path in shadow and percent facing the sun within a glare range. An insightful result from the analysis shows the large difference in the shadow-based analysis metric when using paths generated from a road surface (as is the case with road centerlines) compared to on the sidewalk. Further analysis and comparisons using this generation and evaluation technique will include terrain with more variation such as hills as the angle of view will more drastically change during human mobility.

In the urban design discourse, walkability is a frequently used term and method of analyzing the built environment, and subsequently, as a means to evaluate the difference between design options. With the broad scope in definition and metrics being used, this paper introduces two more: metrics that are related to the directionalality of movement and not just distance between a single source and different destination. This is combined with element of time in which both the position of the sun and direction of the persons movement are incorporated. When sampling the sun position at 60 minute intervals, a route that takes 15 minutes to complete has no change in the sun vector. The evaluations provided by the framework presented here would then be logical only within that sampling period. Future work will investigate the impact of using the change in sun vectors over the duration of the jogging path.

While the shade and glare metrics can be applied to any path, the method for jogging paths is logical to the assumption that

a person starts and ends at their own home. As seen in the literature, often people will intentionally jog around a designated area such as a park, in which this current method does not account for. The motivation here is that an environment should not limit the potential for numerous variants of 15 minute jogging routes to beginning within a park.

Acknowledgements

REFERENCES

1. Andrews, G. J., Hall, E., Evans, B., and Colls, R. Moving beyond walkability: On the potential of health geography. *Social Science & Medicine 75*, 11 (2012), 1925–1932.

2. Balaban, O., and Tuncer, B. Visualizing and analyzing urban leisure runs. In *ShoCK!–Sharing of Computable Knowledge! Proceedings of the 35th International Conference on Education and Research in Computer Aided Architectural Design in Europe*, vol. 1, Faculty of Civil and Industrial Engineering Sapienza University of Rome (2017), 533–540.

3. Birsak, M., Musialski, P., Wonka, P., and Wimmer, M. Automatic generation of tourist brochures. In *Computer Graphics Forum*, vol. 33, Wiley Online Library (2014), 449–458.

4. Birsak, M., Musialski, P., Wonka, P., and Wimmer, M. Dynamic path exploration on mobile devices. *IEEE transactions on visualization and computer graphics 24*, 5 (2017), 1784–1798.

5. Brown, G., Schebella, M. F., and Weber, D. Using participatory gis to measure physical activity and urban park benefits. *Landscape and Urban Planning 121* (2014), 34–44.

6. Deelen, I., Janssen, M., Vos, S., Kamphuis, C. B., and Ettema, D. Attractive running environments for all? a cross-sectional study on physical environmental characteristics and runners' motives and attitudes, in relation to the experience of the running environment. *BMC public health 19*, 1 (2019), 366.

7. ESRI. CityEngine. `https://www.esri.com/en-us/arcgis/products/esri-cityengine/overview`, 2020. [Online; accessed 7-January-2020].

8. Ettema, D. Runnable cities: How does the running environment influence perceived attractiveness, restorativeness, and running frequency? *Environment and Behavior 48*, 9 (2016), 1127–1147.

9. Forsyth, A. What is a walkable place? the walkability debate in urban design. *Urban design international 20*, 4 (2015), 274–292.

10. Gemsa, A., Pajor, T., Wagner, D., and Zündorf, T. Efficient computation of jogging routes. In *International Symposium on Experimental Algorithms*, Springer (2013), 272–283.

11. Huang, H., Lin, N.-C., Barrett, L., Springer, D., Wang, H.-C., Pomplun, M., and Yu, L.-F. Automatic optimization of wayfinding design. *IEEE transactions on visualization and computer graphics 24*, 9 (2017), 2516–2530.

12. Lindsey, G., Wilson, J., Rubchinskaya, E., Yang, J., and Han, Y. Estimating urban trail traffic: Methods for existing and proposed trails. *Landscape and Urban Planning 81*, 4 (2007), 299–315.

13. Loepp, B., and Ziegler, J. Recommending running routes: framework and demonstrator. In *Workshop on Recommendation in Complex Scenarios* (2018).

14. Mathew, C. T., Knob, P. R., Musse, S. R., and Aliaga, D. G. Urban walkability design using virtual population simulation. In *Computer Graphics Forum*, vol. 38, Wiley Online Library (2019), 455–469.

15. Matsuoka, R. H., and Kaplan, R. People needs in the urban landscape: analysis of landscape and urban planning contributions. *Landscape and urban planning 84*, 1 (2008), 7–19.

16. Mead, M. N. Benefits of sunlight: a bright spot for human health, 2008.

17. Miranda, F., Doraiswamy, H., Lage, M., Wilson, L., Hsieh, M., and Silva, C. T. Shadow accrual maps: Efficient accumulation of city-scale shadows over time. *IEEE transactions on visualization and computer graphics 25*, 3 (2018), 1559–1574.

18. Schwartz, M., and Das, S. Interpreting non-flat surfaces for walkability analysis. In *Proceedings of the symposium on simulation for architecture & urban design*, Society for Computer Simulation International (2019), 287–294.

19. Storandt, S. Region-aware route planning. In *International Symposium on Web and Wireless Geographical Information Systems*, Springer (2018), 101–117.

20. Stroobant, P., Audenaert, P., Colle, D., and Pickavet, M. Generating constrained length personalized bicycle tours. *4OR 16*, 4 (2018), 411–439.

21. van Renswouw, L., Bogers, S., Lallemand, C., and Vos, S. Exploring the value of user-generated app data to design and improve urban running environments. *Crossing Borders in Research on Sport and Physical Activity*, 95.

22. Willamowski, J., Clinchant, S., Legras, C., Michel, S., and Shreepriya, S. Running tour generation for unknown environments. In *International Conference on Human-Computer Interaction*, Springer (2019), 528–535.

23. Zuniga-Teran, A. A., Stoker, P., Gimblett, R. H., Orr, B. J., Marsh, S. E., Guertin, D. P., and Chalfoun, N. V. Exploring the influence of neighborhood walkability on the frequency of use of greenspace. *Landscape and urban planning 190* (2019), 103609.

Build2Vec: Building Representation in Vector Space

Mahmoud M. Abdelrahman[1], Adrian Chong[1], Clayton Miller[1]

[1]**Department of Building, School of Design and Environment (SDE),
National University of Singapore (NUS), 4 Architecture Drive, Singapore 117566.**

ABSTRACT

In this paper, we represent a methodology of a graph embeddings algorithm that is used to transform labeled property graphs obtained from a Building Information Model (BIM). Industrial Foundation Classes (IFC) is a standard schema for BIM, which is utilized to convert the building data into a graph representation. We used node2Vec with biased random walks to extract semantic similarities between different building components and represent them in a multi-dimensional vector space. A case study implementation is conducted on a net-zero-energy building located at the National University of Singapore (SDE4). This approach shows promising machine learning applications in capturing the semantic relations and similarities of different building objects, more specifically, spatial and spatio-temporal data.

Author Keywords

Graph embeddings; node2vec; STAR; Feature learning; Representation learning

1 INTRODUCTION

The amount of data generated during the last two decades exceeds that which has ever been generated in history. This growth is due to the radical evolution of Internet of Things (IoT) networks of interconnected objects that are used in sensing the surrounding environment (sensors) or controlling the physical world (actuators) [10]. Also, more powerful computational power and algorithms have helped in managing and processing these data to extract information [11, 19]. In the building industry, IoT, computational resources, and algorithms have pushed our understanding of different scales and interactions of the built environment. Scales, in this context include city, group-of-buildings (district), building, and human scale in addition to the network of interactions between those different scales [20]. In this research, we focus on the extraction of spatio-temporal data from buildings and the representation of them in an embedded vector space using graph-embeddings [27, 24, 26]. We refer to spatial data such as, spaces, walls, doors and windows; and temporal data as IoT sensors such as, indoor environmental data (temperature, humidity, noise), energy consumption data (equipment load, lighting load, HVAC) and occupants data, which could be considered as a movable sensors [16], (presence, thermal

comfort, location over time). This paper aims to introduce *Graph-embeddings* as an effective method to capture the spatial and temporal complexity within the buildings and the representation of them in a single vector space for the aim of improving prediction, classification, and recommendation accuracy. Potential applications are presented in the Discussion Section.
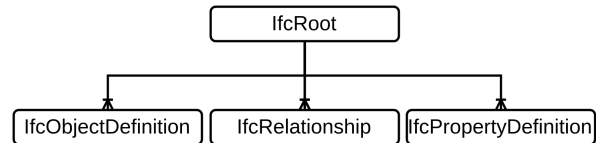
Figure 1. Ifc abstract structure
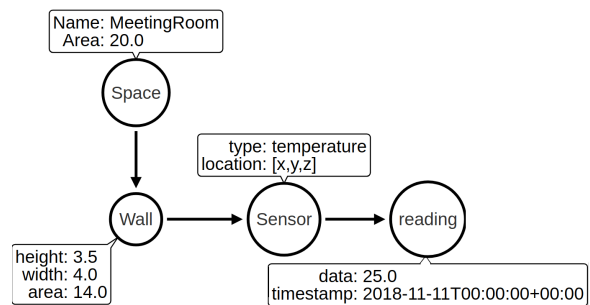
Figure 2. Example of attributed graph, where nodes and edges can have additional attributes

There are many ways to represent the spatial data of the buildings, amongst which, the Building Information Model (BIM) is the most well-known. BIM models are used in integrated design that enables multiple stakeholders to work on one embedded platform throughout the building life-cycle. This technology enables creating a rich repository of data of the building during its life-cycle, including spatial representations [6, 17]. BIM models are a static representation of the building; they do not include the temporal data from IoT sensors. Several studies use different methods to fuse the BIM models with IoT data by using middle-ware or representational state transfer (RESTful) API that connects the virtual BIM sensor with the corresponding IoT node [12, 3].

There are many software/tools available for creating, editing, and managing BIM models. In this study, we use the Industrial Foundation Classes (IFC) which is an open, international
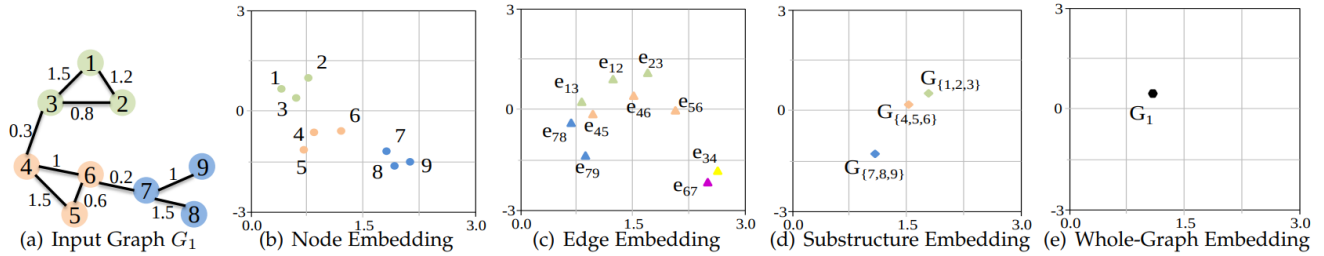
Figure 3. Graph embeddings convert the graph structure (vertex and edges) into vector representation. Image adapted from [4]

standard (ISO 16739-1:2018) digital description of the built environment [1]. The IFC model consists of objects built in a hierarchical structure, which represents the physical components of the buildings (IFCObjects), the relationships between different objects (IFCRelationship), and the properties of objects (IFCPropertyDefinition) as shown in Figure 1. The hierarchical structure of IFC enables extracting the entity data as a graph. A *graph* is a data structure that consists of a diagrammatic representation or objects and the relation between them in the form of *nodes* (referred to as a vertex) and *edges* connecting those nodes [22, 8]. However, a simple graph that consists of vertex and edges is not sufficient to represent the complexity of the BIM data (relations and attributes), and the IoT temporal data. Thus, an attributed graph is more suitable to handle such data. An attributed graph can handle rich information by enabling attributes on both edges and nodes (Figure 2). Previous work has introduced methodologies to convert the BIM model to graph theory using an online tool called BIMServer [13, 14].

Graph embeddings is a method to convert graphs into vectors. An example is seen in Figure 3. For a weighted graph G, where $G(V, E, W)$ is a graph and V denotes vertices, E denotes edges, and W denotes weights:

$$GraphEmbedding(G(V, E, W)) \rightarrow \mathbb{R}^n$$

Most machine learning models work with the feature vector representation of data where an instance is multi-dimensional and represents features in different ways (numeric data, binary, categorical,etc.). The motivation behind learning embeddings is to perform important tasks on networks such as classification [25, 5, 18], prediction [9] and clustering [21]. For example, predicting the most probable label of a node in a graph could be used as a recommendation system. Cai et al. conducted an extensive survey on graph embeddings and applications [4]. For this research, we will be using two types of graph embeddings, namely, node2Vec [9], and Spatio-Temporal Attentive RNN (STAR) [23]. The main difference between these models is that **node2Vec** works well with static graphs, i.e., not temporal and it does not support directional graphs, while **STAR** model can capture the temporal attribute changes of the graph which makes it ideal for situations where sensors' temporal data included within the graph network.

2 METHODOLOGY

2.1 Case study

Our case-study is the SDE4 building, a net-zero energy building located in the School of Design and Environment (SDE) at the National University of Singapore (NUS). We selected this case study for several reasons:

1. Smart building - There are sensors deployed in every space measuring indoor, outdoor environmental data and indoor localization.

2. Data Openness: There is to access both spatial data (BIM models) and temporal data (sensors/actuators/weather stations) and thermal comfort data (user feedback).

3. Variety of HVAC systems: A hybrid cooling system is adopted and it is made up of AC and natural ventilation systems, which give richness to the model.

2.2 Data sources

There are three main sources of data used in this study: (1) Spatial data: using the IFC file of the BIM model as seen in Figure 4, (2) Temporal data from the sensors in the form of sequential snapshots, and (3) occupant comfort feedback data from a experimental implementation.
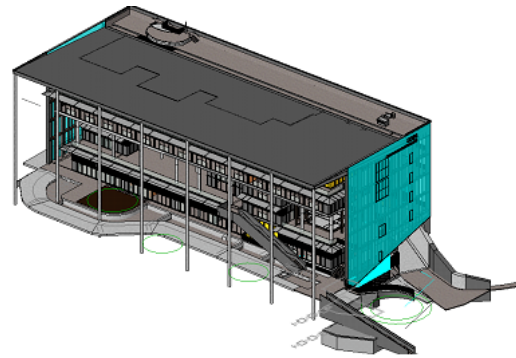
Figure 4. The SDE4 case study building BIM model representation

The IFC file for the SDE4 building is converted from an *Autodesk Revit* model of the building and doesn't include data about the sensors' locations. Additionally, occupants are treated as movable sensors (nodes) whose feedback is the prediction target on hot encoder i.e. comfortable = [1,0,0], uncomfortable=[0,1,0] and neutral=[0,0,1]. These data are not available in the initial version of the IFC file, therefore a preprocessing step of the occupants data should be first performed.
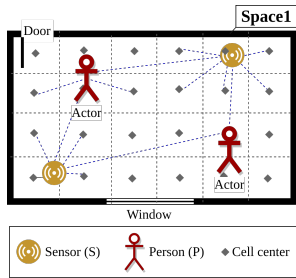
Figure 5. Spaces are discretized and objects are dynamically connected to each other.
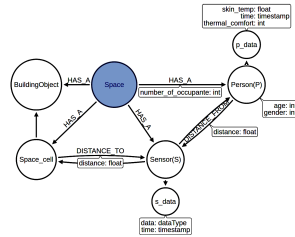
Figure 6. A temporal graph is dynamically changing at each time step

Figure 7. Snapshot of the animated demo of the attributed graph schematic design after discretizing the space into cells (link to animated online version found in the text)

There are two types of sensors deployed into the building spaces: (1) Sensors from the Building Management System (BMS), including in the thermostats that measure the ambient temperature, humidity, and CO2 levels and (2) other non-permanent IoT sensors deployed within the building in different spots that measure air temperature, humidity, light levels, noise levels, and CO2 levels.

The user feedback data were collected during an experiment conducted in the building that included over 30 participants that gave high-frequency subjective comfort feedback using micro ecological momentary assessments on a smart-watch including their indoor location, and thermal sensation feedback [15]. The location of each test participant was collected using the YAK mobile indoor localization app [2].

2.3 Framework implementation example

A sample of the framework is shown in Figures 5 and 6. If *Space1* from the IFC file is targeted, this space is a supertype of some other objects such as *IfcDoor*, *IfcWindow*, *IfcWalls*, *IfcFloor*, and *IfcCieling*. However, an occupant does not have a corresponding *dynamic* IFC object, so a method was developed to trigger the occupant once they moves into the space. The space is first (discretized) into two-dimensional structured cells using a spatial-discretization algorithm [7]. Spatial discretization divides the space shape into finite elements in the form of graph objects. Secondly, the relation between a moving person in space and the adjacent cells is created at each time-step as shown in the snapshot of Figure 7. Similarly, a relation (edge) between each sensor and the adjacent cells is created. An animated illustration of Figure 7 can be viewed online[1].

Two models of graph objects are created for both the node2Vec and STAR methods. The first model is a simple graph that represents the relation between every two objects in the form of an `(object->object)` relation as seen in Figure 8. A live demo of this Figure is found online[2]. To view the demo properly, deselect the *Sphereize data* on the left menu. It can be noticed that nodes 82 and 65 are spatially close to each other, this means that they share the same topological structure, or in this case, *IfcSpace*. A higher resolution image of Figure 8 can be viewed online[3]. The second model

consists of a multi-dimensional adjacency tensor, where the third dimension captures the temporal attributes of the graph, as seen in Figure 9. This tensor is then fed into a gated recurrent neural network first to capture the critical features and then to ignore useless features.
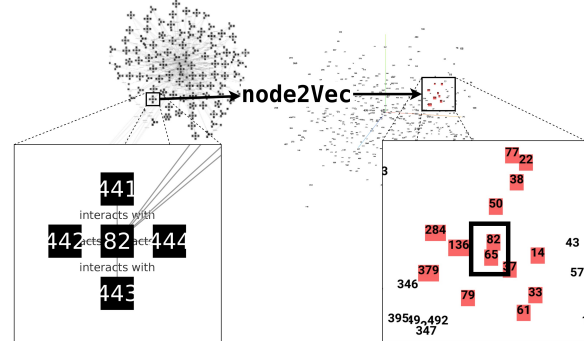
Figure 8. Feature learning using node2Vec for the SDE4 building (link to a live online demo found in the text)
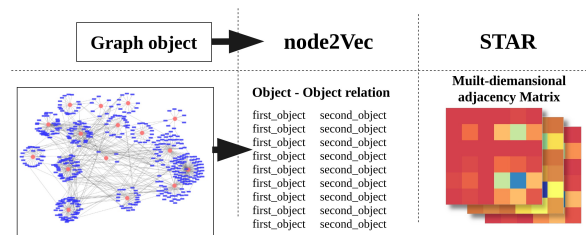
Figure 9. Converting the graph object into different formats for the feature learning process (link to a higher resolution online version found in the text)

3 CONCLUSION

In this paper, we introduced and implemented a framework of an embedding graph representation of buildings into a lower-dimensional vector. This process could enhance the machine learning capabilities of the complex relationships between the spatial model, temporal data sources, and dynamic occupant behavior. This application is challenging as not many simulation tools can translate the data from these sources into

---

[1] https://youtu.be/4iFuQvKG_Wg

[2] http://tiny.cc/ia0qiz

[3] http://tiny.cc/o874lz

a comprehensive model. Vector representation is important as a feature input in the learning process to predict, classify, or cluster relative nodes, values and/or edges attributes. We identified two possible algorithms to be used: node2Vec and STAR and show them in the context of the SDE4 building. This work is a preliminary step towards the conversion of BIM models into the vector space.

REFERENCES

1. Industry Foundation Classes (IFC) - buildingSMART Technical.

2. Abdelrahman, M. M., Jayathissa, P., and Miller, C. YAK: An Indoor Positioning App for Spatial-Temporal Indoor Environmental Quality Research, 2019.

3. Bottaccioli, L., Aliberti, A., Ugliotti, F., Patti, E., Osello, A., Macii, E., and Acquaviva, A. Building Energy Modelling and Monitoring by Integration of IoT Devices and Building Information Models. In *Proceedings - International Computer Software and Applications Conference*, vol. 1 (2017), 914–922.

4. Cai, H., Zheng, V. W., and Chen-Chuan Chang, K. A Comprehensive Survey of Graph Embedding: Problems, Techniques and Applications. Tech. rep., 2017.

5. Chang, S., Han, W., Tang, J., Qi, G.-J., Aggarwal, C. C., and Huang, T. S. Heterogeneous Network Embedding via Deep Architectures.

6. Eastman, C., Liston, K., Sacks, R., and Liston, K. *BIM Handbook: A Guide to Building Information Modeling for Owners, Managers, Architects, Engineers, Contractors, and Fabricators*. 2008.

7. Fischer, C., Nanz, G., and Selberherr, S. Finite difference, boundary-fitted grid generation for arbitrarily shaped two-dimensional simulation areas. *Computer Methods in Applied Mechanics and Engineering 110*, 1-2 (1993), 17–24.

8. Gross, J., and Yellen, J. *Handbook of graph theory.* 2004.

9. Grover, A., SIGKDD, J. L. P. o. t. n. A., and 2016, u. node2vec: Scalable feature learning for networks. *dl.acm.org*.

10. Gubbi, J., Buyya, R., Marusic, S., and Palaniswami, M. Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems 29*, 7 (2013), 1645–1660.

11. Hilbert, M. Big Data for Development: A Review of Promises and Challenges. *Development Policy Review 34*, 1 (1 2016), 135–174.

12. Isikdag, U. BIM and IoT: A Synopsis from GIS Perspective.

13. Ismail, A., Nahar, A., and Scherer, R. Application of graph databases and graph theory concepts for advanced analysing of BIM models based on IFC standard. Tech. rep.

14. Ismail, A., Strug, B., and Ślusarczyk, G. Building knowledge extraction from BIM/IFC data for analysis in graph databases. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10842 LNAI, Springer Verlag (2018), 652–664.

15. Jayathissa, P., Quintana, M., Abdelrahman, M., and Miller, C. Indoor Comfort Personalities: Scalable Occupant Preference Capture Using Micro Ecological Momentary Assessments. *preprint* (2020).

16. Mahdavi, A., and Taheri, M. An ontology for building monitoring. *Journal of Building Performance Simulation 10*, 5-6 (2017), 499–508.

17. Pärn, E., Edwards, D., and Sing, M. The building information modelling trajectory in facilities management: A review. *Automation in Construction 75* (3 2017), 45–55.

18. Perozzi, B., Al-Rfou, R., and Skiena, S. DeepWalk: Online Learning of Social Representations.

19. Sagiroglu, S., and Sinanc, D. Big data: A review. In *Proceedings of the 2013 International Conference on Collaboration Technologies and Systems, CTS 2013* (2013), 42–47.

20. Talari, S., Shafie-Khah, M., Siano, P., Loia, V., Tommasetti, A., Catalão, J. P. S., and Tah, J. H. M. A Review of Smart Cities Based on the Internet of Things Concept.

21. Tang, M., Nie, F., and Jain, R. Capped p-Norm graph embedding for photo clustering. In *MM 2016 - Proceedings of the 2016 ACM Multimedia Conference*, Association for Computing Machinery, Inc (10 2016), 431–435.

22. West, D. *Introduction to graph theory.* 1996.

23. Xu, D., Cheng, W., Luo, D., Liu, X., and Zhang, X. Spatio-Temporal Attentive RNN for Node Classification in Temporal Attributed Graphs. Tech. rep., 2019.

24. Yang, Z., Cohen, W. W., and Salakhutdinov, R. Revisiting semi-supervised learning with graph embeddings. In *33rd International Conference on Machine Learning, ICML 2016*, vol. 1, International Machine Learning Society (IMLS) (2016), 86–94.

25. Zhang, H., Shang, X., Luan, H., Wang, M., and Chua, T. S. Learning from collective intelligence: Feature learning using social images and tags. *ACM Transactions on Multimedia Computing, Communications and Applications 13*, 1 (11 2016).

26. Zhang, Y., Xiong, Y., Kong, X., and Zhu, Y. Learning Node Embeddings in Interaction Graphs.

27. Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M. Graph Neural Networks: A Review of Methods and Applications. Tech. rep.

VR & AR

Reviewing Building Design Modeling in Virtual Reality

Will Wang[1]

[1]Pelli Clarke Pelli Architects
New Haven, CT, USA
will.wang6@gmail.com

ABSTRACT

New virtual reality software geared towards the architecture, engineering and construction industry are eyeing a larger market share amongst other digital tools. The niche for design architects, however, remains stagnantly insignificant. The prevailing trend in virtual reality development is towards coordination, review, or project management. True design tools with the capability to create and manipulate geometric forms with a fine level of dimensional control or topological precision are nearly non-existent. This paper describes the challenges faced in a virtual reality tool made to integrate directly with a dedicated 3d modeling software, explores the potential alternatives to the current user interface by a method of replicating design tasks in said tool environment, and offers conceptual interface wireframes upon which future virtual reality design tools can be built.

Author Keywords

Virtual reality, User interface, CAD, Digital modeling

ACM Classification Keywords

Human-centered computing ~ User interface design

1 CONTEXT

Virtual Reality (VR) technology has ballooned in the past decade thanks to the hardware advancement and multiplied computing power. It catches the attention of architects perhaps more than others, because architects are particularly interested in simulating their designs before it could be constructed. In comparison to many design fields, architecture is constrained by the scope of its prototypes and mock-ups. It is a long-held belief that VR may grant architects an opportunity to experience the design in full effect before the "product" can be realized.

However, it is also a popular opinion that the practice of architecture is at odds with technology integration. While many industries advances with technology [11], architecture is slow to adapt. "Sophisticated tools for modeling and representing buildings already exist. How we use those tools, and how they dictate what the outcome will be, is a design challenge." [6] As the profession becomes increasingly "...dismayed by the elusiveness of success…[and] diminished impact" [4] a renewed attempt at marrying it with the ever-expanding selection of technologies is due. VR is a fit next candidate as it simmers and brews into maturity within the technology sphere.

A previous major success with technology in architecture was the migration from analog drafting to Computer-Aided Design (CAD) software. When computers replaced traditional drawing utensils, newer forms were unlocked in addition to production boosts. VR again teases a new mode of designing that may very well "afford" [9] an Architecture the world has never witnessed before.

Although VR technology has been heralded as enabling architects to have phenomenal understanding of the building forms in an immersive environment [3, 1], this "higher situation awareness" [7] depends on having real-life sizes of objects and context, which can only be modeled from numerical description of rather accurate dimensions. Current VR authoring tools often exploit the mirroring of analog motions, resulting in a more free-form sketching. These free-forms often lack the rigor of a scaled model. Meanwhile most VR software struggle with creating an intuitive interface that can reproduce the CAD drafting experience.

Given the broad nature of architects' work as generalists, the profession is always compelled to look for tools that help "...reserve mental resources to the core of the task at hand" [5]. If the interface is clumsy and difficult, it quickly turns even the first adopters away. The exhausted designer will not spend more time with the tool.

The core of the challenge is that "[e]xisting 3D CAD modeling tools are designed for a WIMP based point-and-click input" [8], which is windows, icons, menus and pointers. Modern architects are trained to manipulate geometries in the workflow optimized for the WIMP model. When VR is introduced, "legacy bias" will play a huge role. A user will carry over old habits from old technology in interacting with new technology, rendering new technology difficult to use even if it is objectively more user-friendly. [8]

Some prior studies have been conducted to investigate the user interface (UI) of the VR environment. They in fact serve as evidence that VR technology has NOT been tailored to the use of architectural designers.

For instance, AIR-MODELLING [2] was a focused study on capitalizing on the richness of gestures in a virtual setting. The study laid out the limits that "[m]ost works merging [VR] with CAD have centered their attention on the latest stages of product design, such as assembly modelling and visualization, staff training and maintenance". Even though the research set out to break free from those, the work acknowledges the lack of surface or volume modelling, ambiguous translation to 3D from 2D sketches, low fidelity or accuracy in shape-making exhibited by VR modeling interfaces.

The same void in modeling interaction traces back to an earlier study (Seichter, 2003) where the creation of shapes was only briefly described in a box extrusion case. Emphasis was put on having simultaneous viewers of a design shape. Another research (Marchand et al, 2018) confirms that the use of VR technology in design concentrates on viewing/reviewing ideas rather than the authoring of formal outcomes thereof.

Voice control may offer a novel angle of approach. Although having "impactful" [13] results, interaction with voice can be difficult to implement in the typical architect office. Design modeling is a mostly silent task. Without changing the infrastructure - in effect office layouts and worker proximity - voice control in VR seems to be beyond immediate practicality.

In order to investigate the most applicable, quickest method of adoption, this paper looks at using commercially available hardware, which is controller enabled, as opposed to fully gesture-based. As modern CAD software often include a "dense set of toolbars and menus" [12], the goal is *not* to make an intuitive modeling environment anymore. The challenge *is* to map old WIMP interface into a new, immersive one so designers utilize the already learnt. "By relating a new interface technique to something with which they are already familiar, the user starts with a grasp of what to do." [18] This approach likely avoids the experience of "negative transfer of training" [14].

Beyond a relative barrier-free UI, there will be tacit knowledge in using the VR tools. Newcomers will encounter a learning curve in adapting their motions to the behavior of the virtual modeler. After some practice, these motions will become trained and "feel" natural to the user.

2   METHODOLOGY

To understand how a VR environment can recreate the design process in the WIMP model, the author first finds a project workflow that represents, to the largest degree, the common workflow of a design practice. The author documents each interaction and attempts to execute them in the VR environment. Each "road block" in the VR environment will prompt a wireframe of an alternative interface that addresses the problem. This wireframe serves as a broad framework within which novel VR CAD

modeling software can be created. This framework extends to augmented reality UI as well.

2.1 The Workflow

The example project is a hypothetical high-rise office building, which is a common and repeated building type. The critical design task associated is the shaping of the exterior envelope along with managing a working central core, colloquially the "core-n-shell" design. The author assumes a 3D context digital model has been provided.

The first step is to create simple massing that helps gauge the scope of building area. The city block boundary provides a rectangular footprint within which the building can occupy. A vertical "extrusion" of this footprint provides the volume from which a quantity take-off can be done on gross floor areas. A desire in the building shape is to create an Entasis so a degree of tapering must be achieved. To have a smooth reading of the facade, the four corners of the building is filleted. A further articulation is made so that on each elevation, the left side of the facade is pushed inwards so there is a shadow line running from bottom to top of tower. The basic shape to reproduce is diagrammed in Figure 1.
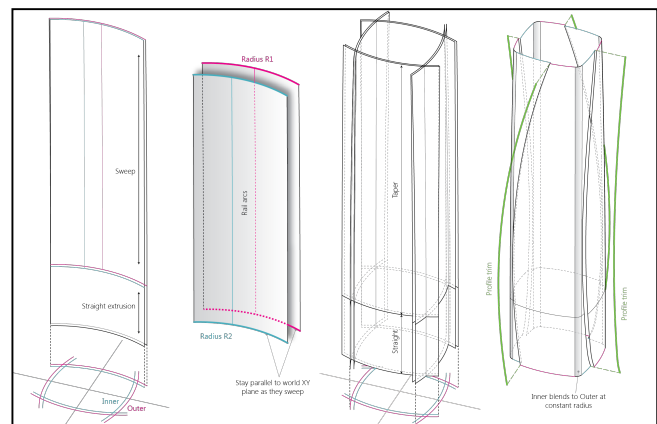
Figure 1. Case study shape generation

Second step is to manage a core design that works for the tower. High-rise cores most likely contain fire stairs, elevators, vertical chutes, mechanical spaces, and restrooms. Each component has a rather strict set of criteria and consequently can be designed through organizing given templates. For example, elevators are often in banks of 6-8 cars and the sizing of fire stairs is directly calculated from occupancy. In terms of shape creation, core design is primarily rectangular elements and straight, extruded forms. Most input actions will be drawing outlines of shapes that will be turned into straightforward 3d objects. The operations concern less with scale or proportions. Rather they are 2d drafting exercises. The study assumes a working core arrangement is provided and used as basis for simple 3D "extrusions" only in VR.

The next step is to turn exterior shape into sensible building cladding. Panelization often involves scripting tools. Their interface can be entirely different from that of traditional

CAD. However, emerging visual scripting (VS) tools bring WIMP model back into the game. VS tools boils down functions into discrete "black boxes" that perform tasks in the same way CAD modeling commands affect model geometry. If the VR tool can successfully capture interactions of WIMP, it can interface with VS tools just as well. Because of the VS tool's often flat "canvas", most manipulations of the script are similar to that of managing core layouts described above.

The last step in designing the high-rise is to visualize the scheme. Photorealistic rendering of architecture is rarely done in real-time. Vast majority of modeling software isolates the process to optimize flow and performance in geometry authoring. When a certain design intent is modeled, geometries are piped to a dedicated engine, whether embedded or stand-alone, to actively process the visualization. There are often few interactions within the 3d modeler other than clicking through steps of exporting objects. Controls of this step is minimal, and can be accommodated by the point-and-click actions in a stock VR UI.

This study intentionally leaves out the documentation of design schemes, which has more to do with designing 2D graphics and adhering to conventions rather than spatial exploration. Architectural drawings are either drafted in the 2D space or extrapolated from a 3D model. As modern documentation workflow matures, fewer professionals are actively designing through "flat media" drawing. Even when they do, the benefits of an immersive environment diminishes. VR experience is rooted in the 3D space.

2.2 The Virtual Environment

The study has collaborated with Mindesk VR to reproduce the design scheme in its environment. Mindesk VR is arguably the only tool in active development that targets direct integration with Rhinoceros, this paper's chosen CAD environment prevalent among architects due to its robust geometric manipulation and analyses. Other VR tools may have better developed UI such as Gravity Sketch or Arkio, both of which the author has experienced first-hand, but they serve purposes rather outside of the rigor associated with architectural modeling in CAD software.

There are a few vendors of VR hardware. Vive and Oculus are two that have more sophisticated controllers, affording more interactions. They share some major features. Both have an omni-direction input. On Oculus it's a thumbstick and on Vive it's a touchpad. This omni-direction input on either controller is click-enabled. Both controllers have a front trigger where the index finger tends to naturally rest, as well as a squeeze button on the side of the controller where either the middle or the ring finger often rests. Oculus controllers have three other buttons on each hand while the Vive ones have two. The study is conducted on the Oculus headset.

3   RESULT

Table 1 shows all actions taken to arrive at the desired design form and its current accessibility in the VR tool under study. Numbering only implies order of first appearance in the process.

Actions	Done in VR
1 – Offset curve by distance	No
2 – Draw arcs	No
3 – Mirror objects	No
4 – Copy/move objects	Yes
5 – Extrude curves	Yes
6 – Sweep forms	No
7 – Extend curves	No
8 – Fillet surface	No
9 – Split curves/surfaces	No
10 – Object layer organization	No
11 – Snap to axis	Yes
12 – Extract sub-objects	No
13 – Join curves/surfaces	No
14 – Cap open surfaces	No
15 – Interface with visual scripting tool	Yes

Table 1. VR command mirroring

It becomes apparent that most modeling commands needed for the study are absent in the VR tool. The current toolset employs an expanded tray triggered by left thumbstick press to house a handful of commands. This tray can only be of a certain size before it falls out of the field of view. Therefore, the number of commands accommodated is very limited. There are over 1800 actions in Rhinoceros and it is impractical to array all of them in a tray display. A more approachable method is to index them

The study proposes to use the thumbstick's ability to rotate as a way to scroll through large number of indices. Commands are organized alphabetically and indexed by first letter. Thumbstick rotation is smooth but a haptic feedback can be implemented to indicate clicks. For example, a user may rotate the thumbstick within an angular domain, staying within the same letter index. A haptic "shake" happens when the user rotates past 20 degrees entering the next letter index. There are 26 letters and an option for commands starting with a numeric value. Each angular "segment" on the thrumstick's rotary domain, therefore, occupies about 360/27=13.3 degrees. As a user rotates the thumbstick, a display of the current letter selection should be produced near the controller as an indicator. The user can also tilt the thumbstick directly to

the letter domain if s/he remembers where it is for faster access instead of always spinning from beginning (top). See Figure 2-1.
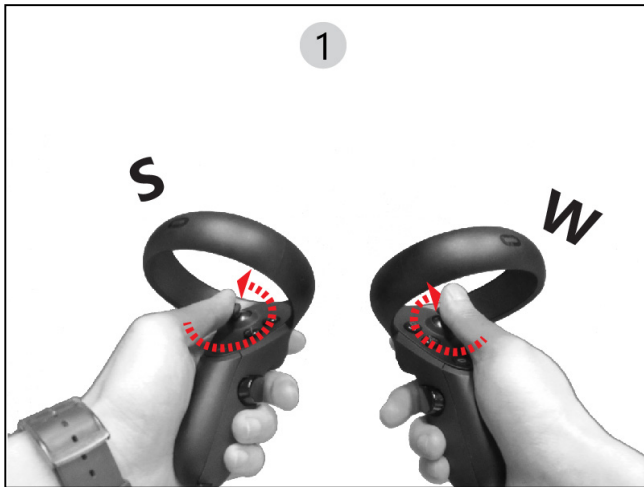
Figure 2-1. Thumbstick indexing

As soon as a letter is selected, the user can still be faced with a large number of commands. The second thumbstick can be programmed to index the second letter of the command. For example, if one wants to create a sweep surface, s/he can spin the left thumbstick to letter "s", receiving a long list of potential actions such as "split", "smooth", or "save". With a second thumbstick spun to "w", the list length shrinks quickly to single digit. This shortened list of "candidate" commands can now be placed in visual tray for the user to pick the desired modeling action. See Figure 2-2.
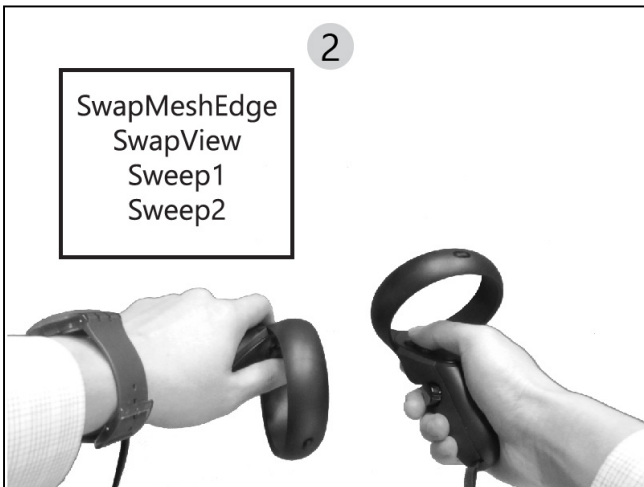
Figure 2-2. Docked result commands

The filtered few commands can also be docked around the virtual hand. There are precedent VR applications where some clickable items are arrayed around where the wrist is. The user can lift up an arm as if looking at a wristwatch to interact with options. This is a great way to expand the limited real estate on VR controllers and should be

implemented widely. Options can be made so that a user can save a few common commands around the virtual wrist for fast access.
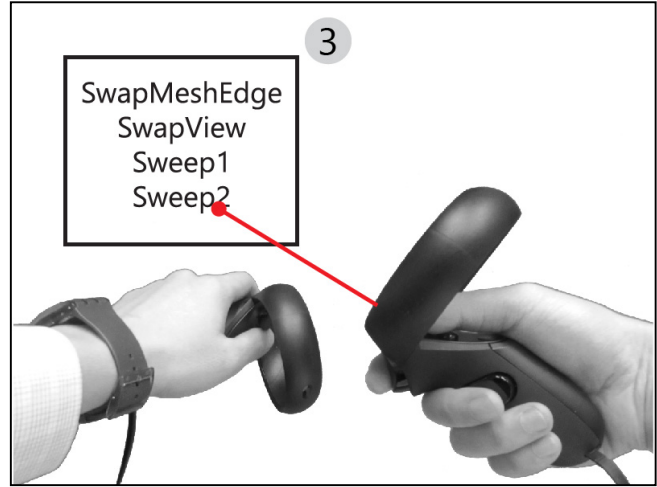
Figure 2-3. Selecting desired command

Complex CAD commands often include actionable options post-call. For instance, the process of rebuilding a surface with prescribed domain and fitting requires additional parameters after the command is triggered. Conventionally (as in WIMP), these additional inputs can be interfaced with different key presses in the command prompt. In the VR environment, these options should be mirrored. A dashboard that includes the parameters near the hand can be displayed, waiting for user input. See Figure 2-3 and Figure 2-4.
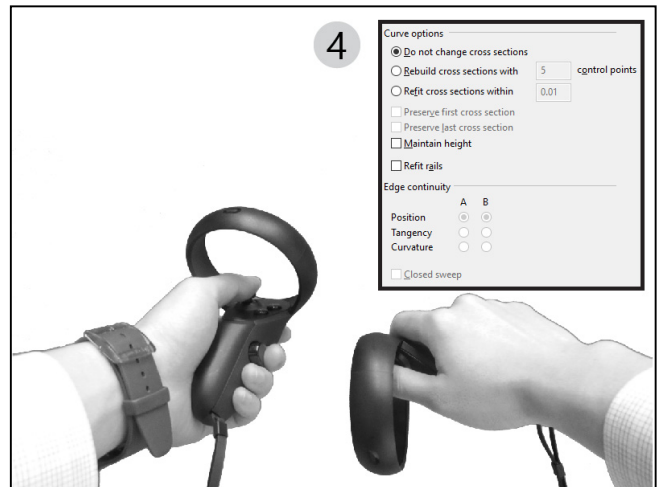
Figure 2-4. In-command parameter interface

As of this writing, MindeskVR has not attached any functions to the lower buttons on each hand controller (labeled "X" and "A" on Oculus). They can be used as the "shift" or "ctrl" keys. The "alt" key has been accounted for with the left trigger on the controller in the context of the study-specific 3D modeler. These function keys become helpful for toggling different behaviors during a modeling

action such as snapping to cardinal direction axes or multiple select/de-select.

Other interactions related to modeling commands rely on the computer mouse. Movement in VR can simulate mouse actions rather easily. In fact, Mindesk VR already implements the many object snapping of the CAD cursor in its environment. However, a variable pointer-tip is recommended to expedite object selection process.

WIMP interactions use collapsed perspective views of the 3D model so a cursor may be hovering over multiple objects in distinct 3D space but coincidental on a 2d monitor. A user can usually scroll through ambiguous selections and pick the object desired in this situation. With VR, the user must truly move controllers to the actual position in space of the target object. This creates difficulty in selecting tiny or complex parts nestled together.

Variable pointer-tip can answer to this and it has already been implemented in tools like Gravity Sketch. It is a resizable orb near the hand that selects anything intersecting it, and is resized via the spinning of the thumbstick. In this study, as thumbstick motions alone are reserved for command indexing, user can hold one of the function keys ("shift", "ctrl", or "alt") as s/he spins thumbstick to activate pointer-tip altering.

4 DISCUSSION

The proposed UI fills a gap in the VR software landscape. The technical achievement of representing 3D objects in a VR headset serves only as step one in transforming the ways of architectural designers. The participation of VR technology in design modeling remains disjointed, albeit valuable, if the use of goggles is limited only to the end process of the workflow.

It has been confirmed with the creators of MindeskVR that the software currently targets design review, which is evident in its void of comprehensive command-mirroring. This leaves the study at a speculative stage. Given the findings as conceptual wireframes in the study, the author acknowledges that a learning curve will be inevitable. The index system may be intuitive to use but a designer may still need hours of practice to acquire tacit knowledge of the tool UI. The strength of this interface and the integration provided by an app such as MindeskVR nevertheless remains that the designer can follow previous habits of modeling actions.

Although some apps may provide a virtual keyboard that the user can punch in individual letters, offering an even more straightforward mirroring of the WIMP model in VR UI, the author believes that it takes too long of a forearm motion to achieve a single letter input. It is an inefficient solution when commands are executed often. The proposed UI takes advantage of the smaller movements of the thumbsticks and key presses to index commands. It will be slower than direct search method as in a command prompt of a 3D modeler, but with practice, a user is likely to build up muscle memory of the command locations, hence a faster action sequence.

Aside from the interface, a few non-trivial constraints are observed. Compared to sit-down sessions modeling with keyboard and mouse, movement-based modeling introduces a level of fatigue. Vertigo is another problem that is caused by vastly different physiological conditions of individuals. These could be hurdles rather difficult to overcome. With 3D modeling, each design exercise can run easily over an hour. During the study, VR sessions were often shorter than 20 minutes and yet some level of discomfort was experienced. However, improving hardware is believed to ameliorate VR sickness, as enhanced picture and frame rates reduce the causes that are "conflicts between the sensory information... and vestibular systems" [17]

The paper approaches the challenges of VR in architecture from a specific angle in order to produce a viable framework quickly. The process of modeling the hypothetical massing may not be completely representative of the entirety of an architect's 3D modeling experience. More studies of similar breadth or deeper dives into generalization should be conducted to construct an even more generic guideline for VR UI design to follow in the future.

5 CONCLUSION

Whether or not the modeling of buildings will one day migrate fully to the VR environment is up to the test of time. However, given the infancy of current technology and more importantly its user experience, much is to be done before an architect can intuitively construct virtual designs while immersed.

There is still significant roadblocks to integrating VR technology in the workflow of building designs, not because of the lack of technical capability but a cumbersome user experience, which is largely a UI problem. Sparse interest from software developers due to small market size or low economy of scale has likely led to the lack of rigorous 3D modeling VR software. Businesses oriented towards the building industry are more inclined to develop technology for actual construction and overall project management, both of which center around review and coordination rather than precision-based modeling.

The solution to that is to have architects more engaged in advancing technology. Architects need to be part of the making of their own tools. Although they may lack the means to fully code up working prototype, they should explore venues of collaboration with software engineers to investigate how emergent technologies can be substantially re-tooled for their own purposes. A "harmonious balance" [15] can be achieved in situating technology in this creative field. The interface study described in this paper kick-starts the process much needed to truly bring VR into the everyday architectural design.

REFERENCES

1. Achten, Henri, Walther Roelen, Jan-Thijs Boekholt, Arthur Turksma and Joran Jessurun, Virtual Reality in the Design Studio: the Eindhoven Perspective, *eCAADe conference 1999 proceedings,* Liverpool, UK, p169-177.

2. Arroyave-Tobon, Santiago, Gilberto Osotio-Gomez, and Juan F. Cardona-McCormick, AIR-MODELLING: a Tool for Gesture-based Solid Modelling in Context during Early Design Stages in AR Environments. *Computers in Industry*, 66, 2015, p73-81.

3. Bartosh, Amber and Laura Clark. Mixed Reality Visualizations of Urban Data, *Technology | Architecture + Design*, 2019, 3:1, p89-101.

4. Celanto, David. Innovate or Perish: New Technologies and Architecture's Future, http://www.harvarddesignmagazine.org/issues/26/innovate-or-perish-new-technologies-and-architectures-future, (2007, accessed Sep 4 2018).

5. Coomans, M.K.D. and Harry Timmermans, Towards a Taxonomy of Virtual Reality User Interfaces, *proceedings of IEEE Conference on Information Visualization*, 1997, p279-284.

6. Deutsch, Randy. Convergence: The Redesign of Design, West Sussex: John Wiley & Sons Ltd. 2017.

7. Donalek, Ciro, S G Djorgovski, Alex Cioc, Anwell Wang, Jerry Zhang, Elizabeth Lawler, Stacy Yeh, Ashish Mahabal, Matthew Graham and Andrew Drake, Immersive and Collaborative Data Visualization Using Virtual Reality Platforms, *IEEE International Conference on Big Data proceedings 2014*, Washington, DC, p609-614.

8. Kahn, Sumbul, Bige Tuncer, Ramanathan Subramanian, and Lucienne Blessing, 3D CAD Modeling Using Gestures and Speech: Investigating CAD Legacy and Non-legacy Procedures, *CAAD Futures 2019 proceedings*, Daejeon, South Korea, p624-643.

9. Leach, Neil. There Is No Such Thing as Digital Design, *Paradigms in Computing*, Dr. David Jason Gerber and Mariana Ibanez (e.d), Los Angeles: eVolo Press, 2014, p149-158.

10. Marchand, Emmanuel Beaudry, Tomas Dorta, and Davide Pierini, Influence of Immersive Contextual Environments on Collaborative Ideation Cognition - Through Design Conversations, Gestures, and Sketches, *proceedings of 36th eCAADe*, vol 2, 2018, p795-804.

11. McKinsey Global Institute. Digital America: A Tale of the Haves and Have-mores, https://www.mckinsey.com/industries/high-tech/our-insights/digital-america-a-tale-of-the-haves-and-have-mores, (2015, accessed July 3 2019).

12. Nanjundaswamy, Vishwas & Kulkarni, Amit & Chen, Zhuo & Jaiswal, Prakhar & Shankar, Sree & Verma, Anoop & Rai, Rahul. Intuitive 3D Computer-Aided Design (CAD) System With Multimodal Interfaces. *Proceedings of the ASME Design Engineering Technical Conference*, 2013, Vol. 2

13. Osking, Hunter and John A. Doucette, Enhancing Emotional Effectiveness of Virtual-Reality Experiences with Voice Control Interfaces, *International Conference on Immersive Learning*, Springer, Switzerland, 2019, p199-209.

14. Pausch, Randy, Denni Proffitt and George Williams, Quantifying Immersion in Virtual Reality, *proceedings of SIGGRAPH 1997*, p13-18.

15. Roudsari, Mostapha and Chris Mackey, The Tool(s) Versus the Toolkit, *Humanizing Digital Reality*, edited by K. De Rycke et al. Springer Nature Singapore Pte Ltd, 2018, pp 93-101.

16. Seichter, Hartmut, Augmented Reality Aided Design, *International Journal of Architectural Computing*, issue 1, vol 4, 2003, p449-460.

17. Ruddle, Roy A, The Effect of Environment Characteristics and User Interaction on Levels of Virtual Environment Sickness, *IEEE Virtual Reality*, 2004, p141-148.

18. Sherman, William R. and Alan B. Craig, Understanding Virtual Reality: Interface, Application, and Design, San Francisco: Morgan Kaufmann Publishers, 2003, p284

An Interactive and Responsive Virtual Reality Environment for Participatory Urban Planning

Helmut Schrom-Feiertag[1], **Martin Stubenschrott**[1], **Georg Regal**[1], **Thomas Matyus**[1], **and Stefan Seer**[1]

[1]AIT Austrian Institute of Technology, Vienna, Austria, helmut.schrom-feiertag@ait.ac.at

ABSTRACT

Urban simulation using Virtual Reality can provide an interactive and responsive environment for participatory planning to evaluate alternative designs, regulations and environmental policies. The objectives therefore are that the virtual world is perceived as real and enables people to feel connected to the world and to bring VR technology to all citizens so that they can help to shape the city in a participating way. Therefore, we present a virtual reality environment with multi-modal traffic simulation for participatory planning of urban areas. Using a gamified approach with story-telling enables to engage and guide participants through the environment accompanied by virtual questionnaires for in-situ feedback. This paper demonstrates the feasibility of virtual reality environments to perform participation via immersion and interaction for gathering immediate and authentic feedback guided with a systematic approach to simplify analysis. We describe the technology, developments and theories behind virtual environments and simulation. Finally, we provide insights from our evaluation of the proposed urban simulation used in two real world projects. The combination of highly immersive and interactive virtual environments with traffic simulation was appreciated by the participants and resulted in highly valuable feedback for professional planners.

Author Keywords

Urban planning; virtual reality; traffic simulation; participation; human computer interaction;

ACM Classification Keywords

I.6.3 Applications; H.5.1 Multimedia Information Systems: Artificial, augmented, and virtual realities; H.1.2 User/Machine Systems: Human factors; J.5 ARTS AND HUMANITIES: Architecture

1 INTRODUCTION

Today, urban development is no longer conceivable without transparent information and active involvement of the affected population. Participation can create acceptance and if used systematically and at an early stage it can lead to better results for more sustainable planning: Citizens and actors in urban development have undisputed expertise in all issues that affect their direct living and working environment. This everyday knowledge may represent only a part of reality, but at the same time it is detailed and founded on experience and can represent an important input for decision-making within participation processes.

Improving the participation for a diversity of users is a viable strategy to ensure and intensify early participation of the public in urban planning and development processes and helps to raise acceptance by involving all citizens. As local governments grow more and more interested in civic participation, it becomes important to explore available methodologies addressing challenges related with participatory processes. The participation of citizens in the design of public space is fundamental. Today's city challenges at the urban block scale can be more effectively addressed through participation and the collection of experiences from people who live, work, and play in that space. This shifts the emphasis from top-down design to human-centred problem solving [13].

Virtual Reality (VR) offers the possibility to walk through virtual worlds delivering a comprehensive computer-generated atmosphere that enables a contextual preview of a life-like setting. VR also provides a high level of immersion and presence, and with embedded VR questionnaires and additional feedback channels new forms of citizen participation are possible. This will completely change participation: Citizens no longer look passively via a screen at digital content but become part of the virtual world which is perceived as real and enables people to interact with and feel connected to the world. In combination with traffic simulation, virtual reality environments (VRE) furthermore enable the evaluation of newly planned road environments, whether they are perceived as safe and pleasant or whether additional measures are required.

To unfold the full potential of VR technologies, we developed an integrated application for participatory planning of public spaces that uses a combination of VREs with traffic simulation to allow realistic experiences of traffic situations in public space.

This led to the research question: Is the developed application suited for participatory urban planning? A study was therefore carried out to examine the suitability and added value of the application in two urban planning projects.

The following Section 2 gives an overview of related work in the field of VR, traffic simulation and participation. In Section 3, the overall concept and methodology used for VR and traffic simulation are described. Then the developed interaction methods to best support citizens by the usage of the VRE for participation are explained in Section 4. The demonstrator, participants and results of user feedback are presented in Section 5. Finally in Section 6 the findings are discussed and conclusions are drawn.

2 RELATED WORK

The application of VR in planning processes has been investigated internationally in many projects and studies [5, 10, 11], some studies have dealt with the application of VR in real planning projects (e.g. [16]). VR was examined in the context of evaluation procedures for architectural competitions [17], for the mediation of planning scenarios in traffic projects [3], for interdisciplinary planning in the field of urban mobility [9, 13], for the investigation of scenarios in urban cycling or to research children's traffic behaviour [2]. These studies show the high potential of VR to make planning more accessible to various stakeholders and to increase the social and ecological sustainability of urban planning [16].

In [13] a VR game with interactive elements was developed to shape the urban public space for a project that aims to create a large public space that prioritise the people instead of vehicles and to design to the needs of the users: spacious, pleasant spaces full of vegetation with dynamic uses. The objective was to bring VR technology to all citizens so that they can help to shape the city in a participating way.

There have been some attempts to couple traffic simulations with high-quality VR-capable 3D game engines. PTV Vissim 10 has added support for displaying its network including simulated vehicles [19]. Synchronised co-simulation of SUMO and the Unity game engine has been done by [18] for a vehicle-in-the-loop test environment for autonomous driving with microscopic traffic simulation. In [14] a virtual reality cycling simulator has been developed.

These solutions allow manually controlling one vehicle. The focus of this paper, however, is to create a virtual environment for human-centric urban planning with the point-of-view of a pedestrian.

3 URBAN SIMULATION

Performing studies in a virtual environment requires realistic simulation of the relevant objects in order to obtain a good correlation between the virtual environment and reality [8]. In the context of urban planning which is the focus of this paper, a lively city with realistic simulation of traffic and pedestrians is key for an immersive urban environment and requires the following steps (see Figure 1):

1. The *representation of the urban environment* is created both as a 3D model for VR and as a 2D representation for the traffic simulation. The VR application was implemented with the widely used game engine Unity[1].

---

[1]https://unity.com/

2. *Multi-modal traffic simulation* of the scenario including pedestrians, cars and cyclists is done using an external simulation program at real-time.

3. *Synchronisation* of the Unity game engine with the traffic simulation provides real-time interaction and a responsive environment.

3.1 Virtual Environment

The representation of the environment in VR requires a 3D model of the geometry. With the increased application of BIM and GIS, 3D models become more an integral part of modern urban planning [1] allowing their direct usage for our purpose. The 3D model is imported in the Unity game engine and a 2D representation is derived from it which is used as the base layer for the traffic simulation which must currently be configured in a separate program. It is essential to use exactly the same geometric set-up in the game engine Unity and the traffic simulation software. Inconsistencies can lead to unrealistic behaviour like agents walking through obstacles or cars driving through the user, thus detracting from the users' experience.

Origin-destination (OD-) matrices are added to define the traffic demand of vehicles, cyclists and pedestrians. Furthermore, the different types of agents used in the simulation (pedestrians, cyclists and cars) need the definition of additional information which is described in the next section.

The combination of VR with dynamic traffic simulation enables the mediation of different designs. Different perspectives can be explored and through traffic simulation different situations can be evaluated, therefore achieving a better understanding of planning variants and a higher acceptance for a fair distribution of space between all road users.

3.2 Simulation of Road Users

For the realistic representation of road users, microscopic approaches represent the current state of research (see [4]). Microscopic models represent each individual road user and describe the actions depending on the infrastructure and the movements of other road users. The interactions between different types of road users (pedestrians, cyclists and vehicles) can be modelled in today's traffic simulations but is usually limited to spatially discrete locations such as intersections, cycle paths and protective paths. Innovative street design concepts for active forms of mobility are increasingly based on mixed traffic arrangements. Since the traffic flow in mixed traffic arrangements does not provide for the spatial and temporal separation of the different modes, new approaches for the realistic representation and experience of road users in VR have to be integrated.

We implemented an approach which uses social forces to describe the movements of the pedestrians and their interactions (see [6]). Pedestrians can manoeuvre freely on the two-dimensional plane, whereas vehicles follow predefined paths, obeying the rules of a queuing model. Depending on the use case, cyclists can be modelled as agents moving either in the two-dimensional space, similar to pedestrians or in the one-dimensional space, similar to multi-track vehicles.

Step 1 – 3D Environment   Step 2 – Simulation Scenario   Step 3 – VR Simulation
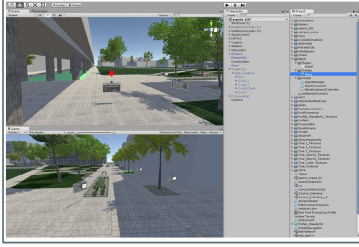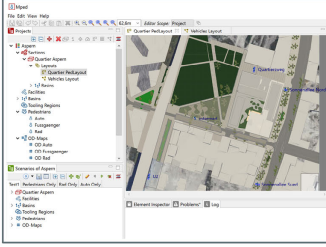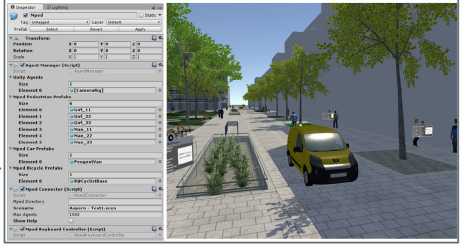
Figure 1. Overall concept of the urban simulation showing the workflow with step 1) creation of the 3D environment and 2D ground plan, step 2) setup of the traffic simulation scenario and step 3) running the VR environment in Unity synchronised with the traffic simulation.

Road users, pedestrians, cyclists and vehicles are detecting each other and adapt their movements in order to avoid collisions. Of course, the participant is part of the simulated world and all other road users react to him, as well.

3.3 Synchronisation of VR and Traffic Simulation

Creating an immersive virtual environment requires two-way interaction between simulated agents in the traffic simulation and the visual representation in the VR. Since the existing traffic simulation software is run as a separate process independent from Unity, a special *Connector* was developed as a Unity asset (see Figure 2). The connector is responsible for starting, stopping and updating the traffic simulation transparently and keeps Unity in sync. This includes:

- Time synchronisation ensures that the traffic simulation runs at the same speed as the Unity simulation.

- The camera position (which equals the position of the person in the VR) is continuously sent to the traffic simulation. Simulated pedestrians can react accordingly and deviate from their path and vehicles stop instead of running over the participant.

- The position and orientation of all agents in the traffic simulation is sent to Unity. If a new agent is created in the simulation, a new instance out of a set of predefined prefabs is created. A prefab represents a game object with all its components, property values and required child-objects. Likewise agents in the virtual environment are removed when needed.

- Animations (moving hands and feet, turning wheels) are synchronised with the agents' velocities.

Figure 3 shows the traffic simulation and the representation of it in the VR environment side-by-side. Note that the 2D representation of the simulation is usually hidden but only shown here for illustration purposes.

4 PARTICIPATION

VR has the potential to provide experiences and deliver outcomes that cannot otherwise be achieved. However, interacting with immersive applications is not always easy and it is not just about providing an interface for the user to perform his tasks. It is about supporting users in an intuitive way to create a positive experience and not frustration ([7]).
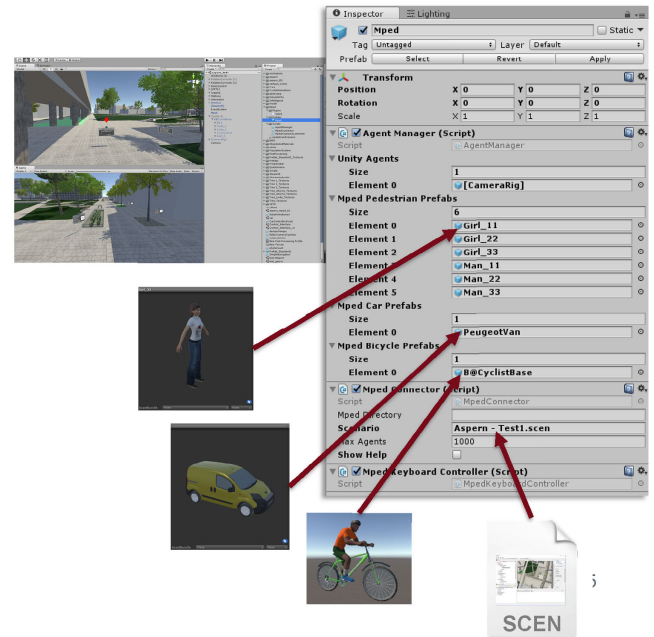
Figure 2. Set up of the traffic simulation asset in Unity with multi-modal agents.
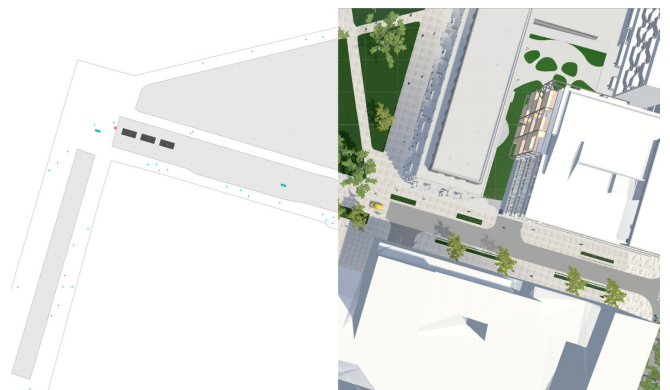
Figure 3. Top view of the environment in the simulation framework in 2D (left) and in VR in 3D (right)

115

Figure 4. Birds-eye view.

Figure 5. Child perspective in the VRE.
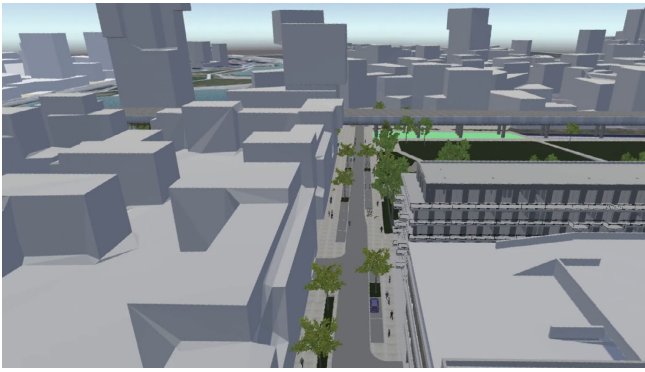
Thus, we had to intuitively communicate to users how the virtual world and the interaction works to support participation through VR in an elegant and comfortable manner for a wide audience of non-experts.

Through the combination of interactive VRE and multi-modal traffic simulation, street designs were made tangible and new opportunities for citizen participation were created. To create a highly interactive VRE we developed and integrated intuitive information visualisation, navigation and interaction as well as the possibility to collect feedback directly in VR.

A great challenge in participatory planning is to include persons with various backgrounds and diverse knowledge in the process ([15]). Therefore, we developed appropriate interaction methods for the use of VR to support such diverse backgrounds and to efficiently gather important user feedback. For navigation in the virtual world, real walking within a small area and teleportation across larger areas was implemented. Special attention was paid to the function for changing one's own body size, as this effect could be used for two important aspects of planning. On the one hand an enlargement of the own (virtual) body (Figure 4) is a good method to get an overview of larger areas. Also distances become shorter, so that users only have to take a small step in VR to walk a whole block. By shrinking the virtual body of the user to the size of a child, one can take the perspective of a child (Figure 5) and better experience the viewpoint of children.

For object manipulation, lifting, moving, re-positioning and the creation of objects (trees and benches) were realised. Virtual questionnaires and a virtual photo camera in VR were implemented as feedback possibilities for the users.

In first user evaluations we could observe that without guidance users often wandered to parts of the 3D model that were not interesting in terms of user participation or otherwise distracted by trying to interact with objects that were not in focus of the participation, e.g. trying to open car doors. Also if all interaction possibilities were presented at once most users were overwhelmed by the possibilities.

Thus, we identified the need to provide a story and gamification elements to guide and engage non-professionals in the design of public space overcoming the limitations of conven-
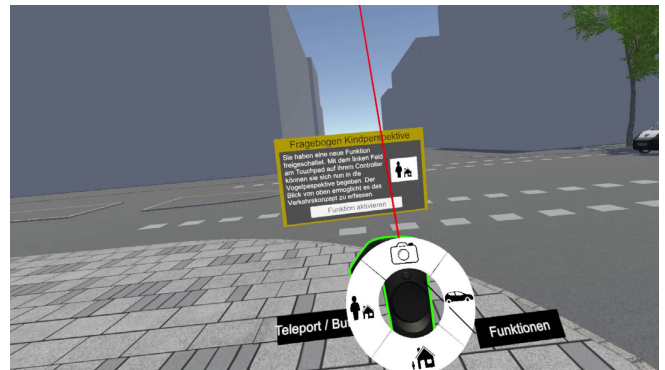
Figure 6. Controller with enabled features shown in the VRE.

tional participation methods. We implemented certain gamification elements to guide the user. Users could unlock one feature (e.g. take pictures, switch to bird's-eye view, etc. cf. Figure 6) after the other at certain points along the desired path, marked with a red arrow. To unlock a new feature the user had to answer a questionnaire embedded inside the VR (Figure 7). To implement the questionnaires we used the VRate framework of [12], a Unity3D asset that supports the integration of subjective experience questionnaires directly in VR.

By implementing these gamification elements, we could ensure that the user walks the right path and experiences a similar 3D environment. By gradually unlocking features the user is not overwhelmed by the range of features available at once. Moreover, designing these features as gamification elements engages the user to answer all questionnaires which supports to retrieve valuable user feedback in the participation process.

In order to provide information to the users, there were interactive info boxes with explanations of the design and function for the different architectural features. These info boxes had interactive buttons to highlight the particular elements with glowing outlines (Figure 8).

5 DEMONSTRATION

To evaluate the proposed VR-application for participatory planning we used it in two real projects, a new city area in aspern, Seestadt (Vienna, Austria) and the new railway station in Kapfenberg (Styria, Austria).

116

Figure 7. VR questionnaire to immediately gather feedback.

Figure 8. Highlighting elements in the VRE.

5.1 Technical Setup

For VR experiences we used a commercial off-the-shelf VR system, the HTC Vive and the HTC Vive Pro[2]. To control the avatar in the VRE, we used the HTC Vive controllers. We decided to use the HTC Vive, as its controllers offer an intuitive way of interacting with our system and the tracking, especially the possibility for walking around in VR, called room-scaling, is very promising. To select a desired function the users had to use the touchpad of the HTC Vive Controller (Figure 6). The VR system was connected to a gaming laptop, an Alienware 17R4 equipped with a Nvidia GTX 1080 graphic card.

5.2 Demonstration aspern Seestadt

For the application of VR in the participatory planning of public road space, the neighbourhood of "Quartier am Seebogen" in the northern part of the apsern Seestadt in Vienna was used. This area is particularly suitable for the application of VR because it is located between the tendering and construction phases and because the street space has already been designed. The future buildings in the neighbourhood have also been determined and gradually integrated into the 3D model (Figure 9).

The demonstration of the new city area "Quartier am Seebogen" took place in cooperation with the local district management, as their premises was used for the workshop. Therefore, the participation setting was very realistic and corre-

---

[2]https://www.vive.com

sponded to a real information event for citizens. Following media were used in this information event: classic plan presentations and renderings printed on paper and a HTC Vive and a VR application combined with the traffic simulation and developed in Unity. An example of a user in the VR can be seen in Figure 10. On these two days 30 persons tested the VR application, 13 male and 17 female. After the VR experience user feedback was gathered by using pen and paper questionnaires.

5.3 Demonstration Kapfenberg

To investigate urban development, we used the redesign project of the railway station including a new barrier-free access, traffic, cycle and footpath concept, bus terminal and the connection of a public space via an entrance portal to the railway station. First we made a demonstration for the city planning office of Kapfenberg and the Austrian Federal Railways ÖBB and participants had the opportunity to test the prototypes extensively and to give feedback. Subsequently a real world citizen participation event was conducted where citizens could inform themselves about the new planned railway station. Here, the following mediation technologies were used: Classical representation with CAD plans and the VR application. At this demonstration in Kapfenberg 29 people participated, 15 male and 14 females.

5.4 Insights

Overall 59 people participated in the two demonstrations and tested the VR application. The use of VR was easy to manage in terms of integration into the participation processes. Also the usage of VR was overall perceived well by the participants. In the post experience questionnaire we used a 4 point Likert scale from 1 totally agree to 4 totally disagree. That is a scale that forces an opinion and is good for recording opinions about a new application that the user has experienced. All participants strongly agreed that participation was fun (mean: 1.00, sd: 0.00) and that VR makes participation more easy (mean: 1.47, sd: 0.68). In both demonstrations the users agreed that the usage of VR is helpful for participation processes (mean: 1.32, sd: 0.58).

We could not observe negative effects of cybersickness and also the answers in the post experience questionnaire reveal no strong negative effects (disorientation - mean: 3.55, sd: 0.79; dizziness - mean: 3.65, sd: 0.70; nausea - mean: 3.92, sd: 0.43).

In contrast the users stated that the system was easy to use (mean: 1.24, sd: 0.47) and that they felt as being part of the virtual environment (mean: 1.70, sd: 0.77). Also, especially interesting for the simulation community is the fact that users strongly agreed that the virtual people, cyclists and cars have increased the impression of reality (mean: 1.37, sd: 0,56). Apart from user experience and acceptance practical aspects of administering and designing such participation workshops are important.

We could successfully show the feasibility of combining VR and a virtual traffic simulation and its application in a participation workshop. Nevertheless, following aspects have to be considered to successfully conduct such participation

Figure 9. VRE of the public road space in [blank for review].

Figure 10. Showing a VRE user in front of a projection screen.

workshops. We would like to emphasise that using VR needs enough time as the VR equipment has to be adapted for each participant and also the exploration of VR itself needs time. Especially interesting was the fact that often user wanted to stay longer in VR and the facilitators had to ask them to end their experience after a certain time. Thus, it is important to set and communicate a certain time limit for VR exposure to allow enough people to experience the VR environment and prevent possible negative virtual reality sickness effects. Also, it is important to consider using VR as an additional media but not as the only option. In the aforementioned demonstrations we also used analogue materials (folders etc.) to ensure that people can be informed and participate even if they do not want to use the VR setup.

6 CONCLUSIONS

We have presented an approach to combine traffic simulations with participatory elements like guided tours and using questionnaires for urban planning in a virtual environment.

Overall most participants were able to use the VR application after a short introduction. The combination of highly immersive and interactive virtual environments with traffic simulation was appreciated by the participants and resulted in highly valuable feedback for professional planners. Especially the usage of a scientific validated and accurate multi modal traffic simulation is an important aspect to ensure a realistic experience of public spaces and traffic situations.

At this point we want to give answers to the research question whether the developed application is suitable for participatory urban planning. When looking into the possibilities of using VR in participation processes we provide the following recommendations: For information VR is well suited because it allows for additional information qualities and enables information through tools and interactive applications that is not possible in conventional presentation methods. For consultation VR is well suited because it provides the possibility to enable additional feedback through interactive tools in the application. In principle VR is well suited to support cooperative processes, however, certain points must be considered: Meaningful cooperation is only possible if the right tools are provided and the right questions are asked, as regular people are not designers. In addition, the high immersion of the proposed designs creates a feeling of high realism. But if the designs shown in the VR are not realised in reality, this may lead to disappointment.

In future work we will further develop the simulation framework especially by adding more diverse 3D models of pedestrians and other characters to the simulation. Especially additional characters e.g. playing children on a playground, people sitting on benches etc. will be implemented to increase the realism of the environment.

REFERENCES

1. Alsaggaf, A. I., and Jrade, A. Benefits of integrating bim and gis in construction management and control (2015).

2. Bakar, N. A. A., Zulkifli, A. N., and Mohamed, N. F. F. The use of multimedia, Augmented Reality (AR) and Virtual Environment (VE) in enhancing children's understanding of road safety. In *2011 IEEE Conference on Open Systems* (Sept. 2011), 149–154.

3. Chun, W., Ge, C., Yanyan, L., and Horne, M. Virtual-Reality Based Integrated Traffic Simulation for Urban Planning. In *2008 International Conference on Computer Science and Software Engineering*, vol. 2 (Dec. 2008), 1137–1140.

4. Chun W.; Ge C.; Yanyan L.; Horne, M. Virtual-reality based integrated traffic simulation for urban planning. In *International Conference on Computer Science and Software Engineering*, vol. 2 (2008), 1137–1140.

5. Echevarria Sanchez, G. M., Van Renterghem, T., Sun, K., De Coensel, B., and Botteldooren, D. Using Virtual Reality for assessing the role of noise in the audio-visual design of an urban public space. *Landscape and Urban Planning 167*, Supplement C (Nov. 2017), 98–107.

6. Helbing, D., and Johansson, A. *Pedestrian, Crowd and Evacuation Dynamics*. Springer New York, New York, NY, 2009, 6476–6495.

7. Jerald, J., LaViola, J. J., and Marks, R. Vr interactions. In *ACM SIGGRAPH 2017 Courses*, SIGGRAPH '17, Association for Computing Machinery (New York, NY, USA, 2017).

8. Li, H., Zhang, J., Xia, L., Song, W., and Bode, N. W. Comparing the route-choice behavior of pedestrians around obstacles in a virtual experiment and a field study. *Transportation Research Part C: Emerging Technologies 107* (2019), 120–136.

9. Lopes, C. V., and Lindström, C. Virtual Cities in Urban Planning: The Uppsala Case Study. *J. Theor. Appl. Electron. Commer. Res. 7*, 3 (Dec. 2012), 88–100.

10. Nguyen, M.-T., Nguyen, H.-K., Vo-Lam, K.-D., Nguyen, X.-G., and Tran, M.-T. Applying Virtual Reality in City Planning. In *SpringerLink*, Springer, Cham (July 2016), 724–735.

11. Portman, M. E., Natapov, A., and Fisher-Gewirtzman, D. To go where no man has gone before: Virtual reality in architecture, landscape architecture and environmental planning. *Computers, Environment and Urban Systems 54*, Supplement C (Nov. 2015), 376–384.

12. Regal, G., Schatz, R., Schrammel, J., and Suette, S. Vrate: a unity3d asset for integrating subjective assessment questionnaires in virtual environments. In *2018 Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, IEEE (2018), 1–3.

13. Sanchez-Sepulveda, M., Fonseca, D., Franquesa, J., and Redondo, E. Virtual interactive innovations applied for digital urban transformations. Mixed approach. *Future Generation Computer Systems* (Sept. 2018).

14. Schramka, F., Arisona, S., Joos, M., and Erath, A. Development of virtual reality cycling simulator. *Journal of Computers 13*, 6 (2018-06), 603 – 615.

15. Schrom-Feiertag, H., Lorenz, F., Regal, G., and Settgast, V. Augmented and Virtual Reality Applied for Innovative, Inclusive and Efficient Participatory Planning. In *Proceedings of 7th Transport Research Arena TRA 2018*. 2018.

16. Stauskis, G. Development of methods and practices of virtual reality as a tool for participatory urban planning: a case study of Vilnius City as an example for improving environmental, social and energy sustainability. *Energy, Sustainability and Society 4*, 1 (Apr. 2014), 7.

17. Suneson, K., Allwood, C. M., Heldal, I., Paulin, D., Roupé, M., Johansson, M., and Westerdahl, B. Virtual Reality Supporting Environmental Planning Processes: A Case Study of the City Library in Gothenburg. In *Chalmers Publication Library (CPL)* (2008), 481–490.

18. Tettamanti, T., Szalai, M., Vass, S., and Tihanyi, V. Vehicle-in-the-loop test environment for autonomous driving with microscopic traffic simulation. In *2018 IEEE International Conference on Vehicular Electronics and Safety (ICVES)* (Sep. 2018), 1–6.

19. Traffic Simulator for Virtual Reality with PTV Vissim. `https://www.ptvgroup.com/en/solutions/products/ptv-vissim/areas-of-application/vr-traffic-simulation/`, Dec. 2019.

Assessing the Emotional Impact of Cognitive Affordance in the Built Environment through Augmented Reality

Stefan Florescu[1,2], Diana Elena Nistor[3]

[1]Bartlett School of Architecture
University College London
London, United Kingdom
stefan.florescu.13@ucl.ac.uk

[2]AHMM Architects
London, United Kingdom
sflorescu@ahmm.co.uk

[3]Carol Davila University of
Medicine and Pharmacy
Bucharest, Romania
diana.nistor@stud.umfcd.ro

ABSTRACT

Today, Augmented Reality (AR) technologies allow us to enhance our surroundings with virtual visual information in order to help us perform certain actions: from navigation to play; from highly advanced professional use to mass consumer use. But what actually happens when you augment a physical object with virtual information? How does this affect the user? The research proposes to answer a very particular question related to the level of well-being within the built environment: *Does a higher level of cognitive affordance correlate with a higher level of well-being?*

The paper presented here is part of a much more detailed master thesis done at the Bartlett School of Architecture UCL and focuses on the technical aspect of the research: software and hardware, physical and augmented reality within the built environment. The subject of the thesis was driven by a specific situation in Switzerland, used here as a case study for the proposed methodology. The project puts forward the idea of an augmentable architecture that through AR will be enhanced visually (by adding cognitive affordance), thus offering the image of a much more familiar environment and implicitly inducing a new level of comfort or well-being for users.

In order to test this hypothesis, an experiment was conducted over a small number of participants experiencing an AR environment using a video see-through device and an AR smartphone application developed by the authors. During their experiences, their emotional activity was recorded in subjective and objective manners, through questionnaires and electroencephalogram (EEG) readings respectively.

The thesis talks about affordances, emotions through technology and simulations, describes the AR environment and its related technologies, human factors and human computer interaction, and focuses on how to identify and quantify the well-being emotions within the experiment. The self-assessed method was correlated with the electrophysiological readings in order to determine if the expected emotions were present during the AR simulation.

Author Keywords
Augmented reality; Cognitive affordance; Augmentable Architecture; Urban-Rural Simulation.

ACM Classification Keywords
I.6.1 SIMULATION AND MODELING

1 INTRODUCTION

Today, by *augmented reality* we understand a technology that integrates in real time digital information with the real-world environment. Through the addition of stimuli that respond to all senses, AR enhances the experience of the user by superimposing virtual worlds on top of the physical environment using mobile devices or special headgears.

Although AR technology was originally developed to facilitate and ease specific industrial tasks, now it is on the verge of transforming more domains such as art, entertainment, construction, science and engineering, health and medicine, aerospace, defence and education [1].

In architecture and construction AR capabilities are yet reduced and focus on visualising and construction management. This paper aims to analyse the use of AR technology in architecture from a different point of view, namely focusing on the effects and changes over the user.

The paper will contextualise the intersecting fields of psychology, AR environments, emotions measurement techniques, and architecture in order to see if certain emotions are present with the mixed reality environment. The hypothesis will be tested through an experiment that uses two different methods for quantifying the emotional state of the participants. The paper will focus on the methodology used in the experiment, namely by presenting the physical and AR environment, how those two were created and the hardware and software used for assessing participants' emotions.

The design project on which this paper is based proposes a new architectural typology in Albinen, a place where new settlers will move temporarily as part of their application process. Located in the middle of the existing village, this new architectural assemble will use AR technology in order to make a step by step transition from urban to rural landscape. The virtual environment will allow newcomers

to discover gradually and understand the beauty of this village by making a visual transition between urban-like forms and the vernacular architecture of alpine chalets. Initially encountered as a fully immersive AR experience, their perception of surroundings will gradually merge with the reality of the physical environment (Figure 1).

2   CONTEXT

The subject of this paper was driven by the increased fascination of the general public with the AR technology and by a very specific situation in Switzerland, in the village of Albinen, a rural settlement with decreasing demographics, where newcomers are being paid by local authorities to move in and settle [2].

The immigrant condition in Switzerland is very specific; foreigners that immigrate here have an increased rate of higher-education qualification and more than half used to live in urban environments [3]. At the same time, Switzerland has an alarming rural to urban migration. Swiss villages are threatened to perish while big cities cannot cope with the growing population caused by rural to urban migration and international immigration [3].

Albinen is one of the many Swiss villages threatened to parish due to a drastic decline in population. Local authorities adopted an uncommon strategy in order to tackle this widespread problem by offering a settlement grant to those willing to move there [2]. Although having a predicted high rate of success, the main concern that fears local authorities is that newcomers, most probably with an urban background, will not fully accommodate to their new rural lifestyle and will live there only for how long they contractually committed, thus not solving on long term the village depopulation issue.

3   HYPOTHESIS

Architecture and AR technology can be an answer to this challenge of accommodation uncertainty. AR technology can be calibrated to augment the rural landscape with urban elements in a very specific manner, assuring a gradual transition from city to village. In this way, the new settlers can comprehend and experience the new landscape gradually, until they are fully accustomed to this new lifestyle, thus creating a genuine bond with the rural environment, not financially stimulated, and minimising the chance of moving out after the contract will end.

But what is actually happening when the physical is being augmented with virtual visual information? - it obtains a new level of cognitive affordance.

This paper proposes to test the core idea of the above speculation and to analyse if a new layer of information added through AR will have any emotional impact over the architectural-space user. Therefore, the scientific question of this paper is: *Does a higher level of cognitive affordance correlate with a higher level of well-being?*

Through a specific methodology described in the following chapter, this paper aims to describe the process used for analysing the psychological and emotional effects of virtual augmentation over the user and how those influence our experience in the built environment.

4   METHODOLOGY

In order to test the hypothesis, a series of experiments were conducted over a group of participants. In order to increase the reliability of the results, a scientific control experiment was done alongside the actual experiment.

The experiment consists in each participant experiencing
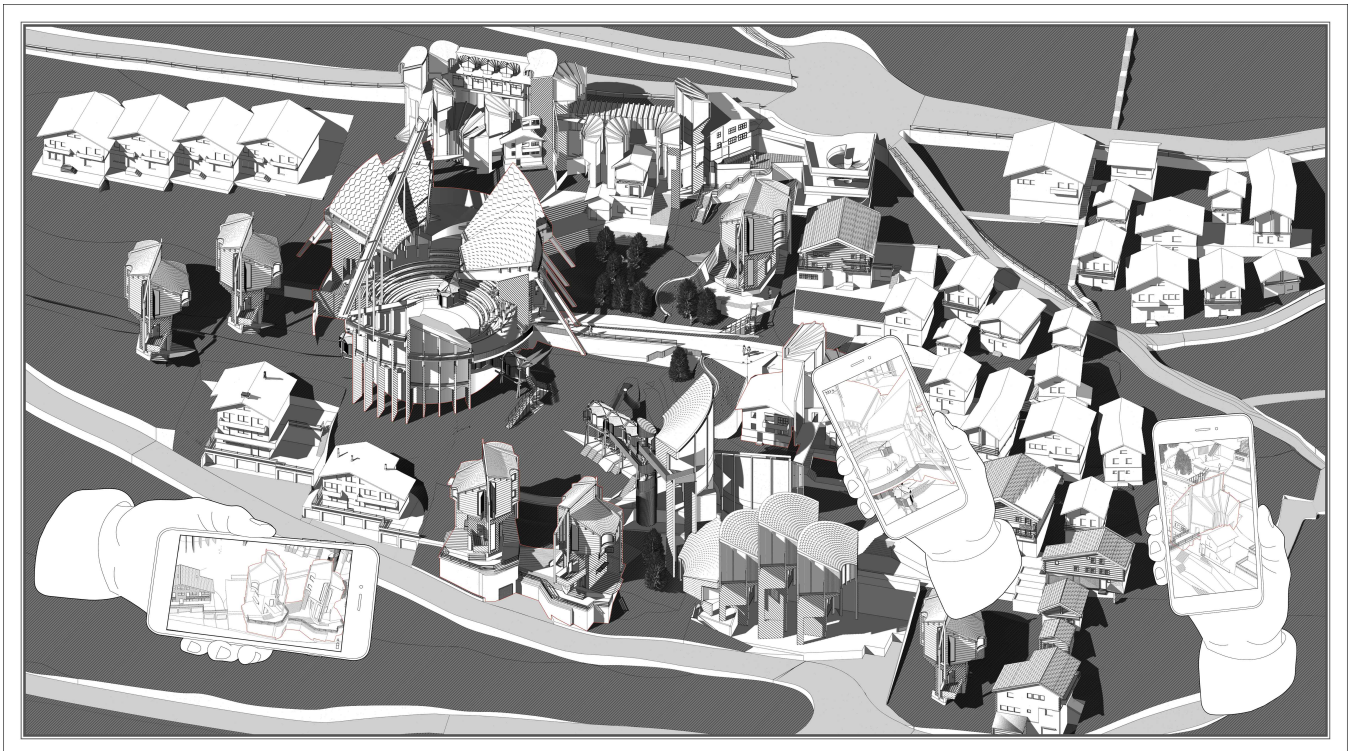
Figure 1. The New Village of Albinen – aerial perspective of the design proposal for augmentable architecture.

various AR environments over a physical model while their brain activity is being recorded. To ascertain that conditioning does not occur, each participant will handle the AR environment individually and will experience one by one, in no particular order, the control test and the test regarding cognitive affordance.

The experiment aims to analyse in both objective (through EEG recordings) and subjective (through questionnaires) manners if the participants' well-being emotions are triggered by the experiment. The control experiment is held in order to eliminate the emotions generated by the unexpected, when participants will handle for the first time an AR environment.

For an easier understanding, in the following subchapters we will refer to the control experiment as *The Control AR Test* and for the actual experiment as *The AR Test*. The sequence in which the experiment was conducted is as follows: Pre-AR Test Questionnaire, either the Control AR

Figure 2. Experiment testing ground; the physical model (centre), EEG device (left), video see-through AR device (right), Pre-AR Test Questionnaire (far right).

Figure 3. Participant exploring the AR environment while being EEG monitored.

Test or the AR Test, Post-AR Test Questionnaire.

4.1 Participants

In order to produce a meaningful result, a sample of 12 research participants was selected to be tested. For consistency, all participants were architecturally trained students. Prior to the experiment participants were questioned to ascertain general data about their *age, gender, and experience with AR environments*, thus assuring that the chosen sample of testers is homogenous. Beside these basic enquires, other questions strictly relevant for the experiment were asked as follows: *Do you suffer from epilepsy? (A: Yes or No)*, *Which of the following describes best the landscape of the model? (A: Rural Area or Urban Area)*, *With what geographic region can you associate this landscape? (A: Eastern Europe, Western Europe, Asia)*, *What is the function of the building in the foreground? (A: Residential, Commercial, Agricultural, Civic, Religious)*.

Although 12 was considered a suitable number of participants for producing relevant results for this paper hypothesis, in order to produce results that have statistical significance the experiment has to be carried out on a much larger scale, on a bigger sample of participants.

4.2 Physical Environment

The experiment was conducted within the Bartlett School of Architecture, UCL in London. The testing area was a rectangular room containing only a chair and a table. The physical ambient (Figure 2) was as neutral as possible, with white walls, a white table and a dark-coloured carpet. With no visually dominant element within the testing ground, participants can generate genuine reactions, not altered or influenced by any external factors. At the same time, in order to keep the electrical noise at minimum, no wall-plugs or light bulbs were closer than 5 meters from the testing area, allowing the EEG device to record accurate results.

Once the participant finished the Pre-Test Questionnaire, the physical model (that later on will support the virtual augmentation) was brought in. The physical model was a 3D printed object made in one piece using white PLA filament (Figure 4, top). The model depicted simple architectural volumes with no details that resembled Swiss vernacular architecture. The model was designed as a diorama and allows participants to see the entire rural landscape, thus compensating for not being a fully immersive AR experience. In the foreground the dominant volume depicts the massing of a new building typology designed for the newcomers in order to allow them to adapt to this new lifestyle.

4.3 Hardware and Software

The AR environment was presented to participants using a Video See-Through display, namely an iPad Pro. The device has a back-facing video camera of 12Mp with a field of view (FOV) of 60 degrees when the device is held horizontally. The video information is being displayed on a large 12.9-inch screen with a resolution of 256 pixels per inch and a refresh rate of 120Hz.
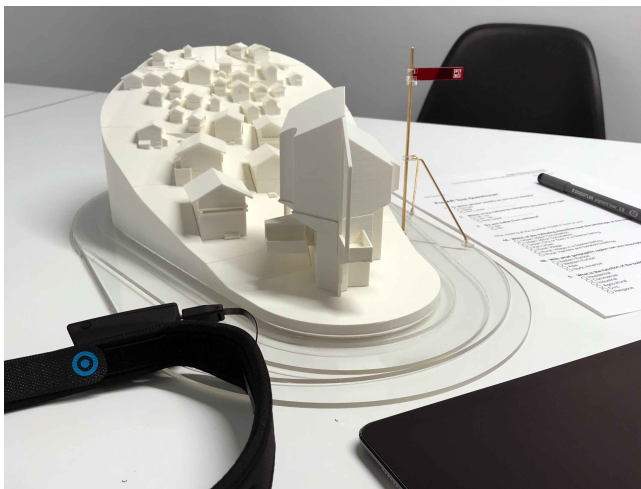
The video signals from the camera and the computer-generated imagery are combined by an IOS application specially designed for this test, in order to produce the Augmented Reality. The application was created using Vuforia Model Target and Unity, allowing the iPad camera to detect and track the 3D printed physical model and overlay virtual 3D elements onto it.

The hardware technology involved in this experiment allows participants to have 360 degree of freedom over the physical model, thus enabling them to explore the virtual environment from any desired angle. Although not a fully immersive experience, this AR application manages to engage the user through its interactive manner and through the large device display. But even with such display and a high frame rate of 60fps, Breaks in Presence (BIP) occur. Some of the reasons for which these BIP happen could be the absence of sound or the use of a Video See-Through device instead of a Binocular Augmenting-Display device.

The emotional development of participants was measured by monitoring the brainwave activity. The electrical discharge created by neurons when changing emotional states was measured with MyndBand BLE Brainwave Headset (Figure 3), a portable electroencephalogram (EEG) device with 3 gold-plated dry sensors. MyndBand Brainwave Headset uses the referential measurement technique. Thus, apart from the 3 dry sensors, this EEG device contains another referential electrode placed on the earlobe, a neutral location from an electrical point of view. The device compares the signal received by the dry sensors with a signal recorded by the referential electrode and manages to eliminate the electrical noise. The 3 dry sensors were placed on the forehead of participants. This location allows the EEG to measure brainwave activity of the pre-frontal cortex, the area where higher thinking and emotions occur. The MyndBand Headset is able to detect brainwave activity from 0.5Hz up to 40Hz, thus recording the whole spectrum of emotional states, including *pleasure*, *satisfaction*, *relaxation* and *well-being*. The raw data collected by the sensors is then transmitted via Bluetooth to a MyndPlay Pro software which records and displays as graphs all frequency bands of the participant brain. Being light and ergonomic, this EEG device allows the user to interact freely with the physical and AR environment while safely measuring the brainwave signals.

4.4 Augmented Reality Environment
The virtual environment unfolds to participants into two tests: the Control AR Test and the AR Test.

The Control AR Test augments the foreground of the physical model with an avatar of a girl hand-waving to the user (Figure 4, middle). This type of information identifies as false affordance. After the Control AR Test, the surprise emotion generated by the first interaction with the AR environment within the test disappears; and this without being yet exposed to any cognitive affordance.
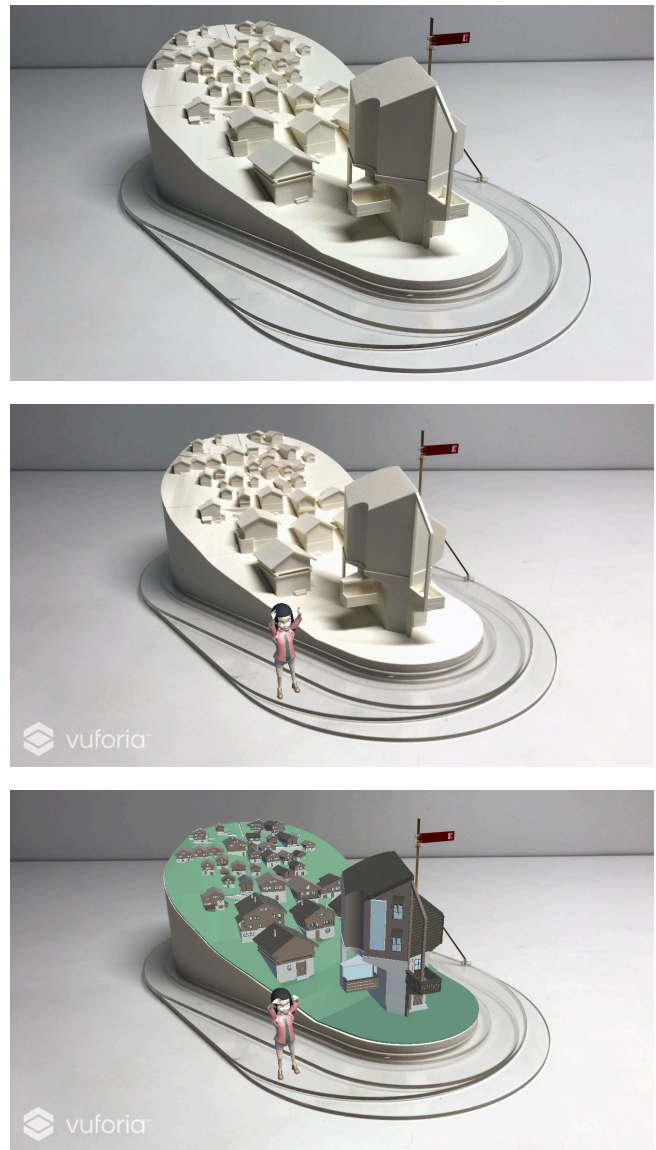
Figure 4. Participant's view of the experiment.; the physical model (top), the control AR test (middle), the AR test (bottom)

The AR Test changes only one variable to the control experiment. Beside the waving girl, the main experiment adds to the physical model a virtual layer abundant in architectural information: architectural elements (void-solid, windows, doors, roofs), architectural details (vernacular architectural elements, balustrades, fenestration) and materiality (timber, stone, glass) (Figure 4, bottom). This AR information translates into an influx of cognitive affordance. Now, participants perceive a higher level of cognitive affordance than before the virtual augmentation.

The virtual environment is based on a conceptual design of a centre for the newcomers that are eligible to apply to

settle in Albinen. Families that are in the final stages of selections are invited to spend 3 months on this campus in order to accommodate and understand rural life. The experiment depicts one of the houses in which the newcomers will live within those trial months. The architectural style is inspired by local vernacular architecture with modern urban influences.

4.5 Measurement Technique

This experiment attempts to assess well-being induced by a higher level of cognitive affordance. This is achieved in both quantitative and qualitative ways, through answering the Post-AR Test Questionnaire and through the data obtained by the MyndBand EEG Headset respectively. Data will be collected only from the AR Test.

Well-being is an emotion that covers a large number of feelings, but all from the same spectrum. According to the Oxford Dictionary, *well- being* translates as '*a state of being comfortable, healthy and happy, in a pleasurable environment*'. Within the experiment *well-being* deals with the comfort of having more knowledge about the surrounding environment; as an antithetic example people do not feel comfortable in darkness because of lack of information.

Thus, *well-being* was measured by reading and interpreting the presence of higher Alpha Brainwaves and lower Beta Brainwaves (frequencies between 10-16Hz) in each participant's brain activity. If during the experiment brainwaves within those frequencies are recorded, most likely emotions related to well-being were experienced by the participants.

To ascertain the emotional state during the experiment the Post-AR Test Questionnaire contained questions about the binary emotions from the same spectrum such as *pleasure* and *enjoyment*.

4.6 Post-AR Test Questionnaire

After finishing the AR experiment participants were asked to answer a short set of questions: *How comfortable do you feel within this environment?* (A: *0-100*), *What other emotions did the AR environment induce?* (A: *Well-being, Anticipation, Anger or Surprise*), *What AR architectural element do you remember?*.

5 RESULTS

The Pre-AR Questionnaire had six questions, three general and three experiment specific. The general questions were asked in order to determine the nature of the study group. The first and second questions ('*To which gender identity do you most identify?*' and '*Which of the following categories describes your age?*' respectively) had equal percentages for each answer, suggesting that the sample of participants is homogenous, thus making the results more reliable. The third question ('*Do you suffer from epilepsy?*') was mandatory and was conditioning participation to the next phase of the experiment; all 12 participants responded

'*No*' allowing them to be tested with the EEG device. The following questions had to be answered while analysing the physical model. Questions *4A* and *4B* asked about the landscape and the geographic location depicted in the model. Participants had various answers but with the highest score were '*Urban: City or Town in a Sparse Settings*' (33,33% of answers) and '*Western Europe*' (58,77% of answers) respectively. The last enquiry of this questionnaire was '*What is the function of the building in the foreground?*' and although it received all given answers, the response that has the highest score was '*Civic*' with 58,33% or 7 votes.

The Post-AR Questionnaire had seven questions: four related to the AR experience and three identical to the Pre-AR Questionnaire. The first one asked '*How comfortable do you feel within this environment?*' and participants' average answer was '*84*' out of 100. The second question was also related to the emotions within the AR environment ('*What other emotions did the AR environment induce?*') and was answered with '*Well-being*' (50%), '*Anticipation*' (33%) and '*Surprise'* (17%). The third question was open-ended, allowing respondents to provide their own answer, and asked '*What AR architectural element do you remember?*'. Although all the answers were different, most of them depicted architectural details such as roof tiles, fenestration or balustrades.

The next set of questions was the same as *Questions 4A, 4B and 5* from the Pre-AR Questionnaire and were asked in order to see if the higher level of cognitive affordance affected participants' perception in the AR environment. Although identical questions, the answers were different. While Q4 maintained the same answer as favourite, but with a better score (83,33% instead of 58,33%), Q3 and Q6 received a different response: '*Rural: Village*' (58,33%) instead of '*Urban: City or Town in a Sparse Settings*' (33,33%) and '*Residential*' (75%) instead of '*Civic*' (58,33%).

At Q5 ('*Do you consider there is enough architectural detail provided by the AR environment in order to support your previous answers?*') 83,33% of participants answered with '*Yes*' , reinforcing the conclusion of the questionnaires results and answering the hypothesis: the level of detail did actually influence participants in their perception, thus impacting their emotions. This will be discussed in further detail in the *Discussion* chapter.

The EEG results were displayed by the MyndPlayer Pro software into three different graphs: raw EEG data, MyndPlayer Emotion Graph and Brainwaves Frequency Bands. EEG readings are based on interpreting three variables: frequency, voltage and morphology. Frequency describes the repetitive activity in Hz of the waves and it can be rhythmic (waves of constant frequency), arrhythmic (EEG activity with no stable rhythms) and dysrhythmic (EEG activity characteristic to unhealthy subjects). Voltage refers to the peak of the EEG activity and it can be
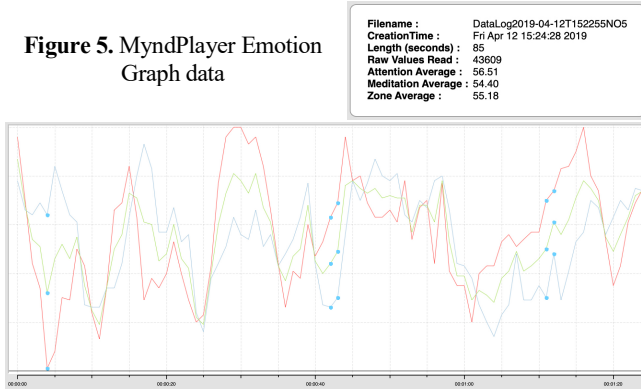
Figure 5. MyndPlayer Emotion Graph data

Figure 6. MyndPlayer Emotion Graph data
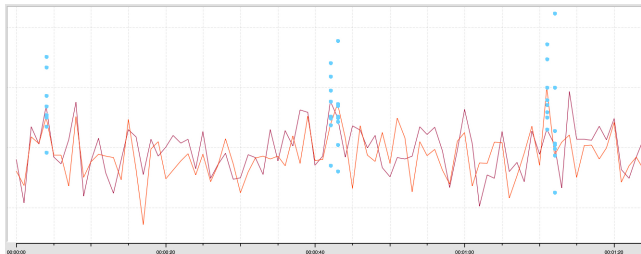
Figure 7. MyndPlayer Brainwaves Frequency Bands Graph

attenuated (a reduction in the amplitude of the waves resulted from low voltage), hypersynchrony (a rhythmic activity with an increase in voltage) and paroxysmal (a rapid increase in voltage that has an abrupt return to lower voltage activity). Morphology represents the shape of the waveform and it can be monomorphic (one dominant EEG activity), polymorphic (distinct EEG activity that combines into a complex waveform), sinusoidal (waves similar in shape with sine waves) and transient (an isolated pattern distinct from the general EEG activity) [5].

The first graph describes the raw EEG activity (µV/t). The second graph describes the strength of emotions over time and represents an interpretation of the raw EEG activity by the MyndPlayer Pro software (Figure 6). This shows *Attention* (which can be correlated with fully aware mind open to emotions such as surprise or anticipation), *Meditation* (which can be correlated with a relaxed mind and with emotions such as serenity and joy) and *Zone* (which represents a combination of both mental focus and mental calm, vital for emotional balance, peak performance, comfort and well-being) [6]. Each participant registered a score of these variables that represents their average activity (Figure 5). The average emotional activity of all participants was the following: *Attention* 53,7% of time, *Meditation* 57,1% of time and *Zone* 55,2% of time.

The last graph shows the frequencies of the brainwaves (Hz) in relation to time (s). Only highAlpha and lowBeta waves were displayed in this graph as they are the only relevant waves for this experiment (Figure 7). Although these graphs (Hz/t) cannot be interpreted as an average activity of all participants, by analysing the individual

results it can easily be seen a direct correlation between the paroxysmal voltage and the answers given at Q3 from Post-AR Questionnaire, together with observation made during the experiment. In the following chapter the paper discusses in further detail these aspects.

The results obtained from the EEG readings during the Control AR Test and the AR Test were analysed using a T-Test in SPSS statistical analysis software. The experiment had the same group of participants and two different EEG readings (from AR Control Test and from AR Test), thus the Paired Samples T-Test was chosen to analyse the Attention, Meditation and Zone numerical data obtained. The result of the T-Test showed a statistical probability that having a higher level of cognitive affordance will affect emotions, as follows: $p=.035$ for Attention, $p=.001$ for Meditation, $p=.051$ for Zone. Except for Meditation, the p value suggests little statistical significance to indicate a positive correlation in these tests. Although the EEG results match with the Questionnaires results and clearly indicate that the emotional state of the participants was influenced by the amount of cognitive affordance presented through the AR environment, the high p values obtained in the T-Test could be explained by the small sample of participants tested.

6 DISCUSSION

6.1 Cognitive affordance changed participants' mind

The Pre-AR Test Questionnaire had three specific questions which participants had to answer before experimenting the AR environment. They assessed participants' degree of understanding over the physical model and constrained them to analyse the depicted scenario on different architectural levels: background, context and foreground. The physical model was composed of simple massing volumes, specially designed not to indicate any available answer.

The first question was related to the architectural background of the model and tested if the information provided by the physical model was sufficient to trigger what in cognitive neuroscience is called pattern recognition. The question was asking what kind of landscape best describes the model and it had multiple choice answers. Although all participants were architecturally trained and familiar with the architectural concepts behind each answer option (urban conurbation, city in sparse setting, peri-urban or rural village), the cognitive information on the model was inadequate in order to activate the encoding visual pattern for each participant. Thus, the responses received were diverse, without major differences in percentage: city or town in a sparse setting 33%, village 25%, village in a sparse setting 16%, peri-urban 16% and hamlets and isolated dwellings 8%.

The second question was '*With what geographic region can you associate this AR landscape?*' and was assessing the understanding of the architectural context of the model.

Based on their architectural knowledge, participants were trying to associate the geometric forms of the model with different architectural styles, specific to each geographical region described by the multiple-choice answer: Eastern Europe, Asia, Western Europe or North America. In this case the pattern recognition to which participants have undergone refers to the recognition-by-components theory (RBC) which suggests that in order to recognise objects people decompose complex volumes in basic 3D shapes (cylinders, cones, etc) named geons. Irving Biederman, the author of this theory, suggested that all the objects can be created by combining 36 types of geons [7]. The physical model was composed of simple massing volumes that definitely resembled a residential landscape, but due to the insufficient cognitive information provided, the geons that participants detected (cubes and prisms in this case) were too generic to accurately indicate the geographical area of origin. Thus, different responses were recorded: 25% answered Eastern Europe, 58% Western Europe and 16% Asia.

The third question was assessing the level of perception for the main architectural piece of the model by asking '*What is the function of the building in the foreground?*'. This question subjected participants to the same cognitive neuroscience process as the previous question. The peculiar shape, specially designed not to be easily associated with any architectural functions, created difficulties in decomposing the volume in recognisable geons, despite participants' architectural training. Thus, the answers to this enquiry were also diverse and checked all possible answers. Although all multiple responses of this question were picked at least once, more than half of participants (58%) answered with '*Civic*'. In the absence of enough information, the unrecognisable massing of the building made participants use architectural deduction, thus selecting from the multiple choices an option that they thought will suit the landscape.

The AR environment presented within the experiment added an influx of cognitive affordance to the physical model in the form of architectural elements (void-solid, windows, doors, roofs), architectural details (vernacular architectural elements, balustrades, fenestration) and materiality (timber, stone, glass).

Despite the fact that the Post-AR Test Questionnaire reproduced the same questions addressed in the Pre-AR Test, the same group of participants responded differently after the AR experience.

For the first identical question, 66% of respondents picked a different answer. Not only that more than half of participants opted for a different response, but the most selected answer changed as well, registering an even higher percentage than the main answer from the Pre-AR Test Questionnaire (58% instead of 33%). The second question registered a 25% change in answers while the third question

recorded the highest percentage of different responses between pre and post AR experience.

The question related to the function of the building in the foreground registered an 83% change in answers. At the same time, the favourite response changed as well and recorded a higher rate of selection than the first time (75% instead of 58%). Moreover, the post-AR answer '*Residential*' coincides with the design intentions of the author for the foreground object.

All these changes in the respondents' answers suggest that the cognitive affordance added by the AR environment generated enough information in order to trigger participants' pattern recognition. The supplemental information provided by the AR together with participants' architectural knowledge helped access the recognition-by-components pattern, allowing subjects to make more informed choices.

The different results given by the same group to these identical questions indicate that a higher level of cognitive affordance play a very important role in decision making and mood.

6.2 Participants' emotions within the AR environment
In the previous subchapter we saw that the amount of cognitive affordance influences decision making and mood. In the following paragraphs, the thesis will analyse how emotional state changed and what feelings were present while being exposed to this influx of new information.

At declarative level, participants' answer to Q1A describe an emotional state with positive valences within the AR environment, with an average of the level of comfort of 84%. The following results supported this figure, participants choosing '*Well-being*' and '*Surprise*' as other emotions induced by the AR.

These subjective responses were reinforced by objective figures obtained through EEG readings, as presented in the previous chapter Results. Interpreting the EEG readings in comparison with the questionnaire answers produced interesting results.

By analysing the individual results from the brainwave frequency graph, it can easily be seen a direct correlation between the paroxysmal voltage and the answers given at Q2 from Post-AR questionnaire, together with observations made during the experiment.

For example, the graph below (Figure 8) represents the highAlpha and lowBeta waves of Participant 12 and at
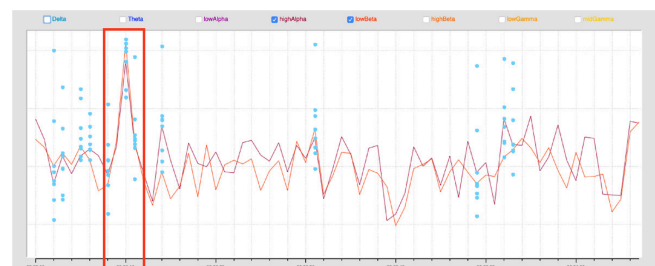

Figure 8. Participant 12 MyndPlayer Brainwaves Frequency Bands Graph.

127

second number 10, the EEG recorded a transient paroxysmal wave. This represents a peek in the participant emotional response and correlates with the physical reaction of exclaiming '*I can see the tiles on each building roof*' together with the answer given for Q2, '*Tiles on roofs*'.

Another similar example was encountered at Participant 6. In the brainwave frequency chart, we can easily identify an intensified alpha and beta wave activity at second 119. This corresponds with the moment when the participant was analysing the architectural details of the facades and matches perfectly with the answer provided at Q3, '*Balustrades*'. But the detection of intensive emotions during the experiment was not mandatory for inducing a general well-being state over the user. According to his questionnaire answers, Participant 10 rated the comfort within the AR environment with 80 and ticked '*Well-being*'. His EEG scores from the second graph recorded the highest values among all participants with attention at 74, meditation at 62 and zone at 68. Although his excellent scores relate to a general emotional state of well-being and comfort, his alpha and beta wave activity recorded no transient paroxysmal waves, the graphs recording a normal pattern activity.

Given all of the above, the paper proved the centre question of the hypothesis; a higher level of cognitive affordance does correlate with a high level of well-being and it does influence the general emotional state of the user together with its decision making.

7   CONCLUSION

The specific situation of the Swiss village of Albinen and the design project set the core research of this paper. The research design was based on the fact that AR adds a new layer of information as cognitive affordance, creating an immersive virtual environment which translates into a larger spectrum of emotions experienced by the user than in the real-world surroundings. The paper validates the relationship between the level of cognitive affordance and the well-being emotions within the built environment. Although the central hypothesis was proven, the results of the T-Test together with the very small sample of participants used for this demonstration suggest that further studies should be conducted in order to make this idea relevant for research.

As previously stated, AR technology was used in this paper as a tool and tools play a determinant role in achieving any goal; powerful tools increase the accuracy of the results. In this paper, a general-purpose device with AR capabilities was used in order to conduct the experiment. Although it produced eloquent results, the way in which these could be influenced by a more powerful technology was debated in the previous chapter *Discussion*. But even with the most sophisticated AR technology available now, can we

produce this kind of impact over the user on a day to day basis?

The current AR technology is still in the infant stage and indubitably the current systems represent only the tip of the iceberg of what is to come. Today, even the most advanced AR devices still have bulky form, limited display capacities (small FOV, low resolution, insufficient framerate and problems with the vestibulo-ocular-reflex) and still induce VIMS. But looking into the near future, these are very likely about to change.

Together with the technological advancements, neuroscience and psychology are evolving as well, making emotional states and moods easier to understand and quantify in the near future.

In conclusion, taking into consideration all points mentioned above, AR can manipulate the cognitive affordance around us in order to enhance or induce certain emotional states. The current growth of AR technologies, together with the advancements in psychology and neuroscience, make the idea of having a built environment that adapts to your emotions and needs closer than ever.

REFERENCES
1. Aukstakalnis, S. *Practical Augmented Reality: A Guide to the Technologies, Applications and Human Factors for AR and VR.* Pearson Education, Inc., USA, 2017.
2. Swissinfo.ch, Albinen cash incentive given green light https://www.swissinfo.ch/eng/business/attracting-new-residents-_move-to-albinen-and-get-_chf25-000-/43719788 As of 15 April 2019.
3. Swissinfo.ch, Swiss population getting larger, older, more diverse https://www.swissinfo.ch/eng/society/digging-into-_data_swiss-population-getting-larger--older--more-diverse/41129138 As of 15 April 2019.
4. World Bank | Switzerland Data, https://data.worldbank.org/country/switzerland As of 23 April 2019.
5. Gert et al editors, *Rhythmic EEG activities and cortical functioning: Proceedings of the 1st International Symposium on Event-related changes in Cortical Rhythmic Activities-Behavioural Correlates...1979, Developments in neuroscience; v. 10. Elsevier*, Amsterdam, Netherlands, 1980.
6. Myndplay.com, MyndPlay explained https://myndplay.com/myndplay-explained As of 15 April 2019.
7. Sternberg, R. J., *Cognitive Psychology, 6 edition.* Ed. Cengage Learning, Belmont, Calif, 2011.

Generative Design

A Multi-Objective Optimization Workflow based on Solar Access and Solar Radiation for the Design of Building Envelopes in Cold Climates

Abel Sepúlveda and Francesco De Luca

Tallinn University of Technology
Tallinn, Estonia
{absepu, francesco.deluca}@taltech.ee

ABSTRACT

This paper proposes a workflow based on multi-objective optimization for the design of the building envelope. The design variables of the workflow are the sorting method of sun vectors to calculate the solar envelope, the building orientation and the number of floor plan divisions. The objectives are the maximization of the building volume, the maximization of the ratio of the windows that fulfill the solar access requirement and the minimization of the mean incident solar radiation through the windows to reduce the risk of overheating during the warm season. The workflow has the potential to be adopted by designers being integrated in the Grasshopper plug-in for Rhinoceros that is a widely used design platform. The proposed workflow shows a successful way to deal with complex multi objective design goals. Multi-objective criteria based on volume and solar access achieve the best trade-off solutions with volume ratios higher than 20% and ratio of the windows that fulfill the solar access requirements higher than 57%.

Author Keywords

Solar envelope; environmental analysis; building envelope; urban planning; building design; solar access; optimization

1 INTRODUCTION

Cities currently account for over 70% of global carbon emissions and up to 40% of the total world energy consumption. By 2030, the 80% of the world population will live in cities [16]. These high levels of pollution have irreversible repercussions on the climate change due to the global warming effect. The main cause of the high carbon emissions is the dependency of energy production on fossil fuels. Project Horizon 2020 created by the European Union proposes energy efficiency requirements for the nearly zero energy buildings (nZEB). Additionally, it is well known that indoor comfort in buildings is a key factor for human well-being [15]. In terms of indoor comfort, it is not only important the thermal but the visual comfort. The visual comfort includes view out, glare protection, daylight provision and solar access (SA) in buildings [11].

SA is one of the first criterion that architects might must consider in the early stages of the building design. SA

regulations influence building density and maximum buildable volume. The solar envelope (SE) is a method introduced by Knowles to calculate maximum volume new buildings cannot exceed to guarantee required solar access on neighboring facades [13]. The shape and the size of new building envelope have crucial influence on energy performance and indoor comfort levels. It is not only important how the new building will influence the urban environment but the other way around. The solar collection envelope (SCE) is the minimum height of each considered point of the buildable zone that fulfills the SA regulation [2] [9]. Most of the SA regulations specify the minimum quantity of sun hours per apartment during a specific period [6] [12]. There is no specification in terms of quality of the sunlight; hence, this allows flexibility to generate SEs even for the same regulation and urban environment.

Essentially, two different approaches exist to calculate the SE: calculating the maximum height to each grid point of the plot (additive algorithm) or deleting grid volumes from the theoretical buildable block (subtractive algorithm). Subtractive algorithms have been proved to be high time-consuming [21] but useful to use as building massing technique [5]. On the other hand, the additive algorithm defined by Knowles [16] and firstly implemented by Capeluto [2] is the most used for the SE generation due to its simplicity. In recent years, this additive algorithm is implemented in several tools like LadyBug [14] and DIVA4Rhino [20], both Grasshopper plug-ins for Rhinoceros software. Nevertheless, these tools do not calculate urban context-dependent SEs. Indeed, the urban context has been taken into account by De Luca [8] and Capeluto in different SE generation methods [1].

Recently, a novel method based on solar ordinances for urban planning has been developed. The method uses multiple criteria to generate the SE for different SA regulations [6]. The lack of existence of tools that could take into account the urban environment and different solar ordinances to generate the SE, motivated researchers to develop an open-source plug-in for Grasshopper called "Solar Envelope Tools" (SET) [18] based on the previous method mentioned [6]. Another previous research showed

how different pattern layouts influence in terms of total floor area and SA performance of the building masses located in urban areas in Tallinn [4]. However, the performance of the building masses generated from each pattern layout differ from the building envelope performance. Moreover, De Luca, Nejur and Dogan proposed a method for the assessment of optimal building clusters for direct solar access in urban environments in Estonia, preferred type of floor plan layout (interior and exterior) and buildable floor area [7]. In fact, the method for the optimization of building clusters gives detailed information about the building envelope design. Nevertheless, in both mentioned investigations, a unique SE per urban environment was considered. These facts reduce the flexibility of the architectural design process at building envelope level and interior floor layout. Indeed, there is a lack of research about potentialities of different SEs per urban environment for the building envelope design.

The aim of this research is to study how SEs based on different criteria influence the building envelope design for a preferred floor plan and different objectives. The possible objectives are: (1) maximization of total floor area, (2) maximization of SA performance of the new building envelope; and (3) minimization of incident solar radiation in the warm season of the façade area that fulfill the minimum sun hours according to the SA ordinance. Two urban areas, one in Tallinn and another in Warsaw are considered as cases study. The outputs of this workflow are the optimal method to generate the SE with SET, the number of floor plan divisions, orientation/size of the building and zone of the façade that fulfills the SA requirements for a given multi-objective criteria. Implicitly, this workflow also provides flexibility in the creation of the interior floor plan of the building.

2 METHODOLOGY

In order to study how SEs generated by different criteria influence the building envelope design and use this information to optimize the performance of the new building, the workflow shown in Figure 1 is defined. This workflow has 4 main phases: SEs generation using SET, building envelope generation (using in-built Grasshopper components and Python functions), solar analysis of the building envelope using SET and the assessment of the building performance based on multi-objective criteria.

2.1 SEs generation with SET

The first step in the SE generation using SET is to extract climate information about the location of the case study (latitude, longitude, UTC and hourly direct normal solar radiation for the annual period) using the component "Location Data Reader". After this, it is necessary to calculate the sun vectors of the analysis period using the component "Analysis Period" (SunVectGen). Using SunVectGen, we can introduce a specific time step and multiple analysis periods. When the selected time step is different than 1 (hourly), SunVectGen uses a linear interpolation to calculate the direct normal solar radiation in each time step.
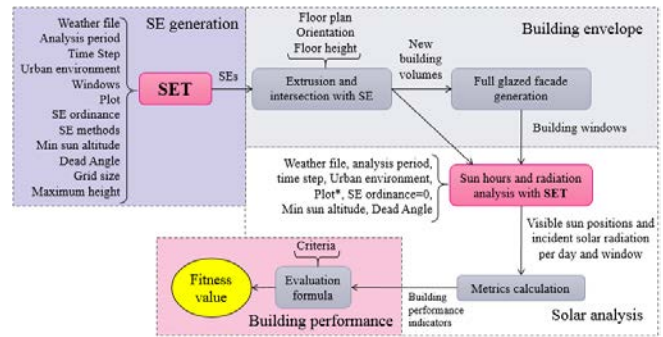
Figure 1. Workflow for the building performance assessment.

The main feature of SET is the possibility to filter and select sun vectors that fulfill different conditions. The filtering process consists of identifying the sun vectors of the analysis period that are not blocked by the surrounding buildings (Context) and fulfill the minimum defined solar altitude (MinAltitude) and dead angle (DeadAngle). This filtering process is run for every existing window introduced as input. Once the filtered sun vectors are identified, a sorting and selection process can be performed taking into account the specific SA ordinances at hand and sorting methods. The SA ordinance of each case study determines the SA minimum requirements for the new building. The sorting methods that can be selected through the input SunVectSel, allows taking into account different sun light quality criteria (Table 1). Consequently, the sorted and selected sun vectors can be used to generate the SE using the component "Solar Envelope Generator" (SolEnvGen) which gives as output: (1) the points of the final SE; (2) the SE points for each window. Additionally, the sun vectors filtered and selected for each window can be visualized using the component "Sun Path Generator" (SunPathGen).

VectorsType	Description
1	Larger solar altitude first
2	Larger solar incident radiation first
3	Sun vectors out the plot with large solar altitude first, then vectors through the plot with large solar altitude first
4	Sun vectors out the plot with large solar altitude first and then vectors through the plot with larger incident solar radiation first
5	Sun vectors out the plot with large solar altitude first and then vectors through the plot close to the corners of the plot first.

Table 1. Available sorting methods in SunVectSel.

The workflow has been tested on two existing urban areas: one located in Tallinn, Estonia (59°23'52.6"N 24°41'04.6"E)

(Figure 2) and another in Warsaw, Poland (52°16'42.6"N 20°55'14.5"E) (Figure 3). An identical plot 90x90 m in size is considered to build a new residential building in both locations. According to the Estonian Standard for daylight in dwellings, a new building cannot decrease SA of existing apartments more than 50% of actual sun light hours, and the new apartments must have at least 2.5 hours per day of direct sun light in at least one room. Both requirements need to be fulfilled every day of the period from 22th April to 22th August [12]. The SA requirements in Poland are similar for existing and new residential buildings: each apartment must have at least 3 hours of direct sun light on 21th March and 21th September between 07:00 to 17:00 [17].
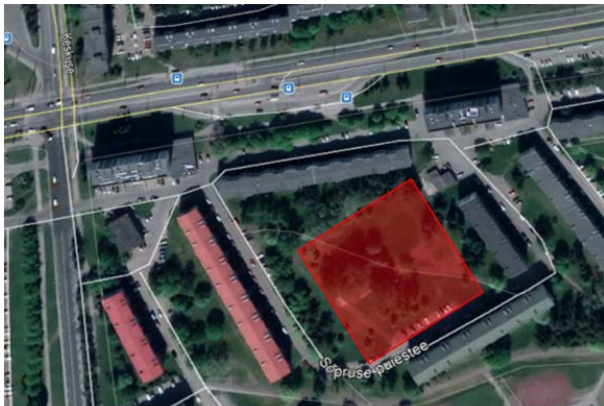
Figure 2. Real (up) and virtual (down) urban environment for Tallinn case study [19].

The choice of a suitable time step is essential for the accuracy of the generated SE and the required computational time [6]. Thus, a time step of 2 and 30 are used for Tallinn and Warsaw, respectively. The minimum angle between a sun ray and the building facade (window dead angle) is set to 10 degrees for both urban areas [3]. The minimum sun altitude is set to 12 and 10 degrees for Tallinn and Warsaw, respectively [11]. All the available sorting methods are taken into account for both urban areas. The distance between grid points on the buildable plot is set to 9 meters in order to achieve a good balance between SE accuracy and computational time cost [7]. Finally, the maximum buildable height was set 30 and 50 meters for Tallinn and Warsaw, respectively.

2.2 Building envelope generation

Once the different SEs have been created, the new building volumes can be generated extruding a selected floor plan.

The following step is to generate optimal windows for the new building facades. Nowadays, facades with high window to wall ratios (WWRs) are a common solution at northern latitudes due to the lack of sun hours during the winter and the importance of the view to the outside. The component used to select the sun vectors requires windows as simple surface geometry, which are used for sun hours and radiation analysis. The distance between window centroids is set to 3 m. The floor plan considered for the main cases study is a linear floor plan typology defined by ASHRAE [10] with different number of divisions (3, 4 and 5 divisions) (Figure 4). Moreover, different orientations (from south to south-west every 45 degrees counterclockwise) are chosen in order to study its effect on the building performance.
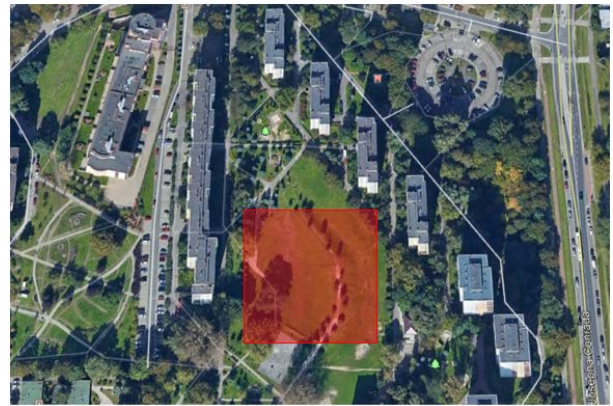
Figure 3. Real (up) and virtual (down) urban environment for Warsaw case study [19].
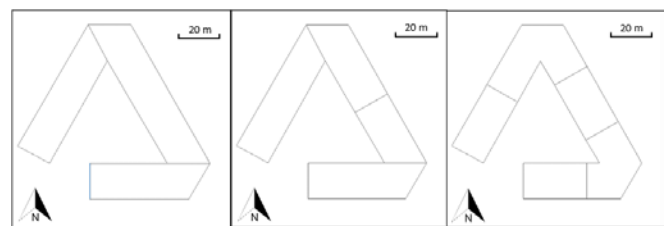
Figure 4. Different number of floor plan divisions (south oriented): 3, 4 and 5 (from left to right).

2.3 Sun hours and radiation analysis

The third phase consists of sun light hours and incident solar radiation analysis of the new building. The objective of this part of the proposed workflow (Figure 1) is to study how the urban environment affects the SA and solar radiation received by the new building. In this analysis, surroundings and new building must be provided as input in SunVectSel

component as urban context. Windows created in the previous phase are also used. To use SunVectSel for this sun hours and radiation analysis is necessary to define an auxiliary plot different from the plot used for the main cases study defined in the first phase of the workflow since we are not interested in to generate any SE in this step. However, the rest of the settings (dead angle, minimum solar altitude, weather files, time steps, etc.) are the same of those defined in the first phase of the workflow.

The outputs of this phase are the building performance indicators calculated in relation to the visible sun positions and the incident solar radiation for each time step and window. These performance indicators are:

- Volume ratio (V_r) (-) as the relation between the total volume of the new building and the volume of the SE,

- Window ratio (W_r) (-) as the relation between the number of windows of the new building that fulfill the SA requirements and the total number of windows,

- Mean incident solar radiation ($mISR$) (W/m^2) as the mean (for the analysis periods defined in section 2.1) incident solar radiation received by the windows of the new building that fulfill the SA requirements.

Thus, knowing V_r, W_r and $mISR$ for the new building we can study how its volume is fitting the SE volume, which ratio of the total facade area is fulfilling the SA requirements and the level of solar radiation exposure of the facade.

2.4 Building performance assessment

In this last phase of the workflow, the building performance indicators are used to calculate the objective variables, which are the variables of the proposed objective linear function F (1) that represent the building performance. Thus, it is possible to compare building designs through the value F (fitness value). The objectives of the evaluation formula (1) are: maximization of V_r' (OV) (2); maximization of W_r' (OW) (3); minimization of $mISR$ (OI) (4). The values of these objective variables are between 0 to 1 (5) and they represent the level of fulfillment of the building i within a set of n possible new buildings (1, 2,..., i,..., n).

$$F(Vr_i', Wr_i', mISR_i') = \alpha Vr_i' + \beta Wr_i' + \gamma mISR_i' \quad (1)$$

$$Vr_i' = Vr_i / \max(Vr_i) \quad (2)$$

$$Wr_i' = Wr_i / \max(Wr_i) \quad (3)$$

$$mISR_i' = 1 - mISR_i / \max(mISR_i) \quad (4)$$

$$\alpha + \beta + \gamma = 1 \quad (5)$$

The coefficients α, β and γ are the weight factors whose values depends on the chosen criteria. For the main case studies, different criteria were taken into account to evaluate the fitness value of each building design (Table 2). The fitness value is always between 0 and 1. Value 1 is associated to an ideal building performance and that would mean that the new building fulfill all the objectives previously mentioned at the same time.

	OV(α)	OW(β)	OI(γ)
C1	1	0	0
C2	0	1	0
C3	0	0	1
C4	0.5	0.5	0
C5	0.5	0	0.5
C6	0	0.5	0.5
C7	0.3	0.3	0.3
C8	0.7	0.3	0
C9	0.3	0.7	0
C10	0	0.7	0.3
C11	0	0.3	0.7

Table 2. Coefficients of the objective function for different criteria. OV=Volume maximization; OW=Solar access maximization; OI=Solar incident radiation minimization.

The workflow is applied to two case studies (Warsaw and Tallinn), different sorting methods (Table 2.1), one floor plan typology with different number of divisions (3, 4 and 5) and 8 different orientations (from South to North-West every 45 degrees). Moreover, these 240 cases are evaluated according to 11 different multi-objective criteria (Table 2.4) in order to evaluate which criteria offers the best trade-off building performance in each case study.

3 RESULTS

In this section, the dependency between the design parameters (sorting method (SM), number of floor plan divisions, orientation and multi-objective criteria) and the building performance indicators (V_r, W_r and $mISR$) are analyzed.

3.1 SE volumes

The volumes of the SEs after considering different SMs can be seen in Figure 5. In general, SMs 1, 3 and 5 allow larger SE volumes than SMs 2 or 4. This fact was expected because the sorting methods 1, 3 and 5 are based on the selection of sun vectors with larger solar altitude and/or whose projection are outside the plot. On the other hand, SMs 2 and 4 take into account sun vectors with high incident solar radiation which normally have lower sun altitude for the surrounding façade facing east and west, limiting the SE volume.

In Warsaw case, the SE volume is between 136770 m^3 and 120793 m^3. In Tallinn case, the maximum SE volume is 153975 m^3 (SM 3) and the minimum is 78510 m^3 (SM 4). Furthermore, the SE volume is more sensitive to the SMs used in Tallinn than in Warsaw case. This is due to the combined effect of different aspects as surrounding urban environment, plot orientation, latitude, solar radiation,

analysis period, time step, dead angle and minimum solar altitude (Figure 6).
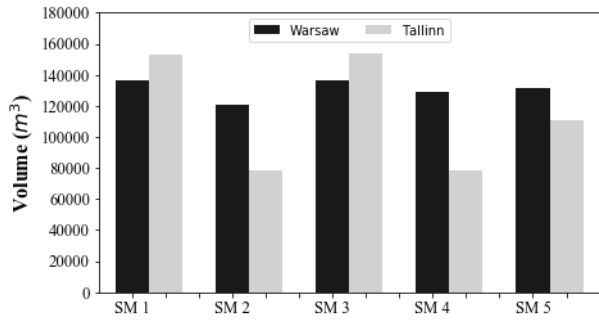
Figure 5. Volume of the SEs using different sorting methods for the cases study.
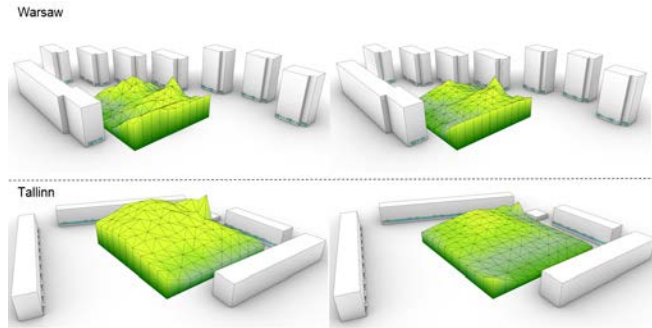
Figure 6. Examples of the generated SEs for both cases study using SM 1 (left) and SM 2 (right).

3.2 Warsaw case

The values of the objective variables for different criteria used in Warsaw case are shown in Figure 7. Different groups of criteria gives the same optimal solution. The first optimal solution is generated by volume-weighted (C1, C5 and C8) and trade-off (C7) criteria. The second optimal solution is obtained regarding SA and $mISR$-based criteria (C2, C3, C6, C9, C10 and C11).The third optimal solution is generated by the criterion C4 that is based on 50% V_r and 50% W_r. In addition, the mean incident solar radiation is minimal in these three solutions ($mISR'$=1.0). The first solution has maximum volume and the second one has maximum SA. Moreover, the third solution is a trade-off between both objectives.

The design parameters and the performance indicators for each optimal solution can be seen in Table 3. South-West is an optimal orientation in Warsaw case (Figure 8). The first and the third solutions have 5 number of floor plan sections while the second one has only 4 divisions. SM 2 is not recommended for any criteria. The consideration of SMs 1, 3 and 5 allow the maximization of SA (W_r=0.63) when the floor plan has 4 divisions (Figure 8b). However, if the number of divisions is increased to 5, the use of SM 5 allow the maximization of the volume (V_r=0.21) (Figure 8a). Indeed, the use of SM 3 and 5 divisions for the floor plan allow a good trade-off between volume (V_r=0.2) and SA (W_r=0.58) (Figure 8c). The mean incident solar radiation

for all the solutions is around 100 W/m². Furthermore, the best trade-off solution can be either the first or the third one since there is no much difference in absolute terms for the volume and SA indicators (V_r and W_r).
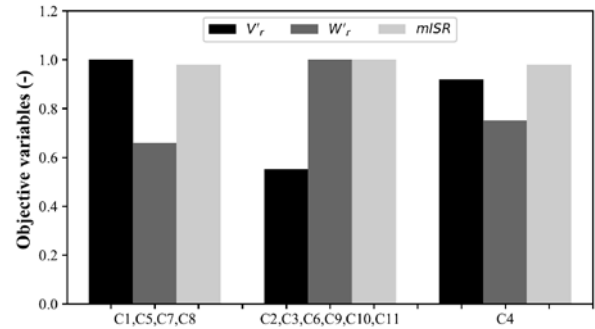
Figure 7. Objective variables values for different criteria in Warsaw case.

Criteria	nFPd	Ori.	SM	Vr	Wr	mISR
C1,C5,C7,C8	5	SW	5	0.21	0.57	100.58
C2,C3,C6,C9, C10,C11	4	SW	1,3,5	0.17	0.63	100.03
C4	5	SW	4	0.20	0.58	100.71

Table 3. Design parameters and performance indicators using different criteria for Warsaw case.
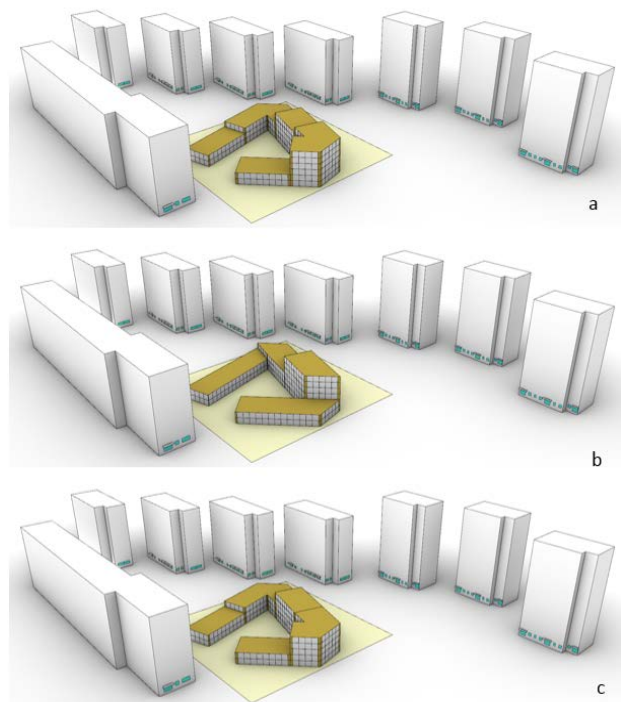
Figure 8. Optimal combinations in Warsaw case.

3.3 Tallinn case

The values of the objective variables for different criteria used in Tallinn case can be seen in Figure 9. In these case, five different group of criteria give 5 different optimal solutions:

- The first optimal solution is generated by the 100\% volume-weighted criterion (C1),

- The second optimal solution is a 100\% SA-based criterion (C2),

- The third optimal solution is generated by a group of criteria based on different trade-off considerations (C3, C5, C7, C11),

- The fourth optimal solution is generated by a group of criteria based on different trade-off between volume and SA (C4, C8, C9),

- The fifth optimal solution is generated by a group of criteria based on different trade-off between SA and $mISR$ (C6, C10).
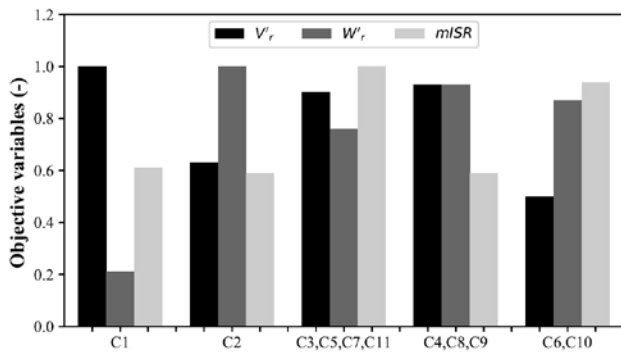
Figure 9. Objective variables values for different criteria in Tallinn case.

The design parameters and the performance indicators for each optimal solution are shown in Table 4. In general, SM 2 and 4 are not recommended for any criteria. The maximization of the building volume (V_r=0.24) is possible by orienting west a floor plan with 5 divisions and considering SM 5 to generate the SE (Figure 10a). Nevertheless, the combination of the SM 1 or 3 with 3 divisions of the floor plan and NW orientation, achieve a low volume fitting with the SE volume V_r=0.18) but high performance in terms of SA (W_r=0.87) and $mISR$ (190.14 W/m^2) (Figure 10e). Moreover, the south orientation combined with a low number of divisions of the floor plan (3 and 4) are associated to high W_r (between 0.84 and 0.97) when SM 1 is considered (Figure 10b 12d). The only difference between the fifth and the third optimal solution is the number of divisions of the floor plan (4 and 3, respectively). This increment from 4 to 3 divisions of the floor plan generates a self-shadowing phenomenon: the 3% of the whole facade pass to do not fulfill the SA requirement because V_r increases a 5% (Figure 10c 12e).

Finally, the third (Figure 9c) and fourth solutions show good trade-off performance. Nevertheless, since the deviations in terms of $mISR$ are around 10 W/m^2, the fourth optimal solution is the best one (Figure 9d): south-oriented building whose floor plan is divided by 4 sections (SM 1 used for the SE generation) achieve a V_r of 0.23 and the 89% of the facade is fulfilling the SA requirements with a mean incident solar radiation of almost 198 W/m^2.

Criteria	nFPd	Ori.	SM	Vr	Wr	mISR
C1	5	W	5	0.24	0.66	197.39
C2	3	S	1,3	0.19	0.92	197.7
C3,C11,C5,C7	4	NW	1,3	0.23	0.84	188.78
C4,C8,C9	4	S	1	0.23	0.89	197.89
C6,C10	3	NW	1,3	0.18	0.87	190.14

Table 4. Design parameters and performance indicators using different criteria for Tallinn case.
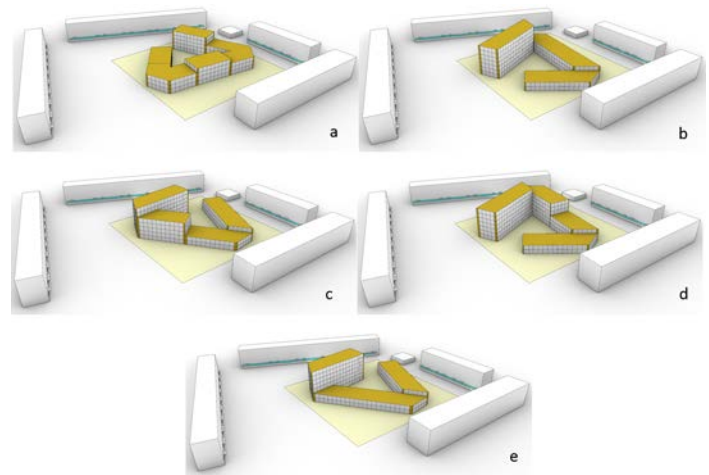
Figure 10. Optimal combinations in Tallinn case.

4   CONCLUSION

Solar access is one of the key criteria for visual comfort that can also affect the thermal comfort and energy efficiency in buildings. There are different ordinances that define the minimum SA requirements for new and existing residential buildings. These requirements are related to the quantity not the quality of direct sun light. Different SEs might fulfill the same SA requirements for a case study. This fact provide flexibility in the early stages of the architectural design process. Previous research just only focused on one SE per urban area.

The aim of this paper is to study the potentiality of different SEs in building envelope design in cold climates. The methodology consists of a workflow based on multi-

objective optimization. The objectives used are: maximizing the volume of the building, maximizing the ratio of the facade that fulfills the SA requirements and minimizing the solar radiation received by a portion of the facade. Thus, this workflow is used to design a new residential building in two locations with different SA ordinances: Warsaw (Poland) and Tallinn (Estonia). The same buildable plot and floor plan typology are used for both case studies. Moreover, the design parameters used are: the number of floor plan divisions, the orientation of the new building, the sorting method of sun vectors for SE generation and multi-objective criteria. Case study consolidates an insightful comparison of various optimization criteria.

In general, considering a large number of possible floor plan divisions, the volumes of the new building can fit better SE volume. Moreover, the consideration of sorting methods based on larger solar altitudes and sun vectors outside the plot achieve larger SE volume than those based on incident solar radiation. Moreover, the maximum relative difference in terms of SE volume is almost 50% and 12%, for Tallinn and Warsaw case, respectively. Indeed, this difference can be due to the length of the analysis periods.
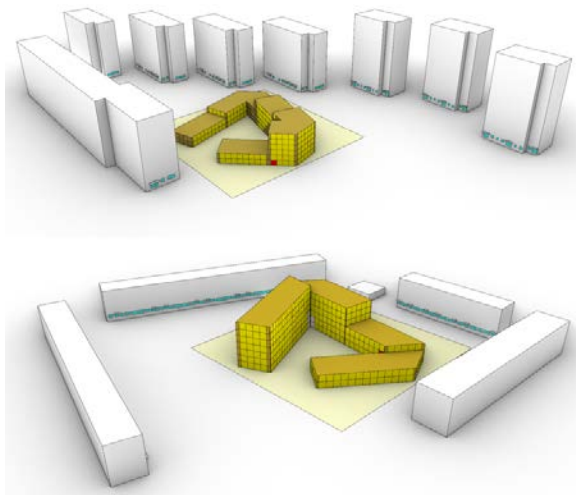
Figure 11. Optimized building envelopes for Tallinn (down) and Warsaw (up) cases. Windows in yellow fulfill the SA requirements and the red window has the maximum mean incident solar radiation.

A suitable choice of the design parameters is critical to optimize the new building envelope. If the preferred criteria to design the new residential building is the volume ratio and the SA of the facades, the use of criteria based on volume-weight for Warsaw and trade-off criteria between volume and SA for Tallinn are suitable ones. The optimized building envelopes for both case studies is shown in Figure 11. The optimal number of floor plan divisions are five and four for Warsaw and Tallinn cases, respectively. The total floor area is 8930 m^2 and 11617 m^2 in Warsaw and Tallinn, respectively. South-west (Warsaw) and south (Tallinn) orientations allow good performance between volume and

SA of the new buildings (Figure 11): volume ratio higher than 20% and ratio of the facade that fulfill the SA requirements of 57% and 89%, for Warsaw and Tallinn, respectively.

This research presents a useful workflow to be used in the early stages of the design of residential buildings in cold climates. How to take advantage of the SA requirements to optimize the building performance using different multi-objective criteria is investigated. This workflow is especially suitable when the SA regulations are strict. First decisions as building orientation, windows location, building footprint and buildable volumes characterize the solar access of the existing and new buildings. Other requirements related to daylight provision, overheating, energy consumption or glare protection could be fulfilled with a suitable design of interior layout and façade. In this paper, different criteria are considered and all the possible building envelope are evaluated and the maximum fitness value searched. Future work is to apply this workflow in a larger optimization system (to evaluate different floor plan typologies) based on genetic algorithm (GA). Additionally, this method will be implemented in a Grasshopper tool for Rhinoceros to help architects and designers to choose between possible building design strategies.

REFERENCES

1. Capeluto, I. G., and Plotnikov, B. A method for the generation of climate-based, context-dependent parametric solar envelopes. Architectural Science Review 60, 5 (2017), 395–407.

2. Capeluto, I. G., and Shaviv, E. Modeling the design of urban fabric with solar rights considerations. Israel: Faculty of Architecture and Town Planning Technion-Israel Institute of Technology Haifa. Recuperado de http://www. ibpsa.org/proceedings/BS1999/BS99 C-22. pdf (1997).

3. Darula, S., Christoffersen, J., and Malikova, M. Sunlight and insolation of building interiors. Energy Procedia 78 (2015), 1245–1250.

4. De Luca, F. From envelope to layout-buildings massing and layout generation for solar access in urban environments. Proceedings of Sharing Computational Knowledge! ShoCK! 35th International Conference on Education and research in Computer Aide Architectural Design in Europe, eCAADe Vol. 2 (2017), 431–440.

5. De Luca, F. Solar form finding: Subtractive solar envelope and integrated solar collection computational method for high-rise buildings in urban environments.

Proceedings of Disciplines and Disruption - 37th Annual Conference of the Association for Computer Aided Design in Architecture, ACADIA 2017, Cambridge (MA), United States (2017), 212–221.

6. De Luca, F., and Dogan, T. A novel solar envelope method based on solar ordinances for urban planning. Building Simulation: An International Journal 12, 5 (2019), 817–834.

7. De Luca, F., Nejur, A., and Dogan, T. Facade-floor-cluster-methodology for determining optimal building clusters for solar access and floor plan layout in urban environments. Proceedings of Computing for a better tomorrow. 36th International Conference on Education and research in Computer Aided Architectural Design in Europe, eCAADe, Lodz, Poland Vol. 2 (2018), 585–594.

8. De Luca, F., and Voll, H. Computational method for variable objectives and context aware solar envelopes generation. 8th Annual Symposium on Simulation for Architecture and Urban Design, SimAUD 2017, Toronto, Canada Vol. 11 (2017), 281–288.

9. De Luca, F., and Voll, H. Solar collection multi-isosurface method - computational design advanced method for the prediction of direct solar access in urban environments. Communications in Computer and Information Science, 17th International Conference on Computer-Aided Architectural Design, CAAD Futures 2017, Istanbul, Turkey 7 (2017), 170–187.

10. Dogan, T., Saratsis, E., and Reinhart, C. The optimization potential of floor-plan typologies in early design energy modeling. Proceedings of BS2015: 14th Conference of International Building Performance Simulation Association, Hyderabad, India, International Building Performance Simulation Association (IBPSA) (2015).

11. European commission, EN 17037:2018 Daylight in buildings.https://velcdn.azureedge.net/~/media/marketing/ee/professional/28mai2019%20seminar/veluxen17037tallinn28052019.pdf

12. Daylight in Dwellings and Offices. 894:2008/A2:2015. https://www.evs.ee/products/evs-894-2008+a2-2015.

13. Knowles, R. L. Sun rhythm form. Cambridge, MA, USA: MIT Press. (1981).

14. Ladybug Tools 2019. https://www.ladybug.tools/ladybug.html.

15. Samuels, R. Solar efficient architecture and quality of life: The role of sunlight in ecological and psychological well-being. In Proceedings of the 1st World Renewable Energy Congress: Energy and the environment, Pergamon Press Oxford, UK (1990), 2653–2659.

16. Sanaieian, H., Tenpierik, M., van den Linden, K., Seraj, F. M., and Shemrani, S. M. M. Review of the impact of urban block form on thermal performance, solar access and ventilation. Renewable and Sustainable Energy Reviews 38 (2014), 551–560.

17. Sokol, N., and Martyniuk-Peczek, J. The review of the selected challenges for an incorporation of daylight assessment methods into urban planning in poland. Procedia engineering 161 (2016), 2191–2197.

18. Solar toolbox 2019. https://www.food4rhino.com/app/solar-toolbox.

19. Google Earth Pro. 2019. https://www.google.com/intl/es/earth/versions/#earthpro

20. Solemma LLC. DIVA. 2019. https://www.solemma.com/Diva.html.

21. Stasinopoulos, T. N. A survey of solar envelope properties using solid modelling. Journal of Green Building 13, 1 (2018), 3–30.

Urban Space Simulation Based on Wave Function Collapse and Convolutional Neural Network

Bo Lin[1, 2], Wassim Jabi[2], Rongdan Diao[1]

[1]Wenzhou University
Wenzhou, China
diaorongdan@126.com

[2]Cardiff University
Cardiff, United Kingdom
{LinB, JabiW}@cardiff.ac.uk

ABSTRACT

In this paper, we propose a pipeline of urban space synthesis which leverages Wave Function Collapse (WFC) and Convolutional Neural Networks (CNNs) to train the computer how to design urban space. Firstly, we establish an urban design database. Then, the urban road networks, urban block spatial forms and urban building function layouts are generated by WFC and CNNs and evaluated by designer afterwards. Finally, the 3D models are generated. We demonstrate the feasibility of our pipeline through the case study of the North Extension of Central Green Axis in Wenzhou. This pipeline improves the efficiency of urban design and provides new ways of thinking for architecture and urban design.

Author Keywords

Urban space synthesis; Wave Function Collapse (WFC); Convolutional Neural Networks (CNNs).

ACM Classification Keywords

J.6 COMPUTER-AIDED ENGINEERING: Computer-aided design (CAD)

1 INTRODUCTION

Around 54% of people in the world live in the urban areas in 2014 and the number of 66% is predicted in 2050 [24]. Modern cities exhibit increasing complexities and dynamics which demand for a fast adaption of urban design approach. However, traditional urban design method is still static, sectorial and time-consuming [18]. Thus, the computational generative urban design has become a hot research topic in recent years. The automatic generation has been addressed using approaches of procedural modeling and example-based modeling. Procedural approaches are based on the manually set design grammar or rule to generate design. In contrast, example-based approaches inspect examples and decompose them into a set of pieces of blocks, followed with a reorganization through algorithms, such as WFC, and the style of output is matched with the example statistically and perceptually [5]. Moreover, Nowadays, the artificial intelligence (AI) is developing rapidly and the application of deep learning technique for generation has been investigated. CNNs are utilized in the areas of computer graphics, such as texture synthesis.

In this work, we put forward an example-based approach for urban space synthesis which leverages the technique of WFC and CNNs. The first step is the establishment of urban design database. The second step utilizes WFC and CNNs to generate urban road networks, urban block spatial forms and urban building function layouts, followed with the evaluation and selection. The last step is the generation of 3D models.

The technique of WFC and CNNs are quite suitable for envisioned task as they learn the distribution underlying a set of images. With these techniques, the manual set of rules or parameters tuning which is tedious for non-experts will be no longer needed for generative design. The advent of AI techniques in urban space synthesis is still in early period, but it offers promising results. More than a simple opportunity, the potential of WFC and CNNs for the urban space synthesis is a major step ahead.

2 RELATED WORK

The first and second part of this section reviews procedural and example-based approaches focusing on urban space, texture and model synthesis. In the third part, we briefly review content generation approaches which leverages the technique of deep learning.

2.1 Procedural approach

In a comprehensive survey on procedural and inverse procedural modelling, Smelik et al. [23] survey procedural methods generating features of virtual worlds, including terrains, vegetation, rivers, roads, buildings, and entire cities. In general, procedural methods rely on the use of manually defined or automatically determined rule sets for content generation. Neil Leach and Weiguo Xu [12, 13] discuss the digital techniques for architecture in young architects' works and student architectural design works. Feng Yuan et al. [27] survey the different digital workshops in China. Yi Shi and Meipo Kwan [22] introduce the development of digital city in three American universities, namely, Massachusetts Institute of Technology, University of Illinois at Urbana-Champaign and University of Oregon. Ying Long [15] discusses the different approaches and applications of big data in urban planning and design and puts forward an approach of self-feedback urban design

based on urban sensors and online platform. Reinhard Koenig [10] focuses on the evolutionary many-criteria optimization to improve urban design synthesis for building layout. In Yufan Miao et al. [18], the focus is on computational urban design prototyping. Decoding Space is used to generates urban layouts including street network, parcels and buildings according to designers' requirements.

2.2 Example-based approaches

Compared with procedural approaches, example-based methods do not need manual rules to generate content. It relies on analyzing the data such as urban morphology or road network to extract template or statistical information. Zheng Xie [26] explores the potential use of space syntax analysis in 3D parametric design. He extracts the correlation between urban morphology information and integration and generate urban form based on this parameter relationship. In Nishida et al. [19], road patches are extracted from real road networks and a new road network is generated. The terrain is taken into account to attach road patches to connector streets from an initial seed point. Heeger and Bergen [6], Portilla and Simoncelli [21] all try to match the large-scale random attributes of texture samples with the new texture. Vivek Kwatra et al. [11] use time as the third dimension to extend the texture synthesis and generate three-dimensional or multi-dimensional texture synthesis. Paul C. Merrell [17] simplifies texture synthesis technology by selecting samples with random shapes and then synthesizing larger texture output. Based on Paul C. Merrell's model synthesis study in 2009, the "Wave Function Collapse" algorithm by Python is published, which can generate similar images according to the input image [16].

2.3 Learning-based approaches

With the development of deep learning techniques, it has been used for procedural and data driven generation. In Yumer et al.'s research [28], the features of a low-dimensional generative model from high-dimensional procedural models is learned and new models can be generated. Nishida et al. [20] focus on interactive building modelling. CNNs are trained to classify the type of a rule snippet and to regress parameter sets. Then, the building mass, roof, etc. can be sketched iteratively. Results are classified and the parameters are inferred by CNNs. Stefan Hartmann [5] publishes StreetGAN. The road network is generated relying on GAN. A similar approach is leveraged by Adrian Albert et al. and he models urbanization patterns based on GAN. Stanislas Chaillou [2] puts forward an ArchiGAN which is a generative stack for apartment building design. A Pix2Pix GAN-model is trained to generate floor plan. Varshaneya V [25] leverages GANs to sketch in vector format and compare the generation results between GAN architecture SkeGAN and a VAE-GAN architecture VASkeGAN.

3 REVIEW OF WFC AND CNNS

3.1 WFC

WFC algorithm is widely used in the field of game and AI and it can randomly generate map according to custom rules. Similar images can be generated based on the input image by WFC algorithm. This algorithm is published on the website of github by Maxim Gumin [3]. He defines local similarity related to the following two aspects. Firstly, each N (feature parameter) ×N pixel module of the output image appears at least once in the input image. Secondly, the probability of appearance of each N*N pixel module in the output image is similar with the probability of appearance in the input image. Different styles of generated images can be generated by modifying the feature parameter N. There are five steps in this algorithm. Firstly, WFC learns the input and count N×N patterns. Secondly, an array with the output dimensions is created. The elements in the array represent the states of the outputs in region of N×N. Thirdly, the wave in the completely unobserved state is initialized. Fourthly, there is a cycle of observation and propagation. Observation includes the search of a wave element with minimal nonzero entropy and the collapse of the element into a particular state depending on coefficients and distribution of input. Propagation means propagating information observed. The cycle is repeated until all the wave elements are with zero or undefined entropy. Fifthly, if the wave element is in the observed state, then output is generated. If the wave element is in the contradictory state, then the work is ended without output.

The algorithm is mainly based on the discrete model synthesis study by Paul Merrell. He explores to propagate adjacency constraints in a simple tiled model [16]. This algorithm is also affected by declarative texture synthesis proposed by Paul Francis Harrison. He explains the adjacency data of tiles through the labels of these tiles' borders and fill the tilemap by backtracking search [4]. Isaac Karth and Adam M. Smith formulate WFC as an instance of constraint solving method for procedural content generation with multiple in-the-wild users in 2017 [8] and develop a system based on WFC trained on positive and negative examples and discuss this system in a general context with example based generators in 2019 [7]. Marian Kleineberg uses WFC algorithm to procedurally generate city based on the tiled model for Procjam 2018 [9]. It is the first step of application of WFC in the urban space generation.

3.2 CNNs

CNNs is a well-known deep learning technique in recent years. It is widely applied in the field of image recognition. Zhang Fang et al. [29] utilize CNNs to measure human perceptions towards a large-scale urban region. Zhou Bolei et al. [30] propose a novel method to recognize city identity through analysis of geo-tagged images and discuss the application of computer vision in urban planning. In 2016, Alex J. Champandard [1] proposes a semantic style transfer approach and turn two-bit doodles into fine artwork by

CNNs. In 2017 Li Chuan et al. [14] put forward the image synthesis method combining Makov random field and convolutional neural networks. This algorithm can be run by both GPU and CPU. GPU synthesis speed is much faster than CPU synthesis. This algorithm uses the several images as inputs to generate a new image and it depends on the original image style, original image's annotation and the target content image with its annotation. This algorithm extracts patch information from the style image and then transfers it to the target image according to the match degree. The matching effect is related to the number of iterations and computation speed. Within a certain range, the higher the number of iterations, the better the match, the longer the calculation time, but the matching effect will not increase after the number of iterations increases to a certain degree. Different number of iterations will result in different effects towards generated images.

4 URBAN SPACE SYNTHESIS BASED ON WFC AND CNNS

In this research, relevant theories and methods such as urban morphology, WFC, CNNs and urban design evaluation are applied. This research establishes a relatively well-developed urban design database, leverages WFC and CNNs for urban space synthesis, establishes the urban design evaluation system, and achieves the final urban design scheme through the interaction between designers and the pipeline (figure 1).

4.1 Establishment of urban design database.

The urban design database is established according to the international typical big data of urban space. The composition pattern of its spatial form is analyzed, contributing to basic type of the spatial prototype, which can be roughly divided into multi-center spatial pattern, grid pattern, annular radiation pattern and axis pattern. According to the classification of urban design cases, various urban design schemes are collected to establish the database. The urban design database includes three sub-datasets: urban road network database, urban block spatial form database and urban building function layout database.

The data of urban road network system is reflected by 2D maps of roads in different levels. The data of urban block spatial form is reflected by 2D maps with different colors to distinguish the square, green space and buildings in different heights. The data of urban building function layout is reflected by 2D maps that show different building functions by different colors.
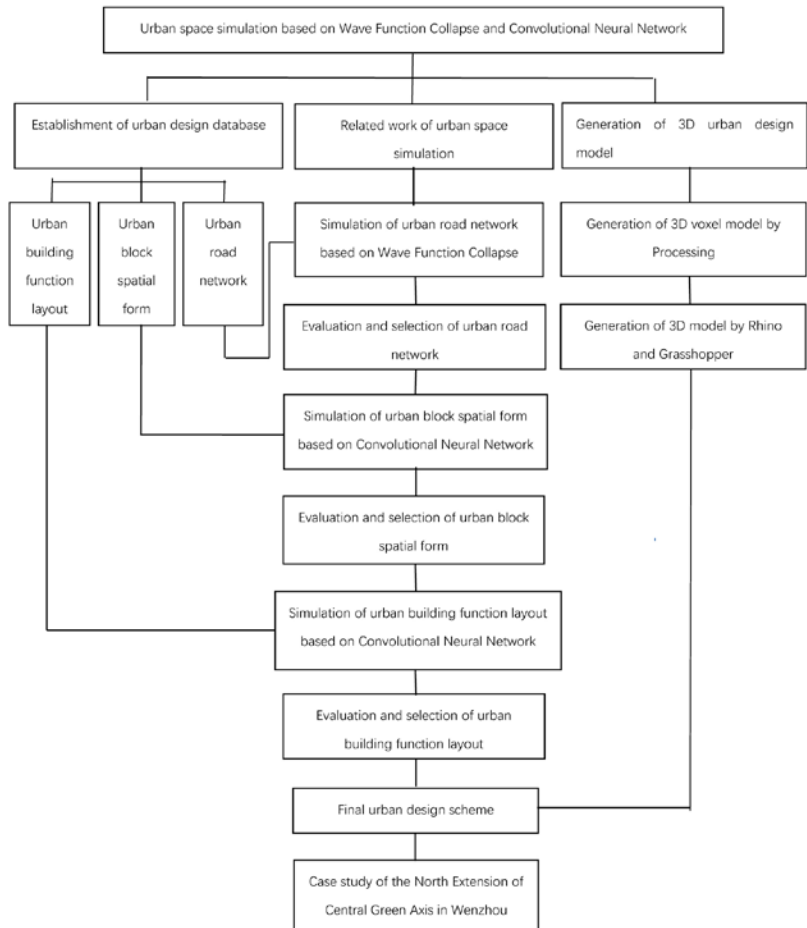
Figure 1. Research framework.

4.2 Urban space synthesis and scheme evaluation and selection.

The urban road network in the urban design database is the data source for synthesis by WFC and the characteristic parameter N is adjusted to generate multiple road networks. Then the generated road networks are evaluated and the optimal urban network is selected. Afterwards, the urban block spatial form in the urban design database is extracted for synthesis by CNNs, and the number of iterations is adjusted to generate multiple urban block spatial forms, followed with evaluation and selection, contributing to the optimal urban block spatial form. Finally, the urban building function layout in the urban design database is extracted, and the number of iterations is adjusted to generate multiple urban building function layouts. They are evaluated and the optimal urban building function layout is selected.

The evaluation and selection of urban design schemes can be divided into three steps. The first step is the evaluation of the urban road network. The second step is the evaluation of the urban block spatial form, and the third step is the evaluation of the urban building function layout. The evaluation indicators of urban road network include the density of the road network, the accessibility of the road network, the capacity of the road network and the integration degree of surrounding sites. The evaluation indicators of urban block spatial form include rationality of urban landmark layout, aesthetic quality of urban skyline, construction land rate, adaptability to topography, greening rate, rationality of urban square, variety of building form and variety of public spatial form. The evaluation indicators of urban building function layout include the diversity of urban building function and the rationality of urban building function layout.

4.3 Generation of 3D urban design models.

The data in three sub-databases are all 2D images, and the urban design schemes generated by WFC and CNNs are also 2D images. In this part, the 2D images are automatically generated into 3D urban design model by algorithm.

The different color in the generated 2D urban spatial block form represent squares, greening and building with different heights. The algorithm written by the software of Processing moves the pixels in generated 2D urban block spatial form to different heights according to their colors, contributing to 3D voxel model, and the 3D coordinate data of each voxel are exported. Then, the data are imported into the software of Rhino and Grasshopper and the 3D voxel model are changed to the 3D block model, which is the final 3D urban design model.

Barcelona	Hangzhou Qianjiang CBD
Paris	Venice
Washington	Wenzhou Ancient Town
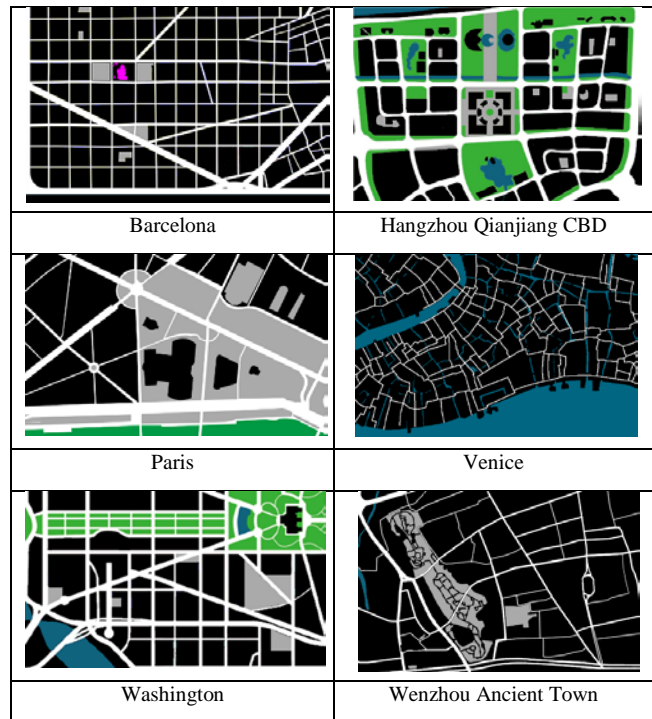
Table 1. Sub-database of urban road network

Barcelona	Hangzhou Qianjiang CBD
Paris	Venice
Washington	Wenzhou Ancient Town
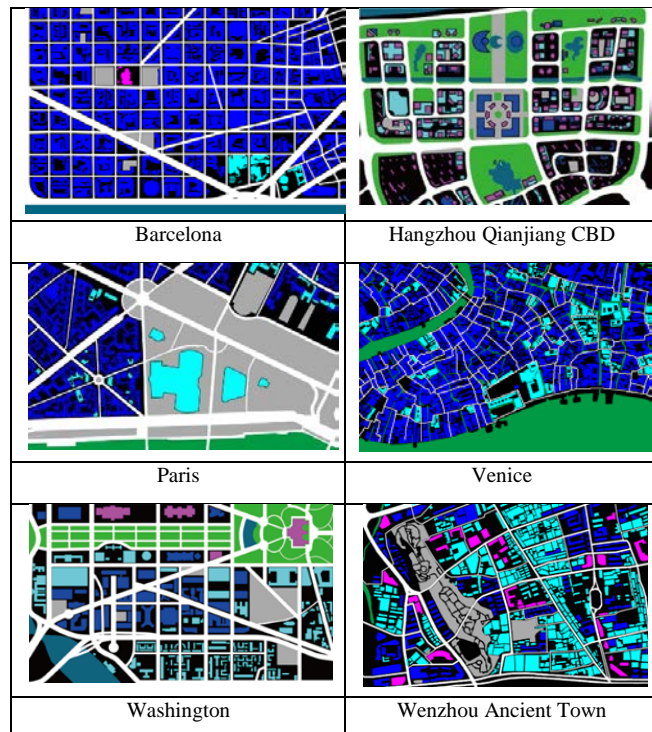
Table 2. Sub-database of urban block spatial form

142

5 CASE STUDY: THE NORTH EXTENSION OF CENTRAL GREEN AXIS IN WENZHOU

In order to showcase the versatility of the urban space synthesis approach. It is validated on the case of the North Extension of Central Green Axis in Wenzhou. The central green axis is the main landscape axis through the north and south of Wenzhou, and the center of public culture and public activities. In this paper, we establish an urban design database, leverage WFC, CNNs and Processing algorithm to generate road network, urban block spatial form, building function layout and 3D spatial model.

5.1 Establishment of urban design database

According to the design vision and context, there are 6 urban space cases in the database, namely, Washington, Hangzhou Qianjiang Central Business District (CBD), Wenzhou ancient town, Venice, Barcelona, Paris. The design vision is the extension of central green axis. The cases with clear axis are selected, such as Washington, Barcelona, Hangzhou Qiangjiang CBD and Paris. In order to strengthen the identity of Wenzhou, the case of Wenzhou ancient town is selected. The selection of Venice is due to the similar topography. Both Venice and Wenzhou are watery place. The urban design database is divided into three sub-databases. Each case contains three parts of information, namely, urban road network (table 1), urban block spatial form (table 2) and urban building functional layout (table 3). Urban road network is reflected on figure-ground map. Urban block form uses different colors to distinguish buildings of different heights, squares, green Spaces and other public Spaces. Building heights can be divided into the 5 categories, namely, 1~3 story buildings, 3~8 story buildings, and buildings with more than 8 stories. Urban building function layout are divided into the following 5 categories, namely, cultural and educational areas, commercial areas, residential areas, office areas, and mix use of commercial and residential areas. Different colors are used to distinguish the different types of urban building functions.

5.2 Urban road network synthesis by WFC and scheme evaluation and selection

The urban road network in the database is extracted as data source for synthesis by WFC. The characteristic parameter N (3, 4 and 5) is adjusted to generate 3 different road networks. For scheme 1, the density of road network is appropriate; the accessibility of road network is high; the fit of surrounding sites is high. For scheme 2, the density of road network is high, the accessibility is appropriate, and the fit of surrounding sites is low. For scheme 3, the density of road network is high; the accessibility is low; the fit of surrounding sites is low. Thus, the scheme 1 is the optimal road network scheme (figure 2).

5.3 Urban block spatial form synthesis by CNNs and evaluation and selection

Urban block spatial forms from urban design database are learned and CNNs is leveraged to generate squares, green

spaces and buildings with different heights. Each data source is learned in 3 different iterations (40, 60 and 80). The 6 cases are learned as data source separately and 18 schemes are generated. 2 of 6 cases are selected as data source to be learned afterwards. There are 15 case combinations totally and 45 schemes are generated. In other words, there are 21 different data sources to be learned and 63 schemes are generated. Then, in order to preserve the ecological structure, the constructions in Yangfu Mountain and Shangdoumen River in the schemes are replaced as green and blue. Finally, the urban block spatial form schemes are evaluated and selected. There are 8 evaluation indicators, namely, rationality of urban landmark layout, aesthetic quality of urban skyline, construction land rate, adaptability to topography, greening rate, rationality of urban square, variety of building form and variety of public spatial form. The highest score for each indicator is 10. The score is provided by urban designers. The final score is the sum of indicators above. Finally, 6 schemes whose final score in top 6 are picked out. These schemes are learned from the case of Wachingtom, the cases of Washington and Hangzhou Qianjiang CBD, the cases of Washington and Wenzhou ancient town, the cases of Washington and Paris, the cases of Hangzhou Qianjiang CBD and Wenzhou ancient town, the cases of Hangzhou Qianjiang CBD and Venice respectively.

Barcelona	Hangzhou Qianjiang CBD
Paris	Venice
Washington	Wenzhou Ancient Town
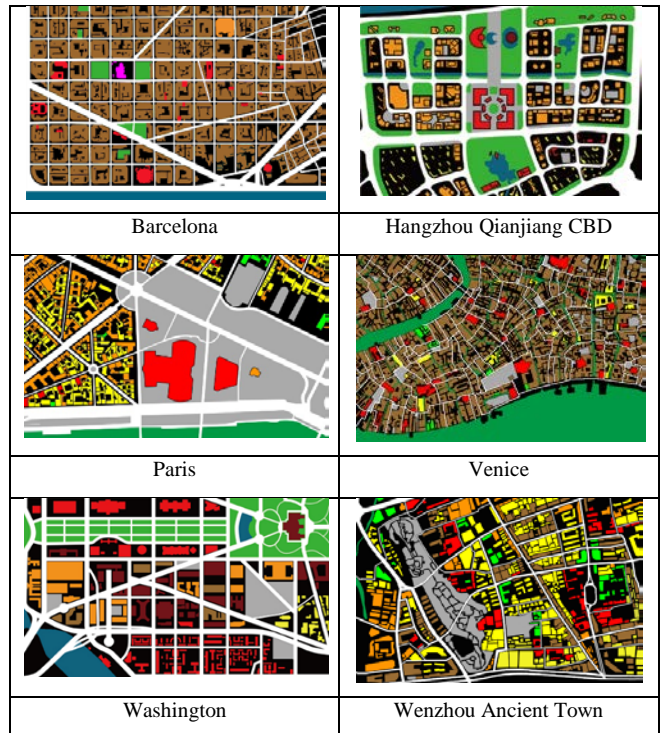
Table 3. Sub-database of urban building functional layout

5.4 Urban building function layout synthesis by CNNs and scheme evaluation and selection

The urban building function layout data of the 6 cases in the database are extracted and CNNs are leveraged to generate the building function layout. The iteration is adjusted as 40,

143
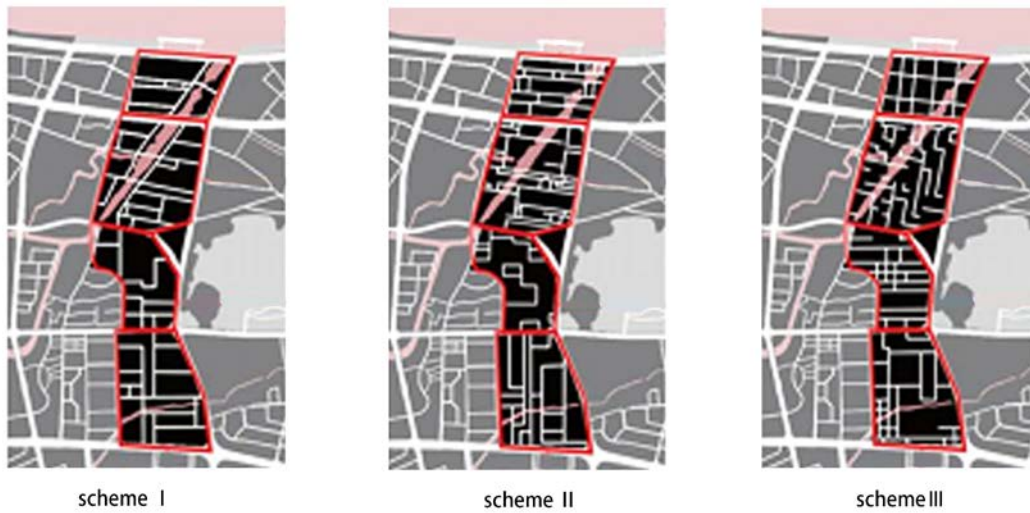
Figure 2. Generated urban road network.
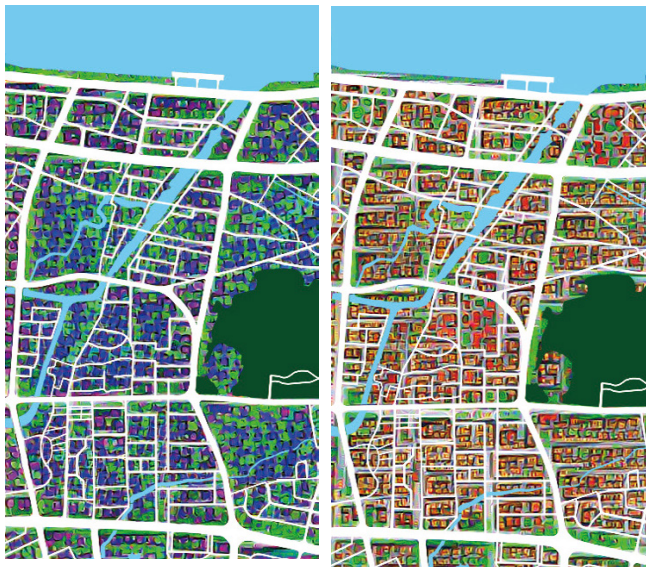
Figure 3. Optimal design of urban block spatial form and urban functional layout.

5.5 Generation of 3D urban design models

The first step is writing the code in Processing. Pixels in the 2D urban block spatial form map generated in the previous stage are assigned to different heights according to their colors and a 3D voxel model is generated by algorithm (figure 4). Afterwards, the 3D coordinate information of each voxel point is exported. In the second step, the 3D coordinate information data obtained in the last step is imported in Rhino and Grasshopper to automatically generate the 3D block model of the North Extension of Central Green Axis in Wenzhou (figure 5).

Figure 4. 3D voxel model.

60 and 80. 18 urban building function layouts are generated for evaluation and selection. The evaluation indicators specified in this paper are as follows: the diversity of urban building function layout and the rationality of urban building function layout. The indicator of urban building functional diversity is the number of building functions and the indicator of the rationality of urban building functional layout depends on the distribution density of buildings with different functions and the coverage of building service radius. Based on the scheme, the designer will make an empirical score. The highest score for each indicator is 10 and the final score is the sum of two indicators. In the end, according to the final score, the best functional layout is generated by Hangzhou Qiangjiang CBD and Venice with iteration of 80 (figure 3).

Figure 5. 3D block model.

144

6 CONCLUSION

We have investigated the suitability WFC and CNNs for urban space synthesis. Firstly, an urban design database is established as the data source for generation. Then the road networks urban spatial forms and urban functional layout are generated by WFC and CNNs and evaluated by designer afterwards. Finally, the 3D models are generated. Our result demonstrates the feasibility of WFC and CNNs for fast prototyping of urban design. It can improve the design efficiency of urban designers and assist government's decision making.

However, we identified several limitations in our pipeline. Firstly, the schemes generated by this method are not accurate enough due to the limitations of computer hardware and algorithm. The boundaries between areas with different colors are not clear in the generated map by CNNs. Secondly, the process of scheme evaluation and selection is still subjective and empirical. The accuracy of evaluation and selection heavily relies on urban designers' ability and experience. Thirdly, the data format of the data source in the database and the generated results are all pixels. The resulting images produced in our pipeline can only be used for a preliminary planning concept and there is still a certain gap from the actual application of engineering projects. Transforming the format of output from raster to vector is a crucial step for allowing the pipeline to integrate with common architectural tools and practices. Fourthly, this pipeline is intended to be used for new spaces and cannot used for the management of the existing urban areas which is highly demanded in urban regeneration.

In order to solve the problems above, there are several interesting directions for future work. We plan to leverage Generative Adversarial Network for generation in order to improve the accuracy of image generated. Then, an optimization code, such as genetic algorithm, will be developed for evaluation and selection. Finally, we will investigate a new algorithm which can directly generate design output in the format of vector.

REFERENCES

1. Champandard Alex J.. Semantic Style Transfer and Turning Two-Bit Doodles into Fine Artwork. https://arxiv.org/pdf/1603.01768.pdf, 2016.

2. Chaillou S. ArchiGAN: a Generative Stack for Apartment Building Design. https://devblogs.nvidia.com/archigan-generative-stack-apartment-building-design/, 2019.

3. Gumin M. https://github.com/mxgmn/WaveFunctionCollapse

4. Harrison, P. F. Image Texture Tools: Texture Synthesis, Texture Transfer and Plausible Restoration. PhD Thesis, Monash University, 1--141. 2005. Retrieved from http://www.logarithmic.net/pfh/thesis.

5. Hartmann S., Weinmann M., Wessel R. and Klein R. StreetGAN: Towards Road Network Synthesis with Generative Adversarial Networks. *Computer Science Research Notes,* 2701 (2017), 133-142.

6. Heeger David J. and Bergen James R.. Pyramid-based Texture Analysis/Synthesis. *International Conference on Image Processing*, pp. 229–238, 1995.

7. Karth, I., & Smith, A. M. Addressing the fundamental tension of PCGML with discriminative learning. ACM International Conference Proceeding Series. 2019.

8. Karth, I., & Smith, A. M. Wave Function Collapse is constraint solving in the wild. ACM International Conference Proceeding Series, Part F130151. 2017.

9. Kleineberg M. https://github.com/marcliberatore/wavefunctioncollapse-1

10. Koenig R. Urban Design Synthesis for Building Layouts based on Evolutionary Many-Criteria Optimization. *International journal of architectural computing,* 13, 3-4 (2015), 257-269.

11. Kuatra V., Schodl A., Essa I., Turk G., and Bobick A. Graphcut textures: image and video synthesis using graph cuts. *ACM Transactions on Graphics*, 22, 3 (2003), 277-286.

12. Leach N. and Xu W. Digital Construction: Young Architects' Works. China Building Industry Press, 2008.

13. Leach N. and Xu W. Digital Construction: Student Architectural Design Works. China Building Industry Press, 2008.

14. Li C. and Wand M. Combining Markov Random Fields and Convolutional Neural Networks for Image Synthesis. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2017, 2479-2486.

15. Long Y. and Cao Z. Methodology and Application of the Self-feedback Urban Design Based on Urban Sensors and Online Platform, *International Urban Planning,* 33, 1 (2018), 34-42.

16. Merrell P. Model Synthesis. the University of North Carolina, Chapel Hill, 2009.

17. Merrell P. and Manocha D. Continuous model synthesis. *ACM Transactions on Graphics*, 27, 5 (2008), 158:1-7.

18. Miao, Y., Koenig, R., Knecht, K., Konieva, K., Buš, P., & Chang, M. C. Computational urban design prototyping: Interactive planning synthesis methods—a case study in Cape Town. International Journal of Architectural Computing, 16, 3(2018), 212–226.

19. Nishida G., Garcia-Dorado I., and Aliaga D. Example-driven procedural urban roads. In Computer Graphics Forum. Wiley Online Library, 2015.

20. Nishida G., Garcia-Dorado I., Aliaga D., Bedrich Benes, and Adrien Bousseau. Interactive sketching of urban procedural models. *ACM Transactions on Graphics*, 35, 4 (2016).

21. Portilla J. and Simoncell E. P. A Parametric Texture Model Based on Joint Statistics of Conplex Wavelet Coefficients. *International Journal of* Computer Vision, 40, 1 (2000), 49-71.

22. Shi Y., Kwan M. The Development of Digital Urban Design in American Universities. *International Urban Planning,* 33, 1 (2018), 28-33.

23. Smelik R. M., Tuenel T, Bidarra R., and Benes B. A survey on procedural modelling of virtual worlds. *Computer Graphics Forum*, pages n/a–n/a, 2014.

24. United Nations, World Urbanization Prospects, 2014 Revision, Department of Economic and Social Affairs, Population Division, New York, 2014.

25. Varshaneya V, Balasubramanian S. and Balasubramanian V. N. Teaching GANs to Sketch in Vector Format. https://arxiv.org/pdf/1904.03620.pdf, 2019.

26. Xie Z. Exploration of the potential use of space syntax analysis in 3D parametric design: the case of Quanzhou. University College London, London, 2011.

27. Yuan F. and Leach N. Digital Workshop in China. Tongji University Press, 2013.

28. Yumer M. E., Asente P., Mech R., and Kara L. B. Procedural modeling using autoencoder networks. *Proceedings of the ACM Symposium on User Interface Software and Technology*, 2015, 109-118.

29. Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. Measuring human perceptions of a large-scale urban region using machine learning. Landscape and Urban Planning, 180(2018), 148–160.

30. Zhou, B., Liu, L., Oliva, A., & Torralba, A. Recognizing City Identity via Attribute Analysis of Geo-tagged Images Recognizing City Identity via Attribute Analysis of Geo-tagged Images. 2016.

Parametric Design and Structural Performance Assessment of a Topologically Interlocked Arch

Niloufar Emami[1]

[1]Louisiana State University
Baton Rouge, LA. USA
nemami@lsu.edu

ABSTRACT

I present research investigating the design and simulation of topological interlocking assemblies (TIA) made of concrete modules to create a self-supporting structure. Several studies have performed finite element method (FEM) analyses of simple interlocking assemblies. While the global performance of the assemblies using different block types has been examined, the effect of geometric variations of the local geometry on the global performance of such assemblies has not been widely investigated. In particular, the angle and number of undulations can be varied to increase the efficiency of these assemblies of osteomorphic blocks. This paper is part of an ongoing project on the design and fabrication of TIA using osteomorphic blocks. In it, I present a series of FEM simulations that investigate the relation between local design of the modules and the global structural performance of the assemblies in an arch.

Author Keywords

Topological interlocking assemblies; parametric and computational design; structural simulation; Finite element method; concrete.

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING (e.g. Model Development). See: http://www.acm.org/about/class/1998/ for more information and the full list of ACM classifiers and descriptors.

1 INTRODUCTION

Topological interlocking assemblies (TIA) are made of solid elements. Their overall structural integrity relies on each element being kinematically constrained by its neighbors. This system establishes equilibrium through compression forces, where the weight of each heavy block is used against itself to maintain it in the span. Given fixed boundary conditions, the assemblies can resist forces without any additional binding material such as mortar—adjacency replaces mortar. In addition, the global geometry of the assembly can be controlled to minimize deflections and increase the efficiency of the system.

Many researchers have studied the design and assembly of topological interlocking elements. Dyskin et al. conducted theoretical and experimental research, suggesting that the properties of topological interlocking, such as increased

fracture toughness, tolerance to local failure, high formability, and low bending rigidity, opens new avenues for large-scale mortar-free construction [1]. They later categorized various edge conditions of these assemblies as either having an external frame, an internal prestressed cable or tendon, or self-weight, which are all auxiliary elements of the structure [2]. Later, researchers investigated cube-shaped topological interlocking systems with an emphasis on those built with aluminum, steel, and PVC [3]. Estrin et al. also studied multi-layer interlocking elements (as opposed to single layer assemblies) while exploring the use of hybrid materials [4]. Their research group recently published a study hypothesizing that the mechanical responses of topological interlocking assemblies are controlled by the geometry of the interlocking elements and their surfaces, as well as their local surface patterns [5]. Oliver Tessmann extensively studied the design of topological interlocking assemblies. He parametrically designed differentiated interlocking elements for planar and curved surfaces [6] [7] [8]. More recently, he published on the application of non-continuous logics in architecture, offering a relevant case study of topological interlocking [9]. Vella et al. created a parametric model inspired by the flat Abeille Vault to create a double-curvature surface [10]. Weitzmann et al. also studied topological interlocking systems as organizational mechanisms that can be used for facades, using two or three different tetrahedra for planar tiling systems and curved surfaces [11]; they later studied how they are used for the construction of building floors and conducted numerical simulation using finite element method (FEM) analysis in DS SolidWorks. They also conducted primary topology optimization analysis to demonstrate the material and weight improvements to a single unit using Altair's Inspire software [12], and more recently, conducted discrete element method (DEM) structural analysis of the modules using Itasca 3DEC [13]. Other research studies on TP systems include in-depth equilibrium analysis of TI for structural assemblies [14]; interlocking cast glass components analyzed by DIANA FEA software [15]; and a theoretical and experimental study of topological interlocking assemblies by using X-joint theory edge modules, which absorb lateral thrusts [16]. The literature reviewed for this work is summarized in Table 1.

Researchers	Module's geometry, material, and assembly	Method	Findings/future directions
Dyskin, Estrin, et al. 2001 [1] [2]	Mechanical properties of tetrahedron-based structures; planar surface configuration.	Theoretical and experimental study.	Significant non-elastic strains under concentrated force; small bending compared to a massive plate with an equivalent thickness.
Schaare, Riehemann and Estrin 2009 [3]	Aluminum, steel and PVC cube-shaped elements; planar surface configuration.	Numerical simulation (using 3DS Simulia ABAQUS) and experimental study.	The assemblies have high energy absorption potential.
Estrin, Dyskin, and Pasternak 2011 [4]	Single layer polyhedral assembly; multi-layer osteomorphic assembly; hybrid structure comprising a layer of metallic cubes and a silicon rubber; all as planar surface configuration.	Prototyping scaled design concepts with various materials can be combined within a single topologically interlocked structure.	Multifunctional hybrids can be obtained that combine various materials within an assembly.
Djumas, Simon, Estrin, and Molotnikov 2017 [5]	Explores the idea of hierarchical structures to topologically interlocked assemblies.	FEM Numerical simulations using ABAQUS/Explicit; fabrication of elements using 3D printing; mechanical experiments under point loadings.	Introducing a secondary surface geometry to non-planar elements creates an additional interlocking effect. Patterning of the interlocking block surface delays the onset of slipping of the central block under concentrated load.
Tessmann, 2012–13 [6] [7] [8]	Differentiated tetrahedron assembly on a planar and curved surface.	Parametric design; fabrication through creating the unrolled cut pattern and using it as a mold to cast plaster.	Potential of the interlocking systems for performative sculptures or retaining facilities for mud floods; identifies need for better sequencing and self-organization of elements within the system.
Vella et al. 2015 [10]	Inspired by the identical tetrahedron ashlars of the flat Abeille vault, creates differentiated modules covering a double-curved surface; the Truchet vault; and osteomorphic blocks.	Parametric design.	Abeille-like configuration as a generic example of TI; the combination of digital design and digital fabrication enables the reanimation of historic construction techniques.
Weitzmann et al. 2015–19 [11] [12] [13]	Tetrahedra projected on planar and curved surfaces; potential use for construction of building floors.	Computational and parametric design, 3D printing modules, FEM and DEM structural analysis.	Existing relationship between geometry and structural performance with an application to the construction of floor systems.
Pfeiffer et al. 2019 [16]	Tackles the idea of removable structures using X-joint theory edge modules.	Computational design, digital machining, and assembly experimentation.	Some X-joints undergo substantial bending stress and more in-depth structural analysis is needed. Hexagonal pattern and special edge module open door to geometric exploration of TIA.

Table 1. Summary of literature on design, simulation, and experimental testing of topological interlocking assemblies.

Many previous studies have focused on parametric design of TIA to create differentiation either in terms of modules' porosities or to accommodate complex curvatures. Some have focused on structural performance assessment of the assemblies through simulation, while others have emphasized prototyping the assemblies through digital fabrication [16]. Yet, the relationship between the design of the modules and the structural performance of the assemblies has not been investigated in depth.

This study employs parametric design and simulation to easily vary module geometry and assess its effect on the structural performance of the assembly. In other words, the focus of this study is to understand the relationship between geometry at the micro level and structural performance at the macro level. Investigation of this part-to-whole relationship can provide a design guide before design parameters are adjusted to accommodate more complex curvatures or digital fabrication.

2   PARAMETRIC DESIGN

Researchers hypothesize that the mechanical response of TIA are controlled by the *geometry of the interlocking elements*, *their surfaces*, and the *local surface patterns* that may be introduced [5]. The geometry of TIA can vary from platonic polyhedral geometries with varying numbers of faces to osteomorphic blocks. Regarding the surfaces, researchers have stated that the contact face between the modules can be planar (e.g. contact faces in platonic modules), or more complex (e.g. contact faces in osteomorphic blocks). Planar contact faces only constrain movement in one direction, whereas more complex contact faces with curved multi-directional undulations can constrain movement in at least three directions [16]. Patterning or itching the contact faces of local surface patterns can further increase the friction between the modules, thus increasing the assembly's efficiency.

In this study, design is considered at two scales in the parametric design process: local and global. At the local level, the geometry and surfaces of the modules is considered. Osteomorphic blocks are selected as the module geometry. The inclining wavy faces in osteomorphic blocks lock the movement of neighboring units in both up and down directions (Figure 1). Then, the blocks are shifted in each row to achieve interlocking [17]. At the global level, an arch is designed as the underlying surface; the number of modules projected on it can vary. I parametrically modeled this catenary arch after an osteomorphic catenary arch by Fallacara [18] using Rhino, Grasshopper and the Kangaroo form-finding plug-in. The catenary arch allows uniform load distribution and after form-finding, the osteomorphic blocks are projected onto the arch. All of the blocks undergo topological deformation, meaning that there are dimensional and angular variations among them. An arch was selected for two reasons: first, the single curvature of an arch seemed appropriate for initial investigation before exploring more complex double curvature forms; second, the single blocks are similar in each row of the arch, while half of the arch is a mirrored replica of its first half. This rationale is in line with one of the goals of the study—controlled variability.

I parametrically modeled a square unit and found the mid points of each edge to construct lines. After the lines were offset, a curve was blended between the two to create the lower cross section of the block. The upper cross section of the block was designed by flipping the direction of the blend. The two cross sections were then lofted, and the blocks populated in the x and y directions. Finally, these blocks were projected onto the catenary arch. The blocks at the base were cut with a plane to create flat edges (Figure 2).
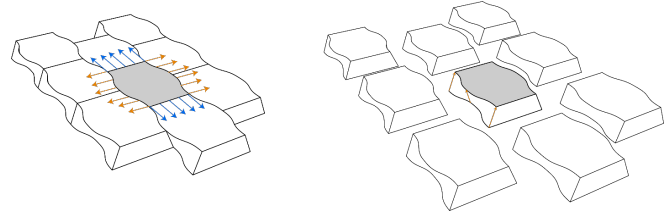
Figure 1. Mechanism interlocking osteomorphic blocks in the x, y, and z directions by shifting each row (left); with undulated contact faces (right).
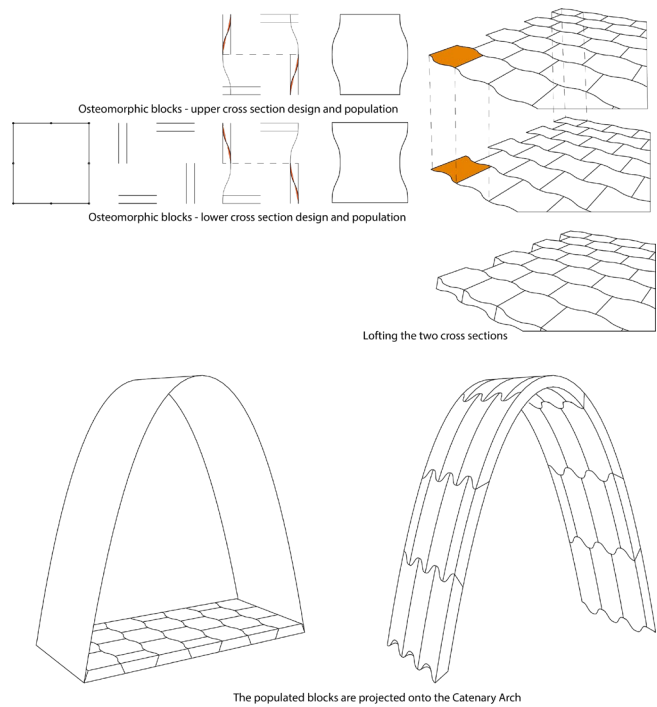
Figure 2. Parametric modeling process for an osteomorphic arch

The parameters that affect the geometry of the interlocking modules vary the contact surfaces of the blocks by changing the offset distance between the two lines (explained in the previous step) and the blend factor between the two offset curves. In addition, the width and length of the blocks can be changed (Figure 3). The constant and variable parameters are summarized in Table 2.

149
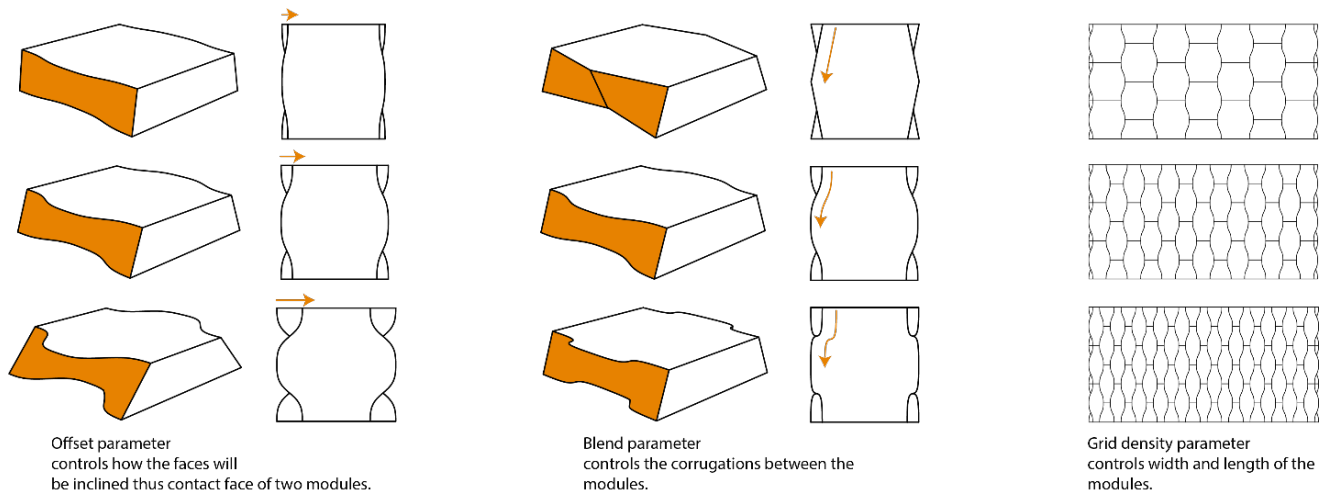
Figure 3. Parameters that can be varied at the module level

Parameter	Value/Range	Constant/ Variable
Catenary span	240 cm	Constant
Catenary rise	260 cm	Constant
Catenary width	120 cm	Constant
Interlocking module thickness	10 cm	Constant
Interlocking module offset	4–18 cm; 1-cm intervals	Variable
Interlocking module blend factor	0%–1%; 0.1% intervals	Variable
X-direction replication (number of units in x direction)	7–19 units, intervals of 2 (always an odd number to maintain a key stone)	Variable

Table 2. Constant and variable parameters in the parametric model.

3 STRUCTURAL SIMULATION

Masonry is a heterogeneous brittle material that consists of units and mortar joints. It can be built out of fired clay, concrete, or calcium silicate bricks. Cracking and crushing may occur in masonry structures' units, mortar, at the brick/mortar interface, and/or all of the above [19].

The discrete/distinct element method (DEM) is one method for simulating the mechanical behavior of masonry structures; it "…is characterized by modeling the materials as an assemblage of district blocks or particles interacting along their boundaries" [19 p. 37]. DEM is an explicit method based on finite difference principles, derived from Peter Cundall's original work [19]. It is presented in the UDEC (Universal Distinct Element Code) and 3DEC software, developed for commercial use by Itasca Ltd [19].

It seems that 3DEC is appropriate for analyzing convex geometries; however, it is not capable of analyzing concave geometries or geometries with curved surfaces—in this case, the osteomorphic blocks. Thus, 3DEC was not used for this analysis.

Instead, I used ANSYS 19.2 software for FEM analysis. Any geometry created in Rhino NURBS modeling software can be imported into ANSYS. If the geometry is initially modeled as an assembly of discrete geometries that each have watertight masses and are assembled to have contact faces, ANSYS recognizes those contact faces and can assign a series of contact types to them. "Bonded" is the default contact type, which is applied to all contact regions when any geometry is imported. Bonded contact does not allow sliding or separation between faces or edges of discrete elements. "No Separation" contact is similar to "Bonded" except that it applies only to faces or edges. "Frictionless" contact models standard unilateral contact (i.e., normal pressure equals zero if separation occurs). In this nonlinear mode, the area of contact may change as the load is applied; thus, gaps can form between the bodies. A zero coefficient of friction is assumed in this model, which allows free sliding. "Rough" is another contact type, similar to "Frictionless." It models perfectly rough frictional contact but does not allow any sliding. Finally, if the "Frictional" contact type is selected, the two contact geometries can carry shear stresses up to a certain magnitude before they start sliding [20].

Before analyzing the arch, I conducted a series of simulations to benchmark the simulation engine, as well as to understand how changing the settings used for contact faces affects the results. These benchmarking simulations are presented in the following subsections. It should be noted that the legends in the simulation snapshots vary from case to case since they are auto generated using the minimum and maximum values of the related result. Thus, though one legend cannot be used for all cases, the maximum result is highlighted as a written text in the related tables.

150

3.1 Benchmarking a beam modeled as a continuous geometry versus a beam modeled as three discrete elements

ANSYS's "static structural" module was used for this benchmarking simulation. Two beams were modeled in Rhino and exported into ANSYS: One had a cross section of 0.5 m by 0.5 m spanning 10 m modeled as one piece; the other a cross section of 0.5 m by 0.5 m comprising three equal discrete pieces connected through their contact faces, spanning 10 m. Once imported into ANSYS, concrete was assigned as the material, and the two edges of the beam selected as support using the "displacement" support, which restricts movement in x, y, and z directions but allows rotation around the y and z axes (degrees of freedom: FFFFRR). Finally, a uniform load of 1 KPa was applied to

the top surface of each beam. The default setting of Bonded was used for contact faces of the three-element beam.

I recorded maximum total deformation and maximum von Mises stress for both beams (summarized in Table 3). Maximum deformation was similar for both beams, and the von Mises stress levels were very close. These results demonstrate that the Bonded contact faces for a geometry made of discrete elements act in a manner similar to a geometry modeled as a continuous geometry.

Maximum total deformation dominated, leading me to assume that the modules will slide off or destabilize before cracking. Therefore, deformation was used to compare the performance of different scenarios in subsequent simulations.
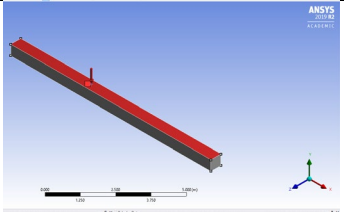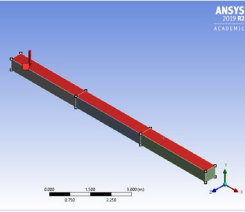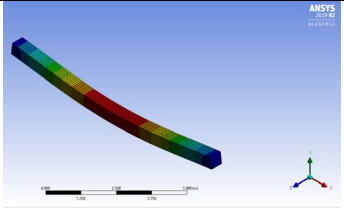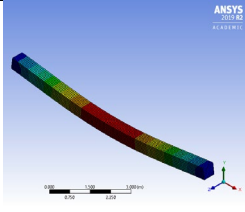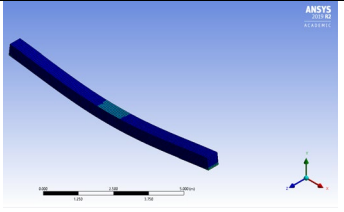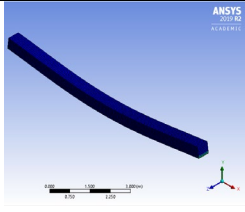
	Beam modeled as a continuous geometry	Beam modeled as three discrete elements (bonded contact faces)
Simulation settings		
Maximum total deformation	0.19 cm	0.19 cm
Maximum von Mises stress	18 MPa – occurring at the support	20 MPa – occurring at the support

Table 3. Benchmarking results of the analysis of a continuous and discrete beam.

3.2 Benchmarking a circular arch made of three discrete elements with different contact types

In this phase, a simple arch made of three discrete elements with contact faces is modeled with a square cross section of 0.5 m by 0.5 m, spanning 10 m with a rise of 5 m. Once imported into ANSYS, I assigned concrete as the material and selected the two faces of the arch as displacement supports (restricted in x, y, and z direction), and self-weight as the load.

The first simulation of the circular arch used the default setting in ANSYS (Bonded modules). After the simulation was completed, the contact faces between the modules were

changed to other types. When any one of the contact regions is selected, ANSYS displays the location of the contact faces, and the two elements that are connected through that face (Figure 4).

I simulated deformation of the arch in multiple simulations using the Bonded, Frictionless, Rough, and Frictional, settings (Figure 5); deformation values for each scenario are summarized in Table 4. Note that the deformation visualizations in these figures are scaled up to allow close inspection. A coefficient of 0.3 was used for analysis with the Frictional setting as in [13].
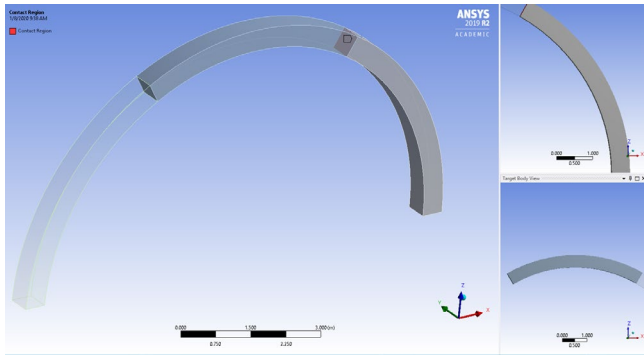
Figure 4. ANSYS displays the elements connected through a contact face when that face is selected.

Contact faces of discrete elements	Maximum deflection (cm)
Bonded	0.022
Frictionless (coefficient = 0), free sliding	0.49
Rough (coefficient = 0), no sliding	0.17
Frictional (coefficient = 0.3), carries shear stress	0.17

Table 4. Deformation of the arch with three modules with different settings for contact faces.

a. Bonded contact faces

b. Frictionless contact faces

c. Rough contact faces

d. Frictional contact faces, coefficient = 0.3
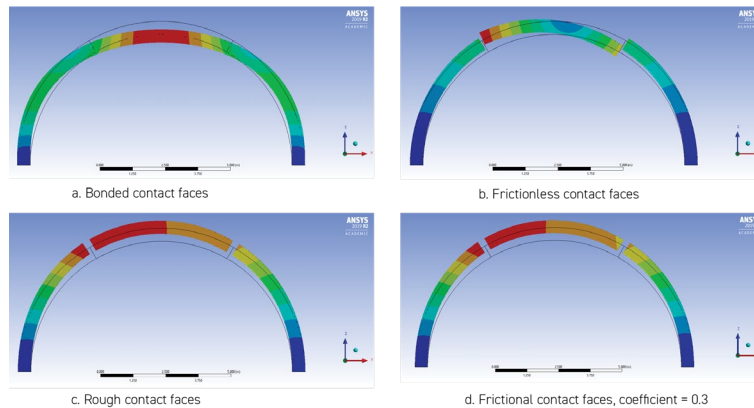
Figure 5. Deformation simulated with a. Bonded, b. Frictionless, c. Rough, and d. Frictional contact faces.

Bonded contact faces undergo the smallest deformation. Rough and Frictional contact faces create similar deformations, both visually and numerically. Frictionless contact faces result in the largest deformation, with free sliding of the keystone visible. Thus, frictionless contact faces (with zero co-efficient) do not seem appropriate for the contact faces, since an important feature of topological interlocking assemblies is constraint by neighboring modules, in which friction plays a role in keeping elements in place. Therefore, Frictionless contact faces were not used in later simulations.

3.3 Benchmarking arches with different numbers of elements using Rough contact

At this stage, I increased the number of elements in this simple arch from 3 to 5, 9, and then 11, to examine how the deformation shape and amount are affected by increasing the number of modules. The contact faces were set to Rough with a friction coefficient of 0, limiting the sliding of the modules. The deformation simulations are summarized in Table 5. It should be noted that the modules of the simple arch are not interlocking elements, and the simulations are done to better understand the effect of increasing the number of elements in an arch using a specific contact face setting.

The results show that deformation increases as the number of elements increases. Knowing that no mortar was used in

this arch (no Bonded contact), these results required further validation using an interlocking geometry for the modules; regular brick-shaped modules without mortar can easily fail under self-weight.
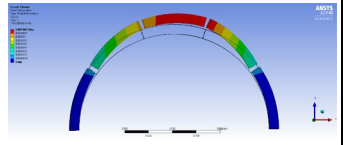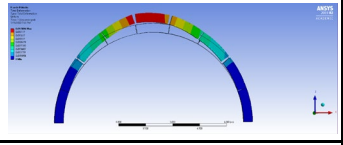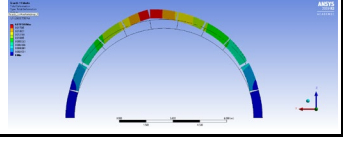
No. of elements	Max. deformation (cm)	
5	0.55	
9	1.7	
11	1.91	

Table 5. Deflection of arches with 5, 9, and 11 modules under self-weight.

3.4 Analysis of the catenary arch

In this stage, I simulated one version of the parametrically designed catenary arch. The curve offset value in the parametric model was set to 11 cm, with a blend factor of 6%. I then simulated the global deformation when the contact faces were set to Bonded, Frictional (co-efficient = 0.3), and Rough (Table 6). The visualizations are presented in Figure 6.

Contact faces	Maximum deformation (cm)
Bonded	4.4 e-3
Frictional (coefficient = 0.3), carries shear stress	0.32
Rough (coefficient = 0), no sliding	2.6 e4

Table 6. Topologically interlocked catenary arch
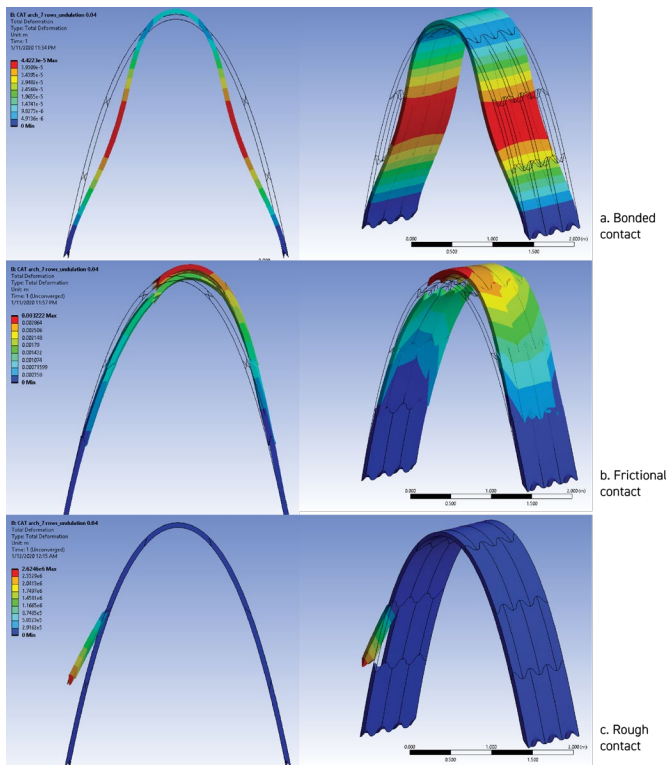
Figure 6. Simulation of a catenary arch. a. Global deformation with Bonded contact faces, b. Local deformation with Frictional (coefficient = 0.3) contact faces, and c. local deformation with Rough contact faces.

As expected, using Bonded contact faces resulted in the smallest deformation (4.4 e-3 cm). The deformation shape presents the global behavior of the arch. When Frictional contact faces were used, deformation increased to 0.32 cm. In this case, the keystone was the module undergoing maximum deformation, causing displacement of neighboring units. Finally, simulation of Rough contact faces revealed substantial deformation (2.6 e4 cm), with one module completely displaced and subsequent failure of the whole arch. This simulation shows the important role of friction in the analysis, even for the interlocked geometry. In future simulations, the global behavior of the structural assemblies will be simulated through Bonded contact faces and the local behavior of the assembly simulated through Frictional contact faces.

4 CONSIDERATIONS AND LIMITATIONS

A few points will need to be addressed in future studies:

Topological interlocking assemblies require fixed boundary conditions (various edge conditions of these assemblies include either the use of an external frame, an internal prestressed cable or tendon, or self-weight). In the simulations for the circular or catenary arch under self-weight, only the base of the arch was set as the support, while all blocks along the edge were left free to slide. In Fallacara's built arch, steel wires are integrated into the top surface of the arch to prevent the modules from sliding (Figure 7), and a beam placed on two sides of the arch to fix the boundary conditions. One of the next steps of this research project will be to model these boundary conditions and examine their effect on the results. Once the appropriate modeling and simulation strategies are identified, design parameters can easily be varied, and the effect of changing local geometry on structural performance closely studied.

Figure 7. Steel wires on Fallacara's arch [18].

On a separate note, one of the limitations of this study is that consideration of all elements was based on simulation, not data experimental studies. Load testing physical assemblies will be a next step to validate the results of simulation.

5 CONCLUSION

This research project is still in progress; the aim is to identify the best simulation method for TIA. This will be one that considers the effect of the local shape of elements on the global performance of structural assemblies. In this work, I have confirmed via a series of simulations in ANSYS FEM software, that:

- when an assembly of modules with contact faces is modeled and imported into ANSYS, different contact types can be selected and applied to it. If the Bonded contact type is selected, the assembly will perform in a manner very similar to one modeled as a continuous geometry in the first place. In other words, failure at the brick/mortar interface will never occur in this type of simulation. Instead, the global behavior of the assembly will be simulated. This is a good simulation strategy for understanding the global behavior of these assemblies.

- the Frictional contact type recognizes the contact faces and allows failure of the brick/brick interface to be simulated. The Frictional contact type with a predefined, non-zero coefficient can carry shear stresses up to a certain magnitude across interfaces before individual elements start to slide. In other words, this setting allows us to see which module starts to slide first, causing maximal deformation or instability of the whole assembly. The local behavior of such assemblies can be simulated through Frictional contact faces.

After completing the benchmarking simulations, the next step of the study will involve varying the design parameters of the topological interlocking modules to understand the effect of changing local geometry on global performance. This project is being conducted in parallel with one that aims to determine how such modules can be fabricated using 3D-printed forms to cast concrete [21]. The ultimate ambition of this research program is to derive the optimized geometric properties of TIA based on consideration of the limitations of their structural performance and fabrication. The results of the present study contribute to knowledge on the design, assembly, and fabrication of topological interlocking assemblies used as self-supporting structures.

REFERENCES

[1] B. A. V Dyskin, Y. Estrin, A. J. Kanel-belov, and E. Pasternak, "Toughhenning by fragmentation - how topology helps," *Advanced Engineering Materials*, vol. 3, no. 11, pp. 885–888, 2001.

[2] A. V. Dyskin, E. Pasternak, and Y. Estrin, "Mortarless structures based on topological interlocking," *Frontiers of Structural and Civil Engineering*, vol. 6, no. 2, pp. 188–197, 2012, doi: 10.1007/s11709-012-0156-8.

[3] S. Schaare, W. Riehemann, and Y. Estrin, "Damping properties of an assembly of topologically interlocked cubes," *Materials Science and Engineering A*, vol. 521–522, pp. 380–383, 2009, doi: 10.1016/j.msea.2008.10.069.

[4] Y. Estrin, A. V. Dyskin, and E. Pasternak, "Topological interlocking as a material design concept," *Materials Science and Engineering C*, vol. 31, no. 6, pp. 1189–1194, 2011, doi: 10.1016/j.msec.2010.11.011.

[5] L. Djumas, G. P. Simon, Y. Estrin, and A. Molotnikov, "Deformation mechanics of non-planar topologically interlocked assemblies with structural hierarchy and varying geometry," *Scientific Reports*, vol. 7, no. 1, pp. 1–11, 2017, doi: 10.1038/s41598-017-12147-3.

[6] O. Tessmann, "Topological Interlocking Assemblies," in *Physical Digitality—Proceedings of the 30th International Conference on Education and Research in Computer Aided Architectural Design in Europe*, 2012.

[7] O. Tessmann, "Interlocking manifold: kinematically constrained multi-material systems," in *Advances in Architectural Geometry*, L. Hesselgren, S. Sharma, J. Wallner, P. Bompas, and J. Raynaud, Eds. 2013, pp. 269–278.

[8] O. Tessmann and M. Becker, "Extremely heavy and incredibly light: performative assemblies in dynamic environments," *Proceedings of the 18th International Conference on Computer-Aided Architectural Design Research in Asia (CAADRIA) : Open Systems*, pp. 469–478, 2013.

[9] O. Tessmann and A. Rossi, "Geometry as Interface: Parametric and Combinatorial Topological Interlocking Assemblies," *Journal of Applied Mechanics*, vol. 86, no. 11, pp. 1–13, 2019, doi: 10.1115/1.4044606.

[10] I. M. Vella and T. Kotnik, "Geometric Versatility of Abeille Vault," *Proceedings of the 34th International Conference on Education and Research in Computer Aided Architectural Design in Europe*, vol. 2, no. Figure 1, pp. 391–397, 2016.

[11] M. Weizmann, O. Amir, and Y. J. Grobman, "Topological interlocking in architectural design," *CAADRIA 2015 - 20th International Conference on Computer-Aided Architectural Design Research in Asia: Emerging Experiences in the Past, Present and Future of Digital Architecture*, no. September, 2015, doi: 10.13140/RG.2.1.3642.3766.

[12] M. Weizmann, O. Amir, and Y. J. Grobman, "Topological interlocking in buildings: A case for the design and construction of floors," *Automation in Construction*, vol. 72, no. June, pp. 18–25, 2016, doi: 10.1016/j.autcon.2016.05.014.

[13] M. Weizmann, O. Amir, Y. J. Grobman, and T. Planning, "Structural Performance of Semi-regular Topological Interlocking Assemblies Faculty of Architecture and," *SimAUD*, no. April, pp. 217–223, 2019.

[14] Z. Ma, "Equilibrium Analysis of Topological Interlocking for Structural," 2017.

[15] F. Oikonomopoulou *et al.*, "Interlocking cast glass components, Exploring a demountable dry-assembly structural glass system," *Heron*, vol. 63, no. 1–2, pp. 103–137, 2018.

[16] A. Pfeiffer, F. Lesellier, and M. Tournier, "Topological Interlocking Assemblies Experiment," in *Design Modelling Symposium (DMS): mpact: Design With All Senses*, 2019, pp. 336–349, doi: 10.1007/978-3-030-29829-6_27.

[17] Y. Estrin, A. V. Dyskin, and E. Pasternak, "Topological interlocking as a material design concept," *Materials Science and Engineering C*, vol. 31, no. 6, pp. 1189–1194, 2011, doi: 10.1016/j.msec.2010.11.011.

[18] G. Fallacara and M. Barberio, "An Unfinished Manifesto for Stereotomy 2.0," *Nexus Network Journal*, vol. 20, no. 3, pp. 519–543, 2018, doi: 10.1007/s00004-018-0390-z.

[19] V. Sarhois, K. Bagi, J. V. Lemos, and G. Milani, *Computational Modleing of Masonry Structures Using Discrete Element Method*. 2016.

[20] Ö. Sunar, "Contact Types and Behaviours in Ansys," 2018. [Online]. Available: http://www.mechead.com/contact-types-and-behaviour.

[21] N. Emami and P. Holmquist, "Cast Stereotomy : a material-based investigation of stereotomic modules," in *Association of Collegiate Schools of Architecture (ACSA)*, 2020.

A Deep Image of the City:
Generative Urban-Design Visualization

Ariel Noyman[1]**, Kent Larson**[2]

[1]City Science Group, MIT Media Lab, Cambridge, MA, noyman@mit.edu
[2]City Science Group, MIT Media Lab, Cambridge, MA, kll@mit.edu

ABSTRACT

Streetscape visualizations are necessary for the understanding and evaluation of urban design alternatives. Alongside blueprints and textual descriptions, these design aids can affect city-form, building-codes and regulations for decades to come. Yet despite major advancements in computer graphics, crafting high-quality streetscape visualizations is still a complex, lengthy and costly task, especially for real-time, multiparty design sessions. Here we present DeepScope, a generative, lightweight and real-time platform for urban planning and cityscape visualization. DeepScope is composed of a Generative Neural Network (DC-GAN) and a Tangible User Interface (TUI) designed for multi-participants urban design sessions and real-time feedback. In this paper we explore the design, development and deployment of the DeepScope platform, as well as discuss the potential implementation of DeepScope in urban design processes. Demonstration and code are available at: https://www.media.mit.edu/projects/deep-image-of-the-city/

Author Keywords

Urban design; GANs; Generative Design; Visualization; Collaborative Planning

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING (e.g. Model Development). See: `http://www.acm.org/about/class/1998/` for more information and the full list of ACM classifiers and descriptors.

1 INTRODUCTION: THE IMAGEABILITY OF THE CITY

"To understand the role of environmental images in our own urban lives (...) we needed to develop and test the idea of imageability (...) and thus to suggest some principles for urban design." [21]

Urban design renderings and streetscape visualizations are essential for designers, stakeholders and decision-makers during city-design processes. These visual aids can clarify the

Figure 1. DeepScope TUI: Multiple users can simultaneously interact and discuss urban design iterations. The table-top is used as both the design space and a schematic urban top-view. The vertical monitor visualizes the DCGAN street view.

outcomes of complex design decisions, such as zoning, building codes or land-use allocations, and can affect urban development for decades to come [1, 37]. The importance of understating the impacts of urban design on the street-level was known to architects and planners for centuries, but tools, mediums and techniques to communicate these effects were often limited [5].

In his seminal 1960 book, Kevin Lynch introduced 'imageability' as a novel approach to visual perception of urban environments [21]. Lynch suggested a toolset for classification of city-form, in which nodes, landmarks, paths, edges and districts reflect the sensation of transitioning through the urban scape. Later, in 'The View From the Road' study [3], Lynch's 'imageability' paradigm was tested using a new medium: Lynch mounted a dashcam to a car, and went on several rides around Boston and other US metros [2]. When later played, these recordings were sped to reflect the overall

Cityscapes — Edited dataset   **DCGAN** — Pix2pix at 16/32/64 filters   **Keras API** — Train/test on VM GPU   **TFjs** — Model   **Node.js** — Model Host Service

CityScope TUI   **CityScope API** — Urban modeler   **cityIO** — CS Backend
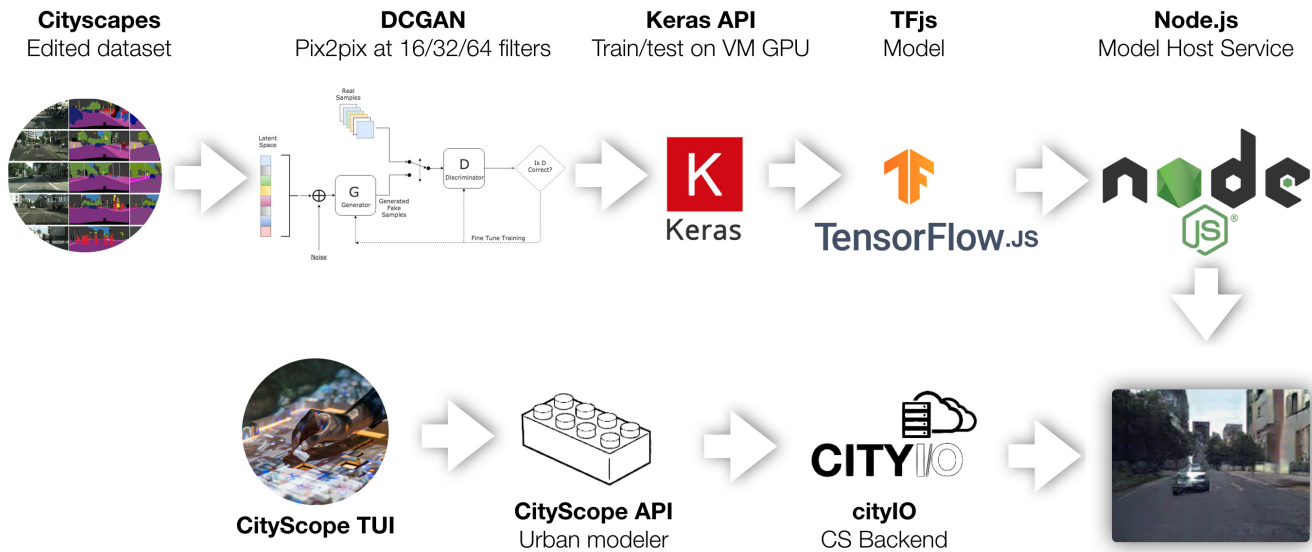
Figure 2. Top row: Model trained on Cityscapes dataset, deployed as node.js app. Bottom row: TUI triggers DCGAN renderings.

'feel' and 'mood' of the road trip; Lynch proposed to overlook fine-grained street elements or the architectural details, and instead focus on the 'imageability' of the urban outline: What is the composition of the built mass? What shapes the street-section? Are there any noticeable landmarks? In the following years, Lynch's innovative documentation techniques became mainstream tools in the field of urban-design [7, 30].

1.1 The Challenge of Visualisation

Despite Lynch's contribution to the perception of cities, documenting existing environments is not sufficient for predicting the impact of future interventions. As Batty concludes, urban visualizations are critical during initial design stages, when the context of the design challenge is only being established, as well as to the generation and evaluation of alternative designs [5]. In the last few decades, advancements in CAD and computer graphics introduced numerous tools to visualize future urban developments [35, 17]. Yet despite their abundance, only few tools offer real-time, realistic urban visualizations during collaborative design processes [26]. Most CAD tools carry complex setups, costly hardware and software, and steep learning curves [39, 22]. These tools often require users to set up many control parameters in virtual environments, such as cameras, lights, materials or shaders. This process might become laborious in complex design scenes, and can gravely affect the outcome, cost and duration of visualization processes [19].

Moreover, common CAD User Interfaces rarely support multi-user collaborative design. This limits decision-makers and stakeholders from taking an active part in iterative design sessions, and forces a synchronous decision making process. Lastly, early urban design stages suffer from lack of design

details, which hinders realistic visualizations. These stages commonly involves crude massing exercises, and lack street-level details, so that visualizations are schematic at best [11, 6].

2 DEEPSCOPE: METHODOLOGY AND SYSTEM DESIGN

This paper presents DeepScope, a collaborative, tangible and real-time urban design and visualization platform. DeepScope allows multiple users to collaboratively perform early urban design and land-use allocation sessions, and observe the outcomes as realistic streetscape visuals. Unlike CAD tools, DeepScope offers minimal setup, simple and cheap hardware and software, and requires no expertise to use.

This section details the main parts of DeepScope: (i) a tangible user interface (TUI) for rapid urban prototyping, and (ii) a Deep Convolutional Generative Adversarial Network (DC-GAN) for visualizations: As users interact with the TUI, a virtual city model is procedurally updated and fed into the DCGAN model. The model then generates a cityscape visualization based on a user-selected view. The rest of this section explores DeepScope TUI, hardware components, and user interaction.

2.1 HCI Platform for Rapid Urban Prototyping

DeepScope Tangible User Interface (TUI) is built for iterative urban design and land-use allocation. This TUI offers a playful, multi-user tangible environment for design that is augmented by real-time visualization.

Traditional Computer Aided Design (CAD) tools were commonly built around a single user with limited inputs (mouse, keyboard) and outputs (monitor, printer). These interface

a. Urban landuse map

b. Selected design area

c. Landuse converted to DeepScope labels

d. DeepScope TUI and "Observer"

e. Observer captures are input for the DCGAN

f. DCGAN Streetview perspective output

Figure 3. DeepScope process: (a,b) designating an urban intervention site (c) translating the site's land-use/zoning bounds and (d) user-interaction into (e) procedural 3D environment and (e) passing it to DCGAN model for generation of a street-view visualization

Figure 4. Multi-user interaction with DeepScope. Depending on scale and extents of urban context, design sessions can accommodate up to 15 users

Group	Classes
flat	road*+; sidewalk*+ parking*+; rail track
human	person*; rider*
vehicle	car*; truck*; bus*; on rails; motorcycle; bicycle*; caravan; trailer
construction	building; wall*; fence; guard rail; bridge
object	pole*; pole group; traffic sign*; traffic light*
nature	vegetation*+; terrain
sky	sky*
void	ground; dynamic; static

Table 1. Cityscapes classes: Marked with 'plus' are labels which can be altered dynamically using CS TUI. Marked with 'star' are labels that are generated dynamically in the 3D model

were not initially conceived as collaborative design tools, even when computer networks became mainstream [38, 5].

In past decades, several TUIs have been developed to facilitate collaborative urban design, augmented by computational analytics. Among these are the Augmented Urban Planning Workbench, the I/O Bulb, The Clay Table and Sensetable [13, 14, 29], all built to allow teamwork and collaboration in urban design processes. In recent years, The MIT City Science group has been developing CityScope (CS): an urban modeling, simulation and collaborative decision-making platform. CS merges TUIs and analytical modules to support a collaborative, evidence-based discourse around the built environment [27]. For the purpose of this research, a CS instance was developed, constructed and tested in an active demonstration space at the MIT Media Lab, Cambridge, MA.

2.2 DeepScope User Interaction

The TUI is composed of three components: (i) a physical urban model, (ii) a scanning module and (iii) a feedback module. The urban model includes an arbitrary grid of tiles, tagged with binary patterns. The tiles are made out of 4x4 LEGO bricks, which were found to induce interaction and creativity during CS design sessions [27]. Each pattern is a 16 bit code of black or white 1x1 LEGO studs, allowing over 65,000 unique pre-defined land uses and attributes. Figure

3.d depicts a user positioning a tagged LEGO brick into the TUI design space.

Each grid-cell pattern represents a different streetscape class: roads, buildings, green-spaces, parking, sidewalks, etc. Each class instance contains additional parameters, such as height, volume, shape, rotation or density. Table 1 specifies the classes and their attributes. When the user shifts a tile, the scanning module detects the interaction through a scanning and networking tool using OpenCV and Node.js. Lastly, a feedback module, containing monitors and projectors, communicates the interaction and analysis outcomes back to the users. This interface has been shown to allow for rapid design iteration, facilitate collaboration and engage users in urban design processes [28].

2.3 Procedural Cityscape Environment

With each interaction, the scanner decodes the new grid-cell patterns and updates the table's data structure. This triggers a regeneration of a virtual 3D environment, in which each grid-cell is represented via its class and additional parameters (see figure 5). As users allocate tiles, the environment is procedurally filled with streetscape elements: a vegetation pattern will create a surface with procedural trees, bushes or live-fences; A sidewalk pattern will produce pedestrians and street-signage, and a parking-lot pattern will be proliferated with parked vehicles. This 3D environment is uniformly hued

157

with RGB values that correspond to input classes expected by the Neural Network model (see section 3). The scanning and 3D scene generation is done on a client-side web-browser using a simple webcam and a WebGL program [25].
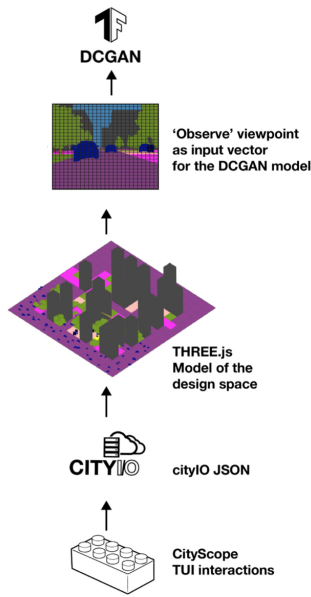
Figure 5. TUI to latent space: TUI interactions are analyzed using OpenCV and streamed as JSON with the webGL app. A 3D model is created based on the JSON array and the Observer viewing angle. Lastly, a snapshot image is fed as an input vector to the DCGAN model.

2.4 Observer

The urban environment designed by the users is constantly 'photographed' by the 'Observer' grid-cell. Similar to Lynch's 'View form the Road' [3], this unique pattern mimics a virtual nomad in the city, and allow users to sets its position, point-of-view and angle. The 'observer' baseline parameters (such as FOV, Frustum and height) were approximated to the camera calibration appendix of the Cityscapes dataset [8]. Additional camera controls were implemented to allow users to move, rotate, pan or zoom the 'observer' by relocating the cell itself and via custom game-pad joystick (see figure 6).
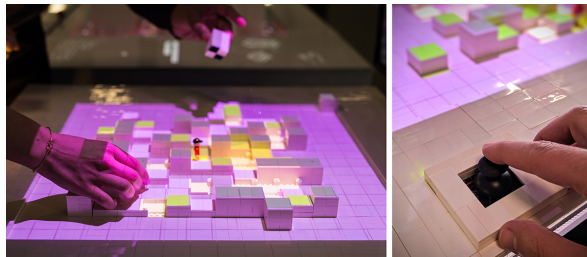
Figure 6. (left) User interaction with grid-cells. (right) 'Observer' viewing angle, depth and position is set via an Arduino Gamepad

2.5 Table-Top Augmentation

The TUI table-top is used as the design space as well as a canvas for visualization. With each design iteration, an illuminated land-use diagram is projected onto the table-top, so that each tile is showing its respective pattern, name or parameters (density, land use, etc.). The Observer position is displayed using perspective cone that indicates its viewing angle and FOV (see figure 7). Together, DeepScope TUI components allow multiple users to design and amend the urban environment and observe the effects of different scenarios on its streetscape.
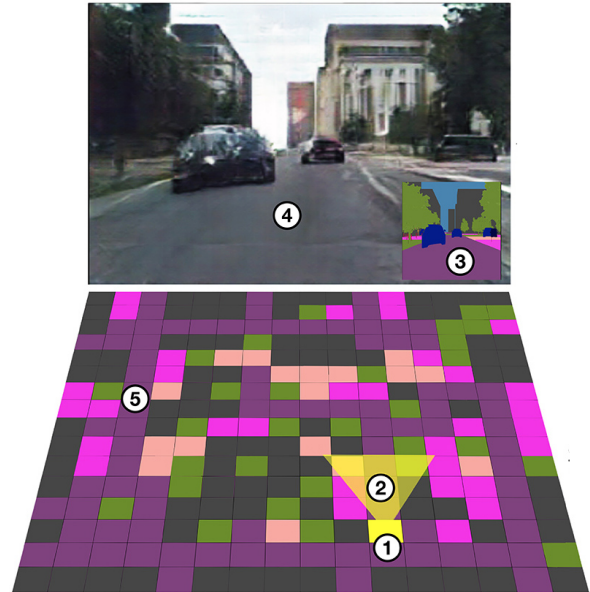
Figure 7. DeepScope TUI: (1) Observer position (2) Observer view angle and FOV cone (3) Observer's 3D street-view as input for DCGAN (4) DCGAN model prediction of street-view (5) TUI interactive grid

3 DEEPSCOPE GENERATIVE NEURAL NETWORK

In order to produce realistic street-view visualization, DeepScope implements a Neural Network (NN) variant called Deep Convolutional Generative Adversarial Network (DCGAN). Following TUI interactions, the Observer's viewpoint is captured and converted into an input vector for the DCGAN (see figure 5). The DCGAN generates an image corresponding to the input vector, where each pixel in the input vector triggers a pixel in the DCGAN output. The resulting image is then drawn onto the DeepScope feedback module. This section explores the dataset, model architecture and NN training.

3.1 Dataset and Model Training

Accurate pattern recognition using NN was already feasible in the late 1980's [18]. However, generating new data that well concatenates a given dataset is still considered a complex problem in Machine Learning [9]. Data generation using NN was greatly advanced with the introduction of Generative Adversarial Networks (GANs)[12]. GANs use two competing NN, Generator and Discriminator, that 'adverse' one another. The Generator attempts to create new data (such

as image, sound or text), and the Discriminator aims to nullify these 'fake' creations by comparing them to ground-truth data. The training is completed when the Generator creates indistinguishable samples that constantly fail the Discriminator [15].

3.2 Image-to-Image Translation

A branch of GAN is Conditional GAN (cGAN), in which both NN are given additional data that focuses the generation on specific targets [24, 33]. A notable use-case of cGAN is a pixel-wise conditional generation of images, also known as Image to Image Translation (I2I), or 'pix2pix' [16, 15]. In I2I, pairs of images are used for training, where the pixel values of one image are used as labels (also known as 'classes') of the other. This allows pixel-level prediction using spatial classification of regions in the image [4, 40]. DeepScope implements a lightweight variant of I2I that is fitted for real-time predictions on low-tier devices.

In practice, cGANs extends the classic GAN zero-sum objective function with additional 'class' data: $\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)}[\log D(x|y)] + \mathbb{E}_{z \sim p_z(z)}[\log(1 - D(G(z|y)))]$. Here, function V of Generator and Discriminator G, D attempts to minimize a delta between ground-truth data x (in this case, the pixel data) and z, which is the accumulated pixel distribution learnt on each training step (see figure 8). Unlike classic GANs, $\log D(x|y)$ denotes that the additional 'class' data y conditions the learning on both data x as well as on y class. In this respect, distributions created by cGAN generator do not only share resemblance to the learning dataset, but are trained to mimic high-level data structure.

3.3 Cityscapes Dataset

DeepScope's DCGAN model was trained on the Cityscapes dataset [8]. Cityscapes is composed of pairs of street-view images taken using a dashcam around 50 European cities, during different seasons, daytime and weather conditions. Each pair includes a street-view image and a corresponding segmented image with 30 semantic labels. These labels represent different streetscape classes, from buildings and roads to license-plates and road signs. For DeepScope, a pre-processing algorithm was designed to remove motion-blur, increase sharpness, saturation and remove color-casting which were common in a shots taken of a moving vehicle.

3.4 Model Architecture and Performance

DeepScope NN architecture was designed to allow fast predictions, minimal setup and high portability. The Generator has 16 layers with a U-Net [32], encoder-decoder structure. For performance purposes, the Discriminator has 5 layers and is using Leaky ReLU activation that has been shown to improve stability in training [31]. Commonly, DCGAN models benefit from high number of filters set to detect patterns on input data [32]. However, added filters increase the model size, which can gravely impact real-time performance and usability in low-tier devices. In order to still maintain attention to fine details, a shallow NN design with a random up-sampling of 150% was designed [16]. This design allows deployment on most client-side browsers or even on mobile devices, as long as Node.js and TensorFlow.js are supported [36].
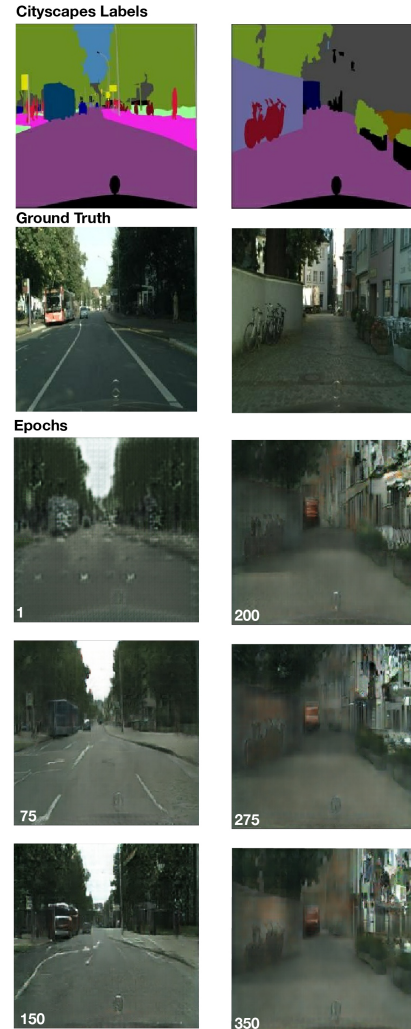
Figure 8. Test samples of different epochs during training: Right column shows quality degradation beyond 200epochs.

3.5 DCGAN Training and Results

As described in Model Architecture and Performance, portability and speed were key factors when balancing between image quality and model size. 20 training sessions were performed with 16,32,64 and 128 filters, with 50 to 2000 epochs. Resulting models were converted to a web format and tested for stability and response time on various client devices. A trained model with 64 filters and 200 epochs showed the best overall results. Models with less filters produced low-quality results; models with 300-2000 epochs demonstrated inconsistencies and 'mode collapse' [4]. Models with more filter were too slow to load and predict in real-time.

3.6 User Interaction Performance Test

In order to avoid interaction latency, two asynchronous processes were used: (i) prediction process and (ii) TUI interaction response. In preliminary user tests, the DCGAN model predicts at ~0.66sec/prediction and the TUI showed a fixed response interval of 50ms. Although the DCGAN slightly

trails the TUI, the observation showed that users tend to focus attention to the TUI before expecting the DCGAN output. In that sense, the overall user experience could be considered real-time with continuous design-and-feedback loops [10].

4 DISCUSSION AND CONCLUSION

This paper described DeepScope, a tangible urban design platform for real-time street-view visualization. Visualizations are created using a Deep Convolutional Generative Adversarial Network (DCGAN) trained on the Cityscapes dataset. A tangible user interface for rapid urban prototyping was created for iterations and feedback. The rest of this section will discuss the strengths, weaknesses, threats and opportunities of this work.

4.1 Strengths

DeepScope is designed to allow experts and non-professionals alike to collaboratively experiment with urban design scenarios and real-time feedback. The platform can augment early stages of cityscape design with vivid street-view visuals. These stages have major impacts on urban form and spatial organization of cities, but commonly lack sufficient design representation [5]. Unlike traditional CAD tools, the complexity of creating a 3D urban scene is carried out by DeepScope pre-trained NN. Designed for the web, DeepScope is 'platform-agnostic' and requires minimal computational resources, making it more accessible and portable for public participation. Lastly, the 'unpolished' nature of the GAN outcome allows designers and regulators to focus on the overall 'feel' of the 'Image of the City', instead of highly-specific design details [21].

4.2 Weaknesses

Despite the promise of generative NN, GANs have several drawbacks. First, GANs require large and properly labeled datasets; as such, creating a new Cityscapes dataset for other geographies will involve significant efforts. Several emerging methods suggest decoupled [40] and label-less learning [20], which can simplify the labeling effort. Nevertheless, dataset collection and partial labeling would still be required. Moreover, GANs tend to be inconsistent during learning process, as explored in DCGAN Training and Results [34]. Lastly, the DeepScope GAN would not be able to visualize non-street view angles: Since the Cityscapes dataset was captured using a vehicle dashcam, only matching angles produce reasonable predictions [33]. This issue is common amongst supervised NN, and requires either non-supervised methods or more extensive datasets.

4.3 Threats

The rising popularity of GANs is greatly attributed to their ability to 'create'. Nevertheless, GANs tend to be unpredictable in their results. When it comes to the design practice, certain degree of 'creative freedom' might be desired, yet unpredicted tools might cause resentment or misleading impressions. In the context of DeepScope, the same street-view angle with the same urban-design setup, might produce different visual results if ran twice. While the authors perceive that as a design feature and manifestation of Kevin Lynch's

'Imageability' concept [21], others might observe this as a sign to an untamed technology. Additionally, NN are strictly bounded by their architecture and training data. Tempered NN or datasets can greatly affect the outcomes of the model and inject bias into the results. With machine-learning tools becoming mainstream in the design industry, these concerns should be addressed by testing, validating and open-sourcing design tools, models and data.

4.4 Opportunities

DeepScope can be improved in several aspects: First, emerging NN architectures and training parameters can improve the DCGAN results. Other methods, such as VAE or auto GANs, can produce finer results with greater control [23]. As well, extending the training datasets to different urban environments could yield more versatile representations. Lastly, the TUI can be improved to include multi-scale environments and more finer-grained editing capability.

4.5 Applications and Real-World Implementations

As mentioned in section 2.1, a prototype of DeepScope was constructed and tested in an active demonstration area at the MIT Media Laboratory in Fall '19. During this period, hundreds of design-professionals and random visitors interacted with the tool; their input was incorporated into the UI/UX design and HCI factors of the tool. The most prevailing comments were gathered to appear in this SWOT Analysis.

Currently, DeepScope development is on two major trajectories: *[i] Implementation of DeepScope as a native CityScope module:* The CityScope platform is in the process of adapting a micro-services architecture, in which modules of urban analytics can be 'plugged' into the system when needed. In this context, DeepScope would be used not as a standalone tool, but rather as an additional analysis layer, side by side with other urban matrices (such as noise, ABM, traffic, etc.)

[ii] Real-world testing and deployment: As a standalone tool, DeepScope is now being tested as part of an urban modelling system used in an international urban design competition for a major European city. In the context of this competition, the tool will be used to evaluate street-language and urban form for different design proposals which are at an early schematic design phase. Using DeepScope, both the design teams as well as the jurors would be able to quickly evaluate early-stage urban design decisions, and amend them during sessions. Lastly, the tool might also be incorporated in public participation process in which participants could get better understanding on the implications of different planning alternatives.

More broadly, DeepScope might hint to a future of insightful CAD tools, spanning beyond digital rulers and drafting aids. Such tools would not only expedite tedious tasks, but might be able to leverage the power of advanced computation and become insightful design 'companions'.

REFERENCES

1. Al-Kodmany, K. Using visualization techniques for enhancing public participation in planning and design: process, implementation, and evaluation. *Landscape and urban planning 45*, 1 (1999), 37–45.

2. Andrews, M. The view from the road and the picturesque. In *The aesthetics of human environments*. 2007, 272–289.

3. Appleyard, D., Lynch, K., and Myer, J. R. *The view from the road*, vol. 196. MIT press Cambridge, MA, 1964.

4. Arjovsky, M., Chintala, S., and Bottou, L. Wasserstein generative adversarial networks. In *International Conference on Machine Learning* (2017), 214–223.

5. Batty, M., Chapman, D., Evans, S., Haklay, M., Kueppers, S., Shiode, N., Smith, A., and Torrens, P. M. Visualizing the city: communicating urban design to planners and decision-makers.

6. Brusaporci, S. The importance of being honest: Issues of transparency in digital visualization of architectural heritage. In *3D Printing: Breakthroughs in Research and Practice*. IGI Global, 2017, 333–360.

7. Carr, S., and Schissler, D. The city as a trip: perceptual selection and memory in the view from the road. *Environment and behavior 1*, 1 (1969), 7.

8. Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., and Schiele, B. The Cityscapes Dataset for Semantic Urban Scene Understanding.

9. Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., and Bharath, A. A. *IEEE Signal Processing Magazine 35*, 1 (2018), 53–65.

10. Deber, J., Jota, R., Forlines, C., and Wigdor, D. How much faster is fast enough?: User perception of latency & latency improvements in direct and indirect touch. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM (2015), 1827–1836.

11. Drettakis, G., Roussou, M., Reche, A., and Tsingos, N. Design and evaluation of a real-world virtual environment for architecture and urban planning. *Presence: Teleoperators and Virtual Environments 16*, 3 (2007), 318–332.

12. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Front Matter. *Environmental Fluid Dynamics* (2013), iii.

13. Ishii, H., Ben-Joseph, E., Underkoffler, J., Yeung, L., Chak, D., Kanji, Z., and Piper, B. Augmented urban planning workbench: overlaying drawings, physical models and digital simulation. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, IEEE Computer Society (2002), 203.

14. Ishii, H., Ratti, C., Piper, B., Wang, Y., Biderman, A., and Ben-Joseph, E. Bringing clay and sand into digital design—continuous tangible user interfaces. *BT technology journal 22*, 4 (2004), 287–299.

15. Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. Image-to-image translation with conditional adversarial networks. *arxiv* (2016).

16. Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), 1125–1134.

17. Kempenaar, A., Westerink, J., van Lierop, M., Brinkhuijsen, M., and van den Brink, A. "design makes you understand"—mapping the contributions of designing to regional planning and development. *Landscape and Urban Planning 149* (2016), 20–30.

18. LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. Backpropagation applied to handwritten zip code recognition. *Neural computation 1*, 4 (1989), 541–551.

19. Lovett, A., Appleton, K., Warren-Kretzschmar, B., and Von Haaren, C. Using 3d visualization methods in landscape planning: An evaluation of options and practical issues. *Landscape and Urban Planning 142* (2015), 85–94.

20. Lucic, M., Tschannen, M., Ritter, M., Zhai, X., Bachem, O., and Gelly, S. High-fidelity image generation with fewer labels. *arXiv preprint arXiv:1903.02271* (2019).

21. Lynch, K. *The image of the city*, vol. 11. MIT press, 1960.

22. Mekni, M., and Lemieux, A. Augmented reality: Applications, challenges and future trends. *Applied Computational Science* (2014), 205–214.

23. Mescheder, L., Nowozin, S., and Geiger, A. Adversarial variational bayes: Unifying variational autoencoders and generative adversarial networks. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, JMLR. org (2017), 2391–2400.

24. Mirza, M., and Osindero, S. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784* (2014).

25. Mrdoob. mrdoob/three.js, May 2019.

26. Mueller, J., Lu, H., Chirkin, A., Klein, B., and Schmitt, G. Citizen design science: A strategy for crowd-creative urban design. *Cities 72* (2018), 181–188.

27. Noyman, A. Powerstructures: The urban form of regulation. Master's thesis, 2015.

28. Noyman, A., Holtz, T., Kröger, J., Noennig, J. R., and Larson, K. Finding Places: HCI Platform for Public Participation in Refugees' Accommodation Process. In *Procedia Computer Science*, vol. 112 (2017).

29. Patten, J., Ishii, H., Hines, J., and Pangaro, G. Sensetable: a wireless object tracking platform for tangible user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2001), 253–260.

30. Pearce, P. L., and Fagence, M. The legacy of kevin lynch: research implications. *Annals of Tourism Research 23*, 3 (1996), 576–598.

31. Radford, A., Metz, L., and Chintala, S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks.

32. Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, Springer (2015), 234–241.

33. Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., and Chen, X. Improved techniques for training gans. In *Advances in neural information processing systems* (2016), 2234–2242.

34. Shin, H., Lee, J. K., Kim, J., and Kim, J. Continual learning with deep generative replay. In *Advances in Neural Information Processing Systems* (2017), 2990–2999.

35. Shiode, N. 3d urban models: Recent developments in the digital modelling of urban environments in three-dimensions. *GeoJournal 52*, 3 (2000), 263–269.

36. Smilkov, D., Thorat, N., Assogba, Y., Yuan, A., Kreeger, N., Yu, P., Zhang, K., Cai, S., Nielsen, E., Soergel, D., et al. Tensorflow. js: Machine learning for the web and beyond. *arXiv preprint arXiv:1901.05350* (2019).

37. Smith, A., Dodge, M., and Doyle, S. *Visual communication in urban planning and urban design*. University College London, Centre for Advanced Spatial Analysis (CASA), 1998.

38. Sutherland, I. E. Sketchpad a man-machine graphical communication system. *Simulation 2*, 5 (1964), R–3.

39. Yan, J. An evaluation of current applications of 3d visualization software in landscape architecture.

40. Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision* (2017), 2223–2232.
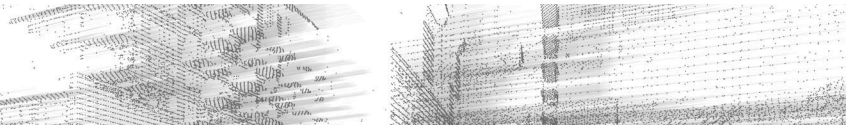
Evaluating QRD Arrays generated with Shape Grammars

Jonathan Dessi-Olive[1] and Timothy Y. Hsu[2]

[1]Department of Architecture
Kansas State University
Manhattan, Kansas, Country
jdessiolive@ksu.edu

[2]Department of Music and Arts Technology
Indiana University Purdue University - Indianapolis
Indianapolis, Indiana, United States
hsut@iu.edu

ABSTRACT

This paper broadens the scope of on-going research that introduces rule-based methods for designing creative acoustic arrays: from single-frequency QRD-based panels to those which account for multiple design parameters such as frequency, absorption, and number of wells. Expanded parameters create more complex arrays that give acousticians a finely tuned and intentional acoustical response as it relates to these input rules. Numerical simulations show three-dimensional directivity responses for these shape grammar generated diffuser arrays. A visual complexity coefficient is presented to quantify the aesthetic component to designs of arrays of acoustic panels. Along with spatial simulated diffusion response, diffusion coefficients, and the visual complexity coefficient, designers and acousticians will be afforded ways to create both acoustical effective and visually interesting arrays.

Author Keywords

Architectural acoustics; acoustic simulation; Schroeder diffuser arrays; shape grammars; design computation.

1   INTRODUCTION

Often due to modest design budgets for acoustical treatment in buildings, acousticians and architects will commonly draw from historical examples, best practices from prior performative success, and known acoustic products. The result of these practices usually deliver visually similar arrangements of the same panel and thus suffer from a lack of visual character, creativity, or complexity. High-budget and custom-designed spaces such as the Elbephilharmonie [8] allow for fabrication of panels specific to the concert space that also enhance the visual character of the hall. Such design scenarios are uncommon however, so new methodologies are needed that encourage designers of acoustic panels and arrays to produce more visually evocative treatments that perform reliably and predictably. This paper builds off previous research on a shape grammar-based methodology for designing arrays of quadratic residue diffusers (QRDs) that provides a means of creative and intentional deployment of acoustical panels that perform acoustically and meet a visual criteria [4]. Commonly, monolithic grids of one-dimensional QRD are applied to a diffusion area. Such M x N arrays (where M is the number of panels in the x direction and N is the number of panels in the y direction) are visually homogeneous and are assumed to have a one-dimensional diffusive response. The design possibilities for QRD-based acoustic arrays have been shown to be visually surprising and unexpected [4] through the application of shape grammars. The objective of addressing design homogeneity in architectural acoustics is extended by demonstrating the following: a clarification and further development of shape grammar rules to include a parametric absorption coefficient to allow for more complex moves and array outcomes; a proposal for an aesthetic evaluation algorithm that provides a visual complexity coefficient; and initial methods to numerically evaluate the acoustic performance of these arrays.

2   REVIEW OF DIFFUSION AND ABSORPTION

Acousticians use different surfaces to control reflections in a room. The porosity, material, construction, and shape of the surfaces in that room are highly impactful parameters that affect the sound propagation and acoustic quality of the space. It is common to see a mixture of diffusion and absorption panels in acoustically sensitive spaces.

2.1 Diffusion and QRDs

Acousticians control reflections by using angled surfaces to direct the energy to specific areas of the room or by using diffusive surfaces to spread the reflected energy to many different directions [3]. This also extends the time response of the reflection as compared to a purely specular reflection. Acoustic diffusers are specially designed devices that employ surface roughness characteristics that cause waves to be reflected back at various non-specular directions. Surface roughness causes phase differences in reflections that are directly caused by the depth of these surface perturbations [9]. One-dimensional QRDs, a special type of Schroeder diffusers [12], are amongst the most common types of diffusers used in practice. Dessi-Olive and Hsu [4] outline the fundamentals of QRD design and the governing equations. Of the parameters that govern these equations, N

and λ_0 are of interest for this study, where N corresponds to the number of wells and is restricted to be a prime number, and λ_0 is the design wavelength. By following the calculations specified in [4], an N=11 QRD panel with a specified design frequency λ_0=1000Hz would have a depth sequence, d_n of:

$$d_n = \{0, 16, 62, 140, 78, 47, 47, 78, 140, 62, 16\} \text{ mm.}$$

When the number of wells becomes large, then unintentional absorption may occur due to viscous effects. When the number of wells is small, then cross modes within the wells can happen. In practice, the well depths range from approximately 2.5 to 5 cm [3]. Historically, diffusion has been deployed in concert halls through different methods including, architectural features, artistic installations, acoustic devices (QRDs or other purposefully built diffusers), and customized panels. In spaces that cannot support such detailed acoustical design, diffusers are most often added to the space after construction.

2.2 Absorbers

Absorptive surfaces turn sound energy into mechanical or thermal energy at impact, thereby reducing the energy of the reflected wave [9]. Absorption is frequency dependent and a vital factor in room design, as it directly relates to the reverberation time of the room. The absorption coefficient, α, is the commonly accepted metric to describe acoustic absorption [3]. The coefficient is a number between 0 and 1 that indicates what percent of energy is absorbed, where $\alpha = 1$ is total absorption and α=0 is total reflection. Typically, alpha is measured to ISO 354:2003 [6] in a reverberation room or with ISO 10534-2:1998 [7]. Typical values for a 1" porous absorber are close to 1 above 1000 Hz and around 0.7 near 500 Hz [10]. Similar 4-inch (about 100mm) panels show an alpha of 1 as low as 200 Hz. Absorbers are used by acousticians to reduce the strength of reflections from a wall or surface. Uncontrolled strong, sometimes specular, reflections can sometimes cause acoustical defects such as focusing and flutter echo [9]. Additionally, absorbers can be used to reduce the overall reverberation time in the room by controlling the reverberant field. This was shown empirically by Sabine [11] when he related volume with absorption. The equation, known as the Sabine equation (below), shows that reverberation time is directly proportional to volume and inversely proportional to the total absorption. Essentially, a larger the absorption area will lower the reverberation time.

$$RT = \frac{0.161 V_m}{S \bar{\alpha}}$$

where,

V_m	is the volume in m^3,	
S	is total surface area,	
$\bar{\alpha}$	is the average absorption coefficient.	

Absorbers generally fall under two categories, porous and resonant absorbers [9] but this paper will only focus on porous absorbers. Porous absorbers use fibrous materials or open-celled materials to trap some of the reflected sound

wave [1]. In these absorbers, the incident energy is turned into mechanical energy and heat, resulting in energy loss. Flow resistance is an important factor that determines absorption as this resistance causes a pressure gradient that is related to the specific acoustic impedance of the material. Furthermore, the thickness of the material is directly related to the absorption. Thick porous absorbers generally provide a greater amount of absorption at a lower frequency than thinner panels. Porous absorbers are considered to be more broadband than resonant absorbers with increasing absorption with increasing frequency.

3 REVIEW OF SHAPE GRAMMARS

Shape grammars [14, 15, 16] are a computational design methodology that have been used in several areas of art, design, and engineering, but their application to acoustics is recent. Shape grammars let designers work visually through recursive rule-based transformations of shape descriptions which operate under the notion of *embedding* [Figure 1]. Embedding remains true for any shape made of points, lines, faces or solids (0, 1, 2, or 3-dimensional shapes). Since shape grammars were introduced by Stiny and Gips [16], they have evolved significantly. Once relying on linguistic analogies [14], shape grammars have since been simplified to *embedding, sets of shape rules, labels, weights, and schemas* [15].
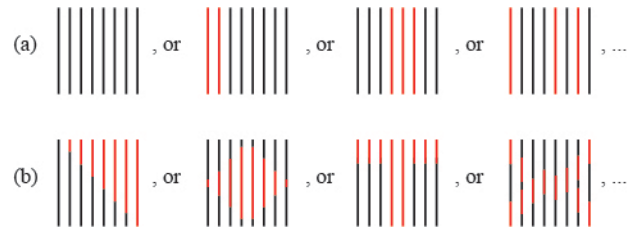
Figure 1. Examples of embedding: (a) Lines identified as themselves (identity); and (b) "Sub-line" configurations of the lines embedded in the field.

4 A VISUAL ALGORITHM FOR ACOUSTIC PANEL ARRAYS WITH DIFFUSION AND ABSORPTION

This paper builds on the authors' previous work [4] that proposed a shape grammar-inspired computational design method for designing arrays of high performance QRD panels that are visually creative. As noted before, despite numerous applications of shape grammars across virtually all areas of design, very little has been done with acoustics outside of this research. The 4-step computing process is specified in this section and illustrated in Figure 2.
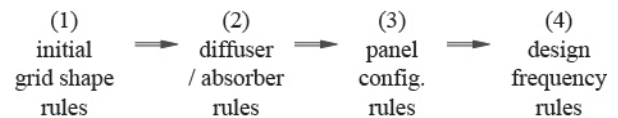
Figure 2. Computing steps for acoustic arrays generated using the shape grammar. Rules are applied reclusively in each step

4.1 Initial Grid Shape Rules

Computing begins when the designer locates one or more square panels in the diffusion area. The initial shape (*i*) is a labeled square that represents a 30cm x 30cm acoustic panel; not yet specifying whether the panel is a diffuser or absorber panel. The circle serves to indicate the center of each panel and prevent unintended scaling. Figure 3 shows a set of *grid shape rules* that aggregate initial shapes. In each rule an initial shape is copied and moved relative to the original. The center label makes it possible to apply grid shape rules in different orientations, according to the "copy and move" schema, $x \rightarrow x + t(x)$. Establishing a grid configuration helps to later categorize and evaluate designs of in terms of their underlying organization, their formal specification, and assumed acoustical response.
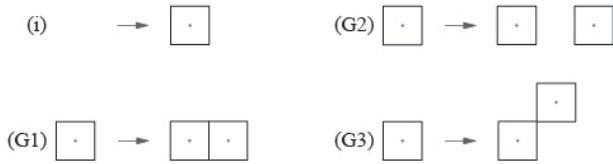
Figure 3. Initial grid shape rules place an initial labeled shapes and aggregate them to make regular or irregular grids.

4.2 Diffuser and Absorber Shape Rules

Once an initial panel configuration has been generated panel-by-panel or with a predetermined M x N grid, *diffuser and absorber shape rules* associate the shape descriptions of panels with a base-type of acoustic device that has known performative properties. For now, *Step 2* specifies either QRDs or basic absorber panels.

4.2.1. QRD Shape Rules

QRD shape rules [Figure 4] replace the circle label on initial shapes with the fins of one-dimensional Schroeder diffuser panel or the L-shaped fins of a two-dimensional panel. In this paper only QRDs with seven wells are specified because they are the most commonly used and commercially available QRD product, but the grammar can support any number of wells. QRD rules include a label system for well-depth sequence: a red circle is in the right well by default, indicating the location of the zero-depth well; a red dash indicates the end of the well sequence.
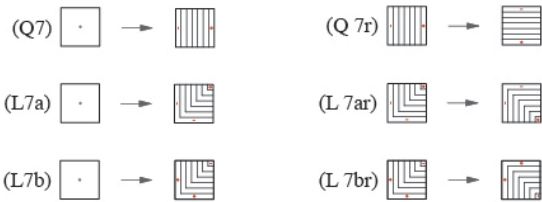
Figure 4. QRD shape rules for one and two-dimensional phase grating diffuser panels with seven wells.

The initial shape (i) does not specify an orientation, so the fins and the well-depth sequence can be placed in several orientations. A common transformation used in physical deployments of QRD panels is a ninety-degree rotation. Figure 4 specifies a set of rules that rotate each of the panel types according to the schema, $x \rightarrow t(x)$.

4.2.2. Absorber Rules

Absorber panels give designers expanded parameters to use for intuitively generating outcomes with diverse acoustical responses. *Absorber shape rules* [Figure 5] work in a similar manner as the QRD panel rules. Initial shapes are replaced with a shape description of an absorber panel of the same size. The red circle label from initial shapes is replaced with a hatch pattern and a blue cross. To account for design processes that start with monolithic M x N grids of vertically oriented QRD panels, rule (Q-A) is provided to replace a QRD N7 panel with an absorber panel of the same size. For now only one generic type of absorber is specified. Later, any *acoustical devices such as resonant absorbers, bass traps, hybrid absorptive diffusers, and other known acoustic products could be included in the grammar to provide substantially more possibilities of designs that are finely tuned and intentional both in terms of acoustical response and visual complexity*.

Figure 5. Absorber panel rules replace either an initial shape or an N7 QRD panel with a shape representation of an absorber panel.
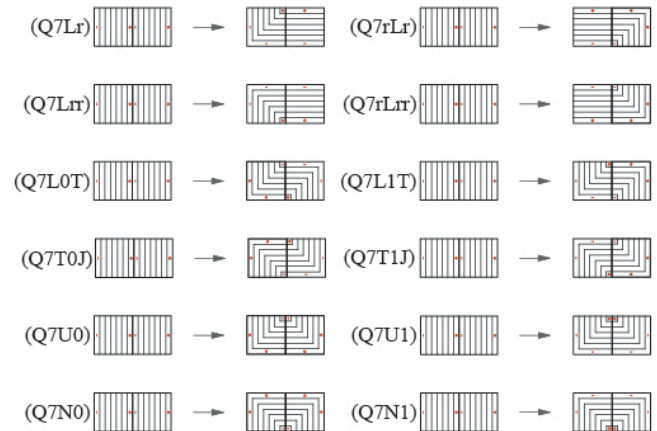
Figure 6. Compound panel configuration rules find two panels in their default orientations to perform transformations including rotating a 1D QRD and/or replacing it with a 2D panel.

4.3 Panel Configuration Rules

Rotation and replacement rules are a powerful means of introducing variation to arrays of panels; however when applied one-by-one, they can be tedious to use. This is particularly true when working with large arrays. A class of *panel configuration rules* lets designers make compound transformations in the array. Multiple transformations are made to groupings of shapes. The rules in Figure 6 make transformations and replacements to pairs of side-by-side,

vertically oriented, N7 QRD panels. In each, one-dimensional panels are rotated and/or replaced with two-dimensional L-shaped QRDs. Though simple, the rules illustrate how expansive the design possibilities are as new rules are added to the grammar.

4.4 Panel Performance Rules

Panel performance rules use weights to specify expected performative criteria of the panels in a design. Tones of color are applied to shape descriptions in a design to indicate diffusion design frequency or absorption co-efficient; both informing the depth of the acoustic panel.

4.4.1. Diffuser Design Frequency

Design frequency rules specify the expected diffusion frequency of the panels by applying a tone to the QRD shape description. Common design frequencies for diffusion surface treatments were assigned a tone of purple. The five design frequency rules in the left column of Figure 7 demonstrate how to apply a tone corresponding to an intended design frequency. On the right are rules that assign a design frequency to 2D L-shaped QRD panels at 500Hz and 2000Hz. The remaining 2D design frequency rules could be specified following the example of the rules below. Applying different tones in a single array presents opportunities to generate multi-band diffuser treatments.
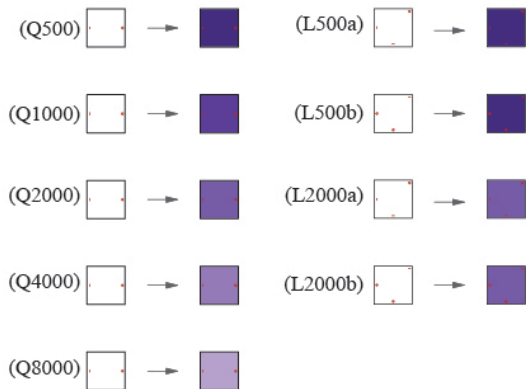
Figure 7. Design frequency rules apply a tone to panels in a design to specify design frequency. For example, (Q500) specifies a design frequency of 500Hz for a one-dimensional QRD panel.

Figure 8. Absorber thickness rules for absorbers with thicknesses of 25, 50, 100 and 150mm.

4.4.2. Absorber Thickness Rules

Absorber thickness rules specify the thickness of absorber panels in a design by applying a tone to the shape description. Absorber rules look for a label (in this case a blue cross) to apply a red tone, corresponding to panel thickness. Four

thicknesses are specified in Figure 8. Additional tones could be used for other thicknesses.

4.5 Example QRD and Absorber Computations

The rules specified above have the ability to generate known and monolithic designs of diffuser arrays, but also surprising and unexpected designs in the language. Figure 9 shows examples of step-by-step computations using rules in the grammar. Beginning with 2 x 2 monolithic arrays, rules are applied to make different designs. In Figure 9(a), rule (QRD 7r) is applied to rotate the one-dimensional panels in the array into a familiar configuration with two-dimensional diffusion. Figure 9(b) uses compound transformation rules to make 2 x 2 diffusion arrays with both one and two-dimensional QRD panels. Figure 9(c) generates a design with all L-shaped QRDs.
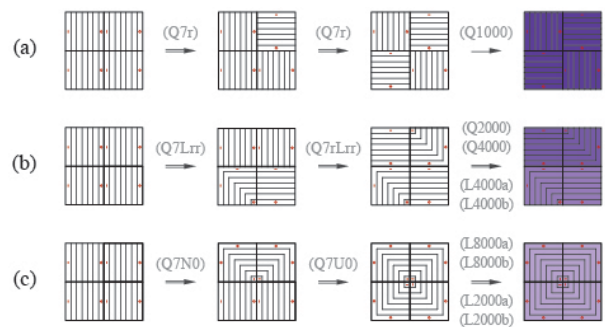
Figure 9. Examples of computing designs of 2 x 2 panel arrays with one and two-dimensional QRD panels. Applying grammar rules produces designs in the language at each step.
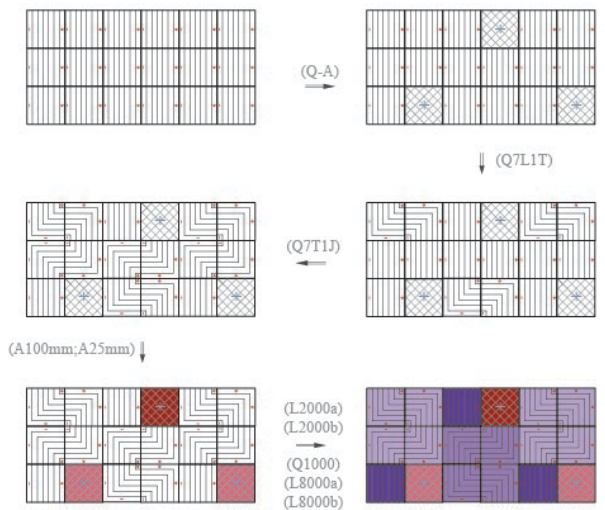
Figure 10. Example computing a 6 x 3 panel array. The design uses a combination of 1D and 2D diffusers absorber panels.

More complex arrays can be generated by applying rules differently - and to larger numbers of panels. Figure 10 shows an example of a 6 x 3 panel array that uses one and two-dimensional QRDs with design frequencies of 1000 Hz,

2000Hz, and 8000 Hz, and absorber panels with thicknesses of 25mm and 100mm. With more iterations, a designer who has experience using the rules, can continue generating visually evocative, irregular and sometimes surprising patterns for acoustic arrays. Once performative assignments are made, a designer can make material and fabrication specifications based on the visual descriptions generated by the grammar.

5   AESTHETIC EVALUATION OF QRD ARRAYS

A major goal of this research is to address the critical issue of design homogeneity in architectural acoustics. There is no single way in which design objects can be ordered in terms of their aesthetic quality. Throughout shape grammar scholarship, several methods have been proposed that specify aesthetic value unique ways. *Shape decompositions,* evaluate shapes by enumerating all the shapes made from a set or rules and the possible sub-sets of shapes in the language. This method is discussed by Stiny in the paper "Two exercises in formal composition" [13]. Recently many shape grammar interpreters have been implemented in computer software to evaluate sets of rules by automatically enumerating their design possibilities and their decompositions [2, 5, 19]. Another method introduced by Stiny and Gips [18] based on the *constructive mode of understanding*, interprets objects in terms of how they are made (rules) and how they are described. The evaluation algorithm E_z assigned high aesthetic value to an interpretation in which, the length of the description of the input was short compared to the length of the description of the output [18]. These evaluation criteria were applied to a grammar for generating plans Palladian Villas [17]. This section proposes an evaluation algorithm - inspired by E_z - that assigns numerical values to shape grammar-generated acoustical arrays as a means of classifying, comparing, and ranking different designs. Designs with the same underlying configuration of initial shapes (i) are given a visual complexity score. The lowest scoring design is the most homogeneous design for a specified configuration. In this case made entirely of seven-well, one-dimensional QRD panels. Higher scores are awarded to designs that introduce change by increasing the number of wells in panels, rotating them or replacing them with a two-dimensional diffuser.

5.1 Aesthetic Diffusion Evaluation Algorithm

For any acoustic array generated by the grammar described here, the aesthetic value of each panel in a design (E_i) is interpreted through a relationship between the evaluative criteria N, D, and T:

$$E_i = N_i \, (D_i + T_i)$$

where:

N	is number of wells in a panel,
D	is the diffusion dimension (1 or 2)
T	is a value indicating how many transformations were applied.

The average of all E_i values in a design describe the aesthetic value of an entire diffuser array (Ē) such that:

$$\bar{E} = \frac{1}{n} \sum_{i=1}^{n} E_i$$

The aesthetic diffusion score increases by making transformations to panels in the array. As rules are applied, the (T) value track how many transformations are applied to a panel in the design. Rules that increase the diffusion dimensions of a panel (D) also impact the score significantly. If a panel had a value of zero, it would be representing an absorber panel, but for now evaluations are limited to designs with diffusers.

5.2 Visual Complexity Coefficient

A visual complexity coefficient, ε is introduced as a coefficient value between 0 and 1, where 1 is maximum visual complexity and 0 is minimum visual complexity.

$$\varepsilon = \frac{\bar{E}}{\bar{E}_{max}}$$

5.3 Aesthetic Diffusion Evaluations of 2x2 QRD Arrays

The following section demonstrates using the aesthetic evaluation algorithm and complexity coefficient for 2 x 2 arrays of diffuser panels. The four designs selected are drawn in Figure 11. The designs in Figure 12a-c are taken from the examples in Figure 9. A monolithic application of vertical QRD panels is drawn in [Figure 12d] and serves as the constant, or lowest scoring possible design. It is assumed every panel has a design frequency of 1000Hz.
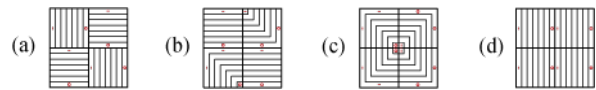
Figure 11. Four designs for shape grammar-generated QRD arrays being considered in this paper for evaluation.

For each design, the panels are numbered and their position is described (x, y), where (0, 0) specifies the position of the bottom left panel in an array. The second column of Figure 12a-d specify the values for N, D, and T and E_i for each design. For the designs considered here $N = 7$ for every panel.

Design (a) gains points for T by rotating two panels, but scores a low complexity coefficient. Design (b) improves its score by applying panel configuration rules that transform every panel in the design by rotating two, and replacing the others with two-dimensional panels. Design (c) demonstrates the maximum possible score within the limited parameters given because every panel in the design is replaced with two-dimensional panels, with an aesthetic value $\bar{E}_{max} = 21$. Design (d) demonstrates the lowest scoring possible design for panels in the language: a monolithic 2 x 2 arrangement.
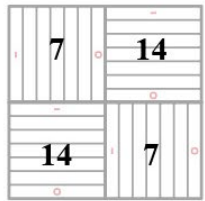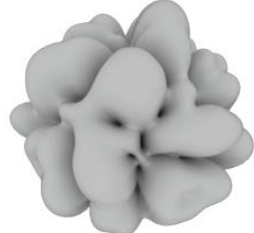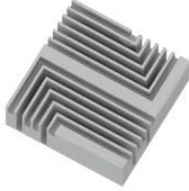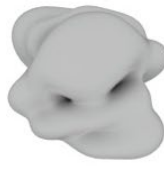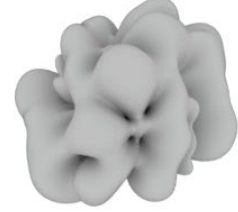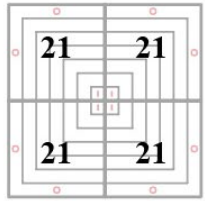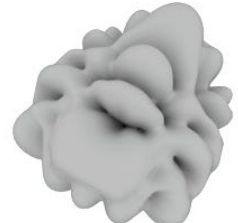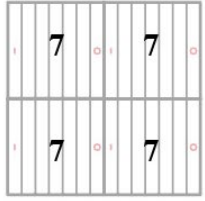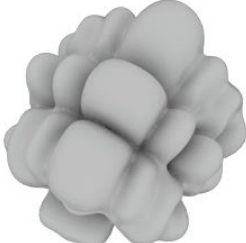
Panel Design	$E_i = N(D + T)$	1000 Hz	2000 Hz
(a)	$\bar{E}_a = (14+7+14+7)/4 = \mathbf{10.5}$ 7 14 14 7 $\varepsilon_a = \mathbf{0.50}$		
(b)	$\bar{E}_b = (21+14+21+14)/4 = \mathbf{17.5}$ 14 21 21 14 $\varepsilon_b = \mathbf{0.83}$		
(c)	$\bar{E}_c = (14+7+7+14)/4 = \mathbf{21}$ 21 21 21 21 $\varepsilon_c = \mathbf{1.0}$		
(d)	$\bar{E}_d = (7+7+7+7)/4 = \mathbf{7}$ 7 7 7 7 $\varepsilon_d = \mathbf{0.33}$		

Figure 12. Representative designs of arrays from Figure 11 were evaluated using the aesthetic evaluation algorithm and using simulation. The simulation produced three-dimensional polar response plots for each panel at 1000Hz and 2000Hz.

6 SIMULATION OF QRD ARRAYS

6.1 Simulation Method

For this paper, the acoustic simulations are limited to diffuser arrays. A perfectly matched layer finite element method (PML-FEM) allows for an accurate detection of the direction and amplitude of the diffused wave. The PML method uses an inner sphere, with a one meter radius, that houses the diffuser array and an outer sphere that corresponds to the wavelength of the test frequency. In this simulation, STL files are created, imported, and meshed. The mesh is set up so the minimum size of the element corresponds to the test frequency's wavelength divided by 120. The maximum size of the element corresponds to the test frequency's wavelength divided by 12. The size of these elements determines the complexity of the simulation. A plane wave

was impinged onto the array at 1000 Hz and 2000 Hz at both 0° and 45° degree incident angles. As the wave was reflected, a far field polar response was extrapolated from the inner sphere surface using PML data.

6.2 Simulation Results

In order to validate this method, a typical N=7 QRD was simulated in order to see if these results would match those found in [3]. The simulation method employed in this paper showed similar polar plots and also resulted in similar diffusion coefficients. 3D graphs were generated and diffusion coefficients were calculated for selected arrays. For 1000 Hz in particular, the 0° and 45° degree diffusion coefficients showed diffusion coefficients between 0.22 and 0.58. Table 1 shows the complete diffusion coefficients for four arrays at 0 and 45 degrees. One weakness of these

diffusion coefficients is that they are calculated along one plane. Thus, to better understand the spatial response, 3D plots shown in Figure 12 were generated and can help designers understand where lobes may exist and how to better shape the diffused reflections.

Panel Design	ε	1000 Hz - 0	1000 Hz - 45	2000 Hz - 0	2000 Hz - 45
(a)	0.50	0.24	0.58	0.25	0.42
(b)	0.83	0.31	0.57	0.44	0.50
(c)	1.0	0.23	0.46	0.34	0.39
(d)	0.33	0.24	0.43	0.39	0.44

Table 1. Diffusion coefficients for the four designs in Figure 12 compared to their corresponding complexity coefficient value.

7   DISCUSSION

This paper has presented several new results for on-going research that proposes a rule-based methodology for designing with acoustic panels. Three goals established the scope of work: to clarify and further develop acoustic shape grammar rules; to propose an aesthetic evaluation algorithm to evaluate and compare designs in the language based on their visual complexity; and propose initial methods to simulate the acoustic performance of shape grammar generated arrays and numerically predict the diffusion coefficients of designs.

A significant development in the grammar, presented in section 4, is the introduction of absorber panels and a parametric absorption coefficient specification to allow for more complex outcomes. The design process was improved so that panel transformations are made after the panel type is specified, making designing this way closer to designing with physical panels.

The aesthetic diffusion algorithm introduced in section 5 provides a means of interpreting designs generated by the grammar in terms of their visual complexity. A scoring system awards points for designs whose diffusers display more visually complex attributes that are directly tied to physical attributes of QRD panels. The minimum score for a design of a given topology of panels represents the design with the least number of wells, each in their default (vertical) orientation. The maximum score is achieved when rules have been applied to every panel in a design, using two-dimensional diffusers. In design spaces with small numbers of panels and limited design rules, the complexity coefficient can be useful to categorize design iterations. For designs of a certain value ε, a finite set of combinations of panels with the same complexity coefficient could be enumerated and explored further [Figure 14]. Ultimately, the aesthetic values of designs determined by the evaluation algorithm serve to later conflate specific aesthetic design decisions using grammar rules with data collected from acoustic simulations. Looking at both aesthetic and performative values of designs, will bring to light how acoustical response of an

array change as its design becomes more complex each time a rule is applied.
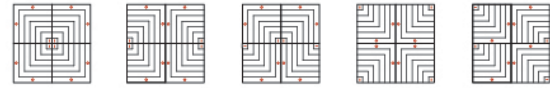
Figure 13. Example designs that all have the same complexity coefficient value and are made of the same panels. Though not yet tested, each would likely have its own unique polar response.

The acoustic simulations in section 6 begin to bridge the gap between applying design rules to QRD panels and understanding the impact those rules have in changing the acoustical response an array. Although quite preliminary, the three-dimensional plots in Figure 13 *show arrays with a higher value of visual complexity have reflections that display increased dynamism and dimensionalit*y. In Figure 13(d), the uni-dimensional nature of diffusion expected by a monolithic array of quadratic residue diffusers is clearly visible. As panels are rotated and replaced with two-dimensional diffusers in Figure 13a-c, the shape of the diffusion response changes significantly. For some, the visual characteristics of the diffuser panel resemble the visual characteristics of the 3D plot. Studying simple examples such as those presented here, can help gain sonic-visual intuition for the form-performance relation of shape grammar rules for QRD arrays. *The 3D plots provide visual evidence that a relationship exists between the appearance of an array and its corresponding response.* A future opportunity may involve designing directly with the 3D plots to allow for semi-customized reflections in rooms that previously could not be achieved. Together, these predictions and rule-based design processes fill the gap in acoustic design between the simple budget solution and fully customized/simulated and expensive solution. This opens the field of room acoustics to wide-ranging visual design possibilities that are fully cooperative and incorporated with the acoustic response.

8   CONCLUSION AND FUTURE WORK

This paper addresses the critical issue of design homogeneity in architectural acoustics, and how the acoustical response of arrays change as precise transformations are made. Shape rules introduce visual sophistication to homogeneous arrays of acoustical panels. The process of generating designs with the proposed method is interactive and non-deterministic; ultimately producing new and visually surprising arrays that can equal or exceed the acoustical performance of traditional 1-D QRD deployments. The QRD was originally chosen as a prototypical acoustical device with known geometries that perform predictably. There are several other known and more complex types of diffusers that could be given shape descriptions for diffuser form rules; their performance specification easily handled through systems of patterns, labels, and weights. A remaining issue with the grammar is diffuser and absorber panels are treated as objects (i.e.

symbols). For computational methods that are limited to symbolic representation, the design possibilities involve simply counting, recombining, or rearranging the objects. Future work on the methodology should strive to make better use of *embedding* to allow for unexpected scaling and overlapping to occur. The grammar should be able to allow for grid distortions; include complex grid patterns and tessellations that have several unique or custom panel shapes; fuse panels together by erasing gridlines made by initial shapes; and suggest a way to make material and fabrication specifications for designs generated this way.

For now, the aesthetic evaluation algorithm only accounts for the visual complexity of diffusers in an array; the design frequency does not impact the measure, nor are absorber panels considered in a rigorous manner, outside of $D = 0$. In future iterations, the algorithm should consider the aesthetic impact of absorber panels; assign higher values of T to certain shape rules that have greater magnitudes of impact on the aesthetic complexity of a design; and more closely study the impact of increasing the number of wells in the QRD panel. Typically the maximum value of N can be determined by how many wells would fit on a panel of a certain size within the recommended range of dimensions for the wells and fins [3]. In cases where more wells may be physically possible to fabricate for a given panel size, a QRD panel would take on absorptive properties. Further studies should investigate how increasing the number of wells in the QRD panel toward N_{max} impacts the acoustical response of the panel. Acoustic simulations should inform such a study to establish an understanding of the form-performance relation in such a panel. Future work in acoustic simulations will extend the FEM-PML method to include absorption and multiband diffusion. Using advanced metrics to describe diffusion in two dimensions will quantitatively complement the visual 3D plot results. Ultimately, this process will populate a database with a large number of diffusion coefficient values and 3D plots. Sorting, selecting, and optimizating algorithms will be developed to systematically categorize acoustic response as it relates to the aesthetic parameters.

REFERENCES
1. Bies, David A., and Colin H. Hansen. *Engineering Noise Control: Theory and Practice, 5th ed*. Boca Raton: CRC Press, 2017.

2. McCormack, J. and Cagan, J. Shape grammar interpreter. United States Patent US 7,502,511. US Patent Office, 10 March, 2009.

3. Cox, T. J. and D'Antonio, *P. Acoustic absorbers and diffusers: theory, design and application,* Third Edition. CRC Press, 2017.

4. Dessi-Olive, J. and Hsu, T. (2019) "Generating QRD arrays with Shape Grammars" in *the 10th SimAUD Symposium*, April 2019, Atlanta, USA.

5. Grasl, T. and Economou, A. "From Shapes to Topologies and back: an introduction to a general parametric shape grammar Interpreter." AI EDAM, 32(2), 208-224. 2018

6. ISO 354:2003 https://www.iso.org/standard/34545.html

7. ISO 10534-2:1998 https://www.iso.org/standard/22851.html

8. Koren, B.S and Müller, T. "Digital Fabrication of Non-Standard Sound Diffusing Panels in the Large Hall of the Elbphilharmonie". *Fabricate.* Eds. Menges, A. et al. UCL Press, 2017, 122-129.

9. Long, M. *Architectural Acoustics*, 2nd Edition. Academic Press, 2014.

10. RPG Acoustics, Absorber™ Panel Performance Data https://www.rpgacoustic.com/documents/2016/07/absp_absorbor-panel_acoustical-data.pdf/

11. Sabine Paul E. "The Beginnings of Architectural Acoustics." *The Journal of the Acoustical Society of America* 7.4 (1936): 242-48.

12. Schroeder, M.R. "Diffuse sound reflection by maximum-length sequences". *Journal of the Acoustical Society of America,* 57 (1975), 149-150.

13. Stiny, G. Two Exercises in Formal Composition. *Environment and Planning B,* 3, 1976, 187-210.

14. Stiny, G. Introduction to shape and shape grammars. *Environment and Planning B,* 7, 1980, 343-351.

15. Stiny, G. *Shape: Talking about seeing and doing.* The MIT Press, Cambridge, MA, USA, 2006.

16. Stiny, G. and Gips, J. Shape Grammars and the Generative Specification of Painting and Sculpture. *Information Processing 71,* 1972, 1460-1465.

17. Stiny, G. and Gips, J. An evaluation of Palladian plans. *Environment and Planning B,* 5, 1978, 199-206.

18. Stiny, G. and Gips, J. *Algorithmic Aesthetics: Computer Models for Criticism & Design in the Arts.* University of California Press, 1978.

19. Trescak, T. et al. "A shape grammar interpreter for rectilinear forms". Computer Aided Design. 44:657-670. 2012.

Multi-Objective Optimization of Robotically Bent In-Situ Reinforcement System

Milad Showkatbakhsh[1], Elif Erdine[2], Alvaro Lopez Rodriguez[3]

[1]Architectural Association (AA)
School of Architecture
London, UK
showkatbakhsh@aaschool.ac.uk

[2]Architectural Association (AA)
School of Architecture
London, UK
elif.erdine@aaschool.ac.uk

[3]The Bartlett School of
Architecture, UCL
London, UK
alvaro.rodriguez.14@ucl.ac.uk

ABSTRACT

This paper describes a novel process towards the application of multi-objective optimization as the form-finding process for the integration of computational design, fabrication, and construction sequences. The design and construction of a doubly curved large-scale prototype made of textile-reinforced GRC shotcrete with a robotically fabricated in-situ reinforcement system serves as the case study for the proposed methodology. Global geometry form-finding process takes into consideration the location and geometrical properties of the in-situ reinforcement rebar system, robotic rod-bending constraints, structural performance, and functional objectives. These criteria are integrated through the application of a multi-objective optimization method in order to formulate multiple trade-off solutions that possess multiple constraints (fitness objectives) which are primarily in conflict with each other and are intended towards automation in fabrication. The primary contribution of the research is the demonstration of a multi-objective optimization methodology that incorporates geometrical form-finding, material and fabrication constraints, and FEA as design drivers during the early stages of design. This optimization method can be further extended and utilized across a multitude of scales in order to save energy, materials, and cost in architectural projects.

Author Keywords

Form-finding; Multi-objective optimization; Data Driven decision; Robotic Fabrication; Robotic rod bending; Glass-reinforced concrete (GRC); Stay-in-place textile formwork

ACM Classification Keywords

I.2.8 Problem Solving, Control Methods, and Search (Heuristic methods); I.2.9 Robotics (Commercial robots and applications); J.5 Arts and Humanities (Architecture); J.6 Computer-aided design; J.6 Computer-aided manufacturing (CAM).

1   INTRODUCTION

The research presented in this paper outlines a novel methodology for the application of multi-objective optimization as the form-finding process towards the integration of computational design, fabrication, and construction sequences. The design and construction of a doubly curved large-scale prototype made of textile-reinforced GRC shotcrete with a robotically fabricated in-situ reinforcement system serves as the case study for the proposed methodology. One of the primary considerations of the research is to allow for the geometrical freedom to create a range of complex doubly curved geometries with the employment of functional, fabrication-related, and structural constraints.

Concrete, being one of the most extensive building materials in construction, can be utilized to create architectural forms with complex geometry. The formal flexibility of concrete raises the question of the application of formwork during the construction process. Furthermore, custom-made reinforcement strategies need to be developed both for the formwork and the reinforcement of the structure. While it is viable to produce formwork with complex geometries via advanced digital and robotic fabrication tools, a key consideration area is the reduction of form-work waste material in manufacturing methods. The potential to incorporate this constraint in the preliminary design process as a design driver will pose advantages in waste optimization as well as production costs. Recent investigations on textile stay-in-place formwork [6] and 3d-printed stay-in-place formwork demonstrate the advantages of employing this method by correlating the geometric flexibility of the formwork with the structural capacity of concrete [15].

Advances in the biological sciences and computation in recent years paved the way to mimic the principles of evolutionary science to solve common real-world problems. This problem-solving methodology comprises search and optimization procedures of single or multiple objectives. Evolutionary multi-objective optimization strategies have been utilized widely since the late 20th century as problem-solving methods. In the 1930s, Sewell Wright was one of the first figures who attempted to apply biological evolutionary principles as optimization processes in solving complex problems [20]. Halfway through the 20th century, John Holland [7], Rechenberg and Schwefel [17] and Fogel et al. [5] respectively developed genetic algorithms (GA),

evolutionary strategies (ES), and evolutionary programming (EP) independently from one another. Their findings eventually led to the creation of a unified field of Evolutionary Computation in the late 20th century [1].

A multi-objective optimization methodology that incorporates geometrical form-finding, material and fabrication constraints, and FEA as design drivers during the early stages of design is demonstrated through in this paper. This optimization method can further be utilized across different scales in order to save energy, materials, and cost in architectural projects. The incorporation of geometrical, material, and fabrication constraints with criteria related to robotic fabrication presents an integrative morphogenetic design methodology [16].

The one-to-one scale prototype presented in this paper is a case study to test the proposed methodology with the design and construction of a complex doubly curved prototype. The prototype, an urban furniture piece that accommodates seating areas, is made of textile reinforced GRC shotcrete with an integrated solution for a reinforcement system that is fabricated via robotic rod bending [3]. The dimensions of the structure are 1,580 mm. width, 3,850 mm. length, with a height that varies between 750 – 1,500 mm. The location of the case study is the outdoor area of Santral Istanbul Campus of Istanbul Bilgi University.

2 EVOLUTION AS A DESIGN METHODOLOGY

In the biological sciences, Genotype is a set of genes or instructions (codes) that performs as a blueprint for the development of the Phenotype. Phenotype is the physical expression (formal and behavioral manifestation) of the Genotype. In the field of evolutionary computation, and more specifically in the context of architecture and design disciplines, a genotype is equivalent to a set of instructions or codes that will produce the geometry, namely the phenotype.

An evolutionary model, as described by Ernst Mayr, includes a two-step process; random variation within the genotype of a phenotype, and the selection of the phenotype through environmental pressures [14]. In line with evolutionary processes in nature, the application of evolutionary computation in design is founded upon these two primary components of variation in the code responsible for generating the geometry (genotype) and the selection of the geometry (phenotype) that fits better in the environmental conditions. In the context of the application of evolutionary computation in design, the environment is equivalent to a set of fitness objectives (e.g. design constraints to be met). Evolutionary algorithm goes through a basic loop. It starts with the generation of an initial random population of solutions. It continues with modifications of genomes through random variations, followed by the evaluation of solutions based on their objective performance. The algorithm concludes with the selection of a group of solutions that correspond to a predefined selection mechanism [4] (Figure1). Through this iterative process of

generation, evaluation and selection, each phenotype will be evaluated based on a set of fitness objectives. Therefore, the formulation of the environment, that is the design problem and the calculation of the fitness objectives, in this process is essential for constructing a successful evolutionary model to produce meaningful design options.
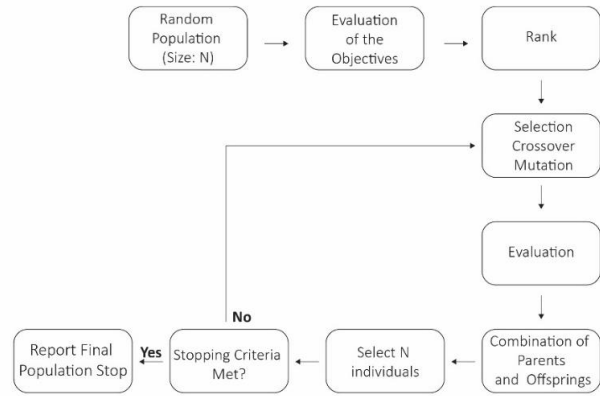
Figure 1. WallaceiX Core Algorithm (NSGA-II) pseudocode. (developed by [2]).

In recent years, evolutionary optimization processes have gained recognition in architecture and related design disciplines both in academia and practice.

Research conducted by Ayman Hassaan et al. explored the application of evolutionary optimization in the design phase by studying geometric formations of a facade at the early stages of design [10]. Yun Kyu Yi implemented NSGA-II algorithm in optimizing building facades with a similar objective to the previous research, nevertheless with a different design methodology [21]. In the building scale, the studies carried out by Machairas et al. employ genetic algorithms to address a set of conflicting objectives such as energy, comfort, and cost in early design phase [9]. In larger and more complex design domains, Makki et al. applied an evolutionary model to generate variations of urban forms that address a set of conflicting objectives [13].

Bionic Partition is amongst other projects carried out by the Living, serving as an example of the application of evolutionary optimization techniques in practice. Functional and structural constraints have been amongst objectives that they sought to address [19].

3 EXPERIMENT SETUP

The presented experiment utilizes multi-objective Non-Dominated Sorting Genetic Algorithm II (NSGA-II) developed by Deb. et al. [2] as the base algorithm upon which the evolutionary simulation is developed. Rhinoceros3D, Grasshopper3D and its plugin 'Wallacei' [11] are used to run the simulation, analyze the results thoroughly, and select the option to be further developed and fabricated. In the conducted experiment, the algorithm parameters within the evolutionary simulation have been set to the following

values. (Table 1) (For a thorough description of the terminology used in the simulation see [12]).

Parameter	Short Description	Value
Generation Size	*Number of individuals per generation*	30
Generation Count	*Number of generations in the simulation*	100
Crossover Probability	*Percentage of solutions that reproduce in each generation*	0.9
Mutation Probability	*The percentage of mutation 1/ (number of variables)*	1/n
Crossover distribution index	*Probability of similarity of the offspring to the parents*	20
Mutation Distribution Index	*Probability of similarity of the offspring to the parents*	20
Random Seed	*Random seed in the simulation*	1

Table 1. The algorithm parameters can be modified in WallaceiX UI first tab. Full description of each setting can be obtained from Wallacei Primer [12].

In order to test the construction of a complex doubly curved prototype, which is made through the application of textile-reinforced GRC shotcrete on an in-situ skeleton acting as the reinforcement system, a preliminary design idea was proposed by the students of the AAVS Istanbul 2019 workshop to serve as a piece of urban furniture with seating areas. In order to successfully fabricate the proposed design idea, a set of design objectives were addressed to calibrate the initial designed form and adjust it for fabrication and construction purposes. Three design objectives were specified to regulate and optimize the initial form-finding process in three categories. These include structural optimization to minimize structural displacements, functional optimization to accommodate for suitable seating areas, and cost optimization to regulate the amount of material used for the in-situ skeleton. The objective of the experiment is to integrate the above stated objectives in the early stages of design in order to realize a full-scale prototype with the given design choices and constraints.

The preliminary design idea was built algorithmically in Grasshopper 3D (visual scripting platform of Rhino3D) to establish the foundation of modelling the *environment* and subsequently the calculation of the fitness objectives (Figure 2). The predefined domains extracted from the fabrication constraints, such as the required curvature of the surface for shotcrete process and the accessibility of the surface by the shotcrete operator, were assigned to the motion vector of the control points of the curves by which the overall morphology is generated. Accordingly, the genotype of the experiment

was formed of 49 genes which modify the motion vectors of the control points of the phenotype. As the result, the size of the design space is calculated to be $9.6 * 10^{75}$. Three fitness objectives were defined to drive the optimization process to calibrate the preliminary design idea. They are:

- Structural Optimization: Minimizing the structural displacement through FEA.
- Functional Optimization: Creating a flat surface to serve as a seating area on the form.
- Cost Optimization: Minimizing the length of the rods to be used for creating the in-situ skeleton.
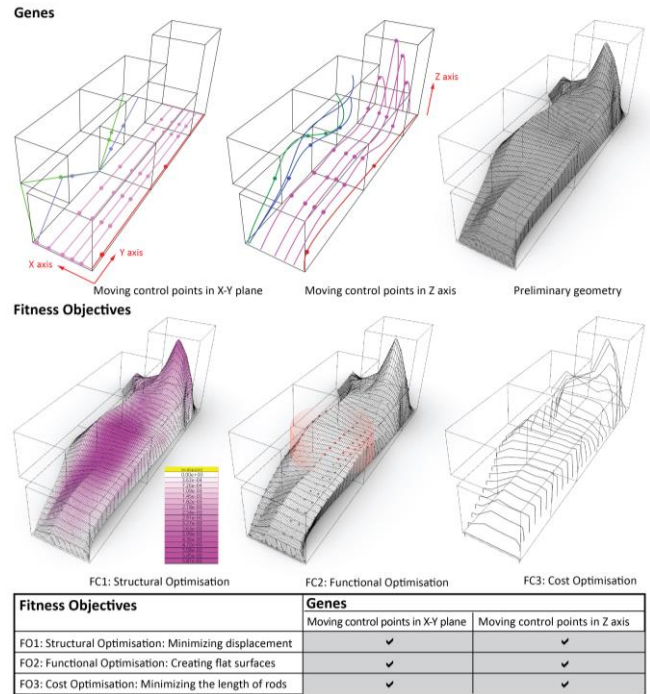
Figure 2. The genotype and the fitness objectives of the experiment.

The first fitness objective was set to minimize the structural displacement of the global geometry through FEA (Finite Element Analysis). As the final prototype was set to be fabricated with textile-reinforced GRC shotcrete application on an in-situ skeleton, the structural model of the phenotype was built as a concrete shell with an approximately 40 mm. thickness. In order to calculate the structural displacement of the generated phenotypes, an FEA analysis was conducted through Karamba3d, a plugin for Grasshopper3d [8]. The numerical value of the calculated structural displacement (in centimeters) was set to be the first fitness objective of the optimization process (Figure 2).

The second fitness objective was determined to evolve the phenotype towards developing flat surfaces to serve as seating areas. A set of control points on the surface of the geometry in a specified spatial domain of the phenotype was selected. The relative differences between the Z coordinates

of the selected points was calculated. The fitness objective was computed to minimize this calculated difference. Thereby, the points on the surface can be moved towards the same horizontal plane, and accordingly the phenotype will develop areas with a flat region suitable for seating (Figure 2).

The third fitness objective was assigned to drive the optimization process to the direction of evolving phenotypes to use as fewer metal rods as possible for the in-situ reinforcement system. The length of the bent rods (in meters) was computed as the third design objective to be minimized (Figure 2).

Due to the complexity of the form, robotic rod bending technique was selected for fabricating the in-situ skeleton onto which textile reinforcement would be placed, followed by the application of GRC shotcrete. Therefore, the angle of the rods to be bent by the industrial robotic arm was an essential factor to consider in the design process. In the iterative process of evolutionary optimization, fitness objective two (developing flat areas) and fitness objective three (minimizing the length of rods) would alter the angle of the rods freely if no constraint had been put in place. In order to automate and streamline the robotic rod bending process, the genes that were controlling the control points of the constructive lines of the phenotype (Figure 2), as well as their location were restricted in a domain which could only produce desirable bending angles. The angles that were divisible by 15 (i.e. 15°, 30°, 45°, 60° etc.) were set as the possible options that the rods could bend and evolve. The algorithmic approach that was used to constrain the generated angles through the iterative optimization process falls outside of the domain of this paper and will be explained in the context of another publication thoroughly.

Given the time constraints of the workshop, the optimization problem was limited to 30 individuals per generation with a total number of 100 generations (in total 3,000 generated solutions). The primary purpose of the conducted experiment was to calibrate the preliminary design proposed by the AA Istanbul visiting school 2019 students, structurally, functionally and materially, all independently from each other through a multi-objective optimization process.

4   EXPERIMENT RESULT AND SELECTED OPTION

By running the multi-objective evolutionary simulation for the main design problem, 3,000 genotypes/phenotypes with three fitness values per solution were produced. The visual analysis and recognition of the fitness performance of each individual, and the selection of the candidate solution for fabrication can be a highly inefficient process, subject to the visual preferences of the user. Therefore, analysis of the data associated with each phenotype plays a crucial role in the selection of the candidate solution to be further developed and fabricated.

The evaluation of the multi-objective optimization process commenced with a set of analyses to examine how the evolutionary simulation addressed the fitness objectives and how successful it performed in its entirety. Figure 3 illustrates the Parallel Coordinate Plot where each line represents an individual in the simulation and each axis is indicative of one objective.
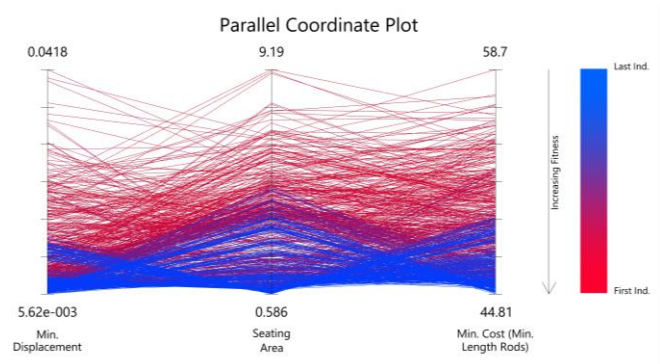
Figure 3. Parallel Coordinate Plot represents the overall performance of the optimization simulation. The graph was drawn by Wallacei Analytics Component.

The color gradient from red to blue illustrates the progression from former to later generations. Wallacei X, the evolutionary solver used for the experiment presented in this paper (the evolutionary solver of Wallacei plugin) optimizes the fitness objectives by minimizing the inputted values through search and optimization based on evolutionary principles. By introducing variations in the genotype and random mutations, individuals with lower numerical fitness values are considered the fittest in each iteration of the evolutionary simulation. These individuals are selected for the next round of the iteration by variation through crossovers, mutations, and selection. Figure 4 demonstrates that the fitness objectives have been successfully optimized towards the end of the simulation, as can be depicted by the blue lines appearing towards the bottom section of the plot. It is understood that the evolutionary process could generate individuals with high performance in all three fitness objectives in comparison to the preliminary design idea proposed.

Figure 4 compares the performance of the optimization run for all three fitness objectives independently and side by side. It comprises four graphs, Standard Deviation Graphs, Fitness Value Graphs, Standard Deviation Trendline (per generation) and Mean Fitness Value Trendline (per generation). Standard Deviation graphs illustrate that as the simulation advances, the population starts to improve for all fitness objectives. The progression of the SD graphs towards left is an indication of the mean fitness value's decrease per generation, thereby illustrating that the overall fitness objective is optimized. This can be cross-checked with the Mean Value Trend Line graphs on the bottom right which demonstrate the decrease in the mean fitness values per generation and subsequently increase in fitness performance by generating optimized solutions. SD value trendlines describe the decrease in the value of standard deviation per

generation which shows that the population is converging towards the optimized solutions by the end of the simulation run. Fitness Values Graphs illustrate that the later generations, illustrated in blue, are accumulated towards the bottom of the graphs, since they have less numerical values, and therefore describe the success of the evolutionary simulation run for all fitness objectives.

One of the challenges of the application of evolutionary principles through multi-objective optimization processes in design is the selection phase where a candidate or a set of candidate solutions need to be chosen for the subsequent stages of project development. In the context of this paper, one single solution was set to be selected to be further developed and fabricated accordingly.

For the purpose of selecting a candidate solution amongst the population of generated options, the population was filtered down into a smaller solution set to examine methodically and select the candidate design to be further developed and constructed. The population of 3,000 individuals was filtered to a sub-set of 86 individuals. These individuals comprised Pareto front solutions of the entire simulation (81 individuals), best performing individuals for each fitness objectives (3 individuals), an individual that perform as equal as possible for all three objectives (1 individual), and finally the individual that has the highest average rank amongst all population (1 individual) (For further description on the last two selection strategies, please refer to [12] and [13]). Due to the morphological similarities of these individuals, their comparison had to be accompanied by the data associated with each individual (Figure 5).

Individual number 24 in generation 89 which is the solution with the highest average rank amongst all population was selected to be further developed and fabricated (Figure 6). Table 2 shows the information of 10 out of 86 chosen individuals in detail. The selected individual, highlighted in orange, is amongst the highest performing individuals in the population. The diamond charts plotted next to each individual in the table demonstrate how well the individual performs for each fitness objective. In the diamond chart, each fitness objective is plotted on one axis where the center indicates best performance and the edges designate worst performance respectively. The selected induvial is ranked high for all objectives equally.

5 POST-RATIONALIZATION FOR ROBOTIC ROD BENDING

There were several post-rationalization adjustments that required to be applied to the selected individual to streamline

the fabrication process through automated robotic rod bending, such as the application of longitudinal rods, the calculation of spring-back in the rods, and a computational process to include the spring-back values in the algorithm in order to reach the target angles without tolerances.
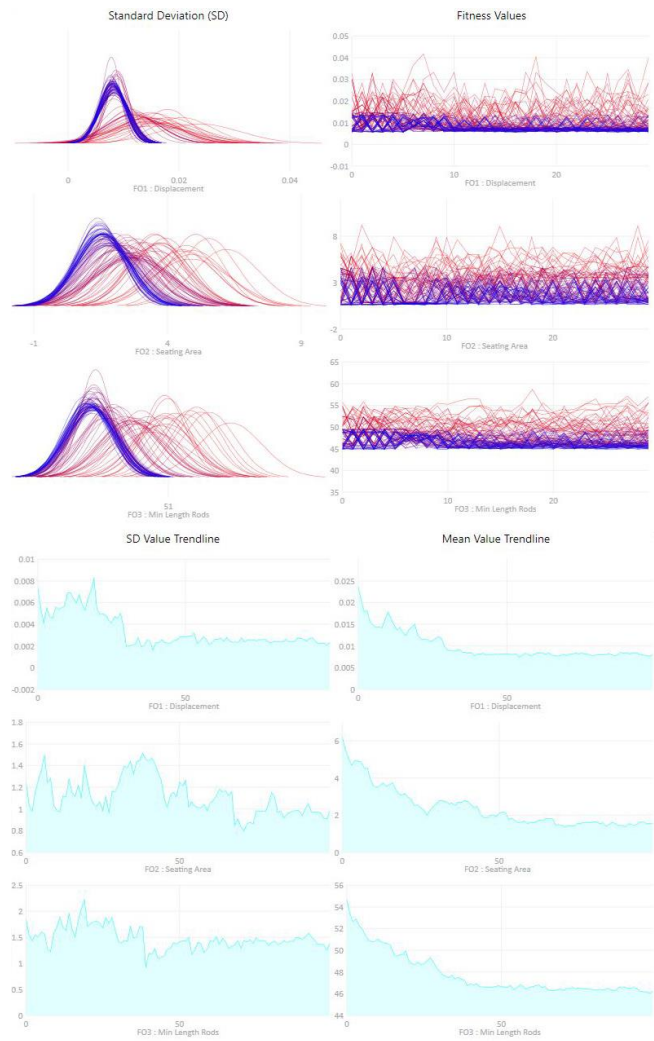
Figure 4. The graphs were obtained from the second tab of WallaceiX user interface.

Figure 5. The subset of the population that was exported from WallaceiX. All the solutions are laid out on a grid containing the data associated with each individual. All solutions are compared morphologically and numerically side by side. The figure shows only 18 out of 81 exported individuals.

Generation 89 // Ind. 24

Displacement
Rank: 543 / 3000
Fitness Value: 0.0064337235

Seating Area
Rank: 861 / 3000
Fitness Value: 1.0258965808

Cost (Min Length Rods)
Rank: 361 / 3000
Fitness Value: 45.2988322963
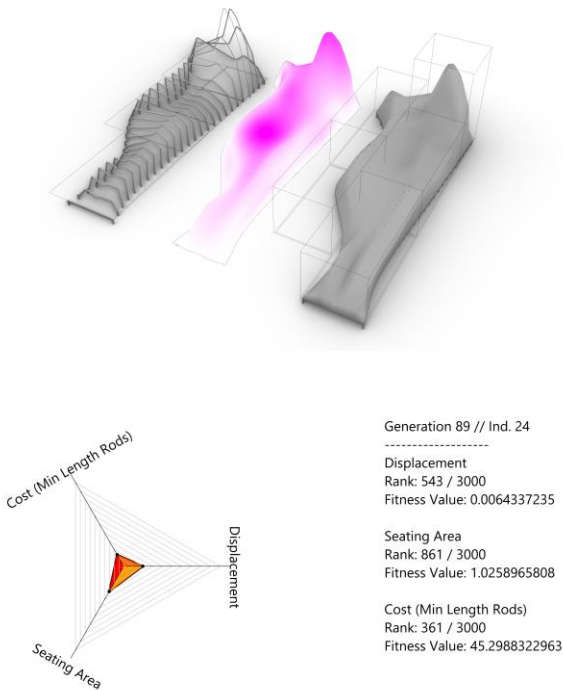
Figure 6. The selected individual with associated data pertaining to three fitness criteria.

Firstly, a set of physical experiments have been conducted in order to test the spring-back of rods with a fixed diameter, 6 mm., fixed length, 2,000 mm., and a fixed jig radius, X mm. The angles that have been tested for spring-back are multiples of 15 (i.e. 15°, 30°, 45°, 60° etc.) (See Section 3). These experiments have been carried out iteratively to accurately record the resulting bending angles against the target angles. It was observed that the deviation of each bending angle due to spring-back was approximately 12%. Hence, a linear regression analysis, which is a supervised machine learning algorithm, was opted for in order to derive a linear function between the resulting bending angles and target angles. The linear regression analysis was carried out in Lunchbox, a plugin for Grasshopper3d [22].

Subsequently, robotic control for rod bending was developed by using a general robotic control software, Robots for Grasshopper [18], and Rhino 6. The computational process was streamlined with all the constraints related to rod qualities, namely the working area, the final bending shape of the rod, and spring-back values. A fully integrated algorithm was developed, defining every step, from the rod division in sets, order of manufacturing and bending angles. Every rod was then identified by the algorithm and this information was used for the correct position and angle of each rod.

Name	Fitness Objective 1 Displacement (cm)	Fitness Objective 2 Seating Area (m²)	Fitness Objective 3 Cost, Rod Length (m)	Description	
Generation 81 Individual 3	0.0056200186 *Rank: 0/2999*	1.382345094 *Rank: 1115/2999*	45.648165637 5 *Rank: 609/2999*	Best Performance for Fitness objective 1	
Generation 94 Individual 1	0.0127958266 *Rank: 2312/2999*	0.585636242 7 *Rank: 0/2999*	48.781808684 *Rank: 2091/2999*	Best Performance for Fitness objective 2	
Generation 99 Individual 2	0.0070189687 *Rank: 930/2999*	2.971071923 7 *Rank: 1871/2999*	44.807899102 5 *Rank: 0/2999*	Best Performance for Fitness objective 3	
Generation 59 Individual 29	0.0077716913 *Rank: 1444/2999*	1.127068148 9 *Rank: 927/2999*	45.570770724 6 *Rank: 573/2999*	Solution with performance as equal as possible for all objectives	
Generation 89 Individual 24	**0.0064337235 *Rank: 543/2999***	**1.025896580 8 *Rank: 861/2999***	**45.298832296 3 *Rank: 361/2999***	**Solution with the best overall rank**	
Generation 99 Individual 5	0.0080857542 *Rank: 1571/2999*	0.729731506 4 *Rank: 283/2999*	46.986499871 6 *Rank: 1465/2999*	In the Pareto Front of the last generation	
Generation 99 Individual 10	0.0110894113 *Rank: 2087/2999*	0.645060567 9 *Rank: 72/2999*	48.776808334 5 *Rank: 2089/2999*	In the Pareto Front of the last generation	
Generation 99 Individual 15	0.0059566409 *Rank: 158/2999*	2.218382404 8 *Rank: 1585/2999*	45.069855818 *Rank: 128/2999*	In the Pareto Front of the last generation	
Generation 99 Individual 20	0.0071210202 *Rank: 1038/2999*	2.949994921 7 *Rank: 1858/2999*	44.889154592 9 *Rank: 2/2999*	In the Pareto Front of the last generation	
Generation 99 Individual 25	0.0064912887 *Rank: 576/2999*	1.036283719 7 *Rank: 864/2999*	46.096287786 7 *Rank: 984/2999*	In the Pareto Front of the last generation	

Table 2. Detailed comparison of 10 selected individuals.

6 DISCUSSION

The research presented in this paper demonstrates the advantages of the application of an adaptable computational model, driven by an evolutionary multi-objective optimization process, to generate morphological variation in response to multiple design objectives. The success of the implemented method in rationalizing a free-form design idea into a manufacturable prototype proves to be an advantageous approach for solving complex design problems. Contrary to the conventional design process of improving a single design solution, the proposed method utilizes an iterative process of developing a population of design candidates, thereby allowing for greater morphological variation within the specific design domain.

The adoption of fabrication-related and structural parameters, coupled with the geometrical freedom and constraints of a selected research agenda, in this case robotic rod bending, has the potential to address multiple performance criteria embedded in the high level of complexity of various design processes. Limitations related to the selected fabrication methods are not within the scope of this paper and will be further described in a separate publication.

Figure 7. In-situ reinforcement system built with robotic rod bending, and the final prototype.

The successful implementation of multi-objective optimization to address geometrical, structural, and material-oriented constraints simultaneously as well as independently in the optimization process can be streamlined in various stages of design, from early design exploration as demonstrated in this paper through to fabrication and construction practices. This optimization method can be further extended and utilized across a multitude of scales in order to save energy, materials, and cost in architectural projects.

The capacity to evolve a population of design candidates that vary in morphological diversity and performance is essential for design problems that cannot have a single optimal solution. This is particularly important when the design problem poses multiple conflicting fitness objectives, thereby presenting the necessity for a population of variable design solutions rather than a single individual or group of individuals. The degree of variation generated through the

177

employment of multi-objective optimization allows for greater flexibility when addressing the fitness objectives. Nevertheless, infinite variation serves little to no purpose, and informed design decisions need to be made when analyzing and selecting the final set of design candidates. Hence, coupling the geometry and data simultaneously during the analysis and evaluation stages of multi-objective optimization is a key step for attaining objectivity in this design methodology. Through the understanding of data that accompanies the geometry, the evolved solution set can become a robust and powerful alternative to a single, preference-based approach.

REFERENCES

1. De Jong, K. A. D., *Evolutionary Computation: A Unified Approach*, Reprint edition, (A Bradford Book, 2006).
2. Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T., A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II, in *International Conference on Parallel Problem Solving From Nature*, (Springer, Paris, France, 2000), pp. 849–858.
3. Erdine, E., Kallegias, A., Devadass, P., Lara Moreira, A.F., Sungur, A., 'Robot-Aided Fabrication of Interwoven Reinforced Concrete Structures', in 2017 Proceedings of the Symposium on Simulation for Architecture and Urban Design (SimAUD). Turrin, M., Peters, B., O'Brien, W., Stouffs, R., Dogan, T. (eds.) 2017 SIMULATION COUNCILS, INC., 77-85.
4. Fogel, D. B., Introduction to Evolutionary Computation, in *Modern Heuristic Optimization Techniques: Theory and Applications to Power Systems*, eds. El Sharkawi, M., and Lee, K. Y., 1 edition, (Wiley-Blackwell, Hoboken, NJ, 2008), pp. 1 – 23.
5. Fogel, L. J., Owens, A. J., and Walsh, M. J., Artificial Intelligence Through Simulated Evolution, *John Wiley & Sons* (1966).
6. Hawkins, W. J., Herrmann, M., Ibell, T. J., Kromoser, B., Michaelski, A., Orr, J. J., et al., Flexible formwork technologies – a state of the art review, *Structural Concrete* (2016).
7. Holland, J. H., Outline for a Logical Theory of Adaptive Systems, *Journal of the ACM (JACM)* **9** (1962) 297–314.
8. Karamba3D, Karamba3D – parametric engineering, (n.d.).
9. Machairas, V., Tsangrassoulis, A., and Axarli, K., Algorithms for optimization of building design: A review, *Renewable and Sustainable Energy Reviews* **31** (2014) 101–112.
10. Mahmoud, A. H. A., and Elghazi, Y., Parametric-based designs for kinetic facades to optimize daylight performance: Comparing rotation and translation kinetic motion for hexagonal facade patterns, *Solar Energy* **126** (2016) 111–127.
11. Makki, M., Showkatbakhsh, M., and Song, Y., Wallacei, *Wallacei* (2018).
12. Makki, M., Showkatbakhsh, M., and Song, Y., Wallacei Primer 1.0, (2019).
13. Makki, M., Showkatbakhsh, M., Tabony, A., and Weinstock, M., Evolutionary Algorithms for Generating Urban Morphology: Variations and Multiple Objectives, *International Journal of Architectural Computing* **0** (2018) 1–31.
14. Mayr, E., *Toward a New Philosophy of Biology: Observations of an Evolutionist*, (Harvard University Press, 1988).
15. Meibodi, M. A., Bernhard, M., Jipa, A., Dillenburger, B., Menges, A., Sheil, B., Et Al., The Smart Takes From The Strong:: 3d Printing Stay-In-Place Formwork For Concrete Slab Construction, In *Fabricate 2017*, (Ucl Press, 2017), Pp. 210–217.
16. Menges, A., Morphospaces of Robotic Fabrication, in *Rob | Arch 2012*, eds. Brell-Çokcan, S., and Braumann, J., (Springer, Vienna, 2013), pp. 28–47.
17. Rechenberg, I., Cybernetic Solution Path of an Experimental Problem, (1965).
18. Soler V. Visose/Robots.; 2019. https://github.com/visose/Robots.
19. The Living, *Bionic Partition* (2016). https://autodeskresearch.com/projects/bionic-partition-project
20. Wright, S., The Roles of Mutation, Inbreeding, Crossbreeding, and Selection in Evolution, in *6th International Congress on Genetics*, (Brooklyn botanic garden, New York, 1932), pp. 356 – 366.
21. Yi, Y. K., Building facade multi-objective optimization for daylight and aesthetical perception, *Building and Environment* **156** (2019) 178–190.
22. LunchBox, *PROVING GROUND* (2015). https://provingground.io/tools/lunchbox/

Case Study of Collaborative Urban Design based on Generative Modeling and Simulation

Xiao Wang[1], Peng Tang[2]

[1] School of Architecture, Southeast University
Nanjing, China
wxhit2010@gmail.com

[2] School of Architecture, Southeast University,
Key Laboratory of Urban and Architectural
Heritage Conservation, Ministry of Education
Nanjing, China
tangpeng@seu.edu.cn

ABSTRACT

Mature model generative tools and stable collaborative platforms bring up new design methods and work diagrams. It helps to explore the practice mode that the conventional design combined with the advanced information technology. This paper introduces two actual projects as examples to illustrate the application of collaborative design in traditional architecture and settlement conservation and urban design. Using the Esri CityEngine platform, these cases conducted by generative modeling and real-time simulation have realized the collaborative work of all participants and stakeholders. In the conservation plan of traditional architecture and settlement, a quantitative method of feature analysis and generative design was developed to regenerate the district. The referable design schemes were able to be generated by quantitative and interactive ways based on the information of quantifying and evaluating the features of historical characteristics. In the urban design case, the scheme by generative modeling and simulation was based on the rules from multidisciplinary analysis results. The site model can be simulated, and construction indicators can be calculated in real-time. By adjusting related factors and parameters, the shape and quantity can be optimized and shared interactively and immediately.

Author Keywords

Shape grammar; CityEngine; urban modeling; generative design; collaborative design.

ACM Classification Keywords

I.6 SIMULATION AND MODELING: Model Development; J.5 ARTS AND HUMANITIES: Architecture; J.6 COMPUTER-AIDED ENGINEERING: Computer-aided design

1 INTRODUCTION

Urban design at city or block scales requires the participation of various stakeholders and specialties because of its complexity. But in conventional design processes, participants are subject to inevitable restrictions such as communication delay, design platforms disunion and model drawing format conversion.

Digital technology has brought technological changes for urban design, and it is gradually improving the status mentioned above in terms of design methods, communication methods, and collaboration methods. In particular, the support of collaborative design by generative design and real-time simulation technology dramatically improves the objectivity, accuracy and efficiency of the design. For example, integrating scientific insights derived from the use of data science was shown to be advantageous for urban design and governance processes [1].

At the same time, the increasing capability of computing power for handling large amounts of data expands the research area about both urban design and traditional settlement. And complex problems can be explored and solved [2, 3]. For detail design problems, computational methods were developed for rapid urban design prototyping [4, 5]. In Grasshopper, it can generate urban layouts, including street networks, blocks, parcels and buildings based on specific requirements [6, 7, 8]. In the GIS-based platform, there are also support tools defined to formulate, generate and evaluate urban designs [9, 10].

However, there is some limited critique for collaborative urban design processes utilizing parametric tools in the literature. For example, focusing on the quantitative characteristics of a model is risky due to using parameters. Those qualitative characteristics should also be considered [11]. Thus, researchers try to develop a new way [12] to evaluate the qualitative characteristics of the designs, as a subjective process may rely on human cognition or group decision [13, 14]. Konieva developed an intuitive browser-supported interface to incorporate various types of expert inputs and ways of representing the information to take a first step towards facilitating collaborative decision-making processes [15].

As a modeling tool, CityEngine provides designers with a platform for digital generation and simulation of design problems due to its programmable nature [16]. By the

external data and its default functions as well as the CGA rule language based on shape grammar, CityEngine can realize the automatic generation and optimization of various elements of space at different scales. Moreover, the interactive interface and the interoperable data enable different parties to participate jointly in the progressive design of the scheme. The earlier research was adopting the estimated results displayed in 2D density maps in ArcGIS to 3D visualization in CityEngine [17]. Also, CityEngine was used to simulate two scenarios of possible vertical development of Surrey city in Canada from the year 2011 to 2040 [18].

CityEngine helps to embed regulation information in a web interface. Various sectors, such as planners, property owners, and stakeholders, are allowed to access the codes even from long distances [19]. Collaborative in urban design is a complex process with various parameters and data [20, 21]. CityEngine may be applied for collaborative and interactive design mode in real projects [22].

In this paper, the Eris CityEngine platform is adopted for two projects: traditional architecture and settlement conservation and renovation of Gunanjie street in Yixing and urban design of the core area of the Zidong region in Nanjing. They are taken as two cases to illustrate the application strategies of the generative modeling and real-time simulation in traditional heritage conservation and modern urban design at different scales, in different design contexts and on different design purposes.

2 METHODOLOGY

2.1 Rule-based Generative Modeling

Rule-based morphological model generation adopts a shape grammar implementation based on Computer Generated Architecture (CGA) language in CityEngine.

The shape grammar of CityEngine is a unique programming language that the user can specify to generate architectural 3D content. The model can be generated by shape grammar, which is used to describe the growth and partition relationship of the morphological elements as the mutual attribute constraints. Designs made with this shape grammars are a result of a passing of rules from a parent shape to its children, which would account for the capability of a procedural tool to produce designs rapidly at many scales [23]. The following rule derivation illustrates the process: on the left side, the start shape is shown, and on the right side, the resulting generated model is displayed (Figure 1).

Modeling steps include external data import, such as importing base maps, terrain data and adding the basic streets, city blocks and parcels 2D data. The next step is modeling by CGA rule to move 2D data into 3D models, and adding rooftops, furniture, vegetation, and other 3D assets. The last step is to texturize the buildings and facades. And several analyses can be conducted if necessary.
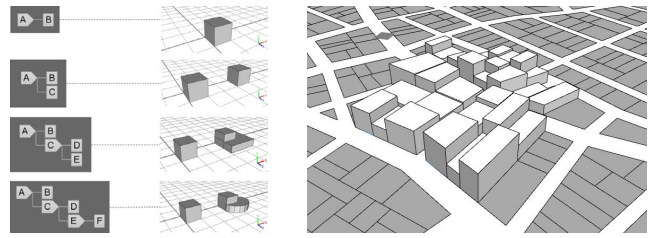
Figure 1. Generative Modeling Step by CGA and the Schematic Result of Buildings.

2.2 Collaborative Process

Rapid modeling, real-time simulation and web sharing using CityEngine platform have various vital advantages for the collaborative design process as follows:

- Rapid and interactive modeling: Because the data structure is simple, the amount of calculation is small, and the accuracy can be adjusted according to the need, so the modeling speed is fast.

- Real-time simulation: Parameter adjustment is convenient in the control panel. Morphological and related operational data are immediately visible.

- Open-access data utilizing: Support GIS and other data import and even different simulation data can be used through the same mapping method. Based on the digital chain of data/information sharing, it forms an optimized closed-loop of each link of the design and participates in the collaboration of each subject.

- Web sharing: Various sectors such as planners, property owners, and stakeholders are allowed to access the codes even from long distances.

3 CASE STUDY ON TRADITIONAL CONSERVATION

3.1 Project Introduction

It is a traditional features conservation and renewal project of Gunanjie street in China. Gunanjie Street is located in the northeast of Dingshu Town in Yixing City, Jiangsu Province. Taking the advantages of convenient water transportation, Gunanjie Street is the birthplace of purple clay teapot, which is the traditional style to brew tea from the 10th Century, Song dynasty. Gunanjie Street developed gradually with the purple clay industry. The research area in this study is the core part of Gunanjie, in which most of the production, commerce, and residence activities of residents are performed (Figure 2).

Figure 2. 3D model of Gunanjie Street scanned by DJI drone.

In the past decades, the purple clay industry was declining, and the teapot was back to the individual handicraft production. The change of teapot production and the rise of cultural tourism increased the demand for business, tourism, and living of local people. This change also magnified the conflicts between the conservation of traditional architectural features and locally spontaneous development suffered from lack of guideline and reference. Thus, the overall traditional features are gradually eroding.

Through the government's coordination, designers' guidance and residents' spontaneous renovation, the project enables all subjects to participate in the protection and construction. Thus, digital technology is no longer just a tool as an isolated solution for designers but a hub for linking all stakeholders.

3.2 Traditional Features Analysis

Traditional features analysis contains how to obtain, analyze, as well as share knowledge, information, data and cognition of different participants [24, 25].

The government provides master planning and related graphic materials, including floor plan and property information. In this stage, officers, as well as designers, come up with a principle to control the future outline of the research object. Then, building samples were selected and presented in the questionnaire for the traditionality assessment investigation. Through the survey, residents give the judgment for each building as the attributes label.

Designers first scanned 2D and 3D information of the site by drone. Meanwhile, they obtained the geometry information of the district, combining with the building photos to collage the street façade. On the other hand, a survey was conducted. The typical building was decomposed into a variety of elements and combination modes as shape grammar, coding for analysis with survey result. Then using a knowledge discovery tool, designers can extract the rules of traditional features.

After that, all the data in different formats need to be integrated into lists or databases. So, they can be read and used easily by the computer automatically for the next stage.

3.3 Generative modeling and simulations

In this stage, elevation information and aerial photos of the site are used to create terrain. After importing the necessary road and plot data, a preliminary spatial structure is obtained. The plots are divided into building footprints according to the texture characteristics obtained from the survey and analysis (Figure 3).

Figure 3. Pre-processing of Generating in CityEngine.
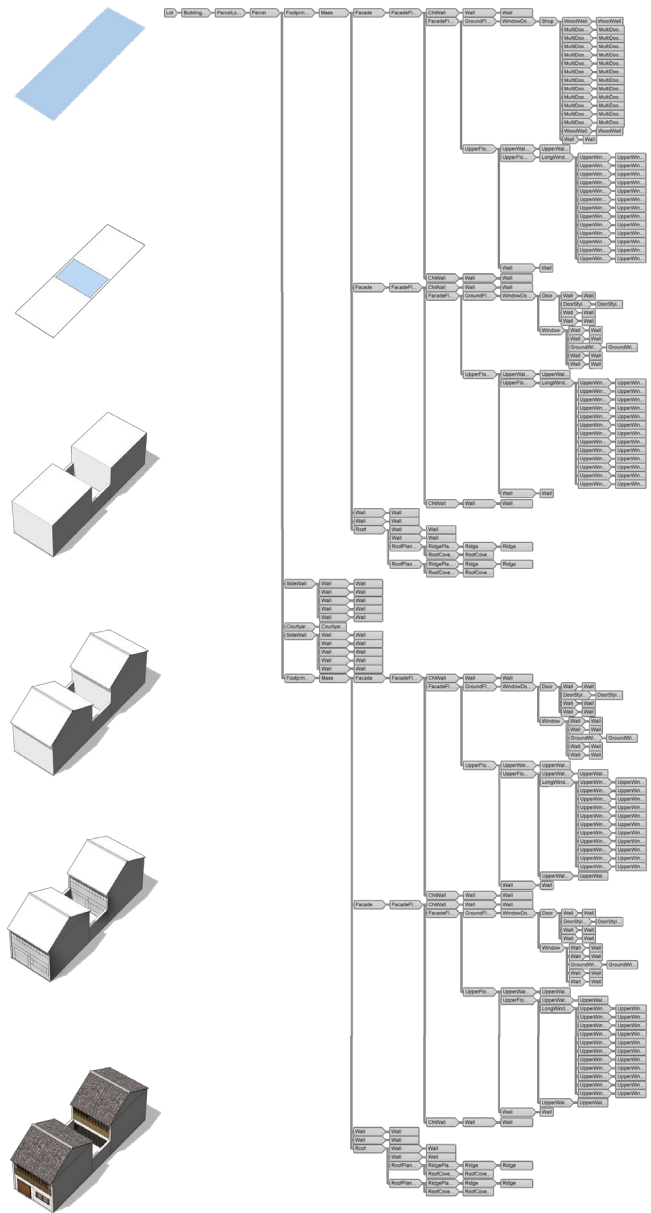
Figure 4. Generate Step of a Typical Dwelling Building by CGA and Its Morphology Hierarchy Structure.

Through the CGA language to operate form according to morphological rules. Based on the previous analysis and research, the reconstruction of the single building or dwelling focuses on the volume's relationship and the group layout. The tree diagram is used to describe the growth and partition relationship of the morphological elements as the mutual attribute constraints. Moreover, the geometric constraints of the elements themselves are obtained by statistics. This kind of rule also conforms to the logic of CGA grammar. For a single building, the entire volume was divided into subitems at different levels according to its morphology and construction characteristics. For a set of buildings for a big family, several kinds of dwellings and shops were generalized

181

according to the dimension and combination mode (Figure 4). Thus, the whole district with traditional features can be simulated (Figure 5) as a 3D scenario in CityEngine.

3.4 Collaborative and Interactive Conservation

In this process, information and data from all parties are aggregated into the platform to be presented and utilized. This digital chain links the complete analysis, design and implementation. The output results and presentation methods can be various and adjusted according to the users and using purposes.
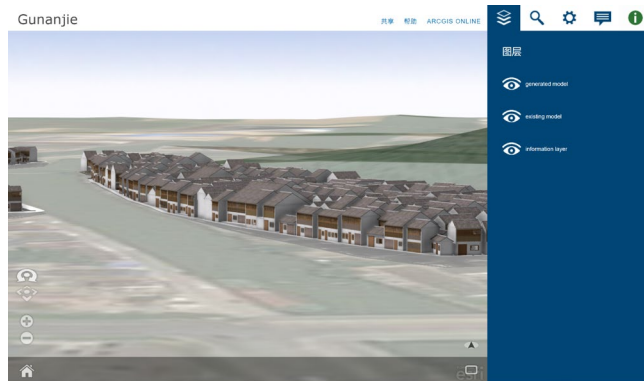
(a) Overall Scene of Project.
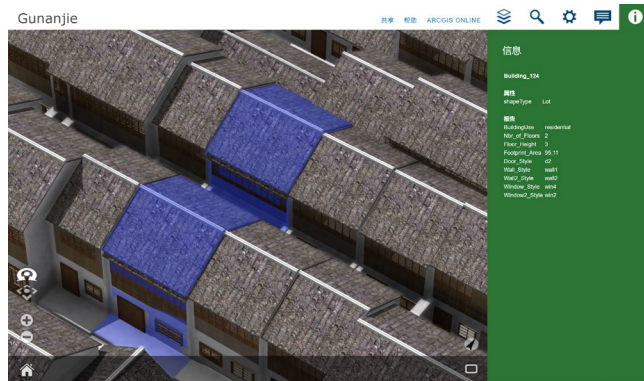
(b) Selected Building Information.

Figure 5. Web Scene of Project.

For making guidelines, the generation result can be printed to establish generally accepted conservation plans for officers and house owners. It is can also be the reference to detailed design in the next stage. With web sharing, all stakeholders or people who are interested in (residents, practitioners, tourists) can access the project by remote browsing and learning (Figure 5).

Taking advantage of this function, the government can get multiple design schemes in an overall view. For results comparison, it is convenient to get the corresponding index, such as economic indicators, for further policy-making.

Notably, residents can decide the renew plan by menu-based selecting or matching according to their own house situation. The owners of private buildings can choose the provided options to make their personal designs. It is

available in the program to make various types of combinations with different building elements according to the location and dimension of the plot. With the statistics of types and their proportion in this district, the optimal scheme could be generated conforming to the surrounding texture as well as individual demand (Figure 6).
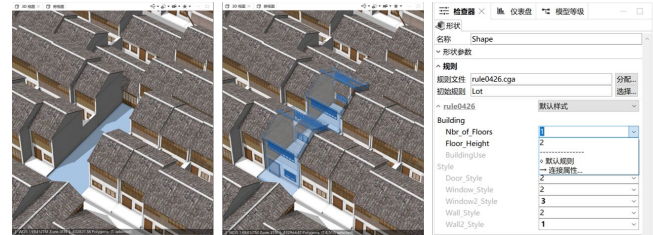
(a) View Window and Parameters Control Panel.
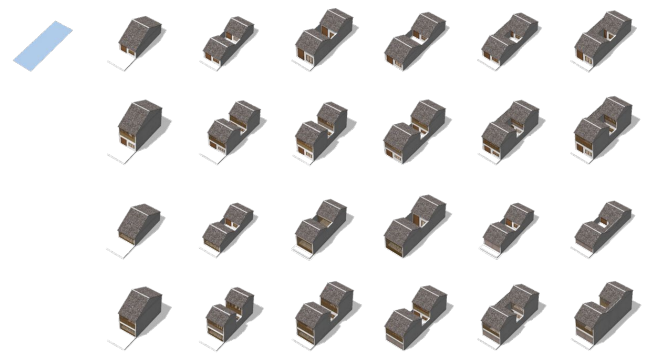
(b) Possible Buildings on One Lot.

Figure 6. Interactive Adjustment and The Various Results.

4 CASE STUDY ON URBAN DESIGN

4.1 Site Introduction

It is an urban design project of the Zidong core area in Nanjing. This city is an important central city in the eastern coastal areas in China. Among the city clusters in the Yangtze River Delta, Nanjing ranks second only to Shanghai in the city scale and is currently the only megacity.

Zidong Area refers to the east of Nanjing, with a total area of about 758 square kilometers. The core area of the Zidong Area is in the middle part (Figure 7). It will become a regional hub to integrate surrounding ecological, industrial, landscape and other resources. It will help improving infrastructure construction and the quality of the space environment, focusing on the features of the urban core area in the hilly areas of the south of the Yangtze River and radiating the surrounding development. All of these will shape an open and inclusive Nanjing-Zhenjiang-Yangzhou integration paradigm.
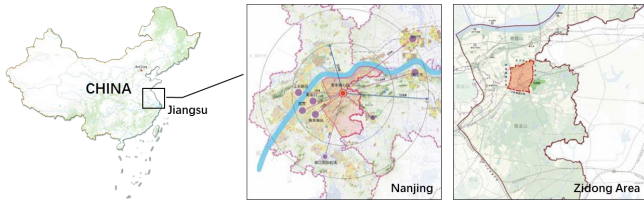
Figure 7. Site Location.

4.2 Synthesized Design Rules from Multiple Disciplines

The aim of the interaction and cooperation of multiple disciplines is to make an effort to create ecological-priority, innovation-driven and future-leading urban morphology and environment in the core area of Nanjing Zidong. Integrated all part work of each discipline, the analysis results, as well as the primary construction indicators provide by the client, become the requirements or rules of the generative design stage. The mainly parametric basis as follows (Figure 8):

- Ecology: Based on the analysis of higher-level planning conditions, physical environment, and site elements, sort out site construction conditions for proper construction analysis. On this basis, a suitable construction scope can be obtained according to the degree of influence of these factors on the construction.

- Performance: According to the cold island rating and wind environment analysis, group pattern constraints are proposed.

- Transportation: The division pattern of blocks by significant transportation networks, influencing factors such as the stimulation of construction by key transportation stations (TOD).
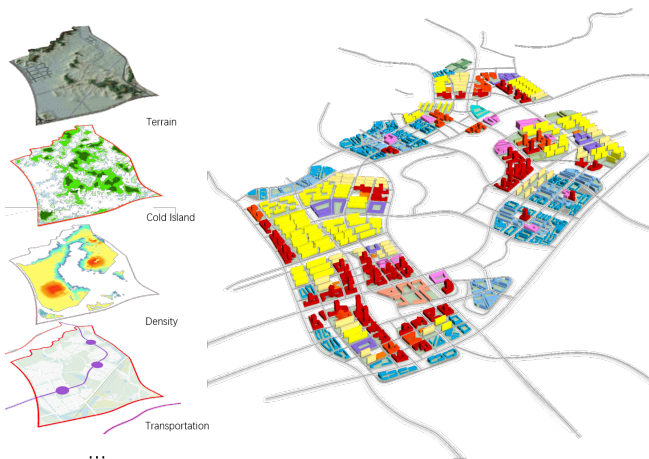
Figure 8. Rule Mapping Layer and Initial Generative Result.

4.3 Collaborative Generative Modelling by CGA Grammar

Apart from the necessary elevation data to make the terrain layer, the initial road network and the constructible area controlled by the ecological factors outline the rough boundary of the construction area of each type of land. After adjusting the road and blocks referring to the

conventional artificial result, a preliminary spatial structure can be shaped. Then the blocks are divided into plots according to the experience parameters about modern city. In the CGA coding process, designers try to solve the plot division, morphological models, attributes control and calculation results of critical indicators output.

In terms of the block division, it is conducted according to the land uses. Typical organization modes of the neighborhood were summarized for extracting the corresponding division methods. They are the commercial-type using skeleton subdivision, the research-type and technology-type using recursive subdivision, he apartment using bisected subdivision, the compound community using offset subdivision, the service-type and support-type using organic subdivision (Figure 9).
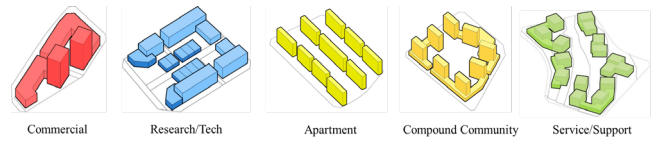
Figure 9. Organization Modes According to The Land Uses.

As shown in Figure 9 above, the block was subdivided into several plots. Then they were subdivided into the building footprints according to the building types corresponding with the land uses, such as the towers with podiums for commercial buildings and the slab building for apartments. From the footprints, building models can be generated with geometric parameters and indicators (Figure 10).
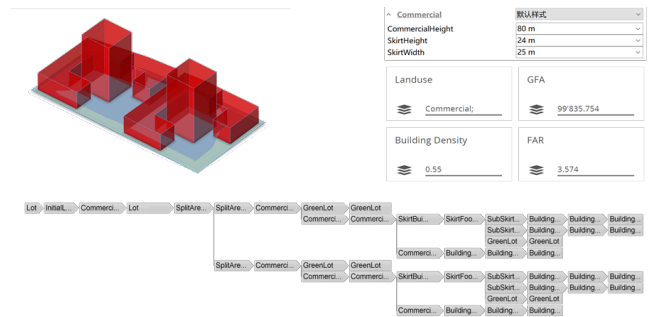
(a) Commercial-Type.
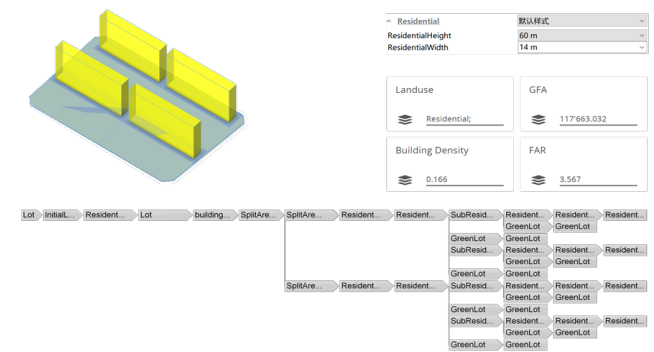
(b) Apartment-Type.

Figure 10. Building Main Parameters and The CGA Hierarchy.

4.4 Interactive Adjustment by Real-time Morphological Simulation

In the early design stage, designers would like to know not only the spatial texture but also the relevant development indicators. It is necessary for the government or for managers to control the direction of the design in the real project. Because the significant indicators such as development capacity and FAR are calculated and presented in real-time in this platform, designers can master the macro situation of the scheme. After selecting the single object, they can also find detail information on a micro-scale. By adjusting related factors or initial shapes, the whole form and quantity can be optimized interactively and immediately if the design conditions changed. Besides, the individual building can be adjusted by changing the parameters asymptotically. It can also support the comparison among multiple random schemes (Figure 11). Eventually, all the 3D models and calculated results can be exported to local for utilizing in other specialized software, platform or operation processes.

In the process, information and data of all parties are aggregated into the platform to be presented. The rapid generation and simulation make the process that usually required multiple parties to play for a long time on the platform quickly completed. The computer simulation results in practice provide scientific theory and computing power for support and reference.

(a) Adjustment to Meet Target Indicators in Real-Time (Scheme 1).
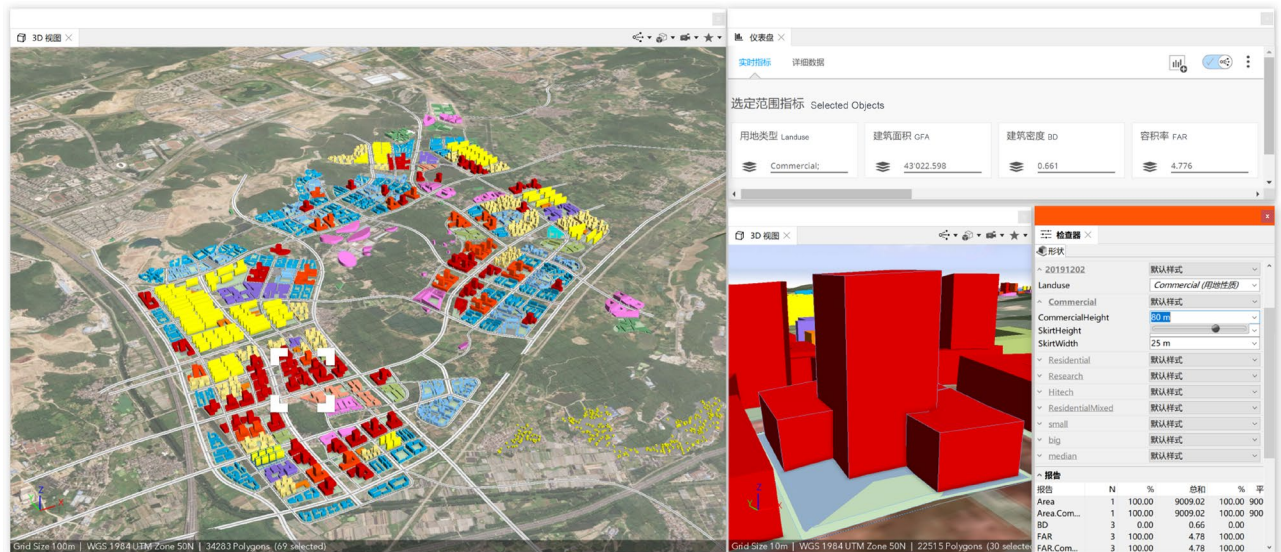
(b) Adjusting Shape Parameters in Detail (Scheme 2).

Figure 11. Interactive Interface of CityEngine.

184

5 DISCUSSION

In the conventional design process, the text and drawings are all the carriers of expression and presentation of scheme information. In the process of computer-aided design, numbers and data are another way of recording and sharing scheme information. The transmission and interoperability of such information is the key to collaborative design. Efficient, accurate and visualizable are the vital advantages. The massive amount of calculations and various formats are also challenging to use. It also has higher requirements for the quantity and quality of data, the scientificity and accuracy of data processing methods, and the operability of operations.

Based on the digital generation or data-driven design and real-time simulation, fast and timely feedback improves the efficiency of collaborative design. Especially in large-scale urban design, the macro and complexity of the problem make it unnecessary to cover all the data in the design process. Taking the CityEngine as an example, the existing framework process can deal with the problems better that need to be solved in the early stage of design. After the professional analysis of each discipline, the data and requirements will export as images. They are imported into the platform as a mapping layer. The data processing and calculation are relatively simple, and the visualization is more readable compared with the conventional design methods.

6 CONCLUSION

This study illustrated two cases to explain an objective and effective generative modeling and simulation workflow for collaborative urban design in the CityEngine platform. By integrating the massive design conditions as digital generation rules, the generative modeling and simulation are expected to increase the efficiency and objectivity of the decision-making and design process. Furthermore, the sharing and collaborative mechanisms are also helpful in improving the effect of multidisciplinary communication. Compared with the conventional design workflow, the interaction and cooperation of multiple disciplines are also reflected in that adjusting the data of each discipline at the same time, and the results can change in real-time to help to optimize the solution.

REFERENCES

1. Herthogs, P., Clavier, F., Knecht, K., Mueller, J., Drillet, Z., Miao, Y., ... & Schlapfer, M. Big Data Informed Urban Design and Governance. *Future Cities Laboratory: Indicia 02*. Lars Müller Publishers (2019), 99-110.

2. Balaban, Ö., & Tunçer, B. Visualizing and Analising Urban Leisure Runs by Using Sports Tracking Data. *Proc. eCAADe 2017*, 533-540.

3. Biao, L., Zifeng, G., & Yunzhu, J.. Modeling and Realizing Generative Design A Case Study of the Assignment of Ji Village. *Architectural Journal 5*, (2015), 94-98.

4. Miao, Y., Koenig, R., Knecht, K., Konieva, K., Buš, P., & Chang, M. C. Computational urban design prototyping: Interactive planning synthesis methods—a case study in Cape Town. *International Journal of Architectural Computing 16*, 3 (2018), 212-226.

5. Koenig, R., Miao, Y., Knecht, K., Buš, P., & Mei-Chih, C. Interactive urban synthesis: Computational Methods for Fast Prototyping of Urban Design Proposals. *Proc. CAADFutures 2017*, Springer (2017), 23-41.

6. Koenig, R., Standfest, M., & Schmitt, G. Evolutionary multi-criteria optimization for building layout planning-exemplary application based on the PSSA framework. *Proc. eCAADe 2014*, 567-574.

7. Abdulmawla, A., Schneider, S., Bielik, M., & Koenig, R. Integrated Data Analysis for Parametric Design Environment-mineR: a Grasshopper plugin based on R. *Proc. eCAADe 2018*, 319-326.

8. Koenig, R., Miao, Y., Aichinger, A., Knecht, K., & Konieva, K. Integrating urban analysis, generative design, and evolutionary optimization for solving urban design problems. *Environment and Planning B: Urban Analytics and City Science*, (2019), 1-17.

9. Beirão, J., Gelly Mendes, J. D., & Stouffs, R. Implementing a generative urban design model: Grammar-based design patterns for urban design. *Proc. eCAADe 2010*, 265-274.

10. Arnold, J. D. M., & Lafreniere, D. Creating a longitudinal, data-driven 3D model of change over time in a postindustrial landscape using GIS and CityEngine. *Journal of Cultural Heritage Management and Sustainable Development 8*,4 (2018), 434-447.

11. Steinø, N., & Veirum, N. Parametric Urban Design. *Congress Aesop 5*, (2005) .

12. Steinø, N., Benbih, K., & Obeling, E. Using parametrics to facilitate collaborative urban design: An attempt to overcome some inherent dilemmas. *Planum. The Journal of Urbanism 1*, (2013), 1-13.

13. Munemoto, S. Study on evaluation of impressions based on inference of spatial elements using rough set theory-A case of special exhibition of the national museum of ethnology. *Journal of Architecture and Planning (Transactions of AIJ)*, 608 (2006), 235-242.

14. Tang, P., Wang, X., & Shi, X. Generative design method of the facade of traditional architecture and settlement based on knowledge discovery and digital generation: a case study of Gunanjie Street in China.

International Journal of Architectural Heritage 13, 5 (2019), 679-690.

15. Konieva, K., Joos, M. R., Herthogs, P., & Tunçer, B. Facilitating Communication in a Design Process using a Web Interface for Real-time Interaction with Grasshopper Scripts. *Proc. eCAADe and 23rd SIGraDi 2019*, 731-738.

16. Esri CityEngine: Advanced 3D city design software. https://www.esri.com/en-us/arcgis/products/esri-cityengine/overview . As of 23 March 2020.

17. Xu, Z., & Coors, V. Combining system dynamics model, GIS and 3D visualization in sustainability assessment of urban residential development. *Building and Environment 47*, (2012), 272-287.

18. Koziatek, O., & Dragićević, S. iCity 3D: A geosimualtion method and tool for three-dimensional modeling of vertical urban development. *Landscape and Urban Planning 167*, (2017), 356-367.

19. CityEngine Resources. https://doc.arcgis.com/en/cityengine . As of 23 March 2020.

20. Sakai, Y., & Tsunoda, D. Decentralized Version Control and Mass Collective Collaboration in design-A Case Study of a Web Application Utilizing the Diff

Algorithm and Automated Design Generation. *Proc. eCAADe 2015*, 207-214.

21. Salgueiro, I. B., & Ferries, B. An "environmental BIM" approach for the architectural schematic design stage. *International Journal of Architectural Computing 13*, 3-4 (2015), 299-312.

22. Schnabel, M. A., Zhang, Y., & Aydin, S. Using parametric modelling in form-based code design for high-dense cities. *Procedia engineering 180*, (2017), 1379-1387.

23. Litovitch, R. Investigating CityEngine as an urban geodesign change model for transit-oriented development planning and design along Winnipeg's future Eastern Rapid Transit Corridor. (2018).

24. Wang, X., Tang, P., & Shi, X. Building Facade Description System and Historical Feature Analysis Method by Rough Set Theory and Algorithmic Language: A Case Study on The Traditional Architecture and Settlements in Gunanjie Street, Yixing, China. *Proc. iaSU 2016*, 179-186.

25. Wang, X., Tang, P., & Shi, X. Analysis and Conservation Methods of Traditional Architecture and Settlement Based on Knowledge Discovery and Digital Generation - A Case Study of Gunanjie Street in China. *Proc. CAADRIA2019*, 757-766.

The Use of CA to Generate Informal Architectural Systems

Song Jie Lim, Varvara Vasilatou and Shih Hsin Wuu

Architectural Association School of Architecture
London, United Kingdom
{songjielim88, barbara.vasilatou, annwuu714}@gmail.com

ABSTRACT

The aim of this study is to demonstrate the incorporation of specific architectural intentions in Cellular Automata (CA), for the generation of mass-customised building blocks in Jakarta's *kampungs* (informal urban settlements).

The primary contribution of this research lies in the horizontal and vertical programmatic relationships within the CA simulation. Through field study in *kampungs*, context-specific incremental aggregation was analysed and codified into bottom-up rule sets that align with local practices of self-built expansions. This is combined with top-down strategies that improve generated morphologies' overall environmental performances.

The generative tool's capacity to be tailor-made for different urban zones is introduced to highlight potential application and value to the field.

Author Keywords

Simulation-based Generative Design; Simulation-based Design Tools and Methods; Cellular Automata; Informal Urban Settlement.

1   INTRODUCTION

The research ambition lies in exploring the potential of CA to include programmatic rule-sets, with the purpose of generating alternative settlement design.

Cellular Automata (CA) is common generative design tool, based on a grid layout. It consists of cells that represent predefined states relating to a set of rules. The aggregation process for subsequent cell-units depend on the rules and the state of the neighbouring cells. Its ease in use derives from its structural set up and the simplicity of the applied aggregation rules. The system is characterized by hierarchy, heterogeneity, partial autonomy, self-organisation and adaptability [5].

*Kampung*s, on the other hand, are the informal urban settlements of Jakarta. [13] "Informal", here, refer to its self-built nature, describing an open-ended design process that is not authoritarian and prescriptive, accommodating individual needs, family models, as well as allowing for adaptation and changes through time. Even though these attributes accommodate demographic and environmental changes, the lack of regulation produce poor quality of life, characterised by poverty, unsafe and unsanitary conditions.

The local government of Jakarta currently adopts a 'Western' approach to social housing provision in the form of *Rusunawa*. The one-size-fits-all solutions are overly simplistic and detached from context, omitting local practices of incremental expansions. This results in many interventions failing as overcrowded vertical slums.

The sequential aggregation of *kampung* houses and adaptive change-in-use are essentially based on spatial relationships between programmatic functions. As the hierarchical, rule-based organisation correlates to the characteristics of CA, the research paper aims to generate alternative settlement design, using an adopted computational model with encoded practices of *kampungs*.

Since Conway's Game of Life, current precedents of CA implementation in the design process are limited to form-finding via simple rule-sets. Few precedents have explored the capability of CA for programmatic distribution in architecture.

Whereby 'Algogram' integrated CA in the preliminary design stages for automated concept diagrams [7], Herr's research suggests the use of CA for higher levels of design complexity, in addition to generative form-finding [6]. Dincer and his team's work further demonstrates the feasibility of CA as computational tool for programmatic distribution in mass housing, integrating user feedbacks [3].

Through a combination of top-down and bottom-up definitions, a revised CA model (that generates mass-customised building blocks) attempts to tackle complex programmatic distribution, retain the diversity and variation of *kampung* houses whilst improving overall environmental performance.

2   DESIGN AMBITION

The CA research aims to address two critical issues concerning *kampungs*, namely:

A.   the lack of urban regulations which leads to poor living conditions, and

B.   the inadequacy of standardised housing provisions that fails to account for ingrained practices of future expansion (*figure 1*).

Addressing A, top-down strategies are introduced to:

1. **Transform the horizontal urban sprawl into taller housing blocks with strict building envelope.**
By accommodating the same neighbourhood population size with smaller footprints, enable wider streets to improve air flow rate for better thermal comfort.

2. **Elevated building mass.**
Informed by the climatic design of Malay houses, the aggregation sequence is initiated from the first-storey, with the building elevated on pilotis. Besides mitigating future risk of inundation, an open-ground condition maximise air-flow rate which in turn reduces humidity levels.

Responding to B, bottom-up strategies are implemented to:

3. **Retain social-cultural programmatic units, spatial relationships and organisation.**
Based on field research, incremental growth patterns of the *Kampung* and indigenous vernacular, *Malay* houses *(figure 1)* are analysed, alongside regional distribution of family sizes and programmatic requirements (*Table 1*).

4. **Increase street frontage for informal economy.**
Larger vertical circulation cores take into consideration the importance of multi-functional domestic frontage, pocket spaces that can be adapted to suit different *warungs* (small, family-owned informal businesses). Extended provision enables larger percentage of inhabitants to engage in trading which would otherwise be limited to ground tenants.

5. **Create spatial porosity within building blocks.**
Undefined voids allow for the future expansion of smaller households (nuclear and basic family types).

6. **Allow mass-customisation**
Rather than a singular solution, the proposal generates varied design outputs according to initial placement of circulation spaces, specification of family type, number (10) and predicted program requirements. Combined with urban configuration analysis, this permits the selection of optimal block.
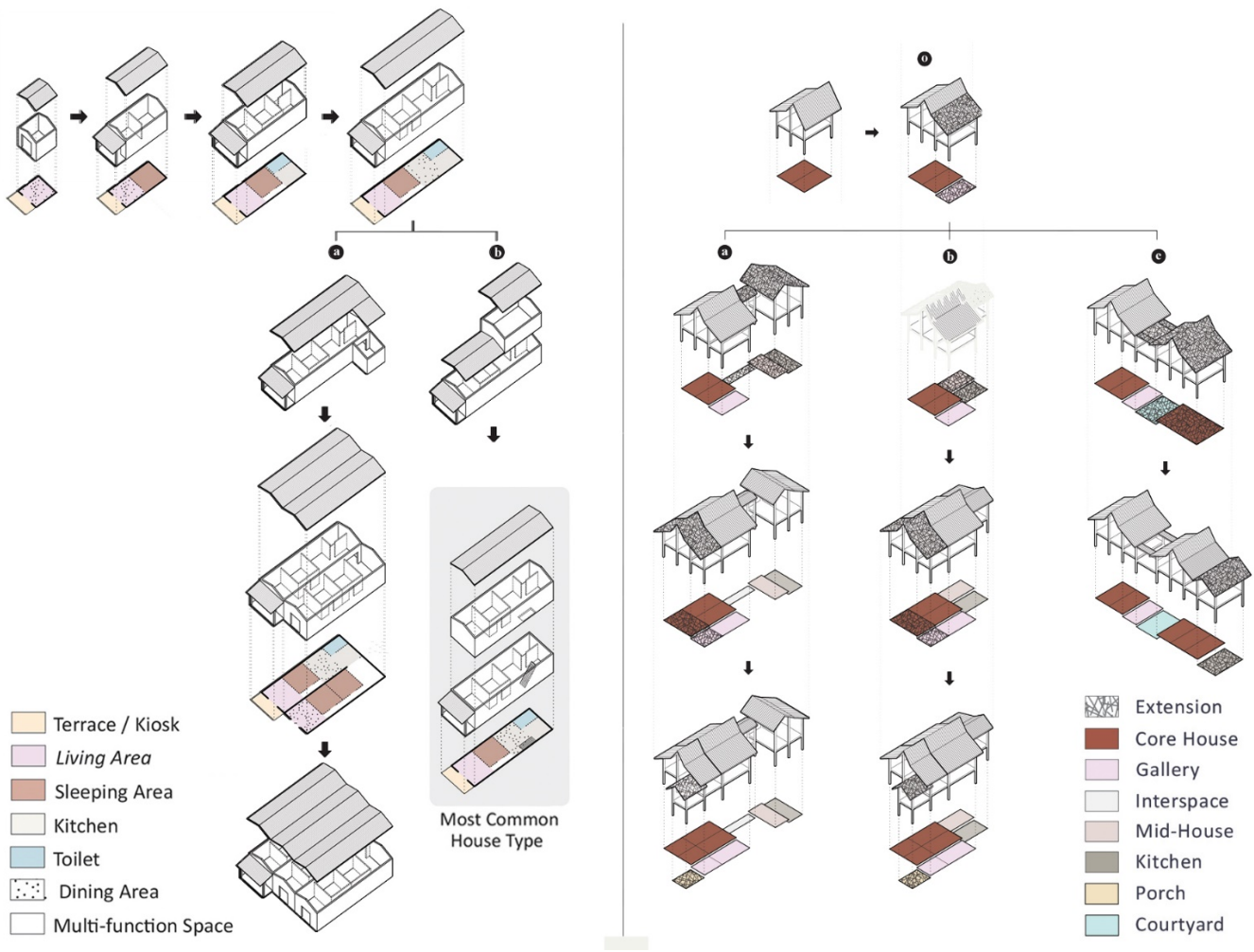
Figure 1. Incremental growth of Kampung house (left) and traditional Malay house (right). [9]

3. METHODOLOGY

3.1 Revised CA Model

The proposed computational model leverages the CA's strengths in defining spatial relationships between neighbouring cells as well as its ability to integrate design elements, such as rulesets, to generate customised outputs.

Unlike most precedent simulations that solely relied on local cell interactions and decentralisation of elements, the adopted computational approach introduces a top-down force as the centralized control, to perform the decision-making process. [1] The developed system consists of two interrelated processes, top-down and bottom-up. The former dictates the overall spatial organisation, whilst the latter provides variation and irregularity *(figure 2)*.
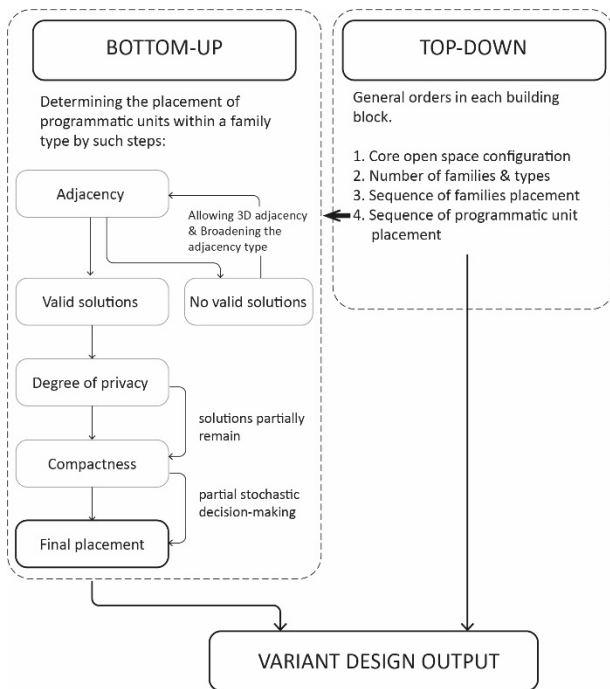
Figure 2. Methodology graph

This programmatic aggregation process exhibits a higher level of complexity that is unprecedented, tailor made to generate context-specific housing solutions.

At the local scale, the algorithm incorporates programmatic units with varied configurations and specific neighbouring relationships. At the building block scale, it generates three-dimensionally arranged clusters, as house units, to accommodate different family types and retain a reasonable privacy gradient. This formulates a consistent spatial order within the building block, whilst providing a great variety in the final output. The heterogeneity in the outputs aligns to the needs of self-built practices and was unlikely to be achieved by the strict use of former CA methodologies. The following sections elaborate on both approaches to provide a comprehensive overview of the computational process.

3.2 Top – Down Process: Units definition & Sequence rules

The prevailing kampung house type and its corresponding growth pattern were identified through field research. In relation to the kampung house type, three common family types and their spatial requirements were also analysed, to form the basis for the computational simulation. *(Table 1)*

Family Type	Family Size	%	Spatial Needs	Unit	Unit Privacy
A **Basic**	2-3	20%	Living	1	Semi-Open
			Kitchen	1	Open
			Dining	1	Semi-Open
			WC	1	Closed
B **Nuclear**	4-10	50%	Warung	1	Open
			Living	1	Semi-Open
			Kitchen	1	Open
			Dining	1	Semi-Open
			Bedroom	2	Closed
			WC	1	Closed
C **Extended**	11-20	30%	Warung	1-2	Open
			Living	2	Semi-Open
			Kitchen	1	Open
			Dining	2	Semi-Open
			Bedroom	4	Closed
			WC	2	Closed

Table 1. Family Types, Distribution and Program Mix

Sequence	Unit Program	Unit Dimension	Privacy Domain	Function Adjacency
Start **(0)**	Circulation Core		0 (Open)	-
A **(1)**	Warung (kiosk)		0 (Open)	Start
B **(2)**	Living Area		1 (Semi - Open)	Start + A
C **(3)**	Kitchen Area		0 (Open)	A + B
D **(4)**	Dining Area		1 (Semi - Closed)	B + C
E **(5)**	Sleeping Area		2 (Closed)	B + D
F **(6)**	Water Closet		2 (Closed)	B + C + D + E

Table 2. Programmatic Setup

The aggregation sequence is based on the degree of privacy defined for each programmatic unit and follows the sequence from the least to the most private ones. *(Table 2)* As such, the *open* units are always located around the cores (initial open spaces); whereas the ones indicated as *closed* are mostly placed at the periphery of the building plot.
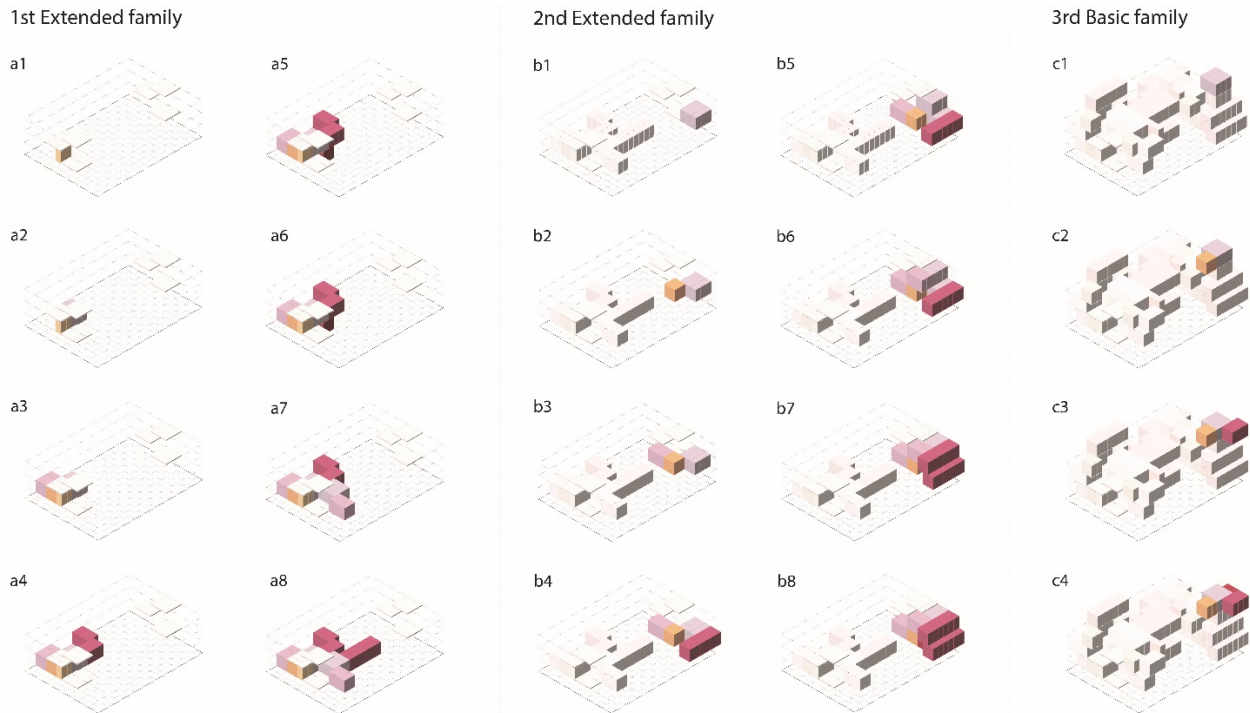
Figure 3. Sequential placement of programmatic units. The first formed household (a1-a8) performs a horizontal distribution; The second one (b1-b8) starts forming three-dimensional arrangement from step b5, due to the rule that the second living room should be adjacent to the first living room in the extended family and no unoccupied cells around it are detected.

In kampung communities, the social and its related spatial structures are comprised of two hierarchies - Rukun Warga (RW), a larger urban patch, and Rukun Tetangga (RT), the community of approximately ten house units. Respectively, the proposed building block consists of ten families. Different family types are placed sequentially and based on the likelihood of future expansion. Type-C families are the most complete ones; thus, they are positioned first. Whereas, Type-A families are placed lastly, on top of the concurrent building incrementation. *(Figure 3)*

The aggregation process for the 10-family building block is as follows:

1. According to site constraints and associated environment risks, an initial algorithm returns with a range of possible solutions for the plot size and height limit.

2. The placement of the first two cores is initiated, serving as both social spaces and vertical circulation to access each household. Two starting cells are defined to generate a series of cores, arranged three-dimensionally.

3. The aggregation of the Type-C family house is implemented by the sequential placement of its programmatic units. *(Table 2)*

4. The output is assessed based on the local rules of aggregation, as stated in the following section (3.3).

5. Steps 3-4 are repeated for the next Type-C family, till the amount of Type-C families reach the predefined total number (3 families).

6. Following the strategy stated in steps 3-5, the algorithm generates house units for Type-B (5 in total) and Type-A (2 in total) families.

7. Once the amount of Type-A families reaches the target, the simulation stops, providing one building block solution.

The creation of three-dimensional configurations relies upon the locations of the cores. Thus, the identification process of their valid locations, namely ones surrounded by unoccupied cells and on different elevations, is vital. Prior to the household placement, the algorithm detects if at least nine cells (size of living area) are available around the previously aggregated core. If not, the system proceeds to the next core location, on either the same or upper level, to start the aggregation (see *figure 3, a8-b1*, where the second extended family is placed around the second core).

3.3 **Bottom – Up Process: Local Aggregation Rules**
The aggregation strategy is formed by local rules that define the process of placing each programmatic unit sequentially, following assigned criteria with a degree of flexibility. In order to enable cells to recognize their relationships with each other, the detection area *(figure 4)*, namely the cells that share edges with the lastly placed unit, requires clear definition.

190
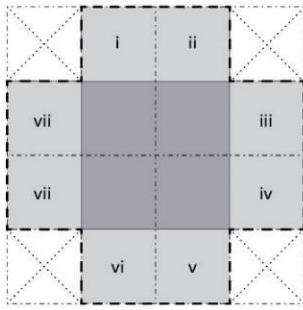
Figure 4. Detection area, shown in light grey. The cells in dark grey indicate the last, placed unit.

For the placement of each programmatic unit, the final position is selected via four steps:

1. **Programmatic adjacency** *(figure 5)*. The process initiates with the recognition of unoccupied cells adjacent to preceding programmatic unit. If all adjacent cells are occupied on the same level, the algorithm seeks solution on upper level *(figure 3, b4-b5)*.

2. **Valid solutions** *(figure 6)*.
   For each starting point, the algorithm generates possible unit placements, considering its defined dimensions *(Table 2)*. The position is stored for further assessment only if no overlapping is detected with the existing configuration. If no valid solution is formed, the algorithm broadens the permitted types of neighbouring units.

3. **Degree of privacy for each unit type** *(figure 7)*.
   Based on observation of existing kampung houses *(Table 2)* and contemporary living standards, each cell of the valid solutions is scored in relation to its proximity to the core open space / vertical circulation space *(figure 7)*. The scores of the cells are summed-up for each valid solution. The valid solutions are then split into three categories, open, semi-open and private, related to their final score. The algorithm retains only the valid solutions that fall under the same category as the unit-to-be-placed.

4. **Compactness** *(figure 8)*.
   Solutions with more unoccupied adjacent cells are considered less efficient and ranked worse, to reduce building block's fragmentation. This evaluation criteria improves the house-unit's space usage efficiency. However, as creating a degree of porosity within building block is of primary concern, the most compact solution is not always selected. Instead, the first 20% solutions with better compactness are considered valid and the final programmatic unit placement is finally selected randomly.

Define the possible starting cells (i,ii,iii,iv) for the positioning of the 'Living Area' unit, according to its function's adjacency 'Warung' (kiosk).

Figure 5. Adjacency

Unoccupied Cells for Possible Solutions. List the possible solutions (1,2,3,4,5,6,7,8) for the starting cells (i,ii,iii,iv).
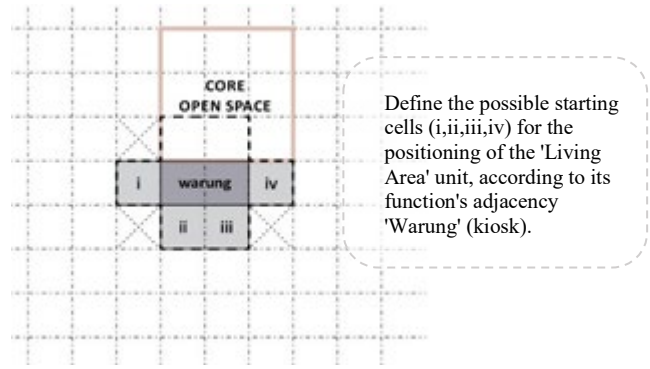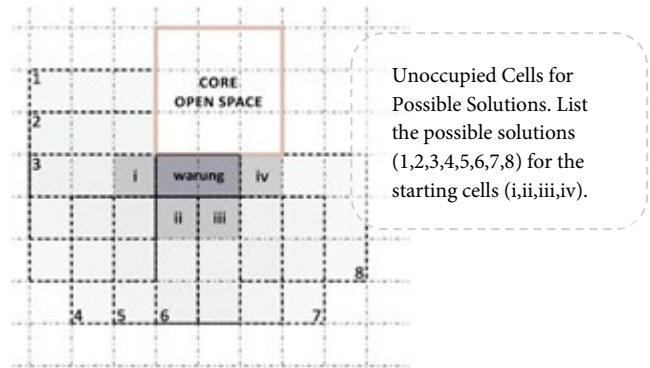
Figure 6. Valid solutions

Ranking possible solutions for 'Living Area' unit (1,2,3,4,5,6,7,8) based on degree of privacy (0,1,2). Suitable solutions (4,5,6,7) are kept for the next assessment.
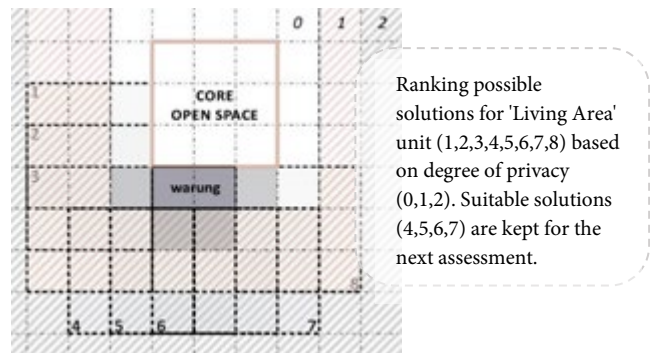
Figure 7. Degree of privacy

Ranking of the possible positions (4,5,6,7) according to the number of adjacent occupied cells - connected edges. Ultimately, the final solution is selected randomly among solutions 5 and 6.
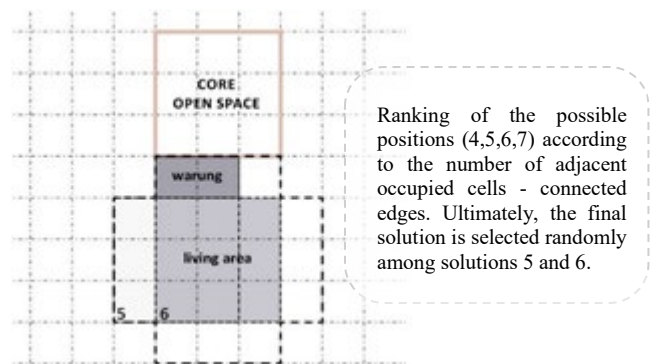
Figure 8. Compactness

191
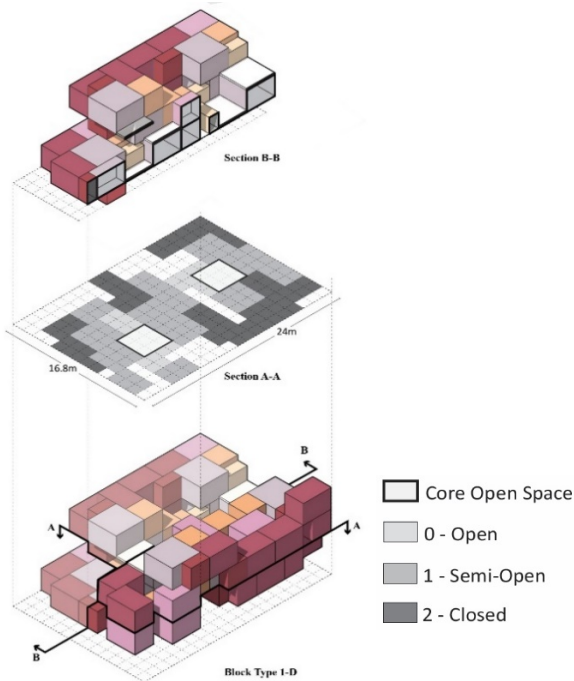
Figure 9. Gradient of privacy in the output building block

4 EVALUATION OF RESULTS

4.1 Building Types and Spatial Attributes

The generated individuals present clear spatial arrangements that formulate a gradient of privacy *(figure 9)*. This reflects on the aggregation strategy of sequentially placing house units according to their related degree of privacy.

To demonstrate the computational outputs, 16 different building blocks are presented *(figure 10)*. Patterns appear clearly between the initial placement of the cores and the emergence of unbuilt voids. This leads to the categorization of the outputs into four distinct building types. When the cores are populated at the periphery, the formation of emergent voids is detected on the opposite sides of the block, in each case (A, B, C). However, when the initial cores are placed within the inner area of the block, the emergent voids are scattered and random, creating porous inward-looking building morphologies (D).

With regards to the evaluation of the building blocks, initially a density comparison is conducted. The outputs are examined in comparison to each other and to existing site conditions, as documented by site visit. The results of Kampung Melayu are presented indicatively, being the oldest and most characteristic kampung of Jakarta. (*figure 11, table 3*). As expected, the public space generated as emergent void or core open space comprise 30-33% of the total plot and is doubled from the existing conditions of mere 16%, while the m² of built area/person are kept on same levels. Type D is the only building type that presents the highest coverage of the plot with 80%, which indicates that emergent void is fragmented within the building block.
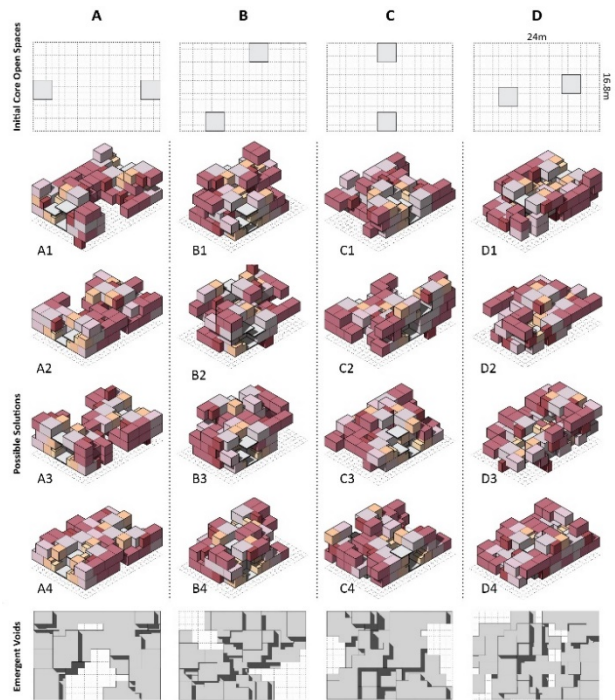
Figure 10. Building block types. Relationship between cores and emergent open spaces
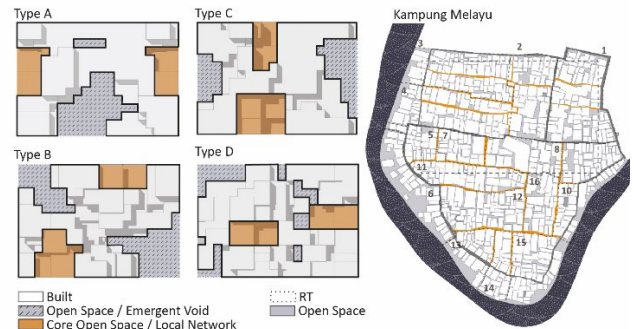
Figure 11. Building block types, Density Comparison

Building Type	Built Plot Area		Core Open Space Area		Void Area	
	%	m²/ person	%	m²/ person	%	m²/ person
Type A	70	3.1	13.6	0.6	16.4	0.73
Type B	67.5	2.8	20	0.71	12.5	0.97
Type C	67.2	3	16.8	0.76	16	0.72
Type D	81	3.4	13.9	0.58	15.1	0.5
Kampung Melayu RT	84	4.15	11	0.55	5	0.12

Table 3. Building block types, Density data. This is an evaluation on the upper levels, as the ground level area is retained unbuilt.

4.2 Porosity Evaluation

The building block's capacity to accommodate future household expansion is calculated through identifying the amount of unoccupied roof cells (1.2*1.2m unit). A proposed strategy of enabling vertical extensions when neighbouring cells are occupied prevents encroachments beyond the building footprint which would compromise environmental performance.

Block	a*	b*	a/b
A1	180	228	0.789
A2	181	234	0.774
A3	194	247	0.785
A4	218	271	0.804
Block A avg.	193.25	245	0.788
B1	179	219	0.817
B2	186	226	0.823
B3	176	226	0.779
B4	176	227	0.775
Block B avg.	179.25	224.5	0.799
C1	176	217	0.811
C2	194	229	0.847
C3	184	235	0.783
C4	190	244	0.779
Block C avg.	186	231.25	0.805
D1	211	264	0.799
D2	195	257	0.759
D3	234	267	0.876
D4	188	255	0.737
Block D avg.	207	260.75	0.793
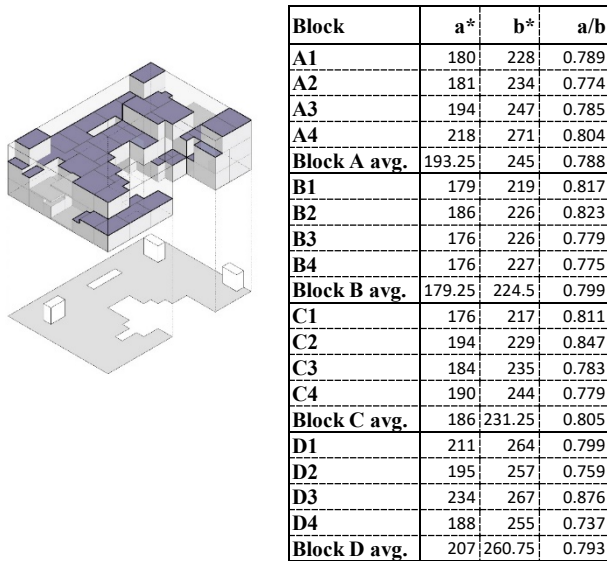
Figure 12. Porosity. **Table 5.** Building block porosity level.
a = unoccupied roof cells, (1.2*1.2m units).
b = building footprint (1.2*1.2m units).

Block Type D contains more unoccupied roof space. However, the ratios of unoccupied roof space to building footprint area indicate that some blocks with smaller footprints (C2, B1) have strong capacity to address future household growth whilst providing a better ratio of open space / built environment.

4.3 CFD Analysis

Rather than evaluating the building block's performance in CFD individually and in isolation, different urban configurations (clusters of 7) are tested in order to approximate more realistic conditions. The implemented evaluation focuses on the air flow rate across the cluster. The existing urban patch is also modelled and tested for comparison.

For the set up the evaluation, the urban patch (similar in size to one RW) is created to place building blocks. This resulted from related research yet focusing on urban scale. Each patch can be populated with maximum seven blocks separated by 2.4-3.6m-wide streets, similarly to local networks. The patch is evaluated via CFD analysis four times, each time populated with the same block type (D, A&C, B), and the last with random block types. The analysis is in turn conducted for an existing kampung patch (Melayu), demonstrating an ill-ventilated urban scenario due to open space scarcity. The later analysis shows the

significance of investigating a strategy that addresses the lack of ventilation, whereas the former present the potential of this system to improve it through configuring the urban void. *(figure 13)*

The current evaluation can serve as a preliminary study for the next phase of urban research. From this simple set-up, some geometrical principles that facilitate a well-ventilated environment can be abstracted, as follow:

- Streets of 2.4m width perpendicular to the wind direction are not ventilated.
- The open spaces located at the beginning of the streets, transform their shape into conical, that results in wind velocity increase (venturi effect).
- The combination of open spaces from neighboring building blocks, causes accumulated urban void, that results in better air ventilation in the urban cluster.
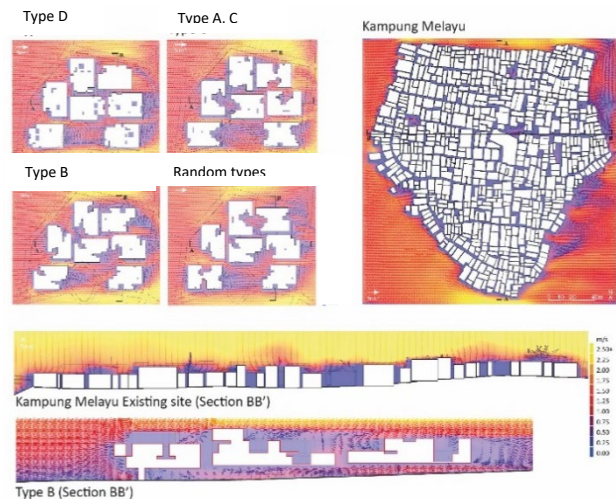
Figure 13. CFD analysis on building block configuration and Kampung Melayu

5 CONCLUSION

The presented computational model has the capacity to simulate the growth patterns of *kampungs* in Jakarta. Whereby the definition of context-specific aggregation logic can be modified to suit different informal urban settlements around the world, large number of building block simulations are necessary in order to verify the standard deviation in built morphology and emergent voids. Further top-down specification can be implemented, based on flood risk, environmental performance and population density to generate contemporary housing alternatives that are better suited for addressing climate change and population increase.

In this study, the low-lying kampung houses which make up for Jakarta's congested urban sprawl is reimagined as an elevated, porous community-scale block, through design specifications that sought to reduce flood risk and improve thermal comfort through increased urban ventilation. Later, the aggregation methodology can be developed on two

scales. On the building scale, an expanded ruleset can simulate future self-built practices based on initial inputs of family sizes. On the urban scale, a selection of mass-customised building blocks can be tested on an integrated system using evolutionary algorithms that seeks out optimal urban configurations.

From the bottom-up approach, relationships between contextual programmatic functions are transcribed into aggregation logics with a degree of flexibility. The creation of a CA approach with high degree of variation simulates local practices of self-built and user agency for continual upgrading. The provision of contect-specifc base units depending on family sizes, collectively generate diversity within the larger neighbourhoods. These outputs not only challenge the local housing provisions that are identical and without potential for future expansion, they demonstrate an alternative approach that is capable of varied growth patterns with high levels of mass customisation.

The primary distinction of the CA study is in the adoption of a concurrent methodology which combines top-down and bottom-up approaches. In comparison to former precedents that utilise CA for form-finding, the proposed computational model showcases the inclusion of programmatic relationships, to simulate self-built organisation and programmatic distribution.

As the framework is based on common principles found in most informal urban settlements, it has the potential to be adapted for other places in the world. Notwithstanding the added function's substantial requirements for multiple trials, feedback loops and bespoke evaluation, the proposed process recognises the dialectical nature of designing 'self-built' systems. Rather than designing a singular housing solution, the generated alternatives value the informal as potential organising principle, aiming to retain the socio-cultural characteristic of Jakarta's *kampungs*.

ACKNOWLEDGEMENTS

REFERENCES

1. Anzalone, P. & Clarke, C. *Architectural applications of complex adaptive systems*. Connecting Crossroads of Digital Discourse. ACADIA 2003 (2003), 324-335.

2. Coates, P. and Schmid, C. *Agent Based Modelling.* In: Brown, A, Knight, M and Berridge, P. (eds.) *17th eCAADe, Architectural Computing: from Turing to 2000*, (1999), 652-661.

3. Dinçer, A. E., Çağdaş, G. & Tong, H. *A Computational Model for Mass Housing Design as a Decision-Support Tool*. 12th International Conference on Design and Decision Support Systems in Architecture and Urban Planning, 25-27.08.14. Proceedings (2014), 271-291.

4. Heylighen, F. *Self-organization in Communicating Groups: The Emergence of Coordination, Shared References and Collective Intelligence.* In: Massip-Bonet À., Bastardas-Boada A. (eds.) *Complexity Perspectives on Language, Communication and Society. Understanding Complex Systems*. Springer, Berlin, Heidelberg (2013).

5. Hoekstra, A.G., Kroc, J., Sloot, P. *Simulating Complex Systems by Cellular Automata*. Springer-Verlag Berlin Heidelberg (2010).

6. Herr, C.M. & Ford, R.C. *Cellular Automata in Architectural Design: From Generic Systems to Specific Design Tools*. In: Chang, T.W., Moleta, T.J., Park,D. (eds.) *Automation in Construction, Volume 72, Part 1,* (2016), 39-45.

7. Herr, C.M. & Karakiewicz, J. *Algogram: Automated Diagrams for an Architectural Design Studio.* In: Dong A., Moere A.V., Gero J.S. (eds.) *Computer-Aided Architectural Design Futures*, CAADFutures (2007), Springer, Dordrecht.

8. Krawczyk, R.J. *Exploring the Massing of Growth in Cellular Automata.* Generative Art Conference (2003), Milano, Italy.

9. Lim, J. Y. *The Malay House: Rediscovering Malaysia's Indigenous Shelter System*, Institut Masyarakat (1991).

10. Batty, M. *Cities and Complexity - Understanding Cit-ies with Cellular Automata, Agent Based Models and Fractals*, Cambridge, Massachusetts, The MIT Press (2005).

11. Ernawati, R., Santosa, H. R., & Setijanti, P. *Facing Urban Vulnerability through Kampung Development.* Case Study of Kampungs In Surabaya, Indonesia. Vol. 1, N° 1, (2013), 1-6.

12. Funo, S., Yamamoto, N., & Silas, J. *Typology of Kampung Houses and Their Transformation Process, A Study on Urban Tissues of an Indonesian City.* Journal of Asian Architecture and Building Engineering, JAABE Vol. 1, N° 2, (2002).

13. Sihombing, A. Kampung Kota. *Locality of Places and Spaces*. International Research Journal of Architecture and Planning, Vol. 1(1), (2015), 2-10.

14. Panahikazemi, L. & Rossi, A. *Spatializing the Social: Computational Strategies for Intervention in Informal Settle-ments.* USA: Int. Journal for Housing Science, Vol.38, N° 2, (2014), 95-104

An Adaptive Workflow to Generate Street Network and Amenity Allocation for Walkable Neighborhood Design

Yang Yang[1], Samitha Samaranayake[2], Timur Dogan[3]

[1]Cornell University
Ithaca, NY, USA
yy848@cornell.edu

[2] Cornell University
Ithaca, NY, USA
samitha@cornell.edu

[3] Cornell University
Ithaca, NY, USA
tkdogan@cornell.edu

ABSTRACT

This paper proposes a novel generative workflow for walkable neighborhood design. Key components of the workflow include automating the process of parsing the map data, building contextual models with population and amenity data, conducting an integrated mobility simulation, and generating a street network and amenity allocation plan accordingly. The proposed framework is versatile and adaptive by allowing designers to tune simulation parameters and customize the decision-making process. The applicability and effectiveness of the workflow are tested in a pedestrian-oriented neighborhood design case study. Three scenarios that adapt to different design goals and boundary conditions are presented. This research equips designers with capabilities to co-design of mobility solutions and urban form early on in the design process. Further, it can be leveraged by more stakeholders in sectors such as real estate, public services, and public health to make decisions as the urban built environment has a fundamental impact on all these fields.

Author Keywords

generative urban design; walkability; mobility; simulation.

1 INTRODUCTION

Population growth, urbanization and ever-increasing vehicle use in urban areas have a significant impact on the quality of life and the environment. Increasing traffic-related energy consumption, greenhouse gas emissions, air and noise pollution, as well as lifestyle-related health issues such as obesity and diabetes can be promoted by poor urban design [1]. While these are worrisome circumstances, the need for urban renewal and densification [2] also provides a unique opportunity to rethink planning paradigms and design approaches. Emerging design movements aim to remedy the aforementioned mentioned issues. They [3] promote high density, walkable neighborhoods as one solution for these challenges. Studies have shown that walkable neighborhoods can significantly reduce traffic-related pollution and lower the risk for chronic diseases [3,4], support local businesses, promote tourism, attract investors, higher property values [6] and foster an increase in social capital and political participation [7]. Walkable amenities, one of the most important ingredients of a

walkable city, have also been associated with socioeconomic growth [7,8] and quality of life [10]. Understanding the implications of urban design choices on walkability while incorporating this understanding into early stages of urban design process provides a unique opportunity to address these issues. This is particularly important because street grids hardly ever change once the urban design is set [11].

One of the major challenges in designing walkable neighborhoods is the lack of effective metrics and workflows that can provide measurable and actionable feedback to facilitate design decision making. To evaluate the walkability of cities, researchers have proposed to rank neighborhoods based on the distance and density analysis of points of interest (POI) in the city. These walkability ratings, commonly referred to as Walkscore [12], are computed on a scale of 1-100 and include factors such as accessibility to amenities like grocery stores, restaurants, banks, parks, and schools. Generative design workflows for walkable neighborhoods that leverage this metric have been developed [13]. They usually first generate an urban layout by spatial logics and then optimize the Walkscore through an evolutionary process that places additional amenities until a sufficiently high score is reached. However, the main question regarding the use of Walkscore as a sole metric in such workflow is its insensitivity to the interactive relationship between key urban design parameters including street network, amenity allocation, and population distribution in the model. This can result in several questionable design decisions: Firstly, generating a street network without considering amenity placement or population distribution can be problematic because the latter two factors significantly influence street utilization and therefore play important roles in urban morphology. Secondly, placing or adding services and amenities only to drive up Walkscore may not be feasible as those new services may not be sustainable as demand is spread too thinly. Thirdly, the amounts and categories of amenities to which it is essential to have walking access differ by population groups. Thus, designers should be able to evaluate walkability with demographic-specific metrics. As a result, it is imperative to propose a more integrated

generative urban design workflows that can take into consideration all the mentioned factors.

Additionally, common generative design workflows rely on optimization solvers such as the Genetic Algorithm, which searches for optimal solution evolutionarily based on each iteration's performances on certain metrics. Although such a strategy is widely used in many generative urban design studies [13,14], it remains questionable in terms of speed of convergence and stability [15]. Although such a trial & error approach is unavoidable for certain ill-defined design problems, this paper proposes a simplified, efficient and transparent approach using the Urbano toolkit [16]. In this workflow, the first step is to build a mobility model with urban data such as streets, buildings, amenities and population densities. Then the potential pedestrian volumes of streets and amenities are evaluated by a mobility simulation. The simulation outcome can directly inform the generative process of the street network and amenity allocation. Throughout this process, designers can customize the model and control the simulation by tuning key parameters and changing variable constraints so that different design conditions can be accommodated.

Overall, this paper introduces a novel generative urban design workflow that is sensitive to street networks, amenity allocation, and population distribution. The workflow is implemented in a pedestrian-oriented neighborhood design case study, and its adaptivity is tested by accomplishing different design requirements.

2 METHODS

Urbano allows designers to build mobility models, run the network and amenity analyses within the Rhinoceros CAD platform and the visual scripting environment Grasshopper. The automated modeling and simulation process is used to drive the generative processes described in this paper.

2.1 Data-Driven Modeling

There are three layers of data that are necessary for the mobility model: Street network, amenities (points of interest), and buildings with building-level population information. Knowing the location and boundary of the site, Urbano can import streets, points of interest (POIs) and buildings, along with their metadata, from sources such as shapefiles (shp), OpenStreetMap (osm) [17] or Google Places API [18]. Streets, buildings, and amenities are represented by geometric primitives such as curves or points. Metadata such as names, types, and addresses are attached to the geometric data using serializable dictionaries that can be modified and customized alongside the geometric objects parametrically within Grasshopper or

through the CAD user interface in Rhino. If required information, such as building-level population, is not accessible from the sources, Urbano provides functions that can infer data or can help to synthesize this information using other data sources. For example, it can estimate building-level population size using total building floor area, customized area usage breakdown, and generalized occupant densities [16,17].

2.2 Trip-Sending Simulation

The simulation framework is initially driven by the Activity Demand Profile (ADP). The ADP describes pedestrian activities over time and can be adapted to reflect activities of specific demographics. One way to derive location-specific ADP is to interpret the spatiotemporal distribution of human activities in a local area by measuring the activeness in urban amenities in this area [21]. The main data source for this method is Google Places "Popular Times" data. Table 1 shows a sample set of the integrated ADP which presents the hourly percentage of population that engages in particular activities in one day in the case study area (Figure 5). Each column represents a one-hour time slot in a day, which can be further synthesized into time periods such as morning, noon and evening. Figure 1 is a graph for ADP data using a 24-hour timeline, which depicts a more detailed activity distribution. The y-axis refers to the overall amount of activities, which peaks during the day and dips in the early morning. Each color layer represents the demand pattern of an amenity. For example, banks and post offices tend to stop service in the early afternoon, while bars and pubs become dominant activities at midnight. Nevertheless, ADP data can also be customized by the designer to target an assumed demographic group. In the simulation, one or multiple sets of ADP can be utilized to represent different human activity patterns coexisting in the area.
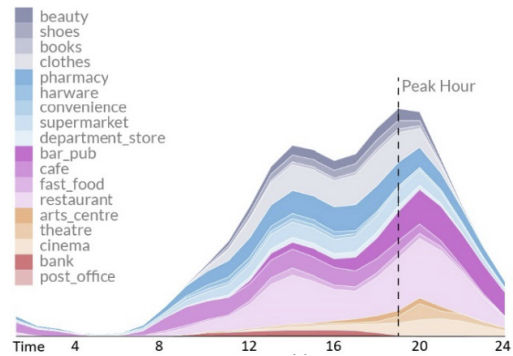
Figure 1. 24-hour timeline representing Activity Demand Profile.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
errands	0	0	0	0	0	0	0	0	0.4	1.1	1.3	1.5	1.6	1.8	1.8	1.8	1.6	1.1	0.3	0.2	0.2	0.2	0	0
restaurant	1.1	0.1	0	0	0	0.4	1.7	5.5	8.3	9.1	9.9	13.3	19.6	22.1	19.5	15.8	14.2	16.8	20.6	21.8	19.4	14	8.2	3.3
grocery	1.8	0.9	0.7	0.5	0.7	0.8	1.6	3.1	5.7	7.6	9.4	11.2	12.9	13.8	14.2	14.1	14.5	14.9	14.3	12.3	9.7	7.3	5	3.2
shopping	0	0	0	0	0	0	0	0	0.8	3.3	7.1	10.7	12.8	14.9	15.1	15	16.1	17.5	16.2	10.7	4.4	1	0.4	0
entertainment	6.2	4	3	1	0	0	0	0	0	0.3	0.4	2.7	5.3	5.7	7.2	7.9	11.6	17.2	24.8	31.6	29.1	25.8	21.8	17

Table 1. Sample of Activity Demand Profile data for the study area. (unit: %)
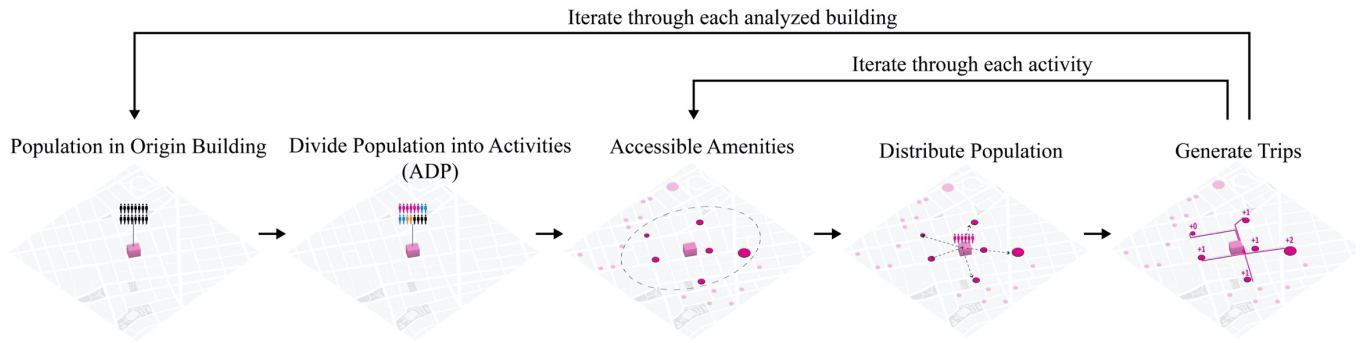
Figure 2. Diagram of the trip-sending simulation algorithm.

The following simulation is based on a trip-sending process, and the concept of each trip is made up of multiple information: the origin building, the destination amenity, the route taken and the corresponding population. The executed trip-sending algorithm is as follows (Figure 2):

1. **Input** model consisting of streets, amenities, and buildings; one or multiple ADPs with same time metrics

2. **For each** time step of ADPs

3.   **Initialize** empty list of trips

4.   **For each** building in the model

5.   Set the building as origin and get its population size

6.   Divide the population into activities according to the percentage data in its ADP

7.   **For each** activity

8.   Search for corresponding amenities as destinations within walking distance (user-defined) using a shortest-path algorithm influenced by *biased routing factors*

9.   Distribute the activity population to destinations according to *biased destination factors*

10.   Generate trips with the distributed population and add the trips to the list

11. **Output** a data tree of trips grouped by time steps

The *biased factors* are inputs allowing for more control by designers. There are currently two types of them in the presented algorithm. The first one is the *biased destination factor*, defined as the weight of a destination which determines the proportion of the population sent to them. The higher the weight, the more proportion of the total accessible population the amenity can receive. This factor can be set according to the quantifiable quality parameters of amenities such as the capacity, popularity or rating. The other one is the *biased routing factor*, defined as the coefficient of calculated street length. This coefficient allows certain street segments to be "shortened" in the simulation so that they can be more utilized in the shortest-path routing process, or "lengthened" in the opposite way. By modifying this coefficient, the simulation can count into more factors influencing the route choice other than distance, such as shade, landscape, facades, urban environment, etc.

Output: Each time step has a distinct set of resulting trips. Trips data can be post-processed into three complementary metrics: Street Hits and Amenity Hits. Street Hits counts how many people use a certain street segment on all trips. Amenity Hits tallies up the total number of people that are sent to a specific amenity on all trips. Moreover, a building-level Walkscore can also be computed according to its original method [12] using data of all trips that originate from a single building. All these metrics will inform the generative design process.

2.3 Generative Process

The primary setup for the generative method is to replace the original design site with a dense grid mimicking a virtual environment that lets people cross freely. This step parallelly establishes a perpendicular coordinate grid representing all potential locations for amenities (Figure 3). Street Hits results from the simulation on this dense grid can reveal people's potential movement trails across the site. Street segments with high Hits can be transformed into new roads in the designed network. Amenity Hits results on the coordinate grid can identify the most profitable locations for amenities that grant most walking access to the population in the model, which can inform the placement of new amenities.
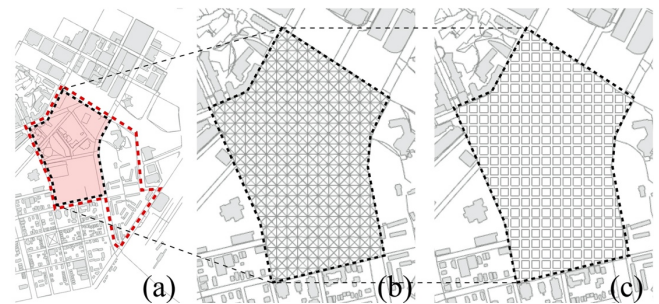
Figure 3. Using part of the case study's site as an example (a), the primary setup is to replace the original site with a dense grid (b) mimicking a virtual environment that lets people cross freely. It parallelly establishes a coordinate grid (c) representing all potential locations for amenities

197
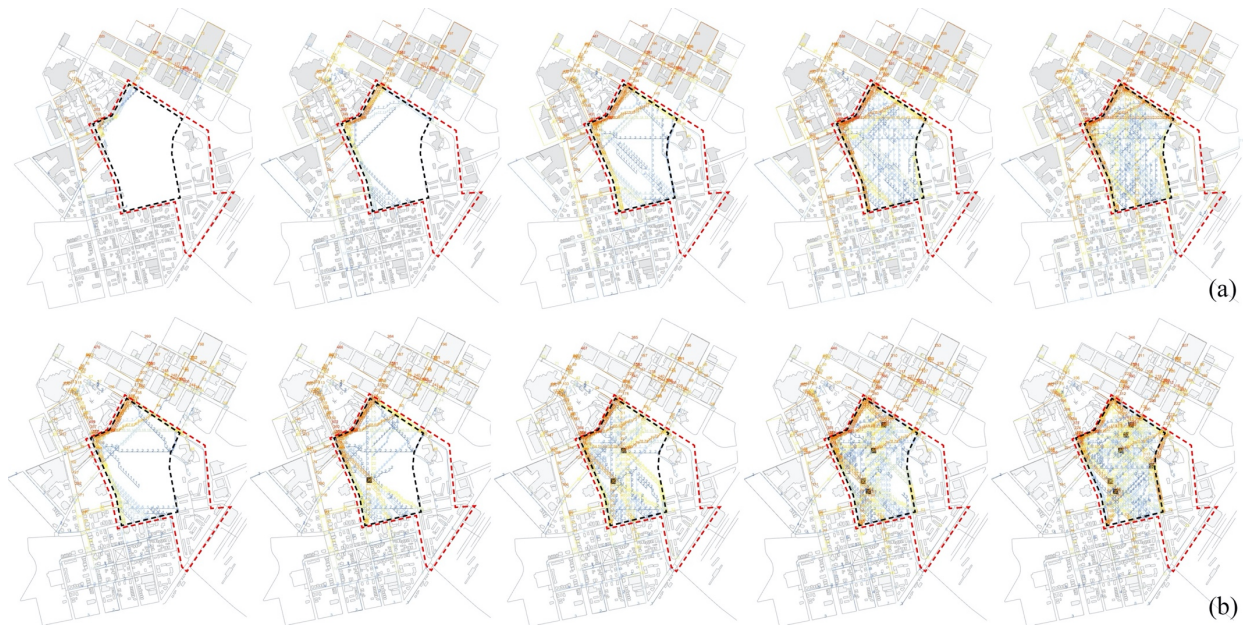
Figure 4. Two series of visualizations of Street Hits output. The first row (a) shows evolving results when increasing the user-defined walking distance. The second row (b) presents changing results when placing new amenities on the design site.

Due to the nature of the trip-sending algorithm, the outcome spontaneously concerns the interactivity between streets, amenities and population distribution in the model. it is also responsive to changes. Figure 4 shows the example of two series of visualizations of Street Hits output from the same simulation (Figure 3) but only with a specific parameter modified. The first row shows when increasing the user-defined walking distance, more trips are generated because more amenities become accessible to all buildings within that distance. The second row presents changing results when placing new amenities on the design site. Besides these, other parameters such as the biased routing and destination factors, or ADPs can all impact the generative process. A highly customizable framework like this enables designers to tune the generative process for specific conditions or goals in design practice as in the case study.

3   CASE STUDY

The site in Figure 5 is located in New Haven, Connecticut. It has residential neighborhoods to the South, a high-density commercial district to the North, institutions to the West, and an industrial area to the East. A highway and a railway adjacent to the East and the North form obstacles isolating the site. The current street network does not connect urban amenities well as Street Hits analysis on the original site reveals that most activities do not take routes across the site (Figure 6). The site is predominantly used as parking lots and is considered as an empty area as the initial condition in this study. However, the site has great potential as it can connect the railway station to the city and fill the gap in pedestrian mobility between different urban areas surrounding it.

The initial site model consists of existing streets, amenities, and buildings. Building types are categorized into

residential and non-residential so that population data can be synthesized accordingly. The overall design objective is to develop a new mixed-use and pedestrian-oriented neighborhood that can alleviate some of the described connectivity issues. All the existing streets and buildings on the site are supposed to be overridden.
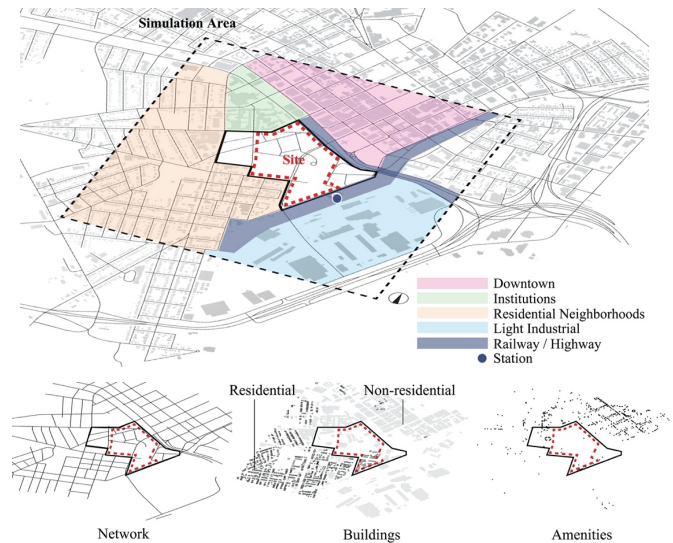
Figure 5. The site and the main components in the initial site model.

Figure 6. Street Hits analysis of the original site shows that most activities do not take routes across the site.

198
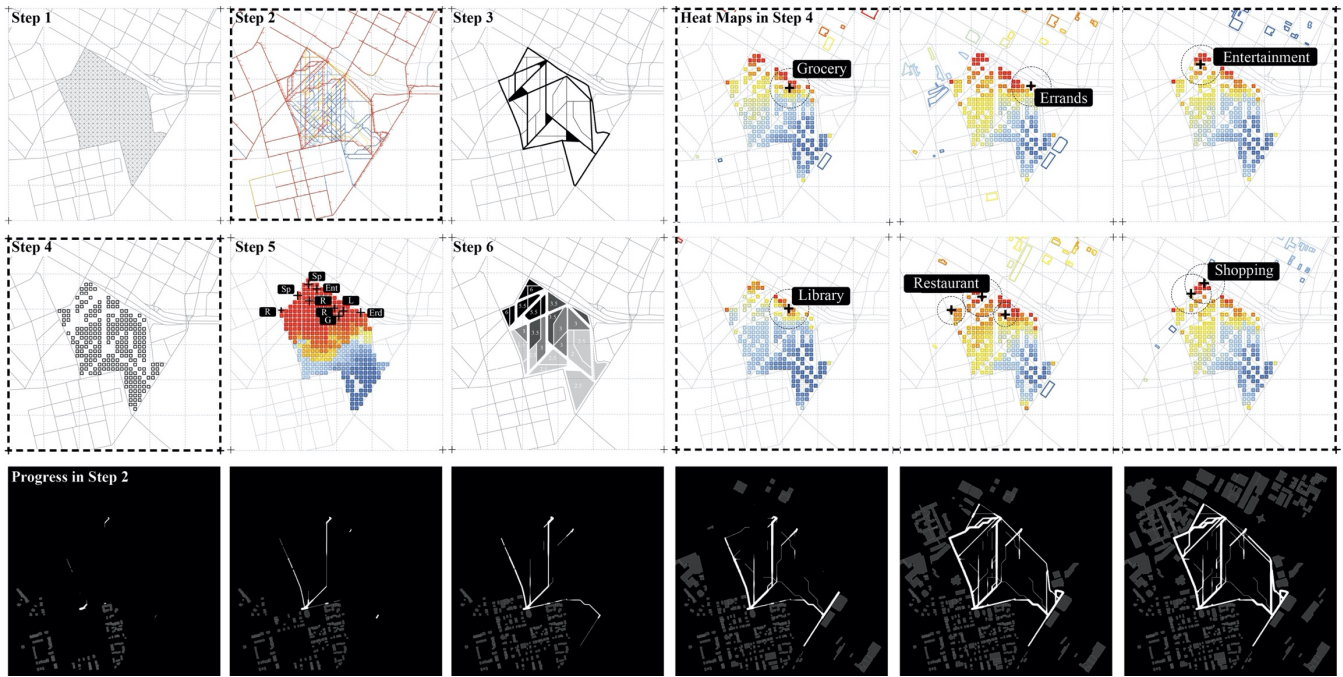
Figure 7. Diagram of the generative workflow of Scenario One.

To test the proposed workflow, three scenarios are generated under distinct design intentions and assumptions. They follow the same primary generative methods but differ in modeling and decision-making process.

3.1 Scenario One

This scenario aims to create a better linking zone. The new network should contribute more efficient passing routes so that increasing pedestrians can go across the site and support new businesses and amenities.

In Figure 7, Step 1 is the primary setup. Step 2 presents the Street Hits result using one normalized ADP data of Table 1. The progress in Step 2 is expanded below to show how the main routes with high Street Hits gradually become visualized during the simulation iteration of all buildings. Using this result, Step 3 generates a new street network by straightening the busiest streets, merging the minor links, and converting the largest intersections into potential plazas. This step is drawn manually at the current stage. Based on this network, Step 4 analyzes Amenity Hits distribution by setting all cells on the 20m*20m coordinate grid as one amenity type. The results are visualized in heat maps highlighting the recommended locations for each amenity type (grocery, errands, library, entertainment, restaurant, shopping). Since the population density is much higher in the northern downtown, all new amenity locations tend to concentrate at the north edge for maximized potential patronage. However, heat maps still vary in color uniformity due to the influences of existing nearby amenities. For amenity allocation, designers can place amenities on the best performing locations, conduct a new simulation for the scenario and evaluate the Amenity Hits of the newly placed amenities. Comparing the resulting Hits

with the Hits of other same-type amenities existing in the model can help designers measure the balance between supplies and demands of new amenities, and then make decisions about their amounts and locations.

One thing to mention is this paper does not focus on the parcellation in the lots or generation of building footprints. However, to present that the previous results can be further developed into actionable design scenarios, the workflow integrates the last two steps. Step 5 computes the final Walkscore based on the coordinate grid, which is used in Step 6 to inform the distribution of development density (FAR) on the generated lots. Since better walkability indicates improved property values, the lots with a greater Walkscore get a higher density. Figure 8 presents an example of how the final masterplan could look.
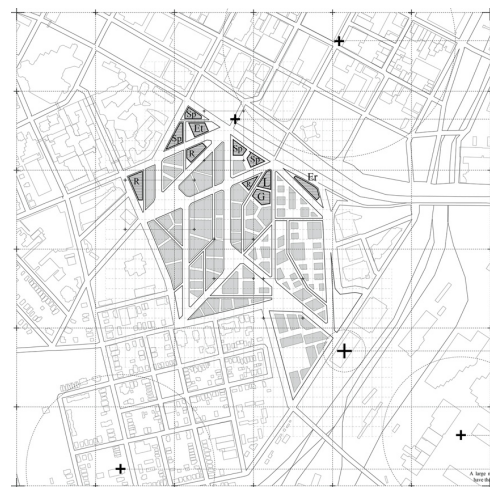
Figure 8. Sample masterplan developed for Scenario One.
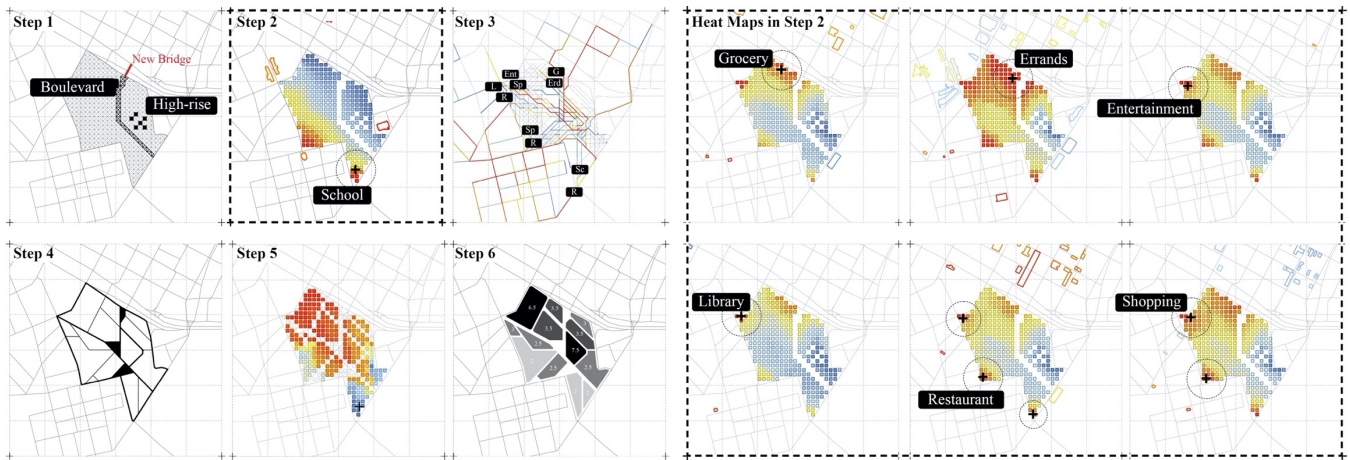
Figure 9. Diagram of the generative workflow of Scenario Two.

3.2 Scenario Two

This scenario intends to include several design conditions: (1) a proposed pedestrian boulevard linking the railway station and the downtown area through the new bridge; (2) a predefined and specifically located high-rise cluster which will hold a high density of population on design site; (3) a prospective new school mostly serving the southern residential neighborhood. The generative workflow is adjusted to address these particular design issues (Figure 9).
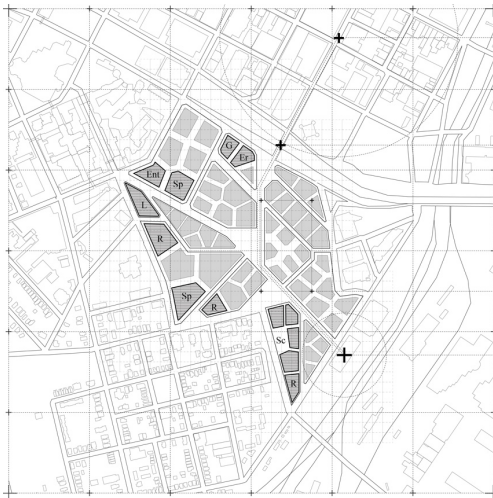
Figure 10. Sample masterplan developed for Scenario Two.

Step 1 models the boulevard and high-rise cluster based on the primary setup. The boulevard is modeled by setting the segments on the grid along the route with a biased routing factor of 0.5. The high-rise cluster is modeled by setting 500 population to each of five high-rise locations on the coordinate grid. Step 2 visualizes heat maps of Amenity Hits. These heat maps differ from Scenario One because the "shortened" boulevard in the simulation attracts more trips crossing the site through this route, and the high-rises also bring more population as consumers. Among all amenities, the analysis of Hits for school differs from the others because it only considers pedestrians coming from the southern residential area. The school's heat map result also

reveals this adjustment as it is best located on the corners that are closest to the neighborhood side. With all amenity allocation decided, Step 3 visualizes Street Hits on the grid, and Step 4 generates the new street network accordingly. The final two steps of Walkscore analysis and lot-level FAR distribution remain the same as Scenario One. Figure 10 presents an example of how the final masterplan looks.

3.3 Scenario Three

Instead of using one normalized ADP data in the first two scenarios, this scenario considers the temporal difference in street and amenity utilization. It aims to create a 24-hour active neighborhood by overlapping the generative results of time steps. However, this scenario only generates street network while the amenity allocation is an input.

In Figure 12, Step 1 specifies an input of the amenity allocation scenario. Step 2 visualizes the Street Hits for three time-periods: morning, noon and evening. Step 3 generates the network by filtering the most vibrant streets at these times and overlapping them together. Step 4 and Step 5 remain the same for Walkscore analysis and lot-level FAR distribution. An additional Step 6 demonstrates people's dynamic movement on the network over time.
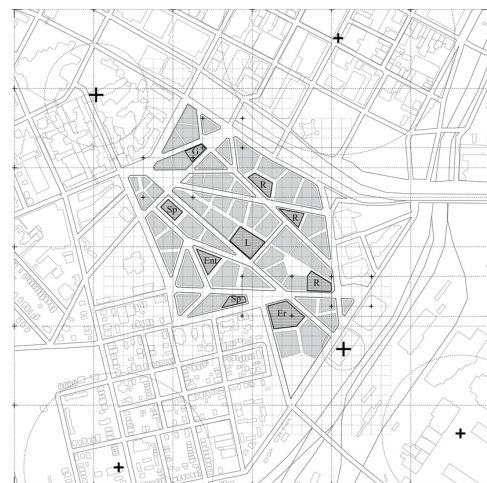
Figure 11. Sample masterplan developed for Scenario Three.
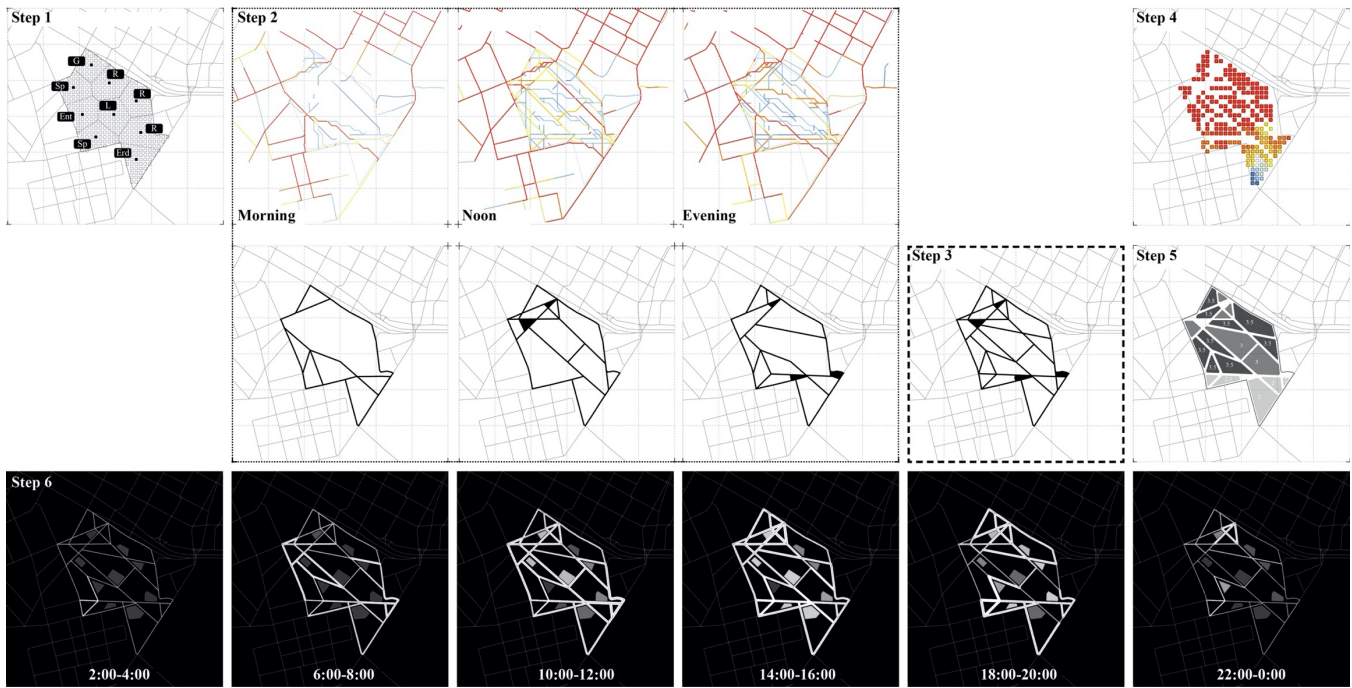
Figure 12. Diagram of the generative workflow of Scenario Three.

3.4 Comparison

To be adaptive to different design goals and conditions, the proposed workflow varies in three scenarios. Firstly, the sequences of decision-making are different. Scenario One generates streets first because it aims to create better links between the surrounding built environment. Scenario Two has more specific requirements for program allocation. Thus, the street network is generated later in the workflow so as to address the new design conditions. Secondly, the generative parameters change. The same one normalized ADP is used in the first two scenarios while the third one uses the ADP of multiple time periods. Also, the biased routing factors and the population distribution are modeled differently in Scenario Two. More parameters such as biased destination factors or walking distance limits have not yet been modified among three scenarios. They are able to allow more precise controls by designers.

4   LIMITATION AND PROSPECT

The limitation of modeling is data quality. For example, high-quality GIS data is only provided in major metropolitan areas. Also, some open data source such as OpenStreetMap has significantly fewer POI entries compared with other sources such as Google data. Consequently, a model that uses data where only a few POIs have been recorded, will yield misleading results. The workflow proposed in this paper will be able to benefit from the ongoing efforts to improve urban data systems.

As for the simulation, there are difficulties in thorough validation, because there is no openly available reference data with which to compare the results. The current ADP data is derived using user-generated data such as Google Places "Popular Times" data, which mostly relies on GPS.

To provide a basic check of the simulation results, five randomly selected samples of cafes and restaurants in the simulation model are used for comparative study. Figure 13 plots their Amenity Hits and their real profiles in Google "Popular Times" data (both normalized and scaled to 1.0) together. There is a certain level of consistency, but exceptions also exist. In the future, more detailed data such as opening hours can be leveraged to improve consistency further. Though "Popular Times" data is indeed an input of deriving ADP, such comparison can still verify the interpretation process of the framework along with other synthesized parameters such as population distribution.
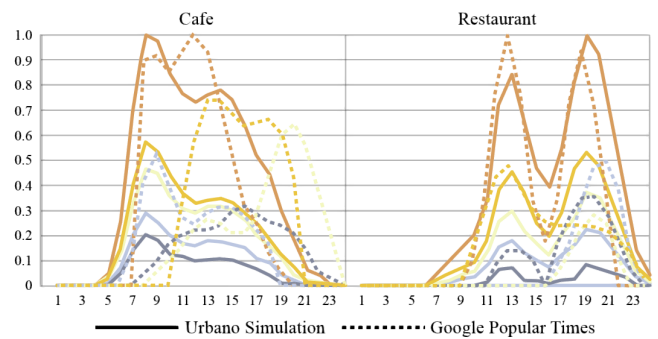
Figure 13. Comparison of five random samples' Amenity Hits results and their real profiles in Google Popular Times data.

In the generative workflow, caution is needed when taking advantage of its adaptivity. Some customized parameters, such as biased routing and destination factors, provide designers with the power to control the generative process, but they also open to the risk of being arbitrary or biased. More sophisticated metrics defining these factors are in demand.

5 CONCLUSION

Design decisions such as zoning, density, program allocation, and the layout of public spaces and streets can have a fundamental impact on the performance of mobility systems. Employing urban planning to mitigate traffic-related problems is widely recognized as an effective strategy. It is expected that this research can equip designers with capabilities that enable the co-design of mobility solutions and urban form, thus motivating the early discovery of cost-effective solutions.

This paper proposes a novel workflow of automating the process of parsing the map data, building contextual models with population and amenity data, conducting integrated mobility simulation, and generating street network and amenity allocation for urban design. The effectiveness and adaptivity of the workflow are tested in a pedestrian-oriented neighborhood design case study by generating three scenarios for different design goals and conditions. This versatile framework can contribute to the design profession and education in terms of increasing awareness and responsiveness to mobility-related urban factors. Moreover, as mobility metrics also have economic and environmental implications, the proposed framework can become more valuable by including other stakeholders in urban development. Practitioners in sectors such as real estate, public services, and public health can leverage the analysis result to make decisions, and designers can benefit from involving a broader range of data and metrics from these fields into the design solution-seeking process.

REFERENCES

[1] A. Loukaitou-Sideris, "Addressing the Challenges of Urban Landscapes: Normative Goals for Urban Design," *Journal of Urban Design*, vol. 17, no. 4, pp. 467–484, Nov. 2012, doi: 10.1080/13574809.2012.706601.

[2] UNEP, "Sustainable Buildings and Climate Initiative," 2015. Accessed: Aug. 05, 2015. [Online]. Available: http://www.unep.org/sbci/AboutSBCI/Background.asp.

[3] T. Haas, Ed., *Sustainable urbanism and beyond: rethinking cities for the future*. New York: Rizzoli, 2012.

[4] L. D. Frank, J. F. Sallis, T. L. Conway, J. E. Chapman, B. E. Saelens, and W. Bachman, "Many Pathways from Land Use to Health: Associations between Neighborhood Walkability and Active Transportation, Body Mass Index, and Air Quality," *Journal of the American Planning Association*, vol. 72, no. 1, pp. 75–87, Mar. 2006, doi: 10.1080/01944360608976725.

[5] I.-M. Lee and D. M. Buchner, "The Importance of Walking to Public Health:," *Medicine & Science in Sports & Exercise*, vol. 40, no. Supplement, pp. S512–S518, Jul. 2008, doi: 10.1249/MSS.0b013e31817c65d0.

[6] S. Claris and C. Luebkeman, "Cities Alive: Towards a walking world," ARUP Publications, 2016. [Online]. Available: https://www.arup.com/-

/media/arup/files/publications/c/cities-alive_towards-a-walking-world_lowres.pdf.

[7] K. M. Leyden, "Social Capital and the Built Environment: The Importance of Walkable Neighborhoods," *Am J Public Health*, vol. 93, no. 9, pp. 1546–1551, Sep. 2003, doi: 10.2105/AJPH.93.9.1546.

[8] T. N. Clark, R. Lloyd, K. K. Wong, and P. Jain, "Amenities Drive Urban Growth," *Journal of Urban Affairs*, vol. 24, no. 5, pp. 493–515, Dec. 2002, doi: 10.1111/1467-9906.00134.

[9] A. Zandiatashbar and S. Hamidi, "Impacts of transit and walking amenities on robust local knowledge economy," *Cities*, vol. 81, pp. 161–171, Nov. 2018, doi: 10.1016/j.cities.2018.04.005.

[10] G. F. Mulligan and J. I. Carruthers, "Amenities, Quality of Life, and Regional Development," in *Investigating Quality of Urban Life*, vol. 45, R. W. Marans and R. J. Stimson, Eds. Dordrecht: Springer Netherlands, 2011, pp. 107–133.

[11] T. Dogan, S. Samaranayake, and N. Saraf, "Urbano: A new tool to promote mobility-aware urban design, active transportation modeling and access analysis for amenities and public transport.," in *SimAUD 2018*, 2018.

[12] M. Brewster, D. Hurtado, S. Olson, and J. Yen, "Walkscore.com: A New Methodology to Explore Associations Between Neighborhood Resources, Race, and Health," presented at the 137st APHA Annual Meeting and Exposition 2009, 2009, [Online]. Available: https://apha.confex. com/recording/apha/137am/pdf/free/4db77adf5df9fff0d3caf5 cafe28f496/paper205082_1. pdf.

[13] T. Rakha and C. F. Reinhart, "Generative Urban Modeling: A Design Workflow for Walkability-Optimized Cities," 2012.

[14] D. Nagy, L. Villaggi, and D. Benjamin, "Generative urban design: integrating financial and energy goals for automated neighborhood layout," in *Proceedings of the Symposium on Simulation for Architecture and Urban Design*, Accessed: Jan. 09, 2020. [Online]. Available: https://dl.acm.org/doi/abs/10.5555/3289750.3289775.

[15] T. Wortmann, C. Waibel, G. Nannicini, R. Evins, T. Schroepfer, and J. Carmeliet, "Are Genetic Algorithms Really the Best Choice in Building Energy Optimization?," in *SimAUD 2017*, Accessed: Jan. 10, 2020. [Online]. Available: https://www.researchgate.net/publication/317277483_Are_Ge netic_Algorithms_Really_the_Best_Choice_in_Building_Ene rgy_Optimization.

[16] T. Dogan, Y. Yang, and S. Samaranayake, "Urbano: A tool to promote active mobility modeling and amenity analysis in urban design," *TAD Translation*, 2020.

[17] "OpenStreetMap." https://www.openstreetmap.org/#map=17/52.90112/-1.22824 (accessed Apr. 01, 2020).

[18] "Places | Google Maps Platform," *Google Cloud*. https://cloud.google.com/maps-platform/places (accessed Apr. 01, 2020).

[19] ASHRAE, "ASHRAE Handbook—Fundamentals (SI)," ASHRAE, Atlanta, 2013.

[20] SIA, *2024: Standard-Nutzungsbedingungen Für Die Energie-Und Gebäudetechnik*. Zürich: Swiss Society of Engineers and Architects, 2006.

[21] Y. Yang, S. Samaranayake, and T. Dogan, "Using Open Data to Derive Local Amenity Demand Patterns for Walkability Simulations and Amenity Utilization Analysis," in *Proceedings of ECAADe 37 / SIGraDi 23*, Porto, Portugal, 2019.

Multi-Objective Optimization for Zero-Energy Urban Design in China: A Benchmark

Thomas Wortmann[1] and Jonathan Natanian[2]

[1]Xi'an Jiaotong Liverpool University
Suzhou, China
thomas.wortmann@xjtlu.edu.cn

[2] Technical University of Munich
Munich, Germany
jonathan.natanian@tum.de

ABSTRACT

Environmental simulation supports the design of more sustainable, zero-energy neighborhoods, especially when leveraged with multi-objective optimization. This study explores the tradeoff between urban density and energy balance—specifically, monthly load match between energy usage and generation—in terms of courtyard, slab, and tower typologies for a hypothetical neighborhood in Shanghai. Using this problem as a multi-objective optimization benchmark, the study compares the evolutionary algorithms HypE and NSGA-II with RBFMOpt, a novel, machine learning-related algorithm. The study concludes that RBFMOpt finds slightly better Pareto fronts and is much more robust, and that courtyard typologies are the most efficient for both low- and high-density neighborhoods.

Author Keywords

Zero-Energy Urban Design; Multi-objective Optimization Benchmark; Model-based Optimization; Monthly Load Match

ACM Classification Keywords

Computing methodologies ~ Modeling and simulation

Applied computing ~ Operations research ~ Decision analysis ~ Multi-criterion optimization and decision-making

Theory of computation ~ Design and analysis of algorithms ~ Mathematical optimization ~ Mixed discrete-continuous optimization

1   INTRODUCTION

The substantial environmental impact of continuing rapid urbanization in China [10] and elsewhere make more sustainable urban design an urgent priority. Energy and other environmental simulations support this design by predicting energy demand. Multi-objective optimization (MOO) builds on this support by automatically searching for efficient design candidates and exploring trade-offs between conflicting objectives [20]. But simulation-based optimization studies on urban, rather than building, form, are rare [9]. Performance evaluations of the often-employed evolutionary MOO algorithms are even rarer, despite an urgent lack of efficient MOO algorithms for fast-paced conceptual design stages and/or time-intensive simulation-based problems [8]. This study explores the tradeoff between

urban density and monthly energy balance in terms of different typologies for a hypothetical neighborhood in Shanghai and compares two evolutionary MOO algorithms with a novel, machine learning-related one.

2   BACKGROUND

This section introduces energy simulation and optimization as design methods for zero-energy urban design, briefly introduces MOO and discusses the MOO algorithms evaluated in the study.

2.1 Zero-Energy Urban Design Simulation

While zero-energy design, i.e., locally balancing energy generation and supply, is well-known on the building scale, relationships between urban form and energy balance have been studied only more recently. In hot climates, this relationship is driven by solar potential [25]. Natanian et al. [21] perform energy balance and daylight simulations for five urban, Mediterranean typologies, using a wide range of urban and building design parameters. Kanters and Wall [16] explore the impact of form, materiality and orientation on the zero-energy potential of Swedish building blocks. In China, Li et al. [19] simulate carbon emissions for two existing neighborhoods in Macao, and Farzana et al. [11] predict the energy consumption of residential buildings in Chongqing. These studies support the design of more sustainable cities by simulating energy demand and/or solar energy potential on an urban scale. Building on such simulations, simulation-based optimization automates the search for well-performing design candidates [9] and supports the co-evolution of design problems and solutions [30].

2.2 Zero-Energy Urban Design Optimization

Compared to the literature on simulation-based optimization on the building scale [9], optimization studies on the urban scale are relatively scarce and recent. In Switzerland, Nault et al. [22] present a design-decision support tool for zero-energy urban design that uses MOO and user-test it with a workshop. Waibel at el. [27] optimize the form and district energy system of four office buildings in Zurich. Camporeale and Mercader-Moyano [4] optimize the energy balance of slab and high-rise typologies in four temperate climates in Argentina and Spain. These studies do not offer performance comparisons of MOO algorithms.

2.3 Multi-objective Optimization

Single-objective optimization (SOO) aims to find the single point (i.e., design candidate) that is the best according to a single performance metric (i.e., objective). By contrast, MOO aims to find the set of points that is the best according to multiple conflicting objectives. (If the objectives are not conflicting, it is better to use a scalarization technique to combine the objectives [8] and to use SOO.) The points in the best set are non-dominated, i.e., it is impossible to improve one of the points' objectives without losses in others. When graphed in terms of the objectives, the set of non-dominated points yields a Pareto front graph (PFG). A PFG visualizes the trade-off between objectives and thus supports designers' decision making [23]. But, typically, the true Pareto front (i.e., the true set of non-dominated points) is very difficult to find due to the often-large numbers of possible design candidates and the exponentially larger difficultly resulting from optimizing multiple objectives. The goal of MOO thus is to find approximations of the true Pareto front. MOO algorithms are evaluated by how effectively and efficiently they find such approximations.

Accordingly, it is important to test MOO algorithms, especially on simulation-based problems that are relevant for design practices. But most papers in architectural and urban design that use MOO do not offer performance evaluations [9]. In, to the knowledge of the authors, one of the most extensive evaluations of MOO algorithms for zero-energy building design, no MOO algorithm converged "completely to the best Pareto front" [14]. A comparison of SOO and MOO methods concludes that the latter are "generally not acceptable in usual engineering problems" due to their inefficiency [5]. This paper offers a simulation-based MOO benchmark as a step to more frequent and rigorous evaluations of MOO algorithms on simulation-based problems. It introduces and tests RBFMOpt (Radial Basis Function Multi-objective Optimization), a novel, machine-learning-related MOO algorithm that potentially is more efficient than evolutionary MOO algorithms.

2.4 Multi-objective Algorithms

The paper compares RBFMOpt's performance with two evolutionary MOO algorithms, HypE [2] and NSGA-II [7]. Both HypE and NSGA-II employ crossover, mutation and selection operators. The difference between them lies in how they select points for the next generation (i.e., optimization iteration): HypE employs the hypervolume metric, which measures the volume of the objective space (i.e., the space that is defined by the objectives) that is dominated by a set of points [31]. NSGA-II employs a less mathematically rigorous combination of selecting points that dominate other points and of avoiding crowding, i.e., avoiding points that are too similar in terms of their objective values [31]. This "crowding metric" aims to ensure a good distribution of points on the PFG.

RBFMOpt aims to find the points on the true Pareto front in sequence. To find a single point on the Pareto front, RBFMOpt calculates a weighted sum of objectives and optimizes the resulting single value with a SOO algorithm, RBFOpt [6]. RBFMOpt approximates the set of non-dominated points by iterating over a series of weights based on a Halton sequence [13]. For each weight, RBFOpt performs a number of optimization cycles. Each cycle consists of several optimization steps that shift from exploring unknown points to exploiting (i.e., aiming to improve) around known, well-performing points. A disadvantage of using a series of weighted sums is that, in its basic form, this approach cannot find concave parts of Pareto fronts. RBFMOpt overcomes this disadvantage with Chebychev scalarization, a special form of weighted sum [8]. RBFMOpt employs the implementations of a Halton sequence and Chebychev scalarization in Pagmo 2, an optimization library for Python [3].

The approach of RBFMOpt as described so far is sound, but hardly novel. What makes it innovative and, potentially, highly efficient is that, unlike HypE and NSGA-II, RBFOpt is a learning algorithm that benefits from known points by constructing and optimizing surrogate models that predict simulation results [6]. RBFMOpt exploits this property by, at each iteration, restarting RBFOpt and feeding it all known points, reweighted according to the current set of weights. In this way, the surrogate model that RBFOpt starts with becomes increasingly accurate with every iteration, allowing more effective and efficient optimization cycles. ParEGO is a MOO algorithm that employs a similar approach [17]. RBFMOpt has greatly outperformed NSGA-II on mathematical test functions [18] and won the "Two Objectives Expensive" track of the Black Box Optimization Competition (BBCOMP) 2019 [12]. This paper presents the first simulation-based benchmark of RBFMOpt.

3 METHODOLOGY

This section discusses the parametric model and its optimization variables, the performance simulations of energy demand and supply, the formulation of the two optimization objectives and the benchmarking methodology.

3.1 Parametric Model and Optimization Variables

The optimization problem considers the design of a zero-energy, residential neighborhood in the metropolitan area of Shanghai. According to the Chinese climate classification, Shanghai lies in the hot summer cold winter zone [15]. This zone is especially challenging for sustainable design due to its large annual temperature variations. Compared to a more desirable mixed-use design, this problem employs a simplified parametric model and simulation. The neighborhood consists of nine blocks that are arranged in a three-by-three gird. Each block is 100 meters wide. The grid's rotation relative to North and the distance between the blocks (i.e., streets) are variable.
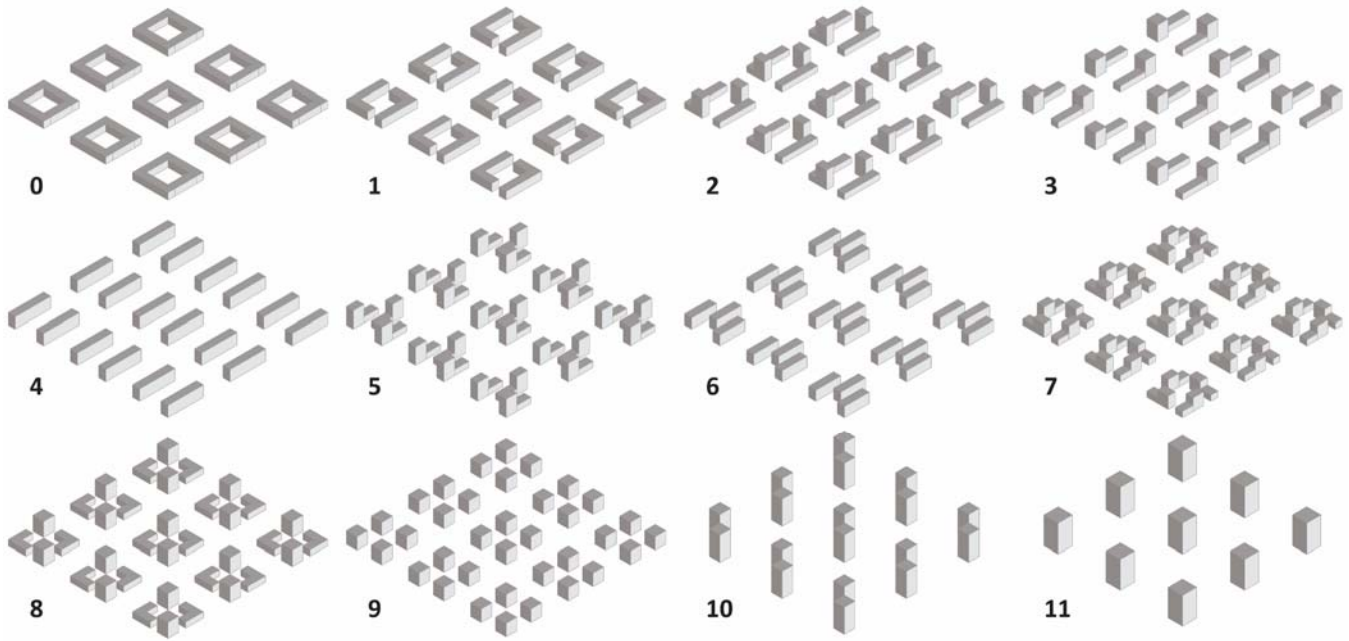
Figure 1. Typologies 0-11 (with Block Distance 75.0, FAR 4.0 and Maximum Slab Floors 5)

Each urban design variant is populated homogenously with one of twelve typologies. These typologies consist of different variants and combinations of courtyards, slabs and towers (Figure 1) with a fixed floor-to-floor height of three meters. The heights of the slabs and/or towers of each typologies vary depending on the floor area ratio (FAR) and a variable maximum height for slabs. The selected typology, FAR and maximum slab height is identical for each of the nine blocks in the grid. A more difficult version of this problem would consider different typologies, FARs and maximum slab heights for each block. Finally, we vary the window-to-wall ratio (WWR) of the towers and slabs, separately for each compass direction. As such, the optimization problem has seven continuous and two integer variables (Table 1).

	Low	**High**	**Integer?**
WWR North (Percent)	30.0	80.0	No
WWR West (Percent)	30.0	80.0	No
WWR South (Percent)	30.0	80.0	No
WWR East (Percent)	30.0	80.0	No
Rotation (Degrees)	-180.0	180.0	No
Block Distance (Meters)	25.0	125.0	No
FAR	1.0	6.0	No
Maximum Slab Floors	1	10	Yes
Typology	0	11	Yes

Table 1. Optimization Variables

Note that the typologies are not in any particular order (i.e., the "Typology" variable is combinatorial, not ordinal).

3.2 Performance Simulations

We use the Honeybee plug-in [24] for Grasshopper, a visual programming tool for parametric geometry generation, to simulate the monthly energy demands for heating, cooling and electricity and potential monthly energy generation for photovoltaics (PV), all in kWh/m^2. We derive the parameters of the energy models from [15]. Anticipating a shift to more sustainable building design, we assume that most surfaces that are not windows and that are not North-facing are covered with PV panels and that PV elements are integrated into windows and sun shades. WWRs vary from 30% to 80%. We map this range on a range of PV coverage from 60% to 40%. (Note that this simplified calculation is based on the average WWRs of the West, South and East facades, and does not consider the facades individual WWRs.) For roofs we assume 70% PV coverage. We assume a solar cell efficiency of 17% and a coverage of 90% (the remainder are for framing and gaps) for the PV panels.

Using the energy demand and supply data for each iteration, we calculate the monthly load matches (MLMs), i.e., the ratios between energy supply and demand. A higher ratio indicates a better balance between energy generation and energy demand. Calculating this ratio for each month instead of over the whole year emphasizes the self-sufficiency of the neighborhood and implies a reduced need for energy storage and/or external energy generation. On an Intel Core i7-8750H CPU with 2.2GHz, one regeneration of the parametric model with the energy demand and supply simulations takes about five minutes on average.

3.3 Optimization Objectives

The two optimization objectives for this MOO problem consist of maximizing the average MLM and maximizing FAR. In other words, this MOO problem studies the trade-off between density and zero-energy in urban design. This trade-off is of high relevance to urban futures in China and elsewhere: On a larger scale, urbanization (i.e., a higher density) can lead to higher energy efficiency, for example through a reduced need for transport. On the smaller scale, a higher density in terms of FAR makes it more difficult to balance energy demand and supply locally [10].

Note that FAR is both a variable that determines the geometry for the simulation and an optimization objective. This unconventional choice is justified by considering alternative approaches: As a single-objective problem with density only as a variable, an efficient algorithm would gravitate quickly to the lowest density, because it results in the highest MLM. To explore the relationship between the two objectives for higher densities, one could optimize a series of single-objective problems with different densities that are held constant for each problem. But this approach, while potentially resulting in a more precise Pareto front, also requires many more function evaluations. As such, the employed problem formulation is a short-cut to quickly generate an approximation of the Pareto front.

3.4 Benchmark Methodology

To compare the MOO algorithms' performance, we run each of them for 192 function evaluations, i.e., model regenerations and simulations. 192 corresponds to about nineteen simplex gradients. (One simplex gradient is the number of variables plus one.) Nineteen simplex gradients are a small function evaluation budget considering the number of variables and objectives. (For SOO energy simulation problems, 100 simplex gradients are recommended [28].) Nevertheless, one single run takes about fifteen hours due to the complexity of the model and simulations. Since long run times are reported as an important barrier to the uptake of optimization in zero-energy building design [1], a MOO benchmark with a relatively realistic simulation model and relatively short run times is especially relevant for design practice. To account for the algorithms' use of randomness, we repeat this experiment five times.
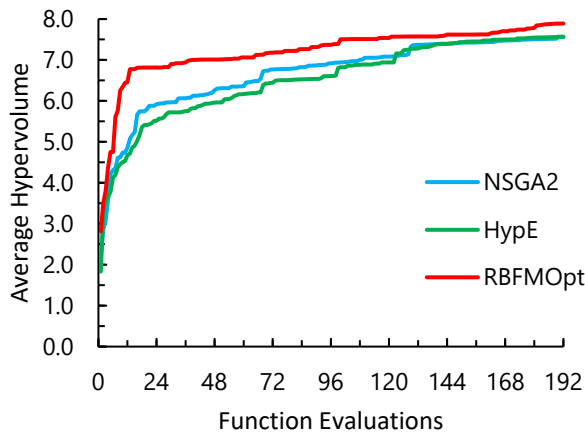
We employ the implementation of HypE in Octopus [26], a MOO plug-in for Grasshopper also used by [4], and the implementation of NSGA-II in the Pagmo 2 library [3]. NSGA-II and RBFMOpt are available in Opossum 2.0 [29], a SOO and MOO plug-in for Grasshopper. We select a population size of 24 for the evolutionary NSGA-II and HypE. This population size results in eight generations for NSGA-II and seven for HypE, since, in Octopus, the first generation's size is twice as large.

For RBFMOpt, we set the number of cycles at three. This parameter determines how long the algorithm spends optimizing each set of weights. For this problem, this results in around eight sets of weights. (Mathematically, each set of weights corresponds to a point on the true Pareto front.) For larger function evaluation budgets, we recommend a larger number of cycles, e.g. six or nine [18]. In that way, the algorithm spends more time on each set of weights, which can result in more accurate approximations of Pareto fronts. Apart from the population size for NSGA-II and HypE and the cycle number of RBFMOpt, we employ default parameters.

4 RESULTS

We assess the convergence and robustness of the MOO algorithms in terms of the found hypervolumes [31]. We also present and analyze one representative PFG per algorithm and the typologies of the design candidates on the best-known Pareto front.

4.1 Hypervolume: Convergence and Robustness

Hypervolume measures the volume of the objective (i.e., trade-off) space that is dominated by a set of points, i.e., design candidates. Mathematically, if one set of points spans a larger hypervolume than another, it cannot be worse than the other set, i.e., the two sets are at least equally good. Conversely, if one set is better than another, it will also span

Figure 2. Speed of Convergence
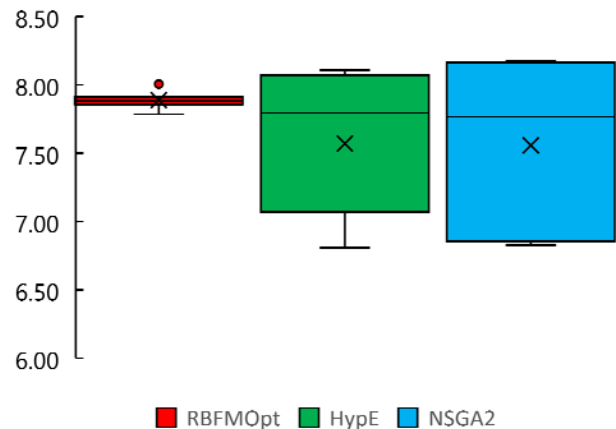
Figure 3. Robustness

a larger hypervolume [31]. As such, hypervolume is not a perfect performance indicator for MOO, but, importantly, it allows performance comparisons with a single number.

Figure 2 presents the average hypervolume of the five runs for each algorithm and function evaluation (i.e., the speed of convergence). NSGA-II and HypE perform very similarly. RBFMOpt performs better, but only by about 4%. This small performance increase relative to other algorithms is similar to BBCOMP 2019, but much smaller than earlier results on test functions [18].

A more decisive difference is evident in Figure 3, which shows the distributions of the hypervolumes after 192 function evaluations: Although the averages are similar, the distributions of NSGA-II and HypE are much wider, probably due to the large amount of randomness present in evolutionary algorithms. In other words, while the performance of HypE and NSGA-II can vary from excellent to worst, RBFMOpt's performance is more robust. This robustness is important especially for scenarios when limited time allows only a single run, such as concept design.

4.2 Pareto Front Graphs

Figure 4 shows the most representative Pareto front (i.e., closest to the average hypervolume) for each algorithm. Note that the fronts for HypE and NSGA-II are above average, while the front for RBFMOpt is below average. It also shows the "best known" front, i.e., the best approximation of the true Pareto from all algorithms and runs. Besides the fact that the fronts found by the three algorithms look similar—which is unsurprising given that the average hypervolumes are similar as well—several observations suggest themselves:

- The "best known" front is quite similar to the most representative fronts. This suggest that, even with the small function evaluation budget, all algorithms were reasonably successful. (But longer runtimes might result in even better approximations and a better "best known" front".)

- This success with a small function evaluation budget suggests in turn that, for a benchmark problem, this problem likely is too easy.

- The found approximations of the Pareto front are concave. Their concavity would make it impossible to find them with a "naïve" approach that uses weighted sums [8]. As such, the results show that the Chebychev weighting employed by RBFMOpt indeed overcomes this challenge.

- The points found by RBFMOpt are not very well distributed, with high-performing points at both ends of the PFG and not many well-performing points in-between. In terms of the hypervolume, this uneven distribution is compensated by RBFMOpt's success at both ends of the PFG. Nevertheless, from the point of view of providing a meaningful selection of options to urban designers [30], this distribution is suboptimal because it does not offer a gradient of FARs. This unevenness is likely caused by the relatively small number of sets of weights that RBFMOpt optimized due to the small function evaluation budget.

- From the standpoint of urban design, the PFGs indicate that one can achieve very high average
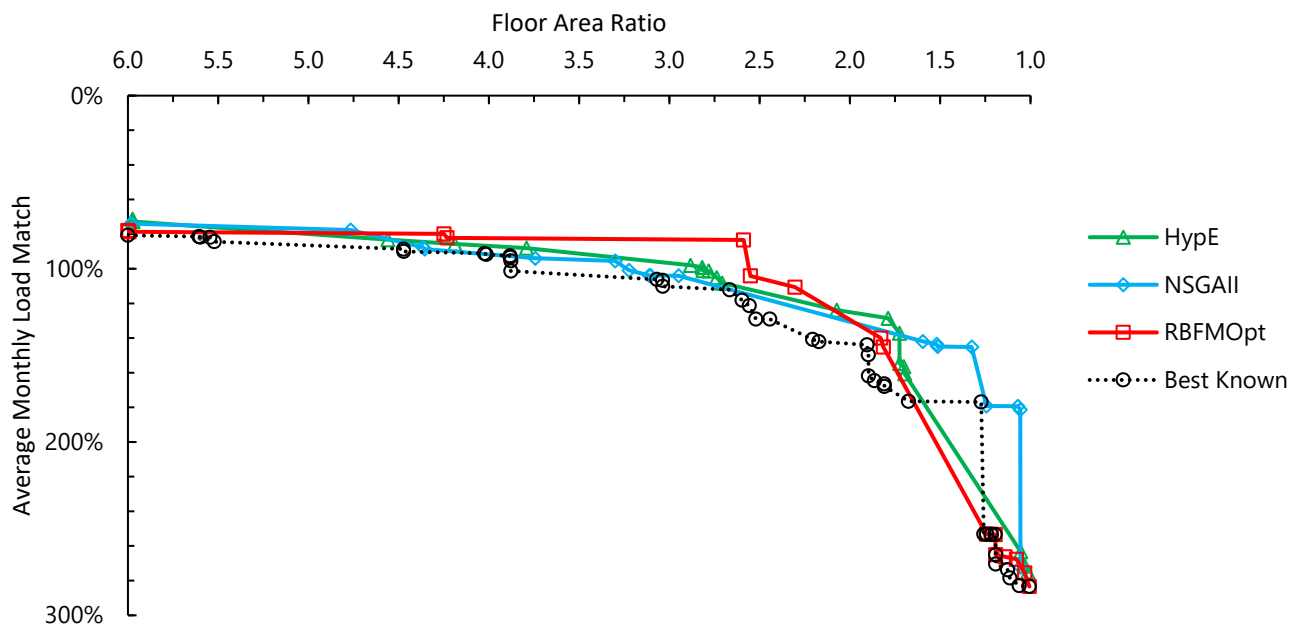
Figure 4. Pareto Front Graphs

207

MLMs up to a FAR of about 1.5. For higher FARs, MLM rapidly decreases to 100% for a FAR of around 3.0, and a respectable 80% for a FAR of 6.0, the highest considered in this study (Table 1). This result demonstrates that, by optimizing typology, WWRs, orientation and other variables, high density, low energy neighborhoods are possible for large blocks in the climate of Shanghai.

4.3 Non-dominated Design Candidates

This section analyses the 41 design candidates on the best-known Pareto front (Figure 5). Strikingly, 33 of these employ the courtyards typologies number 1 and 2 (Figure 1). Of the remaining eight, which all have FARs around 4 or higher, four employ typology number 4, and four typologies numbers 6-9. Typologies number 4 and 6 are slabs, and numbers 7-9 are variations of courtyards with four openings. As such, courtyard typologies are the most efficient for this problem, and slab typologies and courtyard variations are competitive for higher FARs. No nondominated design candidates employed typologies with towers (numbers 2, 3, 5, 10 and 11).

The higher energy balance potential of the courtyard typologies agrees with other studies in hot climates [21]. Courtyard typologies have higher surface-to-volume ratios that allow higher PV energy generation yields. But courtyard typologies also experience the highest self-shading and thus reduce undesirable solar access mostly on east and west facades. This self-shading minimizes energy demand for cooling but also PV energy generation yields. This tradeoff underscores the importance of roof top solar potential in zero-energy balance calculations. In addition to energy balance, it is important to account for the daylight, microclimatic and outdoor thermal comfort 'penalties'

associated with the courtyard's compact form, which are not considered in the present study.

5 DISCUSSION

In its current form, this study is limited by several factors, which we aim to address in future work. Currently, the calculation of the amount of PV coverage is simplified and does not take the individual facades' WWRs into account. In addition, the optimization problem currently considers only energy demand and supply, and no other factors of urban quality, such as daylight and urban comfort. Future work also aims to consider mixes of different typologies, a difficult combinatorial problem.

The benchmark currently is limited by the relatively small function evaluation budget and relatively small number of runs, as well as the—especially compared to SOO benchmarks—relatively small number of algorithms. Although the consideration of both hypervolume and PFGs is informative, future work will consider additional performance metrics such as the binary ε-indicator [8].

In terms of MOO, this study's results are somewhat inconclusive in that all algorithms performed similarly, except for their robustness. Nevertheless, the first simulation-based benchmark of RBFMOpt, a new MOO algorithm, and the small number of simulation-based MOO benchmarking studies make this study a valuable addition.

But although the distribution of non-dominated points requires improvement, these early results for RBFMOpt are encouraging: Outperforming NSGA-II, a popular MOO algorithm in the architectural and urban design communities [9], and HypE, one of the two algorithms of the most popular MOO tool for Grasshopper, in terms of the found
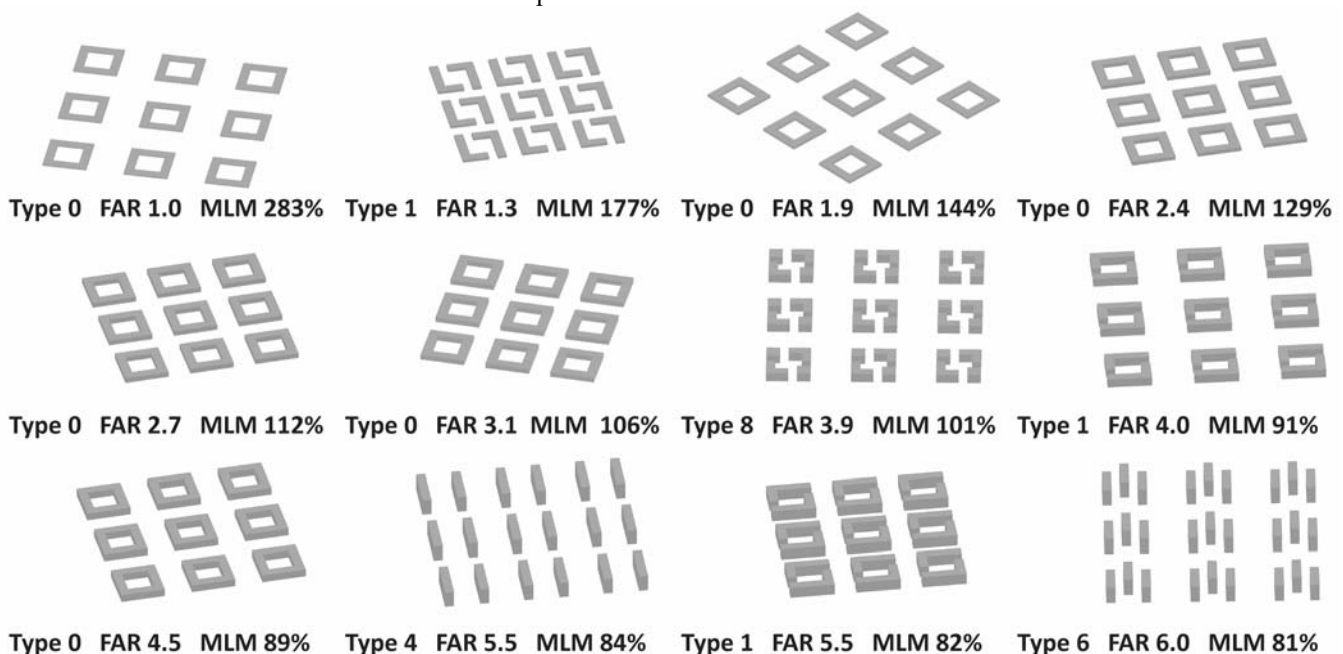
Figure 5. Selected design candidates on the best-known Pareto front (looking North East)

hypervolume is a significant achievement. Similar to RBFOpt, RBFMOpt is very robust in repeated runs on the same problem, which offers greater confidence to designers that can only afford a single optimization run. But although RBFOpt excels at small function evaluation budgets [6], RBFMOpt apparently requires larger function evaluation budgets to find not only large hypervolumes, but also well-distributed PFGs. This constitutes an important finding for the further development of RBFMOpt, since previous tests considered only the hypervolume metric [18].

In terms of zero-energy urban design, the results of this study hopefully will spur the design of neighborhoods with more efficient typologies: Low energy, high FAR ratio neighborhoods are achievable for large block sizes and climates with hot summers and cold winters when employing (mostly) courtyard typologies!

REFERENCES

1. Attia, S., Hamdy, M., O'Brien, W., and Carlucci, S. Assessing gaps and needs for integrating building performance optimization tools in net zero energy buildings design. *Energy and Buildings* 60 (2013), 110–124.

2. Bader J. and Zitzler E. 2011. HypE: An Algorithm for Fast Hypervolume-Based Many-Objective Optimization. *Evolutionary Computation* 19, 1 (2011), 45–76.

3. Biscani, F., Izzo, D., Wenzel, J., Acciarini, G., Marcus, M., Micky, C., Mereta, A., Kaldemeyer, C., Lyskov, S., Corlay, S., Benjamin, Manani, K., Mabille, J., Huebl, A., López-Ibáñez, M., Jakirkham, Jeongseok, L., Hulucc, Polygon, Zajc, L.C., Adler, J., Travers, J. Jordan, J., Smirnov, I., Nguyen, H., Lema, F., O'Leary, E. and Mambrini, A. *Pagmo 2.11.4* (2019), https://doi.org/10.5281/zenodo.3582877

4. Camporeale, P.E. and Mercader-Moyano, P. Towards nearly Zero Energy Buildings: Shape optimization of typical housing typologies in Ibero-American temperate climate cities from a holistic perspective. *Solar Energy* 193 (2019), 738–765.

5. Chiandussi, G., Codegone, M., Ferrero, S. and Varesio F.E. Comparison of multi-objective optimization methodologies for engineering applications. *Computers & Mathematics with Applications* 63, 5 (2012), 912–942.

6. Costa, A. and Nannicini, F. RBFOpt: an open-source library for black-box optimization with costly function evaluations. *Mathematical Programming Computation* 10, 4 (2018), 597–629.

7. Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6, 2 (2002), 182–197.

8. Emmerich, M.T.M. and Deutz, A.H. A tutorial on multiobjective optimization: fundamentals and evolutionary methods. *Natural Computing* 17, 3 (2018) 585–609.

9. Evins, R. A review of computational optimisation methods applied to sustainable building design. *Renewable and Sustainable Energy Reviews* 22 (2013) 230–245.

10. Fan, J.F., Zhang, Y. and Wang, B. The impact of urbanization on residential energy consumption in China: An aggregated and disaggregated analysis. *Renewable and Sustainable Energy Reviews* 75 (2017) 220–233.

11. Farzana, S., Liu, M., Baldwin, A. and Hossain M.U. Multi-model prediction and simulation of residential building energy in urban areas of Chongqing, South West China. *Energy and Buildings* 81 (2014), 161–169.

12. Glasmacher, T. and Loshchilov I. Black-Box Optimization Competition. Retrieved January 12, 2020 from https://www.ini.rub.de/PEOPLE/glasmtbl/projects/bbcomp/index.html

13. Halton, J.H. and Smith, G.B. Algorithm 247: Radical-inverse Quasi-random Point Sequence [G5]. *Communications of the ACM* 7, 12 (1964), 701–702.

14. Hamdy, M., Nguyen, A., and Hensen, J.L.M. A performance comparison of multi-objective optimization algorithms for solving nearly-zero-energy-building design problems. *Energy and Buildings* 121 (2016), 57–71.

15. Ichinose, T., Lei, L. and Lin, Y. Impacts of shading effect from nearby buildings on heating and cooling energy consumption in hot summer and cold winter zone of China. *Energy and Buildings* 136 (2017), 199–210.

16. Kanters, J., and Wall, M. The impact of urban design decisions on net zero energy solar buildings in Sweden. *Urban, Planning and Transport Research* 2, 1 (2014), 312–332.

17. Knowles, J. ParEGO: a hybrid algorithm with on-line landscape approximation for expensive multiobjective optimization problems. *IEEE Transactions on Evolutionary Computation* 10, 1 (2006) 50–66.

18. Roger, K. Tuning and Benchmarking a Blackbox Optimization Algorithm. Yale-NUS College, 2019.

19. Li, Z., Quan, S.J. and Yang, P.P. Energy performance simulation for planning a low carbon neighborhood urban district: A case study in the city of Macau. *Habitat International* 53 (2016), 206–214.

20. Lin, S.E. and Gerber D.J. Designing-in performance: A framework for evolutionary energy performance

feedback in early stage design. *Automation in Construction* 38 (2014), 59–73.

21. Natanian, J., Aleksandrowicz, O., and Auer, T.A. parametric approach to optimizing urban form, energy balance and environmental quality: The case of Mediterranean districts. *Applied Energy* 254 (2019), 113637.

22. Nault, E., Waibel, C., Carmeliet, J. and Andersen, M. Development and test application of the UrbanSOLve decision-support prototype for early-stage neighborhood design. *Building and Environment* 137 (2018), 58–72.

23. Radford, A.D. and Gero, J.S. On Optimization in Computer Aided Architectural Design. *Building and Environment* 15 (1980), 73–80.

24. Roudsari, M.S. and Pak, M. Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. In *Proc. of IBPSA Conference 2013* (2013)

25. Tsalikis, G., and Martinopoulos, G. Solar energy systems potential for nearly net zero energy residential buildings. *Solar Energy* 115 (2015), 743–756.

26. Vierlinger, R. Multi-Objective Design Interface. Technische Universität Wien, Vienna, AUT, 2013.

27. Waibel, C., Evins, R. and Carmeliet, J. Co-simulation and optimization of building geometry and multi-energy systems: Interdependencies in energy supply, energy demand and solar potentials. *Applied Energy* 242 (2019), 1661–1682.

28. Waibel, C., Wortmann, T., Evins, R., and Carmeliet, J. Building energy optimization: An extensive benchmark of global search algorithms. *Energy and Buildings* 187 (2019), 218–240.

29. Wortmann, T. Opossum—Introducing and Evaluating a Model-based Optimization Tool for Grasshopper. In *Proc. of CAADRIA 2017* (2017), 283–292.

30. Wortmann, T. and Schroepfer, T. From Optimization to Performance-Informed Design. In *Proc. of SimAUD 2019* (2019), 261–268.

31. Zitzler, E., Thiele L., Laumanns, M., Fonseca, C.M. and da Fonseca, V.G. Performance assessment of multiobjective optimizers: an analysis and review. *IEEE Transactions on Evolutionary Computation* 7, (2003) 117–132.

Automated Parametric Building Volume Generation: A Case Study for Urban Blocks

Iuliia Osintseva[1], Reinhard Koenig[1,2], Andreas Berst[1], Martin Bielik[1], Sven Schneider[1]

[1]Bauhaus-Universität Weimar
Weimar, Germany
iuliia.osintseva@uni-weimar.de

[2]AIT Austrian Institute of Technology
Vienna, Austria
reinhard.koenig@ait.ac.at

ABSTRACT

The aim of this paper is to demonstrate a new developed parametric workflow for efficient semi-automatic generation of design options for multi-family residential buildings. The generation process is performed by modular components that generate design variants for building volumes automatically but enable the user to combine them in the preferred order, depending on the project goals. Such an approach enables the quick search in the vast solution space and, at the same time, allows the user to influence the generation process and guide it into the desired direction, thus allowing the designer to integrate his experiences and insights. The developed methodology is applied to a case study of a residential building design competition. In this case study, we aimed to 1) generate a multitude of diverse design options for the existing project and thus validate the developed method 2) generate options, that are visually similar to the manual designs submitted by architects and thus validate if the semi-automatic generation procedure is sufficient for including concept-related features and demands into the design process.

Author Keywords

Residential architecture; design automation; parametric design; computational design; parametric block typology.

1 INTRODUCTION

The output of the planning process – the building design – has a great and lasting impact on the social, economic, and ecological performance of the building. Therefore, finding a good solution at the early stage of planning, where most of the design decisions are taken, is crucial. In order to find the best possible solutions for the residential building planning at an early design stage, multiple design options must be created, analyzed, and compared with each other. It is not possible to directly find an individual optimal solution, because various design goals usually contradict each other, like the goals of achieving a maximum of density and maximum of daylight at the same time [12]. Therefore, a good solution will be represented by a trade-off between different design criteria.

Despite various existing approaches, utilizing generative design methods, the current design practice still mostly utilizes manual sketching as a method for the solution search process. As follows, only a relatively small set of design options is explored. Accordingly, good options may not even be regarded. Therefore, even if the final design satisfies the formal demands, there is no guarantee that a better trade-off was not overlooked during the solution search.

In this paper, we present a method for the efficient automatic generation of design options for multi-family residential buildings. The advantage of computational design methods is the enlarging of the design space. However, considering the number of theoretically possible design inputs, parameters, and criteria, the solution space can be infinite, and the computational expenses make it impossible to explore the complete solution space. Accordingly, the aim of this project was to develop a framework that would explore a large number of diverse options yet remain time-efficient. To achieve this, we based the parametric generation framework on the replication of the common existing residential typologies, thus restricting the infinite size of the design space to a smaller set of solutions that are well-established in the architectural practice. Such an approach allows avoiding the generation of non-practical or exotic solutions. The developed method is applied based on the widely used CAD-Software Rhino3d [18] and Grasshopper Plugin [17] in order to provide a seamless transition of the received outputs into the overall design process.

2 RELATED WORK

Although the usage of computational tools in design remains negligible [4], there already exists a multitude of generative tools applicable to different scales of architectural design, utilizing very different approaches on generation. The challenge of creating a highly complex architectural model from a simple set of inputs was addressed by Parish & Mueller [14]. The Decoding Spaces toolbox [1] contains a set of methods for both the generation and analysis of masterplans, starting with street network generation and followed by parceling and buildings placement. Rather than focusing on the search of a single optimized solution, the toolbox instead provides a highly adaptive generation workflow. Wilson et al. [21] developed a CUrBD method as an approach to performance-based master planning, outlining the most relevant trends for categorization of best-

performing design schemes as an instrument for negotiations between multiple stakeholders with contradicting interests. However, those approaches primarily focus on the creation of the masterplans of large areas. Such masterplans demand methods on a subdivision of the area into blocks, parceling, distribution of the density program, defining the uses combination and the allocation of the public spaces, etc. However, most of the masterplans contain numerous blocks that must be further elaborated. For each of the blocks, the solution space remains vast. As for the block scale, various tools and prototypes are available on the market. Many of those tools are online applications [20, 9] and let the user insert their goal criteria and then explore the set of automatically generated solutions. Another group of tools parametrizing the generation process, and automating manual steps is built upon existing CAD software [6]. Such tools allow for a quick adaptation of design to the new variables, but on the downside, explore only one design option at a time. The most recent approaches suggest an application of trained computation models for the solution search [19].

The common feature of most of the above-mentioned applications is their level of automation, which demands the user to insert important parameters and indicate key criteria, like the desired density or the apartment mix to be placed. Further work is taken over by the application, which enables a rapid overview concerning the possible building morphology for a site. In the current project we intended to set up a framework that remains quick while searching automatically for the good design alternatives and as well lets the user interact with the generation process. Such interaction should allow the designer to include his desired experience-based characteristics for a particular project into the generation process and therefore enable the consideration of non-quantifiable parameters, the lack of which is often being mentioned when criticizing computational design tools.

3 METHOD

The presented generation method is based on the replication of the common residential typologies. Most of the residential architecture examples can be categorized into a list of typologies [2, 3, 15]. In this research, we differentiate between three principal typologies (see Figure 1) and consider other residential building forms as combinations of these.

Figure 1. Principal typologies: block, slab and solitary.

Our framework allows us to replicate those three typologies in their basic form and to further perform variations over them to produce diverse yet buildable design options. In the following chapters, we will demonstrate the generation process on the example of an urban block typology.

3.1 Inputs

The inputs needed for the generation can be categorized into three groups: geometric, numeric, and regulations. By the geometric inputs, we mean the construction plot and, if available, the context information, such as streets, public spaces, and the neighboring buildings. Except for the construction plot outline (further referred to as plot) that must be provided by the user himself, the rest of the context data can be taken from publicly available sources, like the OpenStreetMap [13]. Context information is evidently important conceptually, as we evaluate the architectural designs, among other criteria, by their integration into the existing environment. Nevertheless, the context data is also used to derive important factors for the generation, such as the average height of the neighbor buildings. Context data is as well useful for the evaluation of the generation's outputs when it comes to the evaluation of the performance of our design for the surrounding buildings. For example, it is a common practice for the approval procedure of a residential project to provide daylight analysis. The goal of such an analysis is to prove that the daylight situation for the neighbors did not deteriorate significantly because of the new planning. At last, by regulations input, we mean the consideration of the spacing from the plot boundaries. In this case, spacing is defined as a distance from the plot outline till the building edge, that must remain empty [8]. By indicating the geographical location of the plot, we may derive the corresponding spacing indexes and consider them during the generation to ensure that the developed designs remain buildable according to the local building regulations.

3.2 Basic Block

First, the basic form of the chosen typology is applied to the construction plot. In the case of the block typology, the basic block is represented by the offsetting of the plot outline inwards of the plot by the chosen building depth. Depending on the selected building height, the needed spacing is calculated before placing the block. Once the basic block (or their range) is set, we can vary the form to explore the possible design space (see Figure 2).

3.3 Actions

Manipulations of the basic form (further: actions) are developed as modular grasshopper components. Each of the actions performs one type of manipulation over the current block form and can be applied after the basic block or after another action; multiple actions create a sequence. Each action has a set of unified inputs and outputs for the geometry representation as well as a section of user-inputs for manual control over the generation.
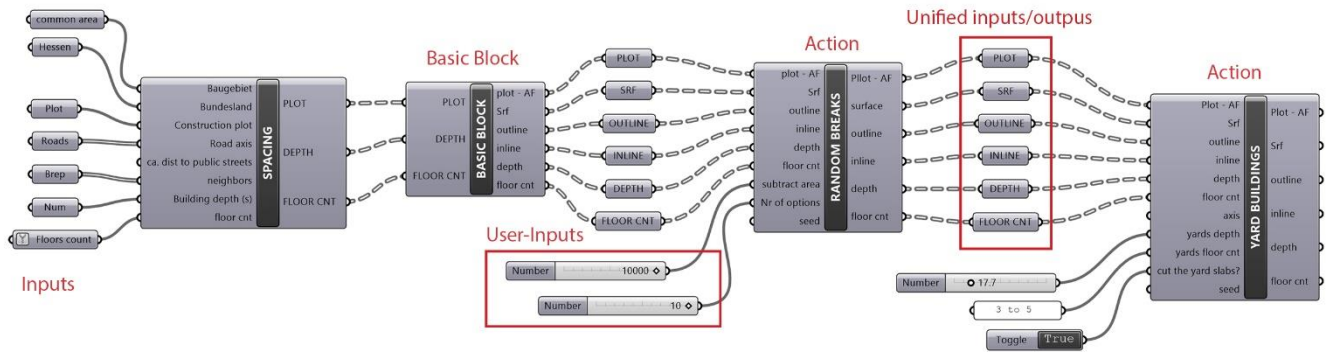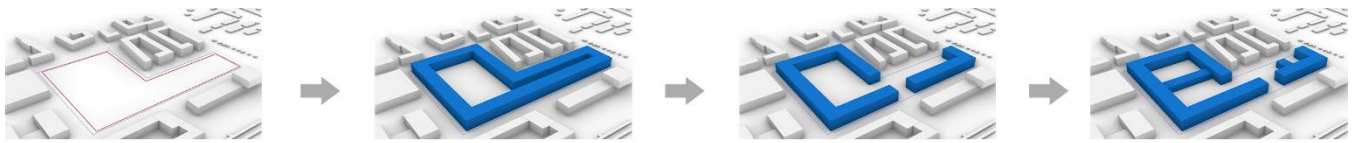
Figure 2. Simplified scheme of the developed parametric flow: Inputs→ Basic form of typology (here: block) →Actions

Because of the unified principal inputs and outputs (see Figure 2), the actions can be applied in any order. The unified inputs/outputs contain the following parameters: plot form after applying the spacing, surface (building footprint), contours of the surface for further facades generation, building's width, and the number of floors per building.

In fact, each action adds or subtracts some area from the current block variation. For example, the action "Towers" splits the block form into smaller fragments and extrudes one or several of them (see Figure 3). Each action runs automatically and is controlled by a set of user inputs (see Table 1). Most of the inputs are numerical and define the amount of area to be manipulated. For the "Towers" action, the user can select the range of areas for possible fragments placement, as well as the maximal number of the floors and the total amount of the area to add. Based on that, multiple options with randomly placed towers will be generated, whereas each of the options will add the desired amount of area to the development.

Figure 3. Comparison of the user-driven tower placement using the red marked allocation attractor (left) with the randomized allocation (right).

Each of the actions can perform automatically and intervene at multiple random locations of the block. However, the user can indicate the intervention placement with an allocation attractor, and the action will be applied at the defined spot. In the example of the "Towers", the action will not place them at different random positions around the block, but instead at the preferred location (see Figure 3).

Type of the user inputs	Illustration
Attractor point(-s) for precise allocation of the needed intervention. Example: placement of "Setbacks".	
Attractor line(-s) for precise allocation of the needed intervention. Example: placement of the "Yard buildings".	
Numeric values to define geometry. Example: width of the "Rooftop terraces".	
Amount of area to manipulate (add or subtract).	

Table 1. Overview of the different possible types of user inputs.

213

Depending on the range of the inputs, each action can perform multiple variations over the same initial block input. Therefore, each next action added to the generation sequence multiplies the number of the output designs. Currently, the framework contains eight actions for the block typology (see Figure 4): "Open Block"(1) deletes one or multiple edges of the block; "Setbacks Edges"(2) and "Setbacks Corners"(3) place setbacks along the block perimeter in order to create plazas or to break the monotony of the long street-aligned facades; "Reduce Height"(4) subdivides the input form into fragments and reduces the density by sinking the height of one or multiple fragments; "Yard buildings"(5) places additional buildings in the inner yard space of the block; "Towers"(6) creates both spacing and solar envelopes [10], using the Ladybug plugin [11] for the given site and places the towers within the previously constructed envelopes; "Rooftop terraces"(7) creates the terraces on top of the block form; "Random Breaks"(8) creates passageways around the block perimeter. The actions listed above are performed automatically; however, it is on the user to choose the sequence of actions as well as to define their inputs. It is as well possible to use multiple sequences of actions in parallel (see Chapter 4, scenario 2). Depending on the selected actions sequences, the design variants can differ greatly. Therefore, the outcomes are more diverse than usual for parametric systems, which typically apply the same process for the changing inputs.

4 DEMONSTRATION

To validate the proposed framework, we demonstrate the possible application on two different scenarios for the same test case, represented by the competition "MK6 Theresienhöhe, Munich. [16]

In the first scenario, we tried to achieve optically similar designs like those created manually by the architects and compare the outputs with the winning competition entries by the density index (FAR). Such an approach allows us to test the methods of the user's interaction with the generation process by trying to follow the formal criteria during the creation of the variants. With this, we can test if such a framework could be useful in the architectural workflow as a sketching tool for the early design phase.

One of the challenges that designers face is the great amount of changing goals and demands in the course of one project. Sufficient interaction methods would allow the implementation of the changes in the course of further project development. In addition, the various analysis, connected to the generation, directly report on the consequences of the morphology interventions, thus creating transparency in the decision making.

The first proposal is the closed block with the small breaks along one façade and a higher fragment at the southern corner. In order to achieve similar outputs, the following sequence was used: Basic Block → Random breaks→ Towers; resulting set: 200 options. Towers; resulting set: 200 options (see Figure 5).
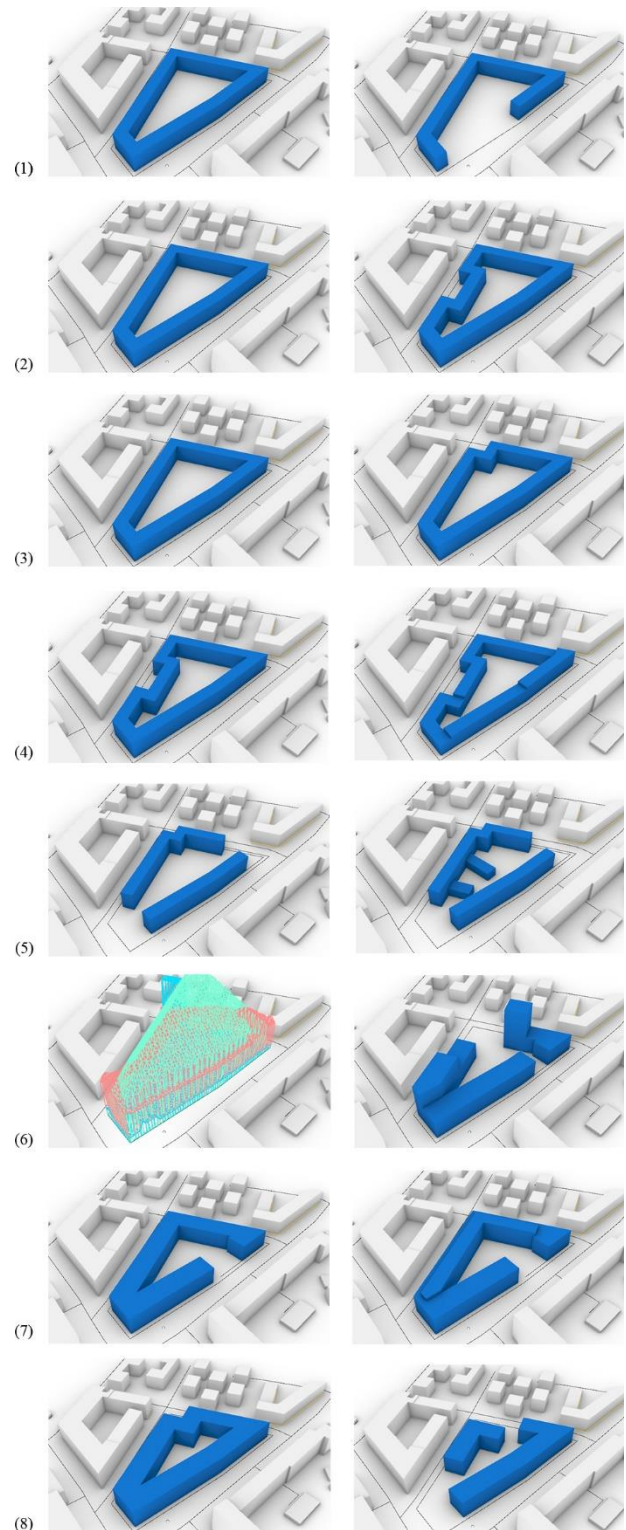
Figure 4. Actions modules, left column: geometry before action, right column: geometry after action. (1) Open block; (2) Setbacks Edges; (3) Setbacks Corners; (4) Reduce Height; (5) Yard buildings; (6) Towers; (7) Rooftop terraces; (8) Random Breaks. The examples of actions application shown above are not restricted by the input form and can be applied to any shape.

FAR = 1.94   FAR = 1.99   FAR = 2.02   FAR = 1.97
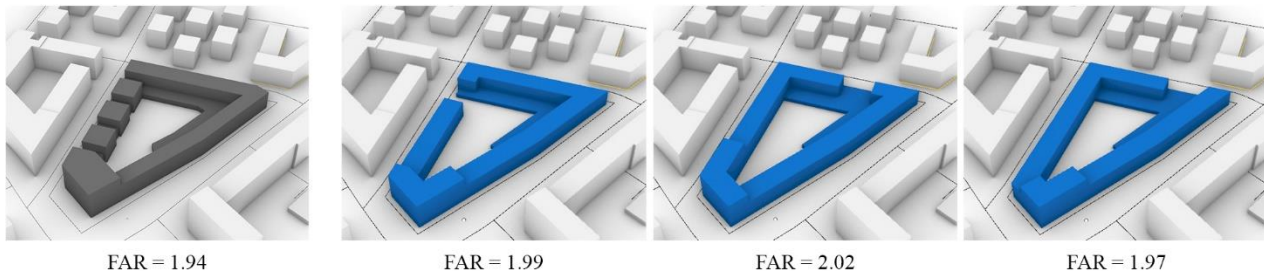
Figure 5. Proposal 1, competition entry [16] and three of the generated alternatives.

FAR = 2.35   FAR = 2.64   FAR = 2.60   FAR = 2.60
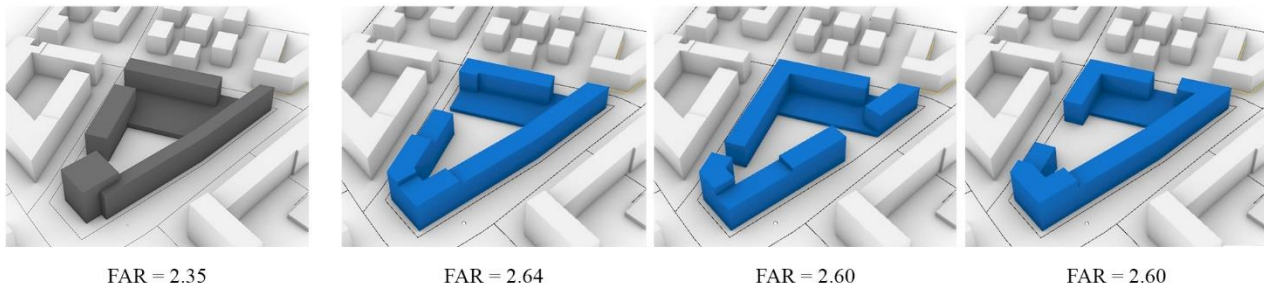
Figure 6. Proposal 2, competition entry [16] and three of the generated alternatives.

FAR = 2.38   FAR = 2.53   FAR = 2.51   FAR = 2.42
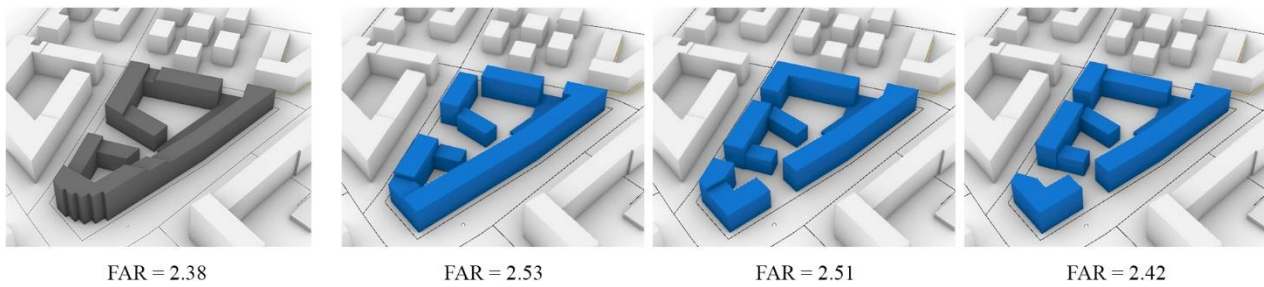
Figure 7. Proposal 3, competition entry [16] and three of the generated alternatives.

Here, the block outline, as well as height distribution was easy to achieve. However, the allocation of the passageways near each other in a sequence did not work (see Figure 5).

The second proposal is the block with multiple broader passageways. The applied sequence is: Basic Block→ Random breaks (delete larger area) → Towers; resulting outputs set: 200 options. In this iteration, it is noticeable that the depth manipulation as an action type is missing (see Figure 6).

The third proposal adds yard buildings to the perimetral block and cuts the southern corner. The sequence of actions: Basic Block → Random Breaks → Yard buildings; resulting outputs set: 200 options. Here as well the precise allocation of the passageways is missing (see Figure 7).

In the second scenario, the goal was to evaluate if the method is efficient in terms of finding multiple diverse solutions in a short time. For this purpose, we used the following sequences of actions (see Figure 8). With the described semi-automatized generation approach, we achieved 10.000 designs in ca. 4 hours. (see Figure 9).
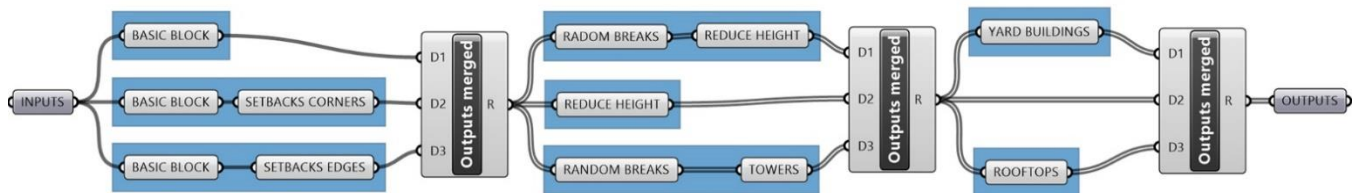
Figure 8. The sequences of actions applied for the second scenario of the test case.

As the goal of this scenario was to generate options as diverse as possible, we adjusted the inputs for each of the actions to ensure that we allow the required diversity. However, it is also possible to use readily combined sequences, where only the context inputs must be adjusted, and thus receive the outputs fully automatized in ca. 15 minutes. As follows, the interactive semi-automated process, where the designer controls the generation process, allows the application of one's experience and conceptual insights for the particular project case. The semi-automated process has the potential to reduce the number of generated designs to a user-friendly subset that does not need to be reduced further. However, by using the fully automated approach, a large number of design variants can be utilized by other participants of the planning process in order to apply additional evaluation criteria that were not included yet. When searching for a good solution among thousands of options, it is helpful to use exploration tools like the Design Space Exploration Framework [7] in order to find solutions with a set of defined properties. However, the goal of the second scenario test was to check if the diversity of the produced options is high enough. There are methods for grouping of similar objects based on pixel representation [7]. However, those methods only consider the two-dimensional space and ignore the third dimension - the height, although the height can have a great impact on the optical building perception. As follows, because of the missing method of distinguishing the visually diverse outputs, we manually selected three subsets of the generation outcomes. To highlight the diversity of the outputs with similar performance, we first filtered three different density groups from the generation population: for the first group the density of the competition entries was chosen (FAR = 2.22),
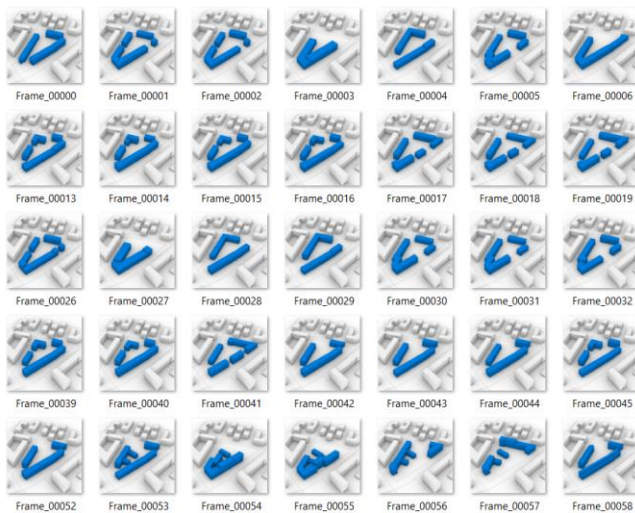
another two groups were with lower density (FAR = 1.70) and high-density (FAR = 3.0). Each filtered subset has a density variation in the boundaries of +-0.05 from the goal value. From each subset (400 options for FAR 2.2, 1200 options for FAR 3.0 and 120 options for FAR 1.7) 5 optically diverse options were selected for demonstration (see Figure 10). The optical distinction between variants, all of which meet the formal criteria of the FAR index, proves that the broad exploration of the possible solutions at the early design stage is essential.

5   DISCUSSION AND OUTLOOK

In this paper, we presented a methodology for the efficient semi-automatic generation of multi-family-residential building designs. The principal goals for the developed methods were a quick generation of design variants (efficiency), exploration of the design space, and extended means of interaction for user influence on the generation process.

We believe that the developed parametric workflow for an automatic generation of design options allows for a more efficient exploration of the design space by avoiding manual form modeling and by a direct connection of the design outputs to the broad scope of spatial analysis and simulations thus showing the performances of the solutions immediately. According to the results of the test case application, we consider the developed framework to deliver a sufficient number of diverse designs (see Figure 10). We as well proved the possibility to consider the formal design aspects during the generation by achieving the outputs that visually resemble the manually designed competition entries.

However, the presented method has several limitations. Because of the nature of the parametric design methods, the design variants are predefined by the type and value-range of the used parameters. There are no surprising solutions as it would be possible by using optimization-based systems with a more advanced data-representation of designs [5]. As well, because of the initial narrowing of the design space to a limited set of possible solutions, better performing variants might be excluded in advance.

To increase the possible design space, we need to introduce the two remaining typologies (slab and solitary) and enable their combinations with the block. We also plan to further advance the options for the user interaction. In order to assure possible integration into an architectural workflow, we as well need to include the next levels of detail, such as the subdivision of the volumes into apartments and the further floorplans elaboration. Those levels of detail are as well necessary in order to conduct precise analysis and simulations for further qualitative design performance evaluation.

Figure 9. A subset of the 10.000 generations for the scenario 2.

Figure 10. Examples of the designs for 3 sub-sets of the generation outputs: (1) FAR = 1.7 +- 0.5; (2) FAR = 2.2 +-0.5; (3) FAR = 3.0 +-0.05.

REFERENCES

1. Abdumawla, A., Bielik, M., Buš, P., Mei-Chih, C., Dennemark, M., Fuchkina, E., Miao, Y., Knecht, K., König, R., and Schneider, S. Decoding Spaces, 2019.

2. a+t research group.Why density? Debunking the myth of the cubic watermelon. a+t architecture publishers, 2015.

3. Buerkin, T., Peterek, M., Petereck, M., Stadtbausteine. Birkhäuser, 2016.

4. Calskan, O. Parametric design in urbanism: A critical reflection. *Planning Practice & Research 32*, 4(2917), 17-443.

5. Dillenburger, B., M. Braach, L. Hovestadt, Building design as an individual compromise between qualities and costs*, CAAD Futures (2009)*, 458 - 471.

6. Finch3d [Webste/Tool]. Retrieved from https://finch3d.com/video-blog/ on 24.09.2019.

7. Fuchkina, E., Schneider, S., Bertel, S., Osintseva, I. (2018, September). Design Space Exploration Framework - A modular approach to flexibly explore large sets of design variants of parametric models within a single environment. *In Proceedings of the 36th eCAADe Conference* - Volume 2 (pp. 367-376).

8. German Building Regulations, retrieved from http://www.bauordnungen.de/html/deutschland.html on 28.09.2019.

9. Kreo [Webste/Tool]. Retrived from https://www.kreo.net/ on 24.09.2019.

10. Knowles, R.L., Berry, R. D., Department of Energy & Solar Energy Research Institute, Solar envelope concepts: moderate density building applications: final report. Solar Energy Information Data Bank, Golden, Colo, 1980.

11. Moudsari, R. S., Pak, M., Adrian Smith + Gordon Gill Architecture. Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. *Proceedings of International Building Performance Simulation Assosiation, Chamberý, France* (2013), 3128-3135.

12. Nagy, D., Villaggi, L., and Baenjamin, D. Generative urban design: Integrating financial and energy goals for automated neighborhood layout. *SimAUD 2018* (pp. 247-254).

13. Openstreetmap [Website]: https://www.openstreetmap.de/, retrieved on 24.09.2019.

14. Parish, Y. I. H., Mueller, P., Procedural modeling of cities, *In Proceedings of ACM SIGGRAPH 2001*, ACM Press, E. Fiume, Ed., (2001), 301-308.

15. Reicher, C. Städtebauliches Entwerfen. Springer, Vieweg (2017), 54 - 81.

16. Results of the residential competition in Munich: https://www.competitionline.com/de/ergebnisse/241367 retrieved on 24.09.2019

17. Robert MCNeel & Associates. Grasshopper, 2014.

18. Robert McNeel & Associates. Rhinoceros 6.0, 2019.

19. Spacemaker [Webste/Tool]. Retrieved from https://spacemaker.ai/ on 24.09.2019.

20. Testfit [Tool]. Retreived from https://blog.testfit.io on 24.09.2019.

21. Wilson, L., Danforth, J., Davila, C. C., and Harvey, D. How to Generate a Thousand Master Plans: A Framework for Computational Urban Design, *SimAUD 2019* (pp. 113-119).

Optimisation of Complex Geometry High-Rise Buildings based on Wind Load Analysis

Erron Estrado[1], Michela Turrin[2], and Peter Eigenraam[2]

[1]Delft University of Technology
Delft, Netherlands
erron_e@hotmail.com

[2]Delft University of Technology
Delft, Netherlands
{m.turrin, p.eigenraam}@tudelft.nl

ABSTRACT

Wind analysis for the structure of buildings is a challenging process. The increasing strength and frequency of wind events due to climate change only add higher demands. In addition, high-rise buildings are growing in number and include many of unconventional shape. Current methods used in practice for calculating structural wind response either do not account for these geometries, such as the Eurocode or are prohibitively time-consuming and expensive, such as physical wind tunnel tests and complex Computational Fluid Dynamics simulations. As such, wind loads are usually only considered towards the end of design. This paper presents the development of a computational method to analyse the effect of wind on the structural behaviour of a 3D building model and optimise the external geometry to reduce those effects at an early design phase. It combines Computational Fluid Dynamics (CFD), Finite Element Analysis (FEA), and an Optimisation algorithm. This allows it to be used in an early design stage for performance-based design exploration in complement to the more traditional late-stage methods outlined above. The method was implemented into a rapid and easy to use computational tool by combining existing plugins in Grasshopper into a single script that can be used in practice on complex shaped parametric high-rise building models. After developing the method and testing the timeliness and precision of the CFD, and FEA portions on case study buildings, the tool was able to output an optimal geometry as well as a database of improved geometric options with their corresponding performance for the wind loading allowing for performance-based decision-making in the early design phase.

Author Keywords
Computational Fluid Dynamics; Optimisation; Finite Element Analysis; Wind engineering; Parametric design; Computational design

ACM Classification Keywords
• Computing methodologies~Simulation evaluation • Applied computing~Architecture (buildings) • Applied computing~Computer-aided design

1. INTRODUCTION

The aim of this research was to develop a computational method to optimise a building's geometry based on the effect of wind on the structure that is sufficiently precise but also quicker than traditional methods such as wind tunnel testing. The method must be able to obtain rapid and dependable results so that it can be used in the early design phase to facilitate design decisions in situations where wind loading is, or should be, of concern. Thus, an array of design options is presented, and adjustments can be made when the detail level of the design is low rather than having to adjust internal structural members when the design is at an advanced stage leading to conflicts between structural, architectural, and mechanical layouts.

Forces due to wind can be critical for a building especially tall, slender towers and with the rise in technology building shapes are becoming increasingly complex. The main methods used in practice for calculating this structural effect of wind are by building codes, the Eurocode in Europe, wind tunnel testing and, to a lesser extent, in-depth CFD simulations but these methods have significant drawbacks.

The Eurocode (EN 1991-1-4:2005) however, only gives calculation methods for regular shapes such as rectangular or circular plan towers up to 200m high. Applying these methods to non-standard geometry buildings produces results that would undoubtedly be inaccurate as it is an approximation. Physical scaled-model wind tunnel testing is the most accurate method and can deal with any given shape. To obtain this accuracy, however, it requires a lot of effort from skilled technicians to properly set up the model, use the wind tunnel, and process the data. This can take a lot of time and expense. Thus, it is usually only done at a late stage for verification purposes.

Computational Fluid Dynamics (CFD) allows a user to simulate wind flow around a model and the effects on the body such as pressure and drag forces regardless of shape. While not so commonly used in professional practice for structural wind engineering it has been widely studied for many years [1-5]. These simulations can take a lot of time and computational power to complete especially when very

detailed meshes and more complicated settings are used to be as accurate as possible. Finite Element Analysis (FEA) software analyses the structural response of buildings and is widely used in practice. It can give very detailed results the level of which, and time needed for simulation depends on the software package and the settings used. Optimisation algorithms allow one to iteratively manipulate input variables until the resulting output value(s) of some algorithm or simulation reaches an optimum. Algorithms used for optimisation include metaheuristics which are widely applicable but require a lot of iterations to converge to an optimum [6], direct search methods which are efficient but not very robust [7], and model-based algorithms which use machine learning methods to give robust performance with few evaluations making it ideal for heavy simulations such as CFD and FEA [8].

The combination of these three technologies gives a Fluid-Structure Interaction based Optimisation (FSIO) method. This is a computational method that simulates wind flow and the resulting pressure on the façade of a building model and translates it to a structural Finite Element Model (FEM) that is analysed with FEA to obtain the resulting structural effects. The combination of CFD and FEA creates a Fluid-Structure Interaction (FSI) method. This is then combined with optimisation where the algorithm manipulates parameters that define the external geometry of a building model that is then analysed by the FSI algorithm and the output, in this case, the mass of the structure needed to resist deflection, is the objective that the optimisation minimizes.

This paper describes the development of this method by combining existing CFD, FEA, and optimisation software into a single computational tool. The testing of the method will focus not on absolute accuracy but on obtaining rapid results with enough precision that it can be used repeatedly and reliably in practice on 3D building models in an early design exploration stage as a complement to the more detailed traditional methods mentioned.

2. BACKGROUND

CFD is an iterative simulation based on the Navier-Stokes equations usually with turbulence model equations. It simulates the flow of a fluid, air in this case, and its impact on an object in its flow. There has been a lot of research into Computational Wind Engineering (CWE), i.e. the application of CFD for wind engineering problems, over the last three decades which is accelerated by the continued advancements in computer technology [2]. However, it is still not widely accepted by many codes as a definitive method for structural wind analysis.

For example, the Eurocode allows for "properly validated numerical methods" as a supplement to its calculation procedure [9] and the Architectural Institute of Japan (AIJ) has published a guide which gives advice on using CFD for wind engineering [4]. The Eurocode calculates wind force (F_w) using Equation 1 where $c_s c_d$ is the structural factor which accounts for the non-simultaneous action of peak wind pressures over the building face and the dynamic response of the structure due to turbulence, c_f is the force coefficient which accounts for the building shape, $q_p(z_e)$ is the peak velocity pressure at reference height z_e. A_{ref} is the reference area on which the force is acting. This equation requires the solution of up to 20 other equations and the reading of many values from tables and charts in the Eurocode which are often only provided for specific standard building shapes.

$$F_w = c_s c_d \cdot c_f \cdot q_p(z_e) \cdot A_{ref}$$

Equation 1: Wind force on a building according to EN1991-1-4:2005

The two most common types of turbulence model are Large Eddy Simulation (LES), which give time-dependent results, and Reynolds-Averaged Navier-Stokes (RANS), which give time-averaged results. Tamura, Nozawa [4] recommend LES for CWE purposes due to its accuracy, however, this is a very computationally expensive procedure. RANS models were used with a gust effect factor for along wind loads which gave results close to physical experiments. Clannachan, Lim [2] after testing various turbulence models on tall building models conclude that RANS is the most practical for typical structural engineering due its efficiency. While they conclude that CFD is not yet able to replace physical testing it has great potential complementary tool in early-stage structural design to assess a broad array of design alternatives.

CFD based optimisation, while applied extensively in aerospace, is rare in building design [10]. This, however, would have great benefits over typical trial and error approaches for finding optimal shapes. However, optimisation only compounds the computational expense of CFD thus, it only practical if a CFD evaluation takes at most a few hours [11]. Bernardini, Spence [11] used surrogate-model based optimisation along with 2D CFD and to find the optimum cross-section of a high-rise building that minimises lift force and drag coefficient with minimal optimisation evaluations. Chronis, Turner [12] used Fast Fluid Dynamics, a lower order CFD method that is faster but much less accurate, combined with evolutionary optimisation to
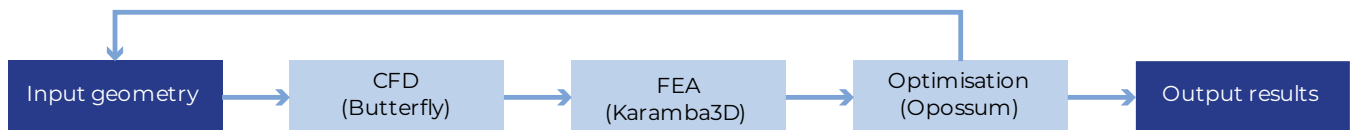
Figure 1: FSIO method flowchart

optimise the shape of a free-form surface based on surface pressure.

There are some shortcomings in CWE techniques compared to traditional wind tunnel tests particularly in modelling the complex airflows induced by buildings [2]. Nonetheless, the benefits of CWE over time-consuming and expensive physical tests continue to inspire more research in the field.

3. RESEARCH METHODOLOGY

The research methodology is divided into four parts: development, CFD validation and sensitivity analysis, FSI validation, and optimisation testing.

These tests were done using various case study buildings. The Commonwealth Advisory Aeronautical Research Council (CAARC) Standard Tall Building Model was used for CFD tests as there are results from physical wind tunnel tests to compare to [13]. In order to verify the method's effectiveness for non-standard geometry buildings, three such models of existing buildings were also used. These are the Absolute World Towers by MAD Architects, Jiangxi Nanchang Greenland Central Plaza by SOM, and Ardmore Residence by UN Studio. Each possesses unique geometric features that test the robustness of the tool and method.

3.1. Development

In this study, the proposed FSIO method (Figure 1) was implemented in Grasshopper, the visual scripting plugin for Rhinoceros3D. It involved the combination of 3 plugins: Butterfly, Karamba3D, and Opossum, with custom scripting in order to produce an easy to use tool where a parametrically defined building model can be input to obtain a collection of variations of that model based on structural performance under wind load.

Butterfly is a Grasshopper plugin and python library developed by Mostapha Sadeghipour Roudsari as part of the Ladybug Tools suite [14]. It allows users to run OpenFOAM CFD simulations from within the Grasshopper environment for cases pertaining to building design such as outdoor airflow. OpenFOAM is a free open source CFD software that has been widely used and validated [15]. Karamba3D is a parametric FEA plugin for Grasshopper developed by Clemens Preisinger in cooperation with Bollinger und Grohmann ZT GmbH. [16]. Opossum, developed by Thomas Wortmann [17] is a model-based optimisation algorithm. It uses Radial Basis Functions (RBF) to generate a response surface thus reducing runtime particularly for heavy simulations [8]. These plugins were selected based on their wide use, previous validation, quick performance, and ease-of-use.

The Butterfly portion was set up according to an outdoor airflow simulation. Modifications were made for the automatic setting of mesh cell size and size of the domain based on the building dimensions. Probe points for sampling pressure on the façade are also created automatically. An algorithm was written to translate the resulting static mean pressure values on a mesh of the building facade to vectors

for point loads and moments which Karamba applies at 6 points to a fixed upright beam formed at the centreline of the building using the dimensions of its structural core creating the FEM. The analysis outputs results for deflection, mass of the structural model, forces, etc. which can be used as the objective for Opossum. The input variables are the numerical sliders controlling the shape of the model being analysed.
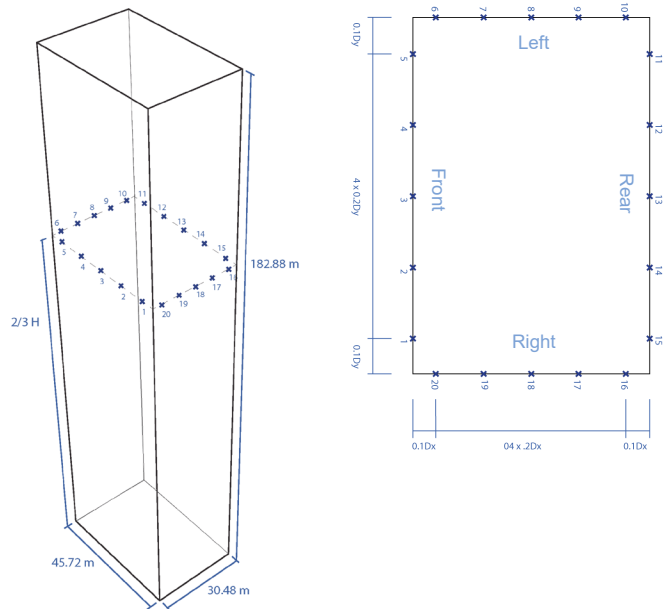
Figure 2: CAARC model CFD setup
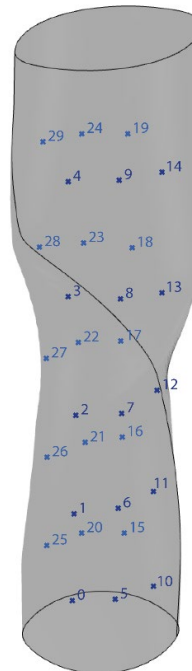
Figure 3: Absolute Tower model CFD setup

221

3.2.   CFD Validation and Sensitivity

A validation study was performed to ascertain the ability of Butterfly to give results within a degree of uncertainty to physical tests. An absolute deviation of ±20% is deemed acceptable for an early stage method as this [18]. For the comparison, the CAARC Standard Tall Building Model was used for which physical wind tunnel test results were obtained from Meng, He [18]. The digital model was set up just as the physical model where pressure coefficient (C_p) was measured at 20 points around the façade at 2/3 of the height and compared to those of the physical wind tunnel tests. The CFD domain size was 900 x 600 x 400 m (Figure 2) and the number of CFD iterations was set at 30000.

Three RANS turbulence models were tested: standard $k - \varepsilon$, Realizable $k - \varepsilon$ [19], and RNG $k - \varepsilon$ [20]. The $k - \varepsilon$ models solve equations for the turbulent viscosity of the flowing air by calculating the kinetic energy, k, and the turbulent dissipation rate, ε [2]. Other settings were kept constant. The resulting C_p obtained was compared between each of the turbulence models and the wind tunnel results at each pressure tap but also the overall shape of the graph was examined to determine its precision in capturing results at each area of the building. Lastly, these results were considered along with the time taken for the simulation to determine how efficient it is.

After a suitably accurate turbulence model was chosen from the validation a sensitivity analysis was performed. This involves changing one parameter while keeping others constant to see how much the simulation time can be reduced while maintaining sufficient precision and accuracy. In this case, the number of CFD iterations and mesh cell size were tested. To evaluate how much the number of iterations could be reduced and its effect on simulation time, simulations were done with 30000, 10000, and 5000 iterations using the chosen turbulence model. All other settings remained constant. The C_p was compared for each simulation to determine deviation. To test the effect of mesh cell size on precision and timeliness, simulations were done reducing cell sizes starting from a value equal to the length of the shortest side of the building divided by 10 [21] then successively reduced by a chosen value of $\sqrt{2}$ based on common practice [22]. This gives resolutions of Coarse, Medium, Fine, Super Fine and XXFine as defined in Table 1.

The building model used was the Absolute Tower. To check precision C_p was measured at 30 points around the façade to capture the effect of its complex geometry (Figure 3). The domain has the following dimensions: windward = 3H, Leeward = 10H, sides = 2.3H, and top = 2.3H where H is the height of the building [21]. Wind speed was set at 30 m/s corresponding to a violent storm on the Beaufort scale and terrain category was chosen for an urban site (roughness length = 1m).

3.3.   FSI

This test sought to compare the values obtained from the Eurocode (EN) method to those obtained from the FSIO method implemented in the script. The Absolute Tower and Nanchang Tower models were used in this test. The wind force (F_w) at 6 heights along the building was calculated using the EN method and the FSIO method and compared. As the EN only gives guidance for regular shapes it was chosen to assume values for a circular cross-section for this calculation as these building models have a smooth cross-sectional shape. Basic wind velocity was taken as 30m/s, roughness length of 1m and A_{ref} was set at 1m². The areas to which the F_w would be applied was obtained in Grasshopper by finding the areas around the point loads of the building perpendicular to the wind flow. This was done for each of the tested case study buildings and the loads compared to those from the FSIO method.

3.4.   Optimisation

The aim of these tests was to determine the extent to which structural performance due to wind can be optimised by making relatively small changes to the geometry of the building. The Ardmore Residence model was used for this study. The model was set up parametrically so that its geometry could be manipulated by the optimisation algorithm. The floor plan shape was built as a curve to be modified by three parameters. Two design variables control the position of each of the wings along the main body. A third slider modifies the edges of these wings from straight to a more angled position. From this curve outline, the massing was extruded to the 136m height of the building. In addition, the building was rotated 45 degrees so the wind would impinge on the building off-axis (Figure 4).

As CFD is very computationally expensive simulation settings were chosen in order to minimize the time for a single iteration. Thus, a coarse mesh setting was used and max iterations for Butterfly was set to 2000.

The output objective was the mass of structural material needed to resist the wind force. This was determined by modelling the core of the building as a hollow concrete (C25/30) beam approximately the dimensions of the core of the existing building.
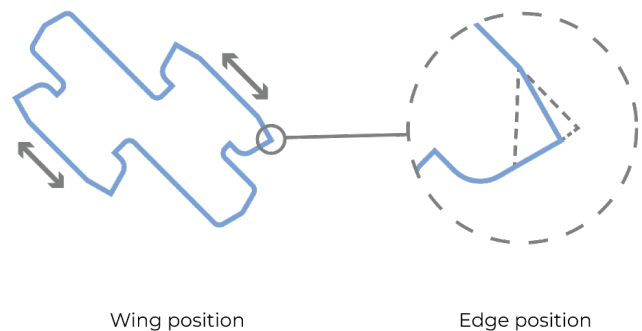
Figure 4: Ardmore Residence model parametric variables

The Optimize Cross-section component of Karamba was used to automatically select from a list of cross-sections for the core varying in thickness (length and width kept constant) that satisfies the deflection limit which was set at 0.66m based on an initial FSI simulation. The core of the building measured 11m by 7m and 0.4m thick. The lists of cross-sections range from 0.10m to 0.59m thick in 0.01m intervals.

In Opossum the RBFOpt optimisation algorithm was used and the max number of iterations was set at 100. The geometry created at each optimisation iteration and its corresponding objective value was recorded in Grasshopper to create a database of design options.

4. RESULTS

4.1. CFD Validation and sensitivity

The RNG $k - \varepsilon$ turbulence model showed the closest Cp values to the physical wind tunnel results (WT) from Meng, He [18] followed by standard $k - \varepsilon$ (Figure 5). The Realizable $k - \varepsilon$ model gave unrealistic Cp values the reason for which was not found and thus was omitted.

From Figure 5 the Cp values of all models are quite close at the front face of the building (pressure tap 1 - 5). On the sides and rear (6 – 20) the values deviate more. RNG $k - \varepsilon$, while with an absolute deviation of 25%, deviates much less from the wind tunnel values compared to standard $k - \varepsilon$ which has a more rounded graph shape. This is likely due to the poor performance of standard $k - \varepsilon$ in predicting regions of flow separation like the building rear and edges [4].

Timewise, RNG $k - \varepsilon$ took 42.6 hours to complete, standard $k - \varepsilon$ took 41.7 hours and realizable KE, 37.4 hours. RNG $k - \varepsilon$, therefore, proved to be the best choice due to its level of precision in capturing the flows at different areas of the building while taking not much more time than other models.

The timeliness of the RNG $k - \varepsilon$ model was further evaluated in the sensitivity analysis. At 10000 and 5000 CFD iterations, the Cp values obtained were nearly identical to that of the 30000-iteration simulation done previously while taking only 15.7 hours and 6.95 hours respectively to complete the simulations. From plotting a graph of the residuals, i.e. the scaled errors of calculated values between successive iterations, the values are sufficiently below the accepted threshold [21] at all tested numbers of iterations used in the validation study (Figure 6). Table 1 shows the settings used, the time taken, and absolute deviation at front and rear of the building from MAD_5, which is assumed to be the most accurate, for each run at differing mesh sizes. XXFine took the most time with 20.3 hours. The coarse mesh size has the lowest time of 5.7 hours. Absolute deviation increases as cell size decreases with the exception of the front at Medium size which is highest. At Fine resolution, the time is greatly reduced while deviation is still around 10% making it an ideal choice to balance time and accuracy for an early-stage study.
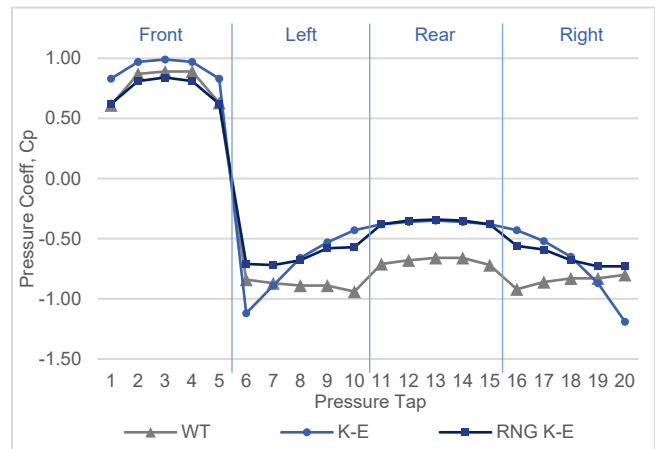
Figure 5: Graph of C_p at each pressure tap from each turbulence model and wind tunnel (WT) for CAARC model
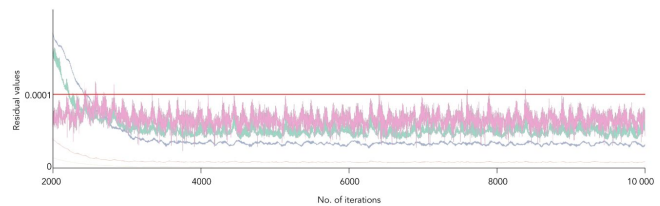
Figure 6: Residuals graph of RNG $k - \varepsilon$ simulation

Test	MAD_1	MAD_2	MAD_3	MAD_4	MAD_5
Resolution	Coarse	Medium	Fine	Super Fine	XXFine
Cell size (m)	4.18	2.96	2.09	1.48	1.08
no. of cells	176545	236050	346647	525640	732422
Time	5.7h	6.3h	8.6h	14.5h	20.3h
Iterations	10000	10000	10000	10000	10000
Dev. Front	18.46%	28.08%	8.71%	0.77%	N/A
Dev. Rear	17.63%	15.36%	10.19%	1.46%	N/A

Table 1: CFD cell size sensitivity analysis results
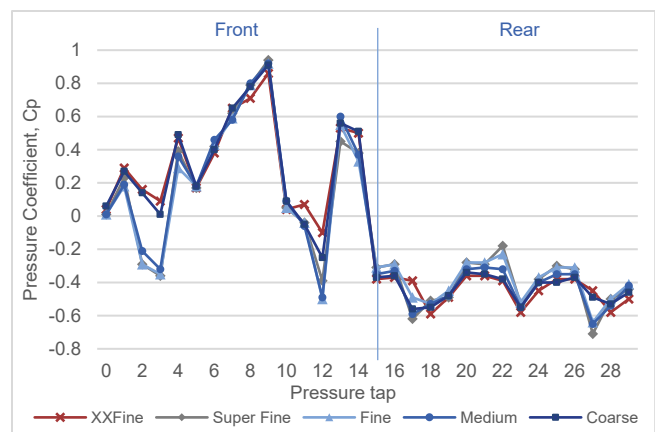
Figure 7: CFD mesh cell size sensitivity analysis
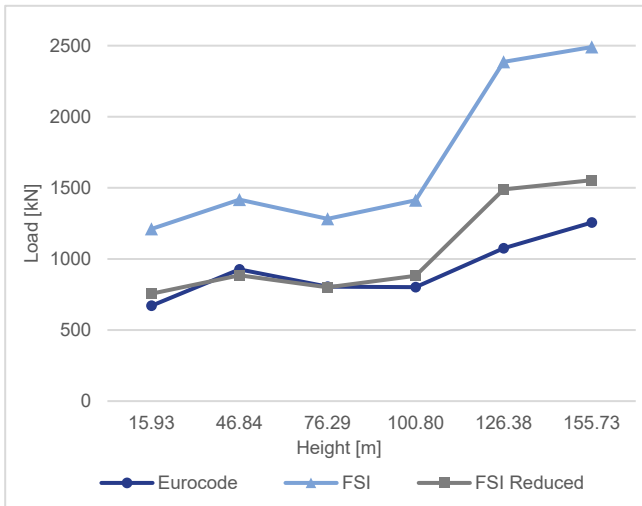
Figure 8: Absolute Tower EN/FSI comparison
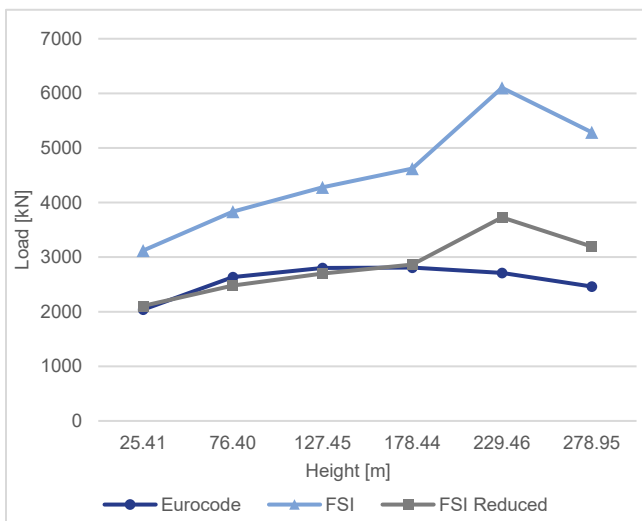
Figure 9: Nanchang Tower EN/FSI comparison
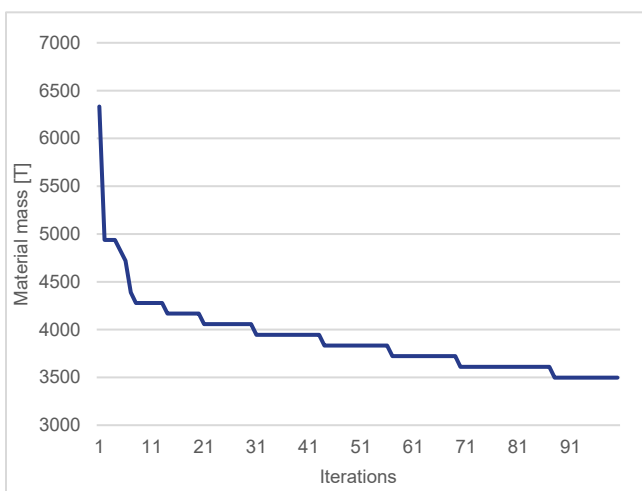
Figure 10: Mass of structural material per iteration

From Figure 7, while there is some deviation the C_p values generally follow the same trend. The exception being around the edges of the building near the twist (pressure tap 1 - 4 and 10 - 12) which can be reasonably assumed to be caused by issues in generating the mesh at those points for Medium and Fine resolution. The result shows that between mesh sizes precision can be maintained. Thus, at an early stage or for optimisation purposes a coarser, faster setting can be used to obtain quick and repeatable results that can reliably inform early-stage decision making. A finer resolution is suitable for a later stage as a final check.

4.2.   FSI Validation
The results for the Absolute Tower model are shown in Figure 8. The EN numbers begin to rise then fall with respect to the decrease in the perpendicular area near the middle of the tower then rise again to a maximum value of 1256kN. The values from the FSI method follow a similar pattern but are much higher as CFD with RANS turbulence models like the RNG $k - \varepsilon$ simulate the mean static pressures across the entire surface. In reality, wind flow in the atmospheric boundary layer is more random and peak pressures do not occur simultaneously over a structure [23]. To account for this the FSI values were multiplied by the structural factor, $c_s c_d$, and force coefficient, c_f, from the Eurocode (FSI Reduced) [9].

As seen in Figure 8, the reduced values from the FSI simulation are now closer in line with those from EN calculations. The discrepancy in values above 100m is accounted for by the unique geometry of the building. At the top of the tower, FSI values are higher than EN as this point has the broad side of the elliptical plan thus causing more pressure than the EN (which assumes a circular plan) has calculated. This is more pronounced since wind speed is highest near the top.

For the Nanchang Tower model, the EN values follow a smooth curve. For the FSI values, the values smoothly increase until a height of 178.44m then jump at 229.46m. This is likely due to the high wind speed and concave façade at this point which leads to a higher pressure as the air would have difficulty flowing around the building at this point. The value then drops back down at the highest point where the wind can then flow over the top of the building (Figure 9).

4.3.   Optimisation
Figure 10 shows the gradual reduction of the objective over each optimisation iteration. A 24% reduction in the material mass was achieved by reducing the building's structural core thickness from 0.40m to 0.30m after 100 iterations. The model-based approach of Opossum seems to have sufficiently converged after the relatively low number of 100 iterations but perhaps could have run longer to confirm since the final line of the graph is quite short.

Figure 11 shows the original vs the resulting optimum geometry, surface pressure from CFD, and FEA results after completion of the optimisation test using Opossum. The

224
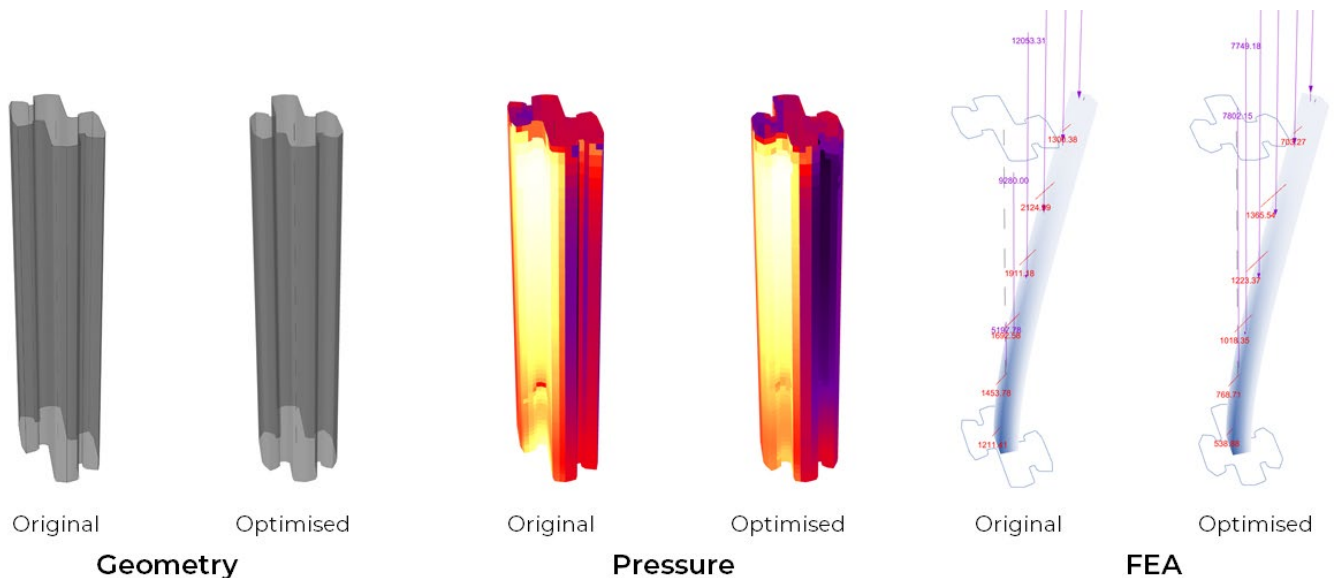
Figure 11: Original and optimised geometry, CFD, and FEA results of Ardmore Residence model

Figure 12: Excerpt of the resulting geometries and the corresponding objective values (mass)

lower wing is moved to the front resulting in a more symmetrical cross-section. The edges were also pulled to a sharper angle. This reduces the flat pressure inducing wall area on the windward of the building and allows the wind the flow more smoothly around the building. The reduced pressure results in lower forces on the core of the building requiring less thickness to resist deflection. The FEA results were also checked with hand calculations. r

This optimal shape is quite different from the original. Using the data recorded from Grasshopper a database of results was created where each resulting shape of the building and its corresponding structural mass calculated by the FSIO method was displayed (Figure 12). This allows a user to make decisions based on performance and architectural factors.

5. CONCLUSION

This paper outlined the development and testing of a Fluid-Structure Interaction based Optimisation (FSIO) method used to optimise the geometry of complex-shaped high-rise

building models based on wind structural response in the early design phase. The aim was to create a computational method to aid in performance-based design exploration as a complement to more traditional methods of structural wind analysis used in the late phases of building design.

The first part involving the validation and sensitivity analysis of the CFD simulations showed that Butterfly using the RNG $k - \varepsilon$ turbulence model was the best performing choice. The validation showed it was able to capture variations in pressure around different areas of the building which is essential for complex geometries in close accordance to physical wind tunnel results. The sensitivity analysis showed that simulation time can be greatly reduced, from 42.6 hours initially to 5.7 hours at most by doing fewer CFD iterations, which had little to no impact on the accuracy, and by using a coarser mesh which has a greater impact on accuracy with about 18% deviation for the coarsest mesh. It follows a similar trend in values over the building area and thus can reliably show relative improvement between simulations of different geometries if CFD settings remain constant.

The second test involved comparing the results from the developed FSI method (the combination of CFD and FEA) to results obtained by Eurocode (EN) calculations. It showed that the FSI results offer near exact values with the EN procedure when combined the structural factor and force coefficient values. The differences in values can be concluded to occur due to the unique characteristics of the geometry of the building showing the FSIO method's applicability to non-standard shapes while the EN assumes a standard shape. The values obtained could possibly be lower if a finer mesh and higher number of iterations for the CFD simulation was used. These were kept low in the tests to save time.

In the optimisation test, the algorithm reduced the objective, mass of structural material, by 24% solely by manipulating the building geometry to reduce wind pressure. However, the optimal geometry was quite different from the original. The database produced contained the geometry made at each optimisation iteration with its corresponding objective value. This can be used to make better decisions for geometry in situations where wind is critical. While the complete FSIO run took 2 days to compute it is believed to still be more efficient and valuable than Eurocode calculations, wind tunnel tests, and detailed CFD simulations in the early stage due to the large number of options it can produce directly by a designer or engineer at no additional cost.

Overall, the method, particularly in the form of a Grasshopper-based tool, is best suited in the form-finding stage of design for complex towers in high wind situations. It can be used to further optimise an initially chosen geometry by making small geometric adjustments to increase wind performance while not completely deviating from the desired form. A final more thorough CFD or wind tunnel test should be done later to verify safety and conformance with code.

Further work should look at increasing the accuracy and speed of the method. Additional CFD software, turbulence models, and meshing procedures could be tested. In addition, the testing of multiple wind directions can have a large impact on the results. The choice of optimisation algorithm in this project was based on research however, tests on different algorithms could be done to find the best performing option for this specific use case.

ACKNOWLEDGEMENTS

REFERENCES

1. Colombo, E., A. Grassano, and F. Perotti, *Comparison of numerical and experimental simulations used to investigate the wind-structure interaction of a skyscraper.* 2006.
2. Clannachan, G., et al., *Practical Application of CFD for Wind Loading on Tall Buildings.* 2009.
3. Fransos, D. and A. Lo Giudice, *On the use of computational simulation in the determination of wind loads on structures: design experiences and food for thought.* 2015.
4. Tamura, T., K. Nozawa, and K. Kondo, *AIJ guide for numerical prediction of wind loads on buildings.* Journal of Wind Engineering and Industrial Aerodynamics, 2008. 96(10): p. 1974-1984.
5. Chang, D. *Aerodynamic Performance Driven Form-Generation for Skyscraper Design.* 2013. Berlin, Heidelberg: Springer Berlin Heidelberg.
6. Wortmann, T. and G. Nannicini, *Black-box optimisation methods for architectural design.* 2016.
7. Evins, R., *A review of computational optimisation methods applied to sustainable building design.* Renewable and Sustainable Energy Reviews, 2013. 22: p. 230-245.
8. Wortmann, T., et al., *Advantages of surrogate models for architectural design optimization.* Vol. 29. 2015. 471-481.
9. NEN, N.S.I., *Eurocode 1: Actions on structures - Part 1-4: General actions - Wind actions*, in *NEN-EN1994-1-4:2005.* 2005, NEN.
10. Thévenin, D. and G.b. Janiga, *Optimization and computational fluid dynamics.* 2008, Springer Verlag: Berlin.
11. Bernardini, E., et al., *Aerodynamic shape optimization of civil structures: A CFD-enabled Kriging-based approach.* Journal of Wind Engineering and Industrial Aerodynamics, 2015. 144: p. 154-164.
12. Chronis, A., A. Turner, and M. Tsigkari, *Generative fluid dynamics: integration of fast fluid dynamics and genetic algorithms for wind loading optimization of a free form surface*, in Proc. SimAUD 2011 p. 29-36.
13. Melbourne, W.H., *Comparison of measurements on the CAARC standard tall building model in simulated model wind flows.* Journal of Wind Engineering and Industrial Aerodynamics, 1980. 6(1): p. 73-88.
14. Sadeghipour Roudsari, M. and M. Pak, *Ladybug: A parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design.* Proc. of BS 2013 p. 3128-3135.
15. *OpenFOAM.* 2019 [cited 2019; Available from: https://www.openfoam.com/.
16. Preisinger, C., *Linking structure and parametric geometry.* Architectural Design, 2013. 83(2): p. 110-113.
17. Wortmann, T., *Opossum: Introducing and Evaluating a Model-based Optimization Tool for Grasshopper.* 2017.
18. Meng, F.-Q., et al., *Sensitivity analysis of wind pressure coefficients on CAARC standard tall buildings in CFD simulations.* Journal of Building Engineering, 2018. 16: p. 146-158.
19. Shih, T.H., et al., *A New k-(Eddy Viscosity Model for High Reynolds Number Turbulent Flows - Model Development and Validation.* Vol. 24. 1994.
20. Yakhot, V., et al., *Development of turbulence models for shear flows by a double expansion technique.* 1992. 4(7): p. 1510-1520.
21. Franke, J., et al., *Best Practice Guideline for the CFD Simulation of Flows in the Urban Environment.* Vol. 732. 2007. 51.
22. Roache, P.J., K.N. Ghia, and F.M. White, *Editorial Policy Statement on the Control of Numerical Accuracy.* Journal of Fluids Engineering, 1986. 108(1): p. 2-2.
23. Cook, N., *Designers' Guide to EN 1991-1-4 Eurocode 1 - Actions on Structures, General Actions, Part 1-4: Wind Actions.* 2007: ICE Publishing.

Rural Urban Transformation: Parametric Approach on Metabolism-Based Planning Strategies in Ethiopia

Diellza Elshani[1,3], Ekaterina Vititneva[1,3], Artem Gilmanov[1,3],

Reinhard König[1,2], Martin Dennemark[1], Sven Schneider[3], Philippe Schmidt[3] and

Abdulmalik Abdulmawla[3]

[1]Bauhaus-University Weimar
Professorship for
Computational Architecture,
Weimar, Germany
diellza.elshani@uni-weimar.de

[2]AIT Austrian Institute of
Technology,
Smart and Resilient Cities,
Vienna, Austria
reinhard.koenig@ait.ac.at

[3]Bauhaus-University Weimar
Chair of Computer Science in
Architecture, Weimar, Germany
sven.schneider@uni-weimar.de
artem.gilmanov@uni-weimar.de
ekaterina.vititneva@uni-weimar.de

ABSTRACT

The emerging countries are confronted by an unprecedented wave of urbanization, which seeks rapid urban planning and sustainable development. A typical example is Ethiopia where the need for economic growth, food security, water, energy and social infrastructure is being fostered. To achieve this Ethiopian government plans to develop a large number of small cities. A computational approach can help to tackle these challenges, by providing several designs in a short manner of time and evaluate their performance.

In this paper we investigate an approach, as part of a big study and tests, to the growth and transformation of an Ethiopian rural zone into an urban one by using generative urban design strategies.

Taking into consideration the materials flow in a system, the crucial part of the transformation is the urban metabolism, which tackles the interactions of natural and human systems. The developed adaptable model tend to transform the rural zones into attractive, interactive and healthy towns that are self-sufficient regarding food, water and energy.

Author Keywords

Generative design; parametric design; urban modelling; urban growth simulation; rural-urban transformation; urban metabolism; self-sufficient town;

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING

1 INTRODUCTION

The rapid urbanization of the last decades brought many people into cities so the need for a more sustainable development world-wide is becoming higher. There is no exception in Ethiopia where this transformation from a mainly agricultural society to industrialization, is bringing changes into the rural and urban spaces.

The emerging countries, including Ethiopia, are confronted by an unprecedented wave of urbanization, which seeks rapid urban planning, and this need for rapid planning becomes even bigger due to the lack of an adequate number of urban planners there. (Dennemark, Schneider, Koenig, Abdulmawla, & Donath, 2017). Therefore, the computational design can tackle easily these challenges, by providing several designs in a short manner of time.

In Ethiopia's current "Growth and Transformation Plan II" in order to foster economic growth and ensure the provision of water, food, energy and social infrastructure, the government plans to develop a large number of small cities.

At the same time, the United Nations are fostering a more sustainable development with the Millennium Development Goals. (National Planning Commission, 2016) It is a big challenge to achieve these goals when it comes to developing countries.

Therefore, in this paper we investigate different approaches to the growth and transformation of an Ethiopian rural zone into an urban one by using generative urban design strategies.

A crucial part of this transformation is the urban metabolism, which tackles the interactions of natural and human systems. So, the core of the city growth

model is the consideration of the flow of resources and the people within the urban system. The developed adaptable models tend to transform the rural zones into attractive, interactive and healthy towns that are self-sufficient regarding food, water and energy.

2 METHODOLOGY

The adaptable city model that is discussed in the paper is developed in Rhinoceros 3D and the visual programming language and environment that runs within the Rhinoceros, Grasshoppers.

The Parametric Approach on metabolism-based planning strategies in Ethiopia is used to understand the relation and interaction of the material flow, space demand, etc. The design process is based on the iterative process, where the designed algorithms make possible the combination of digital computation and the designer's input. The designed algorithm is tested in three different scenarios where urban growth happens gradually again in an iterative process. In addition, the tools and methods for spatial analysis where implemented, as well as the simulation of possible development scenarios

3 IMPLEMENTATION EXAMPLE:

3.1 Water Accessible City *"Drink, grow, produce and play.*
How to enrich water access to growing city in the Wurer area, Ethiopian."

The priority of the project is to secure water supply supporting independence in food and energy provision. Therefore, a feasible water network is proposed to offer clean drinking, domestic, agricultural and industrial water supply.

Although the Ethiopian population has rapidly grown over the past decade (World Bank Group, 2015) and the urbanization rate is increasing (Gebre-ab, 2015, p. 24), the country's economic and spatial transformation has just begun. As a result "61 million Ethiopians lack access to safe water and 65 million lack access to improved sanitation." (https://water.org/our-impact/ethiopia/, n.d.)

"In rural Ethiopia, a Water.org survey found that many women and children walk more than three hours to collect water…. Recurring droughts result in famine, food shortages, and water-related diseases, as people are forced to rely heavily on contaminated or stagnant water sources." (https://water.org/our-impact/ethiopia/, n.d.) In order to design an accessible water city, the Sustainable Development Goal 6 will be addressed as the new development-driver: Ensure access to water and sanitation for all.
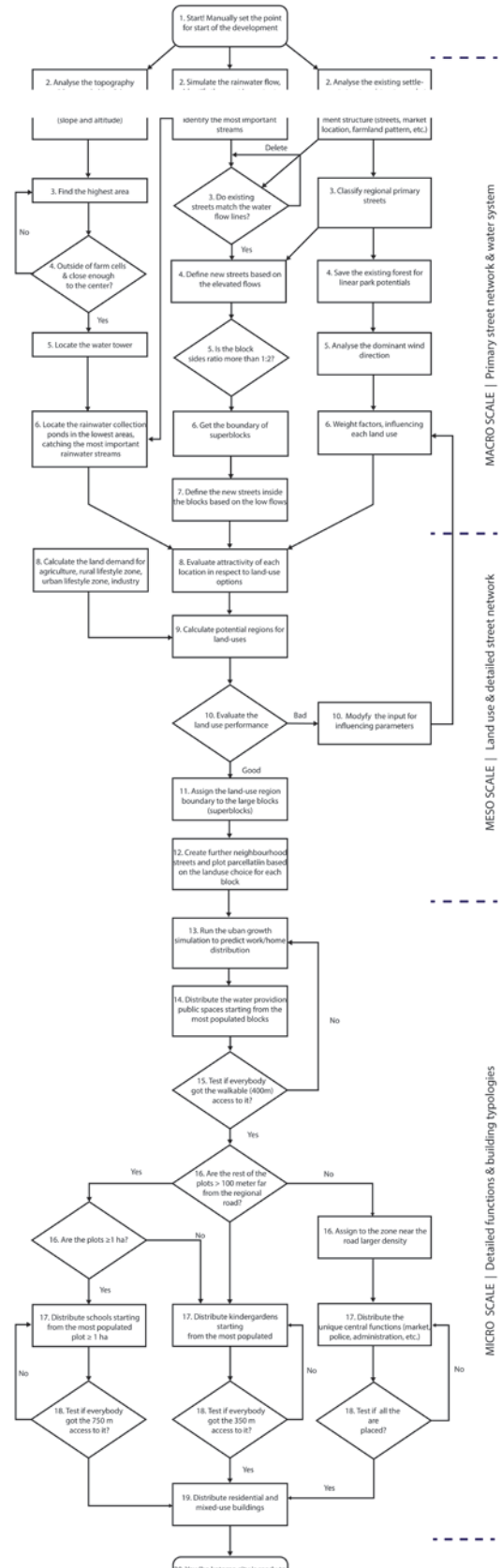
Figure 1. Algorithm of the model

Leading project questions.

1. How to build water infrastructure to offer different water uses, including clean drinking, domestic, irrigational and industrial water provision?
2. How to guarantee sufficient water sanitation (toilet flush and shower)?
3. How will this water system influence spatial planning?

Key water supply principles:

1. The rainwater from public spaces (streets and squares) collected via the water channels, following high- and low-altitude streets. This water collected in the ponds and located in the lowest points. After that, the water will be re-used to irrigate agricultural fields.
2. The existing forest is kept not only as a linear park but also allowing rainwater drainage.
3. The primary source of drinking water is well, supported by a water tower, situated on the high altitude. After treatment, the clean water is transferred through the pipe system to the network of public spaces.
4. Each plot has a public space. The public spaces are designed to provide public access to tap water, shower, and toilet. Also, the biomass and wastewater will be treated in the small biogas units, providing energy supply for the neighbourhood.

Master plan strategy:

The parametric model consists of the following parts:
1. Sets the point for the start of the development.
2. Run the analysis stages, including the slope and altitude of topography, the rainwater flow, and the existing settlement structure (streets, market location, farmland pattern).
3. Locate the tower for the drinking water supply on the highest position outside of the existing farm cells and close enough to the city center. Place rainwater collection ponds for agriculture in the lowest areas, and catch the most important rainwater streams.
4. For the street network, firstly identify existing streets of regional importance. From other streets, which correspond to the rainwater flows are left. They become the base for the rainwater collection system in the city. Following that, define the new streets on the high places, that are found to be dry during the rainwater analysis.

5. Distribute the land use, based on the score which considers the proximity of each location to important factors, like flat space close to the water ponds, market and street network. Afterward, the land demand for each land use is calculated and the potential boundaries are defined.

Figure 2. Street network after analysis.
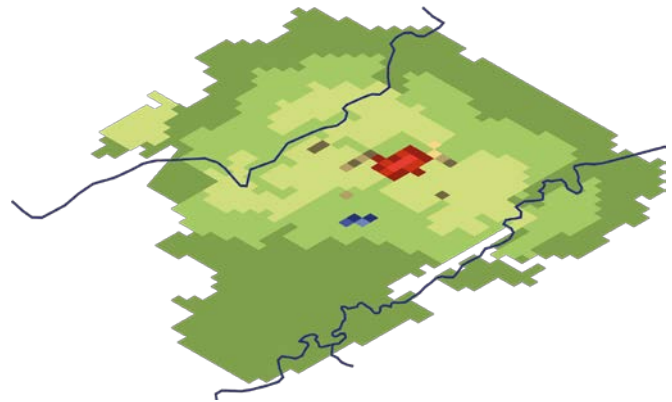
Figure 3. Land-use scenario.**Integration**

After the land use distribution, further steps include plots, building typologies, public facilities, and water systems integration.

1. Run the urban growth simulation to predict work and home distribution. Then the public spaces are allocated, starting from the most populated blocks, with 400 m access to it (Fig. 6).
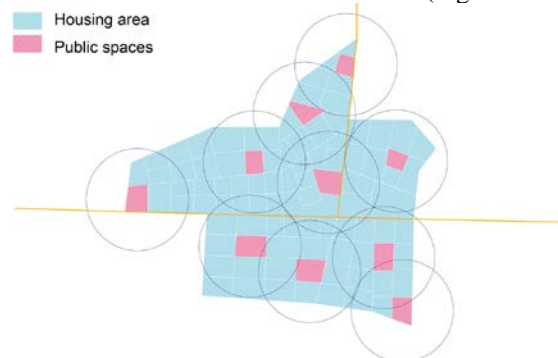
Figure 4. Public spaces distribution.

2. Schools and kindergartens are distributed with the same principle, 100 meters away from regional roads, plots larger than 1ha, with an access radius 750m and 350m respectively. (Fig. 7)
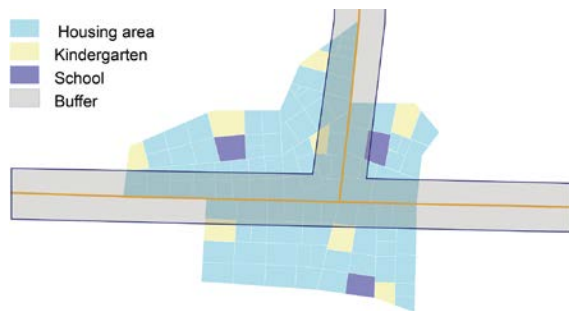
Figure 5. Distribution of school and kindergartens.

3. The market and the church are left in the former places. The city administration and hospital are located along with the most integrated street segments in the space syntax analysis (Fig. 7).

Figure 6. Location of the central functions.

4   CONCLUSION

This model could be used in any plateau region of Ethiopia. An effective and reasonable water layout can reduce the engineering cost, technical requirements and promote the effective use of water network systems. This provides a strong basis for the sustainable development of the urban environment. The access of water systems to each community provides not only effective and safe water but also a leisure space, allowing to gather people in the public space. It helps to create a more communicative environment and a friendlier neighborhood atmosphere. The water usage can also promote communication between people, animals, and the environment. In the context of the water crisis in Ethiopia, water should be considered as one of the main agendas of urban planning. Combining the specific engineering techniques of water infrastructure with the local environment and current development conditions might be the direction for further research.

In the other hand, it shows the possibilities and the advantages of the computational approach in designing and planning rural-urban transformation. Especially when it comes to urban metabolism where the systems of materials interact with each other and the parameters value matter. However, this is only one way of approaching the problem, there are many other studies done in the topic, and it worth to mention that the project is just a piece of work of a big study based on rural urban transformation in Ethiopian context.

REFERENCES
1.   Amare, D., Wondie, M., Mekuria, W., & Darr, D. (2018). Agroforestry of Smallholder Farmers in Ethiopia: Practices and Benefits. Small-scale Forestry.

2.   Dennemark, M., Schneider, S., Koenig, R., Abdulmawla, A., & Donath, D. (2017). Towards a modular design strategy for urban masterplanning - Experiences from a parametric urban design studio on emerging cities in Ethiopia. *eCAADe 2017.* Rome: Conference: eCAADe 2017.

3.   Gebre-ab, N. (2015). Unlocking the power of Ethiopia´s cities. Addis Ababa: New Climate Economy (NCE); Ethiopian Development Research Institute (EDRI); Global Green Growth Institute (GGGI).

4.   National Planning Commission. (2016). Growth and Transformation Plan II (GTP II). Volume I: Main Text Addis Ababa. National Planning Commission.

5.   Schweiz, E. (2019). Worum geht es bei der 2000-Watt-Gesellschaft? local-energy.swiss/programme/2000-watt-gesellschaft.

6.   toolbox.decodingspaces.net. (a.d.). DeCoding Spaces.

7.   Walsh, N. P. (2017). 6 Reasons Cities Are Located Where They Are.

8.   World Bank Group. (2015). Ethiopia urbanization review- urban institutions for a middle-income Ethiopia. Washington, DC: Work Bank.

9.   https://water.org/our-impact/ethiopia/. (n.d.). Retrieved from https://water.org.

SimAUD 2020

Exploring Homeomorphism in Building Plans

Sabri Gokmen[1]

[1] Kadir Has University
Istanbul, Turkey
sabri.gokmen@khas.edu.tr

ABSTRACT

This paper discusses a type of graph called "homeomorphically irreducible tree" (HIT) and its applicability for a formal study of symmetry in building plans. As a theoretical introduction, the mathematical properties of HITs are introduced through different historical building samples all of which display symmetry, proportion and homologous wings in their formal organization. The extracted principles are used to formulate a generative algorithm that reduces graph complexity to simple sequential numeric representation. This method is converted to a "homeomorphic machine" that is explored through generative plans. The aim of the paper is to introduce a new graph-based approach for potential morphological research into architectural symmetry.

Author Keywords

Homeomorphism, Graph, Digital Heritage, Symmetry, Morphology

1 INTRODUCTION

In the 20th century, the advent of computation transformed architectural research drastically by opening up the possibility of incorporating new practical methods with an influx of theoretical knowledge from other disciplines into architecture. Architects started using generative algorithms developed within computer science such as cellular automata, swarm simulations and various form optimization methods to develop new digital design tools, giving a more divergent and open-ended trajectory for the post-digital era [1]. Within this development, one can argue that design and research have become highly individualized and specialized, where core architectural principles were replaced with dynamic computational workflows that are justified by various performative, aesthetic or formal criteria. Before this contemporary development, the previous century also gave birth to various "schools of thought," each proposing an alternative methodology and theoretical insight for the incorporation of computation into architectural research. Among these, *shape grammars* suggest the formulation of shape and generative rules proposing a transformational formalism that is applied to painting, sculpture, decoration, ornament and architecture [2]. Another avenue is called the *space syntax*, often focusing on the analysis of the interior spatial configuration of architectural plans by the perceptive, locomotive or behavioural aspects of humans [3]. Compared to grammars, the syntactic approach relies on discovering patterns among computationally analyzed spatial configurations that were represented as axial maps or graphs of connected lines [4]. The third formulation for architecture and computation was *architectural morphology* that gave primary consideration to the formal organization and physical properties of buildings to develop abstract mathematical models for the spatial and performative analysis of various building types [5]. One of the strengths of this approach is its interdisciplinary agenda, often combining knowledge from biology, computation and mathematics to establish an integrated scientific approach towards the geometric study of architectural form. This places architectural morphology closer to other disciplines in natural sciences due to its systematic approach and kinship to the utility of classification methods, while shape grammars have recently shown powerful computational permutations with "grammar machines" showing applicability with ongoing research in computer science [6].

One of the singularities that still await architectural research is its potential transformation with the advent of artificial intelligence and machine learning methods [7]. This places great concern on the formulation of analytical and computational methods that will be interpreted, modified and applied by neural networks. Current GAN (Generative Adversarial Network) models offer reinforced training methods where an array of building samples can be supplied to a machine learning algorithm to generate new building footprints, stylistic representations and furniture layouts [8]. In this example, a generic building outline informs how an optimal inner spatial division of an asymmetrical shape can be achieved as a top-down and subdivision based process showing great potential for future development.

The following paper aims to contribute to the ongoing investigation on computational morphology by presenting an alternative study combining symmetry and contracted graphs that are studied through historical works of architecture. Compared to other stochastic methods that determine spatial division from an outer boundary, this method is primarily driven by intrinsic generative principles

centered on the notion of symmetry and branching morphogenesis in architecture. The aim of the research is two-fold. Firstly, formal computation is considered within a historical continuum where it can be applied to a broader class of historical works of architecture, potentially drawing links between them. This also requires a re-evaluation of core architectural principles such as symmetry and proportion that can potentially remedy architecture's relationship with natural sciences. Secondly, computational methods need to be formulated primarily according to architectural principles that can overlap with various mathematical and algorithmic applications while conforming to the historical development of architectural knowledge. This aspect can have both theoretical and practical implications to architectural research and influence the formulation of new methods for machine learning in the future.

In this paper, a mathematical term "homeomorphism" is explored through contracted graphs that show potential application into the geometric study of heritage, while offering an alternative rule-set comparable to shape grammars and space syntax. Alternatively, this method will be defined as "formal computation" that neither requires a description of primitive shapes nor axial lines, in contrast, building forms are defined through abstract generative operations that recursively increase formal complexity of a geometric boundary while inner spatial divisions are simultaneously determined. The current application is directed to branching morphologies in architecture to introduce the method using simple numeric representation, while the overall geometric representation of building forms are kept abstract to avoid stylistic representations that could be added in future studies.

2   GRAPHS IN ARCHITECTURAL RESEARCH

Historically, graphs have been used in many other scientific fields and architecture due to their abstract nature and ability to capture complex structures with simple notations [9]. Among many mathematical models that are used for the study of architectural form, graphs most commonly found utility in highlighting connectivity and complexity of plans. In its mathematical definition, a *graph* is a collection of vertices connected with lines that can have diverse topological properties and classifications [9]. For instance, a graph with branching nodes with no edges joining a vertex to itself in a loop is called a *tree*, whereas graphs with cycles define more networked configurations or structures, like a mesh. In architecture, both graph types have been used to study various formal and spatial aspects of plans, such as the topological complexity of skeletonized alphabet plans [10], spatial connectivity and adjacency within rooms of a building [11], and enumeration of networking plan configurations [12]. In these examples, space is often abstracted using nodes for rooms, or the axis for corridors to reveal hierarchical or connectivity structure. Due to this depiction, graphs have been utilized to highlight the organization in continuous structures, particularly plans.

The motivation for the following work is the potential juxtaposition of hierarchical contracted graphs with architectural symmetry that can offer a new method for the morphological and computational study of building plans. Due to the presence of symmetry in these graphs, the primary case studies are selected from historical buildings that exhibit symmetry, patterned wing development and internal divisibility in plans [10]. A first glance at historical works shows that these properties are mostly found among institutional building typologies such as pavilion hospitals, asylums, prisons, military structures and schools. The broader analysis also shows HIT properties among many stylistic examples that exhibit wings. These can be found in modern, Gothic, Baroque and Renaissance architecture that also display symmetry in form and repetitive inner subdivision of rooms that remain secondary to the overall building form.

3   HOMEOMORPHICALLY   IRREDUCIBLE   TREES (HITS)

In Gus Van Sant's 1997 film "Good Will Hunting" an MIT Professor in mathematics challenges his students with difficult math problems that take researchers years to prove [13]. To his surprise, these problems are not solved by any of his elite students, but by an anonymous mathematical prodigy who works as a janitor cleaning the hallways of the classrooms. One of the problems presented in the movie explores the possible connectivity graphs of trees of order ten, of which there are ten possibilities. This mathematical problem is more accurately defined as "homeomorphically irreducible trees" (HITs) that represent unique connectivity graphs described by points and lines [14,15].
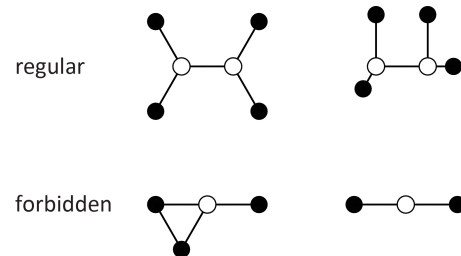
Figure 1. Regular and forbidden cases of homeomorphism. Top row showing two homeomorphically identical trees with six total nodes. Bottom row showing forbidden cases producing a loop (left) and reducible two-legged node (right).

A general characteristic of these trees is their topological properties that are related to symmetry and transformation. Firstly, a homeomorphic tree does not produce any *cycles* connecting nodes back to itself, forbidding triangles or other closed loops emerging inside the network (Fig.1). Secondly, a homeomorphic tree represents all possible topological variations of a unique node configuration. Reflection, rotation or scaling of nodes or trees do not produce another unique tree, as homeomorphism represents node connectivity of a certain mathematical space [16], because changing angles or lengths of a node does not change the topology or connectivity of the tree [14]. Finally, homeomorphic trees appear as *contracted graphs*

where each node has at least three connections to other nodes, rendering nodes with two connections as reducible [15] (Fig.1).

In mathematics, HITs are characterized by single numbers – the total number of nodes in the graph. However, as the number of vertices increases, the unique graphs increase exponentially, making the identification and classification of individual trees difficult. This issue also presents problems for the computation of certain trees, as supplying single numbers lacks internal structure and hierarchy for the graph visualization. A solution to this issue can be mediated by differentiating between nodes of a tree, as internal and external nodes [15]. An internal node (white) inside the graph has at least three connections that can be to another internal or external node, while an external node has only one connection that is to an internal node. Due to the property of *irreducibility* of homeomorphism, HIT gives the option to enumerate the sequential connectivity of internal nodes in the simplified format of $\{i_1, i_2, i_3, …, i_j\}$ [15]. This model captures the valence of internal nodes as a sequence that gives a way of identifying each individual tree as well as offers a way of computing and comparing HITs. With this reformulation the unique trees on the figure 2 can now be redefined in the format of $\{9\}$, $\{3,3,3,3\}$, $\{3*,3\}$, $\{4,6\}$, $\{5,5\}$, $\{4,3,4\}$, $\{3,5,3\}$, $\{3,4,4\}$, $\{3,7\}$, $\{3,3,5\}$ respectively, all of which have a total of ten nodes. This way the numeric representation not only captures the amount of internal and external connections and propagation of the network but also can say a lot about the overall architecture of the graph using simple sequential numeric parameters.
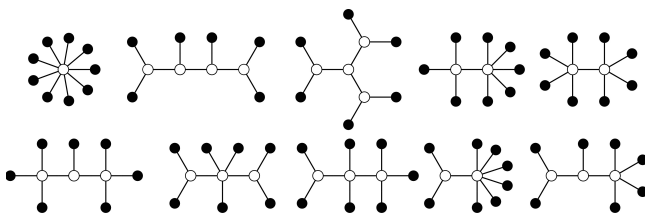
Figure 2. HITs for n = 10. Top row: $\{9\}$, $\{3,3,3,3\}$, $\{3*,3\}$, $\{4,6\}$, $\{5,5\}$. Bottom row: $\{4,3,4\}$, $\{3,5,3\}$, $\{3,4,4\}$, $\{3,7\}$, $\{3,3,5\}$ (For the numeric representation of the third tree, see part 7).

Using figure 2 some of the key visual and geometric properties of the HITs can be discussed. Firstly, HITs with non-constant degree values display a state of equilibrium between nodes while mediating radial and bilateral symmetries (i.e. $\{4,6\}$, $\{3,4,4\}$, $\{3,7\}$) . Another aspect of these trees is the emerging axial propagation that is terminated with radial symmetries. In these trees, the sequential notation becomes more intuitive where the linear progression of the connected internal nodes represents the overall topology of the three. This property also renders notation for identical HITs such as $\{4,6\}$, $\{6,4\}$ as equivalent and signifying both of them becomes redundant.

With the SAS model, the visual representation of constant degree value HITs presents some special cases. For instance, the tree $\{3, 3, 3, 3\}$ for n=10 is represented in two different ways shown in figure 2 since sequential code

of the HITs do not signify the internal connectivity nor propagation of nodes [15]. To differentiate them, the radially symmetric variation of $\{3, 3, 3, 3\}$ is further simplified into $\{3*, 3\}$ by attributing the second number of branches to all primary branches (See part 7 for further explanation of the * symbol). There are similar cases for $\{4,4,4,4\}$ and $\{3,3,3,3,3,3\}$ in n = 14, where the latter shows at least three different visual manifestations. Although the SAS model has limitations to represent radial trees as starting from central nodes lacks directionality, it offers robustness for axial HITs that can be numerically reduced to a sequence.

4   HOMEOMORPHIC ARCHITECTURE

In architectural research, graph theory has become a major tool to study various aspects of plan configuration, adjacency and connectivity of rooms, internal circulation of spaces [10, 11, 12]. In these studies, graphs often reveal hierarchical structures as well as formal relationships between homologous structures. For instance, graphs can be both applied to understand relationships on the *micro* scale- between rooms, or on the *macro* scale between wings that are connected with circulation. The former category often generates cycles due to multiple connections between adjacent rooms violating the *acyclic* nature of HITs. In the latter, branching morphologies that are made of homologous wings are suitable for HIT representation. Due to this morphological characteristic, the primary examples of homeomorphism in architecture are found among historical examples that exhibit symmetry in their plan. Furthermore, typologies with patterned wing development exhibit cases of homeomorphism that in most cases overlap with the building form. These characteristics are mostly found among hospitals, asylums, prisons, military barracks and schools that are often formed with identical wings or blocks (Fig.3). For instance, in radial prisons, such as Pentonville (1844), the wings containing cells reveal homeomorphism of the plan that overlaps with the symmetry axis of doubly-loaded corridors [17]. In asylums and pavilion hospitals, homeomorphism is mostly articulated with patterned wing development or repetition of the pavilion along an axis [10, 17]. In the latter, the "telephone-pole" plan allows the hospital to be easily expanded while offering an organic growth model to increase the capacity towards the open suburban landscape when needed [10].

Further evaluation of historical plans also shows the presence of homeomorphism among stylistic examples found in modern, Gothic, Baroque and Renaissance architecture that also display symmetry in form (Fig.3). In Gothic architecture, homology among wings or building partitions can be observed with repeating bays that exhibit identical vaults. In Salisbury, the main axis of the cathedral branch into multiple transepts while internally dividing these wings to similar vaulting structures. Homeomorphism can also be observed among Palladian villas with the bilaterally symmetrical house branching towards the

landscape. In these structures, symmetry mainly overlaps with internal corridors while each wing undergoes differentiation in the internal division of rooms [19]. In modern buildings such as hotels and offices, homeomorphism results in more repetitive wing plans that are joined at vertical circulation cores.

In figure 3, some of the analyzed historical examples are shown. These buildings are chosen based on the simplicity of their plan and application of homeomorphic principles to exemplify the approach. In this study, each building plan is first traced to reveal the internal skeleton of the building that generally overlaps with the circulation. Then points of convergence are found by intersecting the symmetry axis

that defines the internal nodes of the graph. In the final step, two versions of the HIT diagram are shown. Compared to axial lines in space syntax [4], these graphs do not overlap with the longest visibility lines in space, but instead, define the symmetry skeleton of a building form. Shown under the "symmetry/form" column in Fig.3, the skeleton or symmetry axis of the building form is traced to reveal how it overlaps with the HIT graph. In the subsequent "HIT diagram," the normalized version of the graph is shown to reveal patterned behaviour and repetitive wings of the plan. Among the analyzed samples, Pentonville exhibits a HIT diagram of {5} while Kirkbride Asylum plan shows patterned development in the form of {4,4,4...4} starting from the central block. Salisbury Cathedral can be

Precedent	Abstracted Plan	Symmetry / Form	HIT Diagram

Pentonville Prison, 1844

Kirkbride Plan, 1848

Hotel Nacional de Cuba, 1930

Villa Barbaro, 1560

Salisbury Cathedral, 1320

Herbert Hospital, 1865

Cadillac Palace, 1922

● External Node
○ Internal Node
◉ Root Node
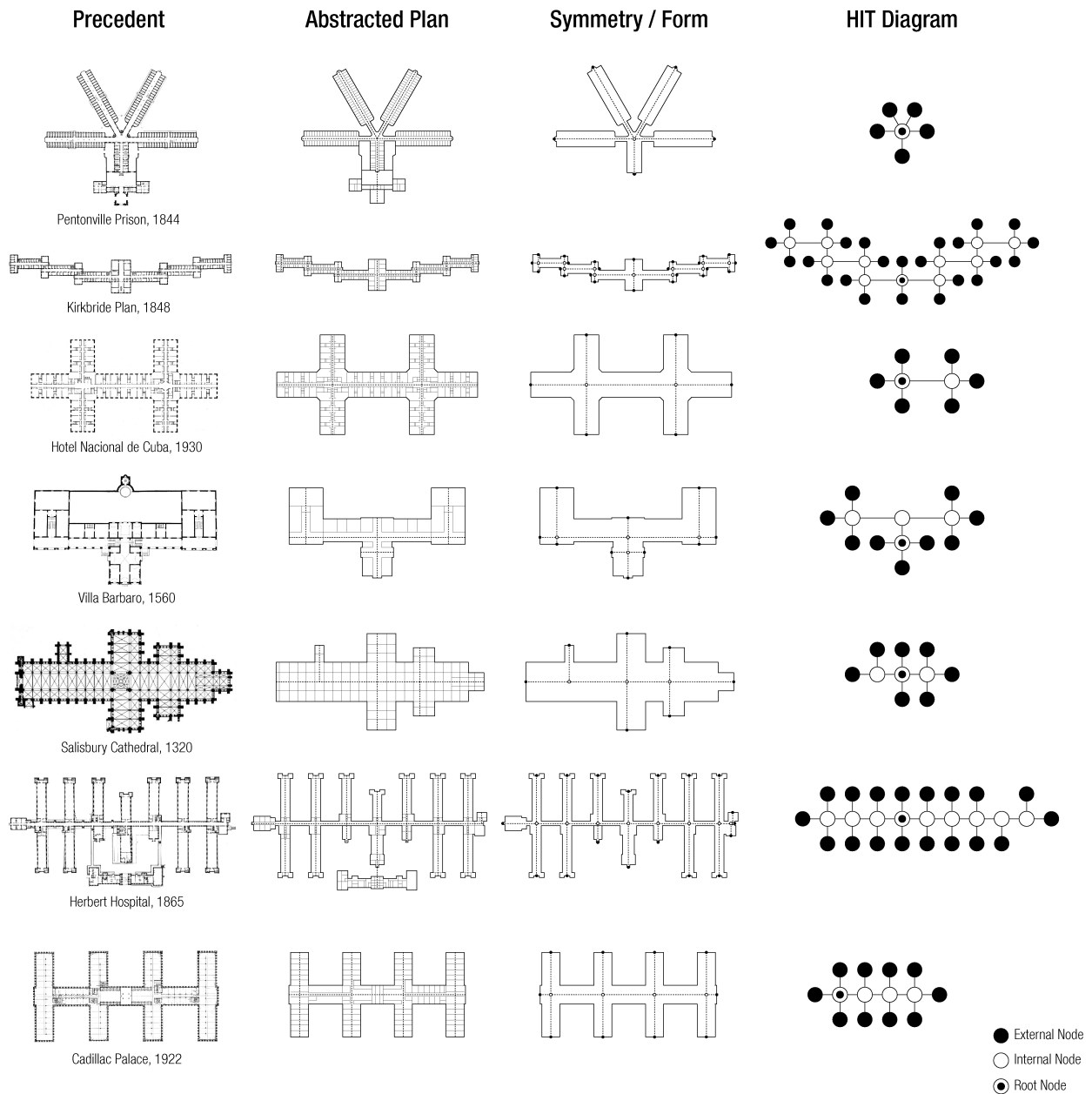
Figure 3. HIT plan analysis of various historical samples.

234

simplified into {3,4,4}, while Hotel Nacional de Cuba ({4,4}), Herbert Hospital ({4,4,4...4,3,3}) and Cadillac Palace ({4,4,4...4}) show similar branching morphologies along an axis. This method shows the macro scale applicability of homeomorphic graphs while the sample sizes in each group can be increased to show more in-depth morphological similarities.

5   COMPUTATION OF HOMEOMORPHIC PLANS

The studied historical examples in the previous part show that in many historical works of architecture the symmetry axis coincides with the homeomorphic inner circulation connecting internal spaces or rooms within blocks. This approach gives a primary consideration for the overall form of the plan particularly expressing a differentiated and highly articulated outer geometrical boundary that separates internal spaces from the external environment. In the following model, this emphasis is transformed into an abstract formal algorithm to consider symmetrical relations in form while issues of scale and measurement in architecture are ignored to visualize the potential computational applications and simplicity of the HIT approach.

To investigate architectural symmetry computationally a simple algorithm structured around the mathematical notion of HIT is developed that can generate typical building plans with axial doubly loaded internal corridors. This recursive algorithm is developed in Python using the principles highlighted in the SAS model where the formal complexity of a tree graph can be represented with a sequence of numbers [15]. In figures 4-5 the procedural development of this algorithm is visualized for the HIT of {3, 4, 6} containing 12 nodes. Starting from the first internal node that acts as the root of the tree, the algorithm uses numeric code to propagate the tree structure in a linear fashion producing a major symmetry axis.
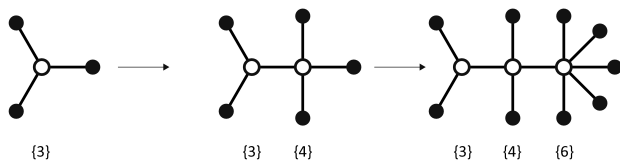
Figure 4. Generative modelling progression of {3, 4, 6}. Internal nodes are marked with white while external nodes are marked with black dots. Each step (numbers in HIT) shows the addition of external branches to achieve the total number of connections to an internal node.

Computation of homeomorphic plans occurs in two stages. In the first stage that is *additive*, the valency values of internal node sequence are used to generate a tree recursively (Fig. 4). For HIT of {3, 4, 6}, the first root node {3} is radially produced in the first step. Then, three new branches are added to produce the second internal node containing a total of four connections {4}. In the third step, five new branches are added for {6} that terminates the additive stage.
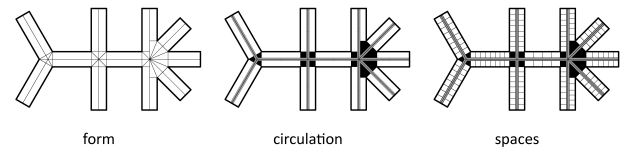
Figure 5. Thickening and subdivision of HIT. Each axis is first offset into a polygon that is divided into three zones – circulation (grey), services (black) and spaces (white), the latter can be further subdivided into rooms using parametric thresholds.

To turn the HIT diagram into a plan, the networking graph lines are offset while forming closed polygons located at the internal nodes (Fig.5). For three and four-legged nodes, the thickening results in uniform polygons. In higher valency nodes, such as five-legged nodes, due to the uniform thickening of the wings, certain asymmetrical polygons start occurring in the plan. For the wings, the circulation is placed along the symmetry axis marked by the HIT graph of the form. This way each wing is composed of two symmetrical halves that are defined by an internal corridor and external-facing spaces or rooms.
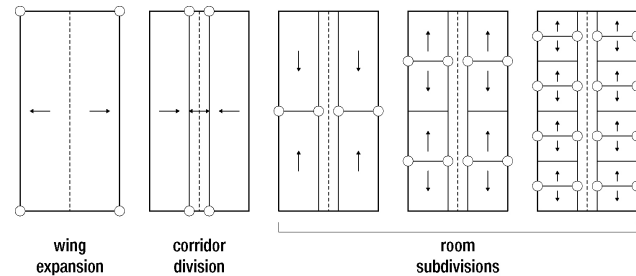
Figure 6. Recursive inner subdivision of wings that determine spaces.

Once the form of the plan is thickened, each emerging rectangular polygon of the wing is recursively subdivided into two halves using a division threshold, progressively producing smaller spaces or rooms (Fig.6). This *subtractive* method takes inspiration from cellular division and packing of cells by developing symmetrical relationships among parts of the plan [20]. All the topological properties and recursive stages of the HIT architecture are controlled by parameters that allow for various configurations and generative studies of emergent forms.

6   ANALYSIS OF HOMEOMORPHIC PLANS

In the next phase, the HIT rules are used to produce permutations of various architectural plans using the existing mathematical codes for 4, 5, 6, 7, 8, 9 and 10 node trees (Fig. 7). To achieve comparable results, the algorithm is supplied with the same parameters for branch thicknesses and lengths, corridor to floor proportion and room subdivision coefficients. The resulting plan forms are analyzed according to their floor area, the number of spaces (white) created, corridor (grey) and service spaces (black), spaces with natural (white) and artificial (dark – grey and black) lighting / ventilation (Table 1). To compare the performative aspects of emerging trees, the ratios of various spaces are used.
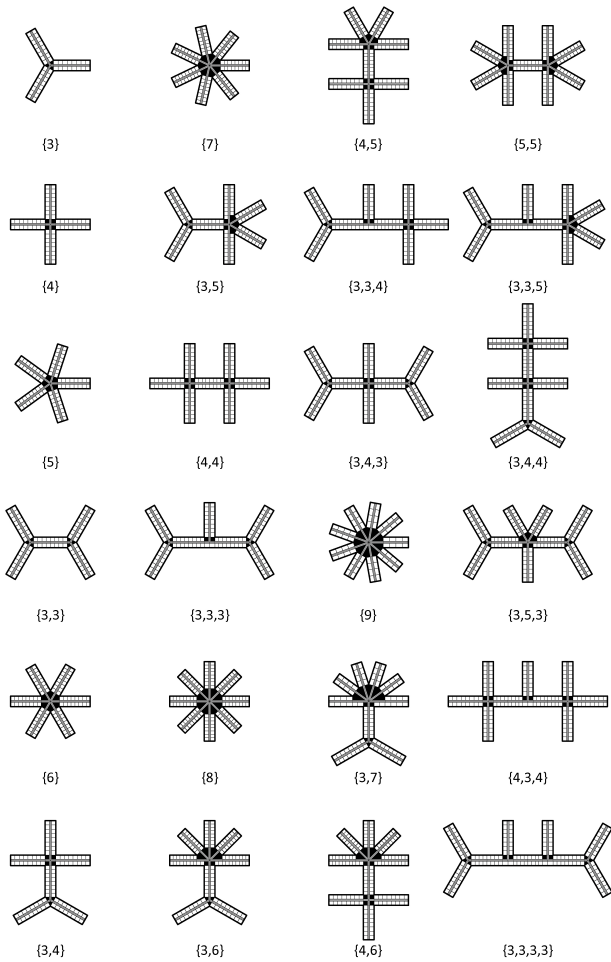
Figure 7. HIT plans showing n = 4 ({3}), n = 5 ({4}), n = 6 ({5} {3,3}), n = 7 ({6}, {3,4}), n = 8 ({7}, {3,5} {4,4} , {3,3,3}), n = 9 ({8}, {3,6}, {4,5}, {3,3,4} ,{3,4,3}), n = 10 ({9} , {3,7} , {4,6}, {5,5}, {3,3,5}, {3,4,4}, {3,5,3}, {4,3,4}, {3,3,3,3}).

The results show that among the produced HIT plans, radial symmetry causes an increase in service areas (Black spaces). As the number of nodes increases, the total outline of the plan form increases allowing for more natural lighting (White spaces). However, the graph variations for the same node count does not play any major role in the number of total rooms created or the overall floor area, except the cases with radial symmetry that produce connections to at least five external nodes. All the possible trees for 5, 6, 7, 8, 9 nodes produce similar number of spaces (rooms), and most of the trees produce similar circulation area / total area ratios, showing that adding new branches does not change the spatial performance of the plans. However, adding more branches often increases the outer surface area for the same node count trees and can improve daylight or ventilation performance. This becomes evident when constant value HITs ({5}, {6}, {7}, {8}, {9}) are compared to the similar node count trees.

In figure 8, the ratios for BT (Black Areas / Total Areas) and DL (Black Areas / White Areas) are plotted in a graph

for the analyzed HITs. The values for single node HITs are connected using a line to show the standard distribution for daylight/ventilation performance. This graph shows that as the inner node count increases, the HIT moves to a lower BT (reducing service areas) or DL (increasing external-facing spaces) gaining better performance for natural light or ventilation (Fig.8). Furthermore, multiple inner node HITs with three and four-legged networks generate a higher performance (Lower BW) compared to five or more legged nodes and single number HITs.

HIT	TN	IN	EN	FB	S	CT	GT	BT	BW
{3}	4	1	3	6.3	48	0.20	0.26	0.02	0.03
{4}	5	1	4	8	64	0.20	0.26	0.04	0.06
{5}	6	1	5	9.4	80	0.21	0.27	0.06	0.08
{3,3}	6	2	4	10.14	80	0.20	0.26	0.03	0.04
{6}	7	1	6	10.8	96	0.22	0.28	0.08	0.12
{3,4}	7	2	5	11.8	96	0.21	0.26	0.04	0.05
{7}	8	1	7	11.9	112	0.22	0.29	0.10	0.15
{3,5}	8	2	6	13.2	114	0.21	0.27	0.05	0.08
{4,4}	8	2	6	13.4	112	0.21	0.27	0.05	0.07
{3,3,3}	8	3	5	13.8	114	0.20	0.26	0.03	0.04
{8}	9	1	8	12.98	128	0.23	0.29	0.12	0.18
{3,6}	9	2	7	14.29	132	0.22	0.28	0.08	0.11
{4,5}	9	2	7	14.88	130	0.21	0.27	0.06	0.09
{3,3,4}	9	3	6	15.53	130	0.21	0.26	0.04	0.05
{3,4,3}	9	3	6	15.6	128	0.21	0.26	0.04	0.05
{9}	10	1	9	13.8	72	0.23	0.30	0.14	0.23
{3,7}	10	2	8	15	100	0.22	0.29	0.10	0.15
{4,6}	10	2	8	15.9	148	0.22	0.28	0.08	0.12
{5,5}	10	2	8	16.2	148	0.21	0.28	0.07	0.10
{3,3,5}	10	3	7	16.9	148	0.21	0.27	0.05	0.07
{3,4,4}	10	3	7	17.2	144	0.21	0.26	0.04	0.06
{3,5,3}	10	3	7	17	146	0.21	0.27	0.05	0.07
{4,3,4}	10	3	7	17.1	146	0.21	0.27	0.05	0.06
{3,3,3,3}	10	4	6	17.6	148	0.20	0.26	0.03	0.04

Table 1. Topological comparison of HITs for node counts between 4 and 10. TN: Total Nodes, IN: Internal Nodes, EN: External Nodes, FB: Form Boundary / Branch Length, S: Number of Spaces, CT: Circulation Area / Total Area, GT: Grey Areas / Total Area, BT: Black Areas / Total Area, BW: Black Areas / White Areas.
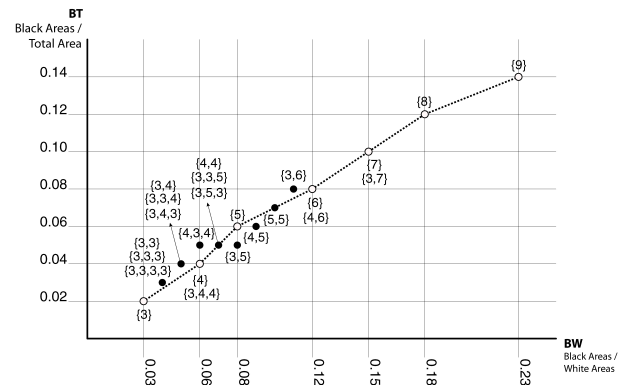
Figure 8. Daylight and natural ventilation performance of analyzed HITs. A dotted line connects single number HITs (White Dots). HITs with more internal nodes move to lower BT and DL areas improving performance.
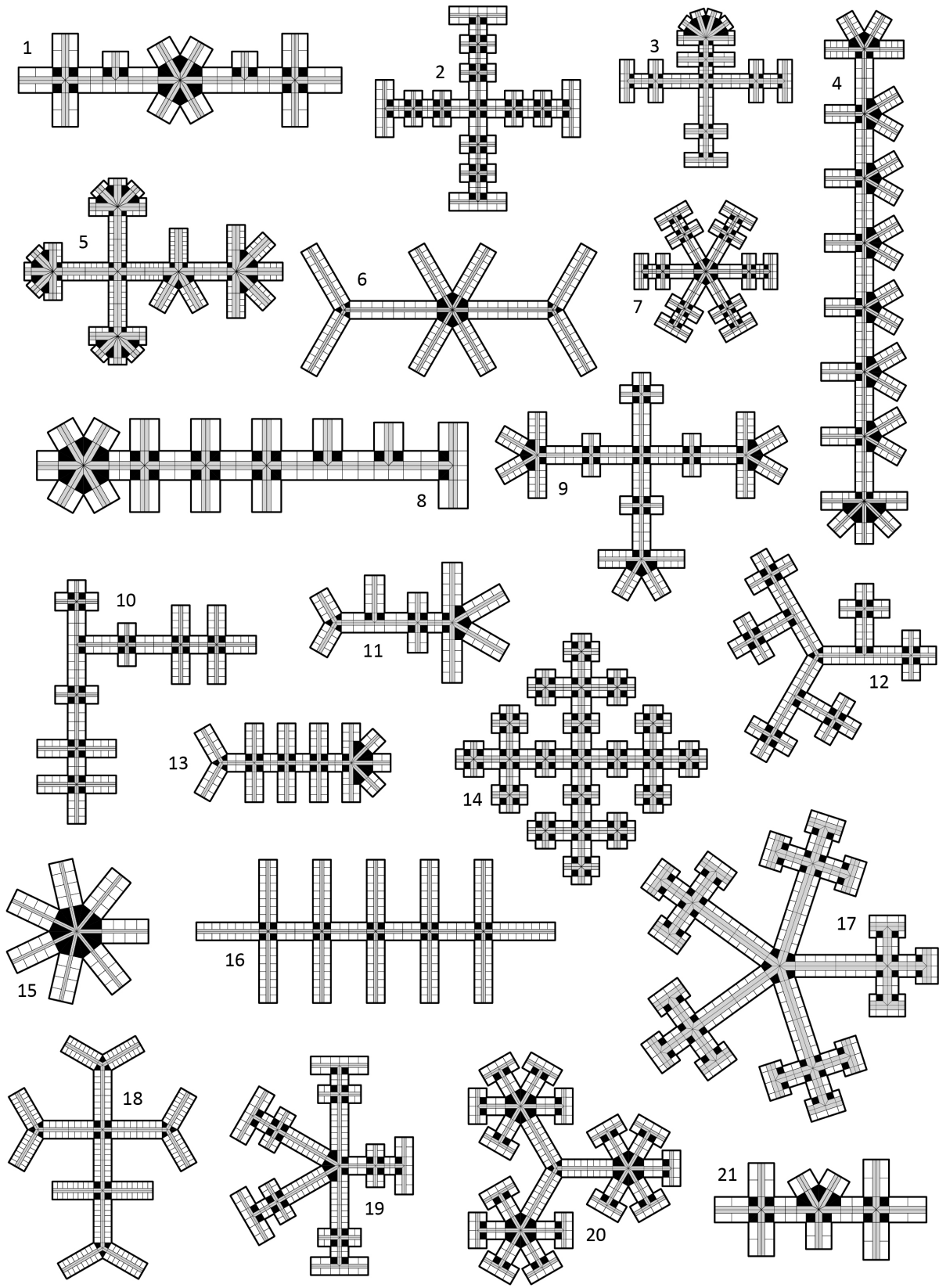
Figure 9. Variations of the HIT algorithm showing diverse architectural plans. The HIT code for generated samples: 1. {4, 3, 6, 3, 4}, 2. {4*, 4,4,3}, 3. {6, 4, 4*,4, 3} , 4. {5, 5, 5, 5, 5, 5, 5, 6}, 5. {5, 5, 4*,6}, 6.{3,6,3}, 7.{6*,4,3}, 8. {6,4,4,4,3,3,3}, 9.{4, 4*,4, 5}, 10.{4, 3*,4,4,4}, 11.{3, 3, 4, 5}, 12.{3*,3*,4}, 13.{3,4,4,4,6}, 14.{4*,4, 4*, 4}, 15.{7}, 16.{4,4,4,4,4}, 17.{5*, 4*,3}, 18.{3,4, 4*,3}, 19.{5*,4,3}, 20.{3*,6*,3}, 21.{4,5,4}.

237

7 HOMEOMORPHIC PLAN VARIATIONS

In the last step, a "HIT machine" is created to show a generative application for architectural plans. As an improvement, the sequential numeric HIT code is arranged with an additional symmetry condition, where an internal node number can be marked (*) to apply the same HIT progression to all the subsequent nodes to generate diverse building forms with local and global symmetries. This way, the form of a building can be represented with numbers, like a compressed DNA sequence, offering various genetic studies and comparisons. In figure 9, homeomorphic form variations of the improved algorithm are shown that can be further elaborated with additional geometric rules.

8 CONCLUSION

This paper discussed the mathematical term "homeomorphically irreducible tree" (HIT) to offer a simplified graph approach for the study of architectural plans. The presented method offers a simplified sequential numeric code for branching architectural morphologies in the form of $\{i_1, i_2, i_3, …, i_j\}$ that can generate wide parametric variability. The results show that HIT algorithm can be both used to study the notion of symmetry among various historical building types and as a generative modelling tool. However, HITs cannot be used to represent plans with courtyards or other circulation systems with circuits as cycles are forbidden in their configuration. While the analyzed historical examples show homeomorphic properties, further investigation into the directionality and symmetry properties of HIT is required to offer a comprehensive analysis of each historical building type. To summarize, this paper presented an alternative use of graphs for architectural research that is described through "formal computation." With this perspective, building plans are primarily studied through their outer form boundary and symmetry axis while inner spatial division can be achieved by subdividing each branch or wing of a building. This way, HITs offer a higher level of abstraction and parametric simplification to generate complex building forms while providing an exciting computational avenue for a genetic study of architectural morphology.

AUTHORSHIP INFORMATION

This paper is a continuation of a journal paper titled "Homeomorphic Architecture," accepted to be published for Environment and Planning B in March 2020. The work presented here expands on the idea by showing analysis of other building types, providing performative comparisons of variations and generating homeomorphic plans.

REFERENCES

1. Carpo, M. *Digital Turn in Architecture: 1992 – 2010* (2013). Hoboken: Wiley.

2. Stiny, G. and Gips, J. "Shape Grammars and the Generative Specification of Painting and Sculpture." In C.V. Frieman (Ed.) *Information Processing* 71 (1972).

3. Hillier, B. *Space is the machine* (1996). Cambridge: Cambridge University Press.

4. Turner, A., Penn, A., & Hillier, B. "An algorithmic definition of the axial map, *Environment and Planning B: Planning and Design* 32,3 (2005): 425-444.

5. Steadman, P. *Architectural Morphology: An Introduction to the Geometry of Building Plans* (1983). London: Pion.

6. Ligler H, and Economou A."From Drawing Shapes to Scripting Shapes: Architectural Theory Mediated by Shape Machine", *Proceedings of the Symposium on Simulation for Architecture + Urban Design (SimAUD)*, (2019).

7. Carpo, M. *The Second Digital Turn: Design Beyond Intelligence* (2017). Cambridge, MA : The MIT Press.

8. A recent promising example to this study is the ArchiGAN that was developed by Stanislas Chaillou as a Harvard GSD master's thesis project.

9. Harary, F. *Graph theory* (1994). Reading, MA: Addison-Wesley.

10. Steadman, P. *Building types and built forms*. (2014). Kibworth Beauchamp, Leicestershire, UK : Matador.

11. Ostwald, M.J. "The Mathematics of Spatial Configuration: Revisiting, Revising and Critiquing Justified Plan Graph Theory." *Nexus Network Journal* 13 (2011): 445–470.

12. March, L. *The Architecture of Form* (2010). Cambridge: Cambridge Univ. Press.

13. Van Sant, G. (Director). *Good Will Hunting* (1997). United States: Miramax Films.

14. Haran, B. The problem in Good Will Hunting – Numberphile. Youtube. https://www.youtube.com/watch?v=iW_LkYiuTKE. Accessed 04 July 2019, (2013).

15. Gerlach, J. R., Dataceutics, Inc. "Generating homeomorphically irreducible trees (Solving the Blackboard Problem in the Movie "Good Will Hunting")." *PharmaSUG 2017 Conference Proceedings*, Applications Development 11 (2017).

16. Gamelin, T. W. and Greene, R. E. *Introduction to topology* (1999). New York: Dover.

17. Kirkbride, T. S. *On The Construction, Organization, And General Arrangements Of Hospitals For The Insane* (1854). Philadelphia: [publisher not identified].

18. Johnston, N. *Forms Of Constraint: A History Of Prison Architecture* (2006). University of Illinois Press.

19. Wittkower, R. *Architectural Principles In The Age Of Humanism* (1965). New York: Random House.

20. Steadman, P. "Why are Most Buildings Rectangular?"*Arq: Architectural Research Quarterly*. 10,2 (2006): 119-130.

Context Specific Evolutionary Design: An Analysis on Computational Abstraction of Modern Urban Complexity

Yufeng Zhai, Elisabeth Riederer

Architectural Association (AA) School of Architecture
London, United Kingdom
Yufeng.Zhai@aaschool.ac.uk, Riederer@aalumni.org

ABSTRACT

Evolutionary design is used to adapt urban systems to predictions such as rapid growing density and effects of climate change scenarios. These effects have weakened the strategies on which ancient cities were built and thrived. Fez el Bali's Medina can be seen as a drastic case and is therefore chosen for theoretical investigations. The Medina nowadays has lost its quality as a functioning system, characterized by a coherent relation of hierarchical order and randomness based on a cultural heritage.

In this paper's architectural approach a city is redeveloped on the basis of the earlier well-functioning ancient city after which an urban patch is then developed further. In order to react to the unpredictable changing conditions, we propose an open system generated by outlining the qualities the Medina was built on and developing this further to be able to react to changes within the city and beyond it.

The key element of this paper is to expand on the level of complexity in Evolutionary Design by operating on the urban scale and contextualized to push its computational potential. Expanding on the application of design strategies of Genetic Algorithms (GA) we incorporate rule-based multi-scale procedural modeling based on the vernacular urban qualities, while examine urban morphological variation evolved in response to conflicting criteria by means of a Multi-Objective Evolutionary Approach (MOEA).

Author Keywords

Evolutionary Design; Multi-Objective Genetic Algorithm; Urban design; Generative Design; Super Blocks; Morphological Variations.

1   INTRODUCTION

Evolutionary Computation to Encounter Urban Complexity

Examining the changes in biological structures has shown how the forms are created through invention and expansion with existing body parts being crucial to innovation [1]. Evolutionary trends more frequently follow available paths so that major changes in evolution occur via mutations in existing genes [2]. The modern city as a complex organism shaped by its dynamics might relate to these evolutionary process to improve adaptation. Evolutionary strategies such as innovation building upon existing structures also seem to be found in well-adapted urban structures. Even though the methods of Evolutionary Computation such as a Multi-Objective Evolutionary Approach (MOEAs) may be seen as an appropriate tool in urban design and have been proven to be an efficient design method to encounter a certain level of complexity, Evolutionary Design is still rarely used in practice. This might be due to the fact that evolutionary models mostly aim to design scripts, which are faster and more effective. Rationalizing scripts might lead to a downside for the design outcome and to questioning the integrity of GA [3]. This development might be initiated by Evolutionary Design most commonly used as a form-finding tool to explore novel forms. These form-finding processes, lack of context, and therefore also practicing abstraction of context specific data as an input for GA. By examining this abstraction further we hope to expand on the topic of context specific Evolutionary Design and trigger its application in practice.

The key element of this paper is to expand on the level of complexity in Evolutionary Design by operating on the site-specific urban scale to push its computational potential. Addressing the application of Evolutionary Design and expanding on the complexity level of Evolutionary Computation as a design method we incorporate complex form generation rule sets and multi-scale procedural modeling while examining urban morphological and topological variations evolved in response to conflicting criteria by means of a MOEA.

The following studies examine the ability of a GA to encounter the degree of modern cities complexity in design solutions. Theoretical investigations are exposed through

redeveloping an urban center to generate optimized design solutions without compromising its cultural values, urban fabric and typology while addressing nowadays conflict of social structures. Due to tourism and overpopulation many ancient cities are affected by social conflicts with different stakeholders fighting for more space respectively. By studying these vernacular well-functioning self-organized urban regions and analyzing the contemporary urban issues within, we are aiming to redevelop and improve the "better equipped" urban fabrics.

Fez el Bali´s Medina to Demonstrate Urban Complexity
Many ancient cities around the world have survived centuries and are well adapted to their environmental context as well as the social economical context through self-organized evolutions. Evolutionary algorithms have the potential to generate the same level of complex well adapted urban design options with only a fraction of the time. This makes it an ideal and efficient design tool for redevelopment in historical urban regions.

One of the typical vernacular self-organized urban regions facing severe challenges nowadays is Fez el Bali (Medina of the city Fez, Morocco). Fez el Bali's Medina builds upon a complexity characterized by a coherent relation of environmental factors and hierarchical order based on cultural heritage and is therefore chosen for investigations. In recent years, the living quality of Medina´s residents has been reduced due to the increasing population density and rapid growth in tourism. Therefore the Medina functions as an admonishment for city developers to consider future developments and to include these in their planning. Multiple statistical data indicates this development and leads to a growing conflict of different stakeholders. Local residents, small business owners providing services and merchandise to tourists and local craft workshops each seek for more space.

2 BACKGROUND

2.1 Fez el Bali´s Medina
Fez el Bali, as part of the Medina of Fez, is located in the north of Morocco. With a population size of 156,000. The Medina covers an area of 280 ha and has a population density (2002) of 710 inhabitants per hectare [4]. It is also one of the world's largest pedestrian-only urban areas.

The medina Fez is dictated by the typical vernacular courtyard housing blocks. This typology is usually a multi-storey adobe house with an inner courtyard for an extended family dwelling. The fortified exterior facade provides strict privacy. The resultant streets often-occurs narrow and over shaded by the high walls.

Social Context and Urban Development
Fez el Bali is characterized by its high social interaction caused by the limitation of traveling as a result of a free-car area and by its limited public space, forcing people to interact more [4].

A rise in population density caused dilapidated building conditions and a deterioration of living conditions which further led to people migrating from the Medinas. Simultaneously and because of still existing high social structures (relationship and support systems) a shift against this trend appeared and new residents settled [6].

Most of the Medinas have become urban cores of much larger urban agglomerations; In Fez, only 12% of the population lives inside the Medina (relatively high value compared with the Medinas of Casablanca and Tangier (only 2%)) [7]. An analysis of the housing issues predicts a rise in population and a rise of tourism that will lead to conflict between stakeholders, seeking space, and local residents striving for residential areas, small business owners providing commercial shops for tourists and manufacturers seeking space to expand their workshops.

2.2 Urban System Modeling
In order to establish the systematic and computational model for the urban generative design we relate to Farzaneh's[24] systematic urban modeling scenarios. As Farzaneh summarized, systematic urban modeling could be categorized into three scenarios:

1. Block dominated; 2. Network dominated; 3. Co-evolved

In the first two scenarios, either the building blocks or the street networks dominate the overall layout of the urban plan while the other emerges as a resultant pattern. Computationally, these two scenarios are easier to be modeled digitally. Block dominated scenarios link the program, land use, density to dimensions and plot geometrical configurations, while the network dominated scenarios focus on the interrelations between the typological nodes and edges of the street network. In the third scenario, coevolved model, the organizational pattern of the street networks and building blocks reciprocate and mutually adapt and modify each other. Which suggests a feedback loop in the generative process.

Given the characteristics of Medina Fes, the computational experiments presented in this paper focused on the block dominated scenarios.

2.3 Evolutionary Computation Literature Review
Evolutionary computation employs analogies to the biological phenomena of evolution and selection to generate algorithms that are used to obtain computational solutions to problems in numerous fields. These computational methods are suitable for finding solutions to complex problems with multiple objectives that may conflict with each other. The methods generate a set of solutions, which satisfy all objectives without any objective dominating and can then be examined for trade-offs [9]. Because of its background the terminology used in evolutionary computation leans heavily on genetics. For instance, a candidate solution to a problem is referred to as an individual; the sum of the defining properties of the individual is its genome, whilst an individual property is referred to as a gene, the value of

which is an allele [10]. Coupled with optimization, the evolved solution set is a robust and powerful alternative to a single, preference-based approach.

The earliest attempt of utilizing biological evolution inspired by problem solving and optimization could be dated back to the pre-digital computing period 1930s, Sewell Wright's works [11].

A major advancement in the evolutionary computation development was only computationally achievable after the advent of the computers. Several scholars emerged between the 1960s and 1970s simultaneously theorized their computational methods of evolutionary computation.

Overall, much of the research and methods published at this stage focused on the single-objective evolutionary algorithms (SOEA). For instance, John H. Holland's on GA [12], Lawrence J. Fogel published a paper 'On the Organization of Intellect'[20] about evolutionary programming (EP), Rechenberg [13] and Schwefel's[l4] on evolutionary strategies (ES).

The first effort towards MOEA design was led by David Schaffer in the 1980s through his published paper 'Multiple objective optimization with vector evaluated genetic algorithms' [14] .

Among all the scholars' research, the most significant contribution to increase the efficiency of the MOEA was developed by David Goldberg who introduced Pareto front evaluation and non-dominant selection strategy into the MOEA. [15] By implementing his strategy, the algorithm was able to incrementally optimize the fitness according to each objective independently, 'spread its population out along the Pareto optimal tradeoff surface' and balance variations towards multi objective while the fitness value could still be incrementally optimized. [21]

The introduction of Elitism and Archive by Eckart Zitzler to MOEA diverged the computational evolution model from the biological evolution model. Zitzler argued that the non dominant individuals (local optimal) in the earlier generations, may still dominate across later generations. Therefore, the elite solutions could be archived directly to the next generation to compete with newer generation. [17] The application of his method can be found in the algorithm Zitzler developed: Strength-Pareto Evolutionary Algorithm, SPEA (1999) and SPEA-2 (2001) [18].

The current state of art of Genetic Algorithms incorporate the strategies mentioned above and have been tested on robust computing efficiency. Some of the most popular GA strategies are Kalyanmoy Deb's NSGA III; Knowles and Conre's Pareto-Archived Evolution Strategy (PAES) and Zitzler SPEA-2.

As GA algorithms typically generate large data sets, one of the key developments in recent evolutionary computation lies in the post-evaluation and statistical analytical tool development. This has made it possible for the general users like architects to evaluate and validate the algorithm result. For this reason, the statistical analytics and visualization tools became essential in the big data analysis.

The recent advancement in machine learning also advanced data analysis, for instance, the utilization of K-mean clustering in Wallacei Analytics[19] providing the users insight on how the GA algorithms developed their novel evolution strategies.

In the field of architectural computation, prior works have demonstrated the problem solving capacity of the MOEA in architecture and urban design. Calixto and Celani [28] contributed the theoretical framework of the MOEA utilization in spatial layout design, while other architects delved into the real world architectural problem solving. For instance, the project by Benjamin et al. [25] on the office layout optimization explored quantitatively optimized the intangible objectives like productivity among other objectives. Schwartz et al. [26] novel approach addresses the life cycle carbon footprint and cost of building redevelopment through MOEA.

In terms of MOEA application on the urban design, Makki et al[27] and Farzaneh et al[16] speculative design experiments evinced MOEA potential, however the projects context were largely underplayed and super block modelling were simplified.

3 COMPUTATIONAL EXPERIMENT

3.1 Constraint and Goal

Inquiry to site context identified the constraints of the site:

1. The complex urban pattern of Fes el Bai is dictated by the vernacular housing blocks. It is sensible to adopt this vernacular typology and urban pattern in the speculative redevelopment of the region to meet the Moroccan family structure and social customs.

2. Density constraints: conflicts between living tourism and industry growth. One of the immediate challenges observed in the Fes el Bali is to find a solution which can achieve reasonably high density and balance stakeholders needs, while still maintaining relatively sufficient area of streets and open public spaces. The maximal floor area ratio (FAR) was chosen as a numerical fitness criteria.

3. topographical and climatic constraints: the sloped site marks 66.6 m elevational change, the appropriate building strategy was needed to address the sloped terrain. In terms of the climate, Fes is Hot-summer Mediterranean climate. self shading provided by the 'Deep canyon' streets and blocks mediates microclimate. However, the internal courtyard receives inadequate solar gain daylighting due to the buildings. For this reason, the environmental objective of the optimization was to optimize the direct sunlight gain from the south during the winter season.

Drawing upon the understanding of the site constraints and literature review of MOEA. The Research question could be

stated as: can MOEA be used as an efficient tool to generate contextualized, climatically and socially adapted urban design solutions for historical urban area redevelopment of Fes el Bali?

3.2 Workflow

To answer the question, two sets of design experiments were set up in a consecutive manner. First experiment focuses on housing block cluster (2x2) generative design with only singular typology considered: the collective courtyard house. Experiment 2 focused on the urban scale superblocks and street network generation. The objectives defined for this experiment were the numeric abstraction of the site context analysis mentioned in the previous chapter.

3.3 Methodology

Experiments presented within this article address the issue of generating urban morphological variation evolved in response to conflicting criteria through the modeling of urban form by means of a MOEA. The computation was driven by the SPEA-2 algorithms, through the plugin called Octopus developed by Vierlinger [22], within the Grasshopper visual scripting environment. The procedural modelling and fitness evaluation were also customized in the Grasshopper environment to streamline the generative process.

3.4 Experiment 1 - Single Block

For initial experiments, an 20x 20m urban patch which consists of 4 square blocks was used for the design experiment. Each block was evenly divided into 4 courtyard houses. Regarding vernacular family structure, social costume and climatic conditions indigenous courtyard houses are a proof of concept and are therefore preferred as the primary typology.

Genes

The genes (building variables) include the number of the building floor (1), street width (2) as well as courtyard ratio (3) and room divisions(4). 1, The lowest floor number 0 leads to the formation of open spaces; thus the more variants of semi-open courtyards could occur. 2, Based on two-storey buildings and street widths of ranges 0.5 to 6 m. 3, The courtyard ratio is controlled by a scaling factor of
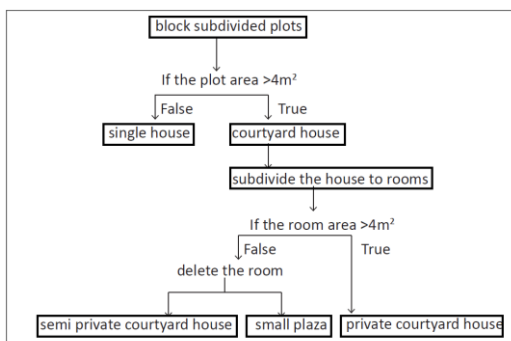
Figure 1. Rule based block generation - Boolean condition

building footprint in the range of 0.2-0.5. 4, 0m building height was also added for every single room to serve as a gene switch which can suppress the extruding building height function. This regulating gene could also lead to opening up the closed courtyard.

Fitness Objectives

1, maximize FAR. 2, maximize street area. 3, maximize south facing direct sunlight exposure (winter solstice 12pm solar vector)

Design Strategy

The ancient Urban qualities are mainly incorporated as ratios in the genes (street to building height and building footprint). We mostly address predictions of rapid growing density and their effects by means of our conflicting fitness criteria (maximizing FAR and solar exposure on the surface areas of courtyards and streets). Therefore Experiment 1 aims to encounter max. density, optimum solar radiation conditions for summer and winter as well as providing sufficient public spaces. The most important urban design considerations based on Fez el Bali´s social and economic context analysis are social conflicts caused by overpopulation and tourism. FAR was chosen as the primary objective to be maximized to accommodate the stakeholders' need for space. To balance the real estate development revenue while ensuring sufficient urban public spaces and amenities we introduced the two conflicting objectives: maximizing FAR and maximizing street area . To counterbalance FAR we identified maximizing street area as the second objective.

In addition to the two objectives described above, the third environmentally driven objective, maximizing the courtyard solar exposure was defined based on the fact that the vernacular multi-level housing type and high-density neighborhood development often compromises the courtyard area, which significantly affects the indoor comfort including direct sunlight gain and daylighting. The population setup of the algorithm was 10 individual solutions per generation by a total of 100 generations. A Rule based procedural modelling was implemented to

Figure 2. Phenotypes and the fitness values
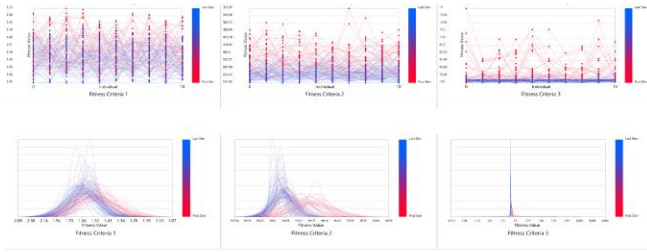
Figure 3. Top chart fitness values (0 marks fittest); bottom standard deviation graph (0 marks fittest)

ensure the algorithm generated architectural solutions are rational and sensible (see Figure 1).

Evaluation

The methods to evaluate the experiment were achieved by implementing a population-based fitness value analysis. By analyzing standard deviation of the individuals' fitness values across the generations provides more insight on the degree of variations and global trend of the optimization (see Figure 3). The standard distribution (SD) graph is generally a bell shaped curve. The wider the bell shape of the standard deviation, the more variations the generation contains). The centerline of each SD curve also indicates the mean fitness value of the generation. If the graph centerlines shift towards optimum direction from older to later generations, it can prove the objective being optimized. By overlaying SD curves across the generations, the exploitation/exploration weights could also be observed. Overall, experiment 1 was successful in terms of achieving sufficient variations while the 3 objectives were optimized incrementally. The graph also reflects clearly that more variations were generated to achieve objective 1 and 2, however, early convergence was observed in relation to objective 3, as the normal distribution curve narrowed down to the mean value in the early generations.
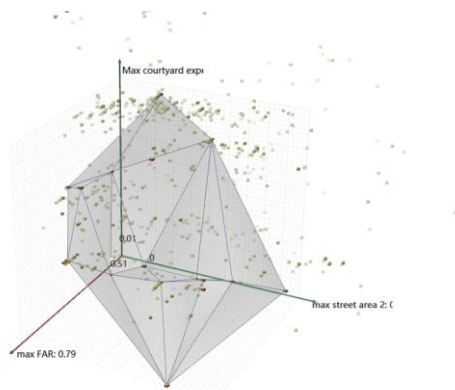
Figure 4. Objective space: solutions fitness value plotted in Cartesian coordinates.

ensure the algorithm generated architectural solutions are rational and sensible (see Figure 1).

Evaluation

The methods to evaluate the experiment were achieved by implementing a population-based fitness value analysis. By analyzing standard deviation of the individuals fitness values across the generations provides more insight on the degree of variations and global trend of the optimization (see Figure 3). The standard distribution (SD) graph is generally a bell shaped curve. The wider the bell shape of the standard deviation, the more variations the generation contains). The centerline of each SD curve also indicates the mean fitness value of the generation. If the graph centerlines shifts towards optimum direction from older to later generations, it can prove the objective being optimized. By overlaying SD curves across the generations, the exploitation/exploration weights could also be observed. Overall, experiment 1 was successful in terms of achieving sufficient variations while the 3 objectives were optimized incrementally. The graph also reflects clearly that more variations were generated to achieve objective 1 and 2, however, early convergence was observed in relation to objective 3, as the normal distribution curve narrowed down to the mean value in the early generations.

3.5 Experiment 2 - Urban Block in Context

Building on the design strategy, Experiment 2 aimed to generate an urban patch which addresses the issue of the land use and additional programs and building types. As the generative process of the whole urban design with multi-objectives is computationally heavy and prolonged, the generative process was divided into two hierarchical steps.

Step 1 – Global Scale Parcel Subdivision

In response to the topography, site, an orientation rule was introduced in the procedural modeling to address the topographic context. The main axis of the parcel always orients with the tangent vector of the nearest topographic contour to reduce elevational change within the building. In addition, Any parcel located on the steep(larger than 20%) terrain will be eliminated. In terms of parcel subdivision and scale variation, a quad-tree subdivision method was implemented. The subdivision logic applied subdivides the rectangular parcel iteratively into 4 parcels and coincides with context geometry. This process is repeated until the children parcels do not coincide with any context geometry (see figure 6). The merit of a quad-tree subdivision method lies in its efficiency to control the subdivision of the urban parcels as well as allowing scale variations. Compared to represent the context of the site, e.g. river, mosque and land use pattern, therefore only a single-objective was needed: maximize the total buildable area. The genes include total number of focal points, x axis spacing between the points, y axis spacing between the points. As a result of this strategy, smaller parcels are commonly preferred around the context while adapting to the geometric feature of the context. For instance, parcels follow the curvature of the meandering

243
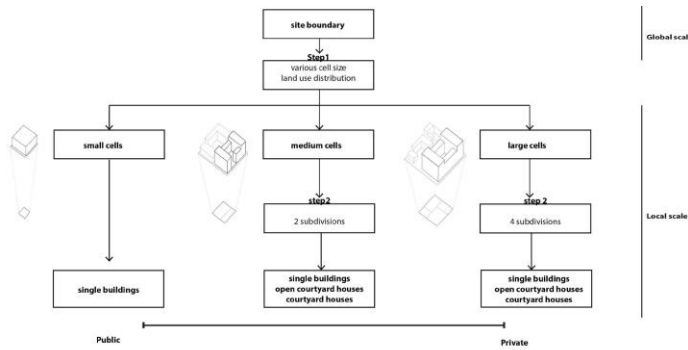
Figure 5. Experiment 2 workflow

canal. This also indicates a denser distribution at the focal points (see Figure 6).

Step 1 Results
The Genetic algorithm-generated result of 1000 solutions (10 individuals x 100 generations) converged into 4 distinctive clusters of solutions, the representative solutions from each cluster were closely scrutinized and compared. In each representative solution, one parcel type predominates. (see Figure 7) Although large parcel predominant solutions provide a larger building area, they perform less in terms of street connectivity and accessibility of each parcel. Regarding Medium parcel and small parcel predominant solutions provide a similar amount of building area. But small parcel solutions provide more open spaces, which serve as potential urban public spaces and can react to future urban growth.
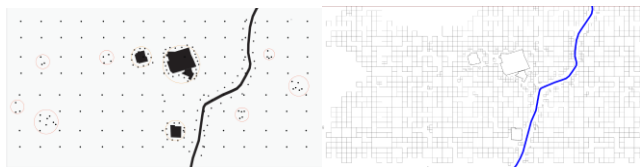
Figure 6. focal point distribution example and quad-tree subdivision
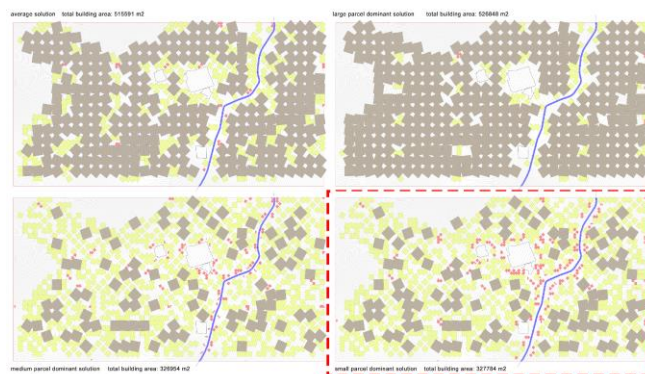
Figure 7. Step 1 GA results(top left:average solution,;top right:max larger parcel area;bottom left:max medium parcel area;bottom right:max small parcel area)

Step 2 – Superblock Generation
Following the urban parcel generative result in step 1, a selected region from step 1 was used as a base for step 2 due

to the limitation of computing power. The systematic modeling and algorithm setup were extracted and modified from experiment 1. Each parcel type was generated in parallel. Parcels were each divided into sub-blocks, subsequently creating secondary road networks within the sub-blocks. The differentiated modeling rule set was applied to the parcel based on the scale: 1, Small parcels were extruded directly as a single storey building, which serves as commercial use.2, Medium parcels were subdivided to one workshop (single building) and one medium scale courtyard house. 3, Large parcels were subdivided into 4 large courtyard houses, while the subdivision also defined the secondary roads between the courtyard houses.

To control the building scale, several Boolean condition statements were imposed in the procedural modeling definition to delete the rooms, which are out of scale. In total, a full range of building variations were introduced, each corresponding to the three programs: commercial use, crafting production use and dwelling. The parcels variations consist of: Small single building-commercial use, Medium single building-crafting & production use, Medium courtyard house-single family dwelling, Large courtyard-collective dwelling Objectives include: 1, maximize FAR. 2, maximize street area. 3, maximize south facing direct sunlight exposure (winter solstice 12pm solar vector)
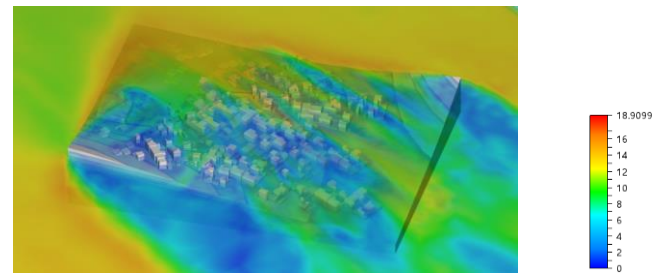
Figure 8. CFD result: air flow velocity and CFD boundary condition setup: inlet air velocity(m/s)

we examined the generated solution and performed Computational Fluid Dynamics analysis (CFD). CFD wind tunnels air inlet set up used prevailing wind direction(north west) and a proximate vertical logarithmic velocity profile to match the elevational change in real urban environment. The simulation result demonstrated clearly that the cross ventilation funnel. Two to three-story building height was in favor to minimize blockage and the stagnant air. It is also worth mentioning that the emerging larger open spaces enhanced air flow. the Building Solar gain can be indirectly measured by the solar exposure hours analysis which measures the number of hours the building can be directly exposed to the direct solar radiation. Both winter solstice and summer solstice day were analyzed to compare the urban layout and building massing configuration in relation to the solar angle. The solar study proved the dimension of the streets and courtyards resulted in adequate shaded outdoor space, which met the initial goal. However, the scale and dimension of the plazas and

244
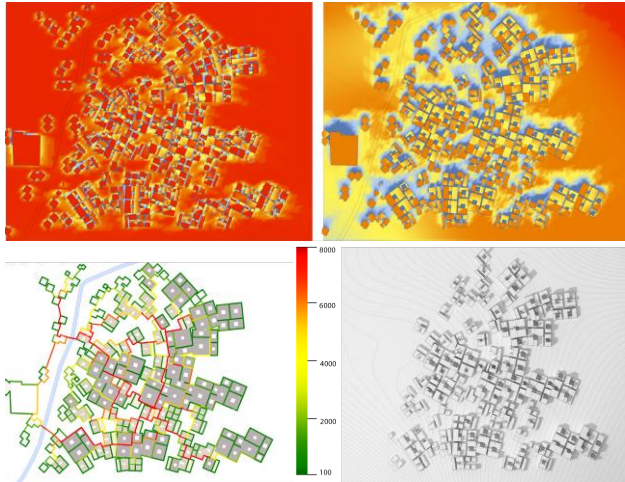
Figure 9. solar exposure hours analysis (unit: h) left: summer solstice right: winter solstice. Street network analysis (Betweenness centrality), building privacy mapping(light gray public, dark gray private)

the building configuration around them still need to be further articulated to prevent overheating.

3.1 Spatial Quality and Accessibility

As previously mentioned, the urban system generative model presented in this paper is predominantly driven by the urban superblock. As previously mentioned, the urban system generative model presented in this paper is predominantly driven by the urban superblock. To quantitatively examine the accessibility of emerging street networks, the network centrality (betweenness) analysis was conducted. From topology point of view, the street network could be abstracted as edges (street segments) and nodes(crossroads). Among all the possible shortest paths between any two nodes, 'The edge betweenness centrality is defined as the number of the shortest paths that go through an edge in a graph or network' [23] it reflects the street segments' popularity ranked by passengers and its potential to attract commercial activities. The analysis result was then superimposed with building privacy mapping. The privacy mapping took into consideration the program as well as the geometric feature on the two assumptions: 1.commercial and workshop buildings are more open compared to housing while 2. a semi-closed courtyard housing is less private than the closed courtyard house. The betweenness centrality analysis indicated the most frequently accessed streets (red) correlate to the distribution of commercial programs and smaller block subdivision. Emerging quality of the street networks exhibit similarity with the existing street pattern of Fes el Bali.

4 CONCLUSION

The computational experiments were successful in the sense that the numeric fitness values were optimized while morphologically, there were still a relatively large amount of variations. The building scale generations also covers the emergence of complex street networks and open spaces. These can be studied further. The demonstrated workflow method and approach showcase that the emergence of complex urban layout could be achieved by the bottom-up generative process. but the critical part is the rules setup, imposed control and Hierarchical order. Although the process and final solution are site-specific, the methodology and workflow could be potentially applied to a much wider variety of urban regions.

LIMITATIONS AND FURTHER STUDIES

The limitations of the process described in this paper were block dominated. The generation of the street network and the adaptations to the topography could be seen as an emergent result from the MOEA. This implies less control of street dimensions and network-hierarchy. Further research could potentially incorporate the street connectivity as a potential fitness objective to direct the MOEA towards a co-evolved urban scenario. Hereby the graph analysis in the computational model also informs the program distribution in the land use pattern as well as the block subdivision. In terms of environmental objectives, only the solar gain factor was considered in the generative process which is far from representing the reality of an urban microclimate. This suggests future research on the integration of additional environmental evaluation of wind criteria, in order to accurately evaluate the adaption of the solutions to the environment.

This design experiment would provide a valuable approach for high-density urban tissues, making use of vernacular topology, street pattern efficiency as well as applying MOEA. The presented strategy aims to prevent build cases such as the modern urban structure in Fez el Bali, deriving from the French colonization, called la Nouvelle Ville. La Nouvelle Ville shows poor adaptation to its environmental conditions. Comparing the two different urban systems in Fez el Bali, the modern urban structure Nouvelle Ville shows no relations to the Medinas urban strategies. In order to adapt urban systems to the complexity of a modern city of rapid growing density and effects of climate change scenarios incorporating vernacular building strategies as well as making use of MOEA were found as an appropriate approach and could be studied further.

REFERENCES

1. Carroll, Endless Forms Most Beautiful: The New Science of Evo Devo and the Making of the Animal Kingdom, p. 289.

2. Sean B. Carroll, 'Evo-Devo and an Expanding Evolutionary Synthesis: A Genetic Theory of Morphological Evolution', Cell, 134 (2008): 25–36, p. 29.

3. Schneider, S., et. al. 'Rethinking Automated Layout Design: Developing a Creative Evolutionary Design Method for the Layout Problems in Architecture and Urban Design' in: J.Gero, (Eds.) Design Computing and Cognition '10 pp. 367-386.

4. Sendin, P. 'The Medina of Fez: Analysis of a Superb Compact Town', 2011. http://www.patriciasendin.com/2011/09/medina-of-fez-analysis-of-superb.html.

5. Johansson, E. 'Influence of Urban Geometry on Outdoor Thermal Comfort in a Hot Dry Climate: A Study in Fez, Morocco', Building and Environment, 41 (2006): 1326–1338.

6. Essbai, S. 'In the Heart of Fes : Getting to Know the Medina', 2015. https://www.lejourna-linternational.fr/In-the-heart-of-Fes-getting-to-know-the-medina_a3032.html.

7. Bigio, A.G., and G. Licciardi. 'The Urban Rehabilitation of Medinas', 2010.

8. Kicinger, R., T. Arciszewski, and Kenneth De Jong. 'Evolutionary Computation and Structural Design: A Survey of the State of the Art', Computers & Structures, 83 (2005): 1943–78.

9. Konak, A., D.W. Coit, and A.E Smith. 'Multi-Objective Optimization Using Genetic Algorithms: A Tutorial', Reliability Engineering & System Safety, 91 (2006): 992– 1007.

10. Luke, S. *Essentials of Metaheuristics: A Set of Undergraduate Lecture Notes*. 2. Washington D.C., 2015.

11. Wright S. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. In: Proceedings of the 6th international congress on genetics, Brooklyn botanic garden, New York, 24–31 August 1932, pp. 356–366.

12. Holland JH. Outline for a logical theory of adaptive systems. J ACM 1962; 9(3): 297–314.

13. Rechenberg I (1965) Cybernetic solution path of an experimental problem. Royal Aircraft Establishment, Farnborough p. Library Translation 1122.

14. Schaffer JD. Multiple objective optimization with vector evaluated genetic algorithms. PhD Thesis. Vanderbilt University, Nashville, TN, 1984.

15. Goldberg DE. Genetic algorithms in search, optimization, and machine learning. Boston, MA: Addison-Wesley, 1989.

16. Farzaneh.Ali, and Weinstock.Michael. 21-26 October, 2019. . Mathematical Modeling of Cities as Complex Systems. ACADIA 19:UBIQUITY AND AUTONOMY [Proceedings of the 39th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA).pp. 554-563

17. Zitzler E. Evolutionary algorithms for multiobjective optimization: methods and applications. Zurich: Swiss Federal Institute of Technology, 1999..

18. Zitzler E., Laumanns M and Thiele L. SPEA2: improving the strength Pareto evolutionary algorithm, 2001,https://pdfs.semanticscholar.org/6672/8d01f9ebd044 6ab346a855a44d2b138fd82d.pdf

19. Makki M, Showkatbakhsh M. Song Y. Wallacei Primer, 2019.

20. Lawrence J. Fogel. On the Organization of Intellect. University of California, Los Angeles--Engineering, 1964.

21. J. Horn. N. Nafpliotis. D.E. Goldberg. "A niched Pareto genetic algorithm for multiobjective optimization." Proceedings of the First IEEE Conference on Evolutionary Computation. IEEE World Congress on Computational Intelligence (1991): 82-87 vol.1.

22. Vierlinger, Robert & Hofmann, Arne.. A Framework for Flexible Search and Optimization in Parametric Design,2013

23. Girvan M, Newman ME (2002) Community structure in social and biological networks. Proc Natl Acad Sci USA 99(12):7821–7826.

24. Farzaneh, Ali. Computational Morphogenesis of City Tissues. PhD thesis The Open University,2017.

25. Project Discover: An application of generative design for architectural space planning ,SimAUD 2017 Conference proceedings,2017.

26. Ho Schwartz Y. Raslan R. Mumovic D. Multi-objective genetic algorithms for the minimisation of the life cycle carbon footprint and life cycle cost of the refurbishment of a residential complex's envelope: a case study. SimAUD 2017 Conference proceedings,2015.

27. Makki, Mohammed & Showkatbakhsh, Milad & Tabony, Aiman & Weistock, Michael. Evolutionary algorithms for generating urban morphology: Variations and multiple objectives. International Journal of Architectural Computing,2018.

28. Calixto, V., Celani, G., "A literature review for space planning optimization using an evolutionary algorithm approach: 1992-2014" (2015)

The Development of Optimization Methods in Generative Urban Design: A Review

Yufan Miao[1], Reinhard Koenig[2] and Katja Knecht[3]

[1]Future Cities Laboratory, Singapore, Singapore, miao@arch.ethz.ch
[2]Bauhaus University Weimar, Weimar, Germany, reinhard.koenig@uni-weimar.de
[3]Future Cities Laboratory, Singapore, Singapore, katja.knecht@arch.ethz.ch

ABSTRACT

Design optimization, as one of the major generative design methods, has been studied ever since the 1960s. However, half a century later, in an urban design context, it is still in the research and experimental phase and rarely employed in actual projects. This paper aims to reflect design optimisation methods through a comprehensive review from the perspective of historical development, competitors' success and future trends. It also proposes a conceptual framework to enhance current optimisation methods with machine learning for generative urban design.

Author Keywords

Design Optimisation, Machine Learning, Urban Design

1 INTRODUCTION

Traditional design methodologies are based on the separation between analysis and synthesis [22]. In addition to the traditonal methods, Computer Aided Architectural Design (CAAD) systems have emerged as a design support tool in the 1970s [36], inheriting from Computer Aided Design (CAD) the capabilities for "creation, modification, analysis, or optimization of a design" [30]. To bridge the gap between analysis and synthesis, modern CAAD tools are expected to provide an integrated, domain-oriented and knowledge-based environment. When the criteria chosen for design can be quantified and expressed mathematically, the design problem can be naturally formulated as an optimization problem [5], which provides an avenue to integrate design analysis with synthesis.

Optimization, as one type of generative design methods, has been applied to architecture since 1960s [26]. However, half a century later, in an urban design context, it is still in the research and experimental phase and rarely integrated into mainstream CAD software [73]. In contrast to design optimization, rule-based modeling as a generative method achieved a greater application in practice through the success of the 3D modeling software CityEngine [20], which dominates the commercial market especially after it having been acquired by ESRI [19] and having been fully integrated

into their Geographic Information System (GIS) products. CityEngine's success originated from its strong theoretical background of shape grammar [69] and L-system [59]. In comparison to rule-based methods, design optimization has such advantages as the integration of design analysis and generation, a large solution space, and disruptive innovation [67], which are necessary for architectural and urban design problems. However, the lack of mathematical formulations for design problems and the issues of inadequate efficiency and effectiveness are shortcomings of design optimization based generative methods [44]. Therefore, a hybrid approach with machine learning could be proposed and studied to take advantage of both.

Recently, with the victory of Google's AlphaGo [28] over human professionals in the Go game, machine learning has ignited people's passion for Artificial Intelligence (AI). New generative methods such as Generative Adversarial Network (GAN) [29], and Variational Auto-Encoder (VAE) [17] have achieved great success in fields such as computer vision. However, such methods have rarely been applied to urban design problems. A remarkable exception in this aspect is the StreetGAN model [31], which applied GAN model for the generation of street networks. However, design results from machine learning are difficult to be interpreted whereas interpretability is of vital importance during the design discussion between designers and stakeholders. Although interpretable machine learning emerged as a hot topic recently [57], it is still at the early stage. In comparison with machine learning methods, design optimization results are easier to be interpreted by user-specified objective functions. However, it is difficult to accumulate design knowledge to improve the search path in future optimization. Therefore, a combination of the latest machine learning methods and optimization methods are expected to complement each other.

Due to the complexity of urban design problems themselves, Evolutionary Multi-Objective Optimization (EMO) methods are usually employed [53]. However, the performance in terms of speed and consistency is often criticized [78]. For architectural design, machine learning based surrogate models for single objective optimization problems have been proposed as an alternative [76]. Nevertheless, in urban design, multi-objective optimization problems are more popular due

to its multi-objective nature [7]. Therefore, to improve the performance, surrogate models could be introduced to partly replace the computationally intensive part of the EMO process.

Latest review papers are often written from the perspective of the authors' background and specific to one aspect of urban design [63, 21, 44]. This paper, on the contrast, aims to provide a comprehensive review from the perspective of the historical development, the current competitors, and the future trend of optimization methods in urban design. Through the review and reflections, a conceptual framework will be proposed to enhance design optimization with machine learning.

In accordance, the following questions are supposed to be answered by the paper:

- What are the advantages and challenges of design optimization in urban design?

- What could design optimization learn from the success of other methods such as shape grammar?

- How can the optimiztion methods be combined with the latest machine learning and surrogate methods to take advantages of both?

The remainder of the paper is divided into four sections, namely, the historical pathway, the peer pressure, the future trend, and finally the conclusion and outlook. In the historical pathway section, the development of design optimization in architectural and urban design is reviewed in chronological order. The major advantages and challenge questions are answered in this section. In the peer pressure section, a comparison between design optimization and other popular generative design methods is presented. What can be learnt for design optimization from other methods is discussed in this section. In the future trends section, surrogate methods and machine learning based generative methods are reviewed. A conceptual framework to enhance optimization techniques with machine learning is proposed. In the last section, the paper is summarized and an outlook on possible future work is provided.

2 THE HISTORICAL PATHWAY

As Confucius said, "Study the past, if you would define the future". An overview of the historical development (table 1) could provide us with a clear picture of the evolution of design optimization methods in architectural and urban design. From this picture, it is expected that the advantages (table 2) and challenges (table 3) of the methods can be better understood from the historical perspective. It is also interesting to see how design optimization methods and techniques evolved to provide the latest cutting edge technology to meet the constantly changing needs in architecture and urban design. The first question is to be answered in this section.

Design optimization started firstly in architectural design before it was scaled up to urban design. In academia, attempts to apply optimization methods to design problems can date back to as early as 1969 from Simon's remarkable paper the "Science of Design" in his seminal book "The Sciences of

the Artificial" [65]. During this phase, there was even no well established theory for architecture yet [75] and design optimization was motivated to be one of the numerical attempts to prove the scientificity of architecture besides its in-born artistic nature. Simon [64] also re-defined design creativity as the richness of a combinatorial space that architects move through and add one element after another to, which coincides exactly with mathematical optimization. Gero [27] reviewed a range of applications of optimization techniques in architecture and urban design. He pointed out that the lack of numeracy in architectural education and the lack of numerical models in architecture hindered the applicability of this method in design. Mitchell et al. [51] also contributed by proposing an optimization method to synthesize small rectangular floorplans. Endeavors continued for decades to develop optimization-based methods for design in the domain of CAAD. In 1980s, Gero, Radford and Balachandran [25, 2, 60] introduced multi-objective optimization in design through several publications of theirs. During this phase, design optimization did solve a few sub-problems of architectural design. However, the employed numerical optimization methods often failed when the design problems could not be mathematically formulated.

In 1990s, more logic-based AI techniques were introduced to weaken the conditions of mathematical formulation while debates about the necessity of such methods for design was heated. Pohl et al. [58] proposed a prototype of an intelligent computer-aided design system that emphasized on the partnership between computer and human. Schmitt [62] pointed out that the research frontiers of CAAD started to shift from design automation to design support and appealed for more insights from human cognition. Malkawi [46] proposed a design-oriented method to evaluate, critique and optimize energy use and design in buildings. Many efforts were made in this decade, but doubts and critiques were ubiquitous from outside the CAAD field.

Flemming [23] strongly defended the field by pointing out the four common fallacies in the critiques. They are 1) treating design as a monolithic and indivisible process, whereas design is iterative and consists of a multitude of subtasks; 2) expecting that architectural practice should be supported as is without challenging the status quo whereas current practice is not without shortcomings, which the design and development processes of CAAD can help identify and address; 3) CAAD being just an application of linguistic analogy to design; and 4) traditional design being a top-down approach of critique that relies on authorities and their theories, in particular Heidegger's, to discredit CAAD systems and approaches thereby disregarding or neglecting empirical evidence from the field, users' needs and experiences.

However, what could not be denied was that the applicability of the developed systems were still very limited. Meanwhile, critical views were also voiced internally from the field and motivated the research agenda to move forward. A famous example is the CAAD's seven deadly sins brought up by Maver [48], including macro-myopia, déjà vu, xenophilia, unsustainability, failure to validate, failure to evaluate, and

248

failure to criticize. Motivated by both external and internal critiques, new methodologies started to emerge such as using genetic programming to explore design spaces [6].

At the beginning of the 21st century, benefiting from the burgeoning computing power, more derivative-free and stochastic optimization methods were realized and applied to complex discrete nonlinear problems. Coates et al. [13] initiated several projects in his Centre for Environment and Computing in Architecture (CECA), including the use of generative algorithms to construct forms for architectural design. Michalek et al. [49] proposed an optimization method for floorplan layout design with both simulated annealing and genetic algorithm. They also proposed an interactive method for architectural layout optimization [50]. Wright et al. [79] tried to optimize building thermal design with the help of a multi-criteria genetic algorithm. Caldas & Norford [8] proposed a design optimization tool based on a genetic algorithm to optimize the environmental performance of buildings. Wetter & Wright [74] made a comparison between deterministic and probabilistic optimization algorithms for non-smooth simulation-based optimization. They drew the conclusion that stochastic or hybrid methods such as particle swarm with Hooke-Jeeves algorithm were the best choices in practice when gradient-based methods failed because of large discontinuities in the cost functions.

For urban design, Derix [16] used Ant Colony Optimization to generate street networks and Quantum Annealing to find out desired adjacencies among different land use units. Moreover, new insights about the usage of optimization methods emerged. Bleiberg & Shaviv [3] used optimization to enhance collaborative design. At the same time, researches in Multi-Objective Optimization also had breakthroughs with remarkable algorithms developed such as SPEA2 [80], NSGA-II [15], and later HypE [1]. However, although more and more physical realities were achieved with the help of CAAD, such as the Science City Zurich, design support tools failed to make significant inroad into design practice [61].

In the past decade, research progress continued in design optimization within the field of CAAD, extending from architectural design to urban design. The project KAISERSROT is a successful example where CAAD was massively used in architectural and urban design practice [68]. In this field, many researches were focused on design space exploration. For design exploration, Janssen [35] proposed an evolutionary system. Turrin et al [71] developed a method with combined parametric modelling and genetic algorithms for design exploration of performance driven geometries. Stouffs [70] also proposed methods to combine generative with evolutionary exploration. Meanwhile, model-based optimization has been proved to be a more efficient alternative to evolutionary algorithms in terms of speed and feasibility [77] when the problem is or can be re-formulated as a single objective optimization problem. For Multi-Objective Optimization problems, hybrid methods with both metaheuristic and model-based optimization would be ideal as was already proved in other engineering design fields [66].

Objects	Developments
1960s-1980s	early ontological attemps; single objective optimization
1990s	methods further developed; logic-based, human-centered methods
2000s	more complex nonlinear problems; meta-heuristic optimization methods; multi-objective optimization methods; simulation-based methods
2010s	data-driven AI methods

Table 1. The historical development of the design optimisation methods.

	Advantages
1960s-1980s	solve sub-problems of architectural design; prove the scientifity of architecture
1990s	attracted more attentions from design field; new methods attempted at the application level
2000s	computing power for more complex problems; applications emerged for urban design
2010s	AI technologies augment human creativities; availability of large amount of data; more statistical models; successful application to design practice

Table 2. Advantages of the optimisation methods in different ages

	Challenges
1960s-1980s	lack of well defined mathematical models
1990s	debates from both the internall and external of the CAAD community
2000s	theoretically thriving but practically immature
2010s	mismatch between the data-driven approaches and the design logic and philosophy

Table 3. Challenges of the optimisation methods in different ages

With the development of spatial analysis techniques such as space syntax [33], an increasing number of quantitative evaluatin methods have been introduced to urban design, which enrich the design requirements design optimization could fulfill. Celani et al. [11] proposed a method to generate urban patterns by combining shape grammars and genetic algorithms. Cao et al [9] used Multi-Objective Optimization methods for land use planning. Koenig [39, 40] proposed a synthesis method for street network and building layouts based on EMO. Moreover, Koenig & Schmitt [41] also proposed the Cognitive Design Computing (CoDeC) framework. Different from earlier design optimization attempts, their methods not only aim at synthesizing urban geometries but also target at tackling a major critique of computational creativity, which is the lack of humanity [14] by reinforcing the role of human intelligence through designers' interactions with the generated urban design. The core of their approach is the EMO-based generative methods which could bring to urban design benefits such as transparency [52] and integrativeness [67]. Moreover, this research aims to address one of the major challenges in EMO for generative urban design, namely the representation problem. Their recent work [37] has successfully applied EMO to the generation of multiple urban design layouts with urban objects such as street network, blocks, parcels and buildings. Despite successful work crossing research fron-

tiers, in practice, there is still a long path to go. The lack of quantified design evaluation measures and metrics is still a big challenge. Moreover, computer generated design solutions are usually simple and only suitable for prototyping. Due to the nature of EMO, highly intensive computation is usually required, even for simple generation, which lags behind real-time. These problems are expeced to be solved with a hybrid approach. The following section will feature a comparison between design optimization and other existing generative design methods.

3 THE PEER PRESSURE

As Confucius said, "When I walk along with two others, they may serve me as my teachers". As one kind of generative design method, design optimization has a great potential to form the basis of future AI-based integrated design support tools. For the time being, it still has a lot to improve to meet the expectations, which can be learnt from its peer generative methods. Generative design methods have already been comprehensively reviewed by Singh and Gu [67]. In their review paper, they identified five major types of generative methods, namely, Cellular Automata, Shape Grammars, L-Systems, Swarm Intelligence, and Genetic Algorithms. This paper was written over 7 years ago and thus does not cover the latest technologies such as machine learning based methods. This section aims to compare existing generative design methods and provide some insights about what can be learnt for design optimization from other generative methods.

Cellular Automata has been successfully applied to architectural and urban design in many aspects [32, 42]. However, due to its bottom-up nature, the outcomes are often complex and difficult to predict. Moreover, to define rules, which are required to guide the generative process, is also unfamiliar to design practitioners. In recent research publications, interests in this method have already faded. On another hand, Swarm Intelligence has also been applied to architectural and urban design applications. The popularly used Agent-based model (ABM) embodies the essence of this method. One famous example is the MATSim [34] as a successful ABM system for transportation planning. However, these kind of models are more suitable for social and collective behavior studies [10, 45]. This can be employed as one dimension of the urban design task but the whole design should go beyond that. For design optimization, agent-based models can be used as for design evaluation and the visualisation of the simulations could help a designer to better understand the future behaviors and movements of people in the designed space.

Both Shape Grammar as well as L-System constitute the so-called rule-based procedural modeling techniques [56], which have been applied for the generation of both architectural and urban design. From shape grammar, two major branches have emerged, namely, Grammars for Designing (GfD), and Grammars Derived from Designs (GDfD) [47].

GDfD is usually achieved by evaluating the rules and shapes derived from the analysis and then selecting the suitable ones for the generation of a new design. Successful studies have been conducted by Duarte et al [18] and Paio et al [55]. Both of them tried to introduce optimization into the generative

process as well, which is a way to combine design optimization and procedural models.

However, to date GfD have achieved the greatest level of success and adoption in practical architectural and urban design. A striking example is the 3D modeling software CityEngine [20]. Due to its integration with GIS products, CityEngine provides a convenient platform for designers by combining location-based data with capabilities for design generation allowing buildings to be modeled with high visual quality and level of detail. CityEngine originally evolved from the first school with the shape grammar of Computer-Generated Architecture (CGA). The limited capabilities of complex shape grammars to support user interaction have in a recent version of CityEngine been solved by introducing interactive design capabilities [20]. However, there is still a drawback that originates from the limitation of GfD approaches in general. The generated designs are usually independent from the urban context rather than from existing designs. They are more suitable for use in a large variety of contexts rather than those with strong local characters. The success story of CityEngine, reasons for its adoption in design practice but also its remaining shortcomings, can help inform future AI-based integrated design support tools.

In comparison to shape grammar, design optimization methods themselves involve highly complex and computationally intensive processes and thus are not ideal for fast and large scale design generation with limited computing power. The advantage design optimization has is the so-called intelligence as a mimic of the human design process. However, the current state of such intelligence is still primitive. Moreover, it is still debatable whether such intelligence is necessary and how the roles human and computer should be defined respectively in the design process. Shape grammar based approaches, on another hand, focus on the generation of large amount of complex 3D geometries which provide users a comprehensive visual thinking environment without interfering with the actual design process. Although this is not aligned with the current heat waves of AI, its success has proved that it is currently the best engineered solutions for design. Design optimization should at least catch up with the speed and scale of shape grammar based software such as CityEngine to be superior. This is also why hybrid methods with machine learning are necessary and will be explained in the following section.

4 THE FUTURE TRENDS

With the success of AI researches, machine learning based approaches are drawing increasing interest from academia. For generative design, these methods are applied in general in two ways. The first way is through model-based optimization whereas the second is through generative models.

For model-based optimization, researchers use machine learning methods to construct the surrogate models to approximate the behaviors of simulations. Remarkable researches in this aspect are from Wortmann [76, 77, 78] where surrogate model based methods were applied to architectural design. Based on his publications, this kind of methods are superior because of the reliable convergence rate and relatively fast

Objects	Advantages
CityEngine	fast generation speed; large scale; GIS data interoperability; 3D visualization and interaction
Design Optimisation	integration of analysis and generation; large number of design alternatives; the emergence of novel solutions

Table 4. A comparison of the advantages of CityEngine and Design Optimisation.

	Disadvantages
CityEngine	separation of analysis and generation; separation of human and computer; difficult to generate novel solutions
Design Optimisation	slow; complex; limited scale; rudimentary results

Table 5. A comparison of the disadvantages of CityEngine and Design Optimisation.

speed. It is also a method popularly employed in energy-driven architectural and urban design [54, 63] where expensive simulations such as computational fluid dynamic simulation is usually popular. This kind of methods are usually applied to single objective optimization problems. For multi-objective optimization problems, different objectives are usually mathematically combined and transformed into a single objective problem. The transformation itself is a difficult task and usually relies on expert knowledge. For urban design, since the objectives can be of a large number, the transformation of multi-objective problems into single objective can be tricky especially when they are correlated to each other.

For generative models, the most popular ones are GAN and VAE. They have achieved great success in the research field of Computer Vision (CV). Therefore, most of the variants of these two methods are designed for image data. GAN model is based on the philosophy of zero-sum game [43], which contains a generator and a discriminator. The former is responsible for generating data with the model learnt from data whereas the latter tells whether the generated data is close to the real data. The model is optimized continuously until the discriminator can hardly tell the difference between generated data and real data. For design generation, streetGAN as a variant of GAN has been successful [31]. This kind of researches have been seldomly employed in urban design but exhibited potential to enhance the current design optimization methods. For single urban geometries such as street networks, streetGAN has already provided a successful example. Similar approaches could be possibly applied to blocks, parcels and buildings. What would be interesting is how the machine learning method could capture the inter-relationship between different types of urban geometries such as that of street network and parcels. VAE, on another hand, adopted a different philosophy. It is to reduce the dimension of the original data and represent the data as an encoder with reduced but key features. Then it uses a decoder to re-generate the original data space. The encoder is optimized through back propagation of the error between original data and generated data. In this process, novel data can also be generated [4], which is important for design. A similar approach that reduces the dimension of data and only depicts key features is the Self-Organizing Map (SOM) which has been proved to be useful for architectural and urban design [24].

Concepts about employing machine learning methods for a recommendation system for urban design were also proposed by Chirkin and Koenig [12]. In their framework, unsupervised learning methods were used to create design quality measures extracted from user preferred designs. These measures were later employed to improve the design draft through optimisation. This research inspired the proposal of a new framework as in figure 1. Similarly, the purpose is to enhance optimization methods with a learning capability. However, different from Chirkin and Koenig's approach, this new framework not only tries to provide alternative methods to optimisation for design generation when different user specifications are input but also tries to avoid excessive usage of optimisation methods when unnecessary. Moreover, as has been found by von Buelow [72], there is a selection problem for the users to choose the right option from massive generated design variants. Also, a performance optimized solutions might not be aesthetically pleasant. For the first problem, SOM as a machine learning method, has been employed as an attempt to reduce the solution space [38]. For the second problem, solutions such as Opossum developed by Wartmann [77] were proposed to provide the designers a visualisation interface to compare the geometries and the Key Performance Indices (KPI) of both manual designs and generated designs. For both problems, machine learning methods are potentially helpful to learn from the users' selection as feedbacks to guide the future generative process.

There are five ways of inputs expected from the users. They are quantitative design objectives, parametric design specifications, geometries, maps or design proposals in image format, and semantic specifications. Quantitative design objectives are well-formulated objective functions and constraints that can be directly input into the optimization process to guide the design generation which is the traditional way of implementing design optimization. The parametric design specifications are expected to be used by the rule-based generators for fast and large scale urban geometry generation. The generated geometry can either be outputed directly or improved further by EMO with specified objective functions after being converted into a genotype in the data structure encoder. The framework should also be able to accept already existing designs in the format of either geometries or images. The geometries can be decoded to retrieve similar genotypes for evolutionary algorithm based geometry generation or similar statistical models for machine learning based image generation. The data assignment and retrieval tasks are handled by a database management system. Moreover, the generated design solutions can be manually reviewed and selected by designers with the selected ones being stored in the database.
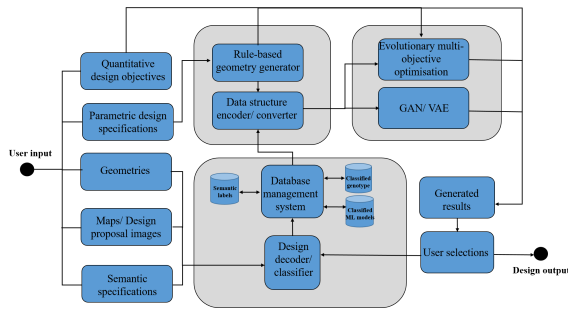
Figure 1. A conceptual framework to integrate design optimization with machine learning (the groups in grey color are expected to run in the background).

5 CONCLUSION AND OUTLOOK

In this paper, key advantages and challenges for design optimization in architectural and urban design have been reviewed from a historical perspective. Moreover, as a generative design method, design optimization and other kinds of generative design methods are compared. What design optimization could learn from other methods is also proposed. Furthermore, the shortcoming of the optimization methods could be enhanced by the latest data-driven approaches such as machine learning. How this could be achieved is explained in the proposal of a conceptual framework.

As this paper constitutes a literature review of different methods and is therefore set at a conceptual level, a concrete description of a possible implementation of the framework is beyond its scope. The implementation of the proposed framework poses several challenges, such as how the data structure encoder be implemented in the proposed framework and how the hybrid database should be designed for the storage of different types of data. These questions need to be answered in the future research and the viability of the conceptual framework also needs to be affirmed by actual implementations.

REFERENCES

1. Bader, J., and Zitzler, E. Hype: An algorithm for fast hypervolume-based many-objective optimization. *Evolutionary computation 19*, 1 (2011), 45–76.

2. Balachandran, M., and Gero, J. A knowledge-based approach to mathematical design modeling and optimization. *Engineering Optimization+ A35 12*, 2 (1987), 91–115.

3. Bleiberg, T., and Shaviv, E. Optimisation for enhancing collaborative design. In *Proceedings of the Building Simulation* (2007).

4. Bozkurt, A., Esmaeili, B., Brooks, D. H., Dy, J. G., and van de Meent, J.-W. Can vaes generate novel examples? *arXiv preprint arXiv:1812.09624* (2018).

5. Braibant, V., and Sander, G. Optimization techniques: synthesis of design and analysis. *Finite elements in analysis and design 3*, 1 (1987), 57–78.

6. Broughton, T., Tan, A., and Coates, P. The use of genetic programming in exploring 3d design worlds. In *CAAD futures 1997*. Springer, 1997, 885–915.

7. Bruno, M., Henderson, K., and Kim, H. M. Multi-objective optimization in urban design. In *Proceedings of the 2011 Symposium on Simulation for Architecture and Urban Design*, Society for Computer Simulation International (2011), 102–109.

8. Caldas, L. G., and Norford, L. K. A design optimization tool based on a genetic algorithm. *Automation in construction 11*, 2 (2002), 173–184.

9. Cao, K., Batty, M., Huang, B., Liu, Y., Yu, L., and Chen, J. Spatial multi-objective land use optimization: extensions to the non-dominated sorting genetic algorithm-ii. *International Journal of Geographical Information Science 25*, 12 (2011), 1949–1969.

10. Carley, K. Sociology: computational organization theory. *Social Science Computer Review 12* (1994).

11. Celani, G., Beirão, J. N., Duarte, J. P., and Vaz, C. Optimizing the "characteristic structure": Combining shape grammars and genetic algorithms to generate urban patterns.

12. Chirkin, A. M., and König, R. *Concept of interactive machine learning in urban design problems*. ACM, 2016.

13. Coates, P., Appels, T., Simon, C., and Derix, C. Current work at ceca. In *Proceedings of the 4th Generative Art Conference (GA2001)* (2001).

14. Colton, S., Cook, M., Hepworth, R., and Pease, A. On acid drops and teardrops: Observer issues in computational creativity. In *Proceedings of the 7th AISB Symposium on Computing and Philosophy* (2014).

15. Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE transactions on evolutionary computation 6*, 2 (2002), 182–197.

16. Derix, C. In-between architecture computation. *International journal of architectural computing 7*, 4 (2009), 565–585.

17. Doersch, C. Tutorial on variational autoencoders. *arXiv preprint arXiv:1606.05908* (2016).

18. Duarte, J. P., Rocha, J. M., and Soares, G. D. Unveiling the structure of the marrakech medina: a shape grammar and an interpreter for generating urban form. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing 21* (2007).

19. ESRI. Esri acquires 3d software company procedural, 2011. Retrieved from: https://www.esri.com/news/releases/11-3qtr/esri-acquires-3d-software272company-procedural.html.

20. ESRI. CityEngine, 2019. http://www.esri.com/software/cityengine.

21. Farahani, R. Z., Miandoabchi, E., Szeto, W. Y., and Rashidi, H. A review of urban transportation network design problems. *European Journal of Operational Research 229*, 2 (2013), 281–302.

22. Fischer, G., and Nakakoji, K. Beyond the macho approach of artificial intelligence: empower human designers—do not replace them. *Knowledge-Based Systems 5*, 1 (1992), 15–30.

23. Flemming, U. Get with the program: common fallacies in critiques of computer-aided architectural design. *Environment and Planning B: Planning and Design 21*, 7 (1994), S106–S116.

24. Fuhrimann, L., Moosavi, V., Ohlbrock, P. O., and D'acunto, P. Data-driven design: Exploring new structural forms using machine learning and graphic statics. In *Proceedings of IASS Annual Symposia*, vol. 2018, International Association for Shell and Spatial Structures (IASS) (2018), 1–8.

25. Gero, J., and Radford, A. The place of multi-criteria optimization in design. *Design Theory and Practice, London: The Design Council* (1984), 81–85.

26. GERO, J. S. Architectural optimization-a review. *Engineering Optimization 1*, 3 (1975), 189–199.

27. Gero, J. S. Architectural optimization-a review. *Engineering Optimization 1*, 3 (1975), 189–199.

28. Gibney, E. Google ai algorithm masters ancient game of go. *Nature News 529*, 7587 (2016), 445.

29. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in neural information processing systems* (2014), 2672–2680.

30. Groover, M., and Zimmers, E. *CAD/CAM: computer-aided design and manufacturing*. Pearson Education, 1983.

31. Hartmann, S., Weinmann, M., Wessel, R., and Klein, R. Streetgan: towards road network synthesis with generative adversarial networks.

32. Herr, C., and Kvan, T. Using cellular automata to generate high-density building form. In *Computer aided architectural design futures 2005* (2005).

33. Hillier, B. *Space is the machine: a configurational theory of architecture*. Space Syntax, 2007.

34. Horni, A., Nagel, K., and Axhausen, K. W. *The multi-agent transport simulation MATSim*. London: Ubiquity Press, 2016.

35. Janssen, P. H. An evolutionary system for design exploration.

36. Kalay, Y. E. *Architecture's new media: Principles, theories, and methods of computer-aided design*. MIT Press, 2004.

37. Koenig, R., Miao, Y., Knecht, K., Aichinger, A., and Konieva, K. Integrating urban analysis, generative design, and evolutionary optimization for solving urban design problems. *Environment and Planning B xx*, xx (2020).

38. Koenig, R., Standfest, M., and Schmitt, G. Evolutionary multi-criteria optimization for building layout planning-exemplary application based on the pssa framework.

39. Koenig, R., Treyer, L., and Schmitt, G. Graphical smalltalk with my optimization system for urban planning tasks.

40. König, R. Urban design synthesis for building layouts based on evolutionary many-criteria optimization. *International Journal of Architectural Computing 13*, 3-4 (2015), 257–269.

41. König, R., and Schmitt, G. Backcasting and a new way of command in computational design. In *CAADence in Architecture*, ETH-Zürich (2016).

42. Krawczyk, R. Experiments in architectural form generation using cellular automata. In *eCAADe* (2002).

43. Luce, R. D., and Raiffa, H. *Games and decisions: Introduction and critical survey*. Courier Corporation, 1989.

44. Machairas, V., Tsangrassoulis, A., and Axarli, K. Algorithms for optimization of building design: A review. *Renewable and sustainable energy reviews 31* (2014), 101–112.

45. Macy, M., and Willer, R. From factors to actors: computational sociology and agent-based modeling. *Annual Review of Sociology 28* (2002).

46. Malkawi, A. M. Building energy design and optimization: intelligent computer-aided thermal design.

47. Mandić, M., and Tepavǒ, B. Analysis of shape grammar application as a tool for urban design. *Environment and Planning B 42* (2015).

48. Maver, T. W. Caad's seven deadly sins. In *Sixth International Conference on Computer-Aided Architectural Design Futures* (1995), 21–22.

49. Michalek, J., Choudhary, R., and Papalambros, P. Architectural layout design optimization. *Engineering optimization 34*, 5 (2002), 461–484.

50. Michalek, J., and Papalambros, P. Interactive design optimization of architectural layouts. *Engineering optimization 34*, 5 (2002), 485–501.

51. Mitchell, W. J., Steadman, J. P., and Liggett, R. S. Synthesis and optimization of small rectangular floor plans. *Environment and Planning B: Planning and Design 3*, 1 (1976), 37–70.

52. Monizza, G. P., Bendetti, C., and Matt, D. T. Parametric and generative design techniques in mass-production environments as effective enablers of industry 4.0 approaches in the building industry. *Automation in Construction 92* (2018), 270–285.

53. Nagy, D., Villaggi, L., and Benjamin, D. Generative urban design: integrating financial and energy goals for automated neighborhood layout. In *Proceedings of the Symposium on Simulation for Architecture and Urban Design*, Society for Computer Simulation International (2018), 25.

54. Nguyen, A.-T., Reiter, S., and Rigo, P. A review on simulation-based optimization methods applied to building performance analysis. *Applied Energy 113* (2014), 1043–1058.

55. Paio, A., Reis, J., Santos, F., Lopes, P. F., Eloy, S., and Rato, V. Emerg.cities4all: towards a shape grammar based computational system tool for generating a sustainable integrated urban design. In *Proceedings of the 29th eCAADe Conference* (2011).

56. Parish, Y., and Müller, P. Procedural modeling of cities. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, ACM (2001).

57. Pena, F. C., and Guerra-Gomez, J. A. Opening the black-box: Towards more interactive and interpretable machine learning. In *Proceedings of the Machine Learning from User Interaction for Visualization and Analytics Workshop at IEEE VIS* (2018).

58. Pohl, J., Myers, L., and Chapman, A. Icads; an intelligent computer-aided design environment. *ASHRAE Transactions (American Society of Heating, Refrigerating and Air-Conditioning Engineers);(United States) 96*, CONF-9006117– (1990).

59. Prusinkiewicz, P., and Lindenmayer, A. *The algorithmic beauty of plants*. Springer Science & Business Media, 2012.

60. Radford, A. D., and Gero, J. S. *Design by optimization in architecture, building, and construction*. John Wiley & Sons, Inc., 1987.

61. Schmitt, G. The impact of computer aided architectural design on physical reality. *International Journal of Architectural Computing 2*, 1 (2004), 31–41.

62. Schmitt, G., and Oechslin, W. Computer aided architectural design futures. In *CAAD Futures*, vol. 91 (1992).

63. Shi, Z., Fonseca, J. A., and Schlueter, A. A review of simulation-based urban form generation and optimization for energy-driven urban design. *Building and Environment 121* (2017), 119–129.

64. Simon, H. A. Style in design. *Spatial synthesis in computer-aided building design 9* (1975), 287–309.

65. Simon, H. A. *The sciences of the artificial*. MIT press, 2019.

66. Sindhya, K., Miettinen, K., and Deb, K. A hybrid framework for evolutionary multi-objective optimization. *IEEE Transactions on Evolutionary Computation 17*, 4 (2012), 495–511.

67. Singh, V., and Gu, N. Towards an integrated generative design framework. *Design studies 33*, 2 (2012), 185–207.

68. Spyridonidis, K.-V., and Vogiatzaki, M. Teaching and experimenting with architectural design. Tech. rep., European Association for Architectural Education, 2008.

69. Stiny, G., and Gips, J. Shape grammars and the generative specification of painting and sculpture. In *IFIP Congress (2)*, vol. 2 (1971).

70. Stouffs, R., and Rafiq, Y. Generative and evolutionary design exploration. *AI EDAM 29*, 4 (2015), 329–331.

71. Turrin, M., Von Buelow, P., and Stouffs, R. Design explorations of performance driven geometry in architectural design using parametric modeling and genetic algorithms. *Advanced Engineering Informatics 25*, 4 (2011), 656–675.

72. Von Buelow, P. Paragen: Performative exploration of generative systems. *Journal of the International Association for Shell and Spatial Structures 53*, 4 (2012), 271–284.

73. Walmsley, K., and Villaggi, L. Generative urban design, 2019. Retrieved from: https://medium.com/autodesk-university/generative-urban-design-43f4ea48039d.

74. Wetter, M., and Wright, J. A comparison of deterministic and probabilistic optimization algorithms for nonsmooth simulation-based optimization. *Building and Environment 39*, 8 (2004), 989–999.

75. Widdowson, W. *Architecture as art: a phenomenological theory of architectural esthetics*. University of Pennsylvania, 1971.

76. Wortmann, T. Model-based optimization for architectural design: Optimizing daylight and glare in grasshopper. *Technology— Architecture+ Design 1*, 2 (2017), 176–185.

77. Wortmann, T. Opossum-introducing and evaluating a model-based optimization tool for grasshopper.

78. Wortmann, T. Genetic evolution vs. function approximation: Benchmarking algorithms for architectural design optimization. *Journal of Computational Design and Engineering 6*, 3 (2019), 414–428.

79. Wright, J. A., Loosemore, H. A., and Farmani, R. Optimization of building thermal design and control by multi-criterion genetic algorithm. *Energy and buildings 34*, 9 (2002), 959–972.

80. Zitzler, E., Laumanns, M., and Thiele, L. Spea2: Improving the strength pareto evolutionary algorithm. *TIK-report 103* (2001).

Hallucinating Cities - A Posthuman Design Method based on Neural Networks

Matias Del Campo[1], Sandra Manninger[1] and Alexandra Carlson[2]

[1]Taubman College of Architecture and Urban Planning,
University of Michigan
Ann Arbor, USA
{mdelc, samdcma}@umich.edu

[2]Michigan Robotics,
University of Michigan
Ann Arbor, USA
askc@umich.edu

ABSTRACT

The main aim of this paper is to demonstrate and interrogate a design technique based on deep learning. The discussion includes aspects of machine learning, 2D to 2D style transfers and generative adversarial processes. The paper examines the meaning of agency in a world where decision making processes are defined by human/machine collaborations (fig.1), and their relationship to aspects of a Posthuman design ecology. Taking cues from the language used by experts in AI such as Hallucinations, Dreaming, Style Transfer and Vision, the paper strives to clarify the position and role of Artificial Intelligence in the discipline of urban design.

Author Keywords

Artificial Intelligence, Neural Networks, Agency, Urban Design, Machine Vision, Machine Learning, Bias, Machine Hallucinations

1 INTRODUCTION

The Map. This icon of urban planning goes far beyond its mere meaning as an abstraction that allows to execute in a controlled manner the materialization of matter and space. It rather represents a vast collection of possible solutions for urban Problems. Considering the gigantic amount of data that a collection of maps spanning more than three millenia[1] represents It appears almost evident to use this enormous repository of urban imagination in the age of big data. A quick search on Google yields 8.920.000.000 images (Yes, that's almost nine billion results!) tagged *map*. The enormous vault that the discipline of urban planning has generated throughout the ages forms THE natural resource of our discipline, waiting to be mined and processed – not to copy or imitate existing urban design solutions, but to find bespoke solutions to specific problems. Urban planners and the respective students learn to differentiate urban textures through visual stimuli, i.e. seeing hundreds and thousands of images of specific maps of cities and projects in order to recognize planning styles[2] later. They learn to differentiate for example between Gothic, Renaissance, Baroque and Modern urban conditions through memorizing geometrical features and patterns. Neural networks (NNs), which are complex functions

loosely based on the structure of the human brain, learn to perform visual tasks in a similar way, by both identifying and extracting meaningful texture and patterns within their input. This paper presents a possible application of a neural network-based image editing technique, called Neural Style Transfer[18], to mesh not only low-level pixel patterns, but also higher-level geometric features, like roads, buildings, etc., between urban maps in an effort to create cities with novel styles. The research on this possibility started as a simple experiment for style transfer between maps in order to explore the opportunities as a design method[3].

Figure 1. Results of 2D to 2D Style transfers based on Nolli plans and an image of the Moon: aspects of Estrangement and Defamiliarization profoundly speak in those results about a design ecology in a Posthuman era.

It is almost impossible to judge maps on a purely pragmatic level. They always simultaneously talk about planning processes, economic environments, material preferences, political conditions and stylistic fashions of the time the urban design was created. Wither this be in the rigorous structure and geometrical purity of Renaissance Ideal cities, as exemplified in the concept of the *ideal town* as proposed by Leon Batista Alberti in *De Re Edificatore[4],* or in the intricate voluptuous geometry of parametrically designed settlements such as Zaha Hadid's *Kartal Masterplan* for Instanbul[5]. In both cases it is not

surprising that the intrinsic matter of urban planning in a large scale involves aspects of ideology and utopia. Both examples mentioned above can be identified as representatives of ideologies that span areas beyond shape and geometry and involve political, social and economic conditions[6]. It might not surprise that in this extent they also represent a vessel and repository of the history of urban planning imaginations, and as such can be considered an enormous mine for new ideas on the nature of the city. Traditionally urban planners are trained during their studies to operate like data miners. Every new project is based on the hundreds and thousands of images ingested during the training received in architecture school. This image-based tradition is exploited in the 2D to 2D style transfer approach presented in this paper. However, it is not only about mining.

2   SCRAPING & PATTERNS

What goes beyond the ability to simply ingest imagery, is the inherently human ability to perform pattern recognition. One of the aspects the human mind is particularly avid about, is to recognize events and objects, separate fore- and background. The ability to even recognize that an error or mistake inhabits the potential for a creative solution to a problem[7]. How can this, computationally rather difficult to grasp problem be harnessed to achieve image to image style transfer? This is where the aspects of the neural network's learned features, or what it has learned are salient pixel patterns within a given image, come into play. We can use trained neural networks to successfully quantify and define textures within images, and in the context of urban maps, we can create a 'city texture' and   hallucinate[8] its specific features in other images of city plans.

There are two main hurdles that need to be taken to successfully apply this technique to urban planning processes. On the one side is the database. What is the Neural Network working with? A couple of lines of code used to collect a dataset of images used in this paper can describe the process that the authors applied:

```
With webdriver.Chrome(executable_path=driver_path)
as wd:
  res = fetch_image_urls(search_term,
        number_images, wd=wd,
        sleep_between_interactions=0.5)
```

The code scrapes the internet for images with particular labels as a first step to create a database as source for any form of style transfer, dreaming or hallucination. This is how we can tap into the existing resources of our own discipline in order to create novel outcomes. The second important aspect in this process is the training of the neural network. Let me mention here a popular example to explain a method of training a neural network. Although handwritten checks are slowly fading out, there

are still massive amounts of them written on a daily basis. In order to facilitate and speed up the process the reading and identifying the written numbers has been handed over to trained neural networks some time ago. NN's take in images of handwritten numbers as input and performs complex thresholding operations on the images' pixels to filter out relevant visual information (e.g., edges, curves), which it uses to ultimately form a prediction for what number is in the image. The accuracy of the network's prediction is evaluated using an error function and a ground truth label for the number captured by the image. The training of the networks starts with human intervention, in that a human tag the images, identifying the numbers (fig.2), as well as in how the error measurement is defined. How else should a machine learn what a six is, an eight, a nine, a zero etc.? How could an algorithm learn the large variety in hand writing styles that can drastically alter the appearance of a four? After being present thousands of hand-written examples, the Neural Network becomes better and better in understanding what the individual numbers are through an autonomous learning process (fig.3). Presto! Automated check reading.
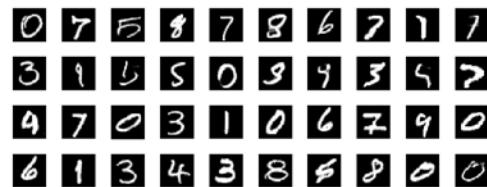
Figure 2. With caption below, be sure to have a good resolution image (see section 5.1 for image preparation instructions).
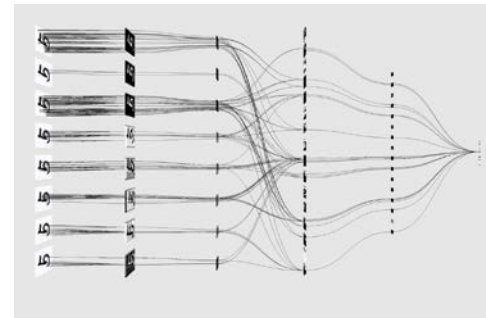
Figure 3. Sketch of a simple neural network to read numbers on a check. The network has 4 layers that allow it to increase the precision in reading the correct numbers. Please refer to the Methods section for more detail.
Image: James Le

Neural networks can be trained to perform much more complex classification tasks, such as differentiating between architectural styles. It is key to collect a large dataset that captures all possible variance of such styles so the NN can learn accurate and representative visual patterns for each class. The dataset used to train a neural network can be thought of as the 'world' the network exists in. We can then use these learned patterns as a way mathematically

represent an image by decomposing it into its base spatial/geometric, style or texture features. For the author's the most exciting discovery in applying this technique is that by manipulating the weights/impact of style and spatial imagery the results produce unexpected, atmospheric and profoundly *other, defamiliarized and estranged* results. Estranged in a good way[9].

2.1 Tipping the Hat to Neuroscience

It is quite fascinating how computer science has adopted the vocabulary of neuroscience to explicate the processes invoked in NN, and the proximity of this language to the wording of architecture when it comes to the imagination of the discipline. Terminology like *Vision* and even *Dreaming* and *Hallucinating* made regular appearance in the manifest heavy postmodern era[10] of the 1970ies, such as in the written works of Hans Hollein (Metaphor and Metamorphosis exhibition) and it was chronicled extensively in Gunther Feuerstein's oeuvre the terminology still provokes the spirit of particularly advanced architecture, albeit in a certain romantic and poetic fashion – which this paper is not about. The plot twist here being how this relationship to the terminology is currently being reinvented for the architecture and urban planning discipline by a series of young practitioners and their allies in computer science and robotics[11]. Instead of adopting the term as a metaphor, it rather borrows the terminology from computer science and more specifically from machine vision research, which has its focus on developing Neural Network solutions for example for autonomously driving cars. The discipline of computer science themselves borrowed the terms *hallucination* and *dreaming* from Neuroscience who developed this terminology in order to explain the behavior of common neurochemical mechanisms and the phenomenological similarities between human dreams and drug-induced hallucinations[12]. In this light it can be stated that a neurochemical mechanism and the synthetic ecology created with computational Neural Networks share similar traits and are closely related, thus the conversation in this paper on *Dreaming*, *Vision* and *Hallucination* in regards of *Imaginary Maps.* Literally discussing machines hallucinating possible solutions. In this paper we lay the ground for a fascinating possibility: a computational method to train neural networks to learn and recognize a variety of urban features, styles and aspects and possible ways to get neural networks to generate novel planning solutions. Another possible application for this approach is the possibility to create an app that is able to analyze urban plans and check them for errors – for example their accordance to code, their energy consumption, or their functionality. The approach however offers an entire set of possibilities that go beyond its application as a mere tool for optimization, thus provoking questions pertaining to the nature of creativity, agency and posthuman culture (Fig.4).

Figure 4. Result of Style transfer between a Dataset of Nolli maps of known cities (Rome, Barcelona, Manhattan, Washington DC) and a 19th century science plate depicting a detail of the moon's surface. These attempts can be considered experiments in Posthuman sensibilities. Can AI's be creative?

In contrast to the approach of other practices and individual researchers working within this paradigm, such as XKool (Wanyu He), Shao Zhang (PennDesign) and MetroDataTech (Tang Ge) -which primarily rely on finding engineering and pragmatic solutions to architectural problems- the approach of the authors is acutely aware of the cultural and discursive dimension of the proposed approach. It is clear, that a conference paper might not be sufficient in length to cover the entirety of the implications in regards of architecture theory within a novel paradigm, thus the authors would like to apologize for the occasional brevity in the argument.

To further lay out the difference of the approach of the authors, and the beforementioned companies and researchers, we would like to propose the following:

There are two main paths of inquiry and critical interrogation: the technical expertise necessary to apply neural networks successfully to obtain comprehensive results in pragmatic problems, such as plan optimization, landscape optimization and the ecological footprint of the design. All of which can be described as tamed problems, dealing primarily with highly specified engineering problems. On the other end of the spectrum AI allows to explore the wicked part of architectural design as well, pertaining to aspects of morphological studies, creativity, style and mood.

In the course of the research conducted by the authors investigating the implementation of AI based algorithms into planning processes the authors made a crucial discovery. In the beginning the sentiment, and prejudice, was that AI can generate everything on its own. It became very quickly very obvious that this is not the case. As described above Neural Networks rely on initial human training to do anything. It is not said however that once sufficient Neural Networks have been trained by humans, they might be able to solve problems entirely autonomous. In the case described in this paper, the notion's focus was on the ability of NN to develop morphologies of architecture entirely independent and divorced from human agency. It did not take long to understand that AI faces a great amount of stereotypical ideas, and fears, based on a lack of factual information. The vast number of Blogs and Internet pages spreading misinformation on the prospects of AI makes finding a proper reference hard[13]. The fear is fueled in addition by comments such as *The development of full artificial intelligence could spell the end of the human race….It would take off on its own, and re-design itself at an ever-increasing rate. Humans, who are limited by slow biological evolution, couldn't compete and would be superseded.* By Stephen Hawking's[14], and Elon Musk's comment on Artificial Intelligence: *"With artificial intelligence, we are summoning the demon."* [15] During the work on the research presented in this paper, which primarily focused on 2D to 2D style transfer, it became very clear how much the behavior of a NN is dependent on the training and parameter tuning conducted by a human being (see argument above). In a sense this means if an AI turns *"demonic"* it does so only because of the training it received − channeling the malignant traits of the human mind[16]. The main stumbling block is the generalization of the problem. In addition, there is a major problem with the assumption that "Jobs get Lost" This generalized assumption needs to be met with a healthy dose of skepticism and needs a critical interrogation. On the one side Jobs transform into something different, on the other hand it is possible to think about jobs in general in an alternative way. The economic concepts that are currently in operation will not be able to cope with the changes at hand pertaining to AI and Automation − as they are profoundly rooted in the 19th century economic thinking of the industrial revolution which are not applicable any more[17]. The way we think about jobs have to transform with the surge of AI and Automation. All of these points describe the ecology of the conversation, the intellectual atmosphere that the research started to touch upon in the initial phase of the project, in regards of grasping the scale of this paradigmatic shift. On another note it can be stated that the term AI is profoundly vague as it describes an entire array of computational techniques such as Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN), and many more.
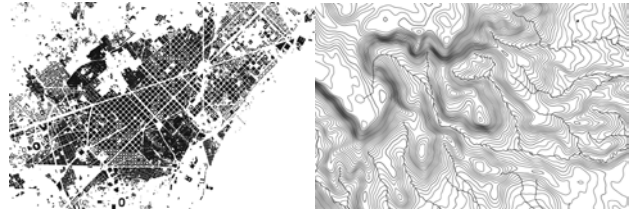
Figure 5. Result of Style transfer between a Dataset of Nolli maps of known cities (Rome, Barcelona, Manhattan, Washington DC) and a 19th century science plate depicting a detail of the moon's surface. These attempts can be considered experiments in Posthuman sensibilities. Can AI's be creative?

3 DYNAMIC STYLES

If we turn the focus back on its consequences for urban design the authors would claim that when these techniques are applied to design, they can blend a chronology of styles to create a *dynamic style* that captures and reflects a variety of design techniques over a period of time including social and cultural evolution (fig.3). Style artifacts can be exaggerated to a point of hyperbole, transforming the natural balance/harmony of human style and design into a pareidolic and compositionally unstable, but novel form rooted in post-human (in the sense that they were not primarily authored by human ingenuity), but humanly accessible, architectural features. An example for this approach was tested by the authors by creating a database of Nolli maps of various cities scraped from the web and applying this *style* to various target images (fig.6). In a playful approach we chose images such as 19th century science plates of the moon, topo lines of alpine areas and random Asian cities as target files. The resulting images serve as a first proof of concept of a possible Neural Network technique for the design of cities. Further work on this technique will be conducted in the upcoming semesters. The goal is to gain better control of the Neural Network by implementing rulesets in order to give more weight to specific solutions. For example, the rulesets defined by Christopher Alexander in *A Pattern Language.* It would be highly interested to see how these rulesets can be applied to an urban condition. The presented 2D image editing method has interesting implications for 2D urban design applications. By employing this technique, it is possible to create style transfers between various city plans, or to hallucinate alien features into conventional maps. Exemplary demonstrated by creating style mash ups between rural areas and city plans, or mutants of various urban conditions. In the following the authors would like to explain the technical background of this approach, explaining the computational methods used in NN.

258

4 BACKGROUND GENERATIVE ADVERSARIAL NETWORKS AND THE 2D VISUAL WORLD

Artificial neural networks are computing systems that are designed to loosely emulate the hierarchical structure of the human visual cortex. A neural network is comprised of processing nodes, called *neurons*, that are organized into groups, called *layers*, based upon how they connect to other nodes in the network. Input information flows through a neural network in a feed-forward, hierarchical manner: Each neuron in the network receives input from neurons in the preceding layer and transforms it into a new representation via a nonlinear function, which acts as a threshold that filters out relevant information captured by its input. This new representation becomes the input to the neurons it is connected to in the proceeding layer. The way in which neurons are connected and transmit information are specific to particular tasks and need to be learned from input data. In this paper, we are interested in purely visual tasks and modeling visual information, so the following sections only consider convolutional neural networks (CNN), which are designed to operate on images. The set of filtering transformations the network performs on images, and consequently the novel ways the network represents salient visual information captured by the images, are learned directly from the image pixel intensities. For example, in image classification, a neural network transforms an input image into a new representation by decomposing it into a set of visual features that makes the semantic image content easy to classify as, for example, 'Street' or 'Plaza'. The visual features that comprise this new image representation could be textural, like tar, concrete, greenery or shadow, or pertain to geometry and shape, like curves or corners. Thus, the 'Street' class may be represented by a set of long, continuous line features combined with stone textural features such as pedestrian crossings or green striped in the middle of the road, whereas the 'Plaza' class could be represented by a set of corners and polygonal features. These visual features are extracted sequentially by the network, where the first layers filter out simple lines, edges and textures, and the later network layers filter out the sets and combinations of these features, such as corners. The final network layer predicts the semantic class label, e.g. 'Street', based upon the set of features extracted from the image by the preceding layers. In this example, the CNN is trained for a discriminative task, and functions as a prediction/classification machine. For this kind of task, the network learns only to model the visual information that maximally differentiates the semantic classes present in the dataset.

5 METHODS -OR: MODELING THE STYLE OF THE REAL WORLD

Independent of the task, neural networks learn how to represent images in terms of color, texture, and geometric structure. These representations can be used to perform image manipulations that result in unique design. In the following subsections we discuss the specifics of the style transfer technique called Neural style transfer[18] , which was used to generate the images in this paper and forms the core result of the presented process. The objective of this image editing method is to alter a given input image so that it captures the style of a second, 'style guide' image without altering the original content, i.e., the geometric/spatial structure of the input image. As previously described, an input image can be decomposed into specific visual features by projecting it into a given network layer, i.e., transforming it into the set of visual features learned by that layer. The network layer representation of the image not only provides information as to what type visual features are present in the image, but also where they occur within the image. Thus, through an optimization process, we can iteratively change the pixel values of our input image such that the network's representation of its style features, like texture and color, resembles the network's representation of the style features of the guide image, while making sure that the network's representation of structural features in the input image, such as outlines of buildings or edges, remain unaltered. This technique allows us to have a quantifiable metric of style that can be used to probe how the 3D nature of buildings, and other architectural components, like streets and buildings, are decomposed and represented in this 2D space. As shown in Figure 6, this new style representation of a building can be fused with other buildings to generate novel architectural types.

Figure 6: Using the same Neural style transfer method, we can apply the architectural style of one building section onto another to produce a novel, different architectural style.

Style transfer in addition to its technical abilities evokes memories to the discussion on style in architecture. It is indeed amusing that the term *Style* returns into conversations about architecture and urban planning via neuroscience and computer sciences, as if it comes back to haunt the discipline and remind them of the importance of its own tradition in this crucial conversation, with proponents such as Gottfried Semper[20] and Alois Riegl[21].

6 CONCLUSION – THE DEFAMILIARIZATION OF THE CITY, OR: AN ALTERNATIVE UTOPIA

As described in the introduction to this paper the urban map is a cultural staple of the architecture discipline. It is the medium that best captures the intentionality of the urban project in an abstract medium as a two-dimensional surface. In architecture discourse the line, the plan, the abstract representation of materiality has played a major role, and it

always has been interpreted as the result of human cognition and mind. This can be illustrated as a core idea in the architectural theory of for example Leon Battista Alberti, as expressed in the *De re aedificatoria*, pertaining to the distinction between "lineament," the line in the mind of the architect, and "matter," the material presence of the building. This particular distinction plays a key role in architectural design, and the conceptualization of the architectural project, throughout the history of western architecture. Le Corbusier described this at the heyday of modernism in the twentieth century like this: "Architecture is a product of the mind." The distinction between mind and matter can be found in Vitruvius, in the distinction between "that which signifies and that which is signified"; at the Accademia di San Luca in Rome, between *disegno interno* and *disegno esterno*; or in Peter Eisenman's distinction between *deep aspect* and *surface aspect* in architecture, to name just three examples[22] that profoundly describe the planning process as a particular ability of the human mind. What position does the discipline have when it comes to understanding the potentialities of applications such as NN's that are able to produce results that question the sole authorship of human ingenuity? Well, there is always the chicken & egg problem: NN's origin in the human mind. That they are able to autonomously generate plan solutions is in itself not yet proof for thinking or even intelligence. However, if we take the philosophical standpoint of materialism it would allow to create an even field between these two thinking processes. In a materialist tradition though itself is just the result of material processes in our brain, neurochemical reactions able to form thought. This was briefly described above in the section explaining the origin of the terminology used in this paper such as *Hallucinating*. If this position is taken, then the conclusion is that AI's can think as much, and form original language[23] or shape[24] as humans can, the only difference being that their neural processes are not based on neurochemical processes but computational processes within another material paradigm. In this paper, we present the possibility to utilize AI applications for the generation of planning processes. In particular the application of style transfers with NNs. This approach on the one side critically interrogates the unique position of the human mind when it comes to creative processes and in addition questions aspects of creativity in planning processes. In a design ecology where the boundaries between human and computational cognition are increasingly blurred, the presented process harvests the multiplicious solutions found by architects throughout the ages and employs mining big data to create possible novel solutions to planning problems.

In an outlook it can be stated that this is only a first attempt in the area of the critical interrogation of planning in architecture in the age of AI. In fact, there is still a lot to be done. The first, alien, results achieved in this paper can only be seen as a first tapping into the potentialities of this approach. From tapping into novel design direction that rather talks about how machines see our world - with all its wonderfully strange results in terms of morphologies, chromatics and possible theories, to profoundly pragmatic approaches. It is feasible to speculate about the pragmatic applications of the findings in this paper. The possibility to create an application as a corrective tool in the planning process. Through datamining (Fig.5) it would be possible to create a NN that can analyze plans to see for example if they comply with local building codes. Or the plans can be analyzed to see if they are functional at all. All of these abilities need to be trained, heavily relying on human judgement at the beginning, but increasing its abilities after a period of training.

Further research needs to be done to dive deeper into the opportunities presented in this paper. In this extent the work on this problem can be considered a work in progress. The refinement of the algorithm allows to continue the conversation laid out in the conclusion of this paper. The authors of this paper have already started refining the approach and are looking forward to the in-depth interrogation of this posthuman design ecology.

REFERENCES

1. The authors refer here to the map of the Babylonian city of Nippur – ca. 1400BC

2. *Style* in this context of conversation is borrowed from computer science to describe a specific computational problem.

3. The full paper contains a glossary explaining the mathematical background of the work done by the authors. By request the authors can provide the code for others to replicate the results.

4. Leon Batista Alberti, *De Rei Edificatore*

5. Zaha Hadid Masterplan Istanbul

6. *Just think about Patrik Schuhmacher's Theoretical oeuvre and the schism it has created in the discipline by provoking with neoliberal statements. In the process creating a counter culture in Digital Design opposing the neoliberal position and adopting instead a leftist, Accelerationist Ideology.*

7. See also Greg Lynn's entire conversation on "Happy Accidents"

8. See Glossary

9. Graham Harman, Weird Realism: Lovecraft and Philosophy, Zero Books, Hants, UK, 2012, p.93

10. Think of Coop Himmelblau's manifests of the 19070ies, or Hans Hollein's flirt with the terms Visionary, also Peter Cook Visionary Architecture, or Gunther Feuerstein's Book: "Visionary Architecture in Vienna 1958 to 1988"

11. See for example the work of Stanislas Chaillou, Daniel Bolojan, Guvenc Ozel, Daghan Cam, Alisa Andrasek and many more.

12. Jacobs, Barry L., Dreams and hallucinations: A common neurochemical mechanism mediating their phenomenological similarities, Neuroscience & Biobehavioral Reviews, Volume 2, Issue 1, Elsevier, London 1978, p. 59-69

13. https://time.com/4742543/robots-jobs-machines-work/

14. Hawkings, Stephen., excerpt from an Interview with the BBC given December 2nd, 2014

15. Musk, Elon., Q&A during the MIT Aeronautics and Astronautics department's Centennial Symposium in October 2014.

16. See also Jeff Larson, Surya Mattu, Lauren Kirchner and Julia Angwin, Machine Bias -

17. See also: N. Srnicek, A. Williams, *Inventing the Future*, Verso, London, New York, 2015

18. Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "A neural algorithm of artistic style." arXiv preprint arXiv:1508.06576 (2015).

19. Kato, Hiroharu, Yoshitaka Ushiku, and Tatsuya Harada. "Neural 3d mesh renderer." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

20. Semper, Gottfried, Der Stil in den Technischen und Tektonischnen Künsten oder praktische Ästhetik: Ein Handbuch für techniker, Künstler und Kunstfreunde (Band 1): die textile Kunst für sich betrachtet und in Beziehung zur Baukunst. Verlag für Kunst und Wissenschaft Frankfurt am Main, 1860, P.13

21. Riegl, Alois, *Stilfragen: Grundlegungen zu einer Geschichte der Ornamentik, Berlin 1893*

22. Hendrix, John S., "Leon Battista Alberti and the Concept of Lineament" (2011). *School of Architecture, Art, and Historic Preservation Faculty Publications*. 30.

23. See for example the *Bob & Alice* experiment by the Facebook AI Research group. Two chatbots were programmed to discuss economic problems with each other. Once the test ran overnight the two bots started to develop their own language.

24. See for example the artwork Portrait of Edmond de Belamy created by Paris based art collective *Obvious* using a Generative Adversarial Network. It was sold at Christies for the sum of $432.000, and was promoted by the auction house as *the first painting solely created by Artificial Intelligence*

25. Gatys, Leon A., Alexander S. Ecker, and Matthias Bethge. "A neural algorithm of artistic style." arXiv preprint arXiv:1508.06576 (2015).

7 GLOSSARY – MATHEMATICAL SUPPLEMENTARY

We provide the mathematical formulation for the different 2D to 3D image editing methods in this section. Each of the following descriptions treats a neural network as a function whose parameters/weights map the input space of images to the output space of labels, or into the activation/feature space of a given network substructure, e.g. a layer or neuron. Additionally, when the term *error* is used, we are referring to calculating the Euclidean distance between two quantities. This is denoted by $| \ldots |^2$ in the following equations and definitions.

Style Transfer

Using the learned representations of a pretrained image classification neural network, VGG-16, we can define a training objective that is based upon the spatial features, or 'content features' of the input 3D mesh, which we assign the name m^c and the 2D 'style features' of the second, user-provided guide image, which we assign x_s.

To make the shape of the generated mesh, m, similar to that of m^c, the 3D content loss can be defined as:

$$\ell(m|m^c) = \sum_{v_i, v_i^c \in m, m^c} |v_i - v_i^c|^2$$

Where v_i is the set of vertices for the manipulated mesh m and v_i^c is the set of vertices for the original content mesh m^c. The style loss is defined to be the same in the 2D image case using the rendered/rasterized image that is output from the 3D Neural mesh renderer:

$$\ell(m|x_s, \phi) = |M(f_s(R(m, \phi))) - M(f_s(x_s))|_F^2$$

Where R is the 3D neural renderer function that projects the 3D mesh m to a rasterized 2D image, ϕ is the viewing angle at which to rasterize m, f_s is the pretrained VGG-16 network (used as a function) that projects the rasterized image into the feature space/representation of a specific network layer, and M is the Gram matrix function, which acts as a metric of style. The feature layers of the VGG-16 network used were *conv1_2*, *conv2_3*, *conv3_3*, and

conv4_3. These two losses are summed to make the final objective, which is minimized via backpropagation to make the output mesh object.

2D to 3D vertex optimization

This method is similar to style transfer, but instead of using the learned feature representations to manipulate the mesh, the training method minimizes the error between the input mesh rasterized into a silhouette image and a user-provided guide silhouette image, $x_{silhouette}$. To rasterize a silhouette image from a mesh, the 3D neural mesh renderer is paired with a neural network that generates a silhouette image from the output of the mesh renderer. The training objective is to minimize the difference between the rendered silhouette image and the reference/guide silhouette image via backpropagation:

$$\ell(m|x_{silhouette}) = |R(m|\phi) - x_{silhouette}|^2$$

3D Deep Dreaming

Deep dreaming, whether it is applied in 2D or 3D, is a visualization technique that allows us to qualitatively determine what visual features a given substructure of the network has learned. Conceptually, it makes assumptions similar to the grandmother cell hypothesis in Neuroscience: there is one neuron that is trained to be the detector for the face of a grandmother. Thus, this process of deep dreaming is more like hallucinating or pareidolia; the network is emphasizing vague pixel patterns in the image if those patterns resemble something that the neuron has learned to detect. In essence, we are seeing what the network is 'seeing' in the image.

To achieve deep dreaming, the training objective is to maximize the activation of a specific neuron, layer, or class by changing the values of the input image pixel intensities over many iterations. In 3D, the same objective is used to manipulate the vertices of a mesh object. Let $f(x)$ be the GoogLeNet pretrained neural network as a function that outputs an activation/feature map for the input image x at the specified neuron. The 3D neural mesh renderer is used to transform the mesh m into an image, which is then fed into GoogLeNet to produce an activation map for the chosen neuron. A neuron in layer *inception_4* from GoogLeNet was used for all of the mesh manipulation. The training objective for 3D deep dreaming to be optimized is

$$\ell(m) = -|f(R(m, \phi))|_F^2$$

where $R(m, \phi)$ is the rasterized image given the input mesh and viewing angle and f is the GoogLeNet pretrained network that projects an image into the representation of the specified neuron.

Building Physics

Energy Co-Simulation of the Hybrid Cooling Control with Synthetic Thermal Preference Distributions

Siliang Lu[1], Zhiang Zhang[1], Erica Cochran Hameen[1], Berangere Lartigue[2], Omer Karaguzel[1]

[1]Carnegie Mellon University
Pittsburgh, USA
{siliang1, ericac, zhiangz, okaraguz}
@andrew.cmu.edu

[2]Paul Sabatier University,
Toulouse, France
berangere.lartigue@univ-tlse3.fr

ABSTRACT

Thermal comfort and energy efficiency are always the two most significant objectives in HVAC operations. However, for conventional HVAC systems, the pursuit of high energy efficiency may be at the expense of satisfactory thermal comfort. Therefore, even if centralized HVAC systems nowadays have higher energy efficiency than before in office buildings, most of them cannot adapt the dynamic occupant behaviors or individual thermal comfort. In order to realize high energy efficiency while still maintain satisfactory thermal environment for occupants indoors, the integrated hybrid HVAC system has been developed for years such as task-ambient conditioning system. Moreover, the occupant-based HVAC control system such as human-in-the-loop has also been investigated so that the system can be adaptive based on occupant behaviors. However, most of research related to personalized air-conditioning system only focuses on field-study with limited scale (i.e. only one office room), this paper has proposed a co-simulation model in energyplus to simulate the hybrid cooling system with synthetic thermal comfort distributions based on global comfort database I&II. An optimization framework on cooling set-point is proposed with the objective of energy performance and the constraints of thermal comfort distribution developed by unsupervised Gaussian mixture model (GMM) clustering and kernel density estimation (KDE). The co-simulation results have illustrated that with the proposed optimization algorithm and the hybrid cooling system, HVAC demand power has decreased 5.3% on average with at least 90% of occupants feeling satisfied.

Author Keywords

Thermal comfort; Energy co-simulation; Synthetic distribution; Hybrid cooling.

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING (e.g. Model Development).

1   INTRODUCTION

Thermal comfort and energy efficiency are always the two most significant objectives in heating, ventilation and air-conditioning (HVAC) operations. However, for conventional HVAC systems, the pursuit of high energy efficiency may be at the expense of satisfactory thermal comfort. Therefore, even if centralized HVAC systems nowadays have higher energy efficiency than before in office buildings, most of them cannot adapt the dynamic occupant behaviors or individual thermal comfort. In order to realize high energy efficiency while still maintaining satisfactory thermal environment for occupants indoors, the hybrid air-conditioning system has been developed for years such as task-ambient conditioning system. Moreover, the occupant-based HVAC control system such as human-in-the-loop has also been investigated so that the system can adapt the system based on occupants' feedback actively or passively. The following sections will introduce recent developments of adaptive thermal comfort and the occupant-based control.

1.1 Adaptive thermal comfort

Since innovations in HVAC are inspired with the targets to improve energy efficiency and improve thermal comfort for individuals, it is of great importance to have comprehensive understandings in these targets.

As defined by ANSI/ASHRAE 55 and ISO7730 [1], thermal comfort is a condition of mind which expresses satisfaction with thermal environment and is assessed by subjective evaluation. For the past 40 years, many researchers have been investigating the principle indicators of thermal comfort. The Danish scientist Fanger believed that thermal comfort was same as neutral state in terms of thermal sensation based on experimental studies. He also derived a well-known equation called predicted mean vote (PMV). Moreover, for indoor environment, since heat exchange and evaporation loss are owing to the difference between thermal environment conditions and human body conditions, for PMV equation, the following six measurable variables are accepted to be the indicators of thermal comfort, which are indoor air temperature, indoor relative humidity, indoor air velocity, mean radiant temperature, clothing insulation and metabolic rate [2].

However, since in regular office buildings, thermal environments are different from such well-controlled

experimental test-bed, many researchers are turning to adaptive thermal comfort model instead of static thermal comfort model such as PMV for years. For instance, one of the milestone projects of adaptive thermal comfort is ASHRAE RP-884 (comfort database I) [3] which collected a total of 22000 sets of data from the real office environments across the world. This project has been widely used to develop various adaptive thermal comfort models, which have been integrated into personalized HVAC controls. With RP884 dataset, Seungjae et al. [4] has proposed a method for learning personalized thermal preference profiles by formulating a combined classification and inference problem with 5-cluster model. However, instead of predicting 7-point thermal sensation described in ASHRAE 55, the paper predicts thermal preferences with 3 classes, namely "want cooler", "want warmer" and "no change" by Bayesian approach. Rather than classification of thermal sensation, it predicts the probability of a test occupant falling into each of the classes by clustering all occupants with Gaussian mixture model (GMM). Moreover, Frederik et al. [5] has also proposed a personalized thermal comfort model using Bayesian network to predict thermal sensation indoors in a specific area such as San Francisco with ASHRAE RP-884 dataset. Moreover, a newly released dataset called ASHRAE global thermal comfort database II (Comfort database II) intends to support diverse inquiries about adaptive thermal comfort in field settings [6].

In all, factors to adaptive thermal comfort in indoor environment can be categorized into environment-related factors and occupant-related factors, as shown in the following tables.

Variable	Unit
Indoor air temperature	℃
Indoor relative humidity	%
Indoor air velocity	m/s
Mean radiant temperature	℃

Table 1. Environment-related factors to thermal comfort

Variable	Unit
Metabolic rate	met
Clothing insulation	clo

Table 2. Occupant-related factors to thermal comfort

1.2 Occupant-based HVAC control

In order to reduce energy consumption of the existing HVAC systems and improve occupant comfort, occupant-responsive HVAC controls have been being investigated for

years. One of the key to designing occupant-responsive HVAC system is to understand the occupant behaviors.

Besides meeting thermal comfort requirements mentioned above, occupant-responsive HVAC system also plays a role in determining the energy consumption of the entire building. For most of the commercial buildings, particularly office buildings and schools, the heating and cooling loads are largely dependent on the occupanct behavioral patterns like occupant presence and activities. However, the conventional HVAC systems have been operated without the ability to adjust supply air rate accordingly. Therefore, much of the energy use for HVAC is wasted, particularly when the conditioned spaces being unoccupied or the operation being under the maximum levels. On the contrary, since occupant-responsive HVAC system can be responsive to the dynamic occupancy profile, it has a large potential to reduce energy consumption.

Occupant behaviors have two distinctive effects on building performances, which are passive and active effects [9]. Passive effects are derived from dynamic occupancy schedules like the presence of the occupants during a day or occupancy behaviors like using the microwave in the lunchtime or doing computer-based work. Active effects are derived from individual preferences of the indoor environment such as personal thermal comfort or occupancy behaviors like turning on/off lights or opening/closing windows based on their own preferences. In other words, to understand passive effects of human on the building systems, it requires objective occupancy information like occupancy schedules or location of occupants. However, to understand active effects of human on the building system, it requires subjective occupancy feedback describing individual preferences such as thermal comfort, visual comfort. Both passive and active effects could play important roles in operations of building systems and building diagnostics**.**

With the comprehensive understanding of effects of occupant behaviors on building performances, it is of great importance to incorporate occupant pattern recognition system with HVAC controls to improve occupant comfort and increase energy savings. Besides, among different occupant-based HVAC systems, the personalized task-ambient conditioning system is not only able to provide occupants with individual control to adapt individual thermal comfort preferences but also ensure that the centralized HVAC system is operated with high energy-efficiency.

The personal comfort system (PCS) from Centre for Built Environment (CBE), UC Berkeley is an innovation to develop a low-energy personalized systems as micro-zones and integrate them into centralized HVAC operations as a macro-zone in open plan office environments [10]. The project has invented the personalized heating and cooling chairs with wireless internet connectivity and tested the performances in different real office environments in

California. PCS adjusts the local thermal environment based on occupants' inputs regarding heating/cooling set-points of the chair. Meanwhile, the whole framework gets further optimized with communication between chairs and the centralized HVAC system by controlling the set-points of the centralized system based on feedbacks of all micro-zones. In the case study, the test energy performances were optimized with the mode of widening HVAC temperature setpoint dead band in conjunction with proposed chairs. In addition, Zhang et al. [11] has developed a task-ambient system heating only the feet and hands, and cooling only the hands and face, to provide comfort in a wide range of ambient environment. The simulated annual heating and cooling energy savings with such task-ambient system is as much as 40%. Last but not least, Lu et al. also [12] has conducted a field study to evaluate the energy and thermal comfort performances of a hybrid cooling system consisting of personalized cooling fans and split air-conditioning system in Shanghai.

Based on the literature review, much more attention has been paid to occupant-based HVAC control, especially task-ambient conditioning system than before. Moreover, compared to static thermal comfort, the adaptive thermal comfort has become more popular, particularly developing thermal comfort model with advanced machine learning algorithms. However, few studies have applied adaptive thermal comfort models trained with machine learning algorithms into the whole-building energy simulation for evaluating the hybrid cooling system such as task-ambient conditioning system. Therefore, this paper aims to evaluate the energy and thermal comfort performances of a task-ambient cooling system where each task system consists of a personal fan and ambient system is a typical VAV system with energy co-simulations. Moreover, this paper has also proposed to use comfort database I&II to create synthetic thermal preference distributions so as to design an optimization control framework for the task-ambient conditioning system.

2 METHODOLOGY

2.1 Development of thermal preference distribution
Even if individuals have different thermal preferences under the same thermal environment in air-conditioned open-plan offices, most of the thermal preference distributions can be approximated as Gaussian distributions where the majority are satisfied while only a few of occupants vote for either being uncomfortably warmer or uncomfortably cooler. Therefore, in order to simulate different thermal preferences in a shared space, Figure 1 shows the diagram of developing the synthetic thermal

preference distributions with the comfort database I&II. As shown in the figure, the pipeline is comprised of clustering of the thermal environments, thermal preference distribution synthesis and sampling as well as calculation of the number of occupants under each thermal preference given the total number of occupants.

The subset of the comfort database I&II was used where a total of 2354 instances were collected. In the subset, data were collected either from hybrid cooling system consisting of ceiling fans and the centralized cooling or from the conventional centralized cooling system. Among the subset, 815 instances were collected with the task-ambient cooling system while 1539 instances were collected with the conventional centralized cooling system.

Since the indoor environments are expected to vary a bit in the database, the unsupervised clustering is implemented so as to cluster the similar thermal environments into a single cluster and see the histogram of thermal preference in each cluster. Instead of using K-means, similar thermal environment conditions. Since the setpoint is optimized with the thermal environment in core zone where discomfort due to non-uniform radiation can be ignored, only air velocity, air temperature and relative humidity were used to represent thermal environment conditions. Then, similar thermal environment conditions were clustered with Gaussian mixture model (GMM) where the number of clusters were selected based on BIC score. After clustering, the histogram of thermal preference under each cluster was developed so that in the energy model, the number of occupants for each thermal preference can be sampled from the synthetic thermal preference distribution with kernel density estimation (KDE) under the given cluster. Moreover, since Energyplus cannot simulate the thermal environment changes after fans were operated, the subset without fan operations and the subset with fan operations were clustered, respectively.

2.2 Thermal preference synthesis algorithm
As mentioned before, the synthetic thermal preference distributions were approximated with Gaussian distribution. Therefore, kernel density estimation (KDE) was implemented based on the dataset. Kernel $K(x; h)$ is a function controlled by the bandwidth parameter h, which can
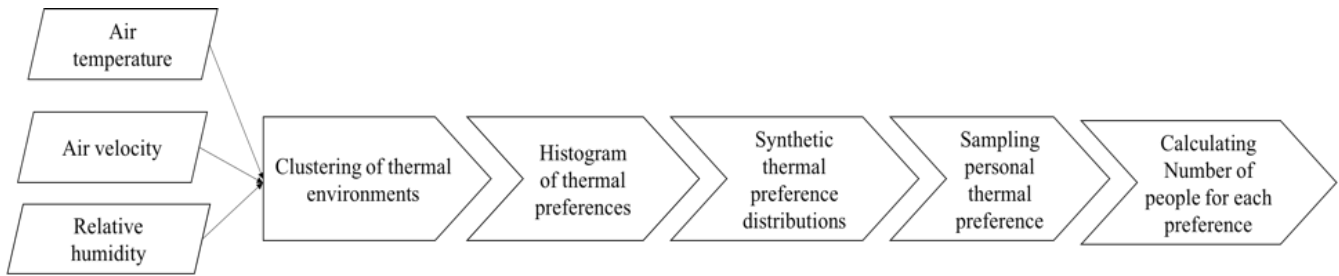
Figure 1 The diagram of developing the synthetic thermal preference distributions

be seen as smoothing parameter controlling the tradeoff between bias and variance in the result. Given the kernel form, the density estimate at a point y within a group of points x_i; $i=1..N$ is given by:

$$\rho_K(y) = \sum_{i=1}^{N} K\left(\frac{y-x_i}{h}\right) \text{ Eq. 1}$$

where h is bandwidth and the bandwidth is tuned with 5-fold cross-validation from 5 candidate values between 0.1 and 1.

2.3 Co-simulation with the proposed framework

The energy simulation was implemented with the co-simulation between energyplus and python. The one-story small office building was simulated in Shanghai from July 1st to August 31st. The total ground floor area is 512 m^2 and 5 thermal zones are built. Moreover, the cooling is supplied with packaged DX cooling coil and the heating is supplied with gas heating coil. The 3D rendering and floor plan are shown in Figure 2 and Figure 3. The co-simulation framework was developed in [13].

Figure 3 Thermal zones of the building

The baseline and optimized simulation were both conducted with dual setpoint schedule. The baseline cooling setpoint schedule is the default schedule where the cooling setpoint is predetermined and fixed to be 22℃ (Figure 4) and the heating setpoint is constant to be 21 ℃. Even if the heating set-point schedule is the same as baseline, the cooling setpoint schedule is based on the proposed optimization framework in the optimized simulation. Meanwhile, the setpoint schedules are the same for all zones so as to ensure the system responses the dynamic setpoint changes in time in both simulations. Moreover, it is assumed all perimeter zones are unoccupied and Figure 5 shows the occupancy schedules of core zone used in both simulations.

Figure 2 3D rendering of the reference building
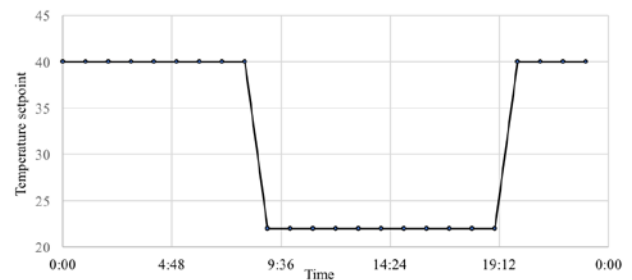
Figure 4 The fixed temperature set-point schedule in baseline

268
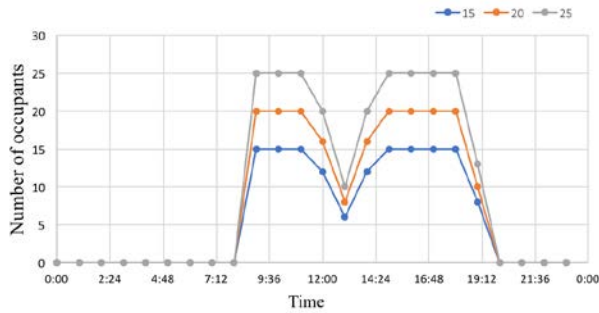
Figure 5 The occupancy schedule in both simulations

Besides baseline setpoint schedule, Figure 6 shows the flow diagram of the proposed optimization framework. As shown in the diagram, two histogram models were implemented to simulate the conditions when the personalized fan is operated or not, respectively. Moreover, the initial cooling set-point is 24℃ and the setpoint is increased by 1℃ or no change when the space is occupied at each time step. However, it is assumed that the reason for turning on fans is only because of feeling warm. Meanwhile, it is also assumed that all the fans will be turned off when determining a new set-point.

The control law of the proposed optimization framework is shown below:

Objective function:

$$F^*: = \min F = \text{sensible cooling loads} \quad \text{Eq.2}$$

subjective to:

% of occupants feeling warm $< \delta$ Eq.3
20 ℃$<$ cooling set-point $t <$ 30 ℃ Eq.4

where δ is threshold tuned with simulation episodes
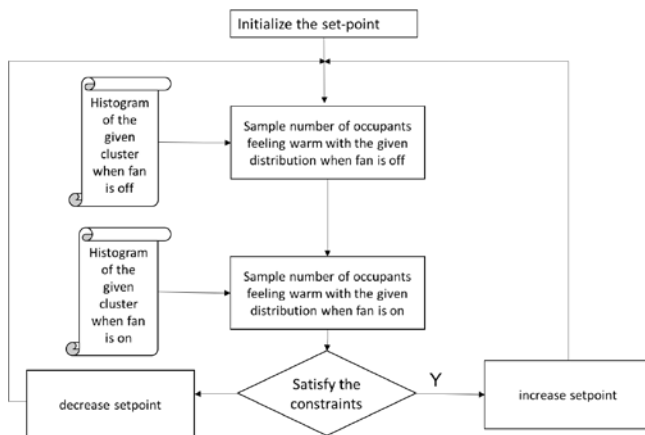
Figure 6 Flow diagram of the proposed optimization framework

As shown in the control law, the cooling set-point is controlled every 30 minutes. In addition, since the objective

function of sensible cooling loads is monotonously decreasing when cooling set-point is increasing, it will reach the optimal state under the constraints after several time steps in the end. Meanwhile, if the number of occupants feeling warm exceeds the threshold or the setpoint exceeds the boundary, the updating will decrease 1℃ for this time step. The threshold can be tuned with benchmark of multiple simulation episodes.

3   RESULT ANALYSIS

3.1 Synthetic thermal preference distributions

As a result, with comfort database I&II, 7 clusters were selected for the subset without fans and 6 clusters were selected for the subset with fans on according to the lowest BIC scores. Table 3 shows the mean of each cluster and the correspondent best bandwidth for KDE without fans. In addition, Figure 7 shows the histogram of thermal preference distribution within each cluster for the subset without fans, respectively. As shown in the table, except cluster 1, different clusters have similar indoor air temperature and indoor air velocity. Since no fans were deployed for the system, the air velocity is smaller than 0.2 m/s. However, relative humidity varies a lot among different clusters. This may result from lack of humidity control in common office buildings. In addition, except cluster 1, thermal preference histograms have shown that the majority vote is "no change" in different clusters. Moreover, compared to thermal preference vote of "want cooler", all clusters but cluster 1 have more votes for "want warmer". This may be because of low air temperature. Therefore, there is a potential to increase cooling set-point to save energy while maintaining occupant thermal comfort.

Velocity [m/s]	Temperature [℃]	Relative Humidity [%]	Optimal bandwidth
0.56	25.78	58.65	1
0.11	23.76	63.34	0.18
0.16	23.96	56.12	0.56
0.12	22.59	67.19	0.56
0.13	23.7	38.1	1
0.12	23.91	58.6	0.18
0.12	23.13	48.85	0.18

Table 3 The centroid of each cluster and the correspondent optimal bandwidth with fan off

269

Similarly, Table 4 shows the means of the clusters and Figure 8 shows the histograms of the thermal preference distribution within each cluster for the subset with fans on, respectively. As shown in the table, the mean air velocity of each cluster is higher than those in most of clusters without fans, which is because of the operation of fans. Meanwhile, the table has illustrated the average value of the mean air temperature in each cluster with fans on is larger than that without fans. In addition, the figure has illustrated that the majority votes within each cluster is "no change", which means such task-ambient cooling system has potential to increase air temperature to save energy while still maintaining high thermal comfort level.

0.47	25.85	63.09	0.32
0.28	24.83	75.68	0.32
0.17	23.5	44.54	0.56
0.19	25.36	63.27	0.32

Table 4 The centroid of each cluster and the correspondent optimal bandwidth with fan on

Velocity [m/s]	Temperature [℃]	Relative Humidity [%]	Optimal bandwidth
0.31	25.98	66.24	0.56
0.34	26.72	58	0.32
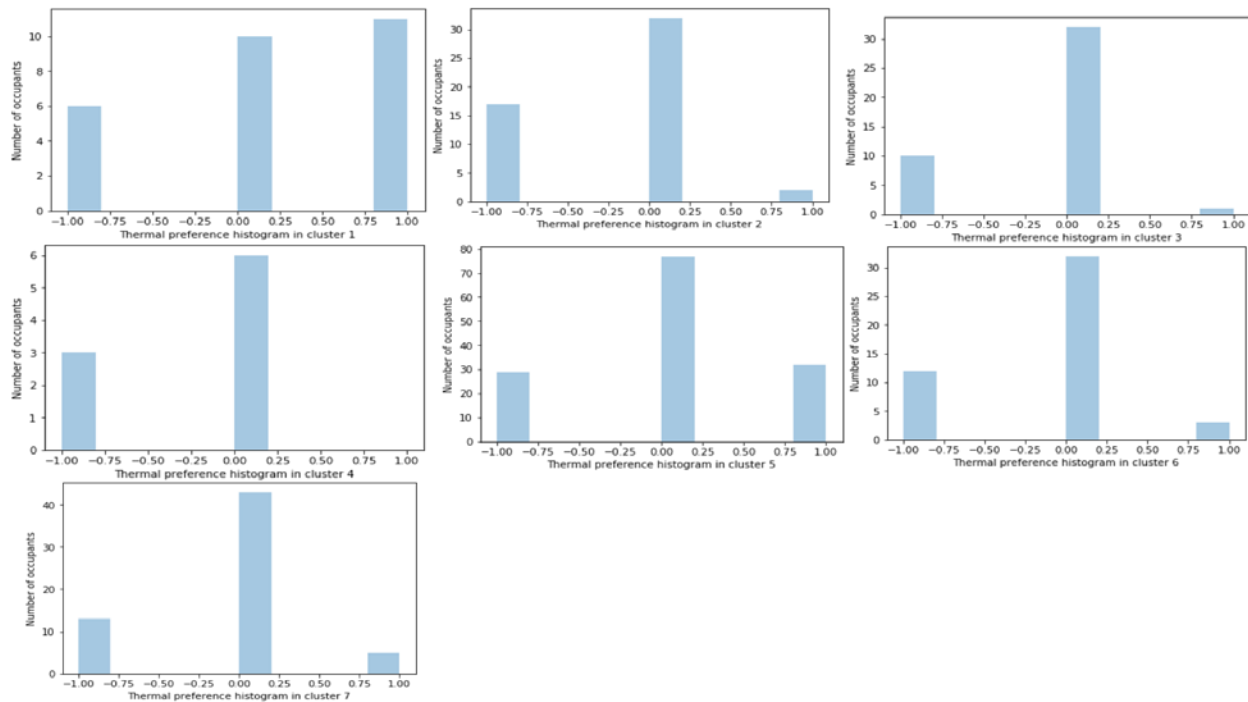
Figure 7 The histogram of thermal preference distribution with fan off
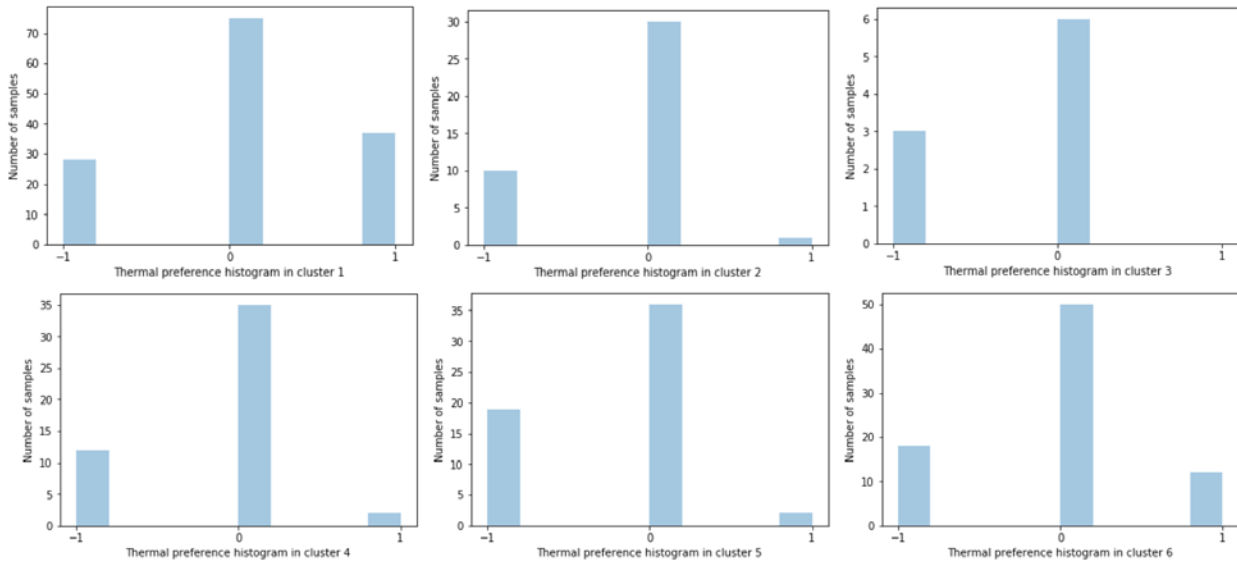
Figure 8 The histogram of thermal preference distribution with fan on

3.2 Energy benchmark of the proposed framework

In terms of energy benchmark between baseline control and the proposed control framework, HVAC electric demand power was used to evaluate the energy performances. Meanwhile, the percentage of occupants feeling warm is controlled within 10%. As a result, Figure 9 shows the comparisons of HVAC electric demand power between baseline and the proposed optimized models with different occupancy schedules. As shown in the figure, the proposed framework has achieved 5%, 5.3% and 5.6% demand power reduction compared to baseline models with 90% of occupants are comfortable when the number of occupants are 15, 20 and 25, respectively. Therefore, it is meaningful to develop the task-ambient conditioning system which not only creates comfortable local environment but also improves the overall energy performance.
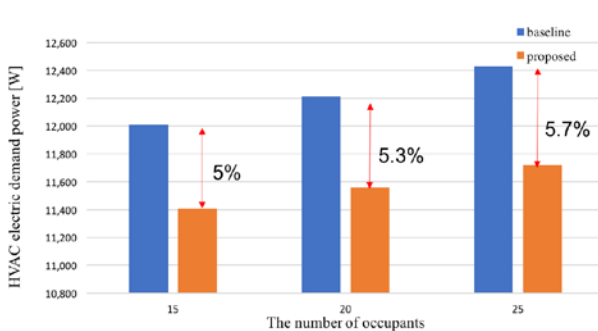
Figure 9 HVAC electric demand power between baseline and the proposed optimized simulations

4 DISCUSSIONS

The simulation study has evaluated the energy performances of the task-ambient cooling system consisting of personalized fans and the centralized cooling system. In order to simulate individuals, have different thermal preferences in a shared office space, the synthetic thermal preference distributions have been developed so as to generate "virtual" occupants with various thermal preferences in the same thermal environments. The results have validated that the proposed optimization framework could achieve 5.3% of HVAC electric demand power savings on average without the compromise of occupant thermal comfort. However, there are still limitations in the energy models. Firstly, due to lack of data, instead of common task conditioning systems such as personalized fans, this energy simulation has applied comfort database where the hybrid cooling system is comprised of ceiling fans and the ambient conditioning system. As a result, the thermal preference distributions based on comfort database may be different from actual task-ambient conditioning systems. In addition, the objective function in the control law can be improved to not only take sensible cooling loads but also take latent cooling loads into consideration.

5 CONCLUSION

This study has proposed an optimization framework to maximize the energy efficiency and thermal comfort with a task-ambient conditioning system by updating the cooling setpoint. In order to evaluate energy performances with the proposed optimization framework, a co-simulation of a typical office building was conducted with Energyplus. Moreover, in order to simulate the fact that different people have different thermal preferences in Energyplus, synthetic thermal preference distributions were generated with kernel density estimation in each cluster based on GMM clustering of the thermal environment given in the comfort database. The results have shown that with the proposed framework, the proposed framework has achieved 5%, 5.3% and 5.6% demand power reduction compared to baseline models with 90% of occupants are comfortable when the number of occupants are 15, 20 and 25, respectively. In future, more

work could be done to optimize the energy performances in the simulation by updating other parameters such as relative humidity.

REFERENCES

1. Standard, A. S. H. R. A. E. (2010). Standard 55-2010: "Thermal Environmental Conditions for Human Occupancy"; ASHRAE. Atlanta USA.

2. Fanger, P. O. (1970). Thermal comfort. Analysis and applications in environmental engineering. Thermal comfort. Analysis and applications in environmental engineering.

3. De Dear, R. J., Brager, G. S., Reardon, J., & Nicol, F. (1998). Developing an adaptive model of thermal comfort and preference/discussion. ASHRAE transactions, 104, 145.

4. Seungjae Lee, Ilias Bilionis, Panagiota Karava, and Athanasios Tzempelikos. A Bayesian approach for probabilistic classification and inference of occupant thermal preferences in office buildings. Building and Environment, 118:323–343, 2017.

5. Frederik Auffenberg, Sebastian Stein, and Alex Rogers. A Personalised Thermal Comfort Model using Bayesian Network. Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, (Ijcai):2547–2553.

6. Licina, V.F. et al. (2018), Development of the ASHRAE global thermal comfort database II. Building and Environment, 142, 502-512.

7. Choi, J.H. CoBi: bio-sensing building mechanical system controls for sustainably enhancing individual thermal comfort. 2010.

8. Ghahramani, A., Castro, G., Karvigh, S. A., & Becerik-Gerber, B. (2018). Towards unsupervised learning of thermal comfort using infrared thermography. Applied Energy, 211, 41-49.

9. Park, J. (2015). Are humans good sensors? Using Occupants as sensors for indoor environmental quality assessment and for developing thresholds that matter.

10. Andersen, Michael P., et al. "Well-connected microzones for increased building efficiency and occupant comfort." Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments. ACM.

11. Zhang, H. et al. "Comfort, perceived air quality, and work performance in a low-power task–ambient conditioning system." Building and Environment 45.1 (2010): 29-39.

12. Lu, S., Hameen, E.C., Zou, Y. (2019) An interactive building control for the integrative HVAC system featuring personalized cooling in office buildings. ASHRAE Annual Conference, Kansas City, USA.

13. Zhang, Z., Chong, A., Pan, Y., Zhang, C., Lu, S., & Lam, K. P. (2018). A deep reinforcement learning approach to using whole building energy model for hvac optimal control. In *2018 Building Performance Analysis Conference and SimBuild*.

Hydrogel-Based Evaporative and Radiative Cooling Prototype for Hot-Arid Climates

Dorit Aviv[12], Maryam Moradnejad[3], Aletheia Ida[3], Zherui Wang[1], Eric Teitelbaum[1],

Forrest Meggers[1]

[1]Princeton University
Princeton, NJ, USA
dorit@princeton.edu

[2]Univeristy of Pennsylvania
Philadelphia, PA, USA

[3]University of Arizona
Tucson, AZ,USA

ABSTRACT

A roof aperture lined with hydrogel membrane is proposed for combined evaporative and radiative cooling in the desert climate. In order to determine the design, materials and predicted performance of this device, several types of digital simulations and physical experiments were performed. The proposed full-scale prototype, planned to be built and tested in Tucson, Arizona, responds to the climate's extreme diurnal temperature gradient through the use of adaptive materials. During the day, the roof aperture acts as a downdraft chimney, trapping the hot dry air passing through it. The funnel-shaped top of the chimney is embedded with a wet hydrogel membrane, which humidifies the air, causing instantaneous cooling and a consequent downdraft airflow into the building's interior. During the night, pockets of enclosed hydrogel encapsulated in the roof's structural frame are exposed to radiative cooling from the night sky and act as thermal storage for additional cooling during daytime. The complexity of the system requires several simulation and physical testing methods to be employed simultaneously: digital simulation tools of CFD, solar radiation analysis, radiative heat loss analysis were employed to analyze the overall geometry's effect on airflow radiant heat exchange; physical bench tests were conducted to analyze the performance of hydrogel membrane and compare it to other materials. A full-scale prototype will be built to validate the results of the model.

Author Keywords

Energy Performance; Airflow Simulation; Responsive Envelope; Material Properties; Hydrogel Membrane

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING

1 INTRODUCTION

A roof aperture prototype channeling airflow into the building's interior is constructed in the Sonoran Desert of Tucson, Arizona in the Southwestern United States to take advantage of two cooling strategies enabled by the extreme dryness of the air in desert climates: evaporative cooling and radiant cooling (see Fig. 1).

First, evaporative cooling is highly effective in this climate because of the low dew point, allowing for significant air temperature drop per mass of water evaporated into the air. The system developed here takes advantage of a combination of evaporative cooling and thermal buoyancy force: the aperture acts as a downdraft chimney that evaporatively cools the air around the chimney's crown by exposing it to a wetted hydrogel membrane surface. The moisture causes the air to instantaneously cool and fall down the chimney into the building's interior volume due to gravity of the denser air mix. Several researchers have described the benefits of downdraft cooling in hot dry climates [1,2]. A downdraft system for the desert has also been exhaustively studied in the context of energy production at the infrastructure scale [3].

Second, the chimney's enclosure is constructed of high-heat-capacity materials in order to gain further passive cooling through radiant exchange with the sky. The low humidity and clear skies in Arizona allow for efficient longwave radiative heat exchange between the cold sky and the thermal mass stored in the roof structure. The challenge this second strategy presents in the context of the desert is the extremely high incident solar radiation that would be absorbed by the thermal mass and cause it to heat during the day far beyond the cooling exchange with the sky. A method to provide shading from direct solar radiation must be provided, while still allowing an optical exposure to the sky. Additionally, the thermal mass must be insulated from heat gain by the hot ambient air during the day.

One solution, proposed by Craig et al [4] suggests the use of roof insulation which provides a barrier for convective heat transfer while still allowing for longwave infrared radiation to infiltrate through it. Raman et al [5] achieved radiative cooling to nearly 5oC below the ambient air temperature under direct sunlight using a thermal photonic approach with a photonic solar reflector that reflects 97% of the incident sunlight while emitting strongly and selectively in the atmospheric transparency window. Ambient convective gains are prevented by an enclosed space between the exterior air and the panel's surface. Using the same strategy, Goldstein et al [6] demonstrate fluid cooling panels that harness radiative sky cooling to cool fluids

below the air temperature. Our approach is similar, as we insulate the thermal mass from ambient air with an infrared transparent membrane, while the interior temperature inside of the chimney is constantly lower than the exterior temperature due to the evaporatively induced downdraft airflow.

The thermodynamics and geometry optimization of a dual-function cooling roof chimney has been initially studied in previous work [7,8] . In this current prototype, we have developed a system which uses hydrogel membrane as the interface for the heat transfer between the exterior and interior environments. The hydrogel membrane acts both as thermal storage material for the building's envelope and as the interface layer through which water vapor is diffused into the air for evaporative cooling.

Hydrogel is a hydrophilic three-dimensional water-swollen polymer. The material chemistry consists of a range of compositions with two basic types: polyacrylamides and polysaccharides. Polyacrylamides are comprised of chemical covalent bonds providing strong and irreversible crosslinks for longevity and durability of the polymer. Polysaccharides are comprised of physical bonds, such as van der Waals forces, and result in weak crosslinks for reversibility and down-cycling of the polymer. Each has their advantages and can be optimized for different functionality with a range of environmental impacts. Some unique benefits of using hydrogel as a building material include the translucent optical properties to provide natural daylighting and the thermal capacitance when saturated with water. In addition, the hydrogel allows for diffusion of moisture for evaporative cooling processes in a passive mode: the moisture will evaporative from the surface area of the membrane based on vapor pressure differentials, thereby intelligently responding to low-humidity environments for passive cooling.

Prior applications of hydrogel in building material systems include the CloudGel glazing system invented by Charoudi for Saint Gobain [9] to provide dynamic natural daylighting and heat transfer control. Another more recent application in the building material market is the hydrogel-lined ceramic blocks from the IAAC group in Spain [10], which provides evaporative-cooling through modular natural ventilation apertures. Early membrane studies for evaporative-cooling with natural ventilation demonstrate the unique correlations between dynamic daylighting and heat transfer with varying saturation rates [11,12]. In these earlier studies, the swelling and deswelling kinetics of hydrogel membranes are considered for providing actuation of ventilation apertures and modulation of airflow and daylighting dynamics. Thermophysical properties of hydrogel diffusion kinetics are studied for other applications [13-16]. As of date, there are very limited implementations of hydrogel in building integration in part due to market constraints (emerging material translated from biotechnology applications), and in part due to limits

on degradation cycling and fatigue (not yet demonstrated for building system applications). However, our work exposes the opportunities for multifunctional environmental response through ongoing research with these membranes.

2   METHODS
Due to the complexity of this dual-cooling system integrated into the proposed full-scale prototype (Fig. 1), multiple digital simulations and physical tests were conducted. The following sections will present various methods: first, bench-scale physical tests for the evaporative cooling capacity of the hydrogel membrane and the radiative cooling capacity of the thermal mass; then, digital simulations for the radiative exchanges and airflow dynamics of the full-scale prototype.
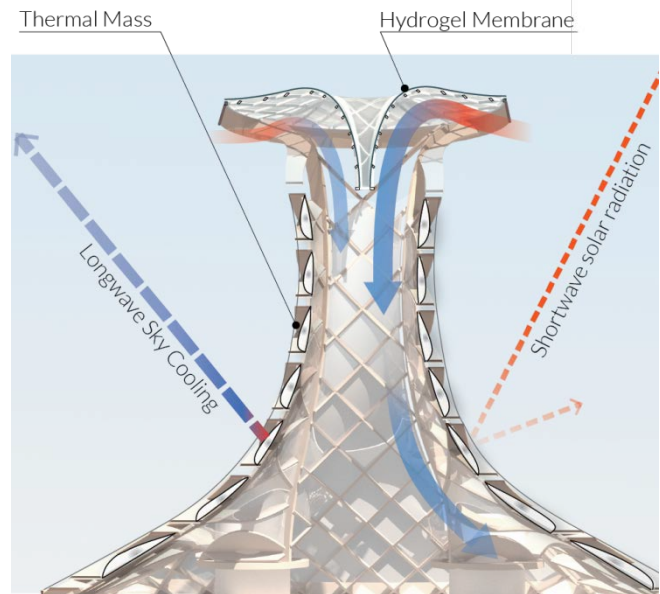
Figure 1. Diagram of the proposed prototype and its different cooling aspects to be simulated and tested: 1. the Hydrogel membrane's evaporative cooling capacity: once the hot dry air is humidified by water vapor diffusion from the hydrogel membrane embedded in the chimney's top funnel, it cools and drops down into the space below; 2. the thermal mass' radiative longwave sky cooling potential and shading/reflection of shortwave solar radiation through shortwave reflecting coating

2.1 Material Testing

2.1.1.   Evaporative Cooling: Hydrogel
The evaporative cooling membrane is designed and engineered to accommodate effective diffusion and re-hydration in the large-scale surface area application for the downdraft chimney.

2.1.1.1. Diffusion Calculations

Diffusion of moisture from the hydrogel membrane is dependent on two major variables: hydrogel mesh pore size and vapor pressure differential with surrounding air. The mesh pore size of hydrogel is controlled through the ratio of crosslink monomers during synthesis as per the Flory-

Rehner theory: the greater concentration of crosslink monomer, the smaller the mesh pore size [17]. Because some structural integrity is needed for maximizing the exposed surface area of the diffusion membrane in this application, there is a balance that is needed between large pore size to accommodate optimal diffusion with small pore size to accommodate higher strength. The polyacrylamide gel synthesized for the current project is a 4% crosslink ratio. Upon full-scale physical prototype testing, the hydrogel synthesis will be modified to optimize diffusion and strength relationships in the membrane application.

The chemistry of the hydrogel also enables the ability for moisture diffusion. In this case, we synthesize a polyelectrolyte gel to enable more ease of release of the water molecules from the polymer chains. Future work may also incorporate chemistry modifications for optimized diffusion rates.

The vapor pressure differential is determined by the surrounding relative humidity (amount of moisture in the air as a ratio of the maximum amount that the air-mix could hold before saturation). The lower the relative humidity, the lower the vapor pressure in the surrounding air, which will enable higher rates of diffusion. Hotter air can hold more moisture, so typically consists of a lower relative humidity, which affords effective evaporative-cooling and diffusion rates. In our application, we assume a fully wetted surface of the hydrogel membrane and calculate based on an hourly rate as follows:

$$g_h = \theta_w A_w (x_s - x)$$

where g_h is the amount of water evaporated per hour, θ_w is the evaporation coefficient of water, A_w is the exposed wet surface area of the membrane, and $x_s - x$ is the difference between the maximum humidity ratio and the actual humidity ratio of surrounding air (temperature-dependent) respectively. The evaporation coefficient for water is assumed at 0.5 based on comprehensive literature review on the subject [18-21].

We incorporate four daily peak relative humidity (RH%) and dry-bulb temperature (DBT) conditions for annual calculations on vapor diffusion potential. The TMYIII climate data for Tucson, Arizona provides the following peak DBT and RH% values:

March 21st (03.21):   24.2 ^{o}C   10% RH
June 21st (06.21):   37.2 ^{o}C   19% RH
September 21st (09.21):   33.9 ^{o}C   27 % RH
December 21st (03.21):   18.3 ^{o}C   33 % RH

The maximum humidity ratio (x_s) and absolute humidity (x) are established from the psychrometric chart according to the recorded climate data points identified above. The diffusion membrane evaporation rate is calculated for each season to provide the potential range for a 1.0 cm^2 surface area:

03.21: $g_h = 0.5(1)(19 - 1.9) = 8.55$ g/cm^2/hour

06.21: $g_h = 0.5(1)(38 - 8) = 15$ g/cm^2/hour
09.21: $g_h = 0.5(1)(35 - 9) = 13$ g/cm^2/hour
12.21: $g_h = 0.5(1)(14 - 4.5) = 4.75$ g/cm^2/hour

The hourly diffusion rate varies from approximately 4.75 g/cm^2/hour to upwards of 15 g/cm^2/hour. We utilize this information to develop the desired thickness and volumetric water holding capacity for the hydrogel membrane. The surface area and density (thickness) of hydrogel membrane is determined in conjunction with the evaporation rates required for effective downdraft cooling in Tucson's hot-arid climate.

2.1.1.2. Hydration and Saturation Rates

The rate of diffusion from the hydrogel membrane will inform the amount of water being used and the cycles for hydration that will be required. Initial saturation studies show that the 4% polyelectrolyte hydrogel will absorb water to full loading capacity at an average rate of approximately 60 g/cm^3/hour under hydrostatic conditions, though the actual rate of sorption varies following an isotherm curve.

Hydration cycles will be actuated through integrated sensing techniques. Dry-bulb, humidity, and photometric sensors will be located at select locations around the diffusion membrane, which is divided into four main segments. The sensors are linked to a control platform and algorithm that will indicate to one of four microfluidic pumps to initiate water-pumping to hydrate the membrane. The actuation of hydration is dependent in part on optical sensing of daylight transmission through the diffusion membrane (higher visible transmission indicates the membrane is saturated, while low visible transmission indicated more water may be needed). The other dependent condition for hydration actuation is the surrounding dry-bulb and relative humidity sensing, which indicates whether additional moisture is needed to enable the downdraft cooling process.

2.1.1.3. Volumetric Expansion and Contraction Kinetics

During the hydration and dehydration studies of the polyacrylamide hydrogel, measures for the respective swelling and deswelling kinetics were established. Maximum loading occurs when the hydrogel has reached full saturation capacity and is at its largest volumetric size. Dehydration to a completely dry state provides the minimum possible dimensions of the hydrogel. The dimensional change ratio is defined as the relation between the difference of these two states to the maximum loading.

The pilot test for deswelling kinetics demonstrates a dimensional change for three volumetric samples over a three-day period. The hydrogel exhibits an ability to modulate by 25% of its original dimensions and up to 50% of its original volume. The dimensional change is reversible. The modulation of both the weight and

dimensions inform the design for the membrane support framework, which requires a flexible mesh fabric to provide a composite hydrogel system (Fig. 2).
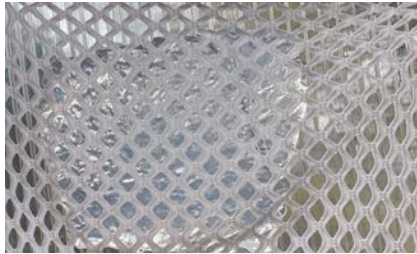
Figure 2. Hydrogel adhered to flexible mesh

2.1.2.   Thermal Storage and Radiant Cooling

Thermal storage and radiant cooling phenomena are assessed for two different translucent materials. Physical material tests are conducted *in-situ* for the materials utilizing thermal sensing and thermography techniques.

2.1.2.1. Physical Material Testing

Material module pilot testing was conducted in Tucson on the rooftop of the building where the full-scale downdraft chimney prototype will be located. The pilot test data compares two heat capacitance and radiative cooling materials: Polyelectrolyte Hydrogel (PEG) and Phase Change Material (PCM) as shown in Figure 3. The PEG was synthesized at the Arizona Material Laboratories and is fully saturated with deionized water. The PCM is 100% paraffin with a 24°C melt point and was obtained from Microtek Laboratories.

Figure 3. Material pilot test setup: a) Polyelectrolyte Gel (PEG) and Phase Change Material (PCM) testing modules; b) FLIR camera used for thermography images; and c) Thermistor sensing for material temperature data collection

Each material module is encased in clear polyethylene and embedded with a thermistor to obtain 1-minute intervals for material temperatures. A rooftop weather station is also providing local micro-climate data for reference. The test-data was collected over a two-day period. The first day enables the heat-capacitance and radiative materials to normalize with outdoor weather conditions. The second day provides the baseline dataset. Initial test-data is provided for one testing period in summer season. Select intervals of

IR-thermography are captured for the material modules during the testing timeframe.

Comprehensive thermal mass and radiative material testing is also currently underway and will be conducted for seasonal comparisons across an annual timeframe. The comprehensive testing matrix consists of nine configurations, each with four material modules (Fig. 4). Each configuration includes four clear polyethylene bags filled with different materials: water, polyacrylamide hydrogel, PCM, and one painted white with no filling. There is one configuration oriented at 180-degrees upwards for full sky-exposure reference. The other configurations are oriented at 90-degrees vertical and 45-degrees diagonal. Each cardinal orientation (North, East, South, and West) is documented concurrently during the testing in conjunction with microclimate data collected with the *in situ* weather station.
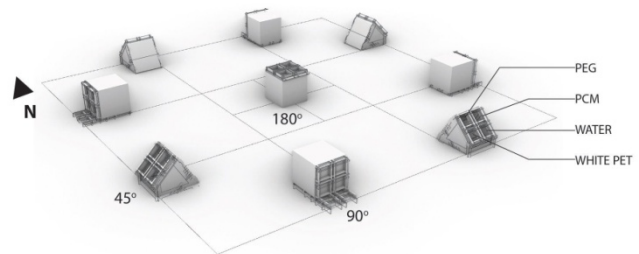
Figure 4. Diagram of the comprehensive heat capacitance and radiative material testing setup

2.2   Full Scale Prototype Simulations

The structure of the full scale prototype is currently under construction. Several simulations were conducted in order to optimize the self-shading of the chimney form and airflow into the building's interior.

2.2.1.   Buoyancy Airflow from Evaporative Cooling

Analytical Model

The convective heat flow induced by evaporative cooling is first analyzed with a simplified analytical model used to determine the basic dimensions of the chimney and is then simulated by a CFD model in order to provide an understanding of the distribution of cooling and air under the chimney, i.e., inside of the building.

The free convection flow estimation can be summarized in Equations 2-5:

The difference between a target temperature (T_{in}) for the interior and outside air temperatures (T_{out}) during the day, together with the air's moisture content [7]. If we change the air's moisture content at the chimney's crown (assuming air as a constant pressure ideal gas), its density will increase by the following relationship:

$$\Delta\rho = \rho_{out} - \rho_{in} = \rho_{in} \times \frac{(T_{in}-T_{out})}{T_{out}} \qquad (2)$$

where $\Delta\rho$ is the density difference between the ambient air and the cooled air; ρ_{out} and ρ_{in} are the densities of the air

outside and inside the chimney, respectively. The heat transfer (Q_{cv}) necessary to cool dry air by a specific temperature difference is given by:

$$Q_{cv} = \dot{v} \times C_{pa} \Delta T \qquad (3)$$

where $\dot{v}$ is the volumetric flow rate per hour, defined by the number of people that occupy the cooled space; C_{pa} is the volumetric specific heat of dry air; ΔT is the temperature gradient between the ambient and the cooled air.

The dense cooled air will drop down the chimney through free convection. We can find the cooled air's flow rate through:

$$Q_{flowrate} = V \times A \qquad (4)$$

where A is the cross-sectional area of the chimney; V is the velocity of the air given by:

$$V = \sqrt{2gl \times \frac{\Delta\rho}{\rho_{out}}} \qquad (5)$$

where g is the gravitational acceleration; l is the height of the chimney; $\Delta\rho$ is the density difference between the ambient and the cooled air; ρ_{out} is the density of the air outside the chimney. The rate of water mass evaporated into the air from the hydrogel membrane surface was analyzed based on the diffusion rate described in section 2.1.1.1. At an average diffusion rate of g/cm²/hour, it is sufficient for the air to be exposed to a quarter of the chimney's top funneled surface area, lined with the hydrogel membrane, (area of 9850 cm²) for the air to cool down by 11°C and drop at a velocity of 1.2m/s. This was the basis for the CFD simulation described in the next section. Additionally, the principle of the Coanda effect was maintained in regards to the surface curvature.

CFD Simulation Model
Computational Fluid Dynamic (CFD) studies were conducted with the Butterfly plug-in for Rhino-Grasshopper and the OpenFoam analysis engine. CFD analyses were used to determine the relationship between the chimney geometry and airflow speeds and morphologies into the space below. The parametric study addresses different heights and diameters of base, neck, and top zones of the chimney. Three iterations of chimney heights were evaluated, including 2-meters, 4-meters, and 6-meters. Three iterations of neck position were evaluated with locations at the midpoint, one-third from the top, and one-third from the base. The top, neck, and base diameters remain constant throughout the study.

The CFD simulation model is set up with the chimney located over the center of a square dimension space and depicts the inlet at the top of the chimney and one outlet at the wall of the space as shown in Figure 5. The analysis planes are established through the mid-sections of the space for plan and section views of the airflow results. The analysis grid was made with snappyHexMesh and blockMeshDict.
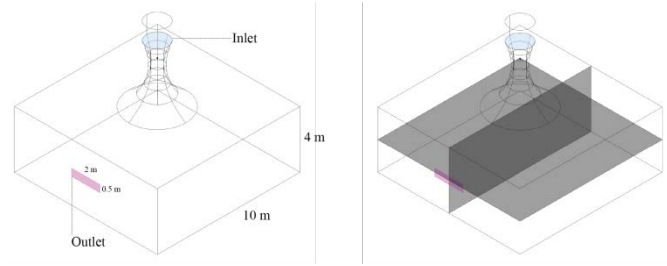
Figure 5. CFD simulation model setup depicting inlet at top of chimney and single outlet in space below (left) with two evaluation planes for plan and section analysis (right)

The current CFD analyses establish basic downdraft airflow patterns with an assumed inlet for downward airflow of θ=180⁰ at a speed of 1.2 m/s from the top of the chimney. Future analyses will make use of local wind speeds and directions entering the top of the chimney from its side-apertures. The current analytical model for the CFD simulation is based on the finite volumetric method for simplification of understanding flow regime into the space below [22-23].

2.2.2. Radiative Heat Exchange

Solar Irradiation Analysis
In order to prevent overheating of the thermal mass by solar radiation gains, it is necessary to provide sun-shading to the material modules while still exposing them as much as possible to the cold sky. A series of radiation analysis simulations were conducted during the design development for the chimney framing system. The solar gain seasonal and diurnal radiation analyses were conducted with the Ladybug-Honeybee plugins for the Rhino-Grasshopper models [24]. The radiation analyses were informative in defining the depth of framing members to provide adequate shading during peak temperature conditions. The overall geometry of the chimney and orientation of framing members was also informed in part by the radiation studies. Ultimately, the radiation analysis for peak summer conditions defined the guide for locating appropriate material modules and establishing heat capacitance properties (i.e. thicknesses for hydrogel modules and melting temperatures for PCM modules). The results of solar irradiation on the chimney surface, in kWh/m² must be then multiplied by a reflectivity ratio of the material, to receive the heat gain of the actual thermal mass surface. Based on previous literature [5,6], we propose a shortwave reflective coating over the thermal mass membrane exposed to solar radiation, which is still emissive in the longwave to allow sky cooling. Thus, the solar heat gain can be reduced by about 90 percent.

Longwave Heat Loss to the Cold Sky
In order to estimate the radiative cooling impact of the sky, a simulation was conducted for test-points along the chimney's surface. Using the surface geometry modeled in Rhinoceros 3D modeling software, the simulation code was developed in Grasshopper algorithmic modeling. This

simulation technique was initially developed in the author's previous work [25] accounting for MRT per point in space. In this case, an array of equally spaced test points was placed on the surface of the chimney. For each point, the view factor of the sky and other surrounding surfaces was computed and the total radiation heat loss or gain from longwave radiation was calculated. Clear skies were assumed for this simulation, although cloudiness downwelling effect can be integrated based on TMY weather files. The clear sky temperature was assumed to vary between 1∘C to 12∘C based on sky temperature models by Berdahl and Fromberg [26] and Garg [27], for low dew point with clear sky. A continuous ambient air temperature is derived from TMY weather files. In the sample results shown here for one hour in the month of June, the air temperature was assumed to be 32∘C. The interior air temperature inside the chimney (and the interior face of the thermal mass) was assumed to be maintained at 27°C due to the evaporative cooling downdraft airflow. Note that longwave radiation was accounted for from the sky and surrounding surfaces. The longwave portion of the solar radiation was included in the previous solar simulation and therefore not included in this simulation.

The results in kWh/m^2 are for a perfect blackbody emitter, while in reality, they must be multiplied by the emissivity coefficient of the surface material for the actual cooling potential.

3 RESULTS AND ANALYSIS

3.1 Material Testing Results

Results from the heat capacitance and radiative material tests are shown for the 24-hour cycle. The IR thermography images depict the fully loaded hydrogel and PCM modules at four-hour intervals from 9:30am on July 12th to 9:30am on July 13th (Fig. 6). The IR images were analyzed with the online FLIR software and assessed with two comparison zones per image: a) thermal mass material zone, and b) surrounding roof zone. The minimum and maximum IR surface temperatures for the material zone of the PEG are 7 ^{o}C at 5:30 AM and 61.3 ^{o}C at 1:30 PM respectively. The minimum and maximum IR surface temperatures for the material zone of the PCM are 6.4 ^{o}C at 5:30 AM and 60.9 ^{o}C at 1:30 PM respectively.

The thermistor sensing data was collected at 1-minute intervals for the 24-hour period and shows the material temperature fluctuations rising towards 60 oC by 2pm in the afternoon and dropping to minimum temperatures between 2am - 8am in the morning (Fig. 7). The minimum and maximum core temperatures for the PEG are 24.8 ^{o}C at 1:30 AM and 59.8 ^{o}C at 4:00 PM respectively. The minimum and maximum core temperatures for the PCM are 12. 6 ^{o}C at 6:15 AM and 60.1 ^{o}C at 1:45 PM respectively.

The PCM exhibits an ability to rapidly release more heat content in comparison with the hydrogel. Conversely, the hydrogel more rapidly increases its sensible heat as soon as

it begins to receive direct solar radiation in the morning timeframe. It is important to note that if shortwave reflective coating is applied to the envelope of the thermal mass, the solar heat gain would be reduced significantly, while the sky cooling will be maintained. The separation of the thermal mass from direct ambient gain, as shown in [5], also contributes to its radiant cooling capacity.
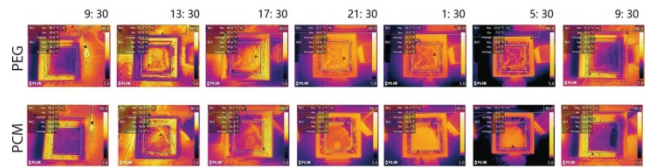
Figure 6. Infrared thermography images for PEG (top) and PCM (bottom) at four-hour intervals during pilot-test period July 12th-13th, 2019
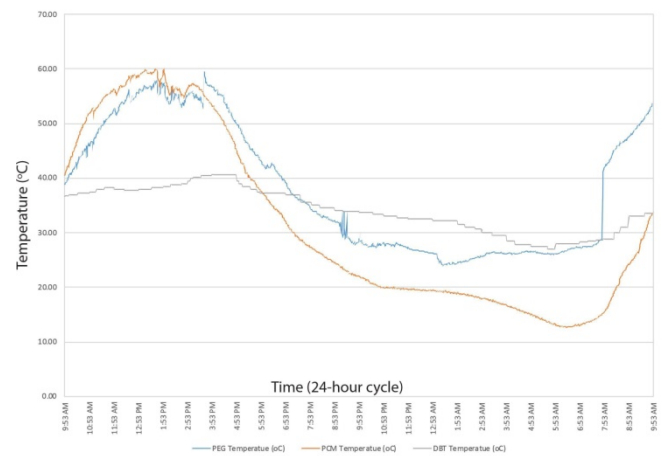
Figure 7. Material thermistor measures and dry-bulb temperature during pilot-test period July 12th-13th, 2019

3.2 Full-Scale Prototype Simulation Results

3.2.1. Airflow CFD results

The results of the parametric CFD simulations show that the tallest chimney height (6-meters) provides the greatest air velocity into the space below. Since the full-scale physical prototype will be constructed with a 4-meter height, the results for the 4-meter height chimney are shown for this study in Figure 8. The optimal neck position for the 4-meter height is located at the midpoint of the chimney to provide greater air speed for ventilation cooling effects throughout the space below. Since the number of outlets, outlet size, and outlet position also influence the airflow patterns and velocities, further studies will be conducted to evaluate optimal spatial choreography of the airflow.

The downdraft air speed maintains the highest velocity in the space directly below the chimney. Vortices form in the spaces of the room to the sides of the chimney aperture and follow natural convective buoyancy patterns. The air speed at the outlet increases in velocity due to the pressure differentials between outdoor and indoor conditions.
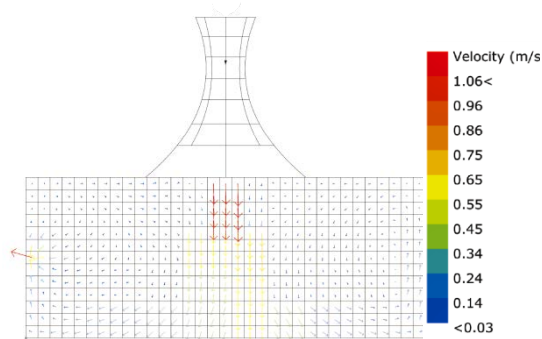
Figure 8. CFD results for optimal 4-meter height chimney

3.2.2. Radiative Heat Exchange Results

Solar Irradiation Analysis
The results show the radiative heat gain pattern of the chimney surface ranging between 0 kWh/m^2 for the fully shaded areas to 9 kWh/m^2 at the upper limit of solar heat gain per day, for completely unshaded areas (Fig. 9). For the actual material surface, assuming a shortwave-reflective coating, we can assume a multiplier of 0.1 for the results, reducing the heat gain significantly.
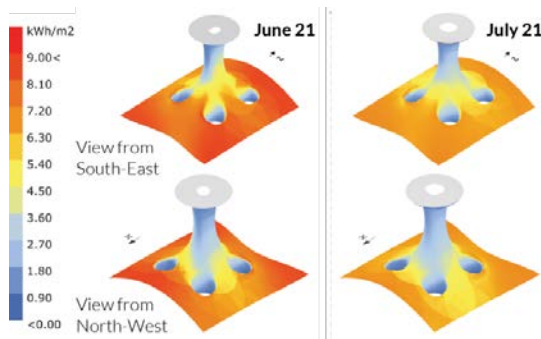
Figure 9. Solar irradiation simulation results for the chimney's surface for the duration of a full summer day of June and July 21st

Longwave Heat Loss to the Cold Sky
Figure 10 shows a sample result for one hour of longwave radiative heat exchange between the surface of the chimney and its surrounding in clear sky conditions in Tucson, Arizona in June. The results show heat gain rather than heat loss at the very top of the chimney due to the self-shading geometry preventing an exchange with the sky. The results also show heat gain at the bottom of the chimney, where its surface is in great proximity to the hot roof . The areas most advantageous for radiant sky cooling are in the middle portion of the chimney, and especially in the north and south portion as seen in Fig. 9. This is due to the surface slope in these portions of the chimney and its relationship to the overall geometry.

The results ranging between 0.0 and 0.09kWh/m^2 are for one hour of exposure only, unlike the radiation results which are for the full day. The heat loss will be greater at night and in the morning hours when the surrounding roof is cold and not contributing heat gain to the surface. The results should then be multiplied by the emissivity ratio of the thermal mass enclosure, which would reduce the total heat loss by about 10 percent.

If we compare the pattern from Figure 10 to the results from the solar irradiation analysis (Fig. 9), we can see that while the middle zone of the chimney is both well shaded from direct sun but still achieves heat loss at lower portion of the south and north segments that are the most advantageous for radiative cooling from the sky are also receiving large fluxes of solar heat gain during the day. For this reason, shortwave reflective coating should be applied to the modules placed in that portion to optimize heat loss and minimize heat gain. The combined results analysis are the guidelines by which the thermal mass will be placed on the final prototype and appropriate wavelength-specific coatings will be applied.
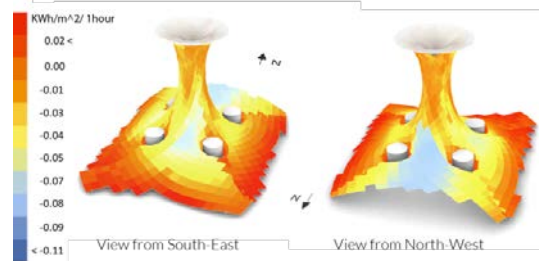
Figure 10. Longwave radiation heat loss/gain simulation results for one hour in the month of June with clear skies.

4 CONCLUSION AND FUTURE WORK

The various analytical, empirical, and simulation testing methods explored in this paper show the cooling potential for a combined radiative and evaporative cooling device in a desert climate. The analytical models for evaporative cooling through diffusion of moisture from the hydrogel membrane demonstrates the capacity of surface evaporation to create a buoyancy force sufficient to cool a significant volume below the chimney. Future CFD simulation models will combine the advanced characteristics of the chimney materials, including: a) moisture diffusion from the cooling membrane at the top (i.e. buoyancy model), and b) heat capacitance material temperature dynamics for the thermal mass modules of the chimney. The empirical test data for the material systems will be integrated into the CFD simulation model. The heat capacitance material tests and longwave radiation heat loss simulations show the combined ability for thermal lag and sky re-radiation to provide additional cooling and attenuate the heat peak during the day.

The next steps include the completion of the full-scale prototype currently under construction. The full-scale prototype will enable the collection of empirical data for downdraft cooling effects and air speeds at the space below the chimney. The empirical data can be compared with the ongoing development of the CFD simulation models. The radiant cooling capacity of the envelope will also be further explored through the full-scale prototype for validation of the simulation results.

REFERENCES

1. Givoni, B. (1993). Semiempirical model of a building with a passive evaporative cool tower. *Solar Energy*, *50*(5), 425-434.

2. Erell, E., Pearlmutter, D., & Etzion, Y. (2008). A multi-stage down-draft evaporative cool tower for semi-enclosed spaces: aerodynamic performance. *Solar Energy*, *82*(5), 420-429.

3. Altman, T., Zaslavsky, D., Guetta, R., & Czisch, G. (2005). Evaluation of the potential of electricity and desalinated water supply by using technology of" Energy Towers" for Australia and America. *Interim Report.*

4. Craig, S., Harrison, D., Cripps, A., & Knott, D. (2008). BioTRIZ suggests radiative cooling of buildings can be done passively by changing the structure of roof insulation to let longwave infrared pass. *Journal of Bionic Engineering*, *5*(1), 55-66.

5. Raman, A. P., Anoma, M. A., Zhu, L., Rephaeli, E., & Fan, S. (2014). Passive radiative cooling below ambient air temperature under direct sunlight. *Nature*, *515*(7528), 540-544.

6. Goldstein, E. A., Raman, A. P., & Fan, S. (2017). Sub-ambient non-evaporative fluid cooling with the sky. *Nature Energy*, *2*(9), 1-7.

7. Aviv, D., & Meggers, F. (2017). Cooling oculus for desert climate–dynamic structure for evaporative downdraft and night sky cooling. *Energy Procedia*, *122*, 1123-1128.

8. Aviv, D., & Kilian, A. (2018). Climate-Adaptive Volume: Solving the Motion Envelope of a Reconfigurable Cooling Aperture for Desert Climate. *Technology| Architecture+ Design*, *2*(2), 186-195.

9. Charoudi, D. U.S. Patent No. 3 953 110 (27 April 1976).

10. Designbloom. "IAAC students create hydroceramic: a passive cooling system for buildings." *designbloom*. (2014). Retrieved from https://www.designboom.com/architecture/iaac-dmic-hydroceramic-passive-cooling-system-09-18-2014/

11. Smith, S. I. (2017). Superporous Intelligent Hydrogels for Environmentally Adaptive Building Skins. *MRS Advances*, *2*(46), 2481-2488.

12. Moradnejad, M., Aviv, D., Ida, A., & Meggers, F. (2019). WATeRVASE: Wind-catching Adaptive Technology for a Roof-integrated Ventilation Aperture System and Evaporative-cooling. *Building Technology Educator's Society*, *2019*(1), 50.

13. Fornasiero, F., Tang, D., Boushehri, A., Prausnitz, J., & Radke, C. J. (2008). Water diffusion through hydrogel membranes: A novel evaporation cell free of external mass-transfer resistance. *Journal of Membrane Science*, *320*(1-2), 423-430.

14. P. M., V., Bayraktar, O., and Picó, G. *Polyelectrolytes Thermodynamics and Rheology (Engineering materials)*. (2014). New York, NY: SpringerLink.

15. Huang, C., Dostalek, J., and Knoll, W. "Long range surface plasmon and hydrogel optical waveguide field-enhanced fluorescence biosensor with 3D hydrogel binding matrix: On the role of diffusion mass transfer." *Biosensors & Bioelectronics*, *26(4),* (2010): 1425-31.

16. Bhattacharyya, D., Schäfer, T., Wickramasinghe, S. R., & Daunert, S. (Eds.). (2013). *Responsive membranes and materials*. Chichester: Wiley.

17. Flory, P.J. and Rehner, J. "Statistical Mechanics of Cross-Linked Polymer Networks: II. Swelling." *Journal of Chemical Physics, 11* (1943): 521-526.

18. Richter, F. "Timescales determining the degree of kinetic isotope fractionation by evaporation and condensation." *Geochimica Et Cosmochimica Acta*, *68(23),* (2004): 4971-4992.

19. Jirka, G. H., & Brutsaert, W. (1984). Measurements of wind effects on water-side controlled gas exchange in riverine systems. In *Gas Transfer at Water Surfaces* (pp. 437-446). Springer, Dordrecht.

20. Bellucci, M., and Trout. "Note: On the Evaporation Coefficient of Water." *The Journal of Chemical Physics*, *21* (2014): Vol.140(15).

21. Eames, I.W, Marr, N.J and Sabir, H. "The Evaporation Coefficient of Water: A Review." *International Journal of Heat and Mass Transfer, 40.12* (1997): 2963-973.

22. Welahetti, P., & Vågsæther, K. (2016). Comparison of OpenFoam and ANSYS Fluent. Computational Fluid Dynamic Simulation of Gas-Gas Single Phase Mixing with and without Static Mixer.

23. Hedayat, Z. H., Samkhaniani, N., Belmans, B., Ayatollahi, H., Wouters, I., & Descamps, F. (2015, December). Energy modeling and air flow simulation of an ancient wind catcher in Yazd. In *3rd International Congress on Civil Engineering, Architecture and Urban Development* (pp. 29-31).

24. Roudsari, M. S., Pak, M., & Smith, A. (2013, August). Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. In *Proceedings of the 13th international IBPSA conference held in Lyon, France Aug*.

25. Aviv, D., E. Teitelbaum, T. Kvochick, K. Bradford, and Forrest Meggers.(2019). Generation and Simulation of Indoor Thermal Gradients: MRT for Asymmetric Radiant Heat Fluxes. In *Proceedings of Building Simulation*. Rome, Italy: IBPSA.

26. Berdahl, P., & Fromberg, R. (1982). The thermal radiance of clear skies. *Solar energy*, *29*(4), 299-314.

27. Garg, H. P. (1982). *Treatise on solar energy*. John Wiley & Sons.

Configurational Optimization of Building Details via Parametric Numeric Simulation: A Case Study of Windows with Vacuum Glass

Magdalena Wölzl[1], Ulrich Pont[1], Ardeshir Mahdavi[1], and Peter Schober[2]

[1]Department of Building Physics and Building
ecology, TU Wien
Vienna, Austria
magdalena.woelzl@tuwien.ac.at

[2]Holzforschung Austria
Vienna, Austria
p.schober@holzforschung.at

ABSTRACT

This paper deals with the integration of highly insulating vacuum glass products into contemporary window constructions. Whereas the research regarding the production of durable vacuum glass products has a relatively long history, there are few research efforts pertaining to the integration of such glass products into window constructions. In the present contribution, we deploy numeric thermal simulation to explore the effect of different design and construction decisions on windows' performance. Preliminary results suggest that an iterative simulation-based approach has the potential to provide detailed predictions of the components' thermal behavior. Thus, initial designs can be reliably appraised and improved alternative designs developed and virtually tested.

KEYWORDS

Vacuum glass; parametric simulation; numeric thermal bridge simulation and analysis; window construction.

1 INTRODUCTION

1.1 Motivation

Window constructions are often considered to be the weaker components of buildings' envelope in view of heat loss, overheating risk, and sound insulation. It has been suggested that between 30% and 60% of buildings' heat losses can be attributed to transparent components [4, 9].

Whereas window construction technology has gone through a number of innovations and improvements in the past, the introduction of vacuum glass brings about new challenges. These include the need for different construction principles in comparison to conventional double or triple-glazed windows. Therefore, recent research and development efforts currently conducted at the TU Wien, Vienna, Austria and the Austrian Forest Products Research Society (Holzforschung Austria) together with stakeholders from industry target the development of contemporary windows with vacuum glazing. Thereby, parametric numeric simulation is deployed to study the thermal performance of windows, both at the overall window level (U-value, heat flow) as well as in evaluation of critical spots and thermal

bridges. The present contribution provides an overview of the methodology and the results of the thermal simulations of windows equipped with vacuum glazing [17].

1.2 Background

Generally speaking, the term vacuum glass refers to two parallel glass panes with a thin evacuated interstitial gap. Equipped with a tight edge seal and a grid of distance-maintaining pillars, such glass products widely eliminate convective and conductive heat transfer processes. The first relevant considerations regarding evacuated glass products date back to 1913 to the German inventor A. Zoller [18]. Since then, several attempts to construct durable evacuated glazing products have been conducted. In recent years, products from South-East Asia seem to have reached an acceptable durability (see [16]).

Most vacuum glass products use similar elements: Figure 1. principally illustrates the constituents and parameters of a vacuum glass panel in a window/frame construction (glass panes, evacuated interstitial space, distance pillars, glass edge seal and cover length). In a previous study, certain characteristics of vacuum glazing and related construction principles were addressed. For instance, it was recommended to maintain a glass edge cover of at least 40 to 50 mm. Moreover, it was suggested to minimize the glass edge seal toward mitigation of thermal bridging effects in glass/frame.

Regarding research on vacuum glass use for windows in Europe, the projects VIG-SYS-RENO [1] and MOTIVE [12] are relevant. VIG-SYS-RENO – as well as a follow-up project called VAMOS [15] – looked into the principle feasibility of vacuum-glass application in existing window constructions, such as casement windows. In contrast, the project MOTIVE addressed new and innovative windows with vacuum glazing products. Both projects utilized state-of-the-art numeric thermal bridge simulation for further development of such windows. This work is being continued in the ongoing project FIVA [4], where thermal simulation efforts constitute one of the main work packages (see Figure 2.).
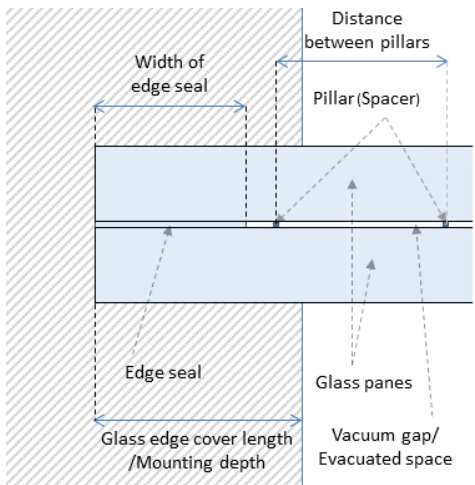
Figure 1. Vacuum glass in a generic frame material: Terminology and Characteristics [14, 17].

Figure 2. Workpackages in the FIVA-project [4]

2 RESEARCH DESIGN

2.1 Evaluated window design concepts

Four different window concepts were selected for detailed thermal simulation (see Figure 3). These are:

- Turn window opening to inside (I.)
- Turn window opening to outside (A.)
- Swing window (K.)
- Sliding window (S.)

Certain aspects of the selected constructions are briefly described below.

Turn window opening to inside (I.)
This window concepts represents the typical Central-European window in view of its operational possibilities. Whereas a high degree of user acceptance can be expected, still the window frame was constructed based on the necessities of vacuum glass. The construction allows for a smooth integration in the internal surface plane of walls. Comparable to typical windows of Central Europe the integration of external shading devices is possible without problems. Due to the usage of special fittings, a small-step opening is possible for ventilation purposes.

Turn window opening to outside (A.)
The opening direction has not been very common in Central Europe in the past decades, however it is well established in

Scandinavia. Due to the integration of special fittings the window can be parrallely positioned to the façade plane for ventilation purposes. While in the fully opened position the usage of external shading seems difficult, this ventilation position might allow for simultaneous ventilation and shading. Regarding the architectural appearance of this concept it can be said, that the very reduced aesthetics might be attractive for planners (Window clearance and glass clearance show same dimensions if looked upon from the inside).

Swing window (K.)
This operational concept is inspired by garage doors of the 60ies and 70ies. The wing swings to the inside via a fixed turn rotation point. Advantages include, that the wing is moved in a parallel position to the ceiling, thus no useful space on the interior is consumed. Moreover, the fixation of the window can be realized in a very simple way, and large window formats are possible. Potential challenges are the implementation of shading and the complex kinematics of the window operation. Moreover, this window potentially requires a motor more than the other window concepts, and the sealing around the rotational point might be challenging.

Figure 3. Window concepts examined in this contribution: From left to right: Turn window opening to inside (I.), Turn window opening to outside (A.), Swing window (K.), Sliding window (S.)

282

Sliding window (S.)

Based on previous design studies pertaining to a sliding movement (from [12]), a further development lead to a combined window opening direction. The window is pushed to the outside and then slid to the side on telescope rails. This concept comes with a set of advanteges, such as easy sealing, and highly-attractive appearance from the outside due to hidden rails as well as from the inside to to easy integration in the opaque wall around. Moreover, the reduced weight of the vacuum glazing allows for easy operation of the window in comparison to heavy weight triple glazing. Challenges pertaining to this type include the integration of external shading, development of proper fittings, and the limited extension range of the telescope rails. Furthermore, telescope rails available on the market come from different domains than the window industry. As such, telescope rails that can be used as window fittings have yet to be developed.

2.2 Simulation environment and input data

For the purpose of iterative simulation-assisted improvement of the different window designs, the state-of-the-art numeric thermal bridge simulation tool AnTherm [1, 2, 10] was selected. This tool relies on the finite difference method to model the thermal performance of 2D and 3D thermal bridges.

To perform numeric thermal bridge simulations of window details, a set of input data information and boundary conditions are needed. Moreover, the window details have to be drafted either in the simulation tool or can be imported from a CAD/BIM environment.

Colour	Material	λ [W.m^{-1}.K^{-1}]
	Aluminium	200
	Compact Foam	0.031
	Glass	1
	Masonry	0.45
	Plaster	0.7
	Purenit	0.096
	Seal	0.3
	Seal encapsulate	0.04
	Silicone	0.35
	Steel	60
	Wood	0.13
	Encapsulate air	0.5555
	Vacuum	0.00000975

Colour	Boundary conditions	$R_{s(H,T)}$ [m^2.K.W.$^{-1}$]
	Inside, +20 °C	0.25
	Outside, -10 °C	0.04

Table 1. Input data for simulations. [5, 11, 13]

Required input data encompass the thermal conductivities (λ values) of materials used in the window construction, the climatic boundary conditions (indoor and outdoor temperature), as well as surface resistance co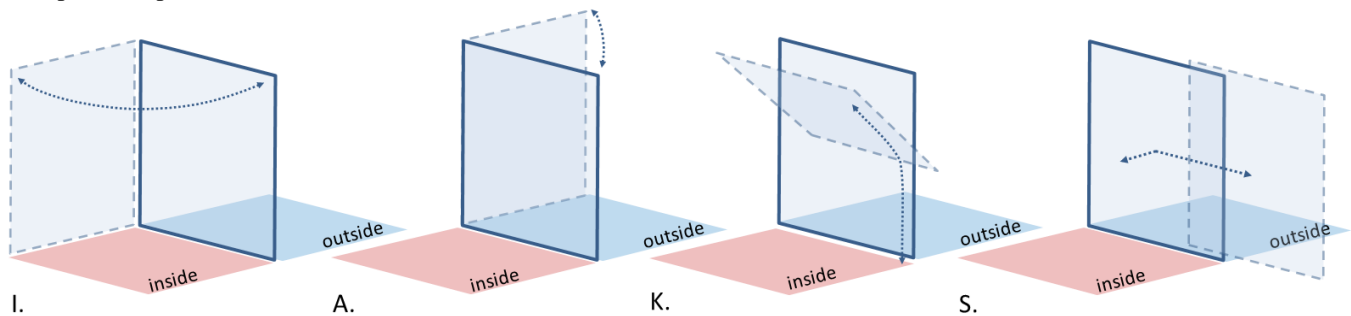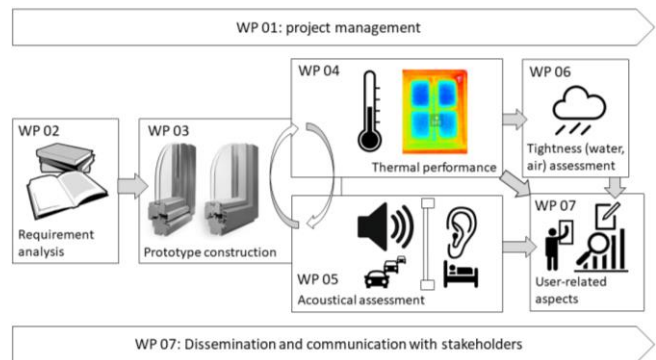efficients. Additionally, encapsulated air/gas layers as well as vacuum layers have to be approximated via effective λ-values, which are derived using normative procedures [5, 8] . Table 1. provides an overview of the input data used for the purpose of the present contribution.

2.3 Derived indicators

Hygro-thermal performance evaluation of window constructions typically concerns surface condensation and mould growth risk. Inside surface temperatures of window constructions may thus not fall below certain values even under extreme outdoor temperatures (e.g., -10 °C). Surface temperatures can be computed and given in absolute values at any point of the simulated construction detail. However, it is more common to use the so called f_{Rsi}-value [7] as per the following equation:

$$f_{Rsi} = \frac{\theta_{si,min} - \theta_e}{\theta_i - \theta_e} \; [-]$$

Here, $\theta_{si,min}$ denotes the lowest indoor surface temperature (°C), and θ_e and θ_i the exterior and interior air temperatures. Threshold values for f_{Rsi} are defined in different standards. In Austria, threshold values of 0.69 and 0.71 have been defined [13] for the evaluation of condensation risk and mold growth risk respectively.

3 RESULTS AND DISCUSSION

3.1 Turn window to inside (I.)

Variant I.1 illustrates the basic design of the turn window opening to inside. This rather traditional window design was equipped with a vacuum glass pane attached to the inner surface of the (wooden) frame construction. Figure 4 illustrates the window detail, isotherm graph, and the surface temperature profile along the inner surface. The analysis points to condensation risk at the interface between the operable window wing and the fixed frame. The lowest temperature is 0.09 °C, while the f_{Rsi}-value is 0.34 (which is well below the required threshold values of 0.69 and 0.71). To improve this situation, the window variant I.2 was equipped with a set of aluminum component adjacent to the inside space, which conduct heat to this critical point. While this measure works fine in the simulation, it may be difficult to realize in terms of a practical construction solution. Moreover, given the simplified approach in the modeling of the boundary conditions, the actual surface temperature at this critical point may be even lower, pointing to a higher risk of condensation risk.

The thermal performance of this variant is by far better than the original one, but still slightly below the threshold values (minimum temperature 9.96 °C, and f_{Rsi}-value of 0.67). Figure 4. illustrates false-color images of both variants and compares the indoor surface temperature profiles.

3.2 Turn window to outside (A.)

Similar to variant I., two different design variants of the variant A are considered, namely A.1 and A.2 Both involve the attachment of the vacuum glass pane to the outer surface of the (wooden) frame. In version A.1, the glass edge cover length was set to 40 mm, which corresponds to the minimum construction necessity for window frame (structural stability, integration of fittings). From the perspective of thermal performance, this length, however, can be considered as insufficient. Results show a minimum inside surface temperature of 8.18 ˚C and a f_{Rsi}-value of 0.61, which is well below the minimum thresholds. In variant A.2, the glass edge cover length was increased to 68 mm , which resulted in an acceptable minimum surface temperature of 10.93 ˚C: Even if the f_{Rsi}-threshold value of 0.71 is not reached, minor improvement of this construction is likely to satisfy the requirements. Figure 5. illustrates the two versions of the window.
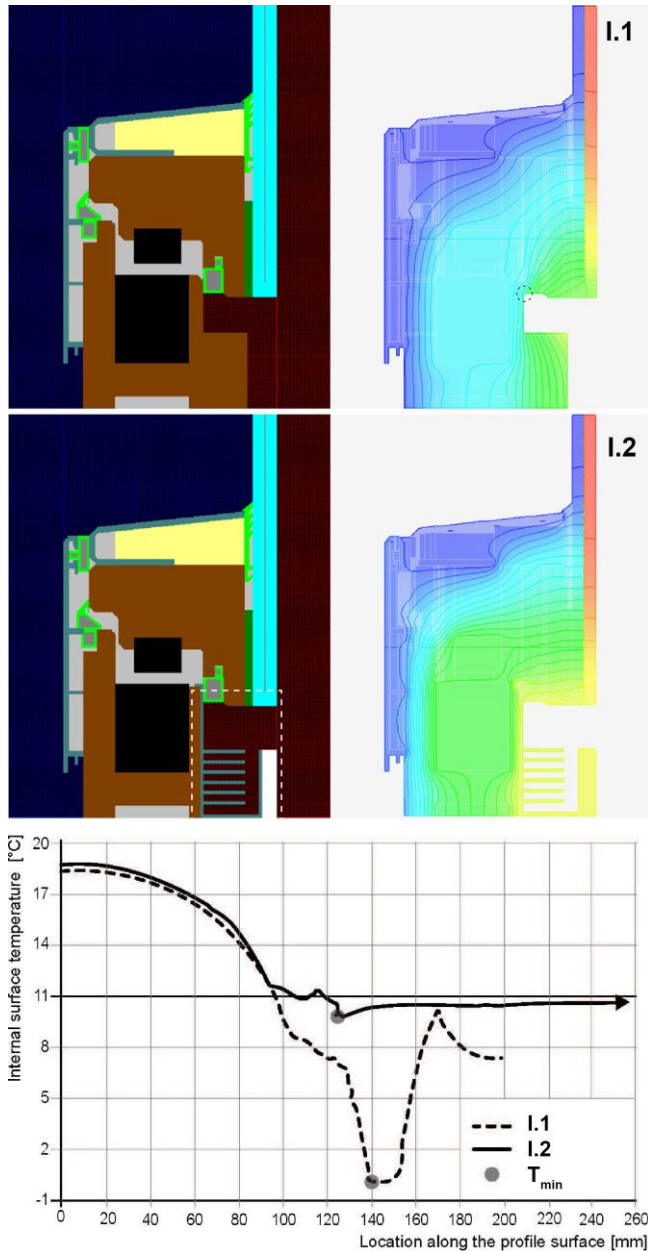
Figure 4. Turn window opening to inside (I.)
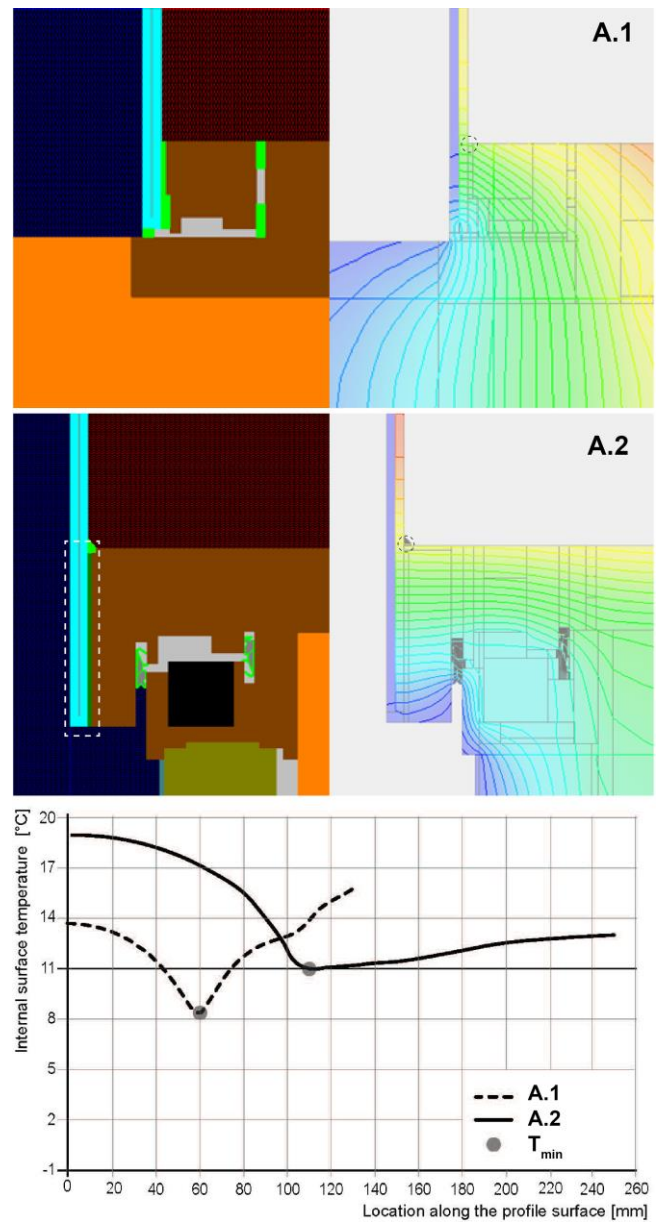
Figure 5. Turn window opening to outside (A.)

284

3.3 Swing window (K.)

For the swing window, two versions of the lower window/frame joint were considered. K.1 displays a f_{Rsi}-value of 0.67 (and an inside temperature of 10.11 °C), which is not too far away from the minimum threshold values for f_{Rsi}. K.2 represents an improved variant, in which a small aluminum sheet is integrated in the frame. This highly-conductive material transports heat to the connection point between frame and vacuum glass, and improves the performance of the construction detail to 11.23 °C minimum indoor surface temperature and 0.71 for the f_{Rsi}-value. As such, the construction detail meets the requirements for condensation risk mitigation and mold growth avoidance. Figure 6. illustrates the two versions of the window.

3.4 Sliding window (S.)

In Figure7., two versions of the construction detail are presented as well: S.1 shows this construction encompassing two seals, which results in a f_{Rsi}-value of 0.62 (minimum indoor surface temperature of 8.6 °C). Variant S.2 integrates a third seal, which hinders the outside air flow into the construction detail. This measure improves the thermal performance of the detail significantly: The f_{Rsi}-threshold value (0.71) is reached with a temperature of 11.42 °C
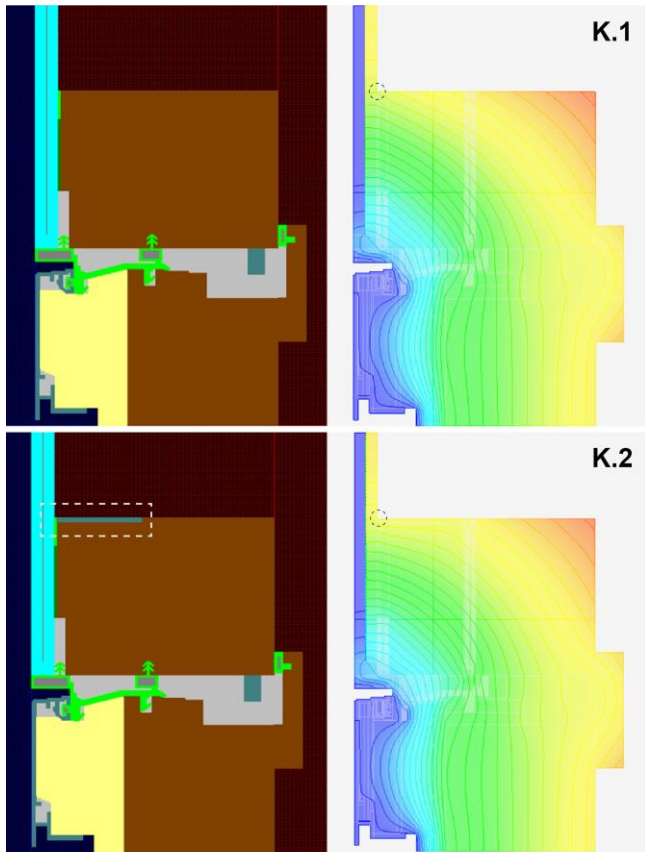
Figure 6. Swing window (K.)
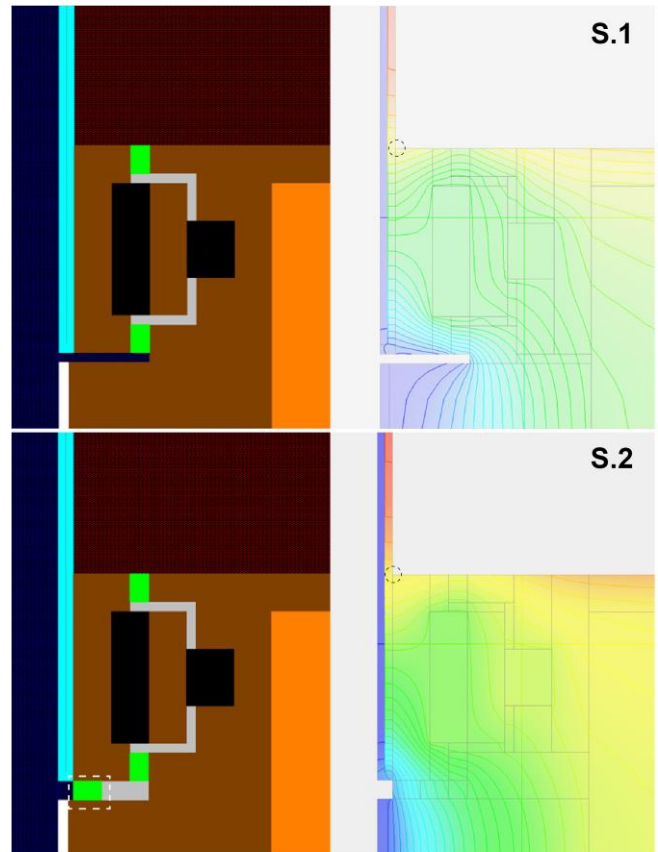
Figure 7. Sliding window (S.)

4 CONCLUSION

4.1 Summary of Findings

This contribution addressed the configurational optimization of innovative windows encompassing vacuum glass products via parametric numeric simulation. Thereby, four different window prototypes were assessed, resulting in a number of observations:

- The positioning of the vacuum glass pane in relation to the frame significantly influences the thermal performance of the window. Generally speaking, the exterior placement of the vacuum glass pane is favorable, as larger portion of the construction maintains higher temperatures. This can be demonstrated, for example, via the comparison of surface temperature profiles of variant I (interior placement) and A. (exterior placement) (see Figures 4. and 5.).

- The relatively large magnitude of heat conduction along the glass pane has to be considered in window constructions. Thereby the thermal bridge of the edge sealing of the vacuum glass has a crucial role, as it represents the thermally weakest spot in the vacuum glass product. Consequently, a large glass edge cover length would be favorable (see Figure 5.).

- The thermal conductivity values of used materials have to be carefully considered. Thermal bridge effects can lead to undesirable heat flows and low indoor surface temperatures. In specific case, however, heat flows may be intentionally directed toward the construction's weak spot (see the impact of inclusion of an aluminum sheet in variant K., as depicted in Figure 6.).

- Generally speaking, the application of highly insulating vacuum glazing can intensify thermal bridging effects and lead thus to lower temperatures in critical locations in the construction. Both exterior positioning of the pane as well as careful design of the glazing-frame details can alleviate the associated risks.

Regarding the four window constructions considered, further observations may be noteworthy:

- **Turn window opening to inside (I.):** In this construction, larger parts of window frame are exposed to the external environment. To mitigate the thermal bridge effect, a highly conductive aluminium component was connected to the inside of the frame, transfering thus heat to the coldest surface spot. It is planned to perform validation measurements for this construction detail in near future with a full scale mock up. Such measurements could be used to calibrate the simulation model for further detailed analysis.

- **Turn window opening to outside (A.):** For this construction the glass edge seal cover is of high importance. If improperly dimensioend, condensation is to be expected at the junction of glass and frame. A larger dimensioning of the frame alltogether (and specifically the glass edge cover) mitigates these effects.

- **Swing window (K.):** This construction displayed a satisfactory performance. Further improvement could be obtained without major redesign via inclusion of a small aluminum sheet to direct heat to glass-frame junction.

- **Sliding window (S.):** In this construction the numeric thermal bridge assessment showed that a significant performance improvement can be realized via a third sealing in the gap between the fixed and the moving part of the window frame. This sealing prevents cold air from entering the construction.

4.2 Limitations & Future Research

The present contribution reported on the results of simulation-based thermal analysis of alternative designs for windows with integrated vacuum glass panes. Certain limitations of this work is briefly highlighted in the following:

- Surface resistance values: The R_{si} and R_{se} values were obtained from standards, as we could not directly measure their values. Generally speaking, physical description of the complex convective and radiative heat transfer phenomena via a few standardized surface resistance coefficients is not satisfactory, given the geometric and material-related variety of the construction elements (glass pane, wooden frames, gaps in fittings and seals, etc.).

- Adiabatic cut-off plane: Frequently, adiabatic cut-off planes are defined in numeric thermal bridge simulation, typically following pertinent standards [5, 6]. However, such standardized conventions insufficiently meet the modeling requirements of windows equipped with vacuum glass. The exact implications of decisions regarding the positioning of adiabatic cut-off planes for the KPI results have not been yet determined.

A number of issues are to be addressed in future research efforts:

- The present contribution relied mainly on simulation of conductive heat transfer processes. However, radiative and convective processes strongly influence the performance of windows. Currently, the potential of a computational environment capable of coupled conductive-convective-radiative heat transfer modeling is being investigated. Using this environment, we intent to further investigate the thermal behavior of window constructions in more detail.

- To add credibility to the computational results, currently experimental measurements on window prototypes in terms of climate chamber experiments are being prepared.

- To foster the acceptability of vacuum glass windows in the market place, provision of common performance

indicators would be beneficial, if not critical. Current standards for the estimation of U-values of window constructions are hardly suitable for components involving vacuum glazing. Developing alternative calculation methods for this purpose is thus one of the points in the agenda of future research in this area.

REFERENCES

1. AnTherm. www.antherm.eu (last checked: Dec 2019).

2. Bucevac, M., Pont, U., Swoboda, S., Mahdavi, A. (2017): *"Data transfer from BIM to building performance simulation tools: A case study"*; Building Simulation Applications - 3rd IBPSA-Italy Conference, ISSN: 2531-6702; Paper ID 74

3. Cuce, E., Riffat, S. B. (2015): *A state-of-the-art review on innovative glazing technologies.* In Renewable and Sustainable Energy Reviews 41, pp. 695–714. DOI: 10.1016/j.rser.2014.08.084.

4. FIVA project, see: https://nachhaltigwirtschaften.at/de/sdz/projekte/fiva.php (last checked: Dec 2019).

5. ISO 10077-2:2017: *Thermal performance of windows, doors and shutters — Calculation of thermal transmittance — Part 2: Numerical method for frames* (ISO 10077-2:2017)

6. ISO 10211:2017: *Thermal bridges in building construction — Heat flows and surface temperatures — Detailed calculations* (ISO 10211:2017)

7. ISO 13788:2012: *Hygrothermal performance of building components and building elements — Internal surface temperature to avoid critical surface humidity and interstitial condensation — Calculation methods* (ISO 13788:2012)

8. ISO 6946:2017: *Building components and building elements — Thermal resistance and thermal transmittance — Calculation methods* (ISO 6946:2017)

9. Jelle, B. P., Hynd, A., Gustavsen, A., Arasteh, D., Goudey, H., Hart, R. (2012): *Fenestration of today and tomorrow: A state-of-the-art review and future research opportunities.* In Solar Energy Materials and Solar Cells 96, pp. 1–28. DOI: 10.1016/j.solmat.2011.08.010.

10. Kornicki, T., Volko, C., Kornicki, M., Pont, U. (2016): *Numeric thermal bridges simulation: Approaching optimized usability for sloped and rounded shapes.* Applied Mechanics and Materials (invited), Special Volume: Energy Saving and Environmentally Friendly Technologies - Concepts of Sustainable Building, 824; p. 527 - 535.

11. Krick, B. (2011): *Wärmebrückenberechnung Zur Zertifizierung Des Fensterrahmens SmartWin Als Passivhaus Geeignete Komponente.* Bericht zur Zertifizierung des Fensterrahmens SmartWin Seite 2 der Firma pro Passivhausfenster GmbH in D-83080 Oberaudorf.

12. MOTIVE project, see: https://nachhaltigwirtschaften.at/de/sdz/publikationen/schriftenreihe-2018-24-motive.php (last checked: Dec 2019).

13. ÖNORM B 8110-2: *Thermal insulation in building construction – Part 2: Water vapour diffusion and protection against condensation.* 2003-07-01, Austrian Standardization Institute.

14. Pont, U., Wölzl, M., Schober, P., Khosravi, S.N., Schuss, M., Mahdavi, A. (2019): *Recent progress in the development of windows with vacuum glass*; in: "Digital Proceedings CESBP2019", R. Cerny (ed.); Matec Web of Conferences, Paper ID 2020

15. VAMOS project, see: https://nachhaltigwirtschaften.at/de/sdz/projekte/vamos.php (last checked: Dec 2019).

16. VIG SYS-RENO project, see: https://nachhaltigwirtschaften.at/de/sdz/projekte/vig-sys-reno.php (last checked: Dec 2019).

17. Wölzl, M (2019): *Fensterkonstrutkionen mit Vakuumglas: Simulationsbasierte Weiterentwicklung von innovativen Fensterkonstruktionen.* Master Thesis, TU Wien, Supervised by U. Pont & A. Mahdavi

18. Zoller, A. (1924): *DE387655C - Hohle Glasscheibe.* Google Patents. With assistance of Dipl-Ing Alfred Zoeller. https://patents.google.com/patent/DE387655C/de, updated on 10/29/1913 (last checked: Dec 2019)

Energy Consumption in Naturally Ventilated DSF and SSFs in Four Different Climate Regions in China

Seyedehelham Sadatiseyedmahalleh[1], Xiaofeng Li [2]

[1] Department of Building Science, School of Architecture, Tsinghua University
Beijing, China
elham.sadatii@gmail.com

[2] Department of Building Science, School of Architecture, Tsinghua University
Beijing, China
xfli@tsinghua.edu.cn

ABSTRACT

The Double Skin façade (DSF) has been popular as a way to improve the energy performance of the building, while its energy performance is still not approved. Numerous researches have analyzed the performance of DSF and Single Skin Façade (SSF) in standard days or typical moments. However, in order to obtain greater accuracy and figure out the best type of facade, the performance of DSF in entire season should be evaluated and compared with SSF. In this study, the heat gains of the optimized DSF and four types of SSFs are obtained to find the total thermal performance of both kinds of façades in the entire cooling and heating seasons in four different climate zones. The quantitative method is applied for evaluating the thermal performance of facades in two seasons. Based on Computational Fluid Dynamics (CFD), a simplified method is used to obtain DSF thermal performance during the cooling season. At the end, according to the total heat gain of each facades, the energy consumption of them in each region were measured. As a result, different performance of facades can be seen for entire cooling and heating seasons in each zone.

AUTHOR KEYWORDS

Double skin façade; Single skin façade; Thermal performance; Energy consumption

1 INTRODUCTION

The field of architecture has been strongly affected by the urgent need to reduce harmful emissions in the atmosphere. The building sector is one of the most energy-consuming sectors in industries all over the world. Buildings have significant potential to minimize emissions through modifications in design and material selection [1]. By introducing this potential, global consciousness on sustainability has increased along with interest in applying passive systems in buildings. In passive buildings, the cooling or heating systems in the interior media of the structure are not regulated. The indoor environment is regulated by the architectural design and structure of the building and its components. At the conceptual level of effective passive design, the elements with crucial influence on building performance are determined, such as orientation and form, glazing type, shading, and wall-to-window ratio. Designers aim to integrate and optimize all these factors to

achieve the best performance. Recently, the Double Skin Façade (DSF) has become popular among all passive design methods. The idea of DSF comes from combining energy efficiency and development with the transparent façade in the modern buildings [2]. However, implementing this method is accompanied by major challenges, due to the complexity of the thermal phenomena and airflow in the DSF, and also the adaptability of combination facades to climatic conditions in various regions [3].

Park, Augenbroe [4] optimized a simple DSF system. Shameri, Alghoul [5] studied the amount of daylight when the DSF was used in office buildings; they concluded that the DSF could not provide sufficient light for offices, and the design of DSFs had to be optimized. Nasrollahi and Salehi [6] studied DSFs to solve overheating problems between two façades using wind pressure. They applied computational fluid dynamics (CFD) to simulate various arrangements of DSFs. Joe, Choi [7] investigated and optimized the various window-glazing type and depths of cavity effects. The cavity depth varied between 10 cm and 200 cm and was dependent on numerous factors such as ventilation, shading, and maintenance. Based on the control logic system, Kim, Lee [8] applied an artificial neural network to manage the cavity airflow. However, the DSF performance was almost dependent on the environmental climate, and the same facade was not appropriate to use in different geographical areas [9].

Numerous researches investigated the effect of DSF in cooling season, and find the positive effect on cooling load by considering night cooling, natural ventilation, WWR, cavity width, different shading device, and low-e glazing [10-20]. On the contrary, there are cases where DSF was found to worsen the buildings' cooling loads [21]. Studies prove a worse situation cooling-wise, there are agreed and well-argued concerns about overheating issues [2, 18, 22-27]. Zhou and Chen [28] resulted that although the high level of the blind temperature increases the cooling loads in the interior of the building, the overheating risk can be decreased by providing good ventilation in the cavity, around the blinds. Moreover, the performance of DSF in heating season was also investigated, and some studies found the positive effect on the heating load [14, 17, 20, 29-

41]. While, Saelens, Roels [42] found that DSF to be only 5% more efficient than non-optimized Single Skin Facade (SSF) and an optimized SSF turned out to perform more efficient than DSF and, Saelens, Carmeliet [11] concluded DSF uses 20% more energy than a conventional SSF in heating season. Similarly, Papadaki, Papantoniou [43] suggests that the DSF is pejorative in terms of heating loads, with a value as significant as 28%.

However, limited researches have been yet undertaken the thermal performance of DSF by considering their performance in both cooling and heating seasons. It should be notice that the facades performed worse on cooling loads, might show positive performance when heating is considered[11, 43]. Some researched used different shading device, such as plants[44], movable integrated glass [20], Photovoltaic Blind[45], pipe-embedded[46], as well as applied the passive and active façades[47] and opening [24] to make DSF appropriate for both two season.

The key challenges in designing facades' system lie in achieving a comfortable temperature and sufficient ventilation in the cooling season while reducing the mechanical heating requirement in heating seasons [7, 24, 35, 48]. However, limited researches investigated the performance of DSF and SSFs in the entire cooling and heating seasons, and they just analyze a few days in each season as a representative of whole the season. To obtain greater accuracy total heat gain for the entire cooling and heating seasons must be used to evaluate energy consumption in the DSF and SSFs.

In this paper, the hourly heat gains of DSF and SSFs located in the south direction were calculated by the quantitative and simplified methods for four different kinds of SSFs (double glazing SSF and triple glazing SSF with internal and external shading) in entire cooling and heating seasons in four different climate regions (extremely cold, cold, hot summer and cold winter, and hot summer and warm winter regions) in China. Then, by the help of these results and calculating energy efficient ratio of different devices, the energy consumption of each type of facades was measured. Finally, it can be concluded that which kind of facades can be consumed less energy during both heating and cooling seasons in each region.

2 CASE DESCRIPTION

2.1 Studied Facades

Optimized structure of DSF for cooling and heating seasons, and four different SSF were analyzed, as shown in Figure 1. The structure of optimized DSF was selected according to the result of Su, Li [49] study, in which the depth of the cavity is 0.1m, the distance between shading and outer glass is 0.7m, height of each opening is 0.1m and the grill angle in the opening is 60°. For the DSF in the cooling season, the cavity is naturally ventilated. The shading device in the cavity and room is the roller blind, while in the outside is the fixed version. The optical properties for each facade are listed in Table 1 [50].
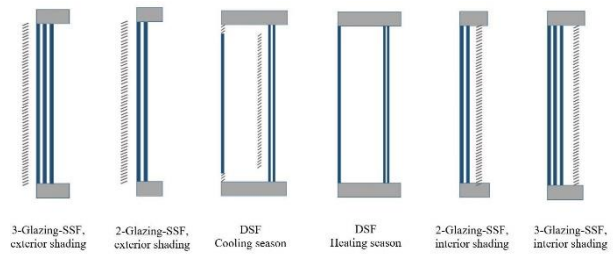
Figure 1. Schematic presentation of 5 different kinds of facades

Surface	Material	δ	τ_{sr}	ε_{sr}	h
1 glazing	glass	5	0.80	0.13	6.31
2 glazing		18	0.64	0.30	3.11
3 glazing	glass	42	0.59	0.35	1.80
Shading device	fabric	-	0.10	0.40	-
Sky	-	-	1	-	-
Other surfaces	-	-	0	-	-

Note: δ is the thickness, mm; τ_{sr} is the transmissivities of the short wave radiation; ε is the total emissivity of the short wave radiation; h is the heat transfer coefficient, $w/m^2 k$

Table 1. Thermal and optical properties setting

2.2 Studied Regions

According to the former studies related to the thermal performance of the building and weather data[28, 51] the following cities in four different climate conditions in China were chosen: Harbin with Extreme Cold, Beijing with Cold, Shanghai with Hot-summer and Cold-winter, and Guangzhou with Hot-summer and Warm-winter weather conditions (Figure 2). The hourly weather data of cities for an entire cooling and heating seasons were obtained from their meteorological database [52]. Based on the results of previous studies [51, 53, 54] and GB50736-2012 (Chinese version) the duration of cooling and heating seasons in each city can be determined (Table 2)
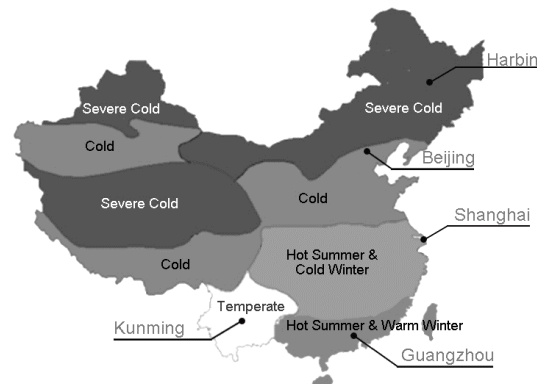
Figure 2. Climate zones in China, including the four major cities

Cities	Cooling Season	Hours per day	Heating Season	Hours per day
Harbin	1 June – 1 September	8 am – 5 pm	15 October – 15 April	24 hours
Beijing	15 May - 15 September	8 am – 5 pm	15 November – 15 March	24 hours
Shanghai	1 April – 31 October	8 am – 5 pm	16 December – 31 January	8 am – 5 pm
Guangzhou	15 April – 15 October	8 am – 5 pm	1 January – 15 February	8 am – 5 pm

Table 2. Cooling and heating seasons in each city

3 METHODOLOGY

3.1 Window Facade Heat Gain Breakdown

The total heat gains through the window facade have the similar pattern in the cooling and heating load conditions, indicating that it has the direct impact on the thermal load of the interior space of the building.

The total window facade heat gain is the sum of a variety of heat exchange mechanisms. Therefore, the break-down of this heat gain is carried out to understand the theoretical background of this phenomena. The solar penetration through the window façade is transferred into the space in three forms, which are conductive, radiant, and convective. The following equation is applied for computing the total heat gain which can be modified for different kind of facades [55].

$$q_t = q_{trans} + q_{rad.inward} + q_{conv} + q_{lr} - q_{cond.frame} \quad (1)$$

Where q_t is the total heat gain of facades, w/m^2; q_{trans} is the transmissivities heat gain, w/m^2; $q_{rad.inward}$ is the inward net infrared heat flow, w/m^2; q_{conv} is the convective heat flow between indoors and outdoors, w/m^2; q_{lr} is the long wave radiation, w/m^2; and $q_{cond.frame}$ is the conductive heat gain through the frame and other divider, w/m^2.

However, $q_{cond.frame}$ and q_{lr} only accounts a small portion of q_t and can be neglected[21, 56]. Therefore, the total heat gain can be calculated by the following equation.

$$q_t = q_{trans} + q_{radinward} + q_{conv} \quad (2)$$

3.2 Calculation of q_{trans}

Transmissivities solar heat gain was calculated by meteorological parameters of the region, as well as geometric and optical parameters of each facade. Equation 3 is related to the q_{trans} of SSF.

$$q_{trans} = J_{s,diff}\tau_t + J_{s,dir}\tau_t \quad (3)$$

where τ is the total transmittances of the facade; $J_{s.diff}$ is the south diffusion radiation intensity on the exterior surface, w/m^2; and $J_{s.dir}$ is the south direct radiation intensity, w/m^2. J_s can be obtained from the meteorological data.

In the case of DSF, the part of reflection added as it has a reflection from the bottom, as shown in Eq(4)

$$q_{trans} = (H\tau J_{s.diff} + H_{dir}\tau J_{s.dir} + H_{re}\tau J_{s.dir})/H \quad (4)$$

Where H is the total height of the façade, m; H_{dir} is the height of receiving direct solar radiation, m; and H_{re} is the height of receiving once reflected solar radiation, m. Figure 3 shows different heights and short wave radiations in the DSF.
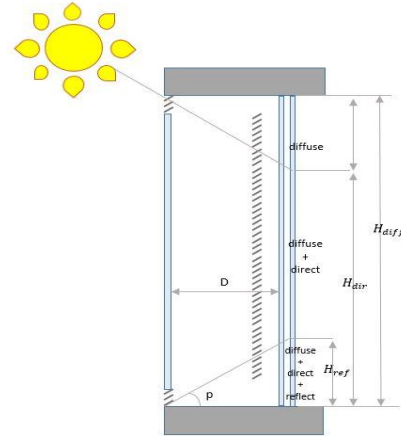
Figure 3. Different heights and short wave radiations in DSF

3.3 Calculation of $q_{rad,inward}$

The solar radiation is divided into three portions, which are transmitted (τ), reflected (ρ), and absorbed (α). The relationship of these three is:

$$\tau + \rho + \alpha = 1 \quad (5)$$

The absorbed part is divided into two parts: $q_{rad,inward}$ and $q_{rad,outward}$, which are energy transferred to the inside and outside, respectively:

$$\alpha = q_{rad,inward} + q_{rad,outward} \quad (6)$$

For SSF, $q_{rad,inward}$ can be calculated by following equations [57]:

$$q_{rad.inward} = J_s \sum \varepsilon_{inward} \quad (7)$$

$$\varepsilon_{inward} = \varepsilon \frac{h_i}{h_{ex} + h_i} \quad (8)$$

J_s is the south total radiation intensity on the exterior surface, w/m^2; h_i is the heat transfer coefficient toward the inside and h_{ex} is the heat transfer coefficient toward the outside, $w/m^2 k$. According to the ASHRAE 2013, h_i is the 8.3, and h_{ex} can be calculated by meteorological data related to mean of velocity in each region by Eq(9).

$$h_{ex}=5.62 + 3.9\ v_{mean} \tag{9}$$

With the help of Eq(7&8), the inward net infrared heat flow in DSF was measured during the heating season, by calculating J_s in the outer surface of the inner glass on the DSF which can be easily obtained by considering the reflection of the bottom and transmissivities effect of the outer glass. However, in DSF during cooling season, the simplified method validated by Xue and Li [58] was uses.

$$q_{rad,inward} = RCI . J_{og}^{2/3} \tag{10}$$

Where RCI is the Radiation Conversion Index. RCI represents the level of the average conversion through inner glass from the solar radiation to the convective heat transfer. RCI is not related to meteorological data and just related to DSF structure. As in this research the DSF structure is fixed, the RCI is the unique value obtained by the following CFD simulation.

3.3.1 CFD Simulation for DSF in Cooling Season
In order to obtain RCI of DSF in the cooling season, CFD simulation was applied. The domain consists of outdoor area and DSF model, while the ventilation inside the cavity is impacted by wind pressure. DSF model consists of outer glass with opening on bottom and top, cavity with venetian blind, inner glass, as well as walls in the top and bottom of the cavity. The DSF models are in the south direction, while it is 3 m above the ground. The outdoor area left and upper boundaries were both 10 m away and above from the outer glass, respectively.
Based on the experimental validation of [59], the porous media was used as a simplified model of the venetian blind. The grid scale of DSF was set at 0.01m and for outdoors was set at 0.05m. The computational total cells were almost 30000. Grid independence was evaluated via grid convergence index (GCI) [60], which is assumed to be less than 5% for double cell number.
Two-dimensional steady-state flow simulations were carried out in the STAR-CCM + 4.04.011. The turbulence model was adopted by the standard two layer k-e model, which can work with low-Reynolds number types (y+ _ 1) and wall function type meshes (y+ > 30) [61]. Thermal radiation was measured by the surface-to-surface model, where the medium filled between the surfaces does not scatter, emit, and absorb solar radiation. For thermal buoyance, Boussinesq assumption was applied. The second-order TVD discretization scheme was used, that were solved by using the SIMPLE algorithm[62]. According to Yaoqing [50], Table 1 gives the optical properties for each case.

3.4 Calculation of q_{conv}
Convective heat gain was obtained by [49, 58]:

$$q_{conv} = U(T_{out} - T_{room}) \tag{11}$$

$$U = \frac{1}{\frac{1}{h_i} + \sum \frac{1}{h} + \frac{1}{h_{ex}}} \tag{12}$$

Where U is the total convective heat transfer coefficient from the inner surface to exterior surface, $w/m^2 k$; T_{out} is the temperature of the outside, $°C$ (can be obtained from the meteorological data), T_{room} is the temperature of inside, $°C$; and h_g is the convective heat transfer coefficient of the glass, $w/m^2 k$.
$\sum \frac{1}{h}$ in SSF is just $\frac{1}{h_g}$, which is different according to type of the glasses (as shown in Table 1), in DSF during the cooling season, based on the Xue and Li [58]'s simplified method is just related to the inner glass; however, in the DSF during the heating season which inlet and outlet are closed, it is related to the inner glass, closed cavity of DSF, as well as outer glass. Convective heat transfer coefficient of the closed cavity can be obtained by Eq(11)[51, 63].

$$h_{cc} = Nu \frac{\lambda}{D} \tag{13}$$

$$Nu_1 = 0.0605Ra^{1/3} \tag{14}$$

$$Nu_2 = \left\{ 1 + \left[0.104Ra^{0.293}/(1 + (\frac{6310}{Ra})^{1.36}\right]^3 \right\}^{1/3} \tag{15}$$

$$Nu_3 = 0.242 \left(\frac{Ra}{\frac{H}{D}} \right)^{0.272} \tag{16}$$

$$Nu = (Nu_1, Nu_2, Nu_3)_{max} \tag{17}$$

$$Ra = \frac{\rho^2 D^3 g\beta C_p|T_{ch} - T_{cc}|}{\mu\lambda} \tag{18}$$

Where h_{cc} is the convective heat transfer coefficient of the cavity, $w/m^2 k$; Nu is the Nusselt number, λ is the air thermal conductivity, D is the depth of the cavity, m; Ra is the Rayleigh number; ρ is the density of the air, kg/m^2; g is the gravity of earth, m/s^2; β is the thermal expansion coefficient of air, k^{-1}; μ is the dynamic viscosity, Pa.s; T_{ch} is the temperature condition on the hot side, $°C$; and T_{cc} is the temperature condition on the cold side; $°C$.

3.5 Calculating the Energy Consumption
Beijing and Harbin, and Shanghai and Guangzhou follow the same way to cool and heat the indoors. In Beijing and Harbin almost officially used Chiller and Gas-boiler for cooling and heating season, respectively. While in Shanghai and Guangzhou for both seasons Air Source Heat Pump (ASHP) is applied. For measuring the amount of energy consumption in each city the energy efficient ratio for Chiller and ASHP need to be obtained by following equations [64].

$$EER_S = \frac{1}{\frac{1}{EER_r} + \frac{1}{WTF_{chw}} + \frac{1}{EER_t}} \tag{19}$$

$$EER_r = \cfrac{1}{\cfrac{1}{COP} + \cfrac{1}{WTF_{cw}} + 0.02} \qquad (20)$$

Where EER_S is the total energy efficient ratio; COP is the coefficient of performance; WTF_{chw} is the water transition factor of chilled water; EER_r is the; EER_t is the energy efficient ratio of terminal (fan coil) system; WTF_{cw} is the water transition factor of condensate water.

WTF_{chw} and WTF_{cw} for the evaluation of the standard operating condition in Beijing is 30 and 25, respectively [64]. The value of these two factors are changed 3% in each degree; hence, the amount for other cities can be obtained according to their temperature difference with Beijing.

And then in the case of Gas-boiler for Beijing and Harbin, the gas consumption should be transferred to electricity, in order to similar the kind of energy for further comparison. Hence, based on the boiler and heater pump efficiency, which are 89% and 90% [65], respectively, the energy efficiency ratio of the Gas-boiler were obtained, and with the help of exergy of the gas [66] the electricity consumption was calculated.

4   RESULTS AND DISCUSSION

Quantitatively analyzed the performance of 5 different facades in four different climates in China. Based on the meteorological data, hourly investigation was employed during entire cooling and heating seasons in Harbin, Beijing, Shanghai, and Guangzhou. The room temperatures were set at 26 and 20 ˚C, in cooling and heating seasons, respectively.

4.1 Heat Gain in Cooling Season

Harbin, Beijing, Shanghai, and Guangzhou were investigated during entire cooling season, as shown in table 2. Figure 4, shows the results of heat gain of these four cities by apply different kind of facades. The most important part that can be get by first glance at the figures is, in all the cities 2.glazing SSF, applies similar with 3.glazing SSF. The reason is, during cooling season the average temperatures of the outside are 23.9, 26.9, 24.1, and 28.9 ˚C in Harbin, Beijing, Shanghai, and Guangzhou, respectively, which do not have a huge different with the inside temperature (26 ˚C).

Therefore, convective heat transfer has a small portion in compare with others. Overall, in all four cities, SSFs with external shading plays the best with almost half of the DSF, and then, SSFs with internal shading gains the hugest amount of heat, with the big difference with other three kinds of facades, that is almost due to the radiant heat gain form the venetian blind. In Shanghai and Harbin, 2. glazing SSFs perform slightly better than 3.glazing SSF as the average temperature of outside is less than inside, 23.9 and 24.1 ˚C, respectively.

4.2 Heat Gain in Heating Season

According to Table 2, the studied facades in studied cities were investigated hourly during whole heating season. During heating season, the venetian blinds located in the internal side of SSF and cavity of DSF are closed; however,
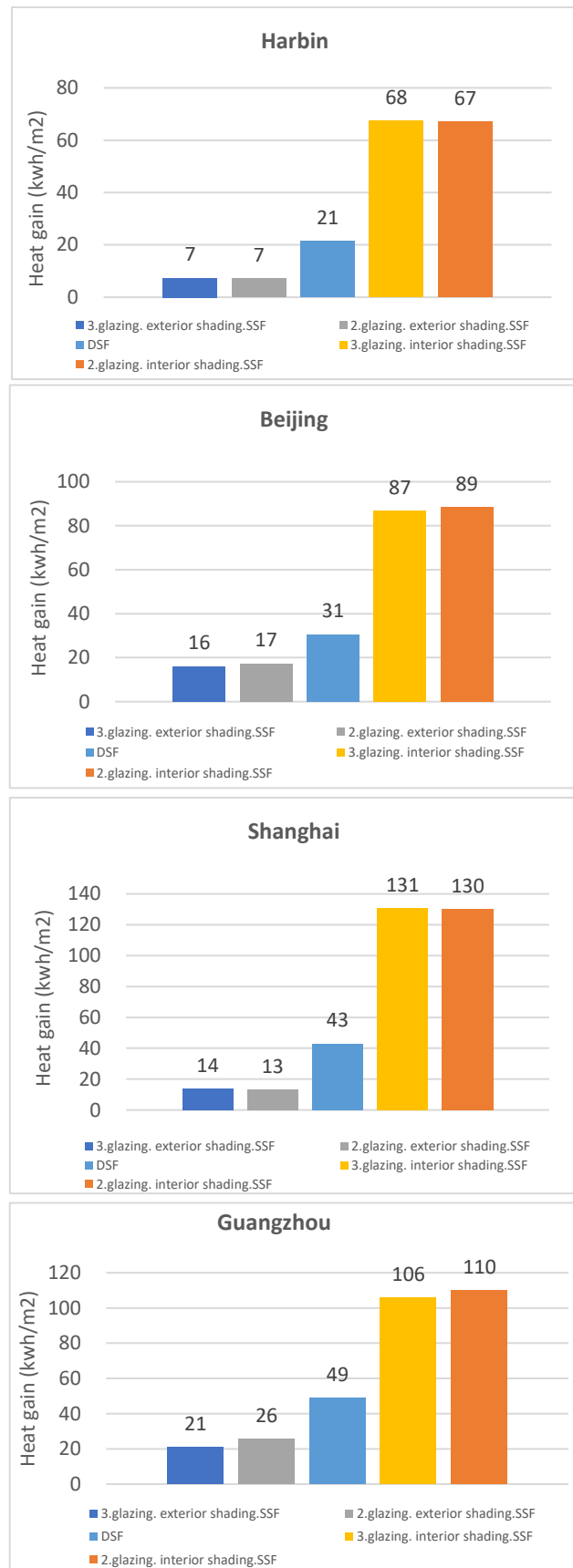
Figure 4. Heat gains in Harbin, Beijing, Shanghai, and Guangzhou during the cooling season

293
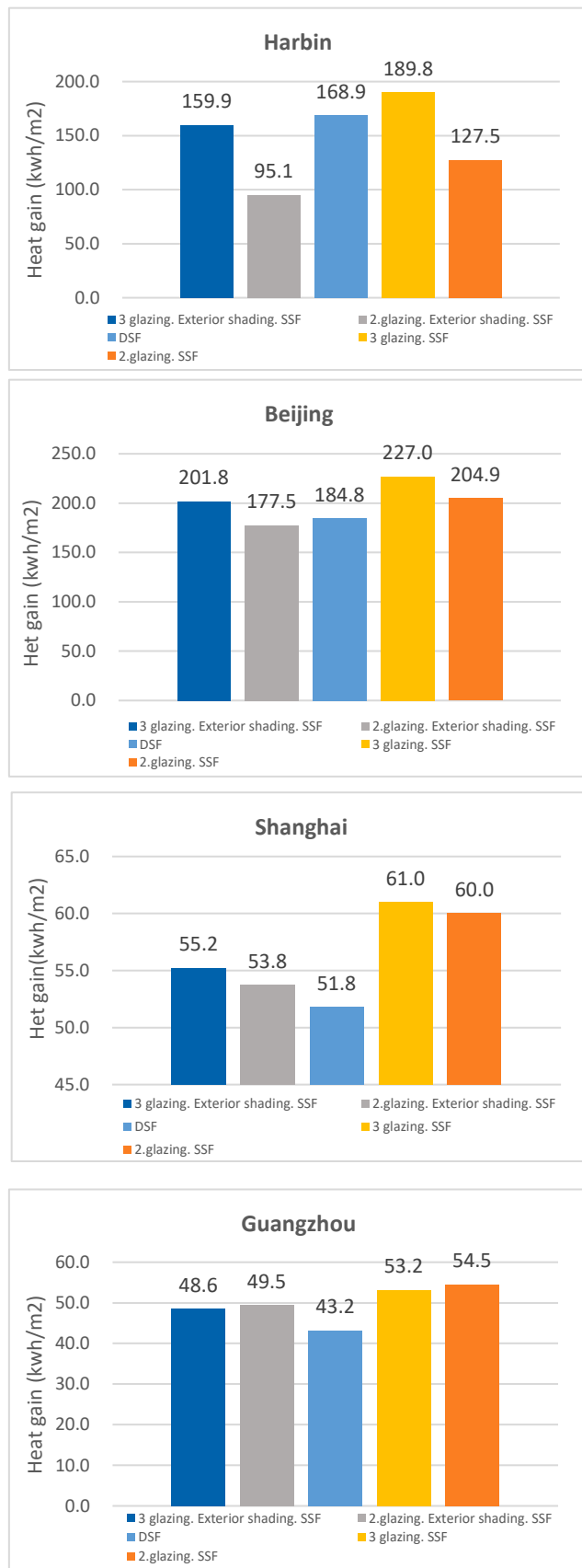
Figure 5. Heat gains in Harbin, Beijing, Shanghai, and Guangzhou during the heating season

for the external one, it is the fixed version and just the angle of blind can be adjusted to receive the largest amount of solar radiation. Figure 5 illustrates the results of heat gain of 5 facades in four studied cities. During heating season, facades which gain more heat play as the best one. The order of performance of each facade in each city are completely different, as the temperature, solar radiation intensity, duration of the heating season in cities have the significant differences. According to this, 3.glazing SSF received the hugest amount of heat in compare with others in Harbin, Shanghai, Guangzhou, with the value of 190, 227, and 61 kwh/m^2 . In Guangzhou, 2.glazing SSF shows slightly better performance from that of the 3.glazing SSF, which is owing to the high average temperature of the outside in this city ($16°C$).

4.3 Energy Consumption During Cooling and Heating Seasons

After measuring the heat gain of each facades in each city, the energy consumption can be estimated by calculated energy efficient ratio of chiller, air source heat pump, gas boiler, as well as exergy of the gas (section 3.5). The outcomes are given in Figure 6. In Beijing, Shanghai, and Guangzhou, both 3 and 2 glazing SSFs with exterior shading play as the best with less amount of energy consumption. In order to compare other three facades together in mentioned three cities, in Beijing 3.glazing interior shading SSF locates in the third place with approximately twice energy consumption of those exterior shading SSFs (90 kwh/m^2), and with the less difference with this façade, the DSF energy consumption is placed ($87kwh/m^2$). In Shanghai DSF consumes around triple less than SSFs with interior shading, and its energy consumption is approximately 0 kwh/m^2 by considering two seasons.

The order of performance of facades in Guangzhou is almost similar to Shanghai, while increasing the consumption for DSF and decreasing for 3 and 2 glazing interior shading SSF. Although, in Harbin the performance is different, 3. Glazing SSF with exterior shading still act as the best one, following with DSF, 3.glazing interior shading, 2.glazing exterior shading, and 2.glazing interior shading. The reason is in Harbin during heating season the temperature difference is so huge (average outside temperature is -9°C) which leads to the large amount of convective heat transfer, therefore, the significant difference between 2 and 3 glazing SSF can be seen.

5 CONCLUSIONS

In this paper, the thermal performance and energy consumption of DSF and four different kinds of SSFs (double glazing SSF and triple glazing SSF with internal and external shading) were investigated in four different climate regions in China, during entire cooling and heating seasons. These regions were selected according to the different weather conditions and they are representative of those climates; Harbin, extremely cold; Beijing, cold; Shanghai, hot summer and cold winter; and Guangzhou, hot summer and warm winter regions. The results are more accurate in compare with results of the typical moments or some standard days.
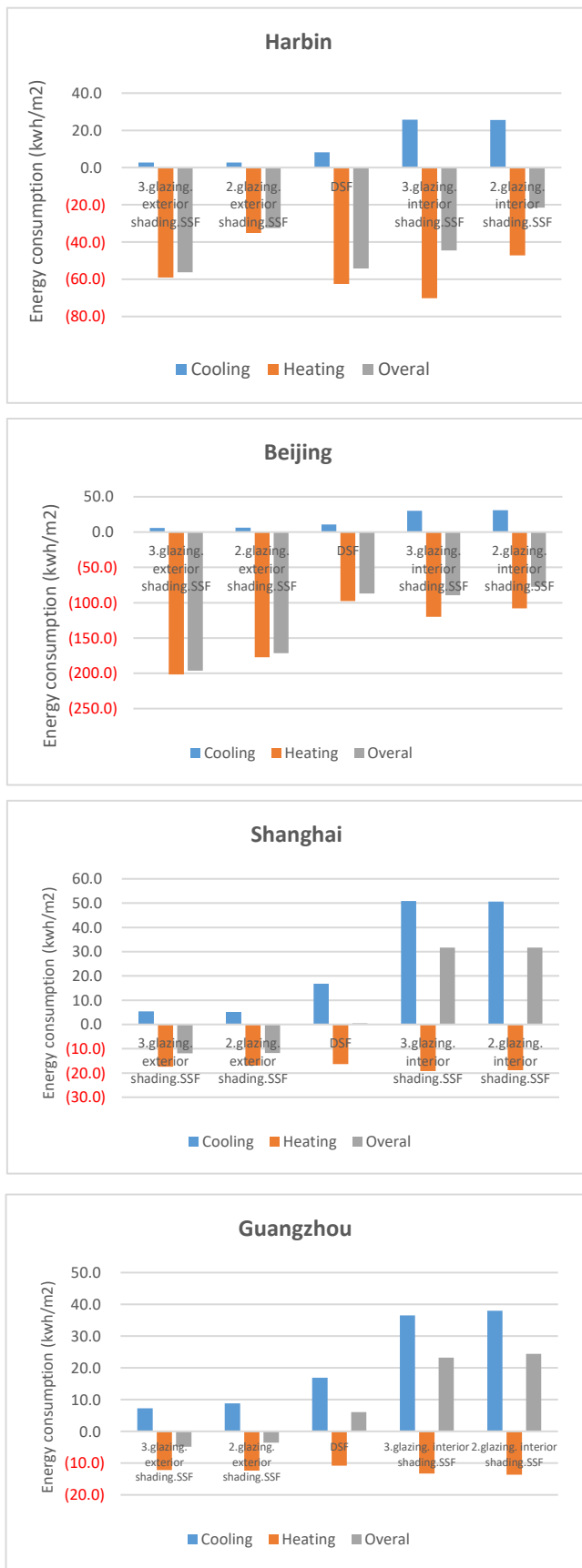
Figure 6. Energy consumption in Harbin, Beijing, Shanghai, and Guangzhou during the cooling and heating seasons

The quantitative and simplified CFD methods were applied to calculate heat gain of facades in mentioned two seasons in each region. And then, the consumption of energy was approximately measured by the help of heat gain results. As a result, 3-glazing-SSF with exterior shading can save the highest amount of energy in compared to the rest in all four regions. Although, DSF has the third level of consumption after two SSF exterior shading in Beijing, Shanghai, and Guangzhou, 3-glazing-SSF with interior shading plays slightly better than DSF in Beijing. The result of this study can be useable for different regions with same climates.

REFERENCES

1. Dewidar, K., et al. The role of intelligent façades in energy conservation. in International Conference on Sustainability and the Future: Future Intermediate Sustainable Cities (FISC 2010). 2010.
2. Barbosa, S. and K. Ip, Perspectives of double skin façades for naturally ventilated buildings: A review. Renewable and Sustainable Energy Reviews, 2014. 40: p. 1019-1029.
3. BBRI. Annual Reports. 2005 [cited 2019; Available from: https://www.bbri.be.
4. Park, C.-S., et al., Real-time optimization of a double-skin facade based on lumped modeling and occupant preference. Building and Environment, 2004. 39(8): p. 939-948.
5. Shameri, M., et al., Daylighting characterstics of existing double-skin façade office buildings. Energy and Buildings, 2013. 59: p. 279-286.
6. Nasrollahi, N. and M. Salehi, Performance enhancement of double skin facades in hot and dry climates using wind parameters. Renewable Energy, 2015. 83: p. 1-12.
7. Joe, J., et al., Optimal design of a multi-story double skin facade. Energy and Buildings, 2014. 76: p. 143-150.
8. Kim, S., J.-H. Lee, and J.W. Moon, Performance evaluation of artificial neural network-based variable control logic for double skin enveloped buildings during the heating season. Building and Environment, 2014. 82: p. 328-338.
9. de Gracia, A., et al., Energy performance of a ventilated double skin facade with PCM under different climates. Energy and Buildings, 2015. 91: p. 37-42.
10. Stec, W., A. Van Paassen, and A. Maziarz, Modelling the double skin façade with plants. Energy and Buildings, 2005. 37(5): p. 419-427.
11. Saelens, D., J. Carmeliet, and H. Hens, Energy performance assessment of multiple-skin facades. HVAC&R Research, 2003. 9(2): p. 167-185.
12. Kragh, M. Facade Engineering and Building Physics. Examples of current best practice and recent innovations. in Integrated Facade Symposium, San Francisco. 2010.
13. Kragh, M., Monitoring of advanced facades and environmental systems. whole-life performance of facades University of Bath, CWCT, 2001.
14. Kim, G., L. Schaefer, and J.T. Kim, Development of a double-skin facade for sustainable renovation of old residential buildings. Indoor and built environment, 2013. 22(1): p. 180-190.
15. Gratia, E. and A. De Herde, The most efficient position of shading devices in a double-skin facade. Energy and buildings, 2007. 39(3): p. 364-373.
16. Gratia, E. and A. De Herde, Natural cooling strategies efficiency in an office building with a double-skin façade. Energy and buildings, 2004. 36(11): p. 1139-1152.
17. Ghadamian, H., et al., Analytical solution for energy modeling of double skin façades building. Energy and Buildings, 2012. 50: p. 158-165.
18. Fallahi, A., F. Haghighat, and H. Elsadi, Energy performance assessment of double-skin façade with thermal mass. Energy and Buildings, 2010. 42(9): p. 1499-1509.
19. Eicker, U., et al., Facades and summer performance of buildings. Energy and Buildings, 2008. 40(4): p. 600-611.
20. Baldinelli, G., Double skin façades for warm climate regions: Analysis of a solution with an integrated movable shading system. Building and Environment, 2009. 44(6): p. 1107-1118.

21. Gratia, E. and A. De Herde, Are energy consumptions decreased with the addition of a double-skin? Energy and Buildings, 2007. 39(5): p. 605-619.
22. Darkwa, J., Y. Li, and D. Chow, Heat transfer and air movement behaviour in a double-skin façade. Sustainable Cities and Society, 2014. 10: p. 130-139.
23. Faggembauu, D., et al., Numerical analysis of the thermal behaviour of glazed ventilated facades in Mediterranean climates. Part II: applications and analysis of results. Solar Energy, 2003. 75(3): p. 229-239.
24. Mingotti, N., T. Chenvidyakarn, and A.W. Woods, The fluid mechanics of the natural ventilation of a narrow-cavity double-skin facade. Building and environment, 2011. 46(4): p. 807-823.
25. Reith, A., A. Gelesz, and G. Pueltz, Evaluation and optimization of a double skin facade with the help of computational simulations. Bauphysik, 2011. 33(2): p. 111-117.
26. Shameri, M., et al., Perspectives of double skin façade systems in buildings and energy saving. Renewable and Sustainable Energy Reviews, 2011. 15(3): p. 1468-1475.
27. Ghaffarianhoseini, A., et al., Exploring the advantages and challenges of double-skin façades (DSFs). Renewable and Sustainable Energy Reviews, 2016. 60: p. 1052-1065.
28. Zhou, J. and Y. Chen, A review on applying ventilated double-skin facade to buildings in hot-summer and cold-winter zone in China. Renewable and Sustainable Energy Reviews, 2010. 14(4): p. 1321-1328.
29. Kolokotroni, M., et al., Environmental impact analysis for typical office facades. Building Research & Information, 2004. 32(1): p. 2-16.
30. Andjelković, A.S., et al., DEVELOPMENT OF SIMPLE CALCULATION MODEL FOR ENERGY PERFORMANCE OF DOUBLE SKIN FAÇADES. Thermal Science, 2012. 16.
31. Choi, W., et al., Operation and control strategies for multi-storey double skin facades during the heating season. Energy and buildings, 2012. 49: p. 454-465.
32. Gratia, E. and A. De Herde, Optimal operation of a south double-skin facade. Energy and Buildings, 2004. 36(1): p. 41-60.
33. Blumenberg, J., M. Spinnler, and T. Sattelmayer. Double-skin façade systems-A comprehensive review on thermal and energetic behavior. in International Conference on Recent Advances in Heat Transfer. 2006.
34. Cakmanus, I. Optimization of double skin facades for buildings: an office building example in Ankara-Turkey. in Proceedings of Clima. 2007. Citeseer.
35. Xu, L. and T. Ojima, Field experiments on natural energy utilization in a residential house with a double skin façade system. Building and Environment, 2007. 42(5): p. 2014-2023.
36. Yılmaz, Z. and F. Çetintaş, Double skin façade's effects on heat losses of office buildings in Istanbul. Energy and Buildings, 2005. 37(7): p. 691-697.
37. Høseggen, R., B. Wachenfeldt, and S. Hanssen, Building simulation as an assisting tool in decision making: case study: with or without a double-skin façade? Energy and Buildings, 2008. 40(5): p. 821-827.
38. Cetiner, I. and E. Özkan, An approach for the evaluation of energy and cost efficiency of glass facades. Energy and Buildings, 2005. 37(6): p. 673-684.
39. Chou, S., K. Chua, and J. Ho, A study on the effects of double skin façades on the energy management in buildings. Energy Conversion and Management, 2009. 50(9): p. 2275-2281.
40. Kim, Y.-M., et al., Contribution of natural ventilation in a double skin envelope to heating load reduction in winter. Building and Environment, 2009. 44(11): p. 2236-2244.
41. Zerefos, S.C., On the performance of double skin facades in different environmental conditions. International Journal of Sustainable Energy, 2007. 26(4): p. 221-229.
42. Saelens, D., S. Roels, and H. Hens, Strategies to improve the energy performance of multiple-skin facades. Building and Environment, 2008. 43(4): p. 638-650.
43. Papadaki, N., S. Papantoniou, and D. Kolokotsa, A parametric study of the energy performance of double-skin façades in climatic conditions of Crete, Greece. International Journal of Low-Carbon Technologies, 2013. 9(4): p. 296-304.
44. Stec, W. and A. Van Paassen, Symbiosis of the double skin facade with the HVAC system. Energy and Buildings, 2005. 37(5): p. 461-469.
45. Luo, Y., et al., A comparative study on thermal performance evaluation of a new double skin façade system integrated with photovoltaic blinds. Applied energy, 2017. 199: p. 281-293.
46. Shen, C., X. Li, and S. Yan, Numerical study on energy efficiency and economy of a pipe-embedded glass envelope directly utilizing ground-source water for heating in diverse climates. Energy Conversion and Management, 2017. 150: p. 878-889.
47. Kilaire, A. and M. Stacey, Design of a prefabricated passive and active double skin façade system for UK offices. Journal of Building Engineering, 2017. 12: p. 161-170.
48. Chan, A., et al., Investigation on energy performance of double skin façade in Hong Kong. Energy and Buildings, 2009. 41(11): p. 1135-1142.
49. Su, Z., X. Li, and F. Xue, Double-skin façade optimization design for different climate zones in China. Solar Energy, 2017. 155: p. 281-290.
50. Yaoqing, L., Practical heating and air conditioning design manual, ed. n. edition. 2008: Beijing: China Construction Industry Press.
51. Wang, Y., Y. Chen, and J. Zhou, Dynamic modeling of the ventilated double skin façade in hot summer and cold winter zone in China. Building and Environment, 2016. 106: p. 365-377.
52. Song, F., et al. Meteorological data set for building thermal environment analysis of China. in Proceedings of the 10th International Building Performance Simulation Association Conference and Exhibition, Beijing, China. 2007.
53. Jiang, Q. and W. Long, Energy efficiency and design optimization of double-skin facade buildings. Heating Ventilating and Air Conditioning, 2006.
54. Wang, Q., Y. Huang, and Q. Cao, Study on saving energy of double skin glass curtain wall in the area being hot in summer and cold in winter [J]. Energy Conservation Technology, 2006. 24(1): p. 46-49.
55. ASHRAE, F., Fundamentals handbook. IP Edition, 21, 2013.
56. Yoon, Y.B., D.S. Kim, and K.H. Lee, Detailed heat balance analysis of the thermal load variations depending on the blind location and glazing type. Energy and Buildings, 2014. 75: p. 84-95.
57. Sanchez, E., et al., Influence of natural ventilation due to buoyancy and heat transfer in the energy efficiency of a double skin facade building. Energy for Sustainable Development, 2016. 33: p. 139-148.
58. Xue, F. and X. Li, A fast assessment method for thermal performance of naturally ventilated double-skin façades during cooling season. Solar Energy, 2015. 114: p. 303-313.
59. Zeng, Z., X. Li, and C. Li, Experimental study on natural ventilation performance of double-skin. Facade under solar radiation. Acta Energiae Solaris Sinica, 2012. 33(11): p. 1930-1936.
60. Roache, P.J., A method for uniform reporting of grid refinement studies. ASME-PUBLICATIONS-FED, 1993. 158: p. 109-109.
61. Yalcintas, M., An energy benchmarking model based on artificial neural network method with a case example for tropical climates. International Journal of Energy Research, 2006. 30(14): p. 1158-1174.
62. Patankar, S.V. and D.B. Spalding, A calculation procedure for heat, mass and momentum transfer in three-dimensional parabolic flows, in Numerical Prediction of Flow, Heat Transfer, Turbulence and Combustion. 1983, Elsevier. p. 54-73.
63. International Standards, I., Thermal performance of windows, doors and shading devices — Detailed calculations. 2003.
64. GB, Economic operation of air conditioning systems 2007.
65. Lam, J.C., et al., Multiple regression models for energy use in air-conditioned office buildings in different climates. Energy Conversion and Management, 2010. 51(12): p. 2692-2697.
66. Xiu, J.Y.Y., Use of equivalent electricity in energy analysis. China academic journal electronic publishing house, 2010. 32(5).

Application of Ontologically Structured Data
for Building Performance Analysis

Dawid Wolosiuk and Ardeshir Mahdavi

Department of Building Physics and
Building Ecology, Vienna University
of Technology, Karlsplatz 13, A-1040,
Vienna, Austria
{dawid.wolosiuk,
amahdavi}@tuwien.ac.at

ABSTRACT

Ontologies are developed and deployed to enhance knowledge and data exchange in a specific field or domain. For instance, building industry benefits from well-structured ontologies such as Industry Foundation Classes (IFC) [6] or green building XML [4] that drive Building Information Modeling software (BIM). In this context, efforts in building performance specification and assessment can also benefit from well-structured ontologies and data schemas. Toward this end, a recently introduced building performance data ontology attempts to identify, categorize, and capture the complexities of building related performance data and its attributes. The present contribution describes a process whereby building-related measured or simulated raw data is "ontologized" toward subsequent utilization in various building performance assessment applications.

Author Keywords

Ontology; building performance data; performance simulation

1 INTRODUCTION

To support information and knowledge sharing amongst researchers and professionals in a specific field or domain, robust and comprehensive ontologies can be of high utility. Building design, construction, and operation domain can also benefit from well-structured ontologies and data schemes.

Recently, efforts have been made to extend building-related ontologies beyond "static" data schemes for constitutive building elements, components, and system. These include, for example, the DNAS occupant-related ontology [5], an ontology for building monitoring data [9], and a building performance indicator ontology [10,11]. Development of the latter two was motivated by the need for a comprehensive ontological schema to represent dynamic building-related data for systematic building performance assessment processes. The ontology for building monitoring data identified, categorized and organized various streams of monitoring data based on former efforts in this area (e.g.,

[7,8]) and define a data schema that is intended to capture all necessary attributes concerning monitored variable.

In a subsequent effort, a building performance indicator ontology and data schema was proposed based on an extensive review of building performance indicators [1]. Thereby, performance indicators were organized in five relevant domains (energy efficiency, hygro-thermal performance, thermal comfort, air quality, visual environment, and acoustical environment). These efforts culminated in a common building performance data (BPD) ontological schema that encompasses the previously mentioned ontologies [12]. This ontological schema attempts to capture relevant characteristics of building performance variables on a descriptive level and in three property domains (spatial, temporal, and frequency). The BPD ontology has undergone initial usability/robustness testing on a large data set [11,12].

In the present contribution we specifically describe a process whereby building-related measured or simulated raw data is "ontologized" toward subsequent utilization in various building performance assessment applications. The following sections discuss in detail the process of preparing raw building performance data for usage in various downstream applications. Also provides an illustrative example where the well-structured performance data is utilized in advanced building performance assessment application.

2 FROM DATA TO APPLICATION

Ontology and ontological schema design is not a trivial task. A comprehensive approach is required to encompass domain specific features, details or elements. A well-designed ontology with a well-structured schema should facilitate the process of identifying attributes, classes, and relations of its instances. Ontology implementation is where the theoretical concepts are put into practice. Depending on a domain or a field of interest that an ontology pertains to, the implementation process might include a list of specific

challenges. Some of the relevant implementation questions are as follows:

i)   What type of data is to be semantically enriched and where does the data originate from?

ii) Should the data be preprocessed in any way (aggregation, segmentation, quality check)?

iii) How can attributes be attached to the data in an efficient manner?

iv) What is the best storage method for the ontologized data?

These questions exemplify the challenges related to the transformation of raw data to ontologized information that is ready to be processing in building performance assessment applications. As alluded to in the introduction, we focus in this paper on a structured process (see Figure 1), whereby building-related measured or simulated raw data is "ontologized" toward subsequent utilization in various building performance assessment applications. The following sections discuss in detail the different stages in this process.

2.1 Data sources

Building performance data (BPD) originates from various heterogeneous sources. One of the most common BPD sources are building control systems, metering devices, or sensors providing information on indoor and outdoor environmental conditions. Likewise, computational building simulation applications can generate relevant performance data concerning the thermal, visual or acoustical conditions in the building. Also, "virtual" sensors can deliver data that is derived using both simulated and measured values of pertinent building performance variables. Due to very different sources of data and the multitude of hardware and software components (e.g., sensors and data logging devices) and corresponding specifications involved, there is a large diversity in the ways raw data is acquired and stored. The acquisition procedure can be influenced by specific manufacturers' standards or even by specific decisions of clients.

In case of simulation sources, specific output formats are prescribed. This multiplicity of possible formats largely impacts the rest of the ontologization process as data from different sources might require preprocessing to match a common standard concerning the temporal nature of the data (i.e., event-based recordings versus regular interval data) or its structural attributes. Some of the common output formats applied to building performance data include: *i)* delimited text files containing tabular data (e.g., CSV); *ii)* database files (e.g., MySQL); *iii)* spread sheet files (e.g., XLS); and *iv)* direct simulation output (e.g., stdout).

In the process of ontological data structuring, sources might have to be handled individually. For instance, semantically
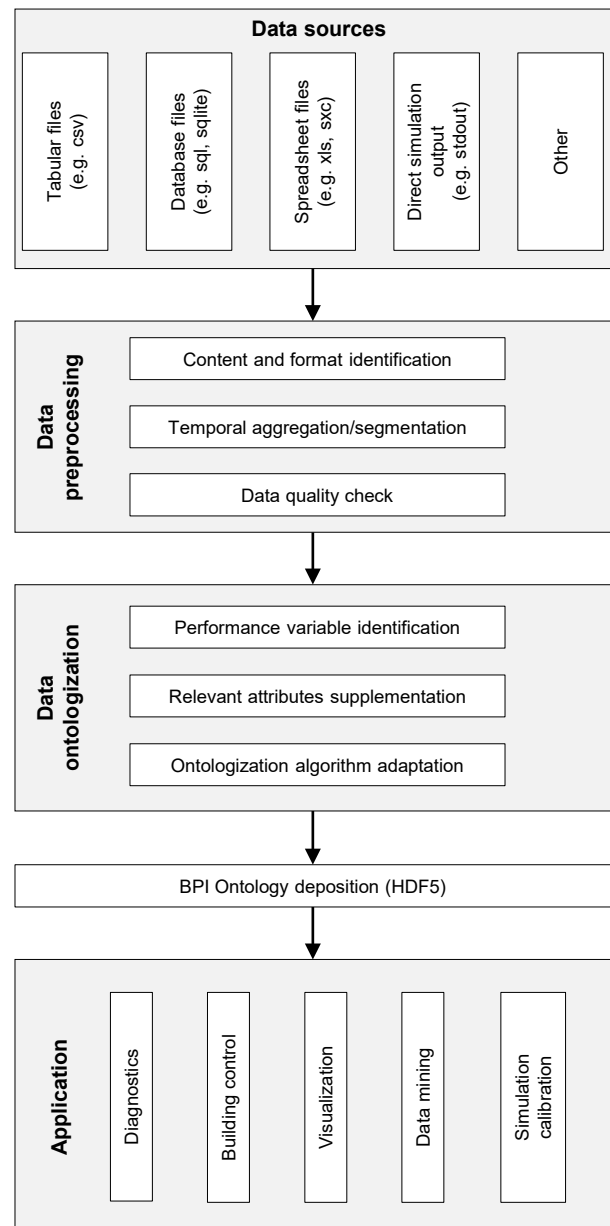
Figure 1. Schematic process overview for transformation (preprocessing, ontologization, storage) of performance-data for use in various down-stream applications (modified based on [12]).

different and heterogeneously structured data can be stored in the same output format. This issue is discussed in the following section.

2.2 Data Preprocessing

Depending on the nature of data source and/or future application, data preprocessing methods should be considered before supplementing properties and storing performance variable values in a repository. The need for spreprocessing is primarily related to the amount and quality of data. In this context, data sets (usually time series of measured performance values) should be examined in terms of data point sampling. There are two common sampling

298

strategies for data recording, namely event-based and frequency-based. In the event-based method, a measured value is only recorded when it indicates a meaningful change in the observed variable (e.g., indoor temperature change above sensor's accuracy or above fixed threshold) or there is a change in state of a device (e.g., open/close window). This type of sampling is increasingly present due to the growing number of wireless and low energy monitoring solutions and the need for an energy-efficient monitoring regime.

The frequency sampling method is based on a fixed time period between recordings of measured values. This type of sampling might seem more straightforward, nonetheless the sampling frequency can vary from device to device. For this reason, both methods face the following (application related) data usability challenge. Namely, many analysis scenarios involving data streams from multiple sources require a common temporal base (i.e., uniform intervals) for relevant data analysis operations. Instance of such operations include, for example, correlation analysis, data aggregation, or multivariate statistical investigations.

Depending on the application, different temporal resolution levels of data samples might be required. Theoretically, the temporal resolution could be kept at a very high level. However, this could lead to data redundancy and unnecessarily high computational loads especially when processing large sets of data. As such, it might be appropriate to follow common practices in a domain and/or application scenario, when deciding about the desired level of data resolution. For example, in case of typical analysis scenarios involving indoor thermal environment, 15-minute time intervals have been found sufficient. Whereas electrical power quality monitoring might require much shorter time intervals due to rapid spikes in the measured variable that would otherwise remain undetected.

The process of sensor-based data collection is prone to errors. Data quality control is a common practice that involves detection of flaws in a dataset to assure its validity and usefulness in future applications. There are a few common quality related challenges concerning data streams from monitoring devices. For instance, outliers are anomalous data points in a population of observations. Approaches to detect and mitigate outliers span from empirical methods (based on value range limits) to those rooted in descriptive statistics (e.g., Tukey's fences test for detection, or moving median for outlier removal).

Gaps in data represent another common problem. Here again, there are several strategies for filling missing data points. In case of small data gaps resulting from the removal of outliers, the commonly used method is the previously mentioned moving median or mean. In case of larger data gaps, a collection of data interpolation methods can be used. The source of such quality issues might be related to sensor malfunctions, temporal power supply interruptions, voltage drops due to a discharging battery, errors in communication with logger or gateway, and other external events.

All of the mentioned preprocessing elements need to be considered when preparing ontologized data for storage, especially when data is to be shared or archived. Hidden flaws and inconsistencies in data sets that are not detected and rectified during the preprocessing phase, might hinder correct utilization of data by future users.

2.3 Identification of categories and supplementation of attributes

The purpose of implementing an ontology-based schema is to give data a meaning and context. The ontology that is based on empirical study of a domain (or a field of interest), outlines the categories, naming convention, and collection of attributes. As with many other cases, also in the case of BPD, the properties of data collections pertaining to performance variables or indicators need to be identified individually. This could become a cumbersome task that potentially involves manual processing of multiple variables, each requiring multiple properties to be identified. Depending on the case at hand, some properties might be available for scraping from the output files of the data sources (e.g., file name, header's content, column name, etc.). Nonetheless, in such an instance, strategies for information extraction need to be developed, resulting in additional efforts. This could be beneficial, but only when processing well-structured data sets. Building performance data, particularly data coming from monitoring systems, often lacks in particular categorical or spatial details. These details need to be supplied by a system designer or individually obtained. As such, the process of assigning relevant properties to data might be prone to human error, especially given the inherent diversity of diverse building performance data.

2.4 Ontologized data deposition

Enriched building performance data must be serialized in a format that meets several requirements characteristic for ontologies in general, as well as particular requirements pertaining to a specific domain. Such a format must be able to map the structure of hierarchical categories and their relations as specified in an ontology definition. It should also allow for assigning properties (detailed in schema) to categories, subcategories, and variables. What characterizes the BPD ontology is that it typically applies to large sets of time-oriented data. The storage format must not only allow for an effective access to categorical data attributes, but also – and more importantly – support efficient queries concerning time-dependent values of relevant variables. This condition excludes some of the simple text-based serialization solutions and requires some form of data base incorporation or data-specific file format utilization.

3 PBD ONTOLOGY IMPLEMENTATION

The BPD ontology and proposed schema were tested as part of a workflow, in which data from multiple sensors, monitoring indoor and outdoor environment, were processed into ontologically structured data. The data reported by each sensor is stored in an individual data base file in a systematic manner. The naming convention for the source files, and the internal data base structure were standardized, so as to

mitigate future conversion processes. However, there are some examples of performance relevant data sets with a non-standard format or source file content that requires individual treatment. In the present implementation effort, an example of such a data source is a sky-scanner. Data from this sensor comes in a form of multiple csv files (one file per each day of recording) containing uniquely structured readings that need processing before being stored in an ontologically consistent manner.

Given the fairly well-structured output files generated by the main data source (sky-scanner) considered in this illustrative implementation, only a modest level of initial preprocessing (as described in the section 2.2) was necessary. However, not only the sky-scanner data, but also data from other sources (e.g., indoor environmental sensors) were categorized and relevant attributes in the temporal, spatial, and frequency domains were organized in the form of a csv tabular file. Several functions written in the Python programming language were created to i) transform performance data, ii) supplement it with relevant properties, and iii) serialize it in a HDF5[16] file. HDF5 file format is a flexible scientific data format designed for storage of high volume of complex data. It is capable of storing data structures in a hierarchical manner. It is also capable of storing attributes assigned to the elements of structure. Such semantically enriched data was then tested in a series of applications. An instance of an advanced application is presented in the following section.

4   ILLUSTRATIVE EXAMPLE

To illustrate how the ontologically structured data, deposited in a HDF5 file, can potentially be used in performance analysis, an interface between populated BPD ontology and Ladybug Tools [15] was created. Ladybug Tools is a collection of free and open source environmental design applications built on top of simulation engines such as Radiance [17] and EnergyPlus [2]. It is integrated with the Rhino [13] 3D modeling software, which supports it in spatial context generation and results visualization. The interfacing is enabled through Grasshopper [3] – a visual programing language and environment for Rhino.

4.1 BPD custom components

Originally designed for 3D algorithmic modeling and parametric design, Grasshopper includes a number of predefined components in various categories that serve this purpose. It also allows for creation of custom components such as Ladybug tools that can be integrated in its visual programming environment. Grasshopper supports a number of programming languages that can be used to create general purpose or task-specific components. A number of custom components written in the Python programming language were created to enable utilization of the BPD ontology with selected Ladybug elements. In general, the created components take the selected performance variable (extracted from HDF5 file based on attribute filtering) as an input, and process it (according to specified input parameters) to produce desired output in a format
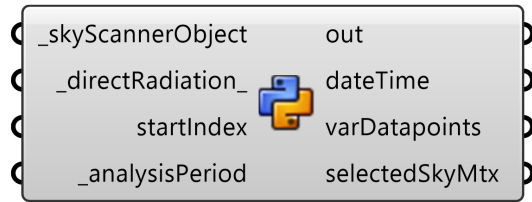
Figure 2. A custom component interface between BPD ontology and Ladybug's Radiation Analysis component.

conforming to Ladybug's requirement. Figure 2 presents one of the created custom components used for extraction and conversion of monitored data (in this case, measured solar radiation data via a sky-scanner) generate input required by Ladybug's *Radiation Analysis* element.

4.2 Environmental analysis – location specific solar radiation studies

In this particular interfacing instance, the aim was to use detailed sky radiation data collected by a sky-scanner and direct solar radiation data (derived from global and diffused solar irradiance measurements) toward solar radiation studies. Sky-scanner regularly measures, for a specific location, luminance and radiance of the sky hemisphere, represented in terms of in a 145 segments or patches. A custom component was created to extract ontologically structured solar radiation data from BPD and process it in terms of the input format of the Ladybug's *Sky Dome* element. The *cumulativeSkyMtx* output is an array of aggregated total, direct, and diffused radiation values per sky patch for a selected time period. Figure 3 presents a visualization of the annual diffuse sky radiation based on the data collected in 2016.

The same structured (*cumulativeSkyMtx*) output can be used as the input for the Ladybug's *Radiation Analysis* component. This component calculates the amount of solar energy falling on specific surfaces or on a group of selected surfaces. It calculates the incident solar radiation on a surface
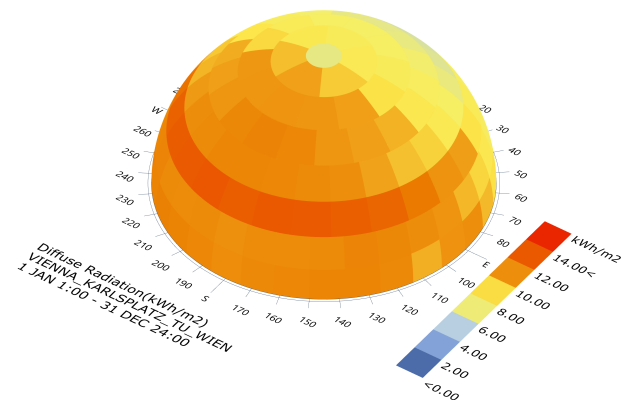
Figure 3. Visualization of diffuse solar radiation generated from sky-scanner measurements in Vienna, Austria.
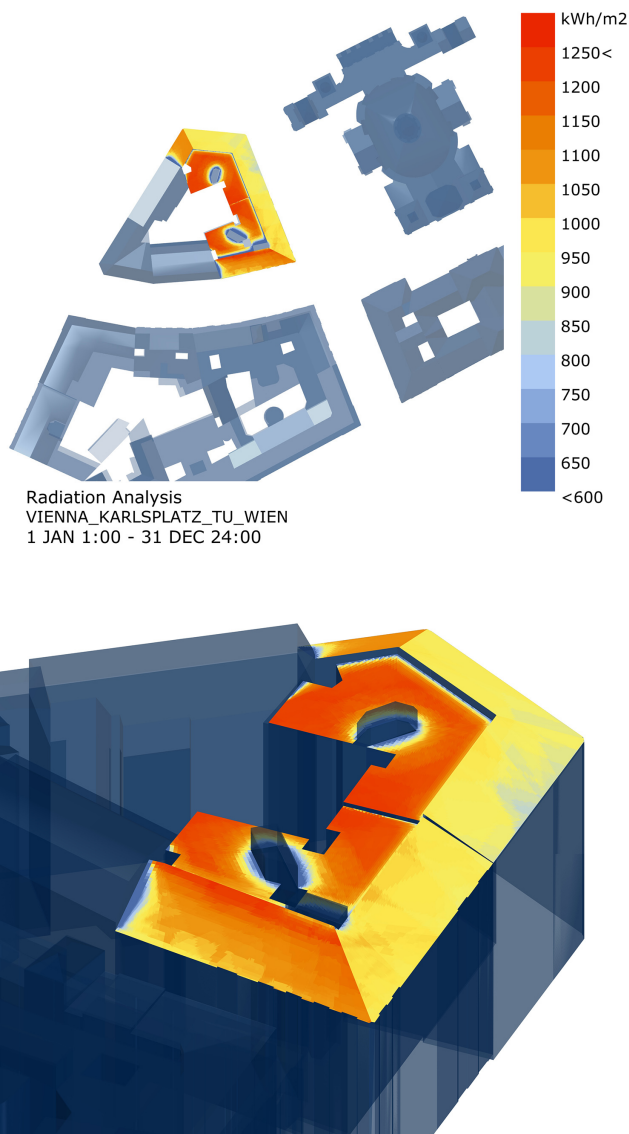
Figure 4. Illustration of a building performance assessment scenario (visualization of incident solar radiation density distribution across a complex roof configuration of an existing building) supported by ontologically stream-lined monitoring-based data.

within a selected time period. It also allows for visualizing the results directly on the 3D model. Figure 4 presents the result of such a study of an existing building in Vienna. The *Radiation Analysis* results can be used in solar heat gain studies or support solar energy systems design.

5   CONCLUSIONS

This paper discussed the process and challenges concerning the transformation of building performance data to an ontologically structured format for use in advanced building performance assessment scenarios. Specifically, we described a process whereby building-related measured or simulated raw data is "ontologized" toward subsequent utilization for analysis purposes. To illustrate this process

and provide a proof of concept, we developed an interface that connects ontologically structured performance data from a building monitoring system to a specific environmental design software. Note that the process of ontologization has the potential to streamline the input data provision not just to single instances of analysis tools, but to a wide range of building performance applications that require structurally similar data types. As a result, the proposed process has the potential to provide effective support specifically in multi-domain integrative and collaborative engineering applications.

REFERENCES

1. Constantinou, N.. A comprehensive multi-domain building performance indicator catalogue. *Master thesis*. Department of Building Physics and Building Ecology, TU Wien. Vienna (Austria), (2017).

2. Crawley DB, Lawrie L K., Winkelman F.C., et al. EnergyPlus: Creating a new-generation building energy simulation program. *Energy and Buildings, 33(4)* (2001), 319-331.

3. Grashopper visual programming language. Available from https://www.grasshopper3d.com/. As of 10 December 2019.

4. Green Building XML schema, Inc. The gbXML schema. https://www.gbxml.org/. As of 10 December 2019.

5. Hong,T., D'Oca, S., Turner, W.J.N. , Taylor-Lange, S.C. An ontology to represent energy-related occupant behavior in buildings. Part I: introduction to the DNAS framework, *Build. Environ. 92* (2015), 764-777.

6. International Organization for Standardization. Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries — Part 1: Data schema (ISO 16739-1:2018), (2018).

7. Mahdavi A. People in Building Performance Simulation. In Hensen, J.L.M., Lamberts R. *Building Performance Simulation for Design and Operation*. Taylor & Francis Group. (2011)

8. Mahdavi, A., Glawischnig, S., Schuss, M., Tahmasebi, F., and Heiderer, A. Structured Building Monitoring: Ontologies and Platform. *Proc. of 11th ECPPM Conference (2016)*.

9. Mahdavi, A., Taheri, M. An Ontology for Building Monitoring. *Journal of Building Performance Simulation*. 10(5-6) (2017), 499-508.

10. Mahdavi, A., Taheri, M. A building performance ontology. *Proc. of 12th ECPPM Conference (2018)*, 385-390.

11. Mahdavi, A., Wolosiuk D. Integration of operational data in building information modelling: From ontology to application. *CLIMA 2019 Congress. E3S Web Conf.Volume 111* (2019).

12. Mahdavi, A., Wolosiuk D. A Building Performance Indicator Ontology: Structure and Applications. *16th International IBPSA Conference (2019)*. To be published.

13. McNeel, R. and Associates. Rhinoceros Version 6. Available from https://www.rhino3d.com/. As of 10 December 2019.

14. Python Software Foundation. Python Language Reference, version 2.7. Available from http://www.python.org/. As of 10 December 2019.

15. Roudsari, M.S., Pak, M. Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. *Proc. of the 13th International IBPSA Conference* (2013).

16. The HDF Group. Hierarchical Data Format Version 5 (HDF5). Available from https://www.hdfgroup.org/. As of 10 December 2019.

17. Ward, G.J. The RADIANCE Lighting Simulation and Rendering System. *Proc. of the 21st SIGGRAPH Conference* (1994).

How People Use Their Thermostats? High-Resolution Data Analysis of Two Years of System Operations in Office Rooms

Negar Sheikh Mohammadi Khamseh[1], June Young Park [2], Enrico De Angelis[3], and Zoltan Nagy[4]

[1]Politecnico di Milano
Milan, Italy
negar.sheikh@polimi.it

[2]University of Texas
Austin, USA
juneyoungpark@utexas.edu

[3]Politecnico di Milano
Milan, Italy
enrico.deangelis@polimi.it

[4] University of Texas
Austin, USA
nagy@utexas.edu

ABSTRACT

This paper presents findings from two years' case study of occupants' thermostat use and related environmental conditions in 53 office rooms long-term data were collected by datalogger measurements of the thermal environment and behavior. It was observed that occupants infrequently interacted with their thermostats and when they did, they increased the setpoint temperature by an average of 1 °C. Overall, occupants interacted with thermostats only 10% of the time, during the monitored period. To study the correlation between the indoor, control and outdoor temperatures with the thermostat setting of occupants, regression analysis was done. The linear regression analysis showed that only indoor temperature had a meaningful correlation with thermostat temperature set by the occupants. To study if there is any correlation between the orientation of the offices and temperature demand by occupants, offices facing the same direction were clustered. The study showed a weak correlation between orientation and thermostat settings. By explaining the nature of human-building interaction, the objective of this research is to achieve a more reliable prediction of occupant behavior, in this case, thermostat setting, as a part of future low energy building design and operation practices.

Keywords
Thermostat use, occupant behavior, user preferences, building energy modeling

ACM Classification Keywords
Human-centered computing ~ User studies
Hardware ~ Temperature monitoring
Hardware ~ Temperature control

1   INTRODUCTION
Often there is a difference between computed energy performance of buildings and actual measured energy use, once buildings are operated. This difference in energy performance measurement is called the "performance gap", which depends on time and environmental conditions. The performance gap may arise from various sources, among which,

there are some factors which have uncertain nature. Occupant behavior is one of the sources of this uncertainty. In recent years many researchers direct their efforts in studying human-building interaction [1]. Occupants are not passive participants and they interact directly with building systems (i.e. thermostats, lighting and etc.) and their presence affects the heating, ventilation and lighting requirements [1].

It is found that building control systems focus primarily on energy savings rather than incorporating results from thermal comfort, especially when it comes to occupant satisfaction [2]. Due to the lack of information, user behavior is usually simplified and exaggerated in building performance simulations through one standard user pattern. To get more precise energy demand simulations, user patterns are needed that capture the wide variations in behavior without making simulations too complex [3]. Factors that affect occupant behavior, in general, include physiological factors, cultural parameters, personal traits, psychological parameters, corporate and social practice, social habits, climate, regulations, policies, and random factors. However, according to Pieter De Wilde the current challenge for energy evaluators is "Getting to the realization that from a building point of view, we are not really interested in 'behavior' – we need to refocus on occupant presence and actions only and stop pretending we can/want to model the factors that cause these."

As building envelopes and mechanical systems become more efficient, the influence of occupants on building energy increases. Meanwhile, trends in teleworking, co-working and home-sharing mean much different occupancy than the standard occupancy schedules. Lastly, expectations for comfort are growing globally, while new technologies may or may not succeed in meeting this demand. The convergence of these trends has demanded a new look at how occupants are incorporated into building design and operation during the building life cycle [4].

Due to the prevalence of technologies of heating, ventilation and air conditioning (HVAC) in buildings, individuals have been able to create a comfortable indoor thermal environ-

ment. "If a change occurs such as to produce discomfort, people react in ways which tend to restore their comfort" [5]. There are two main complications arising with this, one is how should we define discomfort and the other is that occupants don't make logical decisions all the time and their actions are generally stochastic. On the other hand, an increasing number of new buildings tend to provide thoroughly constant and uniformly neutral indoor climate at the expense of high energy consumption [6], which does not necessarily benefit occupants' thermal comfort and health [7] [8].

By measuring variations in energy consumption in identical buildings with different occupants, many studies indicate that different occupants' behaviors affect energy consumption in buildings [9]. A few previous research and case studies have shown that occupancy-based energy management system can offer up to 60% energy savings in commercial buildings [10] [11]. Although it is a huge contributor to building energy conservation, most of the current studies focus only on technical efficiency without considering human dimensions. Low energy buildings often fail to meet expected performance, with occupant behavior contributing decidedly toward building energy consumption and indoor environmental quality [12]. While occupant behavior can significantly affect the design and operation of low energy buildings, it is often under-recognized or over-simplified by energy evaluators. Analyzing the impact of occupants on building's energy consumption level is challenging, due to its complex, stochastic and multidisciplinary nature.

There are two approaches towards modelling the occupant behavior, one is a deterministic approach which assumes that certain situations will lead to certain actions, for example, the window is considered to be open if the temperature threshold is exceeded, or lighting system will be on if lux level is below a certain degree. However, the deterministic approach has some limitations. One is that fixed schedules lead to repeatable and predictable actions. Accordingly, the variations of behavior are lost. Also, it is not accurate to consider that users make perfectly rational choices, towards the control of their indoor environment. The deterministic approach does not consider the personal preferences of occupants. Since, in many cases, personal preferences or habits play an important role in decision making, a probabilistic approach is taken towards occupant behavior. For modelling of user behavior, probabilistic models could be used, which typically use statistical data to estimate the likelihood (or probability) that certain action occurs, correlating observed behavior with indoor climate-related variables such as the indoor temperature [13].

Current limitations in study of occupant behavior include: (1) a collection of suitable data for behavior understanding and modeling, (2) an ontology specific and broad enough to represent occupant behavior in buildings, (3) the evaluation of applicability of behavior models, (4) quantifying the effect of energy-related occupant behavior on building energy performance, and (5) providing metrics and insights to incorpo-

rate sustainable behaviors into robust buildings and smart communities [14]. From the point of view of some experts working on the topic, some of the challenges include: "The challenge is the lack of a standard method to collect high-quality data and of the availability of model algorithms" Yiwen Jian. Mikkel Kjaergaard underlines the need in "Respecting privacy while developing new sensing technologies for the many types of occupant behaviors" and Clair Das Bhaumik suggests "Effective communication and knowledge sharing between the different disciplines involved (e.g. social scientists and engineers)."

This paper presents findings from two years' longitudinal case study of thermal environment and occupants' thermostat use in an air-conditioned office setting in Austin, TX USA (ASHRAE climate zone 2A, hot and humid). Collection of longitudinal data allow one to observe occupant comfort and adaptive behavior as they evolve together across the day and season [9]. Offices were chosen for this research because of their substantial contribution to energy use in the United States, presenting the most common type of floor space in the commercial sector, which contributes to 19% of U.S. energy consumption [10]. The main objectives of the research include a study of (1) the frequency of interaction of occupants with thermostats and (2) the degree of effectiveness of indoor, outdoor and control temperature in predicting the probability of thermostat use.

2   DATA COLLECTION AND PROCESSING
A field monitoring of indoor and outdoor climate conditions and occupants control actions were performed in fifty-three closed offices with active chilled beam units, located in one of the buildings at University of Texas at Austin, in the period from February 2017 to November 2018. The building, built in 2013, is a LEED gold-certified building. The gross area is 19,745 [m^2]. All offices have single occupancy and are situated on the third floor of the 7-story building, along the periphery of the building, facing all four directions. Envelope details are not available. Also, the detail information, i.e., age and gender of occupants was not included in data collection.

Environmental condition data were recorded every ten minutes in each office: indoor room temperature, thermostat temperature, control temperature and outdoor temperature. As some data during the period of 10th March 2017- 30th April 2017 was missing, this period was excluded from the analysis. The collected data were recorded in the imperial unit, (Fahrenheit degrees, later converted to Celsius). Considering that the main aim of the research was to model the thermostat use of occupants, non-occupancy hours were excluded from the study and the data from working hours, from 8 am to 5 pm were analyzed. The data collected in 10 minutes intervals, was converted to hourly values and datasets were analyzed, to monitor the thermostat setting behavior of occupants in certain environmental conditions, to find if there is any correlation between the action of occupants and environmental variables.. Figure 1 shows the variation of hourly

values of all the four variables, outdoor temperature (OT), indoor room temperature (RT), thermostat temperature (TT) and control temperature (CT) in one of the rooms (room 13), from April 2017 to October 2018. TT indicates thermostat dials, which is located right next to the room doors, where light switches are placed. CT is internal control variable, which is a target supply air temperature, and it is automatically calculated temperature from other variables through a feedback loop set by the Building Automation System (BAS) provider.
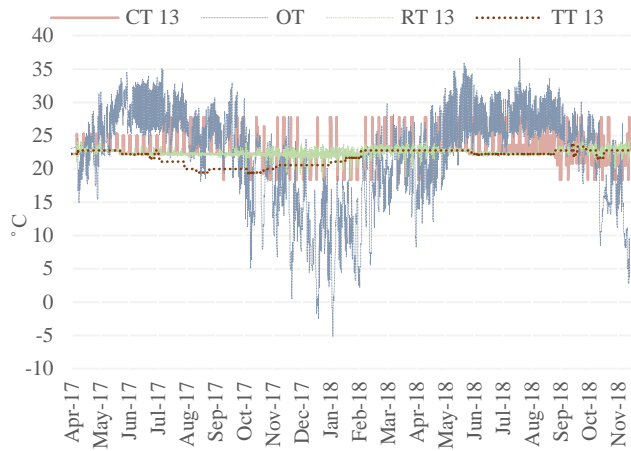
Figure 1. Temperature variation during the measured period (For interpretation of the references to colour, the reader is referred to the Web version of this article)

Figure 2 demonstrates the indoor room temperature throughout the measured period in all 53 offices. Although medians of temperature in all rooms are approximately 22 to 23 °C, variations between the offices indicate that different occupants have different preferences.
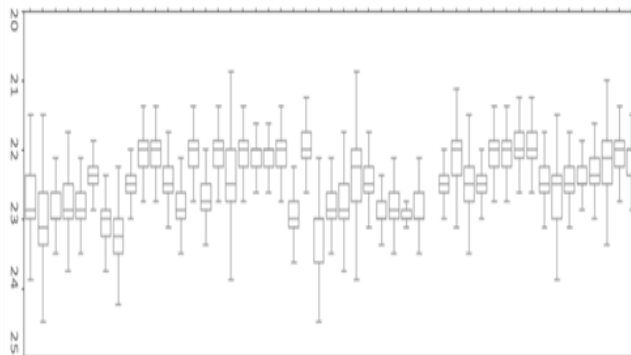
Figure 2. Indoor temperature variation in different rooms

3   ANALYSIS AND RESULTS

3.1 Regression Analysis

The first step towards the analysis of the data was to study the correlation between the indoor, control and outdoor temperatures with behavior of occupants i.e. thermostat setting in this case. Multiple linear regression analysis was performed in each room, in order to study the relationship between all the four variables and to check if thermostat tem-

perature is a dependent variable, connected to independent variables, such as outdoor temperature, control temperature and indoor room temperature. Thermostat temperature was considered as a dependent variable and the other three temperatures were considered as independent variables.

Figure 3 shows the correlation coefficient of independent variables with the thermostat temperature. The results vary from a negligible correlation to a strong correlation in different rooms. In most of the rooms that correlation is very strong, the thermostat temperature was constant throughout the monitoring period, i.e. occupants did not interact with the thermostat. In those rooms, as for any change in outdoor, control and room temperature, thermostat temperature was constant, therefore, it is easy to predict the possible next time step of the thermostat temperature and thus the correlation is strong.
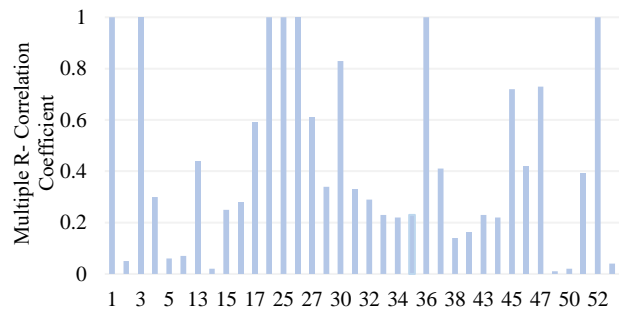
Figure 3. Multiple regression analysis of independent variables (CT, OT, RT) with the dependent variable TT (*x*-axis is Office id.)

Amongst the offices that had approximately medium correlation, room 13 was randomly selected for the future analysis. Figure 4 compares the real thermostat temperature measurements during the studied period, with the predicted values from the multiple regression analysis. As the correlation coefficient in room 13 is 0.45 and is not a very strong correlation, in some events there is up to 3 °C difference between the predicted value and the real value.
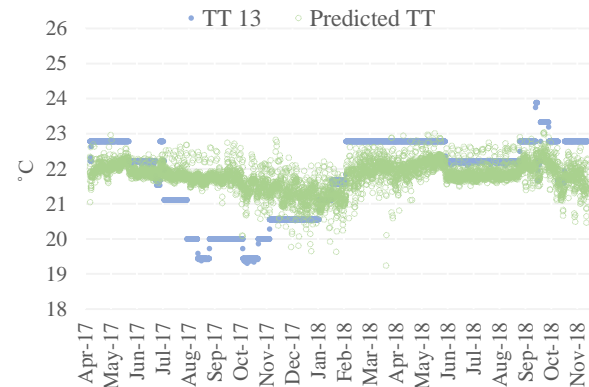
Figure 4. Comparison between real and predicted thermostat temperatures

To study the effect of each individual variable on thermostat temperature set by occupants, single linear regression analysis was performed. Figure 5 demonstrates the correlation between the indoor room temperature and the thermostat temperature. Figures 6 and 7 respectively show the correlation between control temperature and outdoor temperature with the thermostat temperature. Although the R^2 value of RT and TT is 0.22, this value is relatively high compared to other relationships discovered in Figure 6 and Figure 7.
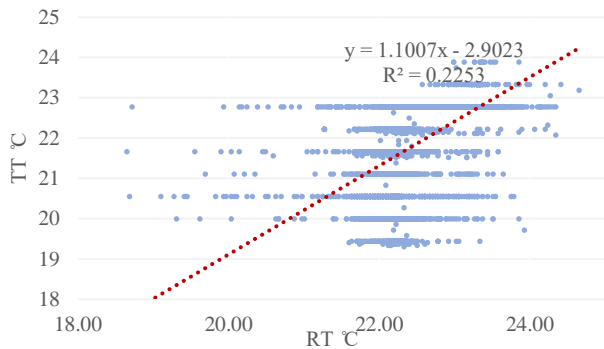
Figure 5. Linear regression analysis RT and TT
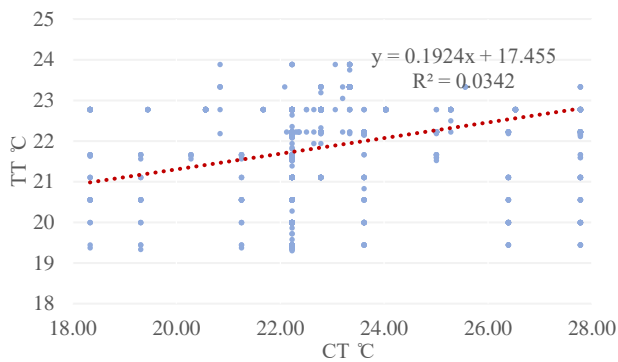
Figure 6. Linear regression analysis CT and TT
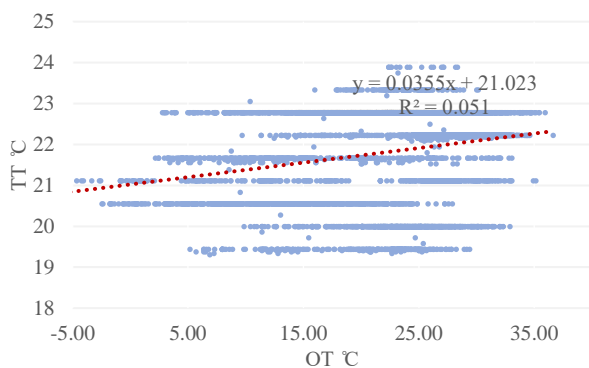
Figure 7. Linear regression analysis OT and TT

The results showed that the indoor room temperature was more correlated to the thermostat temperature as compared to the other two variables. For further study, room 30 was also randomly selected and the same analysis was carried

out. Results presented that in room 30, indoor room temperature had even greater correlation as compared to room 13, i.e. $R^2= 0.66$. In case of control temperature, the values were very close, and the outdoor temperature had $R^2= 0.02$. This indicated that between the three dependent variables, indoor temperature had a greater effect on thermostat use of occupants in comparison to other two variables.

3.2 FREQUENCY OF THERMOSTAT USE

Next step of the analysis included the study of thermostat use of occupants during the measurement period. Figure 8 shows the minimum, first quartile, median, third quartile and maximum values of thermostat setpoints in different offices. Some rooms which had irrelevant data (values outside the range of the thermostat dial) were excluded.
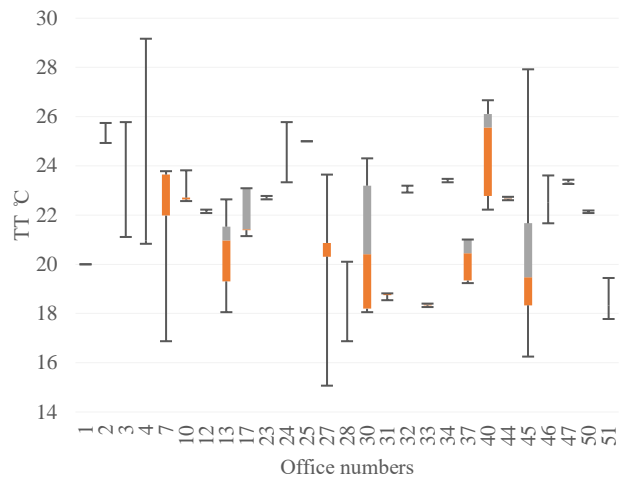
Figure 8. Statistical analysis of thermostat setpoints in the offices

Change in thermostat temperature, ΔTT for the offices in two years was calculated. It was done by calculating the difference in thermostat temperature setpoints in two consecutive time steps. Figure 9 represents this.
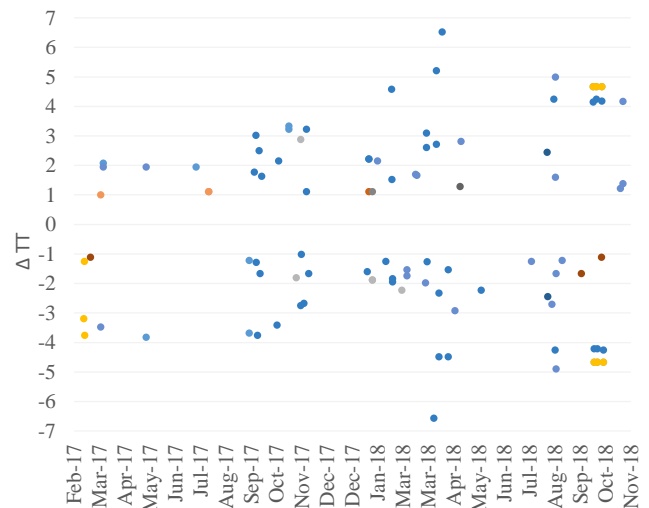
Figure 9. ΔTT in all offices (changes between 0 and ±1˚C were excluded in this graph)

Figure 10 represents the average, minimum, maximum and standard deviation of changes in thermostat dials in different offices.
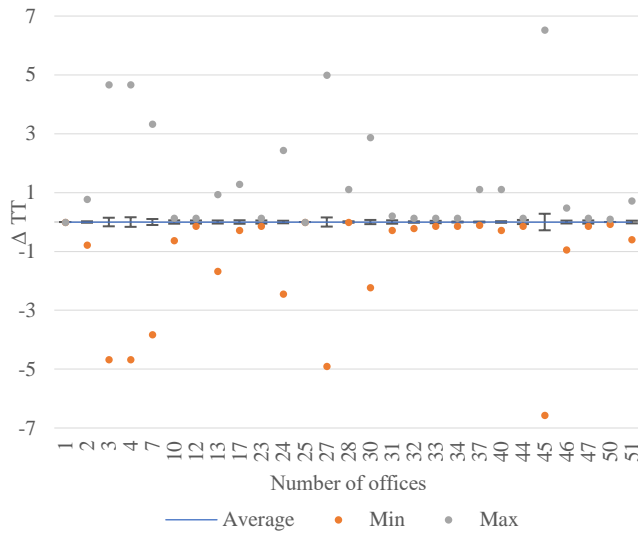
Figure 10. Statistical analysis of ΔTT in all offices

Subsequently, the frequency of each Δ TT for all offices was calculated. Figure 11 displays the percentage of Δ TT frequency distribution in all rooms during the measured period. This indicates that occupants rarely interacted with the heating/cooling systems. Approximately 10 per cent of the time they attempted to change the thermostat setting. The results indicate that 90.24% of the thermostat changes, were 0 degrees i.e. no change and remaining 9.70% of the thermostat changes were +1˚C, with few other exceptions.
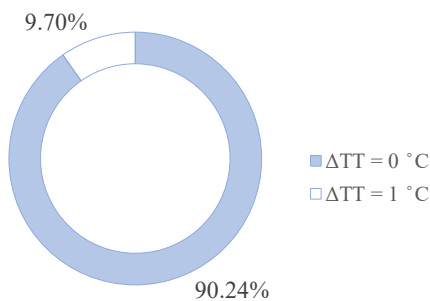
Figure 11. Percentage frequency of ΔTT in all offices

To further analyze the use of thermostats by occupants, the difference between the thermostat setpoints and control temperature was calculated for each time step. The range of differences varied from -14 degrees to +15 degrees. Figure 12 shows the percentage frequency of (TT-CT) in three rooms. Room 7 was chosen because it had the highest positive difference, i.e. occupants demanded warmer temperature than control temperature (set by BAS) and room 37 was chosen because it had the highest negative difference, i.e. occupants demanded cooler temperature. In addition, room 13 was

randomly chosen. It can be concluded that different occupants had different thermal preferences.
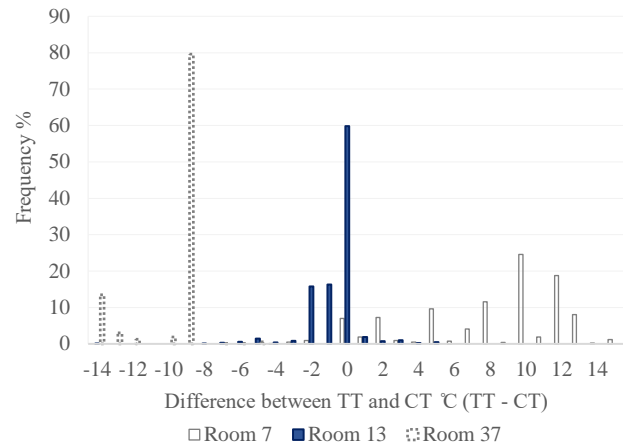
Figure 12. Deviation of occupant thermostat setpoints from BAS control temperature setpoints

3.3 Season Definitions

In order to identify the cooling and heating seasons for future analysis, heating degree days (HDD) and cooling degree days' (CDD) data were studied. Figure 13 shows the number of days in each month, that cooling, or heating is required. Neutral days are those that did not qualify as heating or cooling days. To have a more precise analysis, the sum of the total number of degree days in each month was calculated as shown in figure 14.
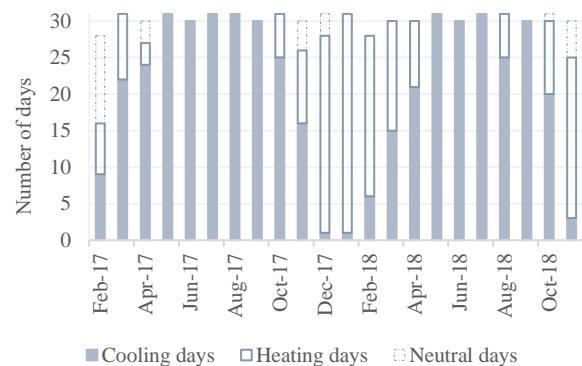
Figure 13. Heating, Cooling and Neutral days
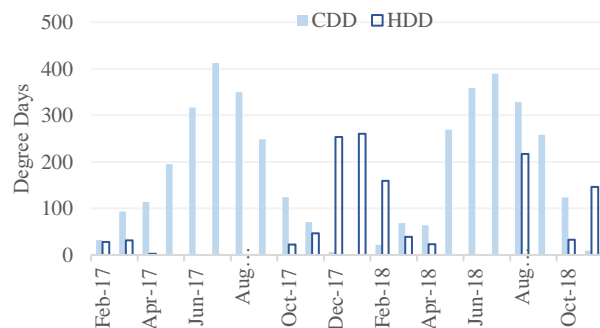
Figure 14. Total number of HDD and CDD in each month

3.4 Seasonal Variation and Thermostat Use

As seasons and temperature change thermal preferences of occupants are expected to change because of factors including changes to clothing level [15]. From the calculated HDD and CDD the heating season was considered to be from 15th October to 15th March and cooling season from 15th April to 30th September.

The average daily control temperature of each office was calculated. Daily running mean outdoor temperature was also calculated by the following approximate equation:

$\Theta rm = (\Theta ed -1 + 0,8\ \Theta ed -2 + 0,6\ \Theta ed -3 + 0,5\ \Theta ed -4 + 0,4\ \Theta ed -5 + 0,3\ \Theta ed -6 + 0,2\ \Theta ed -7)/3,8$ [16]

Where Θrm is the outdoor running mean temperature for the considered day (°C) and Θed-1 is the daily mean outdoor air temperature for the previous day.

Figure 15 shows the scatter plot of average daily indoor control temperature versus running mean outdoor temperature for all heating and cooling days in all offices.
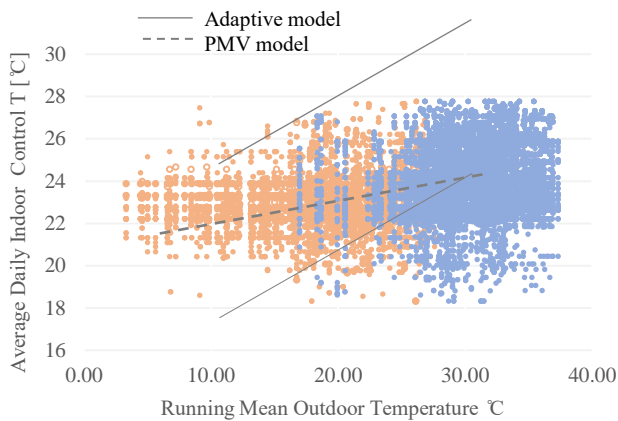
Figure 15. Heat Map (all offices), Average daily indoor control temperatures based on the running mean outdoor temperature for both heating (red) and cooling (blue) season days. The 90% comfort acceptability limits from ASHRAE-55's adaptive comfort model and PMV model (predicted lab based [15]) are included over its prescribed range of temperatures. (For interpretation of the references to colour, the reader is referred to the Web version of this article)

3.5 The Orientation of the Offices

The rooms were clustered based on the direction that they were facing, to study if there is a correlation between the orientation of the room and the temperature demand by the occupants. Average daily thermostat temperature versus prevailing outdoor temperature (calculated by taking the average of previous 30 days' outdoor temperature with equal weighting [17]) was plotted for the cluster of rooms facing each direction.

It showed that there is a weak correlation between the thermostat use of the occupants and the orientation that the office is facing. Among the rooms facing the same direction, the thermostat use behaviors were quite different. Also, the

indoor room temperature analysis of each cluster of rooms, indicated that there is not a significant similarity between the rooms facing the same direction.

As all the 53 rooms (except the four rooms in the corners) have the same amount of exposure, therefore it is safe to state that amount of exposure is not a significant parameter in this study.

4 CONCLUSIONS

In this article, two years' temperature data of 53 private offices located in one of the buildings at the University of Texas at Austin was studied. This longitudinal study of thermostat use in offices offers a better understanding of occupant's temperature preferences. The main questions that were addressed in this research included:

(1) If thermostat temperature was correlated to other variables such as indoor room temperature, outdoor temperature and control temperature. The linear regression analysis showed that indoor temperature, as compared to control and outdoor temperature, had a greater correlation with thermostat temperature set by the occupants.

(2) How frequently occupants interacted with the thermostats? The analysis showed that occupants changed the thermostat setpoints only 10% of the time during the measured period and when they did, they demanded temperature increase of approximately 1°C.

(3) If there was any correlation between the orientation of the offices and temperature setpoint by occupants? The analysis showed a weak correlation between orientation and thermostat settings.

Although the lack of interaction with thermostat could mean that occupants are thermally satisfied, it should be noted that thermal comfort is a function of air temperature, mean radiant temperature, relative humidity, and airspeed. In this study, the focus was only on the temperature. Our results show that there were not many interactions between occupants and their thermostat in this building, which could suggest that occupants may be generally satisfied with their environment. On the other hand, it is unclear whether the lack of interaction stems from generally low desire to adjust settings, which in turn could suggest the potential for energy savings, in particular if rooms are cooler/warmer than the occupants prefer. To address this potential it is important to consider human-building interaction knowledge with respect to occupant comfort and productivity and adopt an occupant-centric approach for building control systems [18].

REFERENCES

[1] T. Hong, D. Da Yan, S. D'Oca, C.-F. C.-F. Chen, S. D 'oca, and C.-F. C.-F. Chen, "Ten questions concerning occupant behavior in buildings: The big picture," *Build. Environ.*, vol. 114, pp. 518–530, Mar. 2017.

[2] J. Y. Park and Z. Nagy, "Comprehensive analysis of the relationship between thermal comfort and building control research - A data-driven literature review," *Renew. Sustain. Energy Rev.*, vol. 82, no. July 2017, pp. 2664–2679, 2018.

[3] D. Aerts, J. Minnen, I. Glorieux, I. Wouters, and F. Descamps, "A method for the identification and modelling of realistic domestic occupancy sequences for building energy demand simulations and peer comparison," *Build. Environ.*, 2014.

[4] A. Wagner, "Occupant behaviour-centric building design and operation EBC Annex 79," *Iea Ebc*, no. October, 2018.

[5] ASHRAE, "Thermal Environmental Conditions for Human Occupancy ASHRAE INDUSTRIAL ADVERTISING POLICY ON STANDARDS ASHRAE Standard Project Committee 55 Cognizant TC: TC 2.1, Physiology and Human Environment SPLS Liaison," vol. 2004, 2003.

[6] T. Hoyt, K. H. Lee, H. Zhang, E. Arens, and T. Webster, "Indoor Environmental Quality (IEQ) Title Energy savings from extended air temperature setpoints and reductions in room air mixing Publication Date Energy Savings from Extended Air Temperature Setpoints and Reductions in Room Air Mixing," *Int. Conf. Environ. Ergon.*, 2009.

[7] E. Arens, M. A. Humphreys, R. de Dear, and H. Zhang, "Are 'class A' temperature requirements realistic or desirable?," *Build. Environ.*, vol. 45, no. 1, pp. 4–10, 2010.

[8] M. Luo, B. Cao, W. Ji, Q. Ouyang, B. Lin, and Y. Zhu, "The underlying linkage between personal control and thermal comfort: Psychological or physical effects?," *Energy Build.*, vol. 111, pp. 56–63, 2016.

[9] T. Buso, V. Fabi, R. K. Andersen, and S. P. Corgnati, "Occupant behaviour and robustness of building design," *Build. Environ.*, 2015.

[10] M. Mysen, S. Berntsen, P. Nafstad, and P. G. Schild, "Occupancy density and benefits of demand-controlled ventilation in Norwegian primary schools," *Energy Build.*, vol. 37, no. 12, pp. 1234–1240, 2005.

[11] X. Lin and J. Lau, "Demand controlled ventilation for multiple zone HVAC systems: CO_2-based dynamic reset (RP 1547)," *HVAC&R Res.*, vol. 20, no. 8, pp. 875–888, Nov. 2014.

[12] K. Schakib-Ekbatan, F. Z. Çakıcı, M. Schweiker, and A. Wagner, "Does the occupant behavior match the energy concept of the building? – Analysis of a German naturally ventilated office building," *Build. Environ.*, vol. 84, pp. 142–150, Jan. 2015.

[13] I. Gaetani, P.-J. P.-J. J. Hoes, and J. L. M. J. L. M. M. J. L. M. Hensen, "Occupant behavior in building energy simulation: Towards a fit-for-purpose modeling strategy," *Energy Build.*, vol. 121, pp. 188–204, 2016.

[14] T. Hong, S. C. S. C. Taylor-Lange, S. D'Oca, D. Da Yan, and S. P. S. P. Corgnati, "Advances in research and applications of energy-related occupant behavior in buildings," *Energy Build.*, vol. 116, pp. 694–702, 2016.

[15] R. J. De Dear and G. S. Brager, "Thermal comfort in naturally ventilated buildings: Revisions to ASHRAE Standard 55," *Energy Build.*, vol. 34, no. 6, pp. 549–561, 2002.

[16] BSI, "Energy performance of buildings-Indoor environmental quality BSI Standards Publication," 2017.

[17] B. Huchuk, W. O. Brien, and S. Sanner, "A longitudinal study of thermostat behaviors based on climate , seasonal , and energy price considerations using connected thermostat data," *Build. Environ.*, vol. 139, no. February, pp. 199–210, 2018.

[18] J. Y. Park *et al.*, "A critical review of field implementations of occupant-centric building controls," *Build. Environ.*, vol. 165, no. August, p. 106351, 2019.

Simulating Invisible Light: Adapting Lighting and Geometry Models for Radiant Heat Transfer

Dorit Aviv[1,2], Miaomiao Hou[2,3], Eric Teitelbaum[1], Hongshan Guo[1], Forrest Meggers[1]

[1]Princeton University	[2]University of Pennsylvania	[3]Tongji University
Princeton, NJ, USA	Philadelphia, PA, USA	Shanghai, China

ABSTRACT

Thermal radiation, being the infrared spectrum of electromagnetic radiation, shares many characteristics with visible light, and thus is highly dependent on surface geometry. Much research effort has been dedicated to characterizing the behavior of visible light in the built environment and its impact on the human experience of space. However, longwave infrared radiation's effect on the human perception of heat within the indoor environment is still not well characterized or understood within the design community. In order to make legible the embodied effect of radiant surfaces' geometry and configuration, we have developed a Mean Radiant Temperature simulation tool which is based on a raytracing technique and accounts for the detailed geometry of the human body and its surrounding environment. This paper is meant to provide an overview of the geometric characteristics of radiant heat transfer with a dual purpose: 1. the integration of these principles into a Mean Radiant Temperature simulation technique in order to better characterize radiant energy exchanges and 2. the development of architectural design strategies based on these principles, which are tested in a case-study prototype. The MRT simulation method and results for the experiment are discussed.

Author Keywords

Longwave Radiation; View Factor; Raytracing; Mean Radiant Temperature.

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING

1 MOTIVATION AND BACKGROUND

Radiant heat transfer may account for more than half of the thermal comfort for occupants in the built environment [14]. Investigation of radiant systems into building are critical for energy savings, since nearly 20% of the total energy demand in the United States is used for heating and cooling the buildings in 2018 [7]. With heightened interest in radiant heating and cooling systems due to their increased efficiency over air systems [11], it has become imperative to develop a simulation tool that makes the consequences of radiation heat exchange legible to designers.

The main method to characterize the radiant heat exchange between a body and its surroundings its representation within the comfort heat balance as Mean Radiant Temperature (MRT). MRT has been taken into account in several thermal comfort model indexes such as predicted mean vote (PMV) and physiological equivalent temperature (PET). Nevertheless, there is a fundamental ambiguity regarding the concept because of the degree of geometric complexity which it represents [9]. The definition of mean radiant temperature in the 2017 version of ASHRAE-55 standard is "uniform temperature of an imaginary enclosure in which the radiant heat transfer from the human body equals the radiant heat transfer in the actual non-uniform enclosure" [1]. The calculation of MRT thus involves the proportion of radiation transferred between every portion of the surface on the human body and all surrounding surfaces, weighted by the view factor between them. Because of its complexity, abstraction and simplification methods for MRT representation have been employed.

While all other environmental comfort variables (air temperature, relative humidity, and air velocity) are dependent on the medium of air, MRT is independent of the air and is in fact a complex representation of the exchange between the body and the surrounding surface geometry and its variable material characteristics and energetic states. A temperature abstraction is necessary to enable direct comparative engagement by researchers and practitioners with the effects of radiant heat transfer in space. Still, the result of using MRT as the only metric for radiation impact on comfort, has led to a generalization and lack of appreciation of the importance of geometry in how it is spatially resolved. Therefore, it is critical to reconsider the fundamental physics of radiant heat transfer for what they are: infrared radiation is part of the same electromagnetic spectrum as the light we see, yet at longer wavelengths.

In this paper, we are proposing an improved method to simulate MRT by building upon existing literature. To highlight its advancements beyond the conventional methods, we reviewed the state-of-the-art approaches qualitatively, and provide a case study where we demonstrate improved MRT accuracy by implementing our method to a recent project in Singapore. Additionally, contemporary 3D-

modeling tools allow for a rigorous representation of the body's surface geometry, as opposed to its abstraction into a platonic shape or a point in space.

2 SIMULATION TECHNIQUE

In the following section, existing methods of geometric representation of light and the human body are described and adapted towards the radiant heat transfer simulation.

The definition of MRT co-evolved along with the mathematical model for view factor. Since 1966 when mean radiant temperature was first defined in a standard by ASHRAE, the continual wording changes from "solid body" which abstracts the human body as a solid sphere to "human body" as well as from "environment" to "enclosure" show the increasing consciousness towards the inherent geometric complexities of both human body and the surrounding environment [2]. At an early stage of its evolution, the definition of MRT was equivalent to mean spherical irradiance, using a small black sphere at the test point to replace human body. The mean radiant temperature T_r at a point P can be derived from the Equation (1), where view factors $F_{p \to i}$ between the point P and all the surroundings are used to weigh the surface temperatures T_i:

$$T_r^4 = \sum_{i=1}^{n} T_i^4 F_{p \to i} \tag{1}$$

Equation (1) is very commonly considered the equivalent mathematical representation of MRT and widely used among textbooks, standards and design guidelines [1] although the underlying assumption is that the human body is viewed as an infinitesimal point.

The improvement proposed here over the series of simplifications historically applied to MRT, is based on state-of-the-art methods used in existing literature and can be considered within three categories: the determination of view factors, adopting ray-tracing techniques from visible light simulation, and using more detailed body geometry. We will first review the most relevant recent advancements within each respective sector before presenting our method in the subsequent sections.

2.1 Determination of View Factors

Rather than simplifying the human body to a single point, there are several successive alternative methods for calculating view factors in a more precise way. One is the Fanger-Rizzo method, which mainly analyzes the geometric relationship in a Cartesian coordinate system. Fanger provided the graph of view factors based on tabulated between a standing or seated human body and a surrounding rectangular plane based on experimental data, which was further developed by Rizzo into algorithms that automatically compute view factors between seated or standing occupant and the orthogonal surfaces of the Cartesian coordinates. Rizzo further developed algorithms to automatically compute the view factor based on Fanger's data [3].

Nusselt Analog is another commonly used method in determining view factors between a point and the surrounding non-orthogonal surfaces. It was first developed by Nusselt to achieve the figure-factor integral. The Nusselt analog utilizes a unit-radius hemisphere centered at the point of interest on which the target surface is projected. Then if the projected surface on the hemisphere is projected onto the plane at the bottom, the measurement of solid angle which is needed for view factor is simplified to the calculation of areas on the plane [9]. As is shown in Equation (2), the view factor $F_{p \to s}$ from the point P to the surface S is a function of the projected area S_2 on the plane at the bottom of the hemisphere.

$$F_{p \to s} = S_2 / \pi \tag{2}$$

Based on the Nusselt Analog method, a method named "Numerous Vectors (NV)" has been proposed in recent years [10]. Rather than projecting surfaces by the Nusselt method, the NV method utilizes the vectors emanating from the center point directly. The next step followed is to distribute points on the unit sphere spatially evenly. Every point on the surface of the sphere is connected to the centered point forming vectors of every direction in a certain resolution. The radiation flux represented by the vectors is thus traced backwards to hit the surrounding surfaces. The view factor can be calculated by Equation (3), where $F_{p \to s}$ is the view factor of the desired point P to a surface, N_1 is the count of the rays hit the surface and N is the total number of rays.

$$F_{p \to s} = N_1 / N \tag{2}$$

The use of vectors allows for a sufficiently accurate and convenient calculation in computer simulation, as the he NV method makes the vector-based representation of view factors feasible and this is the reason for its widely combination with simulation tools. However, the use of this method in the past has been limited to the outdoor MRT and the human body was considered as a small black sphere [16]. Related to this method, the ray-tracing technique can be further used to trace the continual radiant interreflections between multiple surfaces.

The NV method is adapted in the simulation technique proposed in this paper (see Fig. 1).
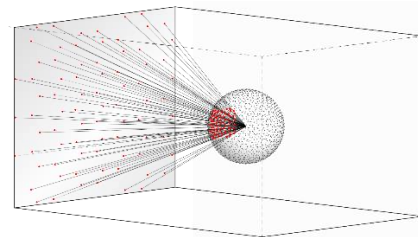
Figure 1: Numerous Vectors method for calculating the view factor.

2.2 Raytracing Technique

The comparison study between visible light and infrared radiation is not rare. The research on the flow of light used

the concepts of scalar and vector illumination since 1960s [13]. Many simulation tools in the computer graphics field have been specifically developed to track visible light in the through raytracing methods. Those simulations include specular, diffuse and directional-diffuse reflection and transmission with specialized operators for each of them. The lighting simulation and rendering system named RADIANCE is the most widely used light rendering systems in the past two decades [30].

The shortwave radiation includes direct, diffuse and reflected light fluxes. When the raytracing technique was first used in the two-pass method, the diffuse part of interreflections was still calculated according to the radiosity method, introduced by Cohen and Greenberg [5], using a simplified coefficient for diffuse contributions. Along with the development of computing capability, diffuse emissions have been incorporated into Monte Carlo simulations of heat transfer and renderings [30],[32].

The RADIANCE rendering equation, which is conducted iteratively for multiple bounces, provides pixels brightness representation of test points based on energy passing through them [30]. It characterizes energy passing through a point in a specific direction, where outgoing radiance from a point on a surface is calculated in terms of incoming radiance over the projected view hemisphere of that point. This is based on the principle of conservation of energy:

$$\varepsilon + \rho + \tau = 1 \tag{3}$$

where ε is the is the portion of radiation absorbed/emitted, ρ is the portion of radiation reflected and τ is portion of radiation transmitted.

For MRT simulation, a test point on the human body can be calculated as the origin for raytracing. If we consider raytracing of four bounces (Fig. 2), the perceived radiant temperature T_r for a point 0 on the body would be:

$$T_{r0}{}^4 = \varepsilon_1 \, T_1{}^4 + \rho_1 \, (\varepsilon_2 \, T_2{}^4 + \rho_2 \, (\varepsilon_{s3} \, T_3{}^4 + \rho_3 T_4{}^4)) \tag{4}$$

where T_i is the surface temperature at point i, ε_i is the surface emissivity fraction and ρ_i is the surface reflectivity fraction. Note that τ was dropped since the energy transmitted through the point does not affect the point's radiant temperature. However, in this equation, for non-opaque surfaces, $\varepsilon + \rho$ does not equal 1.
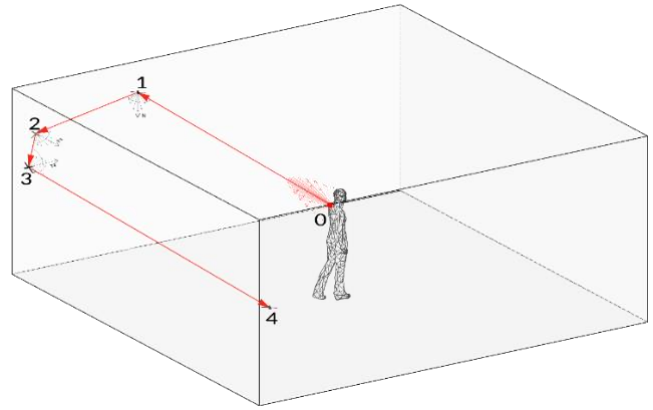
Figure 2: 4-bounce raytracing from a point on the body.

Equation 4 only accounts for specular reflections. For diffuse interreflections, a caching system as used in RADIANCE [30] must be used or the simulation process would require exponentially more computational power.

While reflections between surfaces can be either diffuse or specular, all emissions from a terrestrial surface are diffuse and their intensity is distributed according to the Lambertian cosine law:

$$I = I_0 \cos \emptyset \tag{5}$$

where I is the radiant intensity (w/m^2 sr) I_0 is the radiance normal, i.e., perpendicular, to the emitting surface, and $\emptyset$ is the angle between the emission direction and the normal to the emitting surface. The maximum intensity is in the direction of the normal to the surface and decreases in proportion to the cosine of the angle between the normal to the surface and the emission direction [31]. We use this equation to calculate the intensity of emission for each vector representing a direction of diffuse heat flux from a body to an environment. Fig. 3 visualizes the Lambertian Cosine law in 3D space, as used in our MRT simulation, treating emitted heat fluxes as vector emanating from a sample point on a Lambertian surface. In the context of our raytracing method, we may think of MRT as the mean of all the surface temperatures intersected by the heat flux vectors emitted from all the sample points on the human body.

Figure 3. Visualization of Lambertian emission from a sample point on a discretized human body

	1. Determine the purpose	2. Subdivide surfaces & Space points	3. Distribute rays over the hemisphere, with lambartian distribution	4. Recursive sample	5. Get the outcome of each point	6. Result
Radiance (Visible Light)	2D image: Choose the viewpoint for rendering	Subdivide the surfaces in the scene into patches; test point: one point at each patch	Start from each test point on the surface patches	For multiple diffuse interrefletions, account for both specular and lambartian reflections	Illuminance (lux / cd·sr·m²)	Rendered image
Proposed MRT Simulation	Temperature Value	Body geometry as a mesh with triangulated faces; test point: centroid of each triangle	Start from each test point on the body surface with diffuse lambartian emissions	For multiple interrefletions, account for specular and reflections in the longwave and compute both specular and diffuse reflection if shortwave radiation sources are present	Temperature (°K) per test point, normalized by the subdivided area portion	MRT (°C) for the full body

Figure 4: Comparison of RADIANCE rendering method and proposed MRT simulation method.

We thus have as our MRT equation for N body segments:

$$T_{MRT} = \sqrt[4]{\int_{j=0}^{N_a} dA_j \; \frac{\sum_{i=1}^{N_v} cos\theta_i T_{ri}^4}{N_v/2}} \qquad (6)$$

where N_a is the number of subdivisions of the body surface; dA_j is the fractional area of the body surface with centroid point j out of the full body area; N_v is the number of vectors emitted from point j, θ is angle between the emitted vector and the normal to the emitting surface; T_r is the perceived radiant temperature at point j after multiple bounces, calculated according to Equation 4.

Fig. 4 shows a sequential comparison of the method utilized in RADIANCE as compared to the algorithm used in our own method for calculating MRT.

2.3   Body Geometry

Accounting for mean radiant temperature precisely is largely limited to the body geometry complexity. For thermal comfort calculation, body surfaces of different body parts are related to specific condition of metabolic heat production, sweat evaporation, cloth covering. Furthermore, the body geometry model determines also the body position and movement. All these factors are given major focus in the rising research of localized thermal comfort [35]. Several different methods for simplifying the body geometry model are listed and compared in Table 1.

As described in section 2.1, the simplification method commonly used for body geometry is to assume the human body as a single point [4]. The computational workload is largely alleviated when the view factors between a point and surrounding surfaces needs to be calculated as opposed to complex body geometry.

Apart from view factor calculation, another parameter directly related to the human body geometry, is called projected area factor. The projected area factor is the ratio of the projected area of the human body on the surrounding surface to the "effective radiation area," which is defined as the surface area of the human body directly involved in radiation transfer. These two values are required for the computation of the MRT. In the case of abstracting the human body into a point in a sphere, the projected area factor can be derived from a solid angle method.
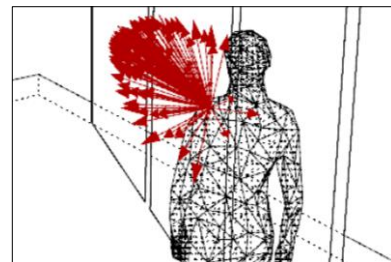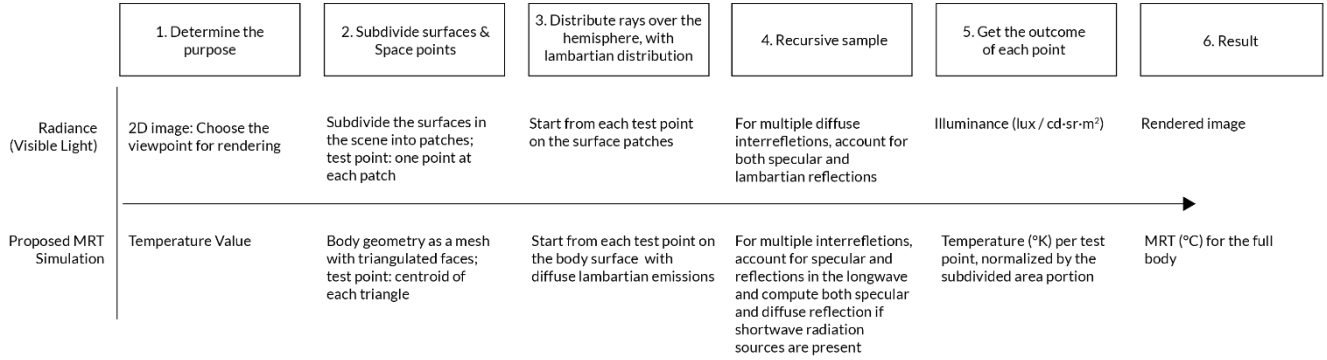
In order to obtain the projected area factor of the human body when projected on its surroundings, Fanger utilized the parallel ray method with the aid of a photographic method. Silhouettes were used in many studies for the comparison of body shape [24]. The body posture (standing or seated), the body shape (for each gender separately) and the clothing condition were taken into account during the experiment. Since the average body shape given as samples varies, different results can be derived in different experiments [28]. Other studies utilized physical models with simple outlines representing the standing and seated human body for experiments [17].

Another prevailing method is to simplify the human body as a single solid figure. This method is also applicable for the derivation of view factors by numerical integration of projected area factor. A seated person is simplified into a sphere, a box or a cube, while a standing person is simplified into a cylinder or an oval cylinder. The projected area factors and the resulting view factors calculated based on these shapes are compared in terms of accuracy [29].

Before the aid of 3D models, manikins were used for thermal experiments. The physical manikin at [6] consists of 16 body segments and each segment is controlled and monitored for calculating the heat transfer. As the increasing computing power contributes to continually emerging simulation tools, different software adopt the simplification method of different degree, since the complete accounting of all view factors between every surface of the human body and the surrounding surfaces is not feasible. Among current simulation tools, there is only a small portion of software have detailed body geometry model including definable human body shape and position [12,21]. One commonly used method to simplify the human body geometry is to abstract different parts of the body into similar solid figures such as cylinders, spheres and cuboids. Many studies divide the human body into 16 to 18 segments for demonstrating the

different sensitivities towards thermal radiation and perception on these segments. One example is the ThermoSEM model representing the human body with 18 cylinders and 1 sphere, each of which has detailed layers for calculating heat transfer between different depth of the body [22]. Despite these simplifications have greater accuracy than previous methods.

Method	Level of Detail	View Factor Method	Cases
Point	Single point	Equation for the view factor between a point and a surface	Chung 2010
Silhouettes	2-dimensional sitting and standing silhouettes	Fanger-Rizzo method	Fanger 1970, Rizzo 1991
Single shape	A cube, box, sphere for a seated person; a cylinder or an oval cylinder for a standing person	Solid angle method using the projected area factor to derive the view factor	Vorre 2015
Several solid figures	Around 16-18 segments of cylinders, spheres and cuboids	Solid angle method using the projected area factor to derive the view factor	Miyanaga 2001, Schellen 2013
Mesh	Mesh with high fidelity to body surface. Discretized for test points	Numerous Vectors method	Author's method

Table 1. Evaluation of simplified body models that can be used to calculate the view factor between the human body and the surrounding surfaces.

In Huizenga et al [34], the authors demonstrate how to derive view factors for MRT from a discretized 3D model. In fact, it has been possible to use detailed body geometry model for decades in ubiquitous 3D modeling software. However, currently available MRT simulation tools still lack detailed body geometry model to achieve necessary resolution demanded by localized thermal comfort research, especially for longwave radiation.

Fig. 7 (from the following case-study) shows the simulation we have developed to integrate 3D-modeled human body geometry into an MRT analysis. For digital simulations, there is no reason to use simplified geometries of cubes or cylinders, since in today's 3D environment, even free-from or highly complex surfaces can be discretized. Using a human body mesh, discretized into triangular segments, with test points at the center of each triangulated mesh, we generate vectors according to Lambertian emissions as shown in Fig. 3, and use raytracing to obtain the radiant temperature for each test point, which can be added and averaged to receive the final MRT per equation 6. We are even able to obtain the effective radiation area of the body using the ray-tracing method, checking for self-intersections of the body which shade it from the surrounding environment (Fig. 5). As seen in Fig. 5, each area of the body may be exposed to the radiant environment to a different degree, depending on the body's position and clothing.

EXPOSURE (EFFECTIVE RADIATION) SCALE

0%   100%
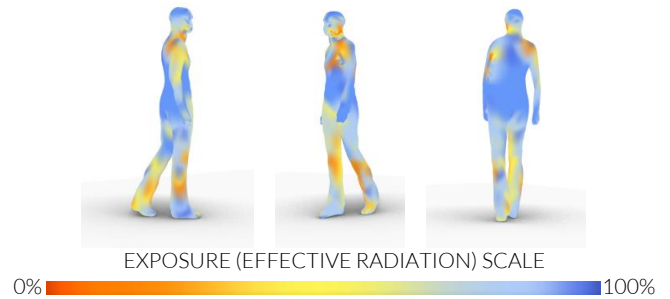
Figure 5: Effective Radiation Area of a human figure while walking.

3 CASE-STUDY

We employed the simulation method for an experimental radiant cooling pavilion we built in Singapore. Radiant cooling panels using an infrared transparent membrane to protect sub-dewpoint cooling surfaces from condensation (cite ASR) are the primary envelope in a specific geometric configuration to maximize radiant cooling. Experimental data is collected allowing the simulation to be validated and better understand the design of surfaces that influence more precisely the effect on the MRT of the occupant.
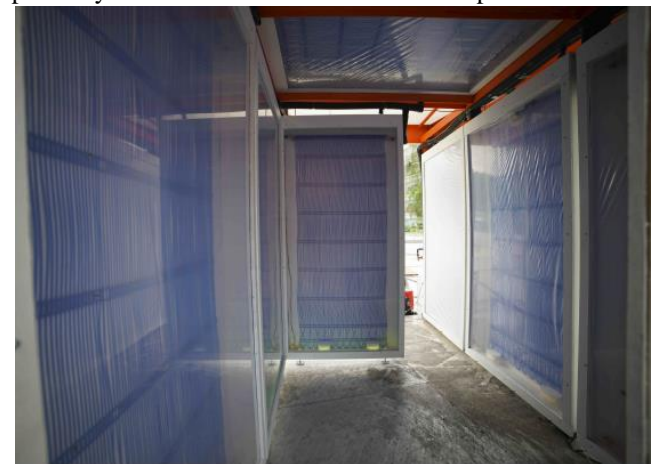
Figure 6. Interior view of the pavilion with modular radiant panels, through cold water supply.

The pavilion was constructed at the United World College, Southeast Asia in Singapore from August to October 2018. The pavilion is enclosed by ten 1.2m x 2.1m (4' x 8') panels; two horizontal panels at the top and eight vertical panels, with north and south facing entrances. The surface of the panels are cooled down by chilled water from custom variable speed chillers to provide radiant cooling. For discussion of the special cooling panel, refer to our previous work [25,26].

For the MRT simulation technique, the relevant parameters are: active cooling panel temperature: 18.7°C; surrounding ground temperature: 39.2°C; non-active pavilion surface average temperature: 31.3°C; Air Temperature: 29.6°C; Skin/clothing average Temperature:30°C.

3.1 Methods for Case-Study Simulation

Two MRT simulations are conducted: The MRT Gradient Map and Full-Body Simulation and Validation:

The MRT Gradient Map uses a single point abstraction of the body, to provide an MRT gradient map for the full space. For this first method, a grid of 750 points is created on a plane at a fixed height of 1m above the floor. At each location on this grid, 1,280 geodesically distributed rays emanate. They intersect the surfaces around them, with assigned known surface temperatures, and the temperature value at each intersection is averaged and recorded as the mean radiant temperature at each point on the grid. A color gradient is then created based on the MRT values. This method used the numerous vectors technique described in the previous section and ray-tracing with multiple bounces for reflective surfaces and Lambertian distribution. The main goal of this first simulation is to evaluate the effect of view factor of radiant surfaces on the perception of comfort at different points in space.

The Full-Body Simulation and Validation provides a detailed human body geometry mesh, where the centroid of each subdivided triangulated part of the body is used as a test point (Fig. 7). A geodesic hemisphere with Lambertian distribution (Fig. 3) is then created and the vectors emanating from the test point are ray-traced through multiple bounces, registering surface temperatures and emissivity/reflectivity/transmissivity values for each intersection. The result is computed based on Equation 6 where the numerous vector method replaces the analytical view factor, and each test point is weighted against its triangulated subdivision on the body to receive the final MRT. We check for a total MRT of the full body as well as local radiant temperature of different zones in the body.

In order to validate the results, an empirical method of sensing MRT is compared to the simulation. A pyrgeometer, a device compiled of, a set of 6 radiometers (Apogee, SL-510-SS; 0.12 mV per W/ m^2; 1% measurement repeatability; 5% calibration uncertainty; +/- °C) oriented orthogonally measuring radiant flux in all 6 cardinal directions, is used. The cubical set of 6 radiometers is considered the ground truth for the measurement of MRT since the sensing element is not sensitive to convection, unlike black globe measurements commonly used [26]. These 6 values were averaged and comprise the MRT at the measurement location.

However, in order to achieve a true comparison of the pyrgeometer reading to the simulation. It is again necessary to abstract the body back to a point in space with geodesic sphere around it, similar to the 6 faces of the device. Naturally, the pyrgeometer is not able to include the view factor of the body upon itself or the body geometry. Therefore, for the second simulation, two results are given- one for a full body, and one for a single point in space -

which could be informative for the significance of the body geometry for MRT calculations.
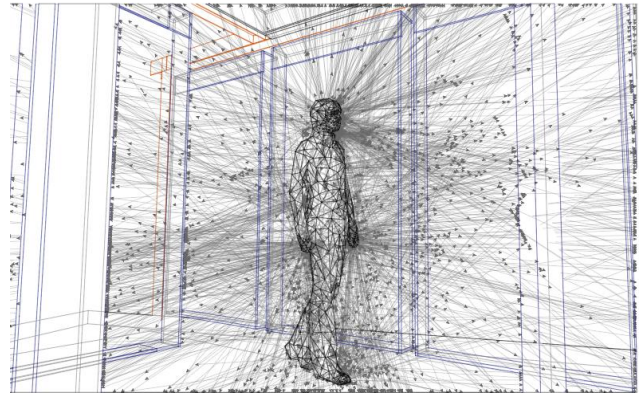
Figure 7. View of MRT simulation in progress: human body mesh inside of the radiant cooling pavilion. Ray-tracing technique from each test point on the body demonstrated.
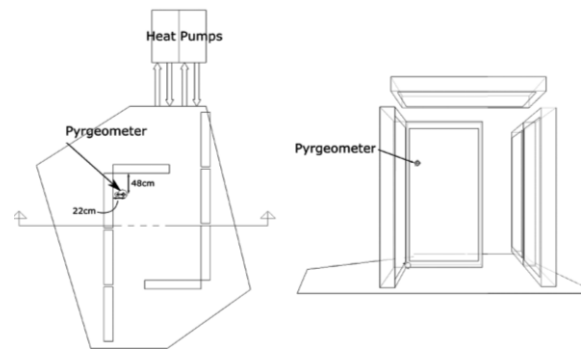
Figure 8. Plan and perspective view diagrams of the pavilion, showing placement of the pyrgeometer.

Figure 9. Pyrgeometers arranged orthogonally on a wooden cube inside the pavilion at the location shown in Figure 8.

3.2 Results for Case-Study Simulation

The MRT Gradient Map results are shown in Figure 10. It was achieved with relatively low-computational time compared to full-body simulation and provides a reasonable estimation of the change in radiant heat transfer intensity through the space of the pavilion due to the differences in the view factor of the active cooling pavilion and hot surrounding surfaces. As expected, once the simulation point is not covered by the active surfaces, the MRT rises significantly. The range of MRT in the space varies by 10°C from 21°C in the most internal corners to 31°C in the most externally exposed areas.

316
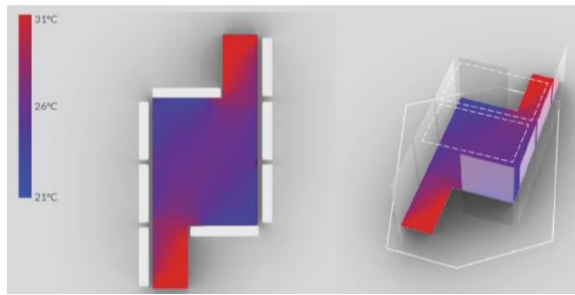
Figure 10. MRT gradient map.
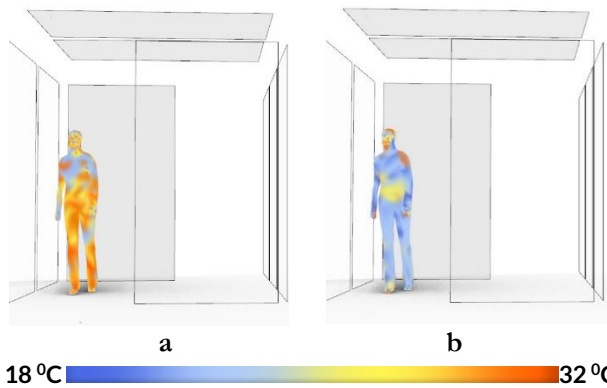
Figure 11. Radiant temperature distribution across the surface of the human body: (a) with inclusion of the body's surface self-radiation; (b) without including the radiation emitted by the body itself.

The Full-Body Simulation and Validation for a geodesic sphere around a point provided an MRT result of 23.23°C, which is in good agreement with the pyrgeometer reading at 23.9°C. However, the full-body MRT result differed from the pyrgeometer by 4.0°C: as it was simulated to be 27.9°C. This difference is caused by the warm surfaces of the body pointing at each other and increasing the average MRT of the whole body. Fig. 11a shows the perceived radiant temperature of different zones of the body when including self-shading of the body, influencing the effective radiation area of each zone. The overall body MRT is the result of averaging all these zones' radiant temperatures, normalized by the surface area of each zone. This is important to recognize as often point-based MRT values are used to calculate the radiant heat transfer from a body. In that cause often the calculated W/m^2 is multiplied by body surface area without recognizing that some of that surface area is influenced by other parts of the body or has a very different effective radiation area. A simulation considering the body geometry but not including the skin surface as part of the environment, closer to the cylinder or other geometry assumptions sometimes used, yielded a much closer MRT result of 24.5°C. Fig. 11b shows the radiant temperature as perceived by different zones of the body in this scenario. Still, this is different than the single point simulation due to the variable view factor influencing heat transfer at different

points. Therefore, for most of the body, the ceiling cooling panels have a smaller view factor than for the pyrgeometer. This demonstrates how simplifications of the body geometry into a point or cylinder are not good estimates of the real radiant heat experienced by a body.

4 CONCLUSION

MRT is an abstraction of radiant heat transfer that enables its direct integration into thermal comfort calculations and their temperature-centric considerations. The inherent complexity is challenging, and though the MRT simplification can be useful it can often mischaracterize actual heat transfer and thermal comfort conclusions. In essence, radiant heat transfer is measured in watts, and those are derived from the energetic state of surfaces, represented by their various temperatures. MRT is a proxy to this, an averaged abstraction that doesn't represent any real measured temperature but serves to unify the variations of surrounding surface temperatures into one mean variable. But in this process the complexity of view factors and geometry that also vary throughout space is often lost. In this paper, methods of light rendering simulations have been integrated into an MRT simulation to enable more complex geometry consideration, and a detailed model of the human body geometry was used instead of a single-point or platonic form abstraction to fully consider the impact of current simplifications. The cooling pavilion case-study provided a unique opportunity to study the influence of various geometric properties of architectural enclosures on the thermal sensation of the human body. The simulation results for that space demonstrated the significance of the proposed MRT simulation improvements. An important take-away of this study is that we should stop thinking about MRT as we do of air temperature. Air temperature is an environmental variable, an intensive property of the material which describes its energetic state under normal pressure conditions, easily measured by simple temperature sensors. MRT, is a representation of the wave interaction of our body with other surfaces through infrared radiation. As part of the spectrum of electromagnetic radiation, we may consider MRT in a similar manner to the concept of radiance- it is a representation of how our body is "illuminated" by all the emissions and reflections of invisible light from the terrestrial and celestial objects surrounding it.

REFERENCE

1. ASHRAE, ANSI. Standard 55-2017:"Thermal Environmental Conditions for Human Occupancy". *ASHRAE, Atlanta USA* (2017).
2. ASHRAE. Thermal environmental conditions for human occupancy ASHRAE standard 55-1966 Tech. rep American Society of Heating, Refrigerating and Air-conditioning Engineers; May 2003.
3. Cannistraro, G., et al. Algorithms for the calculation of the view factors between human body and rectangular surfaces in parallelepiped environments. *Energy and Buildings* 19.1 (1992): 51-60.

4. Chung, Jae Dong, et al. Analysis on the Impact of Mean Radiant Temperature for the Thermal Comfort of Underfloor Air Distribution Systems. *Energy and Buildings*, vol. 42, no. 12, Dec. 2010, pp. 2353–59.

5. Cohen, Michael F., and Donald P. Greenberg. The Hemi-Cube: A Radiosity Solution for Complex Environments.

6. *Proceedings of the 12th Annual Conference on Computer Graphics and Interactive Techniques*, ACM, 1985, pp. 31–40.

7. de Dear, R. J., et al. Convective and Radiative Heat Transfer Coefficients for Individual Human Body Segments. *International Journal of Biometeorology*, vol. 40, no. 3, May 1997, pp. 141–56.

8. EIA, How much energy is consumed in U.S. residential and commercial buildings? https://www.eia.gov/tools/faqs/faq.php?id=86&t=1. As of 14 May 2019.

9. Fanger, P. O. Radiation Data for the Human Body. *ASHRAE Trans.*, vol. 76, 1970, pp. 338–373.

10. Guo, Hongshan, et al. On the Understanding of the Mean Radiant Temperature within Both the Indoor and Outdoor Environment, a Critical Review. *Renewable and Sustainable Energy Reviews,* p.109207.

11. Hatefnia, N., et al. A Novel Methodology to Assess Mean Radiant Temperature in Complex Outdoor Spaces. *32th International Conference on Passive and Low Energy Architecture*. 2016.

12. Imanari, Takehito, Toshiaki Omori, and Kazuaki Bogaki. "Thermal comfort and energy consumption of the radiant ceiling panel system.: Comparison with the conventional all-air system." *Energy and buildings* 30, no. 2 (1999): 167-175.

13. Kubaha, K., et al. Human Projected Area Factors for Detailed Direct and Diffuse Solar Radiation Analysis. *International Journal of Biometeorology*, vol. 49, no. 2, Nov. 2004, pp. 113–29.

14. Lynes, J. A., et al. The Flow of Light into Buildings: *Transactions of the Illuminating Engineering Society*, Sept. 1966. Sage UK: London, England.

15. McIntyre, D. A., and I. S. Griffinths. Radiant temperature and thermal comfort. *Symposium Thermal Comfort and Moderate Heat Stress*, CIB Commission W. Vol. 45. 1972.

16. Miyanaga, Toshiyuki, et al. Simplified Human Body Model for Evaluating Thermal Radiant Environment in a Radiant Cooled Space. *Building and Environment*, vol. 36, no. 7, Aug. 2001, pp. 801–08.

17. Naboni, E., et al. An Overview of Simulation Tools for Predicting the Mean Radiant Temperature in an Outdoor Space. *Energy Procedia*, vol. 122, Sept. 2017, pp. 1111–16.

18. Newling, P. S. B. The Measurement of Mean Radiant Temperature and the Determination of the Amount of Radiant Heat Gained by a Man. *Med. Res. Coun., Lond., RNP Rep*, no. 54/784, 1954.

19. Rakha, T., P. Zhand, and C. Reinhart. A Framework for Outdoor Mean Radiant Temperature Simulation: Towards Spatially Resolved Thermal Comfort Mapping in Urban Spaces. *Proceedings of the 15th IBPSA* (2017): 2414-2420.

20. Rhee, K. N., & Kim, K. W. (2015). A 50 year review of basic and applied research in radiant heating and cooling systems for the built environment. *Building and Environment*, *91*, 166-190.

21. Rizzo, G., et al. Algorithms for the Calculation of the Mean Projected Area Factors of Seated and Standing Persons. *Energy and Buildings*, vol. 17, no. 3, Jan. 1991, pp. 221–30.

22. Roudsari, Mostapha Sadeghipour. Ladybug-primer: Outdoor solar temperature adjustor. https://mostapharoudsari.gitbooks.io/ladybug-primer/content/text/components/Outdoor_Solar_Temperature_Adjustor.html

23. Schellen, L., et al. The Use of a Thermophysiological Model in the Built Environment to Predict Thermal Sensation: Coupling with the Indoor Environment and Thermal Sensation. *Building and Environment*, vol. 59, Jan. 2013, pp. 10–22.

24. Sillion, F., and C. Puech. A General Two-Pass Method Integrating Specular and Diffuse Reflection. *Proceedings of the 16th Annual Conference on Computer Graphics and Interactive Techniques*, ACM, 1989, pp. 335–344.

25. Tanabe, Shin-ichi, et al. Effective Radiation Area of Human Body Calculated by a Numerical Simulation. *Energy and Buildings*, vol. 32, no. 2, July 2000, pp. 205–15.

26. Teitelbaum, E., Chen, KW., Meggers, F., et al Globe Thermometer Free Convection Error Potentials. In Submission. Jan. 2020.

27. Teitelbaum, E., Chen, K. W., Meggers, F., Guo, H., Houchois, N., Pantelic, J., & Rysanek, A. (2020). Globe thermometer free convection error potentials. *Scientific Reports*, *10*(1), 1-13.

28. Thorsson, S., Lindberg, F., Eliasson, I., & Holmer, B. (2007). Different methods for estimating the mean radiant temperature in an outdoor urban setting. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *27*(14), 1983-1993.

29. UNDERWOOD, C. R., and E. J. WARD. The Solar Radiation Area of Man. *Ergonomics*, May 2007.

30. Vorre, Mette Havgaard, et al. Radiation Exchange between Persons and Surfaces for Building Energy Simulations. *Energy and Buildings*, vol. 101, Aug. 2015, pp. 110–21.

31. Ward, Gregory J. The RADIANCE Lighting Simulation and Rendering System. *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques*, ACM, 1994, pp. 459–472.

32. Weik, Martin H. "Lambert's Cosine Law." In *Computer Science and Communications Dictionary*, edited by Martin H. Weik, 868–868. Boston, MA: Springer US, 2001. https://doi.org/10.1007/1-4020-0613-6_9901.

33. Howell, John R. "The Monte Carlo method in radiative heat transfer." (1998): 547-560.

34. Huizenga, Charlie, Zhang Hui, and Edward Arens. "A model of human physiology and comfort for assessing complex thermal environments." *Building and Environment* 36, no. 6 (2001): 691-699.

35. Luo, Maohui, Zhe Wang, Hui Zhang, Edward Arens, Davide Filingeri, Ling Jin, Ali Ghahramani, Wenhua Chen, Yingdong He, and Binghui Si. "High-density thermal sensitivity maps of the human body." *Building and Environment* 167 (2020): 106435.

318

Fine Tuning of Aperiodic Ordered Structures for Speech Intelligibility

Antigoni Karaiskou[1], Martin Tenpierik[2] and Michela Turrin[3]

[1]Buro Happold
Bristol, United Kingdom

anti.kara91@gmail.com

[2] TU Delft
Delft, The Netherlands

M.J.Tenpierik@tudelft.nl

[3] TU Delft
Delft, The Netherlands

M.Turrin@tudelft.nl

ABSTRACT

Speech intelligibility is crucial in many spaces, yet designers often fail to predict the acoustic shortcomings of certain design choices. This paper builds on the potential of hybrid surface treatments showcasing low-frequency absorption to control background noise levels and high-frequency diffusion to improve speech-in-noise perception to introduce a workflow that encodes this information in a format easily perceived by designers. After patterns are being classified based on periodicity into partly periodic, non-periodic or aperiodic, a matrix serves as a rule of thumb communicating to non-experts the critical variables for high-frequency diffusion, such as well depth sequence, scale and profile. These become inputs of a computational process that generates variations to tailor patterns for speech intelligibility. Lastly, plotted graphs that visualize quantitative figures obtained from simulations are marked by a bounding box relative to the effective frequency range for designers to evaluate examined patterns during the process of optioneering. This integrated workflow targets architects and designers that seek for visual feedback to support an iterative exploration of performance driven geometries, while recognizing the contribution of aperiodic order to uniformly distribute the flow of sound energy.

Author Keywords

Speech intelligibility; Surface tiling; Micro-geometry; Correlation scattering coefficient; Quasi-periodic symmetries;

1   INTRODUCTION

Advanced acoustical studies are mainly carried for performance venues rather than daily life environments as those require over-engineered solutions that raise the cost. At the same time, current computational tools are mostly focused on simulations that deliver numerical outputs that do not clearly correspond to geometrical equivalents. The novelty of this paper in the field of simulation lies in suggesting a coherent workflow that encodes complex quantitative results obtained from simulations in a format easily perceived by non-experts, thus greatly contributing to a transdisciplinary design approach to tune ordinary spaces.

Speech transmission in a multi-talker scenario is strongly influenced by the properties of a room requiring successful management of competing factors such as space size, proportion, geometry, material properties and surface details. In that respect, the relationship among sound and geometry is projected on both the spatial context and the surface properties. Nowadays the focus of acoustic simulations has been shifted from evaluating the performance of macro-geometry to obtaining coefficients for micro-geometries as it was proven that adjusting properties at a surface level can improve speech intelligibility for a given macro-geometry.

In the past, most efforts were focused on sound absorption, but over the last decade there has been a growing interest in improving sound performance in terms of scattering by generating complex micro-geometric surfaces through robotic fabrication methods. Most of these examples sought to avoid periodicity by increasing the complexity of the reflected sound field, thus generating complex surface patterns by means of computational resources based on randomness. According to recent findings, quasi-periodic symmetries show superior performance over the common diffusers due to their potential of eliminating periodicity [1]. Such structures seem promising for uniform sound energy dispersion avoiding the formation of bundled or looped reflections.

2   FRAMEWORK

Based on the findings of a MSc thesis for improving intelligibility for a typical absorptive classroom, this paper aims to contribute to the current debate by comparing spectral properties of highly ordered structures that are substantially nonperiodic, or aperiodic to partly periodic alternatives targeting the frequency range for speech intelligibility [2]. The study begins with a preschool environment as a problematic case study for speech intelligibility to define the targeted percentage of absorption and diffusion in the effective frequency range based on certain users' characteristics. The scope of the paper limits itself in tuning micro-geometric surfaces for high-frequency diffusion. Trying to investigate the effect of periodicity for a given pattern, it argues in favor of the theory of aperiodic order, suggesting mathematical structures that possess order without periodicity by regulating physical properties of spatial surfaces [3]. In that respect, it identifies a global set of critical geometrical properties that can influence the sound profile and utilizes Pachyderm2.0RC20 by plotting the

correlation scattering coefficient for comparison purposes. Discussing the results leads to a shift of the global set of properties to a local equivalent relatively to a pattern's periodicity classification.

The outcome if this, is a workflow that suggests a stepped process for tuning micro-geometric surfaces with a given amount of low-frequency absorption for high-frequency diffusion. In that respect, designers need to identify the effective frequency range per task and user group, define the amount of diffusion per frequency range and have it numerically manifested by giving a percentage (%), generate pattern iterations changing critical input parameters based on a local set of properties and compare the generated samples in terms of speech intelligibility by means of plotting the correlation scattering coefficient. Utilizing a computational workflow to quantify scattering seems critical for giving designers a visual feedback on the acoustic performance of their iterative explorations.

2.1 Speech Intelligibility in preschool environments

The first step lies in relating the targeted sound wavelengths for a given user group with the effective frequency range of both absorption and diffusion. To showcase this process, the example of a preschool environment serves as a challenging case study, as this dense and interactive environment poses a threat to intelligibility by running at a high occupancy rate, thus having an acoustic dynamic of high indoor ambient noise levels.

One of the most computationally demanding tasks for the brain is to recognize the message from the din since noise and speech are transmitted as a single stream. Speech-in-noise perception in a preschool environment is strongly related to the user's characteristics with neurological processes of encoding fundamental speech properties securing high signal-to-noise (SNR) ratio. Since cognitive and biological development are correlated, preschool children and those facing learning difficulties are two of the most vulnerable groups [4].

Four categories of uncomfortable sound can be recognized related to sound typology and source identification, namely, threatening sounds -screaming, crying, anger- , high frequency sounds -squeaking, creaking, scratching- , background sounds -fans, radiators, ventilators- and communication sounds [5].

2.2 Acoustic Requirements

Speech power comes from low frequency vowels (125-500Hz), whereas speech intelligibility is given by high frequency consonants (2000-6000HZ) [6]. Consonants consist of low intensity and short duration acoustic cues with a changing spectral content, whereas vowels are of higher amplitude, last longer and show an almost stable spectral content [7].

Harmonics enhance neural encoding by reinforcing the periodicity of a target speaker's fundamental frequency (F0) extracting relevant information from the auditory stream,

thus allowing for speaker localization and analysis of the auditory scene [8]. Several studies proved that the response quality for high frequency acoustically dynamic speech cues, as well as harmonics, is degraded by noise, leaving them more susceptible to masking [7]. Early sound reflections, though beneficial, can affect timbre and localization phenomena if coming from nearby adjacent surfaces, resulting in listeners localizing the talker to another location than where (s)he is standing. Later reverberation is known to mask speech. In that respect, all reflections, but especially the ones arriving after 40 to 50 ms, can be considered a form of noise correlated to the signal [9].

Relevant acoustic measures

Speech intelligibility correlates with various measures, such as reverberation time and signal-to-noise ratio (SNR) to estimate the combined effect of room acoustics and background noise. Either too short or too long reverberation time will decrease speech audibility and speech intelligibility respectively, thus one should consider both speech transmission index (STI) and mid-frequency (500 and 1000 Hz octave frequency bands) reverberation time (T_{30}) to estimate the effective frequency range for either absorption or diffusion.

Measures	Parameters		Comments
Signal to Noise Ratio	SNR [dB]	15	Remove excessive absorption
Reverberation Time	$T_{30[s]}$	0.4-0.6	If too low, SPL can be reduced
Intelligibility	STI	0.6-1	An alternative would be U_{50}
Noise Level	$L_{A,eq}$ dB(A)	< 60	Typical: 80 dB(A)

Table 1. Target values as given from literature

2.3 Effective Frequency Range

Speech intelligibility is related to more than the acoustic measures. The sound level and its characteristics are influenced also by room acoustics, total number of children, pedagogic methods applied, ventilation systems. Typical values of sound pressure levels (SPL) in preschools are in the range of 75-80 dB(A), with children being exposed to higher SPL than personnel, namely 85 dB(A) [5]. Important to note is that there is a correlation between high SPL and frequency spectrum, in that high rates are reached in sound frequencies of 1000-4000 Hz, namely the typical spectrum of children's voices.

Improving speech intelligibility for a preschool environment requires a hybrid treatment, as low frequency absorption prevents amplifying vowel sounds and removes first-order strong reflections, whereas high frequency diffusion delays and temporally diffuses early reflections. For absorption to prevent amplifying vowel sounds of active children, lower frequencies, especially those at 125 Hz, should have the

shortest reverberation time possible, in order to approximate the reverberation graph of a forest, which is argued to achieve high intelligibility properties [10]. The exact amount of absorption and scattering per effective frequency bands is given in Table 2 and is based on the findings of the aforementioned MSc thesis assigning high values for low frequency absorption and mid-to-high values for high frequency diffusion after conducting a Geometric Acoustics Modelling and Simulation study [2].

	Absorption (%)		Diffusion (%)	
Frequency range	Low	High	Low	High
Effective %	**90**	25	25	**50**

Table 2. Effective values of absorption and diffusion per frequency range

3 FINE TUNING OF MICRO GEOMETRIES

After defining the effective frequency range based on the preschool case study, absorption and diffusion targets need to be mapped to equivalent design rules. Though various databases deducted from current literature allow designers to pick and choose among design rules for low-frequency absorption, there is no equivalent workflow for tuning micro-geometric surfaces for high-frequency diffusion. Thus, the study aims to utilize outputs closing this gap.

The most typical device for diffusion throughout history is Schroeder's Quadratic Residue Diffuser (QRD), a phase grating device composed of a periodic arrangement of wells of equal width varying the depths, whose sequence is based on number theory. However, this design is stigmatized by various setbacks, such as unattractive visual appearance, specular reflections at high frequencies when increasing the well width beyond a certain value and non-uniform distribution of the reflected sound field due to the periodic repeating sequence. Certain modulation schemes around this type allowed for both spatial and temporal dispersion, causing a complex reflected wave front, while achieving high absorption at low frequencies due to the resonance created by the deep well system. However, these diffusers still have the issue of periodicity.

3.1 Mitigating the Problem of Periodicity

Periodicity can be problematic for scattering directivity as it is dependent on incidence angle and frequency. Thus, changing the positions of the sound source or altering the effective frequency range based on user characteristics and functional requirements can cause problems.

The concept of randomness for useful scattering
To eliminate periodicity, random sequences are utilized generating random or pseudo-random designs by arranging elements in a non-repeating way [1]. Patterns that resulted from random sequences showed the potential to increase the frequency range of useful scattering [11]. However, at high frequencies the dispersion by a modulated array can often be worse than for a periodic arrangement, meaning that custom designs need to be tested for their diffusion performance [12]. Such random patterns are also hard to generate or predict concerning their performance, due to them lacking a degree of order that defines their physical properties making them cost prohibitive.

Spectral properties of aperiodic order
A promising alternative to current methods that mostly yield numeric performance indicators for estimating scattering performance, is to analyze the spectral properties of diffraction, which might be better suited for interpretation by non-experts.

The spectral measure of diffraction is given by autocorrelation, which is interpreted as a Fourier transform, namely a weighted sum of periodic functions representing the harmonic perception of sound. In that respect, the diffraction pattern can be critical for encoding information about the spatial autocorrelation, whereby point spectra indicate order, which can be measured by analyzing the contribution of each frequency to the autocorrelation [3].

Physical qualities of quasi-periodic formations
Another optimum solution for uniform scattering performance could be a single, long number sequence with good aperiodic autocorrelation, as it gives just a single period without repetition. In that respect, typologies that, though globally unstructured, show a high degree of regularity locally seem promising towards achieving non-repetitive sequences. Highly ordered structures that possess order without periodicity, hence being substantially non-periodic or aperiodic, exhibit surprising properties linking to many areas such as dynamical systems, harmonic analysis, spectral theory and number theory.

Recently, a study explored the effect of different properties, such as structural height, arrangement and profile shape on random- and normal- incidence scattering coefficients for Penrose tiling patterns. Further, it tested fractal order, an inherent property of quasi-periodic formations, which showed high scattering capacity for a wide frequency range as well as high uniformity of scattering directivity [13].

3.2 The Mathematical Model of Surface Tiling

Since recent literature argues in favor of random, non-periodic and quasi-periodic formations, this study tries to fit in this context by comparing tiling arrangements that fall under these categories.

Tiling is defined as a countable group of closed sets that covers the plane in a non-overlapping manner [14]. In a mathematical sense this is represented by a finite set of shapes arranged in a system that tessellates a surface via geometric patterning, whereby local rules give the constraints on the way neighboring tiles can fit together [15].

In terms of periodicity, a tiling is defined as periodic if there exist, among the symmetries of the tiling, at least two translations in non-parallel directions. Meaning that one can

outline a region of the tiling and tile the plane by translating copies of that region without rotating or reflecting. When no translations can recreate the same region, a tiling is thought to be non-periodic. A set of prototiles that can never tile the plane periodically is said to be aperiodic, often referred to as tiling of unbounded spaces.

Comparing hierarchical structures of self-similar systems
A performance comparison among three tiling arrangements, as a true representative sample of the general population of those structures, serves to identify important physical properties per effective frequency range. For tessellating the plane, the discretization strategies of Penrose tiling, Voronoi diagram and periodic centroidal Voronoi were applied to parametrically define test samples of 1 m².
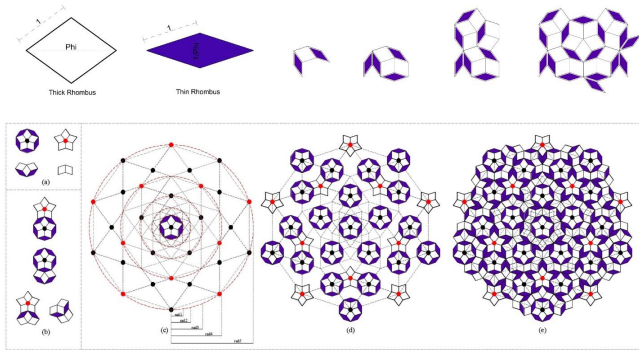

Figure 1. A framework of nested decagrams based on φ, serves as the underlying hidden grid for guiding the construction process of the quasi-periodic system.
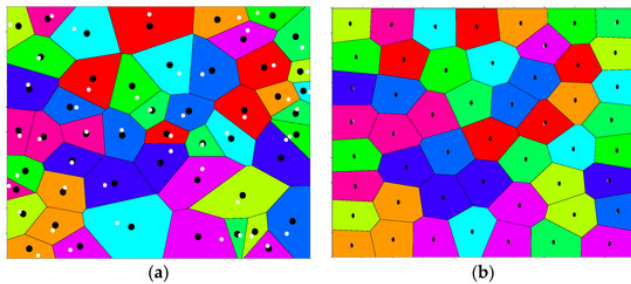

Figure 2. Random and centroidal Voronoi diagram: generators (white dots) and centroids (black dots): (a) random Voronoi diagram, and (b) centroidal Voronoi diagram.

The most famous aperiodic prototiles were invented by Roger Penrose in 1974. As shown in Figure 1, the so called "twin parallelograms" is an aperiodic tessellation of the plane by two prototiles without gaps or overlaps. The pair of rhombuses has equal sides, but their angles differ. The thick rhomb has angles of 72° and 108° and the thin rhomb 36° and 144° [13]. An interesting property is that the ratio of the two is the golden mean. Further properties involve self-similarity found mostly in fractals and finally a fivefold rotational symmetry due to quasi-periodic structure. Though Penrose tiling seems a chaotic sequence, it does represent a very regular local structure by having greater organizational depth [14].

The second typology, known as the Voronoi diagram achieves tessellation of a plane dealing with proximity to a countable set of sites P. The Voronoi region V(P) in the Euclidean plane consists of all the points at least as close to P as to any other site. As shown in Figure 2, the set of those Voronoi regions constitutes a plane tessellation denoted by Vor (P) and goes by Voronoi diagram generated by P. Following the rules for periodicity a Voronoi diagram of a non-periodic site set is a non-periodic tiling of the plane [14].

A special type of Voronoi known as periodic centroidal Voronoi tessellations (PCVTs) is shown in Figure 2, building the third typology for this study. In this case the generators of the Voronoi tessellations coincide with the mass centroids of the respective Voronoi regions and satisfy certain properties that enforce periodicity regarding some unit cells. The most common method for computing CVTs is the Lloyd algorithm [16].

Computational workflow for identifying design variations
The patterns were remapped on surfaces creating three-dimensional geometries based on certain properties. To identify whether a global set of critical properties would be applicable on all patterns whether aperiodic, non-periodic or partly periodic, a computational workflow was developed.

To ensure functional equivalency among the samples following rules were applied for generating the patters and deciding on the well sequence per diffusor. For the aperiodic sample of the Penrose tiling tessellation a grasshopper script in C# generating Penrose tiling was modified, whereby the thick rhombus represented the well of the diffusor. The non-periodic sample was based on Voronoi seeds that allowed for random seed generation to define the relative arrangement of the cells. The scale, namely the total amount of cells was set in respect to the proportions of the previous sample, whereby the 4-, 5- and 6-sided polygons were selected as the wells of the device. For the partly periodic sample, the centroidal Voronoi (CVT), scale and wells were similarly set relative to the non-periodic sample constrains. The test samples are shown in Figure 3.
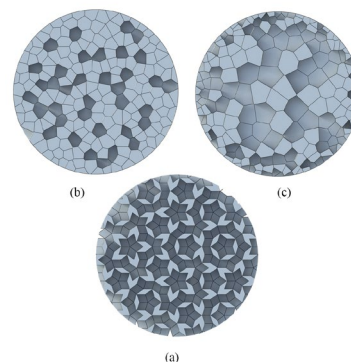

Figure 3. Test samples tessellations. Define wells and number of cells: (a) Penrose tiling, (b) Voronoi diagram and (c) Periodic Centroidal Voronoi

An overview of the simulation alternatives and their variables is given in Table 3. Fractal scale, geometric profile and well depth sequence were identified as global critical properties. The comparison distinguished among the following variables: a large or a small scale, an extruded or a tapered profile, two or various amount of well depths of either 20%, 50%, 150% or 200% depth-to-width ratio.

Critical Properties	Variables		
	Scale	Profile	Well depth sequence
Aperiodic	Cells=2x Cells=x	Extruded Tapered	Two depths Various depths	(20/ 50/ 150/ 200) %
Non-periodic			
Partly periodic			

Table 3. Samples: physical properties and relevant variables

There are certain reasons behind evaluating this global set of properties. The scale serves to investigate whether the same scattering performance can be achieved with half the number of base elements, whereas the profile whether different angles could contribute to spatial dispersion of the sound field. Defining relevant well depths for comparison purposes is a challenging task, strongly related to the targeted frequency range. Varying among 20%, 50%, 150% and 200% depth-to-width ratio was decided based on the following. To achieve soft reflections a depth of 10-30 mm is efficient, whereas to eliminate echoes a greater depth of 50-90 mm is more appropriate [17]. The wavelength of sound (λ) equals the speed of sound divided by frequency (λ= c/f). A well depth that is given as ¼ of the wavelength λ, results in a 180° phase shift relative to a part of the surface without well. This procedure can determine the different depths relative to frequency ranges of 1 kHz - 4 kHz, due to the potential for local sound cancellation and complex reflection patterns [18].

3.3 Comparison Method for Iterative Design
Evaluating diffusion is complex and requires a more thorough investigation since (as absorption) too much or too little can result in acoustic aberrations. Achieving complete scattering at the design frequency and its multiples ensures that energy is not reflected in the specular direction but does not evaluate the quality of the produced dispersion. A diffuser apart from redirection needs to achieve spatial dispersion in order to reduce echoes without moving problems to a different place. Though energy dispersion is the most dominant approach for evaluating current diffusers, the phase in the polar response or the reflected impulse response are equally important. An ideal diffuser produces a polar response that remains constant when changing the angle of incidence and the frequency within its operational bandwidth.

Computational workflows that quantify scattering characteristics can serve designers get a visual feedback on the acoustic performance of their iterative explorations, thus allowing comparison of the scattering effect of the reflected wave front created by various output surfaces within a given room's macro-geometry.

Coefficients per band aperiodicity
The Schroeder Quadratic Residue Diffuser allowed for the first time to evaluate complete diffuse reflection due to periodic phase grating energy that caused grating loops [1]. After this, the most recognized methods for determining scattering coefficients can be summarized as follows: the standard method ISO 17497-1 performed in a reverberation room measuring random incidence values, a laboratory method in a rectangular room measuring normal-incidence values and a BEM-based numerical method in a free field measuring directional values [13]. However, such studies are complex and time-intensive, and their results can be interpreted only by experts.

Another method for evaluating a diffuser performance is estimating the correlation scattering coefficient from polar responses which was originally suggested by Mommertz [19]. The scattering coefficient in that case is given by the correlation of the scattered pressure polar responses from the tested surface and a flat surface. Hence, the measurement gives the dissimilarity between the test and flat surface scattering over a polar response, rather than the energy moved from the specular direction when the surface is moved. For randomly rough surfaces, the ISO scattering and correlation scattering coefficients will be similar, but for diffusers with distinct polar responses this is not the case [12].

Correlation scattering coefficient plots
The sample evaluation in this study utilized the software Pachyderm, version *Pachyderm2.0RC20* using the Finite Volume Method, which, though experimental for reliable quantitative predictions, is useful for comparison purposes of alternatives related to geometrical properties [17]. This is an open source code that recently incorporated a wave-based analysis of Finite Volume Method for estimating the correlation scattering coefficient as an input for Geometrical Acoustics software, although it only considers normal incidence.

For designers to support their iterative explorations with performance measures, the correlation scattering coefficient is thought to quickly visualize results per sample giving an indication of performance for comparison purposes among promising alternatives already checked for compliance if adhering to the rules previously defined as a critical set of global properties found in Table 3.
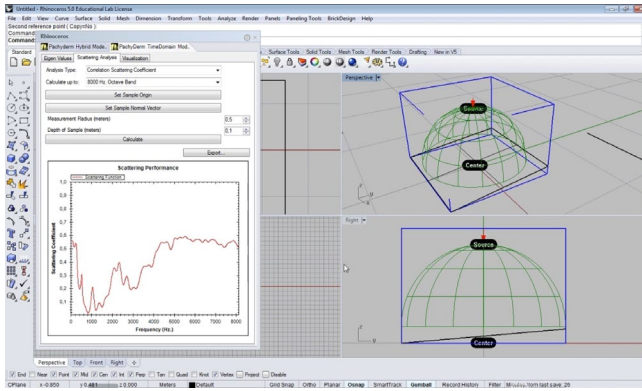
Figure 4. Coherent methodology for plotting the correlation scattering coefficient for comparison purposes

To minimize discrepancies on the obtained quantitative results a coherent methodology guided the various simulations. The samples were placed with their center at the origin (0,0,0) entirely below the given green sphere as show in Figure 4, while cleaning the model from any unnecessary geometry apart from the sample. Next, it was required to set a 'Measurement Radius' wide enough for a sphere of that radius to fully encompass the sample. Like the process followed in the scaled reverberation chamber measurements, the radius of the samples was set to r = 0.5 m with a circular perimeter.

4   VISUAL FEEDBACK

4.1 Rules for Compositional Pattern Producing Networks

The samples discussed in Section 3.2 utilized the global set of critical properties given by Table 3 to generate valuable alternatives, which were then fed into Pachyderm2.0RC20 that visualized the obtained quantitative results. As shown in Figure 5, a bounding box defines the relevant frequency range of the plotted graphs allowing for certain values of the correlation scattering coefficient plots. The relevant plot of the correlation scattering coefficient, as explained in Section 2.3, should be constrained within the range of 2000-7000 Hz obtaining values between 0.5-0.9.

For the well depth sequence of two depths, increasing the depth-to-width ratio up to 150% improved the performance of the samples only for the non-periodic sample, whereas increasing the scale has a positive impact for partly periodic ones. Choosing a well depth sequence of various depths of 200% depth-to-width ratio shows a better performance for the non-periodic and partly periodic samples, whereas the angled profile gives only a marginal improvement over the extruded in all tested cases.

Contributing to the discussion regarding the theory of aperiodic order, only one sample gave a scattering plot fully enclosed within the useful frequency range for speech intelligibility, namely the case of an aperiodic sample (cells=2x, extruded, two depths with a depth-to-width ratio=0.5). Though the correlation scattering coefficient is

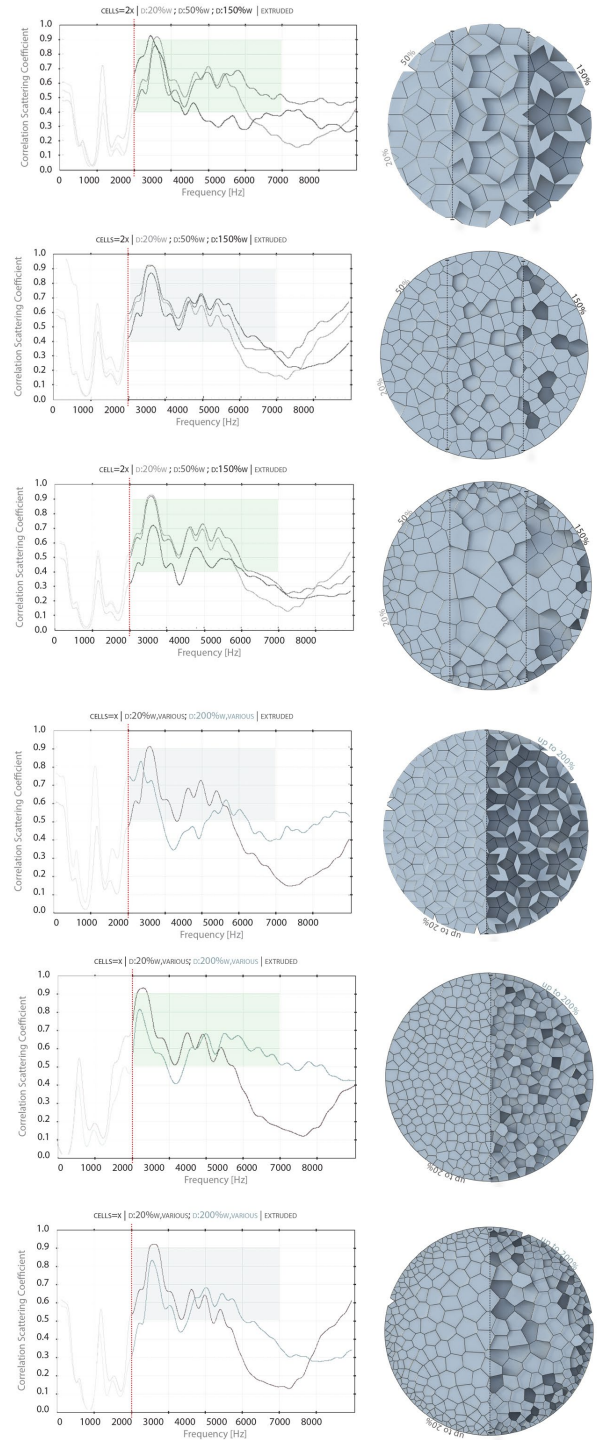**Figure 5.** Iterative sample generation plotting the correlation scattering coefficient for comparison purposes

not enough to prove that the aperiodic samples perform superior to the non-periodic and partly periodic ones, this result is thought to be relevant arguing in favor of the theory of aperiodic order, given the 20 different configurations for assessing the critical properties for the three arrangement types.

Guidelines based on the discussion of the results

The results obtained by this analysis as shown in Table 4 favor simplicity over complexity and are valid for the patterns tested for the purpose of this study, hence a generalization should be avoided. Nevertheless, the critical properties do change based on samples being aperiodic, non-periodic or partly periodic but whether the suggested variables are applicable to all patterns that fall under those categories would require further testing of examples that fall under these classes.

Critical Properties	Variables			
	Scale	Profile	Well depth sequence	
Aperiodic	Cells =2x	Extruded	Two depths	depth-to-width = 0.5
Non-periodic	Cells =2x	Extruded	Various depths	depth-to-width = 2
Partly periodic	Cells =2x	Extruded	Various depths	depth-to-width = 0.5

Table 4. Matrix providing optimum variables per case study

However, there is a clear shift from a global set of properties to a local equivalent based on each pattern's periodicity classification. In that respect the matrix proposed in Table 4 serves in identifying the local set of critical properties for generating valuable design alternatives per pattern. Further guidelines for the digital evolution of compositional patterns followed by designers could serve a performance-related approach for structures that comply with the concept of aperiodic order.

1. Based on a pattern's periodicity classification, the optimum depth-to-width ratio needs to be defined, without exceeding a certain threshold since deeper valleys could potentially increase absorption. Note that the design frequency per device cannot be shifted relative to changing the well depth for normal incidence, as utilized by Pachyderm2.0RC20.

2. The complexity and the number of elements should be decreased by avoiding various depths and increasing the scale where possible.

3. Different geometry profiles are encouraged to be investigated, while being aware that small scale specular reflection is expected for angled or curved surfaces at high frequencies.

4.2 Limitations and Suggestions

The suggested workflow examined the effect of physical properties of aperiodic mathematical structures compared to partly-periodic and non-periodic only on normal-incidence scattering coefficient, as it was found that the scattering coefficient for Penrose-tiling-type diffusers are largely independent from incidence angle [13]. However, further compatible workflows should as well as simulate random incidence scattering coefficients, while testing the uniformity of scattering directivity, given that these properties are critical to evaluate the performance of the reflected sound field.

Towards diffuser optimization the diffusion coefficient, which is frequency dependent, is of great importance, complementary to the scattering coefficient, providing an index to measure its quality. To avoid confusion due to edge diffraction at low frequencies, the normalized diffusion coefficient should be considered evaluating the level of uniform spatial reflections [12].

As far as the patterns are concerned, quasi-periodic formations seem promising for estimating further acoustic properties. Fractal order, which gives higher uniformity into the polar distribution was not considered for this paper but is a critical property especially if considering hierarchical structures of self-similar systems. Structures that allow self-similar systems to emerge offer a possibility to broaden the effective frequency range. Further, utilizing three-dimensional quasi-periodic sequences for well depths seems promising, investigating space-filling rather than plane-filling geometries, such as Ammann's tiles, rhombohedra etc.

5 CONCLUSION

Engineering sound propagation as a function of spatial volume and surface properties can influence the way in which an emitted sound is heard after being projected back into the space. Best practice considers both the macro-geometry of a space, including curved, flat or intersected walls, as well as fine-tuning the micro-geometry of various surface patterns. Certain geometrical properties can allow design products to act as acoustic filters, concentrating and/or dispersing sound in temporal and frequency domains [14]. Modern parametric modeling allows the quick iteration of many design alternatives, supporting a generative design workflow to create variations based on critical properties. To evaluate those based on performance measures, simulations that are strongly dependent on geometry, rather than material properties are required. In that respect, the correlation scattering coefficient can be useful as a comparison method for non-experts to visualize the acoustic performance of their design iterations. As opposed to the assumption that in general more complex surfaces are expected to yield greater scattering, this study argues that a lower level of complexity can still yield a complex wave front if hierarchical structures that exhibit aperiodic order are utilized. The suggested design-oriented workflow identifies the critical properties of those structures and allows designers to run iterations creating and testing different configurations in respect to

uniform scattering across a broad spectrum to ensure intelligibility of everyday spaces.

REFERENCES

1. Ajlouni, R., *Quasi-periodic Geometry for Architectural Acoustics.* ENQ Enquiry A Journal for Architectural Research, 2018. **15**(1): p. 42-61.
2. Karaiskou, A., *Bricks as Spatial Sound Modulators: Towards tuning the geometry.* 2018.
3. Baake, M., D. Damanik, and U. Grimm, *Aperiodic order and spectral properties.* 2017.
4. Thompson, E.C., et al., *Individual differences in speech-in-noise perception parallel neural speech processing and attention in preschoolers.* Hearing Research, 2017. **344**(44): p. 148-157.
5. Waye, K.P., *Sound, mind and emotion research and aspects.* 2009.
6. D'Antonio, P., *Optimizing the signal to noise ratio in speech rooms using passive acoustics.* Proc. Meet. Acoust. Proceedings of Meetings on Acoustics, 2013. **19**.
7. White-Schwoch, T., et al., *Auditory-neurophysiological responses to speech during early childhood: Effects of background noise.* Hearing research, 2015. **328**: p. 34-47.
8. Kraus, N. and T. White-Schwoch, *Neurobiology of Everyday Communication: What Have We Learned from Music?* Neuroscientist, 2017. **23**(3): p. 287-298.
9. Charles, M., *Acoustics : architecture, engineering, the environment.* 1998, San Francisco, CA: William Stout Publishers.
10. Minna, H., et al., *Child & Noise : How does the child percieve the sound environment?* 2017.
11. Reinhardt, D., *The sound of space in 3 robotic prototypes: Introducing 6-axis robotic fabrication to shape macroand micro-geometries for acoustic performance.* AZ A/Z : ITU journal of Faculty of Architecture, 2018. **15**(1): p. 79-92.
12. Cox, T.J. and P. D'Antonio, *Acoustic Absorbers and Diffusers, Third Edition: Theory, Design and Application.* 2016, Boca Raton; Florence: CRC Press: LLC Taylor & Francis Group
13. Hyojin, L., T. Yuzo, and S. Tetsuya, *Acoustic scattering characteristics of Penrose-tiling-type diffusers.* APAC Applied Acoustics, 2018. **130**: p. 168-176.
14. Cáceres, J. and A. Márquez, *An aperiodic tiles machine.* 2002.
15. Ostwald, M.J., *Architecture and Mathematics from Antiquity to the Future Volume II: The 1500s to the Future.* 2015.
16. Zhang, J., Q. Du, and M. Emelianenko, *Periodic centroidal voronoi tessellations.* Int. J. Numer. Anal. Model. International Journal of Numerical Analysis and Modeling, 2012. **9**(4): p. 950-969.
17. Rahm, P., M. Ramsgaard Thomsen, and M. Nguyen, *Humanizing Digital Reality Design Modelling Symposium Paris 2017.* 2018.
18. Reinhardt, D., et al., *Robotic Fabrication in Architecture, Art and Design 2016 / Dagmar Reinhardt, Rob Saunders, Jane Burry, Editors; Foreword by Sigrid Brell-Çokcan and Johannes Braumann, Association for Robots in Architecture; with contributions by Marjo Niemelä.* 2016, Springer: Switzerland.
19. Vorlander, M. and E. Mommertz, *Definition and measurement of random-incidence scattering coefficients.* Applied acoustics. Acoustique appliqué. Angewandte Akustik., 2000. **60**(2): p. 187-199.

A Performance-Driven Approach for the Design of Cellular Geometries with Low Thermal Conductivity for Application in 3D-Printed Façade Components

Valeria Piccioni[1], Michela Turrin[2] and Martin J. Tenpierik[2]

[1]Delft University of Technology
Delft, Netherlands
valeria1.piccioni@gmail.com

[2]Delft University of Technology
Delft, Netherlands
{m.turrin, m.j.tenpierik}@tudelft.nl

ABSTRACT

Additive manufacturing allows the fabrication of complex geometries with enhanced performances, making it interesting for application in façade components. Assessing the performance of non-standard geometries and 3D printed parts requires a combination of digital and analytical methods to retrieve validated models which can guide the design process. In this study a 3D printed mono-material façade component was designed, where the complex geometrical configuration enhance its thermal insulation properties. For this, a digital workflow was developed, encompassing performance-driven design, performance assessment and geometry generation for fabrication.

Analytical heat transfer models, heat flux measurements, and heat transfer simulations with COMSOL Multiphysics were used to assess the thermal properties of different geometrical alternatives. By observing and comparing the results, a validated model was defined to retrieve design guidelines and thermal performance indicators. The results identify porosity as the driving factor for thermal insulation and clarify the nature of the heat transfer in 3D printed cellular structures. Open surface-based geometries were preferred for the good combination of thermal properties and manufacturability. The findings are embedded in a digital workflow in Rhino-Grasshopper, enabling the design of insulating cellular structures to be used in 3D printed façade components.

Author Keywords

Thermal insulation; FEA simulations; Parametric Design; Additive Manufacturing; Façade design; Performance-driven design

ACM Classification Keywords

Applied computing – Arts and Humanities – Architecture (buildings) – Computer-Aided Design

1. INTRODUCTION

Additive manufacturing (AM) is currently being investigated for application in the building industry, suggesting a revolution in the way products are designed, manufactured and distributed to the end users (Gao et al. 2015). The versatile form of fabrication enhances the possibility of integrating performance-driven features in the design, without having to compromise on optimality due to manufacturing limitations. In this prospective, the use of digital tools and simulations is needed to deal with the increased complexity of non-standard components. When compared to traditional manufacturing techniques, AM shows some unique capabilities that suggest a potential application for the built environment. Geometrical complexity can be easily achieved so that optimised and customised shapes are possible. Material complexity is also made possible by manipulating geometry at meso-scale so that materials can be designed to obtain different properties where needed, as presented by Ashby (2006).

The building envelope is in itself one of the most complex parts of the building. Recently, the challenging objectives for energy performance set by new regulations, such as the European EPBD, have redefined the role of façades in the overall building concept, with stringent requirements in terms of materials and performance, especially thermal control. From the scientific community and industries, existing research on AM for the building envelope has proved to be promising for the design of multi-functional façade components with complex geometries for climate control, as presented by Sarakinioti et al. (2018). Among the key requirements of the building envelope, thermal performance is a most promising aspect to be investigated as it can be controlled by the geometry of the component and contribute to savings in terms of material and energy consumption. In particular, thermal insulation is a property which is strongly related to the geometry of the problem and different thermal conductivity values can be achieved with the same material by changing its meso-structure as presented by Gooskens (2016).

This study proposes a novel digital workflow, based on the results of FE simulations, analytical models and physical testing, to design thermally insulating cellular structures and assess their performance. In this context, a research-and-design workflow is proposed integrating digital and physical tools. The workflow is used to design a mono-material façade panel to be additively manufactured, which serves as a filter between the interior and exterior environment. The chosen fabrication method is Fused Deposition Modelling which is a relatively low-cost and

accessible AM technique based on material extrusion (Wu et al. 2016). Among the materials suitable for this technique, polymers were chosen for their low thermal conductivity and the possibility of being printed at different resolutions (Ngo et al. 2018).
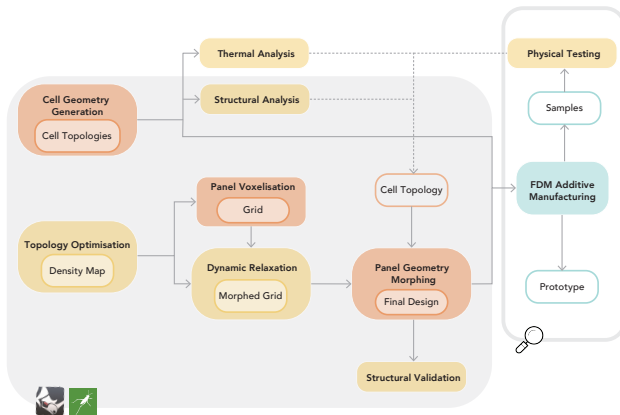

Figure 1: Digital and physical workflow defined for the study, consisting of parametric modelling in the environment of Rhino+Grasshopper, prototyping and physical testing

2. METHOD

The research focused on identifying a method to assess the thermal performance enhanced by the complex geometrical configurations and integrate it into a digital design workflow. The workflow was used to propose the design of a modular façade component, serving as a filter between the interior and exterior environment. The chosen fabrication method is Fused Deposition Modelling which is a relatively low-cost and accessible AM technique, based on material extrusion (Wu et al. 2016). This has proved to be suitable for the production of large scale architectural components in several realised projects, such as the Canal House designed by Dus Architects in 2013 (3DPrint Canal House n.d.). Among the materials suitable for this technique, polymers were chosen for further exploration for the low thermal conductivity, the possibility of being printed at different resolutions and the suitability for the FDM process (Ngo et al. 2018). Considering recyclability, non-toxicity and its low cost, PETG was identified as the material to investigate, combining good thermal properties and low density.

For this study, a combination of research through design and performance-driven design was adopted as methodology. This involved retrieving performance indicators using analytical, numerical and experimental methods and comparing results to verify the reliability of the proposed methods. According to the guidelines given by the results of the validated models, geometrical alternatives were first generated, then assessed and developed into a façade component. Firstly, background studies were used to define the boundary conditions for the design in terms of component scale, production process, material, morphologies and target performances. Then, different geometrical alternatives were explored and their performance was assessed using analytical models, physical

testing and software simulations. The different options were compared and assessed according to their thermal insulation (effective thermal conductivity) and manufacturability with AM techniques. Finally, the chosen cellular geometry was used to design the façade component using topology optimisation principles.

The use of digital tools was essential for the generation of design alternatives and integration of performance-driven guidelines. FE simulations on Comsol Multiphysics were used to study the thermal behaviour in complex geometrical structures which analytical models are not able to describe accurately. To combine these aspects, a digital workflow was defined, in the parametric environment of Rhinoceros 3D and Grasshopper, which includes geometry generation and morphing procedures as well as performance assessment, with a direct connection to the additive manufacturing production process.

3. GEOMETRY EXPLORATION OF CELLULAR STRUCTURES

Geometrical complexity at meso-scale was addressed by studying the properties of cellular solids according to the models proposed by Ashby (2006). On a first categorisation cellular solids can be described as either strut-based geometries, such as lattices, or surface-based geometries, such as foams and shells. Among shells, triply periodic minimal surfaces are known to have interesting thermal insulating properties due to their low surface-to-volume ratio (Maskery et al. 2018). Lattices can be mathematically defined as arrays of points generated by a set of translations in a 3-dimensional space, while shells can be described by an implicit surface equation.

$$r = k \cdot a_1 + l \cdot a_2 + m \cdot a_3 \qquad f(x,y,z) <= t$$
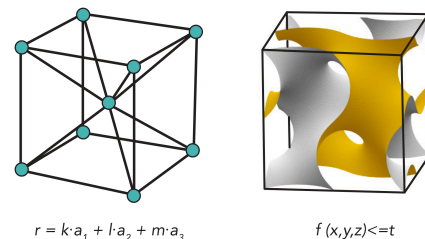
Figure 2: Examples of cellular geometries: lattice and shell

Using their geometrical definition, nine different cell topologies were modelled in the parametric environment of Rhino-Grasshopper and analysed (Figure 3). These typologies were selected according to performance and manufacturability guidelines. In relation to thermal insulation, geometries which take advantage of the low conductivity of enclosed air cavities were chosen along with porous structures which are able to hinder the path of heat transfer. Finally, considering production with additive manufacturing, self-supporting geometries were preferred in order to minimise the use of support material. Moreover, to minimise printing time, topologies which allow for a continuous path of the extruder during printing were selected. According to the findings of Oarakinioti et al. (2018), an elongated cell size was designed with a minimum dimension of 3 cm in the direction of the heat

flow. This was done to minimise convective heat flow within the geometry, while the other dimensions were kept around 15 cm.
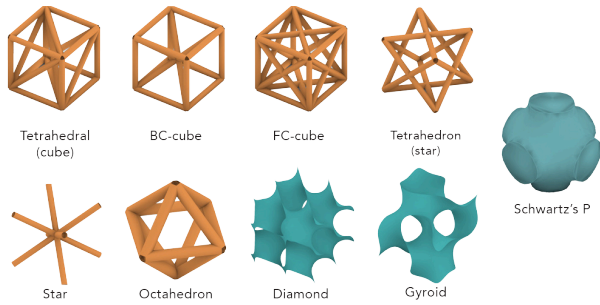
Figure 3: Selected cell topologies for design exploration based on performance and manufacturability parameters

3.1. Analytical Models for Thermal Performance

At first, different analytical models were used to calculate the effective thermal conductivity of cellular structures. These expressions are mainly dependent on the density of the geometry and cannot accurately account for the effect of topology. They were useful to make a preliminary comparison between the different options. The analytical model proposed by Hegman and Babscan (2018) was used to assess the different geometries. This model analyses the problem in terms of two limiting cases, the parallel and series arrangement of the solid and gaseous constituents, and was defined for application to any cellular geometry. It considers only conductive heat transfer mode and was used to retrieve an effective thermal conductivity value. From this model, an effective thermal conductivity for each cell topology was retrieved. The results of the analytical calculations show that thermal performance increases with relative density, with nearly linear relationship (Figure 4). These results are compared with those of simulations and physical tests to assess how geometry influences the two performances in more detail.
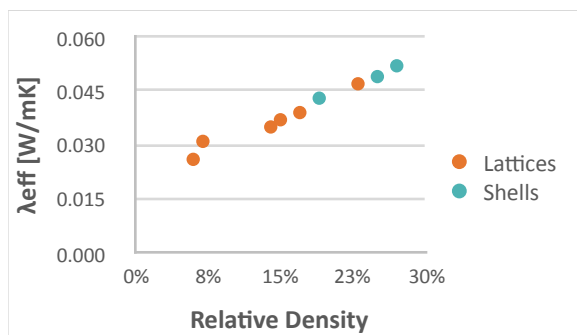

Figure 4: Relation between porosity and thermal insulation

3.2. Physical Testing on 3D-printed Samples

Geometry samples were produced using Fused Deposition Modelling technique with a LeapFrog desktop printer and PETG filament. The size of the samples was defined according to the machine built volume and to the set-up of the heat flux physical test. Geometry samples of 25 x 17 x 1.7 cm were modelled in Grasshopper, corresponding to a 3

x 2 cell array. The chosen cell size is 8.3 x 8.5 x 1.7 cm, with the smallest dimension in the direction of the heat transfer. Shell geometries were found to be particularly suitable for manufacturing with FDM (Figure 5-6). The printing of lattice structures proved to be more challenging due to the horizontal slicing, which results in steep overhanging geometries that require support material. To efficiently produced such geometries power bed fusion techniques could be used or a robotic arm coupled with an extruder for multi-plan printing.

Physical tests were carried out according to the standardised heat flow meter testing procedure (Meng et al. 2015) to retrieve the effective thermal conductivity of the printed samples. Because the surface of the sample was not flat, a VIVAK (PETG) sheets had to be added to make sure the sensors had proper contact and to close off possible air gaps that were present in the geometry sample.

The samples of size 25 x 17 x 1.7 cm were put within a hole in one of the faces of a polystyrene box of 1 m3 volume. Inside the box, a shielded light bulb was switched on, heating up the box until around 46°C and creating a temperature difference with the external environment at room temperature. After thermal equilibrium condition was reached, where net flow of thermal energy between the two systems is constant, steady-state measurements were performed by measuring the surface temperature of both sides of the sample and of the air on both sides with type T thermocouples and by measuring the heat flux on both sides of the sample using two Hukseflux HFP01 heat flux plates. Since the heat flux sensor can only measure a circular area of 32 mm diameter, the measurements were done twice per sample, as in Figure 7: first with the sensor pointed at a solid portion of the sample and then with the sensor pointed at a void part of the sample.

Figure 5- 6: 3D printed samples of Gyroid cell array (top) and Schwartz's P cell array (bottom). Dimensions: a= 25 cm, b=17 c= 4 cm, thickness=1.7 cm

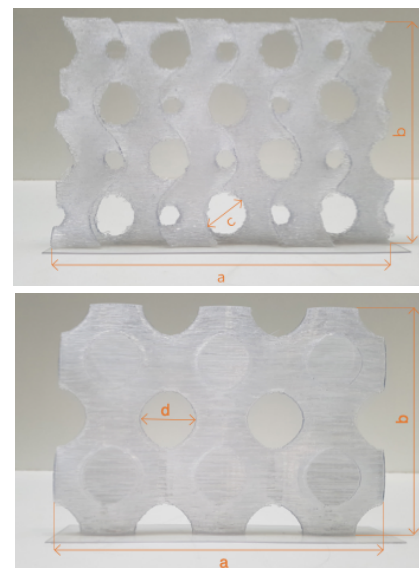| Geometry	Ts_{in} [°C]	Ts_{out} [°C]	Heat Flux (+) [μV]	Heat Flux (-) [μV]	Sensitivity (+) [V/(W/m²)]	Sensitivity (-) [V/(W/m²)]	Q [W/m²]	R_{tot} [m²K/W]	λ_{eff} AV [W/mK]
#1 Diamond	46,4	26,1	3,92	2,98	6 2,23	61,22	55,8	0,364	0,063
#2 Diamond	46,3	28,3	4,58	4,1	62,23	61,22	70,3	0,256	
#1 Gyroid	46,7	28,8	3,64	3,84	62,23	61,22	60,6	0,295	0,071
#2 Gyroid	46,7	28,2	5,58	3,76	62,23	61,22	75,5	0,245	
#1 Schwartz's P	47,8	27,7	3,94	4,04	62,23	61,22	64,7	0,311	0,069
#2 Schwartz's P	50,8	28,9	5,82	5,14	62,23	61,22	88,7	0,247	

Figure 7: Results of heat transfer tests on 3D printed samples

The measurement instruments were connected to a data logging system (Eltek Squirrel 1000 series), transferring the data to the Darca Software. From these heat flux and temperature data (Figure 7), the thermal resistance and thereby also the effective thermal conductivity was calculated according to the following relation:

$$\lambda_{eff} = \frac{d \bullet q}{(T_{s1} - T_{s2})}$$

where T_{s1} and T_{s2} are the temperatures measured at the sample's exterior sides, d is the thickness of the sample and q is the measured heat flow.

According to the results (Figure 7), the three geometries exhibit similar thermal conductivity in the range of 0.07 W/(m·K), which corresponds to a decrease of 65% compared to the thermal conductivity of the solid PETG (λ=0,19 W/(m·K)).

3.3. Software Simulations for Thermal Performance

As a next step, FE simulations were performed using Comsol Multiphysics version 5.4 to gain insight into the thermal behaviour of the different geometries, with three main objectives: compare and verify the results from the physical tests, define a reliable model to test the thermal conductivity of different geometries and retrieve the effective thermal conductivity of lattice geometries that could not be tested. The simulations were performed in different steps, in order to evaluate the relations between geometry and the different heat transfer modes first separately and then combined. Moreover, varying parameters such as cell size, cell topology and cell array were introduced to find the best geometrical combination (Figure 8).

FE modelling of heat transfer through cellular structures proved to be expensive from a computational point of view, due to the complexity of the geometries involved, particularly for shell geometries. The final simulation model took into account conductive heat transfer directly while convection and radiation are accounted for using the simplified methodology of an effective thermal conductivity of air as proposed in NEN-EN-ISO 6946 (Nederlands Normalisatie Instituut, 2008). This model was validated by comparing the results of the simulations with those of the physical tests. For the simulations 'the heat transfer in solids' module was used in combination with steady-state simulations. Two convective heat fluxes through non-solid material were defined at the interior and exterior sides of the geometry, specifying the heat transfer coefficients according to common calculation methods: 7.8 W/m²·K for the interior flux and 25 W/m²·K for the exterior. The sides of the panel were assigned an insulation boundary to represent adiabatic conditions and the front and back of the panel were assigned a fixed temperature of 20°C and 0°C respectively. Furthermore, as thermal conductivity of the PETG filament, 0.19 W/(m·K), was used; the effective thermal conductivity of air according to the aforementioned standard varied depending on the exact dimensions of the cavities (Jóhannesson, 2006).
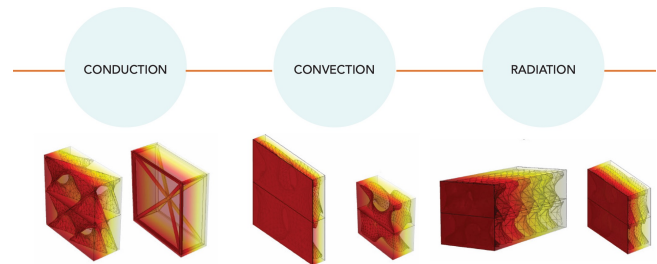
Figure 8: Varying parameters for heat transfer simulations in COMSOL Multiphysics

The simulation results are compared to those of the physical tests in Figure 9. A difference of 10% and 22% is found for the Gyroid and Diamond geometry respectively. This difference is small enough to consider the COMSOL model as a reliable one. For the Schwartz's P geometry, however, this difference increases to 40%. This could be explained by the fact that, contrary to the other two, this geometry was printed with a smaller shell thickness (1 mm instead of 2 mm) as part of the exploration of the manufacturing process. The resulting thermal conductivity from the physical tests would be higher in case the sample was

printed with a thicker shell, thus better aligning with the simulations results. Overall, the results of the simulations prove the reliability of the COMSOL model which is refined enough to approximate the physical heat transfer in cellular structures. The model also allows for a more accurate comparison of the different geometries, as possible inaccuracies in the model would affect the different geometries in the same way. On the contrary, inaccuracy of the measurements of the physical tests can differ from case to case in an unpredictable way.

Moreover, through the different steps of the simulation and testing process (Figure 10), some conclusions could be drawn regarding the thermal behaviour of cellular structures. When small air cavities are involved (up to 2 cm), heat transfer by convection is minimal and can be neglected. Both conduction and radiation contribute to the global heat transfer in cellular geometry, as we can see comparing the results of the simulations where only conduction is considered and those of the simulations where radiation is also considered (Figure 9). Overall, ratio of void-to-solid is the driving factor for thermal insulation. For this reason, lattices proved to be more efficient in thermal insulation.

According to the final simulation model, the average thermal conductivity of such geometries is 0.07 W/(m·K) which is lower than the average for shell geometries, 0,09 W/(m·K). From a geometrical point of view, the thermal resistance of lattices depends greatly on the porosity of the cell structure. The influence of porosity is less predictable in shell structures, where Schwartz's P geometry resulted to have higher thermal conductivity, in spite of its low relative density. In general, differences in results between the geometries are not very significant. Therefore, the choice of the best geometry to be explored in the façade panel design was done considering manufacturing aspects as well.
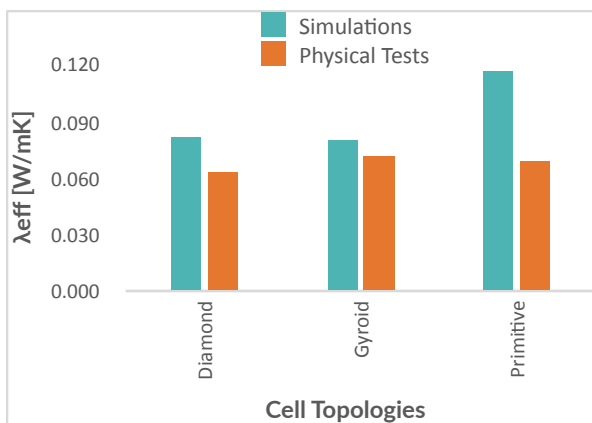
Figure 9: Comparison of simulations and physical tests results
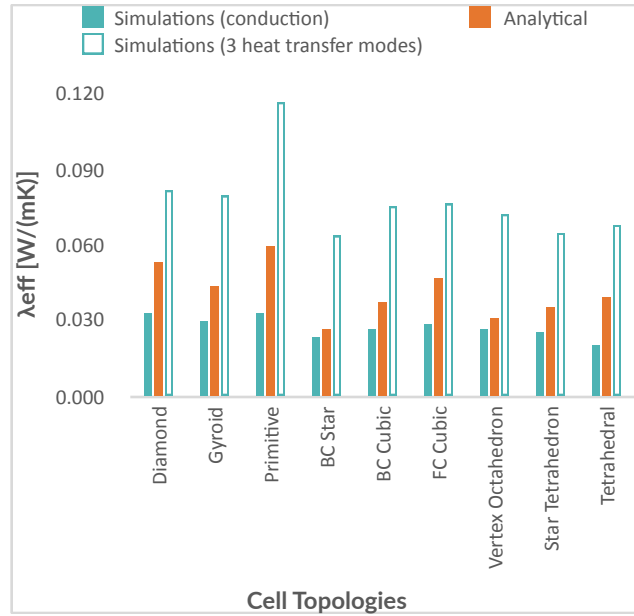
Figure 10: Comparison of results from simulations, analytical models and physical tests for different cell topologies

4. DIGITAL WORKFLOW

4.1. From simulations to performance-driven design
The study of heat transfer in cellular geometries is a complex one, consisting of radiation, convection and conduction. Analytical models are not able to account for convective and radiative heat transfer while finite element analysis is a powerful tool to estimate the thermal insulation property of cellular geometries. However, the use of a stand-alone software, outside the parametric environment of Rhino-grasshopper, resulted in the impossibility of integrating the results directly in the design process. To overcome this, the results of the software simulations and analytical models were compared and an analytical expression was proposed to retrieve an effective thermal conductivity for cellular structures. In particular, two different expressions are proposed for lattice and shell topologies respectively. These expressions are derived from the analytical models presented in Section 3.1 and are adjusted in order to account for radiative and convective heat transfer. This is done following the methodology of NEN-EN-ISO 6946, which consists in assigning a new thermal conductivity value to air. In particular, for lattice structures the model proposed by Ashby (2006) is used, which assumes that, in such geometries, 1/3 of the struts of the cellular structure lie parallel to one Cartesian axis. The expression, including the adjusted thermal conductivity value for air (λg), is:

$$\lambda_{cell} = \frac{2}{3}\left(\frac{\rho_{cell}}{\rho_s}\right) \cdot \lambda_s + \lambda_g$$

where:

- λs is the thermal conductivity of the bulk material [W/mK];

331

- λg is the adjusted thermal conductivity of air [W/mK]; - Vs is the volume of the bulk material [m³];

- Vg is the volume of air [m³];

For shell topologies, the model developed by Leach (1999) is used instead, which unifies the results of a number of approaches to calculate the conductivity of cellular structures. The expression, including the adjusted thermal conductivity value for air (λg), is:

$$\lambda_{cell} = \frac{1}{3}\left(\frac{\rho_{cell}}{\rho_s}\right) \cdot \lambda_s + \left[1 - \left(\frac{\rho_{cell}}{\rho_s}\right)\right]\lambda_g$$

where:

- λs is the thermal conductivity of the bulk material [W/mK];

- λg is the adjusted thermal conductivity of air [W/mK];

- ρs is the density of the bulk material [kg/m3];

- ρcell is the density of air [kg/m3];

The two expressions have been compared to the results of the software simulations and proved to be relatively accurate. In particular, as presented in Figure 11, for lattice topologies a mean approximation error of 10% was found while for shell topologies this accounted for 16%. By looking at the results in Figure 12, it can be noticed that the analytical results do not fluctuate much, giving very similar results for different topologies.

Another approach to retrieve an appropriate analytical expression for thermal conductivity is regression analysis. In particular, the relationship between effective thermal conductivity and cell porosity was estimated by the linear function:

$$y = -0.1347x + 0.1888$$

where:

- x is the porosity of the cell topology;

- the two numeric parameters represent the thermal properties of the solid and gas components.

As can be observed in Figure 11, this relation is accurate for all cell topologies except for the Schwartz's P geometry. A possible explanation for this is the fact that this particular shell topology is the only one featuring closed air bubbles, which may require further refinement of the model created in COMSOL Multiphysics. For the same reason, a considerable discrepancy between the results of the physical test (0,069 W/mK) and those of the simulations (0,116 W/mK) was found for the same cell topology.

Nevertheless, the two analytical expressions described above were found to be useful and accurate enough to be integrated and used within the digital workflow as a preliminary assessment tool for thermal performance. Figure 12 shows how the two expression have been integrated in the workflow. The user is required to input the

material and geometrical properties and of the cellular structure (cell and cavity sizes), while relative density and porosity of the geometry come as output of the geometry generation part.
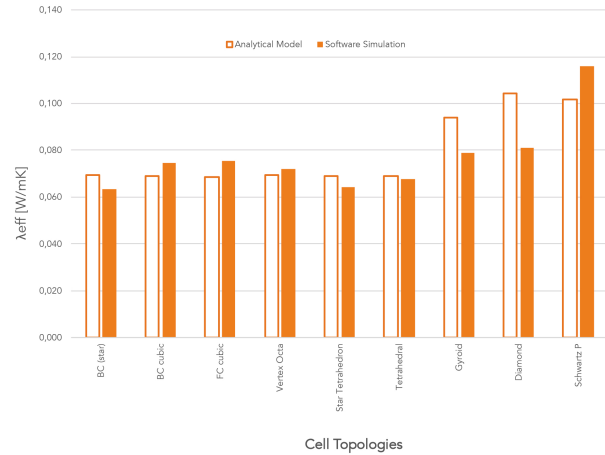
Figure 11: Comparison of simulations and physical tests results
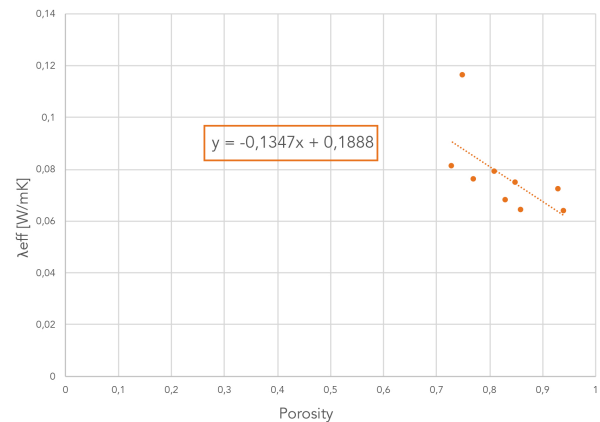
Figure 12: Comparison of simulations and physical tests results

4.2. Design work-flow

As outcome of this study a digital work-flow was defined for designing façade elements with complex geometries for thermal and structural performance. This can be used as a design tool through which the form of envelope components can be generated, according to loads acting on the façade and desired cell geometry. The geometry generated in the parametric model was analysed with FEA using a stand-alone software, COMSOL Multi-physics. In the last phase of the study, the results of the simulations in COMSOL were analysed and compared to the analytical models in order to find a simple analytical expression for the calculation of effective thermal conductivity which could be integrated within the parametric model. The same model was also used to generate sample geometries for additive manufacturing which were then tested for the thermal performance. At the cell scale, the script integrates procedures for generation and performance assessment of both shell and lattice geometries. At the component scale,

for ease of use in the design process, the script is designed specifically for shell geometries to be produced with FDM. However, this could be adapted to lattice geometries with minimal changes in the geometry morphing phase.

5. FAÇADE COMPONENT

The digital workflow was used to design a façade element of 3 x 1.5, in which the chosen cellular geometry is morphed to provide additional stiffness for the most stressed parts of the panel, thus enhancing its structural performance. In the less stressed areas, porosity increases improving thermal insulation. The geometry varies also within the panel thickness according to address the required performance using the same material (Figure 13). According to thermal analysis performed on the element, the component is able to comply with the current regulations regarding maximum deflection and Rc-value, making this a promising direction for further development of a marketable façade component.
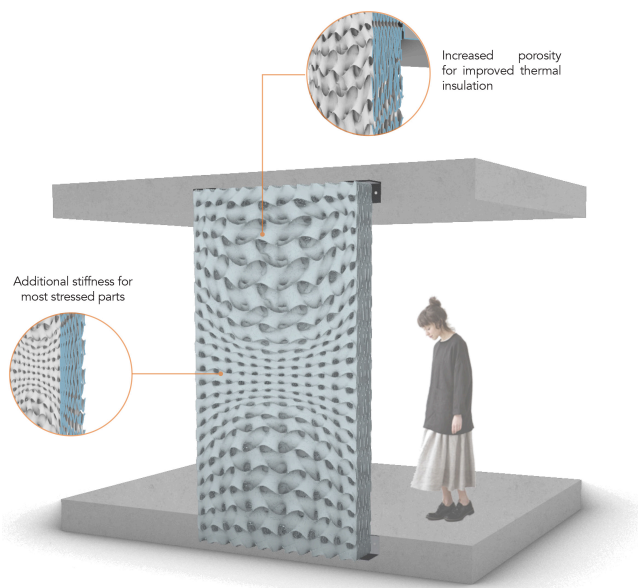
Figure 13: The AM envelope: results of the performance-driven design

5.1. Refinement of cell size

In order to optimise the overall design, the size of each cellular structure was increased to 25 x 25 x 3.4 mm, achieving a more porous yet stiff geometry and, therefore, a lower weight. The relation between cell size and relative density is shown in Figure 14 and depends on the implicit surface equation defined for the gyroid cell:

$$cos(x) \bullet sin(y)+cos(y) \bullet sin(z)+cos(z) \bullet sin(x) = 0$$

Increasing the cell dimension reduces the number of cells needed to fill the panel volume. Moreover, since porosity slightly increases along with the cell size, the weight of the panel can be reduced by more than 8%, reaching 101 kg. Increasing the cell size is also beneficial from a production point of view as printing time is decreased. Considering the

thermal performance, simulations showed that increasing the cell size is beneficial as the relative density decreases. In fact, the thermal performance of the panel benefits from the lower thermal conductivity of air.
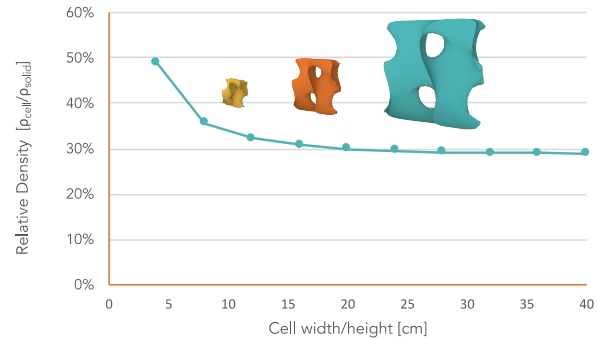
Figure 14: Relation between cell size and porosity in Gyroid cells

To verify the assumption stated above, new simulations were carried out on the bigger cell size geometry in COMSOL. For the simulation, the same set-up described in Section 3.3 was used. An effective thermal conductivity for air was assigned to the material to account for radiative and convective heat transfer, according to NEN-EN-ISO 6946. As expected, the thermal conductivity of the sample slightly decreases, reaching 0.065 W/mK. In order to achieve the prescribed Rc-value of 4.5 m^2K/W the façade panel is required a thickness of 29 cm.

6. CONCLUSIONS

This study aimed at investigating the thermal performance of cellular structures to be implemented into a multi-functional mono-material façade panel produced by additive manufacturing. The research focused on exploring the relation between geometry and performance at different design scales, making use of digital tools, prototyping and physical testing to design non-standard spatially varying geometries. According to the results, a methodology was proposed and embedded into a digital design workflow in Rhino-Grasshopper through which the panel geometry can be generated.

The results of the study showed that cellular geometries can be designed to achieve low thermal conductivity values by controlling the surface-to-volume ratio. Porosity is beneficial for decreasing thermal conductivity. However, heat transfer in cellular geometries also depends on convection and radiation. To minimise convection, air cavities should be kept within small widths and, therefore, elongated cells with small dimension in the direction of heat transfer were designed. Radiation accounts for a large part of the global heat transfer and shell geometries benefit from the shielding properties of the surfaces composing the cell. To minimise material use and cost, research towards further improvement of the thermal insulation could be carried out. Polymers with extremely low thermal conductivity and emissivity could achieve lower thermal conductivity values. The thermal conductivity of the gaseous component could be decreased if air is substituted with a noble gas such as argon and krypton.

The feasibility of the fabrication with AM was demonstrated through different production tests and a 1:1 prototype (Figure 15). This was done using an industrial 3D printer, prototyped by Leapfrog, featuring a massive build volume (55 x 50 x 250 cm) and equipped with a 1.2 mm nozzle extruder fed with PETG filament. The nature of the geometry makes it self-supporting and stiff so that no infill or supporting material is needed. The designed geometry proved to be suitable for large scale additive manufacturing, encouraging application of such geometries to the design of insulating façade components. Improving production efficiency is still required to scale up from prototype to building component. Further research and investigation are needed to define the façade concept as a marketable envelope component. Fast-paced advancements in the field of polymer large scale additive manufacturing are promoting novel applications for the building industry. The optimisation of production processes and tools are expected to lead to more competitive production costs in the short term (Dimensional Research, 2020). Moreover, material costs are also expected to decrease as more options enter the market.
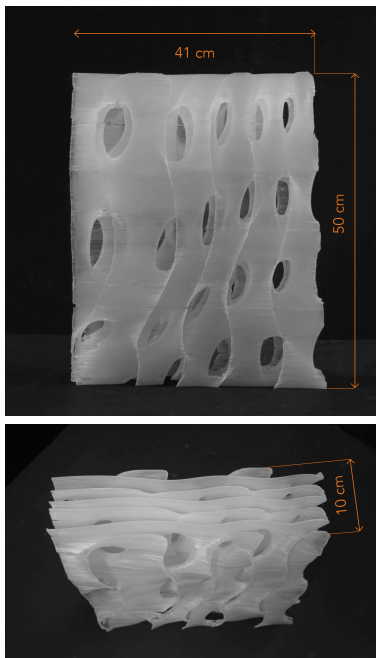
Figure 15: 1:1 scale prototype produced with FDM using PETG to test the feasibility of production with AM

8.   REFERENCES

1. Ashby, M. F. *The properties of foams and lattices.* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 364 (2006), 15-30.

2. Dimensional Research. (2020). *Trends in 3D printing at scale:a survey of 3D printing stakeholders in production manufacturing* [Data set]. Retrieved from: https://cdn2.hubspot.net/hubfs/5442420/Content/Trends-in-3D-Printing-at-Scale.pdf

3. Gao, W., Zhang Y., Ramanujan D., Ramani K., Chen Y., Williams C. B., Wang C. C.L., Shin Y. C., Zhang S. & Zavattieri P. D. *The status, challenges, and future of additive manufacturing in engineering.* Computer-Aided Design, 69 (2015), 65-89.

4. Gooskens, G. 2016. *Printing properties: the potential of functionally graded cellular solids in the building envelope by integration of structural and thermal performance using the capabilities of additive manufacturing.* Repository Delft University of Technology (2016).

5. Hegman, N., Babscan, N. *Specific Feature of Thermal and Electrical Transport in Cellular Media* (2018).

6. Jóhannesson, G. Lectures on Building Physics: heat and moisture transfer. *Lecture notes, Building Physics AF1402.* KTH Royal Institute of Technology (2006).

7. Leach, A. G. Thermal conductivity of foams. I. *Models for heat conduction.* Journal of Physics D: Applied Physics. doi: 26. 733. 10.1088/0022-3727/26/5/003. (1999).

8. Meng, X., Gao, Y., Wang, Y.,Yan, B., Zhang, W. & Long, E. *Feasibility experiment on the simple hot box-heat flow meter method and the optimization based on simulation reproduction.* Applied Thermal Engineering, 48-56 (2015).

9. Nederlands Normalisatie Instituut. *Building components and building elements - Thermal resistance and thermal transmittance - Calculation method.* 2008 NEN-EN-ISO 6946. (2008).

10. Ngo, T. D., Kashani, A., Imbalzano, G., Nguyen, K. T.Q. & Hui, D. 2018. *Additive manufacturing (3D printing): A review of materials, methods, applications and challenges.* Composites Part B: Engineering. 143, 172-196..

11. Sarakinioti, M. V., Turrin, M., Konstantinou, T., Tenpierik, M. and Knaack, U. *Developing an integrated 3D-printed façade with complex geometries for active temperature control.* Materials Today Communications (15): 275-279 (2018).

12. Wu, P., Wang, J. & Wang, X. *A critical review of the use of 3-D printing in the construction industry.* Automation in Construction, 68, 21-31 (2016).

13. *About the 3D printed Canal House: Construction Technique.* Accessed September 15th 2018. https://3dprintcanalhouse.com/

Design Exploration

Floorplan Embedding with Latent Semantics and Human Behavior Annotations

Vahid Azizi[1], Muhammad Usman[2], Samarth Patel[1], Davide Schaumann[3], Honglu Zhou[1],
Petros Faloutsos[24] and Mubbasir Kapadia[1]

[1]Rutgers University, New Jersey, USA, {vahid.azizi; hz289; mk1353}@cs.rutgers.edu
[2]York University, Toronto, Canada, {usman; pfal}@eecs.yorku.ca
[3]Jacobs Technion-Cornell Institute, New York, USA, davide.schaumann@cornell.edu
[4]UHN - Toronto Rehabilitation Institute, Toronto, Canada

ABSTRACT

Floorplans provide top-view representations of buildings that highlight key relationships between spaces and building components. In the last few decades, different approaches have been proposed to compare and catalogue different floorplans for design exploration purposes. Some approaches have considered floorplans as images, while others represented them as graphs. However, both image and graph-based approaches have failed to extract and utilize essential low-level space semantics and structural features. Further, they do not encode information about space utilization determined by people movement and activities in space, which are critical to analyze a building layout. To address these issues, we use deep learning techniques to develop a floorplan embedding – a latent representations of floorplans, which encodes multiple features. Specifically, we propose a novel framework that uses an attributed graph as an intermediate representation to encode space semantics, structural information and crowd behavioral features. We train Long Short-Term Memory (LSTM) autoencoders to represent these graphs as vectors in a continuous space. In addition, we contribute a floorplan dataset augmented with semantic and simulation-generated behavioral features. These representations spark new opportunities for next-gen design applications like clustering, design exploration tools and recommendations. Three different use cases are studied to show the performance of this method.

Author Keywords

Floorplan Embedding; Attributed Graph; Design Exploration; Design Semantic Features; Human Behavioral Features; LSTM Autoencoder

ACM Classification Keywords

G.3 Mathematics of Computing: PROBABILITY AND STATISTICS; I.6.3 SIMULATION AND MODELING: Applications; I.6.5 SIMULATION AND MODELING: Model Validation and Analysis; J.5 ARTS AND HUMANITIES: Architecture; J.6 COMPUTER-AIDED ENGINEERING: .

1 INTRODUCTION

Floorplans provide a well-established mean to represent buildings. As such, they afford a wide range of design activities such as ideation, analysis, evaluation and communication. Computer-Aided Design (CAD) and Building Information Modeling (BIM) approaches support the creation of digital building models, from which floorplans can be extracted. Current approaches, however, do not support the systematic comparison of floorplan features derived from geometric and semantic properties as well as more advanced performance metrics, such as space utilization and occupant behavior generated via simulation.

While research mostly developed in Computer Graphics proposed computational strategies to extract floorplan features by processing them as images [18, 7] or graphs [23, 21], these approaches ignore semantic information that describes the function of each space, as well as time-based analytic of human movement and activities. These high level features are required but they are not enough to capture information about semantic features and the way people behave inside floorplans.

To address this issue, we propose *floorplan embedding* – "latent" representations of building layouts using an attributed graph as intermediate representation that encode not only design and structural features, but also human behavior features. Specifically, a Long Short-Term Memory (LSTM) [14] autoencoder [13] is trained to represent these graphs as vectors in a continuous space. These vector representations are used to cluster and query floorplans with similar characteristics and attributes. Some of the advantages of presenting floorplans as vectors are: (a) compact representation, (b) efficient way to compare designs and fast retrievals, (c) scoring designs and providing feedback/recommendations, and (d) categorizing design according to our target features. In addition, since there is not any dataset augmented with semantics and human behavioral features, we release a novel annotated floorplan dataset with semantics and human behavioral features generated from simulations.

The key contributions in this paper can be summarized as follows: (i) intermediate representation of floorplans as attributed graphs and augmented with crowd behavioral features (ii) novel unsupervised deep learning model to learn a meaningful vector representation of floorplans (iii) creation of a floorplan dataset augmented with semantic and crowd behavioral attributes generated from simulations.

2 RELATED WORK

We aim to empower floorplan representation as continues vectors with encoding the design semantics attributes, low-level structural characteristics and also crowd behavioral attributes. In this section, we will review some of recent works for comparing floorplans and the way they are represented. Prior works can be mainly divided into three categories: image-based, graph-based and symbol-spotting methods.

Several approaches to compare floorplans have been proposed based on conventional image processing techniques [4, 11]. In these approaches features like Histogram of Oriented Gradients (HOG) [5], Bag of Features (BOF) [16], Local Binary Pattern (LBP) [2] and Run-Length Histogram [6] have been used for extracting features from floorplan images. The extracted features are used for comparing floorplans. In [24], a deep Convolutional Neural Network (CNN) is proposed to present floorplans as vectors to address the limitation of conventional image processing techniques for extracting meaningful features. These methods are based on object-centric databases in which floorplans are annotated with furniture or specific visual symbols, and only the visible features from images are encoded. However, these features are not semantics and these methods are not able to capture the design structures.

Graph matching is another way for comparing floorplans. In these methods floorplans are represented as graph and graph matching methods are applied for scoring similarity. Different strategy is proposed for presenting floorplans as graph. In [23], floorplans are modeled as graphs to capture adjacency between the rooms and augmented with arrangement of annotated furniture. In [21], the floorplans are converted to graphs by considering the connectivity, room type and accessibility in graph construction. In [1], a framework for automatic analysis of floorplans is proposed in which floorplans are represented by an attributed graph where attributes for each room are extracted using SURF technique. In [22], three different representation layers are merged to form a graph representation. These representations are area, furniture style, and an adjacency matrix. Since in all of these methods floorplans are represented with graph, the buildings structure like connection of rooms are captured. Some of them add some features to nodes which are mostly extracted features by image processing techniques. In addition all of these methods use graph representation as final representation and do not generate vector representation. Then graph matching methods are used for finding similarity and other type of operations are applied directly over graphs. Since it might be the case that there is no exact matching, subgraph matching [1, 28] is proposed. It means if two graphs are not isomorphic but their parts are, they can be detected with subgraph matching.

Symbol spotting is another mechanism which has been used for floorplans comparison and retrieval. It is a special case of Content-based Image Retrieval (CBIR) [12, 20] which is used for technical document analysis. In symbol spotting, a query is submitted and the system retrieves zones from the document which are likely containing the query. This query could be a cropped image or hand-sketched. Pattern recognition techniques are utilized for this aim, as example moment invariants like Zernike moments in [15]. Reducing search space in spotting symbols is proposed based on hashing of shape descriptors of graph paths(Hamiltonian paths) in [8]. SIFT/SURF [27] features that provide an efficient and scale-invariant are commonly used for spotting symbols in graphical documents. Symbol spotting methods are applied to small datasets which do not have complex images, and they are only applicable for retrieval purpose.

3 METHODOLOGY

The proposed framework is illustrated in Fig. 1. It consists of two components. The first component is for preprocessing floorplans and convert them into attributed graphs. The result of this step is a novel floorplan dataset augmented with design semantic and crowd behavioral features. The second component is the embedding for creating latent vectors of these graphs. These latent vectors represent floorplans as numerical vectors in a continues space. Further details on each component are presented in the following subsections.

3.1 Dataset

HouseExpo dataset [17] is used in this work. It includes about 35000 2D floorplan layouts that are presented in JavaScript Object Notation (JSON) format. They are mostly floorplan segments belong to big buildings. A sparse labeling for layout components is also given for each floorplan. Figure 2 shows an image of a raw floorplan compiled as image. There are about 25 different components types in the whole dataset. Some of them, however, have similar semantics (e.g. toilet and bathroom, terrace and balcony, etc.). We reduce the components types to 11, considering only single type-name for rooms labeled for similar purposes and removing minor components like freight elevator.

3.2 Floorplans to Attributed Graphs

In order to convert floorplans into attributed graphs, first we recognize all potential room boundaries by a series of image processing operations. Then, we assign labels based on provided annotations. Some of the floorplans are not annotated perfectly (e.g. the given coordinates for the labels are not match with room coordinates). In such cases, we calculate overlaps of each room with given label coordinates, and then assign labels to rooms which have the maximum overlap.

The rooms are the nodes in the graph and edges are the potential connections between them. There will be an edge between two rooms if there is an immediate door between them. Edges are formed by image processing techniques. To detect doors, we create a combined image of two rooms and apply blob detection on it. If the number of blobs found is one, it means there is a door between them, otherwise, there is no connection between the two rooms. Up to now, graphs are
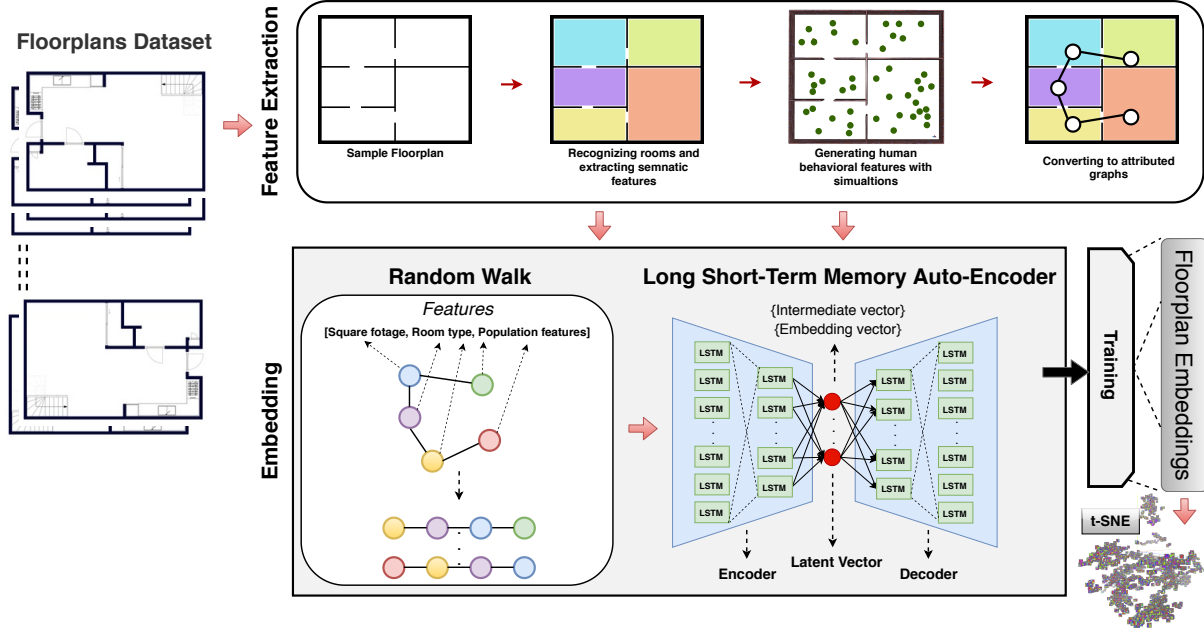
Figure 1: An overview of the proposed framework. In pre-processing, rooms and their design semantic properties are identified, then a 3D model is generated as input for crowd simulator to generate dynamic behavioral features, and at the end, the floorplan is converted to an attributed graph. In the embedding part, first a random walk is performed to convert attributed graphs to a set of sequences. Next, a LSTM autoencoder is proposed for training and floorplan predictions. The encoder has a 2 layer LSTM: first has 84 and second one has 64 LSTM units. The decoder has same architect contrariwise.

representing the structure of floorplans but they are not attributed. Next, we assign to nodes (rooms) their respective semantic features like square footage and room type, which are calculated using image processing techniques. To generate crowd related features, the 2D floorplans are converted to 3D models loadable in a crowd simulator, SteerSuits [25]. The simulator automatically populates virtual agents in each room with the target to exit the floorplan. It then calculates features like maximum, minimum and average evacuation times and traveled distances, as well as overall exit flow for all the agents in room. More details on these crowd aware features is presented in the following subsection. This part is done in an end to end procedure and does not need human interactions.

As mentioned, as a part of this work we are generating a novel floorplan dataset augmented with semantic and crowd behavioral features which is the result of this section. In this dataset, for each sample we have a JSON file augmented with semantic and crowd behavioral features.

3.3 Floorplan Dataset Features

This section presents the procedure for generating a floorplan dataset augmented with semantic and crowd behavioral features. At this point, we have floorplans which are presented as attributed graphs. The formation of a graph captures the structure of a floorplan, and attributes of nodes present its features. The available set of features can be seen in Table 1. These are divided into two groups: design semantic and crowd behavioral features which are generated with simulations. Design semantic features include room types

and square footage which are generated by image processing techniques. Since room types are categorical features, they are presented as one-hot vector with 11 dimension and footage square is represented with an one dimensional scalar. The total dimension of semantic features is 12. All crowd behavioral features are represented with a one dimensional scalar value. In total we have 9 crowd features which lead to have a 9 dimensional vector for crowd features. Sum of all feature dimensions is 21.

However, our approach is not bound to use only the selected features, and more can be adapted into the framework depending on the application. The values of scalar features have different ranges. In order to keep all of them within a same range for robust training, we normalize them between [0, 1] except for features which are presented with one-hot vector. All the outliers are removed before normalization.

3.4 Floorplan Embedding

In this section we discuss the conversion of attributed graphs into continues latent vectors which encode both structure and semantics of the floorplans, as well as their crowd behavioral features. We use graph embedding approach to transform graph nodes, edges, and their features into a vector space (lower dimensional) while preventing any information loss. Graphs, however, are tricky to deal with because they can vary in terms of their scale, specificity, and subject [9].

There exist some approaches to perform graph embedding [10, 26, 19]. However, they are only suitable for unattributed graphs and mostly capture just the graph struc-

Feature Class	Features types	Dimension
Design semantic	Square footage	1
	Room type	11
Behavioral	Not completed agents	1
	Max evacuation time	1
	Min evacuation time	1
	Exit flow rate	1
	Completed agents	1
	Max traveled distance	1
	Ave evacuation time	1
	Avg traveled distance	1
	Min traveled distance	1

Table 1: First column shows the two feature classes, second column shows features available in each class and third columns is their dimension. Dimension for semantics features is 12 and for behavioral is 9, in total the features dimension is 21, all scaled between [0 1].

ture. But in our case, the graphs are attributed. So in order to account for these attributes, we proposed an (Long Short Term Memory) LSTM autoencoder. Auto-encoders [13] are trained to learn the full properties of the data and reconstruct their inputs. They generally have two parts: an encoder that maps the input to an intermediate representation and a decoder that reconstructs the inputs from intermediate representation. The intermediate representations are latent vectors. LSTM is a recurrent neural network (RNN) for capturing long-distance dependencies in sequential data and also supports varying data lengths.

We propose a LSTM autoencoder to learn embedding space in a way that keeps floorplans of similar structure, design semantic and crowd behavioral features, close to each another in the embedding space. Our proposed model is illustrated in Fig. 1. We learn the function for mapping graph G to vector R in a d-dimensional space. In a floorplan (graph) $G = (V, E)$, V denotes its vertex set (rooms) and $E \subseteq V x V$ denotes its edge set (potential doors between rooms). We have unlabeled graphs in our dataset. Each node has a constant dimensional feature-vector F_V.

Adjacency matrix is one of the ways to present graphs as input to algorithms. But since the graphs are varying in number of nodes and edges, presenting them as adjacency matrix opens new challenges. This is because the dimension of the adjacent matrix is different for different graphs. To address this, we

(a)Raw   (b)Semantic Graph   (c)Behaviour
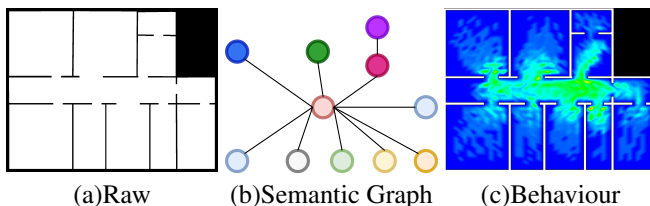
Figure 2: (a) A sample raw floorplan from the dataset. (a) The black pixels are wall (or outdoors) and white pixels are building components (or indoors). (b) These floorplans are converted to graphs and augmented with design semantic and behavioral features in our preprocessing step. (c) Shows the visualization of behavioral features on floorplan.

convert graphs to multiple varying length sequences. This conversion is done by using Random Walk. In a random walk we start from a given source node and the next node will be selected randomly with probability $1/D(N)$, where $D(N)$ is the degree of node N. Each graph is converted to a set of sequences and these sequences are feed to the model for training. Then the average latent vectors of corresponding sequences to a graph are used as a representation vector for the graphs (Equation 1).

$$\Phi(G) = \frac{1}{Nseq} \sum_{n=1}^{Nseq} RS_n \qquad (1)$$

The $Nseq$ is the number of sequences and RS is corresponding latent vector to each sequence.

3.5 Training
Some of malformed graphs are removed from the dataset. We use about 33000 floorplans' graphs which are converted to a set of sequences. We run random walks on the graphs to generate corresponding sequences. These sequences are feed into our model for training, both as input and output since the method is unsupervised. The loss function in our model is Mean Squared Error (MSE) (Equation 2).

$$loss(s) = \frac{1}{|s|} \sum_{i=1}^{|s|} \bar{Y} - Y \qquad (2)$$

Which s is the given sequence and $|s|$ is the number of nodes in the sequence. $\bar{Y}$ is the reconstructed vector for a node and Y is the true vector for that node. Loss function calculates the difference between reconstructed vector(output of decoder) and input sequences which each node has feature vectors F_v. In other words the encoder is trying to reconstruct the node features in the sequence.

4 CASE STUDY AND APPLICATIONS
This section validates and showcases the potential of our embedding methodology with the help of 3 different use cases. For these use cases, we trained three models with the architecture described. The difference between these models is the considered features. First model is trained only with design semantic features, second only with behavioral features and the third with all of the features.

4.1 Pairwise Similarity between Floorplans
In this use case we demonstrate a pairwise comparison between input floorplan and its second nearest neighbour from the embedding space. Underline graphs of both floorplans are compared using a popular graph similarity distance metric, Graph Edit Distance (GED) [3]. The edit distance between $G1$ and $G2$, GED(G1, G2), is the count of edit operations that transform $G1$ into $G2$, where the edit operations on a graph G can be an insertion or deletion of an edge or node. All edit operations have the same constant cost which is 1. If two graphs are identical, their GED is 0.

First, we retrieved a semantically similar design (second nearest neighbour) from the embedding space for the given input floorplan such that they have similar structural attributes but significantly different behavioral attributes, Figure 3a. The GDP value is reported as 0, showcasing that the graphs are
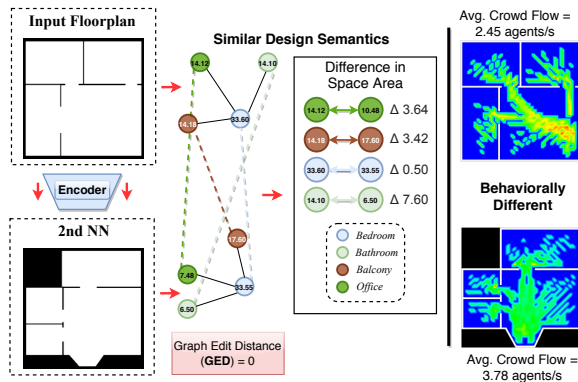
(a) Two floorplans which are structurally similar (Graph Edit Distance (GED) = 0) but have different crowd behavioral attributes.
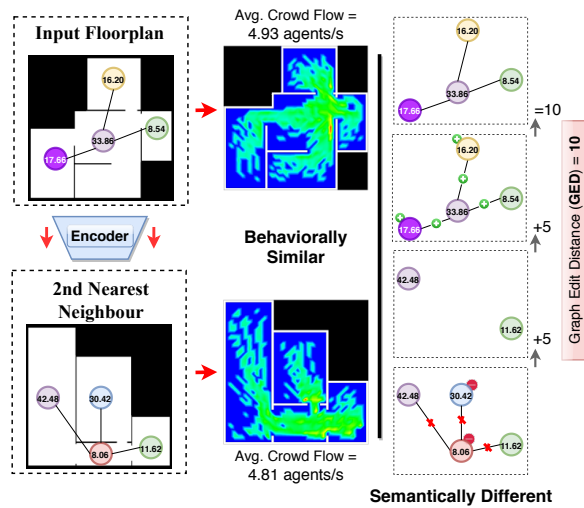
(b) Two floorplans which are behaviorally similar (having similar average exit flows) but design semantically different (Graph Edit Distance (GED) = 10).

Figure 3: Pairwise similarity comparison between input floorplan and queried nearest neighbour from embeddings.

similar in their design semantics. Average exit flow values and color-coded density heatmaps are also shown for both floorplans. Querying floorplan yielded comparatively lower exit flow than its nearest neighbour found from the embedding.

Second, we retrieved a behaviorally similar floorplan (second nearest neighbour) from the embedding space for the given floorplan such that they have similar behavioral attributes but significantly different design semantics. A GDP value of 10 is reported from the graphs comparison, showcasing the two graphs are design semantically different. However, their corresponding floorplans yielded similar exit flow values. A GDP graph transformation as well as density heatmaps are shown, Figure 3b.

4.2 Behavioral and Geometrically-powered Floorplans Retrieval

In this use case we demonstrate the potential of using our embedding methodology to retrieve geometrically-powered (includes static features), behaviorally-powered (includes crowd-based features) and combined, similar and related floorplans.

Figure 5a showcases an example for geometrically-powered floorplans retrieval. This embedding setting contains a 12 dimensional design semantic vector, 11 for room types and 1 for room dimensions. Given an input floorplan, a set of 5 nearest neighbors from the embedding space are retrieved. For an embedding space to be meaningful and valid, a queried graph should have itself as the first nearest neighbor during a retrieval, and therefore, in the figure, the first neighbor is same as the queried floorplan itself. For first query, the second and third nearest neighbours have same structure, the hallway in the middle and two bedrooms, one kitchen and one bathroom around it. Though, they have different design semantic features as annotated in the image. Fourth neighbour has close structure and only with missing the kitchen node. The last neighbour almost has same structure but the middle node is different.

Figure 5b showcases an example for behaviorally-powered floorplans retrieval. In this example the query graph is the same as last one and we retrieve the top 5 nearest neighbours. The model trained by behavioral features is used. All the neighbors have one node in the middle and four in the surroundings with same node degree. The second row shows simulation heatmaps for behavioral features. Since semantic features are not considered in this use case, two of the neighbours have totally different room types and more-less similar square footage. Also, the second nearest neighbour is exactly same as second nearest neighbour in case 1. It shows even without design semantic features, the structure and behavioral features can capture floorplans similarity.

Figure 5c showcases an example of floorplans retrieval from a collectively trained model with both semantic and behavioral features. The features dimension for each node is 21. The top 5 nearest neighbours for the same input floorplan are retrieved from the embedding. Since in this case all of the features are considered, the neighbours must follow similar structure, similar design semantic and also similar behavioral pattern. Regarding structure, they almost have same structure, one node is missing in fourth and fifth neighbours and one node is different in fourth neighbour. But node degrees and structure of the graphs are almost same. The heat-maps visualizing the behavioral features are shown in second row and as it is clear their similarity decreases for later neighbours. Semantic features as annotated in the images are more-less similar. There are some noticeable points in this study. First the floorplan in second rank has the second rank in two later cases. This shows the validity of embedding space which the similarity and dissimilarity between floorplans are captured perfectly in vectors in embedding space. The third neighbours is also seen in last two cases. Semantically it was the third neighbour but behaviorally it was in fourth rank, the accumulation of both class features push it in third rank. This behavior is another justification for embedding space validity and benefit of con-

sidering behavioral features. Behavioral features weakly represent the semantic attributes and integrating them helps to fin the best fit nearest neighbours. Third point is the presence of a new floorplan in fourth rank which is not seen in both cases but the accumulation of features classes as discussed later lead it to get fourth rank. Fifth neighbour was the fourth neighbour in first study and was not seen in second study. It means this floorplan has similar structure design semantic but behaviorally different pattern which their accumulation keeps it in fifth rank.

In order to have an evaluation over embedding space, we find the rank of each floorplan by itself between top 5 nearest neighbours. A meaningful embedding space should lead to a case that each graph has its own as first nearest neighbor. We repeat this process for all three models, model trained with only design semantic features, model trained with only behavioral features and model trained with all features. As Table 2 shows, in all models the query graph by itself is first nearest neighbor. As additional metric, we generate one proxy graph for each floorplan. Proxy graphs are the result of running random walk one more time that makes different sequence for each floorplan. These proxy graphs are not involved in training. We feed them to the trained model and add their representative vectors to embedding space. Then we find the top 10 nearest neighbors for each floorplan. The reason for top 10 nearest neighbours is to make sure we see them between neighbours. Having these proxy graphs in higher ranks shows the validity of embedding space. The result is provided in table 2 for all three models. The reason that all of the proxy graphs are not seen in second rank is because our graphs are unidirectional and we allow having loop in random walk. The randomness in random walk could lead to sequences that is back and fourth between only two nodes or a repeated subsequence because of loop. In both cases the generated sequences are not presenting the graph properly. This situation also has dependency to node degrees in the graphs since next node in random walk is selected based on weighted probability of nodes degree. This can be addressed by running random walk more than one time for each node or avoiding loops, both in training and test time.

4.3 Generation of a Composite Floorplan

Embedding spaces have commonly been used to make predictions and retrieving similar objects given some input criterion or an object. In this use case we demonstrate that our embedding methodology can also be used to generate a compound floorplan with collective features given multiple floorplans with a set of completely or partially different features. Figure 4 showcases one such example. Two floorplans with 2 and 3 different room types respectively, are given as input to our autoencoder, and the first nearest neighbour is a composite floorplan consist of 5 different room types. It not just contains all the features from input floorplans but also maintains a significant proportion of their geometric symmetries.

5 CONCLUSION AND DISCUSSION

This paper presented floorplans as vectors in a continuous space, encoded with both design and crowd behavioral features. Specifically, the framework consists of two compo-

Ranks/Models	Model1	Model2	Model3
First rank(query graph itself)	100%	100%	100%
Second rank(proxy graph)	86%	83%	85%
Third rank(proxy graph)	13%	13%	14%
Fourth rank(proxy graph)	1%	3%	1%
Fifth rank(proxy graph)	0	2%	0

Table 2: This table shows the percentage of the floorplans that have them-self as first nearest neighbours and the rank percentage of their proxy graphs.

nents. The first component is for representing floorplans as intermediate attributed graphs and the second one is the embedding for transforming attributed graphs to a vector. For converting floorplans to the attributed graphs we designed an automated tool based on image processing techniques that gets the image or JSON file of a floorplan as input and convert it to an attributed graph augmented with design semantic and crowd behavioral features generated by simulation. For embedding, we propose an LSTM Autoencoder that converts these graphs to vectors. The floorplans with similar structure, design semantic and crowd behavioral features have close vector in embedded space. These vectors facilitate operation over floorplans in applications like finding similar floorplans, recommendation and etc.

We posit that this contribution paves the way for novel, ground-breaking developments for human-in-the-loop or automated floorplan clustering, exploration, comparison and generation. By encoding latent features in floorplan embeddings, designers can store multi-dimensional information of a building design and performance to quickly identify floorplan alterations that share similar or different features. While in this paper we encode features derived from dynamic crowd simulations of building occupancy, the proposed approach can virtually scale to encode any kind of static or dynamic performance metric.

REFERENCES

1. Ahmed, S., Weber, M., Liwicki, M., Langenhan, C., Dengel, A., and Petzold, F. Automatic analysis and sketch-based retrieval of architectural floor plans. *Pattern Recognition Letters 35* (2014), 91–100.

2. Ahonen, T., Hadid, A., and Pietikainen, M. Face description with local binary patterns: Application to face recognition. *Transactions on Pattern Analysis & Machine Intelligence*, 12 (2006), 2037–2041.

3. Bunke, H. What is the distance between graphs. *Bulletin of the EATCS 20* (1983), 35–39.

4. Chechik, G., Shalit, U., Sharma, V., and Bengio, S. An online algorithm for large scale image similarity learning. In *Advances in Neural Information Processing Systems* (2009), 306–314.

5. Dalal, N., and Triggs, B. Histograms of oriented gradients for human detection (2005).
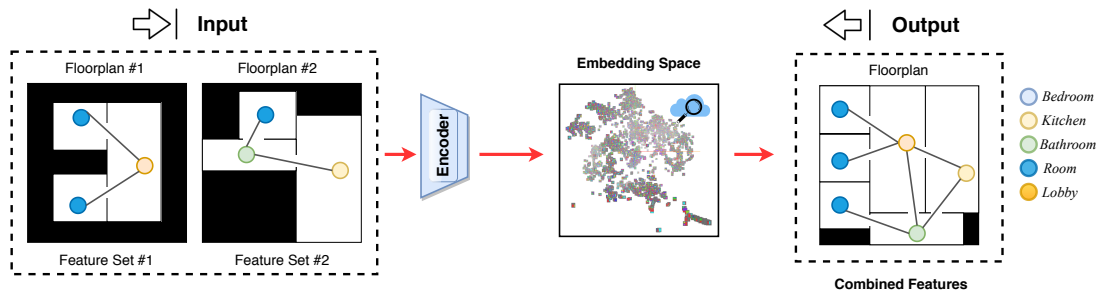
Figure 4: Generation of a single floorplan with combined features from an embedding space given two floorplans with different set of features as input. Nodes of underline graph for each floorplan are color-coded based on room types.

6. de las Heras, L.-P., Fernández, D., Fornés, A., Valveny, E., Sánchez, G., and Lladós, J. Runlength histogram image signature for perceptual retrieval of architectural floor plans. In *Workshop on Graphics Recognition*, Springer (2013), 135–146.

7. Dosch, P., and Masini, G. Reconstruction of the 3d structure of a building from the 2d drawings of its floors. In *Document Analysis and Recognition*, IEEE (1999), 487–490.

8. Dutta, A., Llados, J., and Pal, U. Symbol spotting in line drawings through graph paths hashing. In *Document Analysis and Recognition*, IEEE (2011), 982–986.

9. Goyal, P., and Ferrara, E. Graph embedding techniques, applications, and performance: A survey. *Knowledge-Based Systems 151* (Jul 2018), 78–94.

10. Grover, A., and Leskovec, J. node2vec: Scalable feature learning for networks. In *Knowledge discovery and data mining*, ACM (2016), 855–864.

11. Gupta, N., Das, S., and Chakraborti, S. Extracting information from a query image, for content based image retrieval. In *Advances in Pattern Recognition*, IEEE (2015), 1–6.

12. Heylighen, A., and Neuckermans, H. A case base of case-based design tools for architecture. *Computer-Aided Design 33*, 14 (2001), 1111–1122.

13. Hinton, G. E., and Zemel, R. S. Autoencoders, minimum description length and helmholtz free energy. In *Advances in neural information processing systems* (1994), 3–10.

14. Hochreiter, S., and Schmidhuber, J. Long short-term memory. *Neural computation 9*, 8 (1997), 1735–1780.

15. Lambert, G., and Gao, H. Line moments and invariants for real time processing of vectorized contour data. In *International Conference on Image Analysis and Processing*, Springer (1995), 347–352.

16. Lazebnik, S., Schmid, C., and Ponce, J. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Computer Vision and Pattern Recognition*, vol. 2, IEEE (2006), 2169–2178.

17. Li, T., Ho, D., Li, C., Zhu, D., Wang, C., and Meng, M. Q. H. Houseexpo: A large-scale 2d indoor layout dataset for learning-based algorithms on mobile robots, 2019.

18. Macé, S., Locteau, H., Valveny, E., and Tabbone, S. A system to detect rooms in architectural floor plan images. In *Workshop on Document Analysis Systems*, ACM (2010), 167–174.

19. Narayanan, A., Chandramohan, M., Venkatesan, R., Chen, L., Liu, Y., and Jaiswal, S. graph2vec: Learning distributed representations of graphs. *arXiv preprint arXiv:1707.05005* (2017).

20. Richter, K., Heylighen, A., and Donath, D. Looking back to the future. an updated case base of case-based design tools for architecture (01 2007).

21. Sabri, Q. U., Bayer, J., Ayzenshtadt, V., Bukhari, S. S., Althoff, K.-D., and Dengel, A. Semantic pattern-based retrieval of architectural floor plans with case-based and graph-based searching techniques and their evaluation and visualization. In *ICPRAM* (2017), 50–60.

22. Sharma, D., and Chattopadhyay, C. High-level feature aggregation for fine-grained architectural floor plan retrieval. *IET Computer Vision 12*, 5 (2018), 702–709.

23. Sharma, D., Chattopadhyay, C., and Harit, G. A unified framework for semantic matching of architectural floorplans. In *Pattern Recognition*, IEEE (2016), 2422–2427.

24. Sharma, D., Gupta, N., Chattopadhyay, C., and Mehta, S. Daniel: A deep architecture for automatic analysis and retrieval of building floor plans. In *Document Analysis and Recognition*, vol. 1, IEEE (2017), 420–425.

25. Singh, S., Kapadia, M., Faloutsos, P., and Reinman, G. An open framework for developing, evaluating, and sharing steering algorithms. In *Motion in Games*, Springer-Verlag (2009), 158–169.

26. Taheri, A., Gimpel, K., and Berger-Wolf, T. Learning graph representations with recurrent neural network autoencoders. *KDD Deep Learning Day* (2018).

27. Weber, M., Liwicki, M., and Dengel, A. A. scatch-a sketch-based retrieval for architectural floor plans. In *Frontiers in Handwriting Recognition*, IEEE (2010), 289–294.

28. Wessel, R., Blümel, I., and Klein, R. The room connectivity graph: Shape retrieval in the architectural domain.
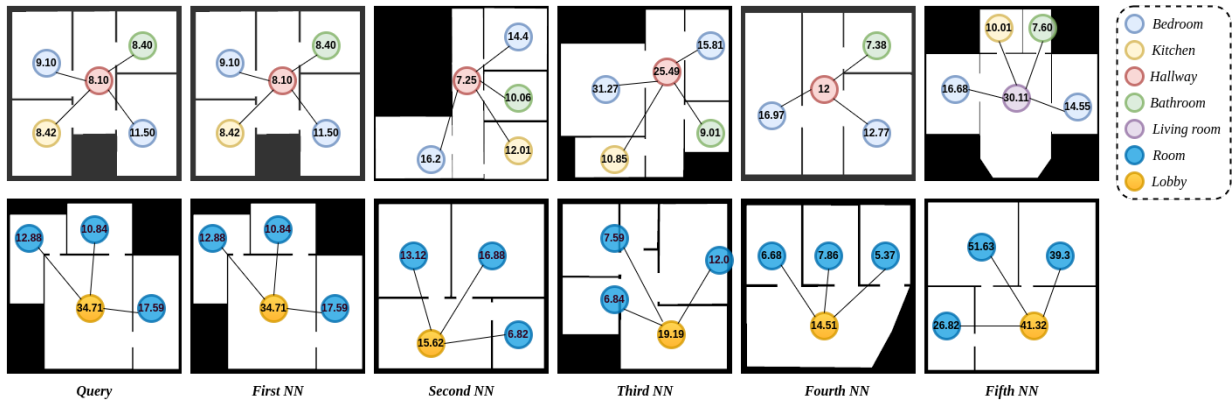
(a) Floorplans retrieval from the embedding with respect to design semantic features alone. Top 5 nearest neighbours of two queried floorplans are shown.
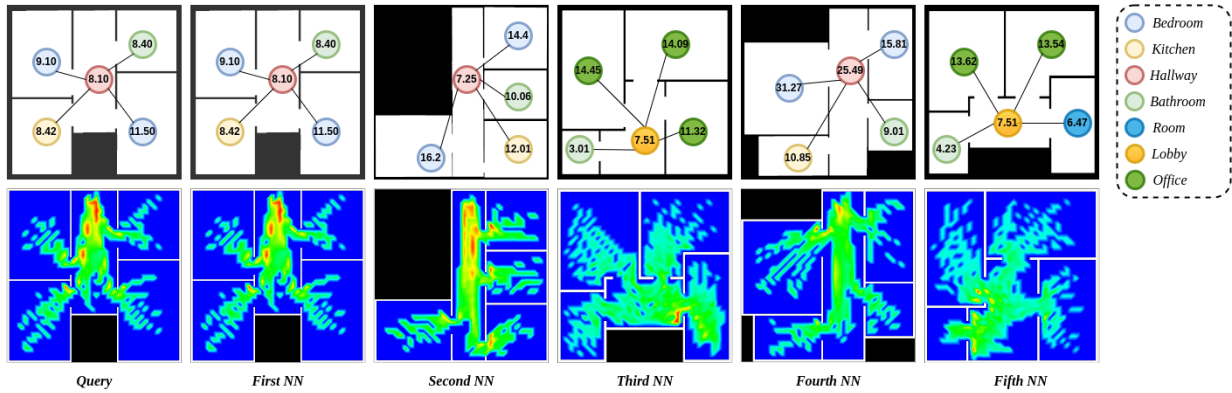
(b) Floorplans retrieval from the embedding with respect to behavioral features alone. Top 5 nearest neighbours of for queried floorplan is shown. The ranking of the neighbours is computed based on differences in their behavioral attributes. Time-based behavioral dynamics for human-building interactions are also shown as color-coded heatmaps, where red areas highlight over crowded regions in space.
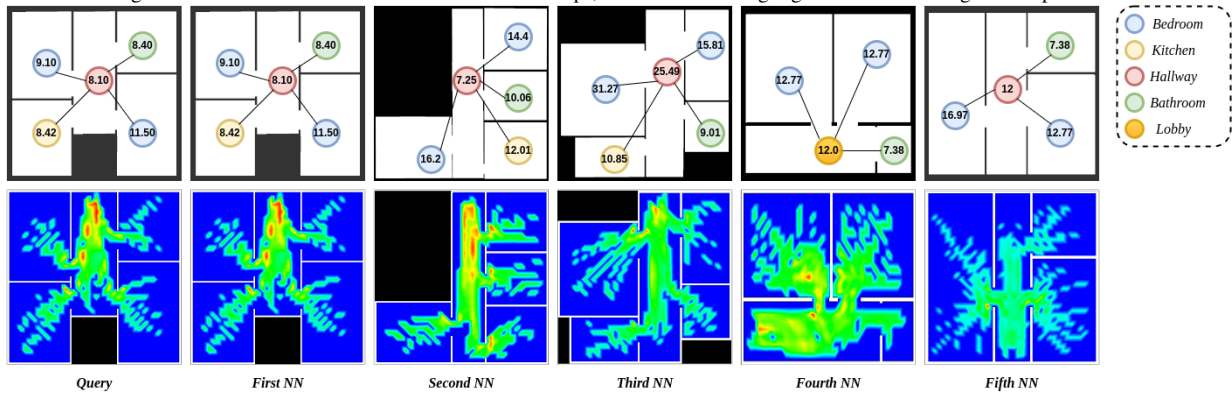
(c) Floorplans retrieval from the embedding with respect to combined design semantics and behavioral features. Top 5 nearest neighbours for the queried floorplan is shown. Time-based behavioral dynamics for human-building interactions are also shown as color-coded heatmaps, where red areas highlight over crowded regions in space.

Figure 5: Retrieval of 5 nearest floorplans for the given design layout from embedding space for design semantic only, behavioral only and combined set of features. We used same input design layout in (a), (b) and (c) to demonstrate how the embedding retrievals vary among different categories of features space.

Building Cluster Optimization to Integrate Energy Performance and Outdoor Thermal Comfort

Francesco De Luca[1], Emanuele Naboni[2], Gabriele Lobaccaro[3] and Abel Sepúlveda[1]

[1] Tallinn University of Technology
Tallinn, Estonia
{francesco.deluca, absepu}@taltech.ee

[2] University of Parma
Parma, Italy
and KADK, Denmark
emanuele.naboni@unipr.it

[3] Norwegian University of
Science and Technology
Trondheim, Norway
gabriele.lobaccaro@ntnu.no

ABSTRACT
Climate change can be tackled by a careful design of urban density, buildings distance and orientation, and materials in order to pursue indoor and outdoor comfort, and low building energy consumption. The research develops a methodology to holistically coupling overheating risks, cooling energy consumption, and pedestrian comfort for the careful design of cluster configurations of office buildings in the city of Tallinn, Estonia. The outcomes can help climate conscious design in Nordic cities. Solutions are provided to reduce overheating up to 23.6% and to increase urban comfort up to 13.5%.

Author Keywords
Urban Design; Energy Efficiency; Urban Comfort; Environmental Analysis; Performance-driven Design.

ACM Classification Keywords
I.6 SIMULATION AND MODELING - Applications; J.5.

1 INTRODUCTION
Buildings form and distance, finishing materials and operation, as well as open spaces' sizes and materials affect local microclimate in cities. Urban morphology affects the magnitude of the meteorological parameters such as air temperature, relative humidity, wind velocity, and mean radiant temperature, which contribute in the definition of outdoor human comfort through thermal indices [6].

Also in Nordic cities, pedestrian thermal discomfort during the warm season is significantly increasing due to a lack of conscious design, especially in new commercial districts [5]. In the near future, urban fabrics will have an even more significant impact due to increasing global temperatures in the built environment. In office buildings characterized by glazed envelopes, the excess of solar gains and increasing exterior air temperature dramatically increases overheating and consequent cooling energy demand during the warm season also at northern latitudes [7]. Studies show the importance of surrounding buildings distance and urban layout in the design of energy efficient developments [8].

In this scenario, the choices of designers and planners can significantly improve both the livability of the urban environment and the quality of the urban microclimate [2].

The present work investigates the performance of office building clusters in Tallinn, Estonia, during the warm season. The aims are: 1) quantify overheating risk and cooling energy needed to maintain occupants comfort; 2) quantify pedestrian thermal stress in commercial districts; 3) find optimal cluster configurations for both performance and trade-offs. The scope is to provide design solutions to help increase the resource efficiency and urban comfort of commercial districts. The novelty of the work lies in the scarcity of overheating analysis for clusters of buildings and the lack of urban comfort studies in Estonia.

2 METHODS
Indoor and outdoor comfort are assessed using the Estonian energy ordinance [3], and the Universal Thermal Climate Index (UTCI) [1], respectively. The first requires that in offices, the temperature does not exceed 25 °C for more than 100 °C·h during the period from 01.06 to 31.08. The latter defines a comfortable condition between 9 and 26 °C equivalent temperature, and cold and hot stress levels.

2.1 Building Cluster
The office buildings cluster used in the study presents three tower buildings. Nine cluster layouts (V1-V9) are used to analyze variations of indoor temperature and cooling energy need, and outdoor comfort (Figure 1). The buildings
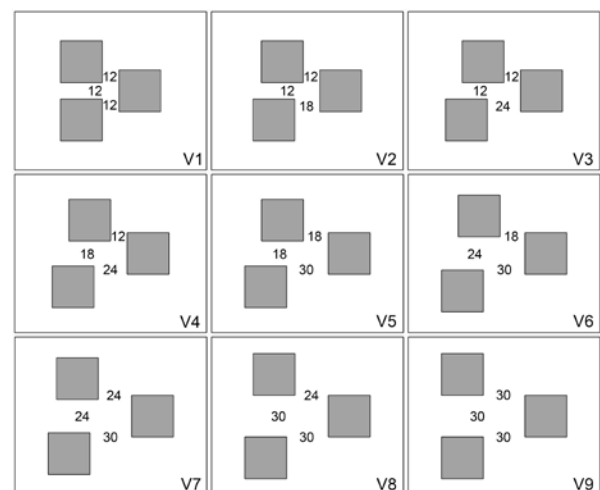

Figure 1. Cluster variations and figures of building distances (m).

Figure 2. The three urban areas of Liivalaia st., Mustamäe st. and Logi st. (left to right) with the building cluster plot (red rectangle).

square footprint length is 30 m, and the height is 100 m. The floor to floor height is 4 m for a total of 25 floors. The cluster plot is 138 m by 114 m in size. The building distances vary not uniformly from 12 m to 30 m (Figure 1). The nine configurations were chosen inasmuch representative of commercial district layouts recently built and under construction in Tallinn.

2.2 Urban Areas

The building clusters are located in three typical urban areas to evaluate the influence of different built environments on indoor and outdoor comfort. The area of Liivalaia st. is located in the high-density city center. The area of Mustamäe st. is located in a mixed-use and medium density quarter. The area of Logi st. is located by the sea in the low density port area. The plot is rotated differently in the three areas to follow the urban morphology (Figure 2).

2.3 Parametric Model

For the study, the three urban environments and the three cluster buildings are modeled using the software Rhinoceros. A parametric model is realized using the tool Grasshopper. It permits to change the location of the buildings for the nine configurations in each urban location (for a total of 27) and to run the following simulations and analysis automatically: 1) urban weather simulation using the Urban Weather Generator (UWG) software; 2) thermal and energy simulations using EnergyPlus software; 3) wind simulations using the software OpenFOAM through the plug-in Swift; 4) UTCI through the plug-in Honeybee.

2.4 Urban Weather

The way buildings are clustered together and anthropogenic activities define urban weather and the amount of heat trapped in urban canyons. UWG uses the urban morphology to calculate the average building height, the site coverage ratio and the façade-to-site ratio. Additionally, it accounts for traffic, building systems, and green areas. UWG permits to obtain weather datasets for the three urban areas through modification of the existing weather data collected in the rural area, to be used for thermal and UTCI analysis.

2.5 Overheating Simulations

Overheating assessment is performed for five floors for each office tower, 1st (ground), 7th, 13th, 19th and 25th floor (Figure 3). The scope is to reduce the time for computationally intensive simulations, and at the same time to guarantee reliability of results. Each floor is divided into

nine thermal zones, eight open offices on the perimeter and one core for a total of 135 zones. The use of multiple zones guarantees a minimum level of accuracy for interior and outer surface temperature simulations. Overheating results are analyzed only for the perimeter zones.

Parameters for the thermal zone, schedules and calculation methods follow the prescriptions of Estonian regulations [4]. Building envelope properties used are standard practice for energy efficient office buildings in Estonia (Table 1). Operable windows and natural ventilation are not used. Air temperature is simulated during the required period every occupied hour. The ordinance allows cooling in office buildings to provide indoor comfort, limiting the temperature exceeding 25 °C below 100 °C·h (degree/hour). The parametric model performs air temperature simulations

Thermal Zone Settings					
People density	0.059 (p/m^2)				
Equipment density	8 (W/m^2)				
Lighting power density	7 (W/m^2)				
Infiltration per floor area	0.000056 (m^3/s/m^2)				
Mech. vent. per floor area	0.002 (m^3/s/m^2)				
Shading/cooling set point	24/25 (°C)				
Building Envelope Properties					
	EW	Win	F	GF	R
U_t **(W/m^{2}K)**	0.15	0.9	A	0.16	0.12
Win	**WWR** 70% - **VT** 51% - **SHGC** 0.25				
Occupancy-Lighting-Equipment schedule					

| Weekend 1-24 | 0.0 |

Table 1. Overheating simulation parameters. U_t=U value total, EW= ext. walls, Win= window, F=floors, GF=ground fl., R=roof, A=adiabatic, WWR=Win-to-Wall ratio, VT=visible transmittance.
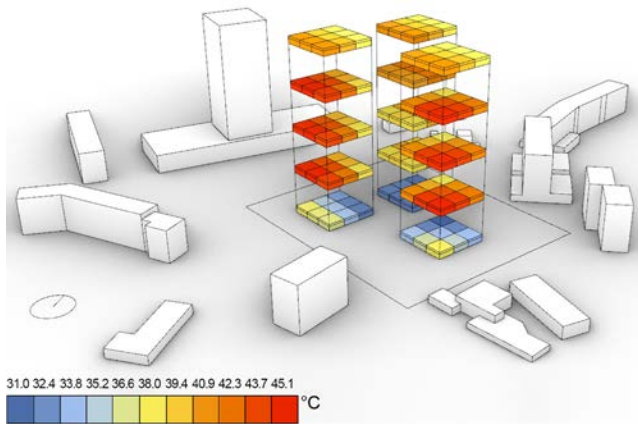
Figure 3. Floors used for overheating simulations for Liivalaia st. V5. Zone average temperature during analysis period.

with and without zone conditioning. Simulations without conditioning allow for determining indoor air temperature, those with conditioning estimate cooling energy demand to fulfil the Estonian energy ordinance.

2.6 Wind Simulations

Computational Fluids Dynamics (CFD) wind simulations are performed following best practices [4]. Buildings within a distance of 500 m from the plot are included in the computational domain. The size of the hexahedral cells varies from 16 m to 4 m for the urban area, 2 m for the surrounding buildings, 1 m for the cluster buildings and 0.5 m for the ground. Wind simulations are performed from 16 cardinal directions with a wind velocity of 5 m/s. Different terrain roughness values (Z_0) are used to account for the different urban morphologies at far distance from the plot.

2.7 Outdoor Comfort Analysis

UTCI assesses the temperature perceived by people based on meteorological input of radiation, air temperature, wind and humidity, and a physiological and a clothing model. The analysis period used is from 03.08 to 09.08. Being the Extreme Hot Week of the urban weather converted STAT file, it is used as a worst case scenario. The time selected is the business and after-work hours from 8 a.m. to 8 p.m.

The workflow used to calculate UTCI is the following. Mean Radiant Temperature (MRT) is calculated, taking into account direct solar radiation, urban environment, the buildings surface temperature and view factors from each of the 1448 nodes of the analysis grid. Building surface temperature is simulated using all the 25 floors of the cluster buildings with a single zone per floor, to limit computation time. Zone conditioning is used to simulate real building use. Hourly wind velocities are calculated multiplying the hourly measured wind speeds by wind factors obtained dividing the simulated velocities by the fixed velocity used for CFD simulations. Air temperature and relative humidity are obtained, as described in 2.4. The pavement material is concrete with a rough finish and an Albedo of 0.30. UTCI hourly values are recorded for every building cluster and every urban area (Figure 4).

3   RESULTS

Results show that temperature degree hour values in all variations and zones exceed by a large margin the maximum limit fixed, up to 168 times. Taking into account the configurations in the same cluster, degree hour increases with the buildings distance. In configuration V9 there is a variation of up to +21.4% in Liivalaia st., +16.4% in Mustamäe st. and +31.2% in Logi st. in comparison with configuration V1. This aspect affects the energy for cooling which in configuration V9 is up to +10% in Logi st. and up to +15% in Liivalaia st. and Mustamäe st. in comparison with configuration V1. The minimum and maximum zone energy need for cooling is 1.8 kWh/m^2 and 11 kWh/m^2. By comparing the same configurations in different districts, degree hour variations are up to +18.9% and +19.1% for clusters V8 and V9 respectively between Logi st. and Liivalaia st. This underlines the importance of the surrounding environment density for building performance.

For outdoor comfort assessment, building cluster variations are compared in relation to the maximum ratio of the outdoor area with no thermal stress (≤ 26 °C UTCI) during the analysis period. Evidence shows that also at northern latitudes urban discomfort is a serious concern since the maximum ratio of the area in the state of comfort ranges between 36.5% and 52.5% (Table 2). Results show an opposite trend comparing overheating. Cluster variations with more considerable buildings distance guarantee a larger comfort. Nevertheless, the smallest and largest plot area ratio in the state of comfort is not always related to V1 and V9, respectively. Additionally, the cluster variations with the worst and best performance are not the same in the different urban areas. The correlation between buildings distance and outdoor comfort is strong for the medium and low density areas and is weak for the high density area.
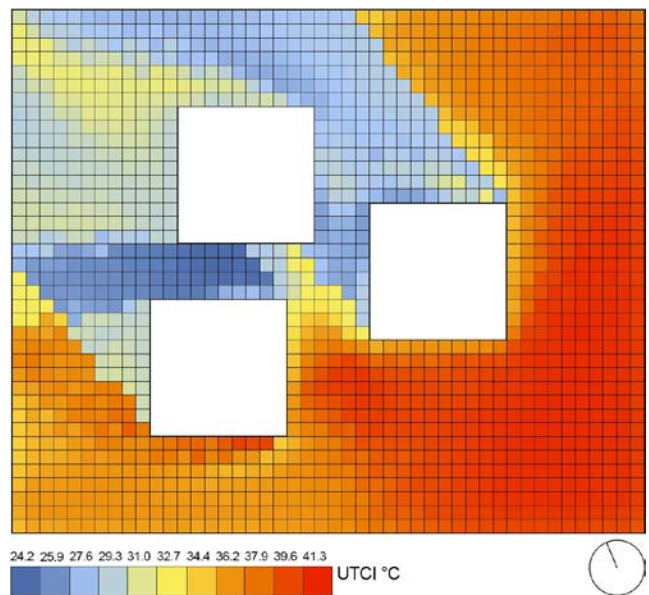

Figure 4. Outdoor comfort map showing UTCI (°C) values at 12 a.m. of the analysis day 04.08 for Mustamäe st. area V2.

V1	V2	V3	V4	V5	V6	V7	V8	V9
44.3	42.3	42.5	45.0	45.4	40.1	47.9	44.0	50.3
36.5	37.0	41.0	40.2	38.8	48.1	43.2	43.8	47.3
39.0	46.4	45.3	47.4	49.0	52.3	51.2	52.5	47.9

Table 2. Max. ratio (%) of plot area in comfort condition for Liivalaia st., Mustamäe st. and Logi st. (from top to bottom rows).

Analysis of integrated performance is performed using the minimum values of zone temperature degree hour, and the maximum ratio of cluster outdoor area with state of comfort. None of the three study areas presents an optimal cluster configuration. Nevertheless, trade-offs are present. (Figure 5). For Liivalaia st. area, variation V1 has the best performance for indoor comfort allowing the smallest overheating (4281 °C·h) and a mean value of maximum plot area with state of comfort (44.3%). For Mustamäe st. area trade-off variations are those with average mean values for both performances, as V3 (outdoor comfort 41% - overheating 5190 °C·h). For Logi st. area a trade-off is V5 that allows the maximum ratio of plot area with state of comfort (49%) and an average overheating (5594 °C·h).
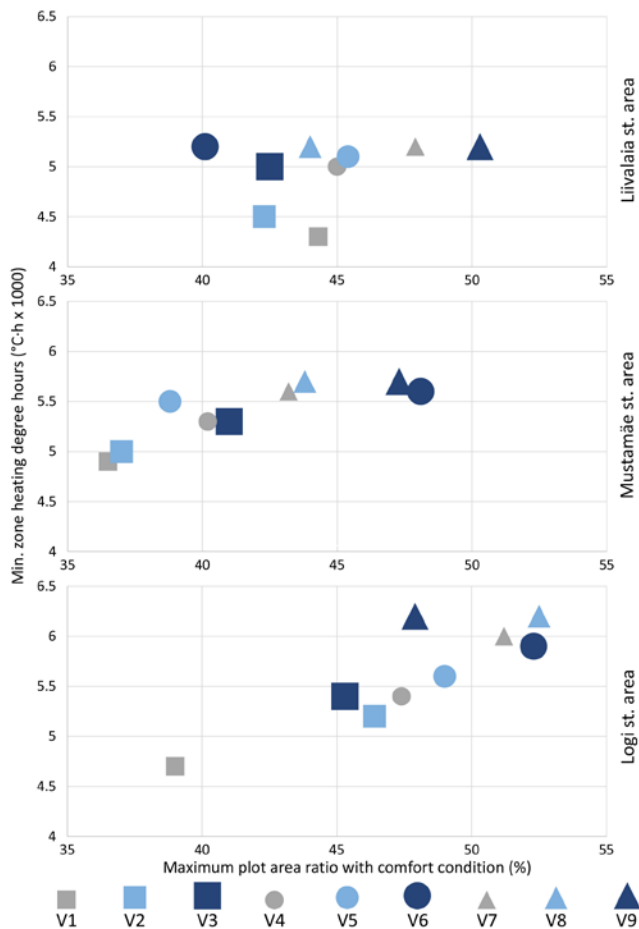
Figure 5. Charts of the integrated performance analysis.

4 CONCLUSION

The paper presents an integrated analysis of overheating and cooling energy, and outdoor thermal comfort during the warm season, for office building clusters located in three areas of the city of Tallinn. The study integrates the urban environment, cluster variations and different simulations in a parametric model. The outcomes of the study are: 1) All the building zones exceed the overheating limit by many times. It is possible to reduce overheating by up to 23.6% and cooling energy, through an environmentally conscious design; 2) Pedestrian thermal stress in commercial districts is a severe concern. The ratio of the pedestrian area within the state of comfort is no more than 52.5% during the analyzed period. It is possible to achieve an increase of 13.5% through performance-based design; 3) For the analyzed cases, no best configuration exists, but trade-off solutions are presented. These can be used by designers to increase the resource efficiency and outdoor comfort, hence to decrease the climate impact of commercial districts.

REFERENCES

1. Bröde, P., Jendritzky, G., Fiala, D. and Havenith, G. The universal thermal climate index UTCI in operational use. *Proc. of Adapting to Change: New Thinking on Comfort*, Windsor, UK, 2010.

2. Coccolo, S., Kämpf, J., Scartezzini, J. and Pearlmutter, D. Outdoor human comfort and thermal stress: A comprehensive review on models and standards. *Urban Climate 18*, 2016, 33-57.

3. Estonian Government. *Minimum requirements for energy performance*. n° 68, Tallinn, 2012.

4. Franke, J., Hellsten, A., Schlünzen, H. and Carissimo, B. *Best Practice Guideline for the CFD Simulation of Flows in the Urban Environment*. COST, Bruss., 2007.

5. Mauree, D., Coccolo, S., Perera, A.T.D., Nik, V., Scartezzini, JL. and Naboni, E. A New Framework to Evaluate Urban Design Using Urban Microclimatic Modeling in Future Climatic Conditions. *Sustainability 10*, 4 (2018), 1134.

6. Park, S., Tuller, S. and Jo, M. Application of Universal Thermal Climate Index (UTCI) for microclimatic analysis in urban thermal environments. *Landscape and Urban Planning 125*, (2014), 146-155.

7. Voll, H., De Luca, F. and Pavlovas, V. Analysis of the insolation criteria for nearly-zero energy buildings in Estonia. *Science and Technology for the Built Environment 22*, 7 (2016), 939–950.

8. Voll, H., Thalfeldt, M., De Luca, F., Kurnitski, J. and Olesk, T. Urban planning principles of nearly zero-energy residential buildings in Estonia. *Man. of Environ. Quality: An Int. Journal 27*, 6 (2016), 634–648.

A Case of High-Performance Building Form Design Workflow Informed by Computational Simulation

Haobo Liu[1], Adam Rysanek[2], Andrea Frisque[1] and Jeanie Chan[1]

[1]Stantec Consulting Ltd.
Vancouver, Canada
{Haobo.Liu, Jeanie.Chan, Andrea.Frisque}@ stantec.com

[2]University of British Columbia
Vancouver, Canada
arysanek@sala.ubc.ca

ABSTRACT

This paper aims to formulate and test a parametric high-performance building design workflow that allows architects to explore realistic space attributes. It also allows architects to assess their environmental performative outcomes such as daylight simulation, solar radiation and occupant visual link in order to inform decision making during the design process. Using the Rhino and Grasshopper platform, a new workflow is proposed for generating and analyzing building forms generatively and extensively to predict their environmental performance. This approach comes from the improved interoperability between the parametric model tools, simulation engines and statistical analysis tools, enabling significant ability to compare energy performance with other performance metrics. The overall framework is divided into four steps: site setup, massing generation, performance evaluation and visualization, and design development. Through a residential building design case in Vancouver, it is anticipated that, by incorporating knowledge about the environmental performance of a design early in the volume-making process, the proposed framework will help designers better navigate performance objectives in the architectural design environment.

Author Keywords

Parametric design; procedural form generation; environmental performance; parallel simulation; building archetype modelling

ACM Classification Keywords

Simulation Theory, Model Development

1 INTRODUCTION

Architects regard themselves as professionals of building form manipulation. Tracing back to the Roman Republic, Vitruvius stated architectural design as a procedure of making space that fulfills the criteria of commodity, firmness, and delight [1]. Since the birth of modern architecture, the organization of building form has become a core research area of design practice [2], and one of the epistemologies regards space as an enclosed volumetric massing [2]. Meanwhile, in mainstream building performance modelling software, the building form is input into the simulation engine as a volumetric zone including its geometrical attributes, programs, behaviours, schedules as well as loads [3]. This connection gives us the opportunity to design an artful volumetric form with consideration via building performance modelling in the preliminary design phase.

In the contemporary world, our increasing understanding of climate change and its future risk, along with the understanding that buildings contribute significantly to greenhouse gas emissions, forces us to assess or re-assess the relationship between building form design, thermal performance, and overall building environmental quality. More municipalities are adopting performance-based codes which require the building to comply with the specific energy target at the time of permit application. For example, in British Columbia, new construction projects must meet the Energy Step Code [4].

In traditional building energy modelling procedures such as the LEED Standard, which uses ASHRAE 90.1 Appendix G [5], the focus has mostly been placed on verifying the final design performance. This is unlikely to cue architects to perform robust and elegant design actions in the design process. However, as the parametric model tools and simulation engines improve, a significant potential is the ability to compare building performance between different form designs. This asks us to change the attitude of the performance model in architecture design, moving from a performance-analysis to performance-informed workflow. It requires, among other measures, fundamental thinking of design strategy, where architects are not yet proficient with the methodology to make this change happen.

The workflow proposed in this paper is for parametric demonstration of various performative outcomes according to basic form exploration in the preliminary design phase. It contains architecture information and enables various volumetric massing iterations with the facilitation of computer-based parametric programs. Performance can be tracked as design iterations are initially developed, helping to draw out more sustainable buildings designs.

2 COMPUTATIONAL BUILDING PERFORMANCE MODELLING

From the scientific viewpoint, it has long been acknowledged that the decision to invest in a particular building design depends on measurable performance

metrics, such as a design's environmental footprint. One of the greatest analysis tools available to professionals is building performance simulation (BPS) tools [3]. These are computer programs that can provide the capacity to simulate building energy physics in detail or estimate the future energy consumption resulting from an installed retrofit. With over three decades of development, BPS tools have become the industry standard for the design, specification, and evaluation of energy supply systems and energy demand reduction measures in new or retrofit building projects. Simulations using tools such as eQUEST, IES Virtual Environement, or EnergyPlus are most widely known in the industry. These tools usually combine digital models of a building with weather data to accurately simulate the thermal behavior of individual zones of a building and interactions between the different building components. In using any of these tools, one may find thousands of different implementable simulation inputs, from building occupant schedules to air-conditioning system configurations. Understanding that optimizing building design using BPS tools is sophisticated and time consuming. Researchers in the BPS industry have pursued a niche computational solution to this problem: the use of integrated workflow execution software, or wrappers, that permit parametric optimization of building design using BPS tools. BEOpt is one of the most well-known wrappers in this domain[6].A typical use of the BEOpt may involve an EnergyPlus model repeatedly evaluating the performance of different glazing systems on a façade until the most cost-optimal glazing system configuration is found, as in the example of façade design in a cold climate [7]. In general, existing optimization wrappers for BPS tools are limited to evaluating parameters that are easily configurable in the BPS tools themselves. As BPS tools were adopted initially by engineering consultants in the buildings industry, it follows that nearly all easily configurable parameters have also been engineering-centric, such as façade material properties, and/or building energy system types. As this encompasses a broad set of variables, it is perhaps simpler to define what parameters have not been easily configurable within the typical BPS-driven optimization process. These are virtually all parameters affecting the preliminary architectural design of buildings: building programming, massing, orientation, and glazing ratio etc. This has been an unfortunate paradigm for the building design process, particularly in light of future building codes. As countries are beginning to adopt increasingly stringent energy performance targets for future buildings [8], it will be contingent on architects, and not necessarily engineers, to identify building forms that can satisfy environmental quality requirements passively [9]. Perhaps in light of this, the typical engineering-centric parametric optimization paradigm is now changing. A new type of parametric modelling software, catering to architectural design, has emerged, and Grasshopper, a visual programming language for the 3D computer-aided design (CAD) software, Rhinoceros, is at the forefront of

this new field. Grasshopper is a generic platform, allowing architects to develop algorithmic processes for preliminary building design as well as connect these algorithmic processes to third party BPS software tools [10]. For example, the ARCHSIM plug-in for Grasshopper, produced by MIT spin-off company Solemma, allows co-simulation between a Rhino Grasshopper parametric architectural design model and EnergyPlus to perform energy simulation [11]. These emerging tools act as the connection between the model and the analysis results in a way that allows the designer to keep manipulating model parametersuntil the desired analysis result is achieved.

Furthermore, several computational plug-ins are available for Grasshopper, such as the Colibri developed in Core Studio in Thornton Tomasetti. Colibri[12] is an open-source tool for investing simulation-based, multi-objective design and decision problems. These tools could be used to wrap an ARCHSIM-evaluated parametric design problem with a parallel simulation solver [13].

3   COMPUTATIONAL BUILDING FORM DESIGN

In the last decades, in accompaniment with the development of computation, researchers have been exploring form language using computer algorithms [14]. Steadman first suggested that if one would be given appropriate geometric definitions of certain classes of plans, systematic methods could be developed for computing all possible plans of each program type [15]. More recently, Steadman proposed a new approach to building design based on generating possible form iterations of building archetypes [16]. In this new method, Steadman assigns a binary code of 0 and 1 to indicate the absence or presence of dimensionless strips of accommodation or open space across each plot. Homeira Shayesteh continued P. Steadman's work and applied it in an urban form generation method for Tehran[17].H. Shayesteh uses the archetypal representation to explore and better understand the relationships between urban built form characteristics, plot size, housing layout, and ground coverage about density. H. Shayesteh investigates the evolution of the stereotypical house form in Tehran over time. Namely, H. Shayesteh analyzed stereotypical forms for housing and developed a model that brings together parameters of an urban structure (e.g. block and plot size,ground coverage ratio) with parameters of the built form (e.g. access frontage, day-lit depth, plan shape). Several years ago, in an effort to explore frameworks of integrating volumetric zoning and energy modelling, Samir proposed a procedure to link the morphological attributes of a form with its thermal behaviour [1]. His framework aimed to provide a feasible way for designers to examine the relationship between architectural forms and the thermal performance of the forms. It was later studied further by Janssen et al., who explored low energy design strategies with a factor in limitations and constraints of both passive and active systems being discussed [18]. Further investigations which look at niching of the genetic algorithm data are needed to

give a more diverse population in the result and thus give the architect more design options [19]. Timur Dogan has recently evaluated new plan typologies concerning exterior morphology and interior organization with their energy performance, but the focus has been on the comparison between real volumetric zoning and the ASHRAE zone method [20].

In the urban scale, Bill Hillier, from University College London in the late 1970s to early 1980s, conceived the term space syntax, which includes a set of theories and techniques for the analysis of spatial configurations [21]. This method later helps urban planners generate designs by breaking down form into components. Most recently, with the development of the Decoding Spaces tool kit [22] and other analytical and generative components in Grasshopper, the computational generation of street network, parcellation and building form based on the urban context and various design goals, become possible. Theresa Fink and Luc Wilson have tried to apply these tools to urban design workflow [23]. For example, the KPF urban interface [24] tests the computational urban design informed by environmental performance metrics. However, in the urban scale, there numerous unknown parameters that could affect the complex design process, such as long-term development, population increase, transportation, and climate change [24].

In this reviewed research, there are readily available algorithms and methods in different areas. But there remains a research gap, in building form scale, of how to integrate computational design platforms into the real design process to generate a more diverse and unique population of building volumetric options and test how this could provide the architect with a greater balance between

performative outcomes of a computational model and design independence.

4 A NEW WORKFLOW FOR HIGH-PERFORMANCE BUILDING FORM DESIGN

A new workflow is proposed here for a computational approach to generating architectural form and predicting their environmental performance. The aim is to deliver a parametric volumetric model that allows architects to explore and demonstrate realistic building attributes and determinants to assess their performative outcomes. The overall framework is divided into four steps: site setup, building form generation, performance evaluation and visualization, and design development. Through a residential building design test case in Vancouver, Canada, the paper evaluates how architects could adopt this workflow in the design process and maintain the feature throughout the whole preliminary design process (Figure 1).

4.1 Step 1: Site setup

In the first step, the existing urban data of the site is imported into Grasshopper in the form of SH files. Other site attributes, including weather, surrounding buildings, site boundary, and set back, are attained and inputted into Grasshopper as site data. This step helps designers collect climatic information and design constraints for later form generation. In this paper, the site is located at Yaletown, Vancouver, a residential community known for its beautiful surrounding landscape, high real estate value, and environmental issues such as summer overheating. The site constraints include 8-meter setbacks from the site boundary, and a maximum gross floor area of 18,000 m^2. The site geographical information is downloaded from the Vancouver Open Data Catalogue[25].The model below generally reflects the contextual features of the site (Figure 2).
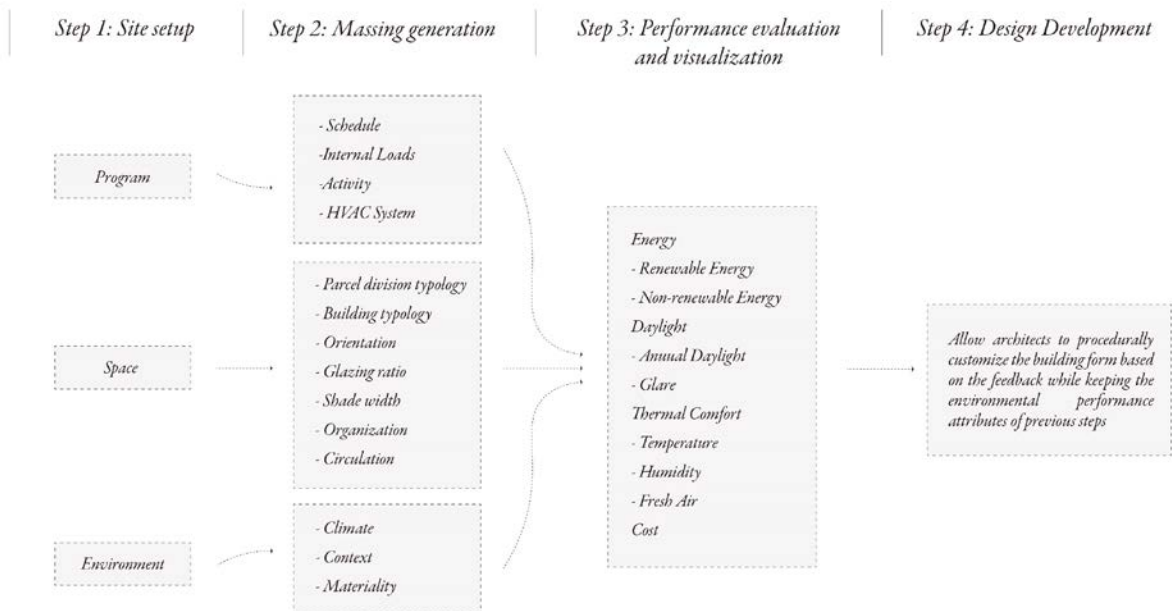
Figure 1. New workflow for high-performance building form design

Figure 2. Input site information and constraints

Figure 3. Parcel division typology

4.2 Step 2: Massing generation

In the second step, the building form is generated from the algorithm based on different variables, such as parcel division, building typology, orientation, glazing ratio and shade depth. Deploying some readily developed tools in Grasshopper, such as the Ladybug tool [26], Decoding Spaces [24] with customized scripts, this step is able to generate various types of building massing.

Three common residential building typologies are established, which includes block type, row type, and tower type. Block type represents low-rise residential townhouses along the street; row type represents mid-rise residential apartments; tower represents high-rises. Along with four parcel divisions methodologies, we received a total of twelve building typologies (Figures 3 and 4).

Building orientation influences street appeal, interior view, and how much solar radiation is captured by the building. Twelve different orientations are set up for each building typology. The basic building orientation is facing towards the south and is rotated incrementally by 30 degrees. However, since block building typology is not applicable for rotation, it only has one orientation (Figure 5).

Window to wall ratio (WWR) represents the percentage area determined by dividing the building's total glazed area by its exterior envelope wall area, which could influence the daylighting access, radiation impact, and heat loss. A varying WWR is studied at 30%, 40% and 50%. The same WWR is applied to each façade following residential building design convention.

External shades block out the extra sunlight and prevents solar radiation from hitting the window surface during summer. This helps bring down the temperature and reduce cooling loads, but it can also block out needed sunlight and increase the heating and lighting energy consumption. Especially under the context of global warming, the impact of shade needs to be studied. The shade simulated here means horizontal shade or outside balcony. The length of the shade input includes 0 and 1.2 meters (see Figure 6).
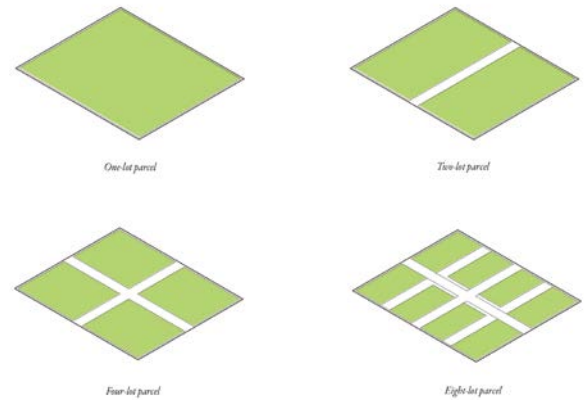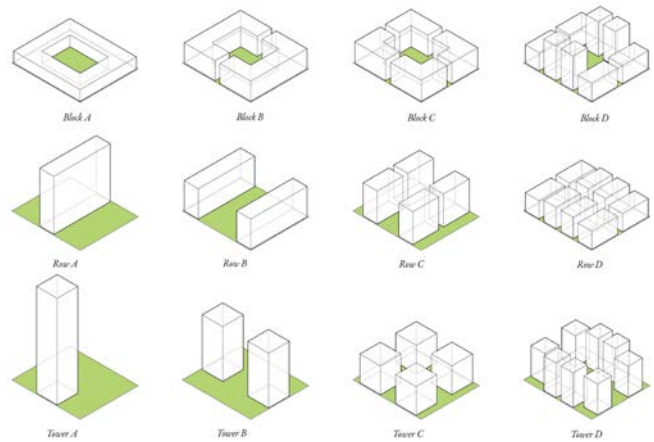
Figure 4. Building typology
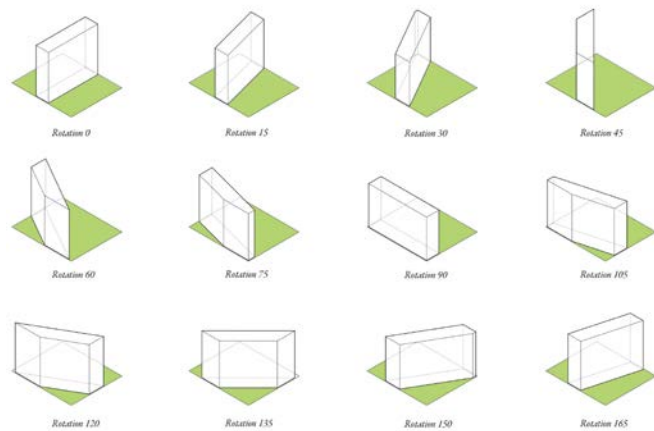
Figure 5. Orientation

352
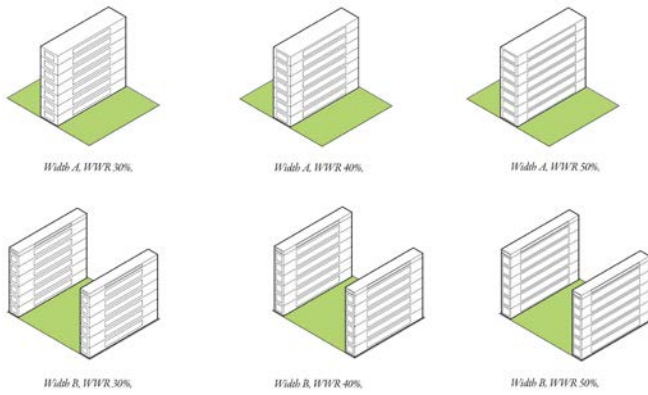
Figure 6. Glazing ratio and shade width

4.3 Step 3: Performance evaluation and visualization

In the third step, the building forms generated from the last step are transferred into performance simulation engines that undertake predictions of environmental performance. For example, EnergyPluscould be used as an energy simulation engine, whereas Daysimcould perform annual daylight simulation. This project intentionally limits the scope of performance metrics to solar radiation and other non-energy rubrics due to the limitation of time. However, customized performative metrics could be set up based on local codes and project targets. For instance, in British Columbia, Canada, specific Total Energy Use Intensity (TEUI) and Thermal Energy Demand Intensity (TEDI) targets should be included as per the BC Step Code [4].

The metrics set up in this design include three categories. This first one is the site requirement, which includes building footprint, building height, and gross floor area. These prerequisites are based on local building codes and building permit requirements. The second group is solar radiation, which contains useful solar radiation (when outdoor T<10°C), harmful solar radiation, (when outdoor T>22°C) and the surface to area ratio. This group is set up to measure solar radiation in both winter and summer. Surface to area ratio is a measure of how compact a building is, and is expressed as the ratio for the external surface area of the building to the treated floor area. The third category is comprised of occupant comfort metrics, such as window area of view quality (targeted area >10%) and public sunlight access. This category evaluates the non-thermal conditions of these design (Figure 7).

After determining the metrics, the algorithm runs the form options with a repeating procedure. All the possible choices are computationally tested under the same gross floor area with different footprint and heights. After running a parametric simulation of 5760 cases within 48 hours through several remote computers, the results are uploaded to a parallel coordinate platform. As shown in Figure 8, each case corresponds to a line on the top and detailed information in the bottom charts. This gives us the opportunity to visually analyses the large data set generated by parametric simulations and interact with the result.

Decision makers could eith use the parallel coordinate to filter out desirable choices or use the slider interface to compare different design input and desing outcomes.
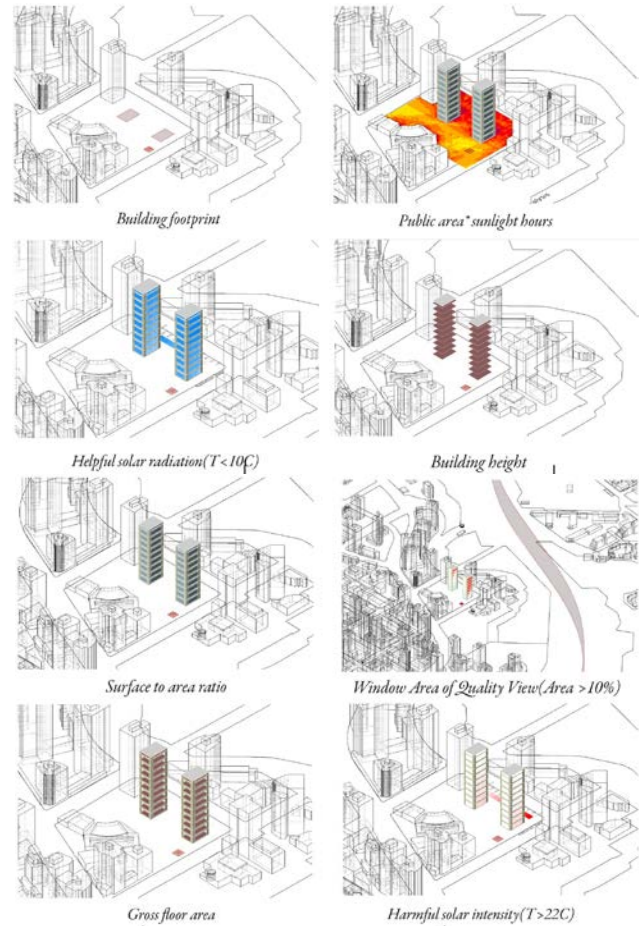
Figure 7. Eight **simulation metrics**
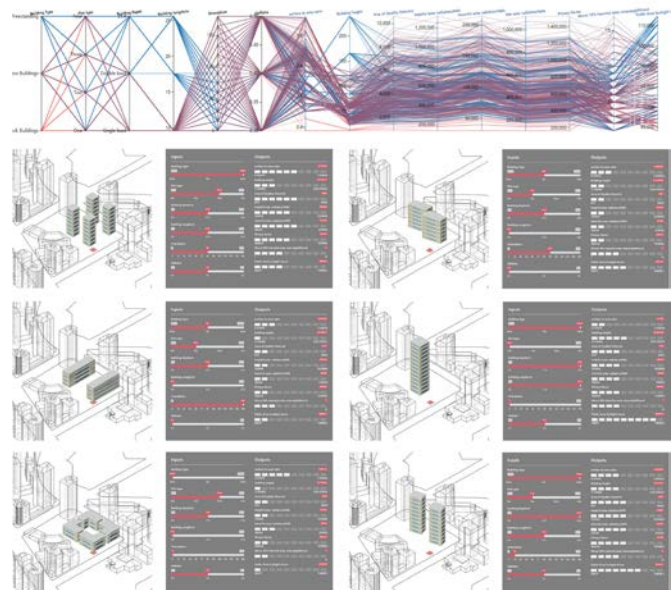
Figure 8. Parallel coordinate platform

353

4.4 Step 4: Design Development

In this step, further design operations are performed to develop the design in detail. The objective is to allow architects to procedurally customize the building form based on the feedback while keeping the environmental performance attributes of previous steps.

Based on the simulation results in Step 3, Two design options, which have relatively better performance in each category, are selected. The selection rubric is to improve every target instead of maximizing only one of them. As shown with detailed performance results in Figure 9, Option A is a row building with 0-degree rotation and Option B is a row building with a 30-degree rotation towards east. The view quality results for Option A and B are similar, but the solar radiation result for Option A is better than Option B, which means that Option B receives more solar radiation in cold winter and less radiation in warm summer. However, as the city requires to maintain an orthogonal urban grid between buildings, Option C is proposed to keep the shape of Option A but rotates the façade by 30 degrees to keep the solar benefits of Option B (Figure 10).
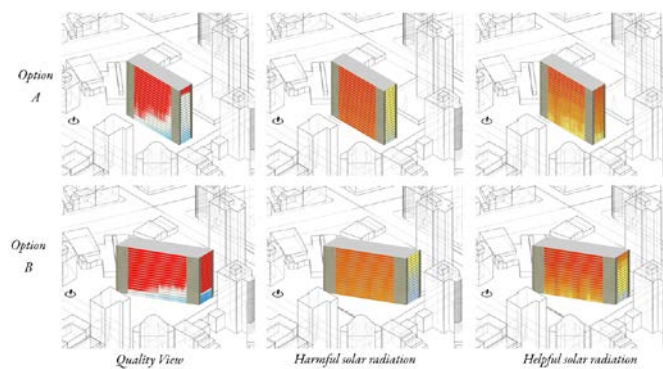
Figure 9. Compare Option A and B
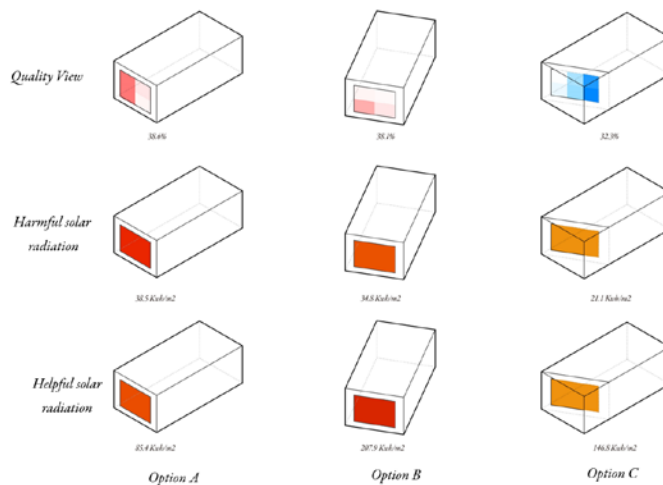
Figure 10. Compare Option A, B and C

In the following design process, some significant operations are made procedurally for the building form. Firstly, during the public hearing and meeting, the neighbourhood residents report a concern that the massing will offend the solar rights for a nearby square and block the neighbourhood view to the surrounding landscape. To make sure the existing solar access and visual link will not be blocked by the proposed massing, the solar envelope is generated through a backward solar boundary tracing method to trim the top part of the mass (Figure 11).
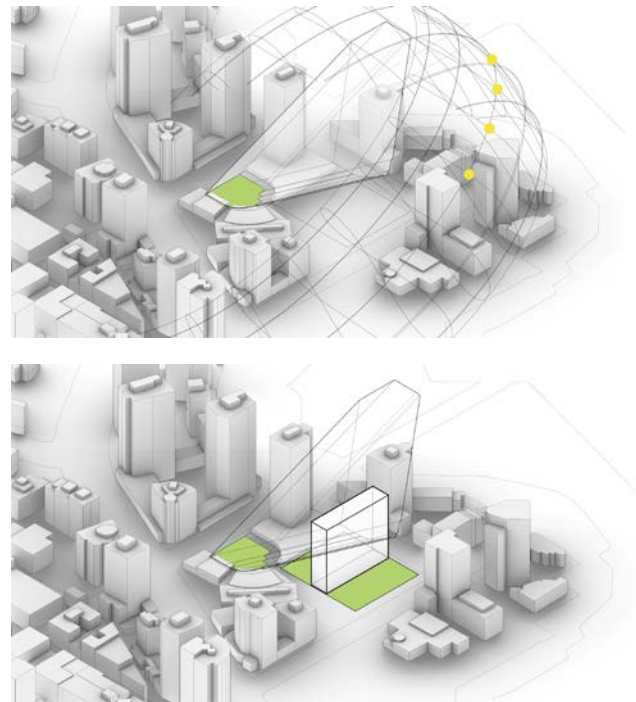
Figure 11. Trim the building to ensure solar rights

Secondly, the idea of bringing the vertical landscape to the midlle-part of the massing through including a sky garden is proposed, which significantly increases the visual link to views for the neighbourhood (Figure 12). Thirdly, the podium program is arranged based on the daylight requirement for different programs. For example, daycare and gym rooms are put on the south side of the podium, but storage and service space are arranged on the north (Figure. 13). Finally, the massing façade is modalized and rotated by 30 degrees based on the previous discussion of Option C.

5 CONCLUSION

This paper explores a procedural simulation workflow of building environmental performance, to facilitate systematic environmental analysis of architectural designs in a parametric demonstration manner. As presented, a parallel simulation and analysis method has been applied to the design process to help designers quickly explore a vast number of potential choices and perform strategic design solutions. The design investigates how simulation tools
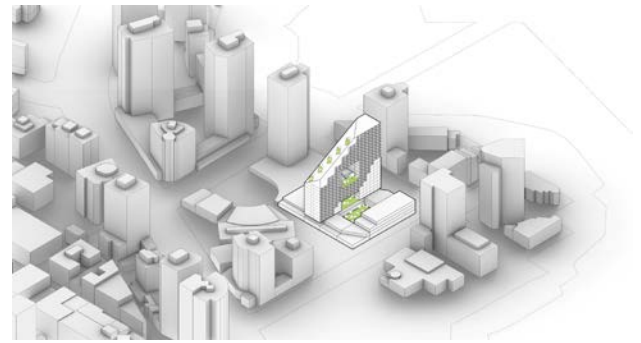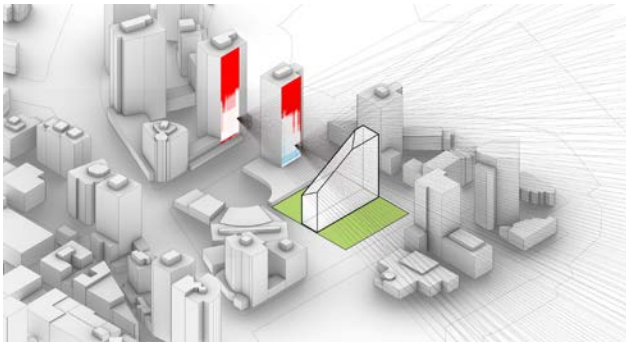
Figure 14. Modalize and rotate the facade
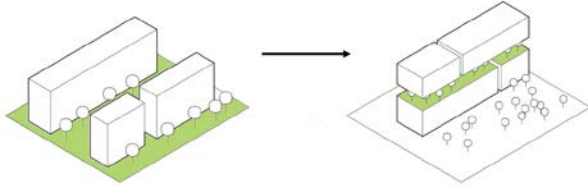
Figure 12. Sky garden to ensure the visual link

Figure 15. Street view rendering
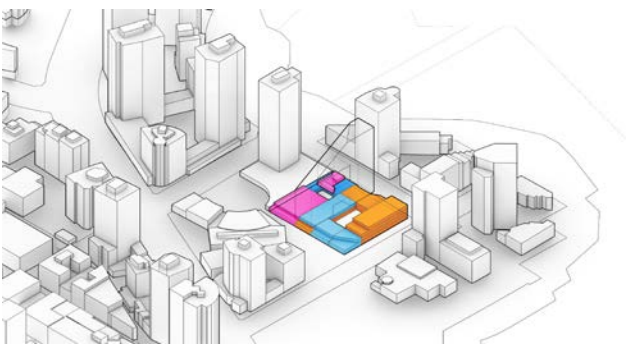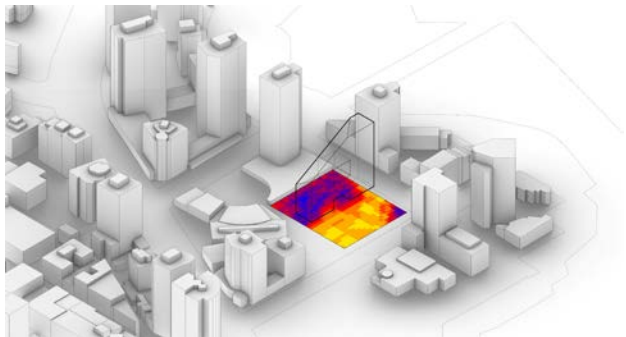
Figure 13. Program arrangement based on the solar requirement

Figure 16. Neighbourhood view rendering

could inform the designer's judgement and guide every move. Undoubtedly, if early simulation becomes part of the standard design process, architects will begin to understand the underlying interactions and will make active and elegant design actions with the confidence that they will lead to a better building (Figure 14 and 15). This research intentionally limits the scope of the relationship between building form design and non-energy performance in the massing design stage, due to the limitation of time. However, we should be aware that the result is limited to pre-defined parameters (selection criteria; typologies, degree of rotation), which might influence the result. Furthermore, research has shown that, in order to get more

accurate result in later design stage, massing method could not represent a high accuracy and resolution for energy simulation[20]. In further research, providing deeper information about the building form could be a focus such as introducing interior floor layout into the parametric simulation process. Also, investigation of the building program, construction material, human activity as well as other important topics is still regarded as necessary.

REFERENCES

[1] E. V. A. M. M. Samir S. Emdanat, "A Conceptual Framework for Integrating Morphological and Thermal Analyses in the Generation of Orthogonal Architectural Designs," 1996.

[2] A. Forty, *Words and buildings : a vocabulary of modern architecture*. Thames & Hudson, 2000.

[3] D. B. Crawley, J. W. Hand, M. Kummert, and B. T. Griffith, "Contrasting the capabilities of building energy performance simulation programs," *Build. Environ.*, vol. 43, no. 4, pp. 661–673, Apr. 2008.

[4] L. Governments, "BC Energy Step Code A Best Practices Guide for," no. July, 2019.

[5] R. and A.-C. E. American Society of Heating, *90.1 user's manual : ANSI/ASHRAE/IES standard 90.1-2010*. American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2011.

[6] C. Christensen, R. Anderson, S. Horowitz, A. Courtney, and J. Spencer, "Energy Efficiency and Renewable Energy Building Technologies Program T OF E N BEopt^{TM} Software for Building Energy Optimization: Features and Capabilities," 2006.

[7] M. Thalfeldt, E. Pikas, J. Kurnitski, and H. Voll, "Facade design principles for nearly zero energy buildings in a cold climate," *Energy Build.*, vol. 67, pp. 309–321, 2013.

[8] K. Lo, "A critical review of China's rapidly developing renewable energy and energy efficiency policies," *Renew. Sustain. Energy Rev.*, vol. 29, pp. 508–516, Jan. 2014.

[9] E. Rey, W. Frei, S. Lufkin, and D. Aiulfi, "An Innovative Architectural Strategy for the Integration of Energy Issues into the Design Process of a Commercial Center in Switzerland," no. June, pp. 83–95, 2014.

[10] "Robert McNeel & Associates. Rhinoceros 6.0," 2018.

[11] C. Cerezo, T. Dogan, and C. Reinhart, "Towards standarized building properties template files for early design energy model generation," *2014 ASHRAE/IBPSA-USA Build. Simul. Conf. Atlanta, GA*, pp. 25–32, 2014.

[12] CORE studio, "Second Colibri Release," 2017. [Online]. Available: http://core.thorntontomasetti.com/food4rhino-release-second-colibri-. [Accessed: 11-Jan-2020].

[13] S. Stevanović, "Optimization of passive solar design strategies: A review," *Renew. Sustain. Energy Rev.*, vol. 25, pp. 177–196, Sep. 2013.

[14] P. Philippou, "Building Types and Built Forms," *J. Archit.*, vol. 20, no. 6, pp. 1127–1138, 2015.

[15] P. Steadman, *Architectural morphology : an introduction to the geometry of building plans*. Pion, 1983.

[16] P. Steadman, *Building types and built forms*. Troubador Publishing, 2014.

[17] H. Shayesteh and P. Steadman, "Coevolution of urban form and built form: a new typomorphological model for Tehran," *Environ. Plan. B Plan. Des.*, vol. 42, no. 6, pp. 1124–1147, 2015.

[18] P. Janssen, K. W. Chen, and C. Basol, "Design Tool Development-eCAADe 29," 2011.

[19] K. W. Chen, P. Janssen, and A. Schlueter, "A Design Method for Multicriteria Optimization of Low Exergy Architecture Optimal Retrofit of Preserved Buildings (OReP)," 2013.

[20] T. Dogan, E. Saratsis, and C. Reinhart, "THE OPTIMIZATION POTENTIAL OF FLOOR-PLAN TYPOLOGIES IN EARLY DESIGN ENERGY MODELING Massachusetts Institute of Technology , Cambridge , MA 02139 , USA," *Build. Simul. Conf.*, 2015.

[21] F. E. Brown, *Space is the machine*, vol. 18, no. 3. 1997.

[22] A. Abdulmawla, S. Schneider, M. Bielik, and R. Koenig, "Integrated Data Analysis for Parametric Design Environment mineR : a Grasshopper plugin based on R Environment Analysis Parametric mineR : a Grasshopper plugin based on R," no. September, 2018.

[23] T. Fink, R. Koenig, and U. Weimar, "Integrated Parametric Urban Design in Grasshopper / Rhinoceros 3D Demonstrated on a Master Plan in Vienna," vol. 3, no. Feng 2009, pp. 313–322, 2018.

[24] L. Wilson, J. Danforth, C. C. Davila, and D. Harvey, "How to Generate a Thousand Master Plans : A Framework for Computational Urban Design," 2019.

[25] City of Vancouver, "Open data." [Online]. Available: https://vancouver.ca/your-government/open-data-catalogue.aspx. [Accessed: 11-Jan-2020].

[26] S. R. M. Mackey C., "The Tool(s) Versus The Toolkit," *Humaniz. Digit. Real.*, 2018.

Simulation-Based Design Optimization for Mixed-Use Recreational Buildings with TEUI, TEDI, and GHGI Targets

Jeanie Chan[1], Haobo Liu[1], Luisa Drope[1] and Andrea Frisque[1]

[1]Stantec Consulting Ltd.
Vancouver, Canada
{Jeanie.Chan, Haobo.Liu, Luisa.Drope, Andrea.Frisque}@Stantec.com

ABSTRACT

Designing buildings in Vancouver, Canada has become an increasingly challenging task due to the new Green Building Policy for Rezoning. This policy mandates that buildings must either meet select 3rd-party certifications (e.g. Passive House) or stringent Total Energy Use Intensity (TEUI), Thermal Energy Demand Intensity (TEDI), and Greenhouse Gas Intensity (GHGI) targets. In our work, we propose to utilize simulation-based tools to analyze design solutions for a new construction project in Vancouver to meet these performance targets. This paper presents how these tools were used in finding the design solutions for a mixed-use recreational facility. Specifically, we performed studies utilizing a whole building energy simulation and parametric simulations to optimize the building's envelope, operation, and heat recovery system to achieve its TEDI targets. Furthermore, we analyzed solutions to reduce the pool's energy use by applying Passive House design concepts.

Author Keywords

Building Comfort & Energy Performance; Whole Building Energy Simulation; Simulation-Based Design Tools and Methods; Parametric Energy Analysis, TEDI, TEUI, GHGI

ACM Classification Keywords

Simulation Tools; Building Performance Simulation; Building Simulation, Energy Efficient Building; Building Envelope; Building Design; Building Energy Saving; Buildng Design Optmization; Buiding Energy Modeling

1 INTRODUCTION

1.1 City of Vancouver's Green Buildings Policy for Rezoning

The City of Vancouver's Green Buildings Policy for Rezoning is part of the Greenest City 2020 Action Plan. This plan was developed from 2009-2011 by local experts and city staff. The goal for Green Buildings is to "lead the world in green building design and construction" [1] with the targets of reducing energy use and GHG emissions by 20% (compared to 2007 levels) and by requiring all buildings constructed in 2020 and after to be carbon neutral.

The Green Buildings Policy for Rezoning offers two paths: Near Zero Emissions Buildings or Low Emissions Green Buildings [2]. Near Zero Emissions Buildings requires the project to meet and apply for Passive House certification or a similar net zero standard. Alternatively, the Low Emissions Green Buildings involves the following building performance requirements: 1) LEED Gold Certification and 2) Performance Limits on the TEUI, TEDI, and GHGI. Currently, the policy has two sets of performance limits, those for buildings connected to a City-recognized Low Carbon Energy System and those not connected (Table 1), which is the most common.

Building Type	TEUI (kWh/m^2)	TEDI (kWh/m^2)	GHGI (kgCO$_2$/m^2)
Residential Low-Rise	100	15	5
Residential High-Rise	120	30	6
Office	100	27	3
Retail	170	21	3
Hotel	170	25	8
All Other Buildings	EUI 35% better than Building By-law energy efficiency requirement		

Table 1. Performance Limits for Buildings not Connected to a City-recognized Low Carbon Energy System

For mixed-use buildings, the TEUI, TEDI, and GHGI limits are a combined area weighted average of the performance limits for each building type. The portions of the building that have a TEDI target must also meet their combined TEDI target [2].

1.2 Review of Tools for Building Design Optimization

Energy analysis using simulation software can be very valuable throughout the design of new buildings to estimate the building's energy use based on the local climate characteristics, HVAC system choices, and geometry. The purpose of energy simulations is to develop an energy balance calculation of the various energy flows in the building. Knowing where the energy is used in the building allows one to identify opportunities to reduce the energy use in a design. Energy efficiency measures (EEMs) can be evaluated using energy analysis to quantify their impact on the building's TEUI, TEDI, and GHGI. The following section illustrates how energy analysis is used in practice during the building design process. A stand-alone energy

model can be used to evaluate more complex EEMs (e.g. pool control optimization) or a parametric simulation study can be used for simpler EEMs (e.g. envelope, lighting, HVAC).

Parametric Simulations for Simple EEM Analysis
In recent years, parametric simulations have made great strides forward in whole-building energy simulation. They are now routinely used in the building design process for new construction. They provide an opportunity to evaluate a vast array of design options and find the optimal solution for any given performance or cost targets.

Parametric simulations have also been used to analyze energy efficiency measures (EEMs). For example, Samuelson et at. [3] uses parametric simulations to analyze EEMs for a prototype multifamily residential building. The authors explored the impact of window-to-wall ratio, glass type, building rotation, building shape, and wall insulation. Using EnergyPlus and jePlus, Guiseppe [4] investigates the energy saving and life cycle economic impact of several EEMs (e.g. insulation types, window performance, heating systems, and thermal solar systems) for a typical Italian single family home. Attia et al. [5] developed a simulation-based decision support tool for the early stage design of zero-energy buildings. It utilizes EnergyPlus-based parametric simulations to optimize passive (e.g. orientation, geometric features, envelope properties) and active (e.g. HVAC, ventilation, photovoltaic, solar thermal) building elements. Lastly, Lee et al. [6] have created DEEP (database of energy efficiency performance) for commercial buildings in California using EnergyPlus. They have explored EEMs covering envelope, lighting, heating, ventilation, air-conditioning, plug-loads, and service hot water.

Stand-Alone Energy Model for Complex EEM Analysis
Indoor swimming pools are known to be high energy consumers and therefore have a great potential for energy savings. High energy use is due to the evaporation heat loss from the pool, high ventilation rates (e.g. 4-6 ACH) to maintain indoor air quality, and high ambient temperatures (e.g. 28-32ºC) required to achieve thermal comfort.

Various strategies to reduce pool energy use have been explored utilizing simulation-based tools. For example, Ribeiro et al. [7] uses ESP-r simulation model to analyze an approach to optimize control of the pool's ventilation and air conditioning systems. Environmental control variables (e.g. ambient air temperature, relative humidity) are optimized during the night, while normal pool conditions are maintained during the day to ensure the users' comfort conditions. Calise et al. [8] uses TRNSYS software to investigate the possibility to perform a retrofit of an indoor swimming pool centre by means of solar thermal collectors and heat pump technologies integrated with the existing plant. Katsaprakakis [9] examines different passive and active heating systems for swimming pools with a TRNSYS model. Passive systems include improving the swimming pools' enclosure and using insulating pool covers. Active

systems include a biomass heater, solar collectors, and a geothermal heat pump co-operating with vertical geothermal heat exchanger. Comino et al. [10] uses a TRNSYS model to determine the energy saving potential of utilizing a hybrid system (i.e. desiccant wheels and indirect evaporative coolers) to condition the air for buildings with high latent loads, such as pools.

1.3 Our Work
In our work, we have developed a process for finding cost-effective design solutions for retrofit or new construction projects [11]. Our process is designed to provide sufficient data and recommendations to inform decision makers of the best design measures or combination of measures to implement. This process builds upon the use of parametric simulations to optimize design by adding financial analysis and a user-friendly and interactive interface to analyze the parametric results.

In this paper, we extend our analysis from [12], where we have applied this process to find design solutions for achieving all three performance limits for a mixed-use recreational facility that includes a pool, gym, multipurpose, and residential spaces. This paper presents the advanced studies we have performed utilizing simulation-based tools to further optimize the building's design. Namely, we have analyzed various solutions to lower the building's TEDI.

2 METHODOLOGY
The purpose of this analysis is to find design solutions for a new construction building to meet the City of Vancouver's building performance targets. The following sections outline our approach to finding these solutions. We first create a building energy model and perform a heating and cooling load analysis. Then, we select the design measures to analyze and evaluate their impact using the energy model or Parametric Design Analysis.

2.1 Building Energy Model
An energy model of the proposed design is created based on the architectural, mechanical, and electrical drawings/documentation of the proposed design.

2.2 Heating and Cooling Load Analysis
The heating and cooling load breakdown of the proposed design was analyzed in order to determine the key energy consumers. This analysis gives us a good understanding of where we can potentially improve the building design with the aim of meeting the performance limits.

2.3 Design Measure Selection
The number of design upgrade measures imaginable are countless; however, many of them do not pass an initial viability test. To establish a list of plausible measures, we have implemented a selection strategy based on several objectives. The following objectives guide the measure selection process and form the basis of their evaluation:

- Provide best value based on professional judgement and experience
- Utilize proven technology

358

- Select measures based on ease of maintenance
- Minimize capital investment where appropriate
- Select measures that save energy in the largest energy consumers

2.4 Design Measure Evaluation

A stand-alone energy model can be used to evaluate more complex EEMs (e.g. pool control optimization) or a parametric simulation study can be used for simpler EEMs (e.g. envelope, lighting, HVAC).

Energy Model Analysis

The design measure is implemented into the energy model of the proposed design, and their impacts are evaluated via the change in results.

Parametric Design Analysis

Parametric Design Analysis (PDA) provides a review of the interactive effects of multiple measures implemented simultaneously. Using PDA limits the need to reduce the number of simulations to a manageable number as the only limitation is computational power. This enables the design team to estimate the impact of all individual design decisions on the overall building energy performance. The analysis consists of the following steps:

Parametric Simulation

Potential design options are created by combining all of the different design parameters that are conceivable for the project. This means that many design iterations are created, typically including envelope performance, HVAC system, lighting design, and any other design consideration of interest, which are combined in various permutations. Whole building energy simulations using each design iteration can be performed using parametric simulations in jePlus [13].

Parallel Coordinates Tool

The parametric simulations generate a very large number of results data as they efficiently perform hundreds or thousands of full individual energy models. To be able to gain information from the data and draw conclusions, the data needs to be processed with a tool suitable for large sets of multi-dimensional data. The parallel coordinates tool can present this data in an interactive way, where sets of parameters can be chosen to view an outcome of interest such as energy use intensity (EUI) or net present value (NPV). The parallel coordinates tool utilizes the open-source JavaScript library d3.parcoords.js.

Design Options Selection

Using the parallel coordinates tool, Design Options (i.e. measure bundles) are selected based on the ability of a combination of individual measures to meet the project objectives, such as, for example, EUI, TEDI or GHGI emissions targets. The power of the parallel coordinates tool lies in the ability to select the target outcome in the interactive data set and get, as a result, all the Design Options that would meet the performance targets.

3 RESULTS

The building analyzed in this study is a recreational facility with a residential tower. It is located in Vancouver, BC. The recreation building is a 14-storey building with a conditioned floor area of approximately 13,220 m². The building accommodates several space types such as a gymnasium, pool, fitness area, office, conference room, restroom, locker room, stairwell, lobby, corridors, residential suite, as well as other spaces to provide service to the building such as storage room, electrical and IT server room, and mechanical room.

The following sections show the analysis performed to find design solutions for achieving all three of the building's performance limits (i.e. TEDI, TEUI, GHGI).

3.1 Building Energy Model

An energy model was created to reflect the geometry and key design features of the preliminary design (Table 2). Figure 1 shows a 3D rendering of the energy model. IES Virtual Environment [14] is the software used to model this building.
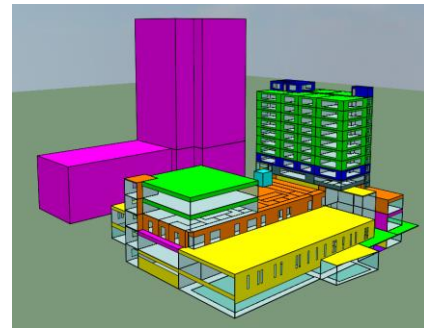
Figure 1. IESVE Model of the Proposed Design

Component	Proposed Design
Envelope	Roof: R-20.8 (RSI 3.7) Recreation Ext. Walls: R-7.3 (RSI 1.3) Residential Ext. Walls: R-5 (RSI 0.9) Windows: U-0.35 (USI 2.0)
Lighting	Fitness: 0.50 W/ft² (5.4 W/m²) Gym: 0.82 W/ft² (8.8 W/m²) Multipurpose: 1.07 (11.5 W/m²) Pool: 1.70 W/ft² (18.3 W/m²) Offices: 0.93 W/ft² (10.0 W/m²) Residential: 0.49 W/ft² (5.0 W/m²)
HVAC: Central System	Recreation heating/cooling: heat recovery chiller and backup gas-fired boiler Residential heating: gas-fired boiler
HVAC: Zone Level	Recreation: VAV with reheat Residential: hydronic baseboards
Heat Recovery	Recreation: heat recovery chiller, HRV Residential: HRV

Table 2. Key Features of the Proposed Design

Building Energy Use

Figure 2 show the energy use breakdown of the proposed design. The TEUI of the building is 279 kWh/m². The largest contributors to the TEUI are electric heating and cooling,

interior fans, and interior lighting. The GHGI is 13 etonnes $kgCO_2/m^2$. The TEDI for whole building and residential spaces are 246 kWh/m² and 29 kWh/m², respectively.
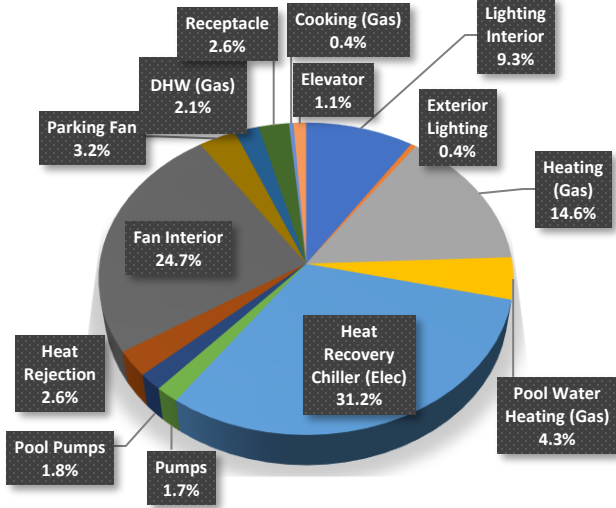
Figure 2. Energy Use Breakdown

CoV Performance Limits for Mixed Use Buildings
This building has both residential areas and non-residential areas, so it falls under the requirements for mixed-use buildings. This means that the TEUI, TEDI, and GHGI limits are a combined area weighted average of the recreational spaces and residential tower limits [2]. The residential tower also must meet the TEDI requirement for "Residential High Rise".

The recreational spaces of this building falls under the "All Other Buildings". Therefore, the performance limits of those spaces are determined based on the Vancouver Building By-law (VBBL) energy efficiency requirements. At the time of the rezoning application, the VBBL energy efficiency requirement is to comply with either ASHRAE 90.1-2010 or NECB 2011 energy code. An ASHRAE 90.1-2010 baseline model was created to determine the TEUI, TEDI, and GHGI targets for the recreational spaces. This baseline model only contains recreational facility space types and does not include the residential portion of the building. The TEUI performance limit for the recreational space types is 35% lower than the baseline model TEUI. The TEDI and GHGI performance limits are the baseline model TEDI and GHGI.

Table 3 show the performance limits and proposed design metrics for the residential spaces, recreational spaces, and whole building. The proposed design is meeting the TEUI and GHGI requirements, but not the TEDI requirements.

Metric	Performance Limits	Proposed Design
TEUI (kWh/m²)	367	279
GHGI ($kgCO_2/m^2$)	70	13
TEDI (kWh/m²)	201 (Building) 30 (Residential)	246 (Building) 29 (Residential)

Table 3. Performance Limits and Values for the Building

3.2 Heating & Cooling Load Analysis
A significant portion of the heating and cooling load is attributed to the pool space in the building. A lot of energy is used to heat the pool water, as well as heat, cool, and dehumidify the space. Figure 3 shows a breakdown of the heating load in the building. 58% of the heating load is from heating the pool water and space heating for the pool area. The space heating load of the building directly relates to the TEDI and GHGI.
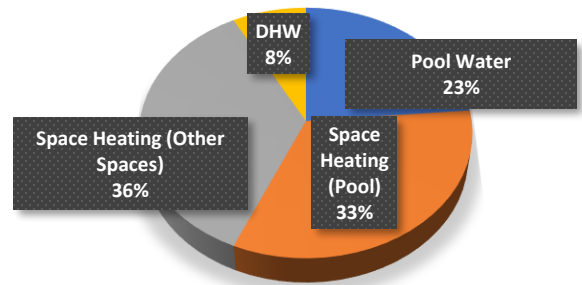
Figure 3. Heating Load Breakdown

This high heating load could potentially result in a high GHGI if a natural gas heating system is used; the eCO_2 emission factor for gas is significantly higher than electricity.

The pool space has such a high thermal demand because it is a large space with high air change requirements (i.e. 6 air changes per hour) and a high heating setpoint of 28°C. It also has high dehumidification load, which results in a significant reheat load. The building TEDI can be improved with additional heat recovery, improved envelope performance, and/or a reduction of the dehumidification load.

The pool water load is further broken down into losses due to convection, conduction, and evaporation (Figure 4). Currently, evaporation makes up 76% of the total pool water load. The high evaporative losses to the pool space result in a high dehumidification load.

Figure 5 shows the breakdown of the building's cooling load. As expected, the largest cooling load is the dehumidification and space cooling for the pool. This shows that there is a significant of heat recovery potential from cooling, where the heat rejected from cooling can be used for heating. The high cooling load affects the TEUI of the building as it results from a significant use of electricity.
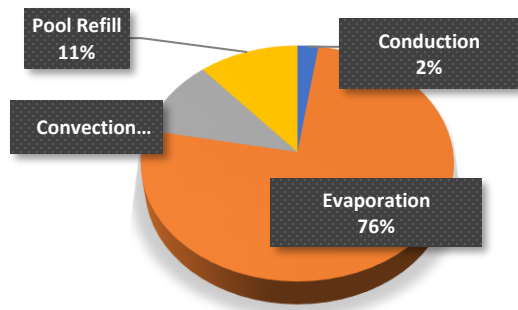
Figure 4. Pool Energy Breakdown
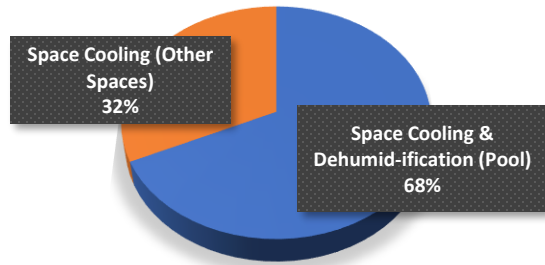
Figure 5. Cooling Load Breakdown

3.3 TEDI Reduction Analysis

This section presents the analysis and results of design solutions to meet the COV's building performance targets Namely, we analyzed design measures that would enable the building to achieve the whole building TEDI target of 201 kWh/m^2. We also explored opportunities to improve the aesthetic of the residential tower envelope while meeting the residential TEDI target of 30 kWh/m^2. The following three categories of measures were analyzed: 1) recreation facility, 2) pool space, and 3) residential tower.

Recreation Facility TEDI Reduction Measures

The measures in this category is specific to the recreation facility only. Since the recreation facility accounts for 65% of the building, these measures will have the greatest impact on the building TEDI value. The aim of these measures is to minimize the heat loss and recover energy wherever possible. The following measures were analyzed:

Decreasing Window Area

The window-to-wall ratio (WWR) of the proposed design is 57%. Heat loss through the windows significantly contribute to the building's TEDI value; however, daylighting is also important to occupant experience and capture solar gains. Therefore, the propose measure must strike a balance between those factors. Strategic placement of windows would also help to decrease window area and capture solar gains where they are the most abundant (e.g. south orientation). Therefore, in this analysis, we proposed an optimized window design that results in a WWR of 40%.

Addition of Pool Cover

Pool covers are an effective way of decreasing pool water evaporation losses. This in turn decreases the latent heat gain to the space, the dehumidification load, and the reheating required after dehumidification. Less heating required would reduce the TEDI. The drawback to this measure is that it can be expensive to implement as it requires the purchase of the cover system, maintenance of the system, additional training of the pool staff, and extra work time every day to deploy the cover.

Exhaust Air Heat Recovery

In order to maximize on the heat recovered from the exhaust air, heat reclaim coils are proposed to be added after the air-to-air heat recovery device. Since this does not reduce heating load of the building, it does not technically reduce the TEDI. Strictly speaking, only the effect of ventilation heat recovery devices is included in the TEDI. However, it can argue that in both cases, heat is extracted from the exhaust air stream and is put back into the building, either directly into the air or via heating coils. Therefore, we believe that the effect of exhaust heat recovery coils should be included in the TEDI. This alternative way of calculating TEDI will be confirmed with the City of Vancouver upon submission of the rezoning application.

Figure 6 shows the estimated impact of the proposed measures on the building's TEDI value. All three measures reduce the TEDI significantly but none of them individually are enough to meet the TEDI target. Therefore, we recommend that at least two of the measure be implemented into the design.
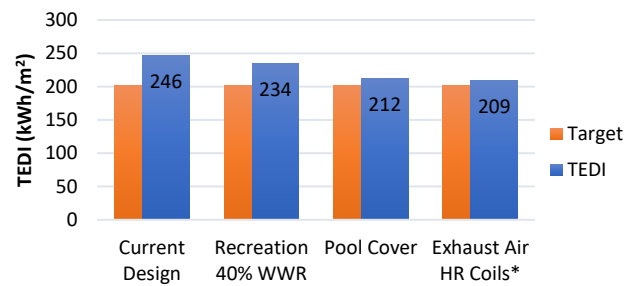
*Includes heat recovered from exhaust air heat recovery coils

Figure 6. Recreation Facility: Design Measure Impact on TEDI

Pool Space TEDI Reduction Measures

This category of measures focuses on the reduction of heating loads in the pool area. The dehumidification process cools the incoming outdoor air to a low temperature, which needs to be reheated before it's supplied to the space. Therefore, if more dehumidification is needed, more reheat is also needed, which increases the building TEDI. In this analysis, we have investigated Passive House design concepts that were applied in the Bambados Pool design [17,18] to lower the dehumidification load in the proposed design.

Condensation Issues in the Winter

The pool space is dehumidified to a relative humidity setpoint to avoid condensation at the windows during winter. This occurs when warm humid air is in contact with a cold

surface. This is especially of concern at the corner of the window frame, which is the coldest part of the window (Figure 7).
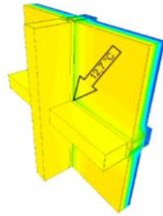
Figure 7. Coldest Point of the Window [16]

The dewpoint temperature (i.e. the temperature at which water condensates) is dependent on the surface temperature and the relative humidity of the air. At a particular surface temperature, there is a maximum relative humidity that the air can be before it starts to condensate. If the surface temperature is higher, the maximum relative humidity can also be higher. Therefore, we proposed to lower the dehumidification needed by increasing the inside surface temperature.

One common way of increasing the inside surface temperature is to continuously blow warm air onto the inside surface of the windows. This results in additional heating and fan energy, which is not captured in the energy model.

Passive House Design Concepts
A more energy efficient way of achieving a warmer inside surface temperature is to have better insulated windows. The Bambados Pool has a Passive House performance envelope, which includes tripled-glazed windows. The pool space temperature setpoint is 32ºC and the estimated surface temperature is 28ºC. Technically, with a Passive House performance window, the relative humidity setpoint for the space can be up to 80%; however, to prevent building envelope material damage, the relative humidity is set to a maximum of 64 %.

During operating hours, the indoor relative humidity is set to 58%. The setpoint is lower than the maximum due to considerations for occupant comfort. When the facility is closed, the air handling unit is switched off and the relative humidity can rise up to 64% depending on the outside air temperature. This has shown to significantly decrease the space and pool water heating load at the Bambados Pool. Furthermore, the increased humidity during unoccupied hours has shown to have the same effect of a pool cover with regards to decreasing evaporative losses. Maintaining the space at an even higher relative humidity was not possible as it resulted in damaging the building envelope.

Demand controlled ventilation was also implemented at the Bambados Pool. Airflow can vary as long as it can maintain the space temperature, relative humidity setpoints, and minimum outdoor air requirements. A minimum air change rate of 15% of the maximum design volume was provided to dehumidify the indoor air and remove pollutants.

Pool Space Design Measure
In the proposed design, the air temperature setpoint for the pool space is 28ºC and the highest relative humidity setpoint is 47% with the proposed windows (U-0.35/USI 2.0). With a Passive House performance window (e.g. U-0.14/USI 0.8), the inside surface temperature is warmer and the relative humidity setpoint can be 64%. Therefore, we propose to implement the following measures for the proposed design:

1. **Passive house windows:** Interior surface temperature of window will be warmer; therefore, the space can have higher relative humidity setpoint
2. **Humidification setbacks at night:** Humidity setpoint is 58% during the day and increased to 64% during unoccupied times. This will act as a "pool cover" and lower evaporative losses of the pool water.
3. **Variable airflow:** Airflow can vary between 15%-100% of the maximum rate (i.e. 6 ACH) as long as it can maintain the temperature and relative humidity setpoints.

The design measure results in an estimated whole builidng TEDI value of 231 kWh/m^2. Furthermore, this measure will provide additional fan energy savings as the supply airflow rates would be lower and there is no longer the need for dedicated blowers for the windows.

Residential TEDI Reduction Measures
The measures in this category is specific to the residential tower only. In addition to the TEDI target for the whole building, there is a TEDI requirement specific to the residential tower (i.e. 30 kWh/m^2). Currently the proposed design meets that target by achieving a TEDI value of 29 kWh/m^2; however, TEDI reduction measures were analyzed to explore opportunities to increase the building window area on the east and west façade. Table 4 shows the various options analyzed for the roof, external wall, windows, west WWR, east WWR, and heat recovery effectiveness. WWR for the south and north side is set to 30%.

Input Parameter	Options
Roof R-value (K·ft²/Btuh)	20, 30, 40 [RSI 3.5, 5.3, 7.0]
External Wall R-value (K·ft²/Btuh)	5, 7, 10, 15, 20 [RSI 0.9, 1.2, 1.8, 2.6, 3.5]
Window U-Value (Btuh/K·ft²)	0.25, 0.35, 0.45 [RSI 1.4, 2.0, 2.6]
West WWR (%)	30, 40, 50
East WWR (%)	30, 40, 50
Heat Recovery (% effective)	0, 75

Table 4. Residential: Design Measures

Figure 8 shows a full set of results of every combination of EEMs analyzed (i.e. 810 simulations). Each line represents one simulation. The first six axes starting from the left represent the input parameters that were varied in the

analysis. The last axis on the right shows the resulting TEDI value of the simulation. A Design Option is a specific combination of measures. TEDI values for the Design Options range from 10 kWh/m^2 to 84 kWh/m^2.
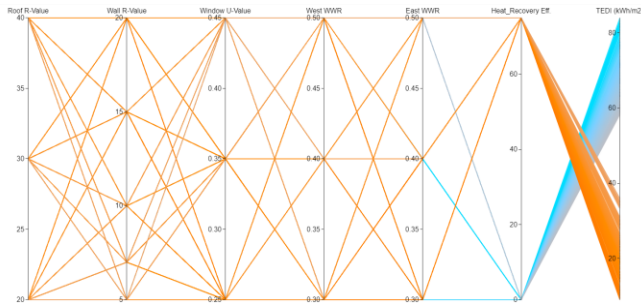
Figure 8. Parametric Design Analysis – All Results

Figure 9 shows the Design Options that meet the TEDI requirement for the residential tower (i.e. 30 kWh/m^2). From the selection options, we can see that heat recovery with a 75% effectiveness is critical to achieving the TEDI target.
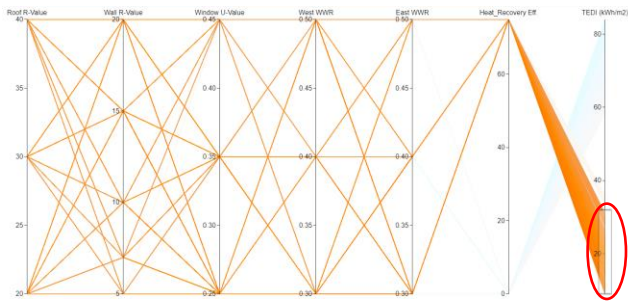
Figure 9. Design Options - Meets TEDI Requirement

Figure 10 shows the Design Options that meet the TEDI target and has the proposed roof and wall values (i.e. R-20 [RSI 3.5] and R-5 [RSI 0.9], respectively). This represents an ASHRAE 90.1-2010 code compliant roof and a typical metal framed external wall assembly with thermal bridging from balconies and other transitions considered. In this scenario, the window must have an overall U-value of U-0.35 (e.g. a high performing double-glazed metal framed window). This is easily achievable in the design; therefore, the proposed design for the envelope is sufficient to achieve the TEDI target.
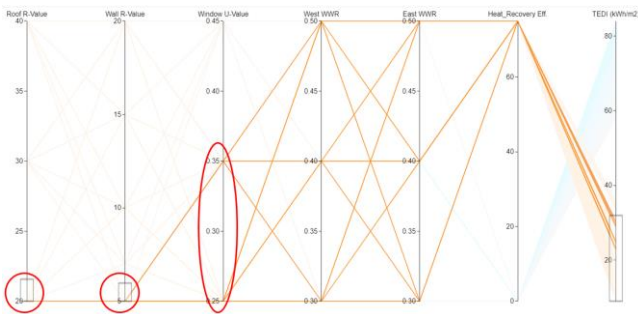
Figure 10. Design Options – Proposed Roof & Wall

Figure 11 shows the case with the proposed envelope and the high WWR values. The resulting TEDI value is 29 kWh/m^2, which meets the target. This shows that the WWR on the east and west façade can be up 50% while still meeting the TEDI target.
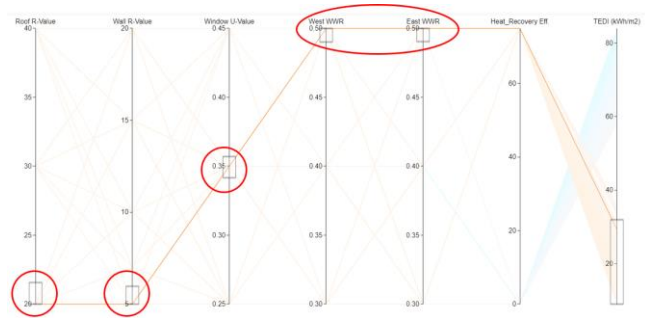
Figure 11. Design Options - Highest WWR

From this analysis, we conclude that:

1. To meet the residential TEDI target, HRV with effectiveness of 75% is required.
2. Acceptable Glazing Percentage:
   - Up to 50% WWR on east and west side
   - Up to 30% WWR on south and north
3. Proposed roof, wall, and window construction are acceptable.
4. Higher WWR values may be possible if thermal broken balconies are implemented.

4 DISCUSSION

After analyzing several TEDI reduction measures, we conclude that it is possible for the proposed design to meet its performance targets if a few of the measures are implemented. In addition to energy savings, life cycle cost implications and ease of implementation should also be considered before making final recommendations. A measure is considered cost-effective if the energy cost savings offset its capital, operational, and maintenance costs.

For the recreation facility, the WWR reduction measure would be the most cost-effective since walls are cheaper than windows; however, the aesthetic of the building would be significantly altered. The exhaust air heat recovery coils measure could potentially be cost-effective as the costs would only include the cost of implementation and minimal maintenance costs. Conversely, for the pool cover, the cost of implementation is low, but the operational costs would be high as it requires additional training of the pool staff, and extra work time every day to deploy the cover. Lastly, the pool space measure would be the least cost-effective as it would include implementing a Passive House envelope for the pool space, which would increase the capital costs dramatically. Furthermore, additional design, monitoring, and analysis efforts are required to ensure that there would be no condensation issues with the proposed control strategy and to further optimize operation.

For the residential tower, the proposed design already meets the residential TEDI target. However, additional design measures would help decrease the whole building TEDI. Similar to the recreation facility, decreasing the WWR would be the most cost-effective measure. However, that may decrease the desirability of the suites. Implementing a thermally broken balcony construction would improve the performance of the external walls and decrease the TEDI. This would increase the capital cost, but it does improve the building TEDI slightly or give the option of increasing the window area for the residential tower.

5 CONCLUSION

The new rezoning policy requirements has significantly changed building design practices in Vancouver. The new policy requires that the building design meet specific TEUI, TEDI and GHGI performance limits. Each limit is intended to drive the building design in a certain direction: The TEDI limit encourages the design to have a high-performance building envelope, a high level of airtightness, and heat recovery ventilation; GHGI favors building systems that use electricity over fossil fuels; and finally, the TEUI limit promotes overall low energy use.

These new targets have also added a layer complexity to the design process. Every measure that is considered must be weighted in terms of energy, thermal demand, GHG emissions, and financial implications. Energy analysis using simulation software is a useful tool is determining which EEM or combination of EEMs are the best for the project by considering all those metrics and ensuring that the optimal design can be achieved.

In our work, we successfully utilized energy analysis to analyze design solutions for a mixed-use recreational facility. We performed studies using a whole building energy simulation and parametric simulations to optimize the building's envelope, operation, and heat recovery system to achieve its performance targets.

REFERENCES

[1] City of Vancouver, "2020 ACTION PLAN: Part Two 2015-2020," Vancouver, Canada, 2015.

[2] City of Vancouver, "Energy Modelling Guidelines, Version 2.0," 2018.

[3] H. Samuelson, S. Claussnitzer, A. Goyal, Y. Chen, and A. Romo-Castillo, "Parametric energy simulation in early design: High-rise residential buildings in urban contexts," *Build. Environ.*, vol. 101, pp. 19–31, 2016.

[4] E. Di Giuseppe, "A parametric building design tool for assessing energy savings and life cycle costs," *Proc. Inst. Civ. Eng. Eng. Sustain.*, vol. 172, no. 6, pp. 283–292, 2018.

[5] S. Attia, E. Gratia, A. De Herde, and J. L. M. Hensen, "Simulation-based decision support tool for early stages of zero-energy building design," *Energy*

Build., vol. 49, pp. 2–15, 2012.

[6] S. H. Lee, T. Hong, M. A. Piette, G. Sawaya, Y. Chen, and S. C. Taylor-Lange, "Accelerating the energy retrofit of commercial buildings using a database of energy efficiency performance," *Energy*, vol. 90, pp. 738–747, 2015.

[7] E. M. A. Ribeiro, H. M. M. Jorge, and D. A. A. Quintela, "An approach to optimised control of HVAC systems in indoor swimming pools," *Int. J. Sustain. Energy*, vol. 35, no. 4, pp. 378–395, 2016.

[8] F. Calise, R. D. Figaj, and L. Vanoli, "Energy and economic analysis of energy savings measures in a swimming pool centre by means of dynamic simulations," *Energies*, vol. 11, no. 9, 2018.

[9] D. Al Katsaprakakis, "Comparison of swimming pools alternative passive and active heating systems based on renewable energy sources in Southern Europe," *Energy*, vol. 81, pp. 738–753, 2015.

[10] F. Comino, M. Ruiz de Adana, and F. Peci, "Energy saving potential of a hybrid HVAC system with a desiccant wheel activated at low temperatures and an indirect evaporative cooler in handling air in buildings with high latent loads," *Appl. Therm. Eng.*, vol. 131, pp. 412–427, 2018.

[11] J. Chan, R. Irwin, A. Frisque, and E. Dunford, "Cost-effective net-zero energy / GHG solutions for retrofit projects using parametric whole building energy simulations," in *eSim 2018 Conference Proceedings*, 2018, pp. 558–567.

[12] J. Chan, A. Frisque, and A. Jang, "Designing to TEDI , TEUI , and GHGI Performance Metrics," in *Proceedings of Building Simulation 2019: 16th Conference of IBPSA*, 2019.

[13] Y. Zhang, "Use jEPlus as an efficient building design optimisation tool," in *CIBSE ASHRAE Technical Symposium*, 2012.

[14] Integrated Environmental Solutions Limited, "IES Virtual Environment," 2019. .

[15] BINE Information Service, "Fun in a passive house pool," Bonn, Germany, 2016.

[16] J. Grove-smith and B. Krick, "Grundlagenuntersuchung der bauphysikalischen und technischen Bedingungen zur Umsetzung des Passivhauskonzepts im öffentlichen Hallenbad," Darmstadt, Germany, 2009.

Effects of Balconies on the Wind Pressure Coefficients of Naturally Ventilated High-Rise Office Buildings

Íris Maria Costa Fajardo W. Loche[1], Kauan Polli de Oliveira[1], Matheus Durante Oliveira[1],

Bruna Costa Fracalanza[1] and Letícia de Oliveira Neves[2]

[1]University of Campinas, School of Civil
Engineering, Architecture and Urban Design
Campinas, Brazil
{i229320, k177642, m203418, b167743}
@dac.unicamp.br

[2]University of Campinas, School of Civil
Engineering, Architecture and Urban Design
Campinas, Brazil
leticia@fec.unicamp.br

ABSTRACT

Natural ventilation and shading are effective bioclimatic strategies for buildings located in tropical regions. However, shading elements such as balconies also impact natural ventilation by influencing wind pressure coefficient (Cp) on building façades. This paper presents an investigation of the effects of balconies on the Cp values of high-rise office buildings, providing a detailed database of wind pressure coefficients. Wind tunnel tests were carried out for fifteen scenarios, considering an isolated building or surrounded by other buildings and combining balconies' location (side façades or frontal and back façades), its dimensions (width from 1 to 5 meters and depth from 0.5 to 2 meters) and 24 wind directions. Results showed that the provision of balconies impact on the Cp values and on the wind pressure difference between inlet and outlet openings (ΔCp) for both windward and leeward façades. For the first and middle floors, ΔCp values decreased when the balconies' depth increased and balconies 1-meter and 5-meter width showed higher ΔCp values than balconies 2.5-meter width. The upper floor showed lower Cp values when balconies' depth and width increased. Also, surrounding buildings highly impacted Cp values, leading to negative values for all façades, including the windward façade.

Author Keywords

Natural ventilation; wind tunnel; balcony; office building

1 INTRODUCTION

Natural ventilation and shading are effective bioclimatic strategies for buildings located in tropical regions. Balconies are horizontal overhanging structures used as fixed solar shading devices that behave as an eave to the lower floor, reducing the incidence of direct solar radiation and improving indoor comfort. Natural ventilation consists on a process that promotes air exchange between external and internal environment through building openings, assisting to provide users a health and comfortable room, without energy consumption. When natural ventilation is wind-driven, the existence of façade elements modifies the wind pressure

distribution on the building envelope, changing incident wind flow's characteristics [1], emphasizing the importance of knowledge of the pressure distribution on the building envelope to assess natural ventilation. The wind pressure on building surfaces can be evaluated through wind pressure coefficients (Cp), which characterizes air flow behaviour in buildings independently from the model scale [2]. Cp values can be obtained through existing databases or through primary sources. The latter considers building particularities and parameters that influence Cp values.

Wind pressure coefficients' databases for simplified building geometries are available on the literature, e.g. the AIVC database [3], the ASHRAE Fundamentals Handbook [4], TPU Aerodynamic database [5]. Analytical methods also estimate mean Cp values on a building surface, such as the model from Swami and Chandra [6] and software tools such as Cp Generator [3] and CpCalc+ [7]. All these sources include Cp values for low and high-rise buildings and most of them provide values for several wind directions, with exception of the AIVC database that presents 45º intervals for low-rise buildings and 90º intervals for high-rise buildings [8]. However, these databases and analytical methods are limited. Among the sources mentioned, only the software tools Cpcalc+ and Cp Generator consider the surrounding terrain type. When it comes to façade detailing and specific geometries (such as balconies), exposure conditions (such as wind direction, existence of surrounding buildings and adjacent obstacles) the use of simplified Cp values may induce to significant mistakes on the results from natural ventilation analysis [9].

Data from primary sources is considered more accurate than existing databases and analytical models since wind pressure coefficients are influenced by a wide range of factors including air flow conditions, urban surroundings and building geometry. One of the most reliable ways to provide primary data is through wind tunnel experiments under controlled turbulence conditions [2]. The wind tunnel experiments are performed using small-scale models and

adjusting wind direction and speed [10]. The results obtained through the experiments may also help to evaluate building thermal energy performance, by using Cp values as input data on building energy simulation (BES) programs such as EnergyPlus.

The EnergyPlus Airflow Network component allows the user to specify whether the wind pressure coefficients are input manually or automatically calculated by the program. To input Cp values manually, the user must determine the external nodes' height and input the Cp values for different wind directions. The airflow is calculated through the interpolation of the specified Cp values for each timestep and wind direction [11]. The Cp values automatically calculated by the software are valid only for rectangular buildings. For low-rise buildings, calculations are performed based on the model from Swami and Chandra [6] and for high-rise buildings, wind pressure coefficients are generated based on the ASHRAE Fundamentals Handbook [4].

Previous studies found that the provision of balconies significantly alters the wind pressure distribution on the building façade. Montazeri and Blocken [12] performed CFD simulations validated with wind tunnel for predicting mean wind pressure distributions on windward and leeward façades of an undefined typology five-storey building with and without balconies. The authors concluded that balconies can lead to strong changes in wind pressure distribution, since they introduce multiple areas of flow separation and recirculation across the façade. Chand et al. [13] analysed Cp values for a five storey-hostel building provided with balconies, varying wind direction from 0º to 45º. Results showed that the provision of balconies alters the wind pressure distribution on windward façades, subjected by wind direction and floor height, but does not introduce significant changes on the leeward side. Kotani and Yamanaka [14] carried out wind tunnel tests to evaluate wind pressure coefficients, also for a five-storey apartment building with and without balconies, varying the balcony type (banister and fence shape) and the wind direction. Results showed that the distribution of the wind pressure coefficients on the building façade depended mainly on the wind direction and not so much on the types of balconies. Castaño [1] performed wind tunnels experiments and concluded that the addition of a balcony on a single floor residence diminished the air flow up to 25%, leading to a small increase on indoor temperature of 0.1ºC.

These findings provide evidence of the variation in the wind pressure distribution on building façades due to the addition of balconies. However, detailed information on how balconies' dimensions, location and different wind directions influence the wind pressure distribution is still lacking, in particular for high-rise office buildings. In Brazil, office buildings were responsible for 14.3% of total energy consumption in 2018 [15], mainly due to intensive use of air conditioning systems. The use of solar shading devices (such as balconies) and natural ventilation may improve building thermal performance, reducing energy consumption. Therefore, this paper aims to provide a database containing wind pressure coefficient values for a high-rise office building with balconies, in order to provide input values for thermal and energy simulations. The database presents Cp values for an isolated building and for a building surrounded by other buildings, considering different balconies' dimensions, location and wind direction.

2 METHODS

This is an experimental research based on wind tunnel experiments.

2.1 Base Case

A base case was determined based on a database containing design parameters of 153 mixed-mode office buildings located in the city of São Paulo, Brazil [10] .The building's fixed parameters are shown on Table 1 and Figure 1.

Building	
Width	11 m
Length	22 m
Height	34.1 m
Number of floors	11
Floor	
Total area	242 m^2
Circulation area	48.4 m^2
Number of offices	4
Offices	
Area	48.4 m^2
Floor to ceiling height	3.1 m
Ventilation type	cross ventilation in adjacent façades

Table 1: Base case - fixed parameters
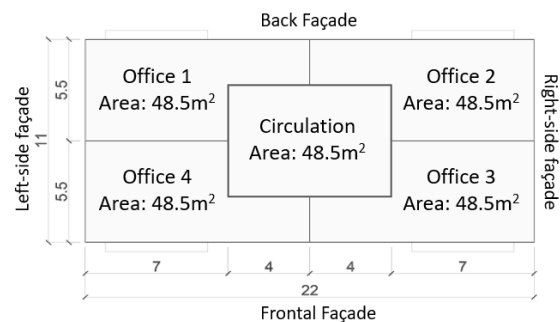
Figure 1: Base case floor plan

2.2 Scenarios

Fourteen scenarios were created combining the base case with variable parameters, such as balconies' location and dimensions and building's surroundings, based on balconies' configurations identified on the databased [10]. A scenario without balconies, named as reference, was created for results comparison (Figure 2, Table 2 and Table 4).
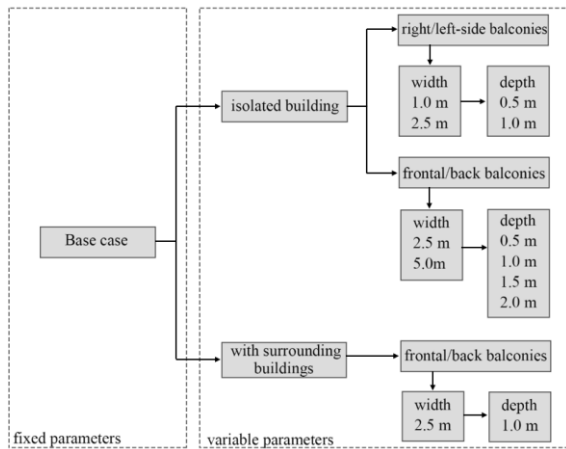
Figure 2: Scenarios flow diagram

Twelve scenarios were created considering an isolated building, with no interference from surrounding buildings (Table 2).

Scenario	Balconies		
	Location	Width (m)	Depth (m)
Reference	-	-	-
1	Right and Left-side façades	1.0	0.5
2		1.0	1.0
3		2.5	0.5
4		2.5	1.0
5	Frontal and back façades	2.5	0.5
6		2.5	1.0
7		2.5	1.5
8		2.5	2.0
9		5	0.5
10		5	1.0
11		5	1.5
12		5	2.0

Table 2: Scenarios - Isolated building

Two scenarios were created considering the building surrounded by other constructions. The surrounding urban design was defined based on the evaluation of the density profile of the region where the buildings from the database developed by Neves et. al [10] were located. Satellite images from Google Maps were used to define a density profile corresponding to 0.34 (ratio of constructed area to total area). The indentation between buildings and the urban design were defined based on the Building Work Code for the city of São Paulo [16] (Figure 3).

The interference of the surroundings on the wind pressure coefficients of the base case building was analysed for two different contexts. For results comparison, that the following assumptions were applied: the building in analysis would have the same variable parameters (Table 4) and both tested surroundings the same density profile (0.34), altering only the surrounding buildings' height. Context 1 was formed by surrounding buildings with the same height as the building in analysis. Context 2 was formed by different surrounding buildings' heights, defined by the height' standard deviation from the database developed by Neves et. al [10] (Table 3).

(a) Floor plan
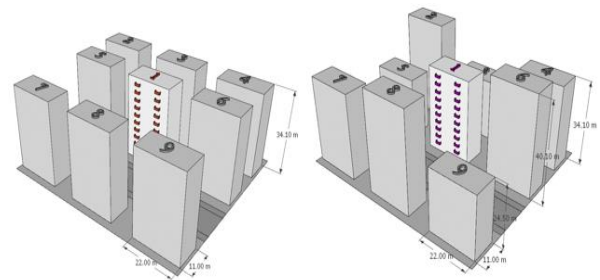
(b) Context 1   (c) Context 2

Figure 3: Surrounding buildings

Building	Height	
	Context 1	Context 2
1, 4 and 7	34.1 m	34.1 m
2, 6 and 8	34.1 m	40.1 m
3, 5 and 9	34.1 m	24.5 m

Table 3: Surrounding buildings height for context 1 and 2

Scenario	context	Balconies		
		location	width (m)	depth (m)
13	1	Frontal and back façade	2.5	1.0
14	2			

Table 4: Scenarios – Building with surrounding constructions

2.3 Wind tunnel experiments

The wind tunnel experiments were carried out at the Institute of Technological Research (IPT), São Paulo, Brazil. The wind tunnel is the biggest of Latin America, with a length of 40 meters and a test section of 6 m². According to Silva (1999), the maximum obstruction area at the test section

should be 7%. Assuming that, the isolated building (used for the reference model and scenarios 1 to 12) was modeled in a scale of 1:50 and the building with surroundings (used for scenarios 13 and 14) was modeled in a scale of 1:100. Thus, the models obstructed approximately 5% of the test section area.

(a) 1:50 scale model   (b) 1:100 scale model

Figure 4: Scale models used on the wind tunnel experiments

The models were built using 3 mm and 6 mm MDF panels. Sixty-four pressure sensors with ±0.05% full scale accuracy after rezero [18] were positioned in two façades of the model – 48 sensors on the façade with balconies and 16 on the adjacent façade (Figure 5). Since the model is symmetric, all the four façades could be evaluated by orbiting the model in 180º. Measurements were taken every 15º, corresponding to 24 wind directions for each building façade.
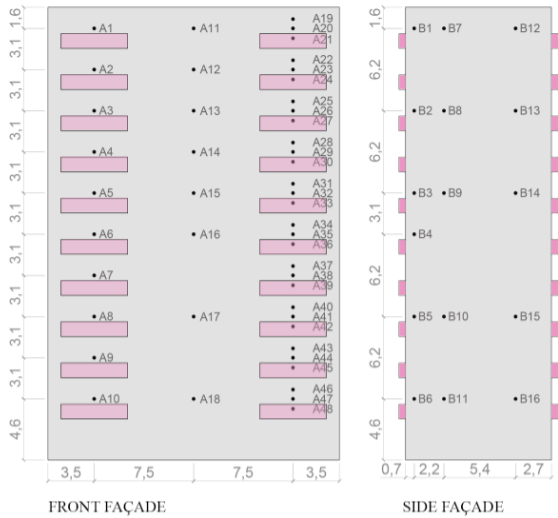
Figure 5: Scenario 9 as an example of sensors position.

The roughness category simulated in the experiments were verified by measuring the wind speed at different heights, by using a hot-wire anemometer sensor. Considering the building's height in analysis (34.1 m at full-scale), the roughness (z0) was 0.64 m for the scenarios 1 to 12 (considering an isolated building) and 0.75 m for the scenarios 13 and 14 (considering the surrounding buildings), representing the roughness category IV, (ABNT, 1988) which consists on terrains covered by several obstacles with little space in-between them (figure 6).
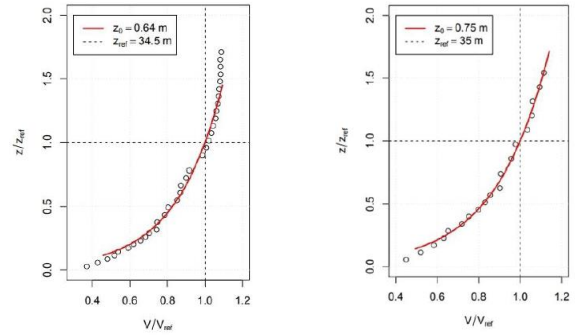
(a) Scenarios 1 to 12   (b) Scenarios 13 and 14

Figure 6: Velocity profile obtained through wind tunnel tests

To create the atmospheric boundary layer, the wind tunnel was provided with roughness elements and spires (Figure 7).

Figure 7: Wind tunnel interior: roughness blocks and spires.

To precisely determine wind pressure coefficients, the environment inside the wind tunnel must be controlled. Turbulence conditions must be assured within the tunnel to guarantee that the flow characteristics around the building are independent of the Reynolds number ($Re > 10^5$) [2]. Tests were performed rotating clockwise the model, from 0º to 45º, in order to ensure a turbulent flow inside the wind tunnel during the experiments. The wind velocity inside the tunnel was increased gradually until the independence of the Reynolds number was achieved, which occurred at 350 rpm (approximately 15 m/s). Also, a beehive-like structure located in the air entrance of the tunnel also contributed to provide a turbulent flow.

2.4 Wind Pressure Coefficients (Cp)

According to Blessmann [17], the wind pressure coefficient (Cp) consists on the ratio between the dynamic pressure at a certain point of the building surface (P_e) and the undisturbed air flow dynamic pressure (P_d):

$$C_P = \frac{P_e}{P_d}$$

Then, to obtain C_P, P_e and P_d, the following equation was used:

$$C_P = \frac{P_{T\,Mod} - P_{E\,Mod}}{f(P_{T\,Ref} - P_{E\,Ref})}$$

The model static pressure ($P_{E\,Mod}$) was measured by a piezometric ring, which calculates the average static pressure

around the test section. The reference total pressure $P_{T\,Ref}$ and the reference static pressure $P_{E\,Ref}$ were obtained using a Pitot Tube. The correction factor for the dynamic pressure (f), which consists on the ratio between the dynamic pressure at the height of the model and the reference dynamic pressure, was 0.79 for the isolated building and 0.52 for the building with surroundings.

Data analysis and graphic workspace software Origin 8 was used for displaying the Cp values obtained by the wind tunnel experiments.

3   RESULTS AND DISCUSSION

A database, containing detailed wind pressure coefficient values for all 15 scenarios and 24 wind directions, was developed to provide full access to all the results, available at:http://www.fec.unicamp.br/~conforto/trabalhos/iris_loche/index.php. This database is a useful source to future research studies and to evaluate natural ventilation through the Airflow Network (AFN), usually coupled with Building Energy Simulation (BES) software, such as EnergyPlus.

To evaluate the impact of balconies on natural ventilation, the difference between inlet and outlet openings' wind pressure coefficient values (ΔCp) obtained in the experiments was calculated and compared. One wind direction was chosen for each scenario to evaluate the impact of balconies for the windward and leeward façades (90º for balconies at right and left-side façades and 0º for balconies at front and back façades). Also, the first, the middle and the upper floors were selected to perform the analysis (Figure 8).

(a) 1- Floor plan – scenarios 1 to 4

(b) Floor plan – scenarios 5 to 14   (c) Selected floors highlighted
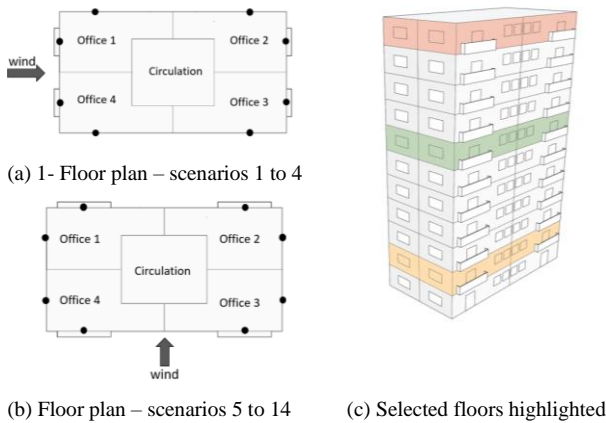
Figure 8: Floor plans, location of Cps in analysis and selected floors.

a) Balconies located at right and left-side façades

The models with balconies located at side façades are represented by the scenarios 1 to 4 (Table 2 and Figure 7).

We chose the wind direction of 90º to observe the wind pressure coefficients' behaviour. Figure 9 shows that Cp values are higher for the scenario without balconies, not only in the windward façade, but also in the sideward and leeward façades of the building. In all the scenarios, Cps have higher values on the upper floors.

Figure 9 shows that offices 1 and 4, located at the windward façade, presented higher ΔCp values than offices 2 and 3, whose openings are located at the leeward and sideward façades. The first floor showed lower values than the middle and upper floors. The first and middle floors showed a higher impact of the balconies' depth than of the balconies' width, since values decreased when the balconies' depth increased. However, on the upper floor, the balconies' width also impacted the ΔCp values, which decreased together with the increase of the balconies' width and depth for the scenario 4. Scenarios 1 and 3, both with balconies 0.5 m depth, showed higher ΔCp values for the first and middle floors for all scenarios, including the reference case, with no balconies. Conversely, in the upper floor, the reference case and the scenario 1 showed the highest values of ΔCp, which decreased when the balconies' depth and width increased.

Figure 9: Wind direction 90º - Wind pressure coefficient values for windward and leeward façades (a) reference model; (b) scenario 1, (c) scenario 2, (d) scenario 3 and (e) scenario 4.
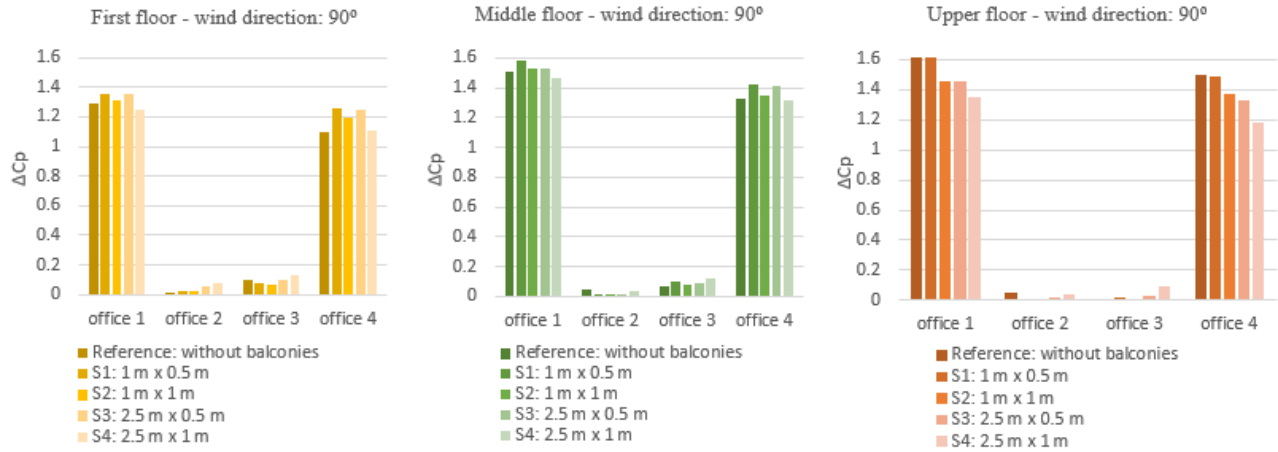
Figure 10: ΔCp for scenarios with balconies located at side façades and wind direction of 90°.

b) Balconies located at front and back façades

The models with balconies located at front and back façades are represented by scenarios 5 to 12 (Table 2 and Figure 7). We chose the wind direction of 0° to observe the wind pressure coefficients' behaviour. Alike scenarios 1 to 4 (Figure 8), Figure 10 shows that Cp values are higher in all façades for the scenario without balconies and on the upper floors for all other scenarios. Figure 11 shows that offices 3 and 4, located at the windward façade, presented higher ΔCp values than offices 1 and 2, whose openings are located at the leeward and sideward façades. The middle floor showed higher values than the first and upper floors. The impact of varying the balconies' depth was higher than the balconies'

width, since values decreased when the balconies' depth increased. However, on the upper floor, the balconies' width also impacted the ΔCp values, which decreased with the balconies' width increase. Scenarios 5 and 9 (balconies 0.5 m depth) showed the highest ΔCp values for the first and middle floors. Conversely, for the upper floor the scenario with no balconies showed the highest ΔCp values, which decreased when the balconies' depth and width increased. Higher ΔCp values were obtained for the first floor of offices 1 and 2, located at the leeward and sideward façades. Higher ΔCp values were achieved for the scenario with no balconies and for the scenarios with shallow balconies.

Figure 11: Wind direction 0° - Wind pressure coefficient values for windward and leeward façades (a) reference model; (b) scenario 5, (c) scenario 6, (d) scenario 7, (e) scenario 8, (f) scenario 9, (g) scenario 10, (h) scenario 11 (i) scenario 12.

370

First floor- wind direction: 0°

Middle floor- wind direction: 0°

Upper floor- wind direction: 0°
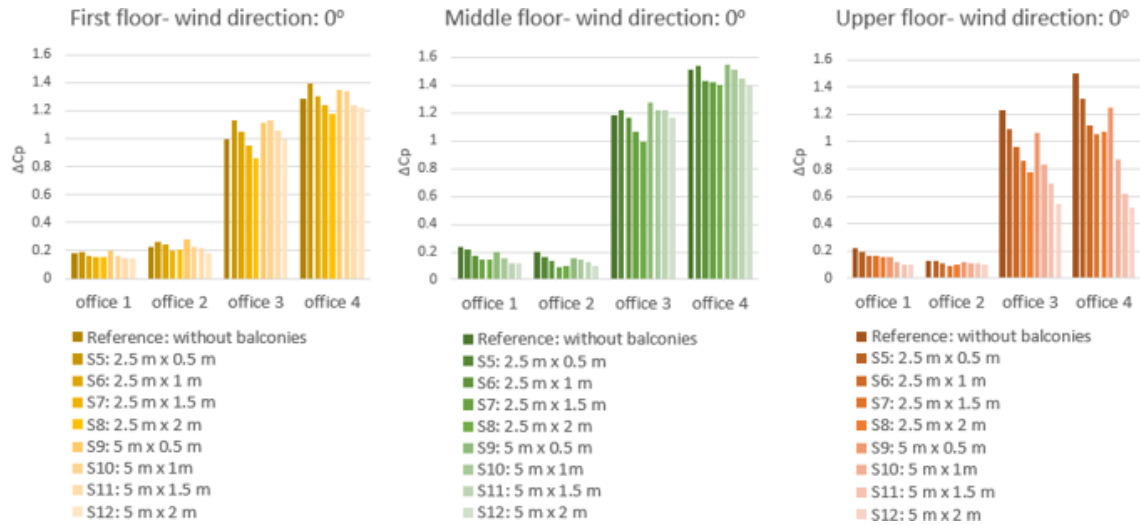
Figure 12: ΔCp for scenarios with balconies located at frontal and back façades and wind direction of 0°.

3.1 Building with surroundings

The models with surrounding buildings are formed by longitudinal balconies and represented by scenarios 13 and 14 (Tables 3 and 4). We chose the wind direction of 0° to perform the analysis. Figure 13 shows that the surrounding buildings caused a decrease in the wind pressure coefficients, showing negative values even on the windward façades. Figure 14 shows that the surrounding buildings' height influences directly on the ΔCp values on windward and leeward façades. Offices 1 and 3 showed higher ΔCp values for scenario 14, mainly in the first floor, due to the influence of buildings 3, 4 and 9, which are smaller in scenario 14 than in scenario 13. Offices 2 and 4 showed higher ΔCp values for scenario 13, since buildings 6 and 8 are smaller.

4 CONCLUSION

A series of wind tunnel tests have been conducted to investigate the effects of balconies on the wind pressure coefficient values of high-rise office buildings isolated and with surroundings, by combining balconies locations and dimensions for 24 wind directions. From this study, the following conclusions can be obtained:
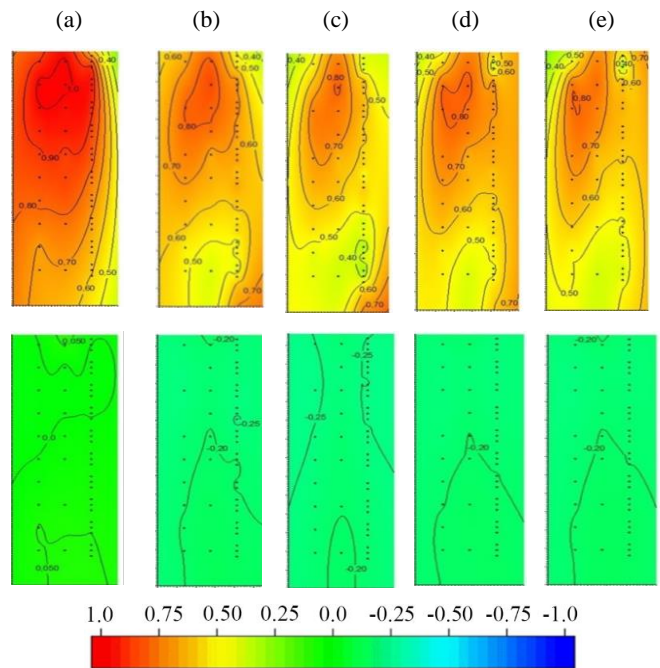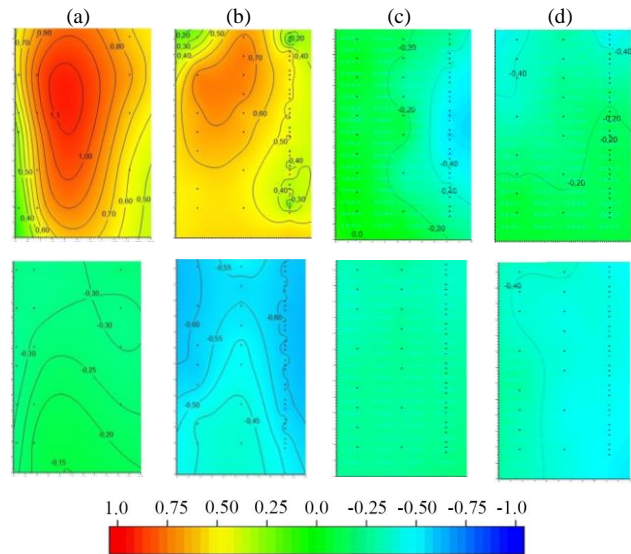

Figure 13:Wind direction 0o - Wind pressure coefficient values for windward and leeward façades (a) reference model; (b) scenario 6, (c) scenario 13, (d) scenario 14.
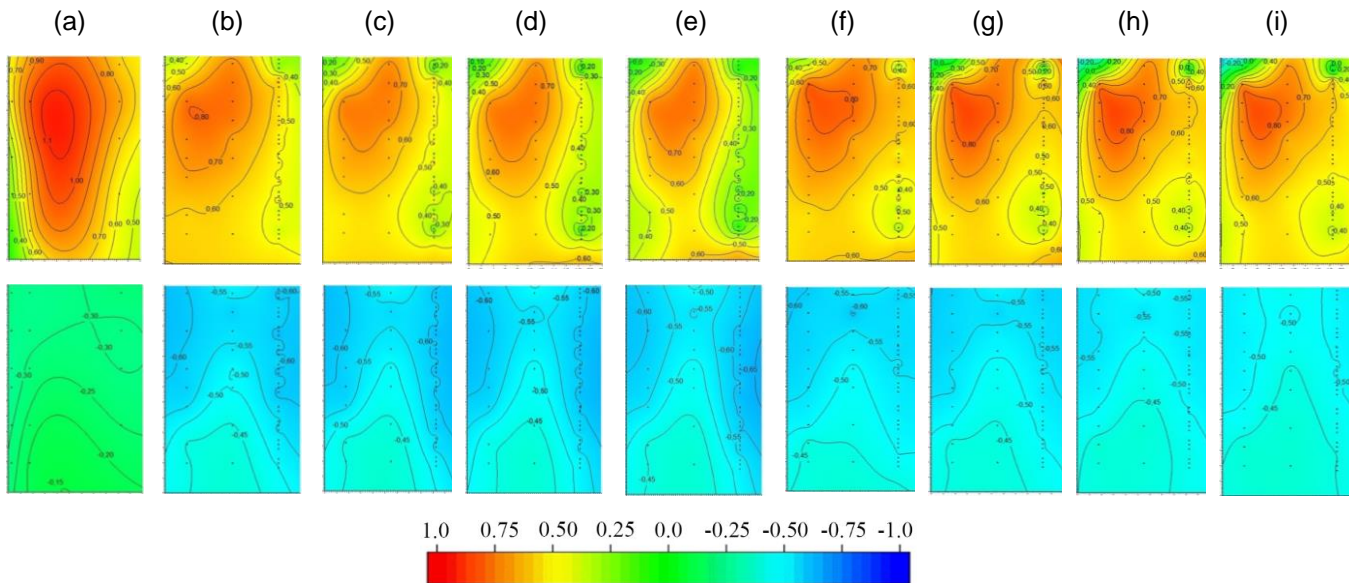

First floor- wind direction: 0°

Middle floor- wind direction: 0°

Upper floor- wind direction: 0°
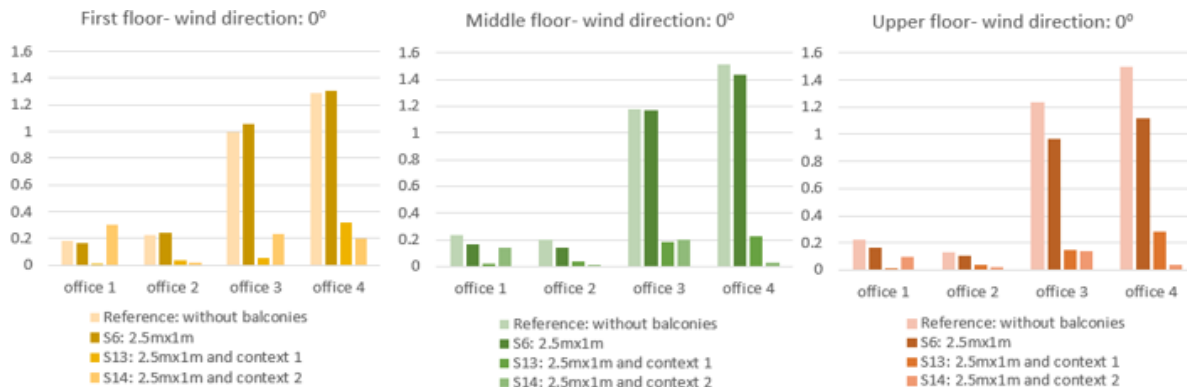
Figure 14:ΔCp for scenarios with surrounding buildings and wind direction of 0°

371

The provision of balconies showed high impact on the wind pressure coefficient (Cp) values for both windward and leeward façades. The difference between inlet and outlet openings' wind pressure coefficient values (ΔCp) for all the scenarios showed higher values for the offices containing openings located at windward façades than the ones with openings located at the leeward façade.

The middle and upper floor showed higher Cp and ΔCp values than the first floor. The upper floor showed high impact when balconies' depth and width increased, resulting lower ΔCp values for deeper and larger balconies. For the first and middle floor, ΔCp values also decreased when the balconies' depth increased and balconies 1-meter width and 5-meter width showed higher ΔCp values than 2.5-meter width balconies.

The distribution of wind pressure coefficient and ΔCp values is highly influenced by surrounding buildings. When surrounding buildings were included, results showed negative Cp values, even for the windward façade. The ΔCp values also decreased substantially when surrounding buildings were considered, mainly when influenced by the higher buildings surrounding the building in analysis. These results reinforce the importance of considering the context in which the building is located when analyzing natural ventilation.

This paper provides a database containing all the results obtained by the wind tunnel experiments, elaborated with detailed wind pressure coefficient values for all the scenarios, according to 24 wind directions. The wind pressure coefficient values available on the database are useful to further the researchers access to natural ventilation, providing input data for simulations in Building Energy Software (such as Energy Plus) reducing results uncertainties and encouraging the use of natural ventilation in new constructions.

REFERENCES
1. Castaño, H. F. M. *Impacto de dispositivos de sombreamento externos e muro na ventilação natural e no desempenho térmico de uma habitação de interesse social térrea.* USP, São Paulo, 2017.

2. Neves, L. D. O. 'Chaminé solar como elemento indutor de ventilação natural em edificações', Universidade Estadual de Campinas, 2012.

3. Knoll, B. Phaff, J. C and Gids, W. F. Pressure Simulation Program. *16th AIVC Conf.*, 1995

4. ASHRAE, 2005. *ASHRAE Handbook: Fundamentals. American Society of Heating, Refrigeratin and Air Conditionig Engineers.* Atlanta, 2005.

5. TPU. *TPU Aerodynamic Database.* Tokyo Politechnic University, 2016.

6. Swami M. V. and Chandra, S. Correlations for pressure distribution on buildings and calculation of natural-ventilation airflow. *ASHRAE Trans.*, vol. 94 (1988).

7. Grosso, D. Marino, M. CpCalc: a program to calculate wind pressure coefficients on a building envelope. *Politecnico di Torino*, 1996. http://porto.polito.it/2579969/. As of 8 Jan 2020.

8. Cóstola, D. Blocken, B. and Hensen, J. Overview of pressure coefficient data in building energy simulation and airflow network programs. *Build. Environ.*, 44, 10 (2009), 2027–2036.

9. Leite, R. *Cidade, Vento, Energia: Limites de aplicação da ventilação natural para o conforto térmico face à densificação urbana em clima tropical úmido.* USP, 2015.

10. Neves, L. Manoel, R. Chvatal, K. and Santesso, C. 'Envelope design of mixed-mode office buildings: theory versus practice. *PLEA Proc. (2017),* 2632–2639.

11. EnergyPlus v9.1.0. Input Output Reference, *Encycl. Ref. to EnergyPlus Input Output*, 2010.

12. Montazeri, H. and Blocken, B. CFD simulation of wind-induced pressure coefficients on buildings with and without balconies: Validation and sensitivity analysis, *Build. Environ.*, 60, (2013), 137–149.

13. Chand, I. Bhargava, P. K and Krishak, N. Effect of balconies on ventilation inducing aeromotive force on low-rise buildings. *Build. Environ.*,33, 6 (1998), 385–396.

14. Kotani, H. and Yamanaka, T. Wind pressure coefficient and wind velocity along building wall of apartment building with balcony. Proc. *IAQVEC* 1, (2007), 147–152.

15. Empresa de Pesquisa Energética. Balanço Energético Nacional. *Relatório 2018, ano 2019*, 2019.

16. São Paulo (Município). 'Decreto n° 57.521', 2016. https://leismunicipais.com.br. As of 08 Jan 2020.

17. Blessman, J. *Aerodinâmica das construções*. URGS. Porto Alegre, RS, 1983.

18. Sensor Solutions. *DTC INITIUM, Network Windtunnel Electronic Pressure Scanning Data Acquisition System.* Hampton, USA, 2016.

Radiation Modeling Strategy for Incorporating Vegetation in Urban Microclimate Simulations

Sarith Subramaniam[1], Mili Kyropoulou[2], Sabine Hoffmann[1]

[1]Technische Universität Kaiserslautern
Kaiserslautern, Germany
sarith@rhrk.uni-kl.de

[2]HKS Inc.
Houston, Texas, USA
mkyropoulou@hksinc.com

ABSTRACT

The presence of vegetation is known to improve outdoor thermal comfort perception and mitigate the effects of Urban Heat Island in cities. The availability of appropriate amount of solar radiation is a critical factor in deciding the selection and placement of vegetation in the form of food crops, turfgrass and ornamental plants. While the selection of vegetation species based on empirically measured radiation is common in agricultural practice and research, there have been limited efforts to leverage radiation modeling to guide urban greening decisions.

This paper proposes a radiation modeling strategy to inform the selection of plants for the purposes of urban greening. The authors review the current practices of modeling photosynthetic radiation prevalent in agricultural research, that are based on the concept of Daily Light Integral (DLI). An annual simulation-based approach for mapping DLI on to urban spaces is formulated and then discussed in the context of an urban neighborhood in Texas, USA. As per the DLI map generated, guidelines are provided for the selection of specific types of grass for urban greening. The computational methodology proposed herein can be applied in context of other urban spaces and climates.

Author Keywords

Photosynthetically Active Radiation; Solar radiation modeling; Urban Microclimate; Vegetation.

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING (e.g. Model Development).

1   INTRODUCTION

The role of vegetation as an instrument of modulating thermal discomfort in urban spaces is widely acknowledged. On a microclimate level, the presence of foliage, in the form of trees and grass, is found to be effective in improving thermal comfort perception. Some of the common metrics that are used to quantify thermal comfort are Predicted Mean Vote (PMV), Physiologically Equivalent Temperature (PET) and Universal Thermal Climatic Index (UTCI), all of which utilize Mean Radiant Temperature (MRT)in their calculation [1]. On a macro-level, the presence of vegetation has shown to be a mitigating factor for the Urban Heat Island (UHI) Effect [2]. An increase in the vegetation profile of an area correlates with reduction in its ambient temperatures in summers [3]. Vegetation in the form of turfgrass or food crops, when applied on pavements and green roofs, reduces the temperature of these surfaces as it has a lower sensible heat flux.

2   PROBLEM STATEMENT

Several approaches - depending on the spatial scope, depth of analysis, and level of detailing desired - are employed for accommodating vegetation in urban and outdoor thermal comfort simulations. Mesoscale simulation tools like Urban Weather Generator typically allow for vegetation to be input as a fraction of the total surface area. Computational Fluid Dynamics-based tools and Urban Thermal Environment tools (like Fluent and ENVI-MET respectively) offer a greater degree of detailing in the form of shapes and surface properties of plants and shrubs [4-8]. Outdoor thermal comfort simulation tools predominantly rely on allocating vegetation as elements of thermal zones to estimate surface temperatures [9].

Irrespective of the methodology employed for incorporating vegetation in the simulation approaches referenced above, the level of detailing rarely, if ever, provides definitive guidance regarding the feasibility of growing and maintaining specific species of plants and turf.

The type and distribution of plant species chosen for greening should be compatible with the geographical and weather conditions of a space. Additionally, the amount of radiation available for affecting plant growth might vary significantly even within a single urban location. This is attributable to urban topology, street canyons and shading profiles from high-rises. One example of a field-study that proposes a species- and region-specific approach for incorporating green roofs can be found in [10]. The following section reviews the current practices for selecting vegetation types followed in botany and agricultural sciences.

3   METHODOLOGY FOR SELECTING VEGETATION FOR URBAN GREENING: STATE OF SCIENCE

The ability to grow any type of vegetation is predicated on the availability of adequate moisture, soil nutrients and light. Of these parameters, the first two can be supplemented through irrigation and fertilization, irrespective of whether

the location is indoors or outdoors. However, for vegetation to be grown on a large scale with the intent of improving thermal comfort and mitigating UHI, the lighting requirements need to be met exclusively through daylight.

The following sub-sections review the current state of science with regards to the factors that influence the selection of vegetation.

3.1 Radiation modeling to estimate Photosynthetically Active Radiation (PAR)

The range of wavelengths, between 300nm to 700nm, of solar radiation that is responsible for photosynthesis is referred to as photosynthetically active radiation (PAR). PAR is expressed as photosynthetic photon flux density (PPFD) in micromoles per second per square meter (μmol/sec-m^2)[11].

Within the fields of botany and agricultural sciences, there is a consensus, that for the purposes of estimating plant growth, the quantum of PPFD can be condensed to Daily Light Integral (DLI). DLI is a measurement of the total amount of PAR delivered over a 24-hour period and is expressed in mol/m^2 per day [12]. The methodology for measuring and maintaining DLI in commercial greenhouses, along with a discussion on its relevance in plant growth, development, yield and quality can be found in [13]. On a much larger scale, studies have provided annual climate-based estimates of DLI in mainland United States [14, 15] and China [16]. An example of the DLI contour map generated for United States is shown in Figure 1.

Till now, there have been limited attempts to leverage DLI estimates to inform the selection of vegetation in urban microclimate simulations. One study, conducted in China, employed field measurements of DLI to guide the selection of plant species in urban spaces [17].

Most species of vegetation have recommended levels of DLI which are required for them to grow optimally in different seasons. An extensive list of studies that correlated plant growth with DLI measurements can be found in [14]. The quantum of DLI affects various aspects of plant growth such as height, stem diameter, dry weight, leaf chlorophyll, leaf area and flowering. DLI recommendations for various plant species, aggregated from several studies over the years, can be found in [13, 18].

3.2 Selection of vegetation species: Suitability for urban greening

In addition to radiation levels, the availability of irrigation and the water requirements of the plant, the soil composition at the site, degree-day requirement, hardiness zone classification and wind conditions are some of other aspects to be considered for selection of plant specifies for a specific location. Examples of climate-specific considerations for selection of food crops and grass types can be found in [10] and [19] respectively.
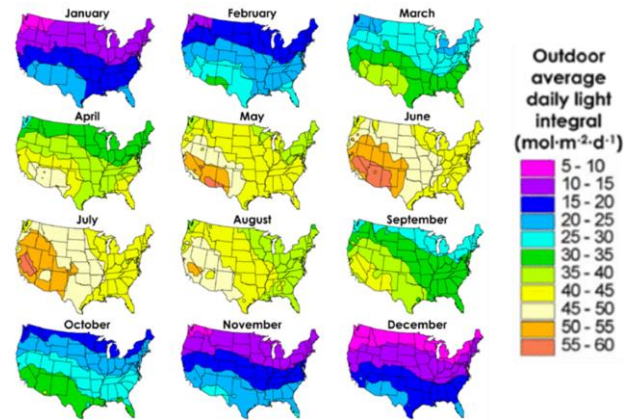
Figure 1. The annual DLI map for mainland United States. Credit: [14]

Depending on whether the plants are being grown at roof or ground level, or both, the depth of the substrate thickness is also a critical factor in the selection of vegetation. In the case of green roofs, several categorizations have been made to classify soil types as intensive, semi-intensive, single-course extensive and multi-course extensive. Intensive and semi-intensive roofs typically feature a substrate thickness of above 6 inches while the other two types have a soil thickness of 3-6 inches. Soil thickness for growing grass is typically six inches or less [20].

The next section proposes a simulation-based approach for the calculation of DLI. The subsequent section describes a case-study where the proposed simulation-based approach was employed to guide the selection of plant species in a commercial neighborhood within a city for the purposes of urban greening.

4 CALCULATION OF DLI IN URBAN SPACES FROM TYPICAL METEOROLOGICAL WEATHER DATA: THEORETICAL BASIS

Past studies featuring the mapping of DLI from weather data have typically involved mapping radiation values obtained from weather stations directly on to the surfaces of a particular area by disregarding natural and urban topography. This approach appears feasible on a very macro-level, as addressed by [14-16] for maps of contiguous United States and mainland China respectively. However, at the scale of a city neighborhood, the calculation of radiation levels needs to be cognizant of both the local climate as well as shadowing and reflection effects from buildings and other urban structures.

For providing DLI estimates for urban spaces, the authors propose a three-step approach for radiation modeling. These steps relate to locating relevant weather datasets for a specific urban location, mapping the radiation obtained from the weather dataset to urban topology and finally estimating DLI values from the mapped radiation. The following sub-sections discuss these steps.

4.1 Selection of climate dataset for a specific location

In their effort to map nationwide radiation data for the maps of United States and China respectively, the studies discussed in [14, 15] and [16] aggregated global-horizontal radiation data of nearly three decades. This data was collected from a database maintained by the National Renewable Energy Laboratory (NREL) and was averaged into monthly DLI values as shown in Figure 1 for the study presented in [14]. A similar approach was followed for creating the maps in the Chinese study as well [16].

For the purposes of mapping radiation data to an urban setting that might comprise of a few city blocks, the authors relied on, and propose the use of, typical meteorological year (TMY) datasets. These datasets contain hourly data for physical quantities such as temperature, radiation, wind velocity and humidity for an entire year. TMY data is usually a composite of typical months from several years. The selection of years chosen for creating a TMY dataset is based on statistical measures of similarity of daily indices such as minimum, mean and maximum dry-bulb temperature, and daily total global and direct solar radiation [21].

TMY data has been employed in building performance simulations for several decades. Successive versions of these datasets, based on their order of release, have been designated as TMY, TMY2 and TMY3 [22]. At present, TMY datasets are generally distributed in Energy Plus Weather (EPW) file formats and are available for over 2100 locations across the world [21]. From the various hourly data that are available in an EPW file, the authors used direct-normal radiation and diffuse-horizontal radiation, both expressed in Watt-hours/meter^2, for the purposes of calculating DLI levels. The rationale for using these data are explained below:

1. Sky modeling: Direct-normal and diffuse horizontal radiation, when considered along with the geographical details of a location such as latitude, longitude and time-zone, can be used to generate a weather-based Perez sky-model [23].

2. Calculation of radiation from Perez sky model: The Perez sky model can be used to calculate incident radiation on surfaces through ray-tracing or radiosity based methods [24, 25]. The incident radiation can then be used to calculate DLI levels.

3. Ability to model shadows and reflections: The use of direct-normal radiation enables the modeling of direct radiation from the sun. This makes it possible to incorporate the effects of shadowing and reflected radiation into the calculation of DLI [24, 26].

4.2 Calculating hourly radiation values from climate data

The Perez sky model discussed in the previous section can be used to calculate the incident radiation for a specific point-in-time through a conventional raytrace simulation. This approach is impractical for calculating hourly values of radiation for the entire year, as it would necessitate 8760 (24 hours for 365 days a year) ray trace calculations. The authors relied on a finite-element approach for calculating the incident radiation that is based on the concept of Daylight Coefficients [27]. This approach has been empirically validated by several studies for the purposes of calculating illuminance from daylight [28, 29]. The two concepts that are central to this approach are:

1. The celestial hemisphere, comprising of the sun and sky, can be subdivided into discrete radiant patches as shown in Figure 2.

2. The total radiation at any measurement point can be calculated by adding up the radiant contribution from individual sky patches. Figure 3 illustrates this concept in the context of an indoor space

4.3 Deriving DLI from hourly radiation values.

The range of wavelengths, between 300 to 700nm, that constitutes PAR significantly overlaps with the range that constitutes the visible radiation. The range of wavelengths responsible for actuating vision is between 380 to 760nm [11, 30]. Visible radiation measured on surfaces is expressed in terms of illuminance. The unit of illuminance is lux [30]. Provided that the spectral power distribution of a light source – regardless of whether it is artificial or natural – is known, it is possible to calculate PAR from illuminance [11]. So, it is possible to leverage the results of a climate-based daylight simulation to calculate PAR.

Conversion formulae exist, for both empirical and simulation-based measurements, to calculate PPFD from illuminance expressed in lux or foot-candles. This conversion is facilitated by the availability of kilolux-to-PPFD conversion factors for different types of light sources ranging from daylight to compact fluorescent lights [11, 13, 31, 32]. The assumptions and methodology to be followed for the calculation of illuminance, and by extension PPFD, in an urban context are similar to those for an indoor space. This is demonstrated through the case study discussed in the following section.
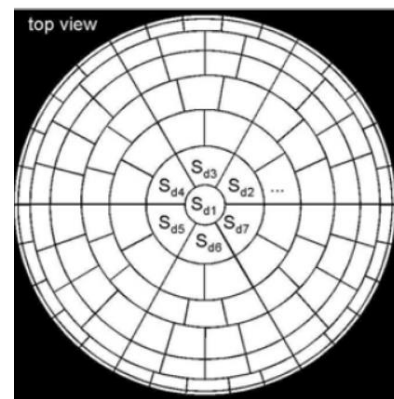
Figure 2. The division of the celestial hemisphere into discrete patches. (Credit: [33])
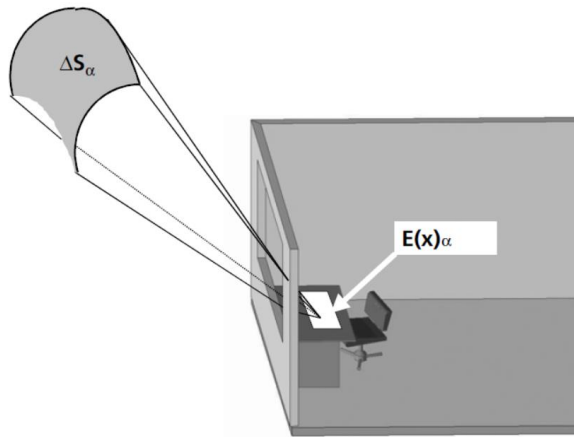
Figure 3. Radiation from a single sky patch measured on a surface inside a room. (Credit: [33])

5 PROTOTYPICAL CASE STUDY: SETUP AND CALCULATION

5.1 Setup

The geographical area chosen for the case study is Dallas, Texas, USA. Dallas has a humid subtropical climate with hot summers. The summer daytime temperatures frequently exceed 38°C [34]. The effects of UHI are prevalent in Dallas city and the greater Dallas-Fort Worth metropolitan area [35-37]. The specific location chosen for analysis is a neighborhood in downtown Dallas. The neighborhood, as pictured in Figure 4, consists of commercial high-rises and streets. The possibility of growing any sort of vegetation is restricted to the roofs and adjoining ground spaces of the commercial high rises.
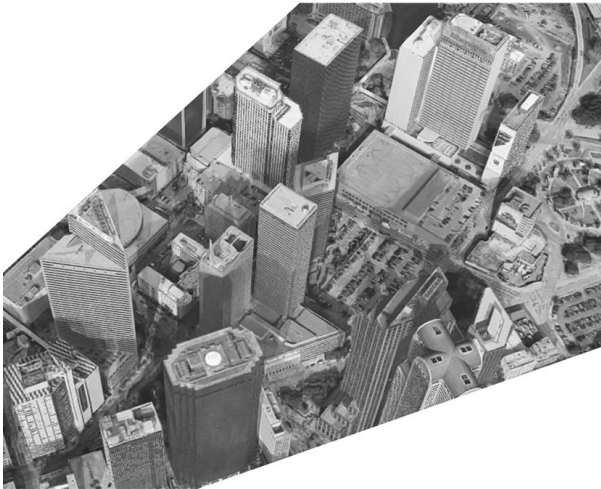
Figure 4. The location in Dallas considered for the prototypical study.

Based on the known effects of vegetation on thermal comfort perception, it is expected that the addition of vegetation will likely have mitigating effect on the daytime air temperatures in the area and thereby improve thermal comfort perception during the warm summers.

5.2 Calculation of Daily Light Integral (DLI)

As discussed previously, conversion factors exist for the purposes of deriving DLI from illuminance from different types of light sources [32]. The value of illuminance may either be measured on field through an illuminance meter or calculated through a daylighting simulation.

In this study, the illuminance was calculated using Radiance, a physically based ray-tracing software. The illuminance values calculated through Radiance have been validated with field measurements by several studies over the past three decades [24, 38, 39]. The hourly annual values of illuminance calculated through Radiance were based on the Daylight Coefficient method [26, 27].

The standard inputs for calculating illuminance with the daylight coefficient method are:

1. CAD model of the location

2. Reflectivity/transmissivity of materials

3. Placement of illuminance sensors within the CAD model and

4. Typical Meteorological Year (TMY) weather data file for the nearest weather station.

The Daylight Coefficient method is applied in this scenario by measuring the radiation from individual sky patch on the entire urban context as shown in Figure 5. The CAD model of the urban context created for the purposes of this study is shown in Figure 6. The placement of the illuminance sensors at roof and ground levels is also highlighted in Figure 6. The TMY data used for the simulations was of Dallas Fort-Worth area [40]. The theoretical basis for the conversion of illuminance, expressed in lumens/sq. meter (lux), to PPFD, expressed in μmol/sec-m^2, can be found in [32]. Based on the conversion formulae provided in [32] and [13] for daylight, the value of hourly PPFD was calculated by multiplying average daily foot candles per hour values with 0.20. The multiplication factor derived from the fact the sunlight has 0.20 foot-candles per μmol.m-2.s-1.

Figure 5. Radiation from a single sky patch being measured in an urban context.
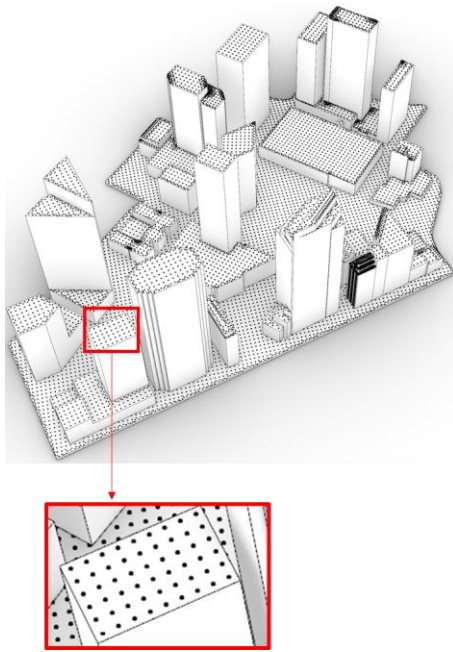
Figure 6. CAD model of the location considered for the analysis. The black dots refer to the positions of the illuminance sensors considered for the daylighting analysis.

DLI was then calculated by multiplying with 0.0864, the total number of seconds in day divided by 1,000,000. The daily values of DLI, obtained for every day of the Typical Meteorological Year, were aggregated and grouped into monthly averages.

6   RESULTS

6.1 DLI values obtained from the simulation
The monthly average values for DLI across the entire year are shown in Figure 7. Key seasonal observations regarding the mapped DLI values in Figure 7 are listed below:

1. Winter (November, December, January, February): There is a significant difference between DLI levels on the roofs and on the ground level. This can be attributed to shading from the buildings and the low angle solar radiation received. DLI for ground level varies approximately from 0 to 15, while that on the roofs varies from 15 to30.

2. Spring/Fall (March, April, September, October): The difference in DLI levels between the roof and ground levels is less pronounced. The general range of DLI is in the range of 15 to 30 and the effect of shading from buildings is noticeable in the periphery of the buildings.

3. Summer (May-August): Due to a prevalence of high angle solar radiation as well as a higher percentage of days with clear skies, the DLI values are in the vicinity of 30 and above.

6.2 Choice of vegetation based on DLI and other considerations
Typical urban greening strategies involve the planting of trees, food crops or grass. Based on the urban topology of the location considered for this study, as well as the commercial nature of the buildings in it, the planting of trees and food crops are impractical. Therefore, the authors considered several varieties of turfgrass for the purposes of vegetating the building roofs and ground locations near them. The net positive impact of turfgrass on thermal comfort has been documented in past studies [41, 42].

According to their optimum temperature requirements, turfgrasses can be classified as cool-season and warm - season grasses. Cool-season grasses have optimum growth temperatures between 15 to 24°C, while warm-season grasses grow optimally between 27 to 35 °C [43]. The climate of Dallas is more suited to the growth of warm-season turfgrasses. Further details regarding the suitability of different turfgrass types to the climate of Texas can be found in [19]. Based on the recommendations provided in [19], Bermudagrass and Zoysiagrass were identified as two probable options for the site being considered. The relative traits of the two grasses are listed in Table 1.

	Bermudagrass	**Zoysiagrass**
Shade Tolerance	Very low to low	Moderate to high
Water Requirement	Moderate to low	Moderate
Drought Tolerance	Very good	Very good
Traffic Tolerance	High	Moderate to high
Cold Tolerance	Moderate	Moderate to high

Table 1. Relative traits of Bermudagrass and Zoysiagrass reproduced from [19]. Traffic tolerance relates to the resilience of the grass from wear and tear due to people walking or running.

A cultivar, derived from "cultivated variety", refers to plant types that are created through selective breeding. Four cultivars, two each of Bermudagrass and Zoysiagrass type, were evaluated for this study. The seasonal DLI requirements of these cultivars are provided in Table 2 [44]. The recommended season for planting warm season grasses is typically in Spring or Fall [43]. Considering Spring to be planting season, a location-specific decision matrix was created for the selection of different types of cultivars. The decision matrix, shown in Table 3, was generated by cross referencing the DLI requirements of the cultivars to the calculated DLI levels shown in Figure 7.
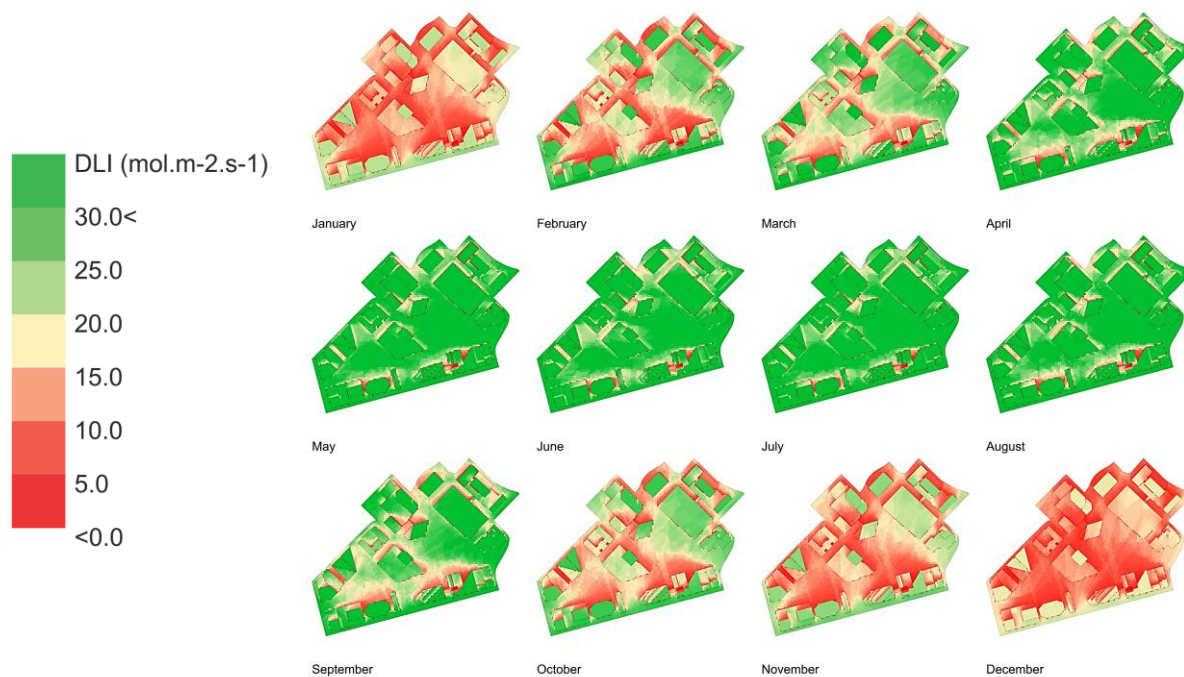
Figure 7. The monthly average DLI map for the Typical Meteorological Year. Top row (from left to right): January to April. Middle row (from left to right): May to August. Bottom row (from left to right): September to December.

	Bermudagrass		**Zoysiagrass**	
	Tifway	Celebration	Diamond	Jamur
Spring	17	14	10	10
Summer	21	20	11	11
Winter	10	8	7	6

Table 2. DLI values, expressed in in mol/m^2 per day, for two varieties each of Bermudagrass and Zoysiagrass. The values, rounded off to the nearest integer, were aggregated from [44].

As indicated by Table 3, the cultivars of Bermudagrass are not suitable for planting on the ground level as the radiation values in Spring are low on account of shading from the buildings.

	Bermudagrass		**Zoysiagrass**	
	Tifway	Celebration	Diamond	Jamur
Roof	Yes	Yes	Yes	Yes
Ground	No	No	Yes	Yes

Table 3. A decision-matrix for the suitability of planting grass on roofs and ground level. The matrix was created by cross-referencing the values of DLI for grass-types (Table 2) with the DLI values mapped for the month of March (Figure 7).

7   DISCUSSION

The preceding sections presented a methodology for selecting the location and the type of plants that can be used for urban greening. The potential benefits of urban greening can be then evaluated by incorporating the characteristics of plants into simulations.

While these simulations are beyond the scope of this study, the following sub-section provides examples of how the effect of vegetation can be evaluated. The sub-sections after that discuss the limitations of this study and some of the challenges encountered in the collection of relevant data.

7.1 Simulation-based studies for quantifying the impact of vegetation on thermal comfort.

The methodology to factor in, and study the impact of, vegetation in urban simulations varies according to the scale and intent of the simulation. For outdoor thermal comfort simulations at a city-block level, vegetation can be incorporated into the simulation model by defining surfaces of thermal zones - that represent buildings or ground - as vegetation. The theoretical bases for the material definitions for including vegetation are discussed in [45, 46]. An urban microclimate study that leverages these types of materials can be found in [47].

The impact of vegetation on thermal comfort at a mesoscale was studied in [6] through a sensitivity analysis with Urban Weather Generator. That study also considered factors like city size, urban morphology and tree coverage. The studies described in [48] and [49] employed ENVI-MET to study the impact of vegetation on thermal comfort expressed in MRT and PET respectively.

7.2 Limitations of this study

The methodology used in this study relied on TMY data to compute DLI estimates. TMY data is generated statistically

378

from long term weather patterns of a location. Empirically measured DLI levels measured at the site will deviate from these estimates based on prevailing weather conditions.

An arithmetic constant was used to convert visible light to PAR. The value of the constant, assumed as 0.2 by the author, has been suggested between 0.18 to 0.22 by various sources [13, 32]. Additionally, the Daylight Coefficient method employed for calculating annual results relies on some simplified assumptions regarding the position of the sun and sky [26].

The DLI values employed for the selection of grass types are based on studies conducted in greenhouses [44]. PPFD levels inside a greenhouse can be fractionally higher or lower than the levels measured outside based on whether the glazing is diffusing or specular respectively.

7.3 Challenges in the collection of relevant data pertaining to DLI

Presently, there are no authoritative and comprehensive sources of information pertaining to DLI requirements for different types of vegetation that may be used for urban greening. Consequently, in this study, a significant amount of effort was devoted to locating DLI values specific to the types of turfgrass applicable to Dallas. The authors expect the lack of readily accessible data to be a recurring challenge in the adoption of any DLI-based recommendation strategy.

8 CONCLUSION AND FUTURE INITIATIVES

This research was an effort to leverage the current state of the art in agricultural practice and building simulations to inform urban greening decisions. The authors presented a review of existing practices in agriculture and botany with regards to the calculation of DLI through field measurements and its role in selection of vegetation. This was followed by a discussion on how a computational approach based on the Daylight Coefficient Method can be employed to calculate DLI through simulations. This approach was demonstrated through a prototypical case study presented in the context of an urban location.

Future initiatives being considered by the authors include urban microclimate simulation-based studies that estimate the impact of vegetation on thermal comfort metrics like PET and UTCI. The authors also intend to perform sensitivity analysis on the impact of plant parameters like height, leaf reflectivity and emissivity on thermal comfort.

REFERENCES
[1] Fabbri, K. *Indoor thermal comfort perception*. Springer, Switzerland, 2015.
[2] Oke, T. R. *Boundary layer climates*. Routledge, 1988.
[3] Susca, T., Gaffin, S. R. and Dell'Osso, G. R. Positive effects of vegetation: Urban heat island and green roofs. *Environmental Pollution*, 159, 8 (2011), 2119-2126.
[4] Bueno, B., Norford, L., Hidalgo, J. and Pigeon, G. The urban weather generator. *Journal of Building Performance Simulation*, 6, 4 (2013), 269-281.
[5] Nakano, A., Bueno, B., Norford, L. and Reinhart, C. F. Urban Weather Generator-A novel workflow for integrating urban heat island effect within urban design process. In *Proceedings of the 14th International Conference of IBPSA Building Simulation* Hyderabad, India, 2015
[6] Salvati, A., Palme, M. and Inostroza, L. Key Parameters for Urban Heat Island Assessment in A Mediterranean Context: A Sensitivity Analysis Using the Urban Weather Generator Model. *IOP Conference Series: Materials Science and Engineering*, 245 (2017), 082055.
[7] Tsoka, S., Tsikaloudaki, A. and Theodosiou, T. Analyzing the ENVI-met microclimate model's performance and assessing cool materials and urban vegetation applications–A review. *Sustainable cities and society*, 43 (2018), 55-76.
[8] Lemercier, C. A. G. Sensitivity analysis of Urban Heat Island parameters based on Urban Weather Generator model (2019).
[9] Mackey, C., Baranova, V., Petermann, L. and Menchaca-Brandan, M. A. Glazing and Winter Comfort Part 2 : An Advanced Tool for Complex Spatial and Temporal Conditions (2017).
[10] Song, X. P., Tan, H. T. W. and Tan, P. Y. Assessment of light adequacy for vertical farming in a tropical city. *Urban Forestry & Urban Greening*, 29 (2018), 49-57.
[11] Thimijan, R. and Heins, R. Photometric, radiometric, and quantum light units of measure: a review of procedures for interconversion [Horticultural science's growth chambers and controlled environments, engineering specifications]. *HortScience* (1983).
[12] Kozai, T., Tsukagoshi, S. and Sakaguchi, S. *Reconsidering the Terminology and Units for Light and Nutrient Solution. Smart Plant Factory.* Springer Nature, Singapore, 2018.
[13] Torres, A. P. and Lopez, R. G. *Measuring daily light integral in a greenhouse*. Purdue University, 2010.
[14] Korczynski, P. C., Logan, J. and Faust, J. E. Mapping monthly distribution of daily light integrals across the contiguous United States. *HortTechnology*, 12, 1 (2002), 12-16.
[15] Faust, J. E. and Logan, J. Daily light integral: A research review and high-resolution maps of the United States. *HortScience*, 53, 9 (2018), 1250-1257.
[16] Sun, Y., Sun, Y., Sun, Z., Li, P. and Li, C. Mapping monthly distribution of daily light integrals across China (2015).
[17] Lin, C.-H., Hsu, C.-Y. and Lin, J.-Y. Using Daily Light Integral Concept to Construct the Ecological Plant Design Strategy of Urban Landscape. *World Academy of Science, Engineering and Technology, International Journal of Civil, Environmental, Structural, Construction and Architectural Engineering*, 9, 7 (2015), 891-897.
[18] Nau, J. *Ball redbook: Crop production*. Ball Publishing, 2011.

[19] Chalmers, D. R. and McAfee, J. *Turfgrass Selection for Texas* https://agrilifeextension.tamu.edu/library/landscaping/turfgrass-selection-for-texas/, 2012.

[20] Shafique, M., Kim, R. and Rafiq, M. Green roof benefits, opportunities and challenges–A review. *Renewable and Sustainable Energy Reviews*, 90 (2018), 757-773.

[21] Hensen, J. L. *Building performance simulation for design and operation*. Routledge, Oxford, USA, 2019.

[22] Wilcox, S. and Marion, W. *Users manual for TMY3 data sets*. NREL, Colorado, USA, 2008.

[23] Perez, R., Seals, R. and Michalsky, J. To all-weather model for sky luminance distribution—preliminary configuration and validation. *Solar Energy*, 51, 5 (1993), 423.

[24] Mardaljevic, J. Validation of a lighting simulation program under real sky conditions. *Lighting Research and Technology*, 27, 4 (1995), 181-188.

[25] Ashdown, I., Jackson, C., Spahn, J. and Saemisch, T. *LICASO and DAYSIM* http://agi32.com/blog/2017/02/11/licaso-and-daysim/, 2017.

[26] Reinhart, C. F. and Herkel, S. The simulation of annual daylight illuminance distributions — a state-of-the-art comparison of six RADIANCE-based methods. *Energy and Buildings*, 32 (2000), 167-187.

[27] Tregenza, P. R. and Waters, I. M. Daylight coefficients. *Lighting Research and Technology*, 15, 2 (1983), 65-71.

[28] Reinhart, C. F. and Walkenhorst, O. Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds. *Energy & Buildings*, 33, 7 (2001), 683-697.

[29] Mardaljevic, J. *Daylight simulation: validation, sky models and daylight coefficients*. De Montfort University, Leicester, UK, 1999.

[30] DiLaura, D. L., Mistrick, R. G., Houser, K. H. and Steffy, G. *The lighting handbook: reference & applications*. IES, New York, USA, 2011.

[31] Ashdown, I. LED Lighting for Horticulture. *LED Professional Review*, 48 (2015), 60-65.

[32] Ashdown, I. *Photometry and Photosynthesis* https://lightinganalysts.com/photometry-and-photosynthesis/, 2019.

[33] Reinhart, C. F. *Daylight availability and manual lighting control in office buildings: Simulation studies and analysis of measurement*. Fraunhofer-IRB-Verlag, Stuttgart, Germany, 2001.

[34] National Weather Service *DFW Climate Narrative*, 2019.

[35] Gallo, K. P. and Owen, T. W. Assessment of urban heat Islands: A multi-sensor perspective for the Dallas-Ft. worth, USA region. *Geocarto International*, 13, 4 (1998), 35-41.

[36] Darby, L. S. and Senff, C. Comparison of the urban heat island signatures of two Texas cities: Dallas and Houston. In *Proceedings of the Seventh Symposium on the Urban Environment* San Diego, California, USA, 2007

[37] Hu, X.-M. and Xue, M. Influence of Synoptic Sea-Breeze Fronts on the Urban Heat Island Intensity in Dallas–Fort Worth, Texas. *Monthly Weather Review*, 144, 4 (2015), 1487-1507.

[38] Ward, G. The RADIANCE lighting simulation and rendering system. In *Proceedings of the 21st annual conference on Computer graphics and interactive techniques* Florida, USA, 1994

[39] Geisler-Moroder, D., Lee, E. S. and Ward, G. Validation of the Five-Phase Method for Simulating Complex Fenestration Systems with Radiance against Field Measurements. In *Proceedings of the 15th International Conference of the International Building Performance Simulation Association* California, USA, 2017

[40] US Department Of Energy *Dallas Fort Worth TMY3 data* https://energyplus.net/weather-location/north_and_central_america_wmo_region_4/USA/TX/USA_TX_Dallas-Fort.Worth.Intl.AP.722590_TMY3, 2019.

[41] Declet-Barreto, J., Brazel, A. J., Martin, C. A., Chow, W. T. L. and Harlan, S. L. Creating the park cool island in an inner-city neighborhood: heat mitigation strategy for Phoenix, AZ. *Urban Ecosystems*, 16, 3 (2013), 617-635.

[42] Takebayashi, H. and Moriyama, M. Study on the urban heat island mitigation effect achieved by converting to grass-covered parking. *Solar Energy*, 83, 8 (2009), 1211-1223.

[43] Christians, N., Patton, A. J. and Law, Q. D. *Fundamentals of turfgrass management*. Wiley Online Library, 2011.

[44] Zhang, J., Glenn, B., Unruh, J. B., Kruse, J., Kenworthy, K., Erickson, J., Rowland, D. and Trenholm, L. Comparative performance and daily light integral requirements of warm-season turfgrasses in different seasons. *Crop Science*, 57, 4 (2017), 2273-2282.

[45] Sailor, D. J. A green roof model for building energy simulation programs. *Energy and buildings*, 40, 8 (2008), 1466-1478.

[46] Sailor, D. J., Elley, T. B. and Gibson, M. Exploring the building energy impacts of green roof design decisions – a modeling study of buildings in four distinct climates. *Journal of Building Physics* (2011), 372-391.

[47] Mackey, C., Galanos, T., Norford, L. and Roudsari, M. S. Wind, Sun, Surface Temperature, and Heat Island: Critical Variables for High-Resolution Outdoor Thermal Comfort (2017).

[48] Rosheidat, A. and Bryan, H. Optimizing the effect of vegetation for pedestrian thermal comfort and urban heat island mitigation in a hot arid urban environment. *Proceedings of SimBuild*, 4, 1 (2010), 230-237.

[49] Ali-Toudert, F. and Mayer, H. Effects of asymmetry, galleries, overhanging façades and vegetation on thermal comfort in urban street canyons. *Solar Energy*, 81, 6 (2007), 742-754.

Programming Spatiality: Simulating Social Interaction as a Parameter of the Building's Geometrical Volatility

Anna Karagianni[1], Vasiliki Geropanta[2], Panagiotis Parthenios[3]

[1]Technical University of Crete
Chania, Greece
akaragianni1@isc.tuc.gr

[2]Technical University of Crete
Chania, Greece
vgeropanta@arch.tuc.gr

[2]Technical University of Crete
Chania, Greece
parthenios@arch.tuc.gr

ABSTRACT

How is social interaction affecting the geometry of a building? This paper presents a simulation experiment on social interaction within the framework of designing a new technology incubator in Akihabara, Tokyo. The pseudo-code simulates the movement of people within the three dimensional grid that represents the building, by tracking the movable wi fi signals and their aggregation in the grid nodes. By experimenting with the set of parameters that accurately simulate human behavior in buildings and explicitly enhance social interaction in it, the authors extract a set of possible heatmaps that are translated into real space. Hence, the paper explores the relationship between social innovation and building design, analysing how inserting information and city data in the algorithmic design process can lead to a more collaborative and inclusive architectural design. Exploring the intersection of algorithmic design and smart users, the authors explore how various characteristics of social innovation can be elaborated in the case study of the technology incubator. The findings suggest that user connectivity can contribute to the geometrical 'decoding' of architectural design.

Author Keywords

algorithmic design; Processing; IoT; building simulation

ACM Classification Keywords

Computing Methodologies; Modelling and Simulation; Simulation Types and Techniques.

1 INTRODUCTION

In the last few decades, 1[1]. In the field of architecture, this practiced has revealed the importance of IoT as key instrument to connect the research on users behavior and the built environment [2]. Considering the above observations, this paper highlights the relationship between social innovation and architectural design, analysing how Wifi connectivity could be a parameter in generating geometries.

The scope of the study is to provide a framework where the behavioral patterns of building occupants, geometry and spatiality of the building are interrelated. This seems crucial in the era of IoT, where the social aspect of buildings emerges as an indispensable layer of built space [3]. In that framework, it needs to be investigated in association with user behavior and geometrical flexibility.

The structure of this paper is as follows: In Section 2, we argue on the possible link - bondage that stands between the topics of contemporary social innovation in urban contexts due to the use of ICT and IoT, and more specifically, in architecture and urban planning. In Section 3, we briefly present the case study that framed our research and explain the key steps of methodology taken. In Section 4, we summarise the results obtained from our analysis and we discuss the conclusions of the work.

2 LITERATURE REVIEW

2.1 IoT, 5G and social innovation in shaping spatiality

Social innovation is today a "contested term" [4], since it is the topic of discussion in many different disciplines and to which there have been attributed more than 200 definitions [5]. This can be ascribed to many facts [6] but the most important is that social innovation rarely happens accidentally; rather, it is the result of actions and projects very carefully planned [7].

In fact, the various contemporary technological developments have allowed "the accumulation and codification of large amounts of knowledge"[8], and large scale of user engagement and co-creation pushing our discipline towards new directions. Urban planning is already improved on multiple scales through practices of social innovation [9], even if most solutions are inherently local [10]. Similarly, architecture is also adopting many of the logics and mindsets of this concept, aiming at improving many of the problems that previously seemed eternally unresolved. This happens mainly through the use of many digital tools, applications [11], with smart technologies that influence urban and architectural functions, all of which foster collaborative innovation, social participation and city change [12].

'Internet of Things', in fact, praised the human factor in the digital environment from its early beginning [13]. It aimed 'to describe a system where the Internet is connected to the physical world via ubiquitous sensors [14], while the operation was already known under the terms 'Pervasive Computing', 'Ubicomp' and 'Ambient Intelligence' much earlier [15]. The biggest advantage of IoT is the ability to interconnect things and to allow for machine to machine communication (M2M) which is a type of communication where no human presence is needed. IoT is responsible for the traffic of massive data and therefore any revolution in the mobile might lead to massive large number of connected devices with high data rates per area ameliorating many practices that today are slower or not very well realized (for example 3d videos, use of multimedia and others). The Fifth generation (5G) will work in 2020 and will offer higher quality of data rates, higher quality, improved wireless communication and a boosted support for the IoT. The expectation in 5G is that every mobile application and services will be incorporated as a major aspect of the cloud computing environment.

Based on the above, the authors argue that the act of using 5G connectivity embodies one of the most important aspects of social innovation, which is citizen empowerment through effective participation. Through integrating this action, and thus framing and defining the unpredictability of user behaviour in any built space, the authors argue that this process might alter the synthetic parameters of the architectural process in 'designing by learning' based on real-time visualisation and interactivity, making design more agile, and targeted to rising needs.

3 CASE STUDY

3.1 Research methodology - The incubator of innovation and its tectonics

This study takes place in Akihabara neighborhood of Tokyo, being firstly introduced in the framework of the Kaizen Studio at GSAPP, Columbia University and lately elaborated by the authors at the School of Architecture, Technical University of Crete. The general objective is the creation of a new technology incubator, an innovative space that nurtures the development and exchange of ideas, concepts and visions, while using the spatial code of information in order to articulate real architectural space. The target is to achieve social innovation through design and innovative geometric analysis, using as a parameter of building design the Wifi connectivity. In fact, given that the wi-fi conveys itself mostly through the presence of people, mapping the virtual space of wi-fi signals across the incubator, will be the means through which interaction will be created between real and virtual, interior and exterior place.

Inside the incubator, the flow of wi-fi signals is translated through the amount of laptops, cell phones and i-pods that people carry with them and the outcome will mimic the conditions for the average user. In addition to the flow, there are info-zones with constant, stable wifi signal, spread around the incubator. Then, within the Processing programming environment, a new generic space is created, where the wifi signals are represented by agents that move along a three-dimensional grid. The nodes of the grid represent the stable wifi signal, while the dynamic - in terms of location - signal is assumed to be brought by users.

A significant aspect of the methodology was the simulation of social interaction versus the simple movement representation within space. In that case, a set of parameters were introduced and functioned in a threefold way: to ensure that user-agents will interact with other users through the wi-fi network; that user-agents are directed towards the nodes of the grid, representing the nodes of interaction between users and technology, thus producing knowledge; and finally that the knowledge accumulated in the nodes is exponentially increased.

The representation of the signal through an envelope shape with physical form contributes to the perception of virtual space as a discrete phenomenon and suggests a correlation between the use of space, the shape of space and the wi-fi signal.

3.2 Case study - the Code

The input of the pseudo - code is a set of possible amounts of wi-fi signals that circulate in the incubator, creating a shrinking and expanding flow. The first part of the code relates to the representation of spatiality: the creation of the

system, the agents and their consequences on it. The system consists of agents that move randomly along the paths of a 3d lattice (Fig. 1).
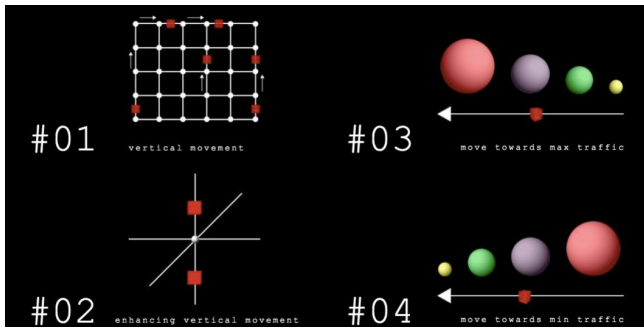
Figure 1. Part I: Code parameters.

During this phase, two basic parameters are created. The first parameter simulates people moving vertically within the building, to ensure an even distribution of people on all building floors. (Fig. 2) The second parameter reflects on a fundamental human need: that of social interaction. In that context, the system simulates people moving towards other people (Fig.3). Since each user carries a wi-fi signal, the aggregation of signals equals the accumulation of people. Under the assumption that people tend to move towards building parts with high density (auditoriums, lounges, etc.), the code simulates these places being located at the intersections of human flows: the grid nodes. When a user reaches a grid, the node accumulates wi-fi signal. Users are programmed to 'prefer' the nodes with higher signal accumulation rates. As a result of the parameterization, we achieve a primitive correlation of the user movement with the geometrical spatiality; the traffic reflects the users' movement and defines the scale and density of the space that will be created around the intersection.
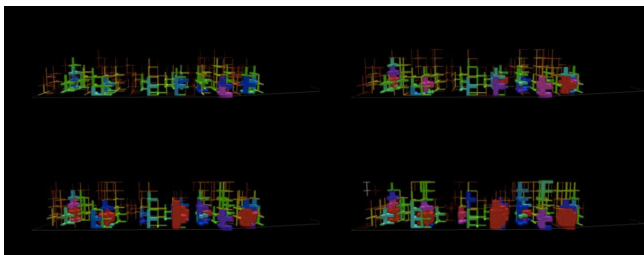
Figure 2. Experiment I.1. Enhancing users(agents) to move vertically.
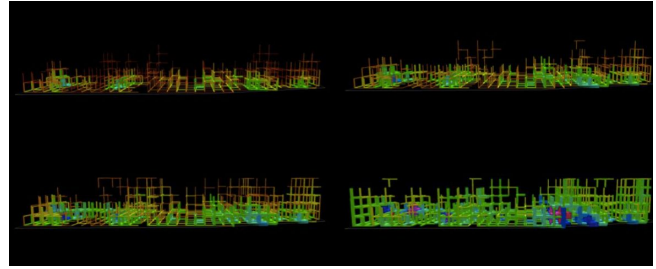
Figure 3. Experiment I.2. Simulating the need for social interaction.

The second part describes the spatio-temporal evolution of the system through shifting the rules and the way that the system is used to define the program, form and density of each space of the incubator (Fig. 4).
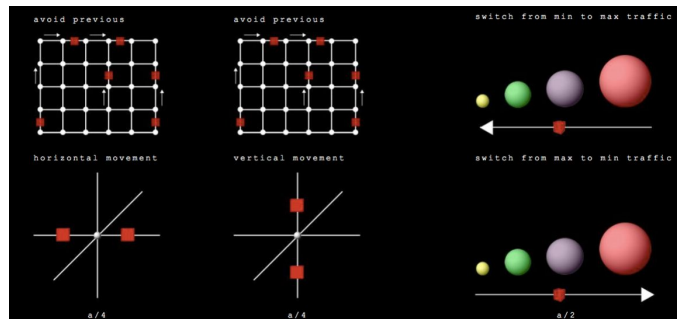
Figure 4. Part II. Code parameters.

Dividing the time into 'a' parts, users start from the bottom of the building, moving vertically in order to inhabit the whole space, then they avoid going to their previous location (Fig. 5), to create the best articulated space. During the second half of the 'a' part, they switch from the most visited places to the less visited ones, reflecting the tendency of people to be isolated after meeting a lot of people, or their desire to meet somebody when they are alone (Fig. 6). They finally occupy a 3d lattice of 36X6X6, arranged at a spacing of 6 m. Combining the above time and space - related parameters, the final output is created.
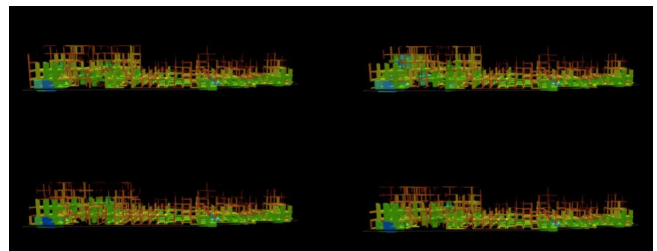
Figure 5. Experiment II.1. Users avoid previously visited nodes.
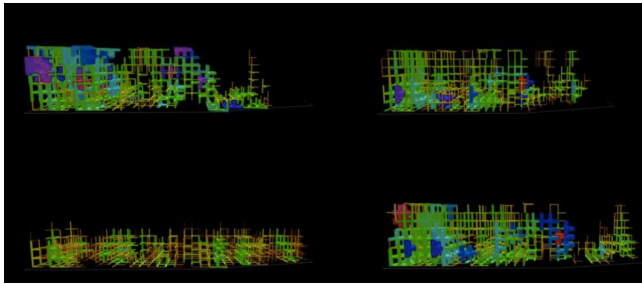
Figure 6. Experiment II.2. Users prefer less visited nodes..

The third part relates to the spatial interpretation through the evolution of the systematic organization of the code. The first assumption is associated with the geometrical tolerance to reach minimum and maximum living dimensions. The signal aggregation at the grid's nodes is translated into cubes of minimum 2 X 2 meters and maximum 8 X 8 meters. Space is generated by establishing a prism to connect the cubes (Fig. 7).
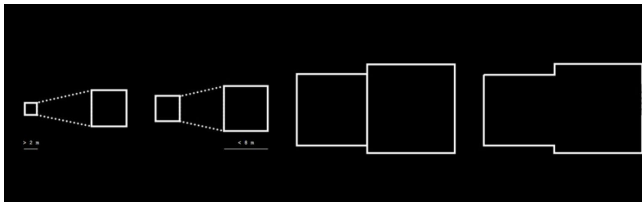
Figure 7. Generating space by connecting the nodes.
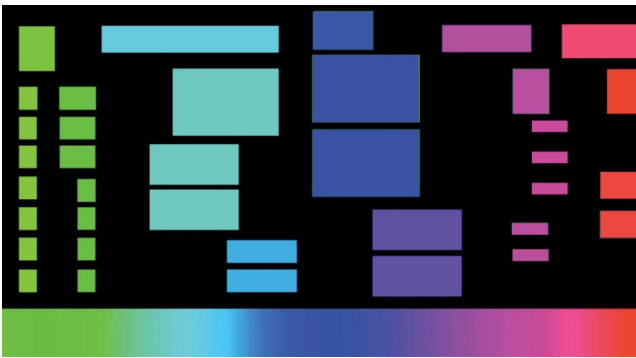
Figure 8. Private - public color scheme.

In order to distribute uses within the building, a color coded scheme is introduced to depict from private (green) spaces to public (red) ones (Fig. 8). The possible variations of the building geometry are then presented (Fig. 9) and the static instance of the building is turned into space (Fig. 10, 11).
The produced spaces of the incubator are of different scales and levels of publicness. Simple rules create a heterogeneous and anisotropic space where the system defines the scale of the space and the transition of one to another and where all spaces, of all scales contribute to one goal: social interaction.

4 CONCLUSIONS, DISCUSSION AND FURTHER RESEARCH

The fact that this code uses the precise reading of wi-fi signals as the design strategy of the building shell opens up new possibilities in two discrete fields: firstly in the simulation as a design strategy and secondly in the interrelation of the physical and social attributes of building.
Among the many strategies, the adoption of simulation-based practices seem an interesting field of action. As technology becomes increasingly sophisticated, and Wifi connection has allowed connectivity to become pervasive leading to the use of big data, many scholars tend to see the city as an 'open platform' for experimentation, hosting 'platform societies' [16] with users leading the way and driving urban changes [17]. In this context, simulation of urban data can be seen as a response to a growing body of smart city literature, a bridge between a top down and bottom up logic of design. In fact, one can imagine that the smart city literature might expand to include also algorithmic methods of designing that simulates urban data as taken directly by users that are in the interior of a building. This can be realized through the use of IoT objects, in several ways and for a number of reasons including methods to familiarise users with an ambient; to increase their awareness with respect to so many aspects of city life and to co-design and explore design alternatives.

In this study, relating the flow of digital signals within the virtual space to the outer, real world is achieved through designing an interactive shell that emerges through the spatial connection of nodes. Extending the idea of interaction, the envelope, consisting of cells, will be able to interact with the wind, light, orientation, sound and temperature or air-conditioning systems used within the building. Sensors measuring the differences of these phenomena will be integrated into the envelope, altering

the dimensions of its components and, consequently its texture. Using its sensors, it translates the events taking place outside into an animated image on its interior side. The envelope will deconstruct and reconstruct an image, associating the incubator with its context.
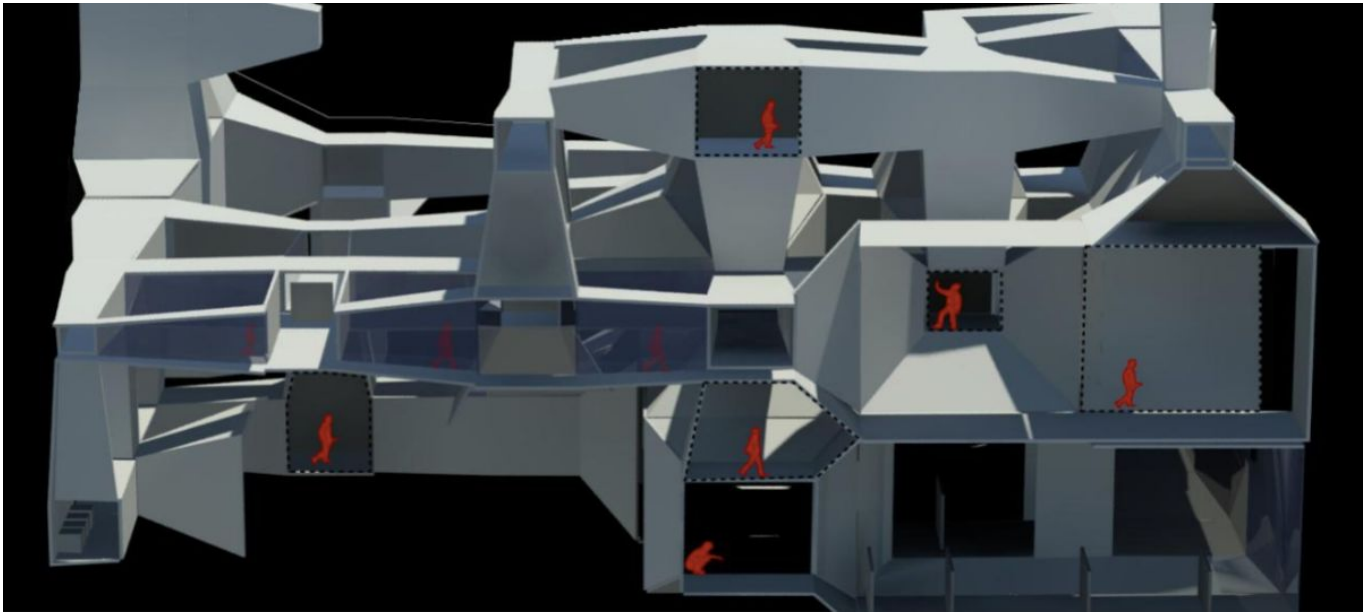
Figure 10. Building Section.   **Figure 11.** Geometric variations.
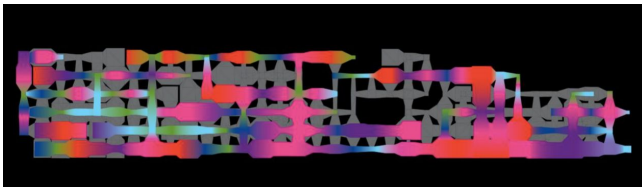
Figure 9. Plan of the generated building geometry.
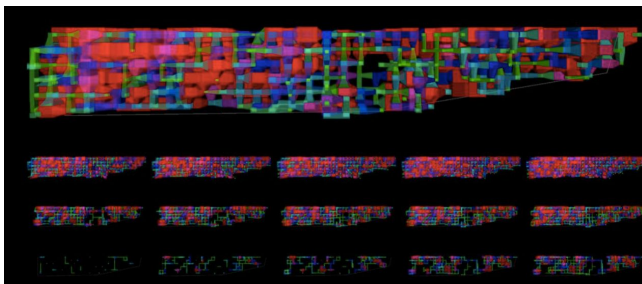
REFERENCES

1. Karagianni A., Geropanta V., Parthenios P. (2019). Exploring the ICT Potential to Maximize User - Built Space Interaction in Monumental Spaces The case of the municipal agora in Chania, Crete [eCAADe Conference Proceedings / ISBN 97891207181/ VOL 2] Portugal 13 September 2019, pp. 597 607. https://ecaade.org/conference/downloads/

2. Thelen, D., Wullink, F., Van Acoleyen, M., Pastoor, V., Thijs, M., (2019) 'The Future of the European Built Environment: A forward-looking description of Europe in 2030 and 2050', Arcadis Netherland B.V., https://www.arcadis.com/media/0/A/3/%7B0A3A3119-619F-4962-BCAD-1DB99CB1D031%7DThe%20Future%20of%20The%20European%20Build%20Environment_2019.pdf [accessed Jan 10 2020].

3. Rüede, D. and Lurtz, K. (2012) Mapping the Various Meanings of Social Innovation: Towards a Differentiated Understanding of an Emerging Concept, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2091039 [Accessed 10 January, 2020].

4. Schachter, E., Mónica, E. and Wallace, M.L. (2015) Shaken, But not Stirred': Six Decades Defining Social Innovation,

http://www.ingenio.upv.es/es/working-papers/shaken-not-stirred-six-decades-defining-social-innovation#.Wfr0–4hx1PY [accessed Jan 10 2020].

5. Phills, J.A., Deiglmeier, K. and Miller, D.T.(2008) 'Rediscovering social innovation', Stanford, Social Innovation Review, Vol. 6, No. 4, pp.34–43.

6. OECD Committee for Scientific and Technological Policy (2011), Fostering Social Innovation to Address Social Challenges, https.//www.oecd.org/sti/inno/47861327.pdf [accessed Jan 10 2020]

7. Bria, F., Almirall, E., Baeck, P., Halpin, H., Kingsbury, J., Kresin, F. and Reynolds, S. (2015) Growing a Digital Social Innovation Ecosystem for Europe, Nesta, London, http://one-architecture.nl/wp-content/uploads/2013/10/2015_The-Hackable-City_A-Research-Manifesto-and-Design-Toolkit.pdf [accessed Jan 10 2020].

8. Angelidou, M. and Psaltoglou, A. (2017) 'An empirical investigation of social innovation initiatives for sustainable urban development', Sustainable Cities and Society, Vol. 33, pp.113–125.

9. Grimm, R., Fox, C., Baines, S. and Albertson,K. (2013) 'Social innovation, an answer to contemporary societal challenges? Locating the concept in theory and practice', Innovation:The European Journal of Social Science Research, Vol. 26, No. 4, pp. 436–455.

10. Komninos, N. (2011) 'Intelligent cities: variable geometries of spatial intelligence', Intelligent Buildings International, Vol. 3, No. 3, pp.172–188.

11. Psaltoglou, Artemis & Angelidou, Margarita. (2019). Social Innovation, Games and Urban Planning: An analysis of current approaches.. International Journal of Electronic Governance. 11. 1. 10.1504/IJEG.2019.10015503.

12. Ashton, K. (2009) 'That 'Internet of Things' Thing' https://www.rfidjournal.com/articles/view?4986 [accessed Jan 10 2020].

13. Goetz, T. (2011) 'Harnessing the Power of Feedback Loops' https://www.wired.com/2011/06/ff_feedbackloop/ [accessed Jan 10 2020].

15. Sinclair, K. (1999) 'Your Building Address as a .com?' http://www.automatedbuildings.com/news/may99/articles/addresses/addresses.htm [accessed Jan 10 2020].

16. Ampatzidou, C., Bouw, M., van de Klundert, F., de Lange, M. and Waal, B.G.M. (2014), The Hackable City: A Research Manifesto and Design Toolkit. Knowledge Mile, Amsterdam. http://one-architecture.nl/wp-content/uploads/2013/10/2015_The-Hackable-City_A-Research-Manifesto-and-Design-Toolkit.pdf [accessed Jan 10 2020].

17. de Waal, M., de Lange, M. and Bouw, M. (2017) 'The hackable city: citymaking in a platform society', Architectural Design, Vol. 87, No.1, pp.50–57.

Investigation on the Impact of Passive Design Strategies on Care Home Energy Efficiency in the UK

Shan Shan Hou[1]

[1]Cardiff University
Cardiff, UK
hous1@cardiff.ac.uk

ABSTRACT

Care homes have consistent high energy requirements 24 hours a day and 7 days a week. It is important to improve their energy efficiency to fulfil the mandatory requirement of achieving nearly zero-energy from the 1st of January 2021 by the EU Energy Performance of Buildings Directive (EPBD) in order to deal with climate change and fuel security. In addition, nearly zero-energy care home design can reduce the running cost in order to reduce the finance burden in an aging society. Reducing energy consumption and carbon emissions in building design requires a holistic systems approach, while passive design which modifies external climate to improve the energy performance of a building is the foundation of this systems approach. Passive design strategies can reduce heating and cooling demands, lead to opportunities to integrate more efficient building systems with lower capacity as well as the potential to use renewable energy supply and storage systems to offset overall energy consumption. This paper aims to explore the impact of passive design strategies on energy performance of care home design in the UK. In this study, a standard UK care home with a modelling framework is developed as the reference case. A series of building energy simulations is conducted with local climate to evaluate the impact of a range of passive design strategies. This study identifies that passive design has an impact on energy efficient design, which accounts for up to 28% of energy saving. Window u-value, infiltration rate and window to wall ratio are the most effective design components, while the u-values of external wall, roof and floor has less impact. Thermal mass does not lead to significant difference in annual energy consumption and thermal performance. The reduction of window g-value can increase the overall energy consumption due to the increase of heating demand to offset the reduced solar heat gain.

1 INTRODUCTION

Energy performance of care homes in the UK varies due to different building fabric and systems, inclusive function, daily occupancy and occupant behavior. Energy surveys were conducted in 52 care homes across UK which defined the benchmark for good practice at $247kWh/m^2$ for gas/ oil

and $79kWh/m^2$ for electricity, and for fair practice is $417kWh/m^2$ for gas/oil and $79kWh/m^2$ for electricity [2]. A survey conducted by Warwickshire county council from 2010 to 2011 concluded that the typical annual electricity consumption for a care home was $75kWh/m^2$ and good practice was $59kWh/m^2$; while typical gas consumption was $555kWh/m^2$ and good practice was $492kWh/m^2$ [26]. Another survey conducted in Scotland reported that the annual electricity consumption ranged from $46.6kWh/m^2$ to $178.8kWh/m^2$, while the gas consumption ranged from $274.5kWh/m^2$ to $557.1kWh/m^2$ [18].

Due to climate change and fuel poverty, countries around the world put forward targets to reduce carbon emissions and improve energy efficiency. The recast of the EU Energy Performance of Buildings Directive (EPBD) required Member States to achieve nearly zero-energy for all new buildings from the 1st of January 2021 and for all new buildings occupied and owned by public authorities to be nearly zero-energy from the 1st of January 2019 [6]. In 2006, the UK Government committed that from 2019 all new non-domestic buildings should also be zero carbon [7]. In addition, 'Energy Positive Building' concept emerges which represents a building can generate more energy than it consumes.

Panagiotidou and Fuller summarized the development of zero energy/ emission buildings concepts, starting from the first conception of a 'Zero Energy House' published by Esbensen and Korsgaard in 1977, to a broad discussion around the concept in 2007 [22]. The concept of nearly zero-energy buildings was introduced in the 2010 recast of the EPBD, and was defined as buildings that have a very high energy performance and the nearly zero or very low amount of energy required should be covered to a very significant extent by energy from renewable sources, including energy from renewable sources produced on-site or nearby.

In order to deliver nearly zero-energy buildings and energy positive buildings, it is important to adopt a holistic systems approach. The systems approach is a sustainable design method that integrates across site planning, architectural design and building engineering, and which can result in a nearly zero carbon building development. This can provide

lower running cost and contribute to reducing the demand on the energy supply infrastructure.

A holistic systems approach consists of 1) reducing energy demand, 2) using passive strategies, 3) applying efficient HVAC systems and 4) integrating renewable energy supply and storage systems [14] (Figure 1).
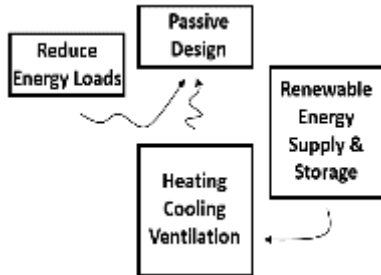
Figure 1. Four aspects of a systems approach [14]

The first aspect is to reduce the energy demand by using energy efficient equipment and lighting systems to reduce the direct energy usage as well as the heat gains to the space. The second aspect is to apply passive design strategies in relation to site planning, building form and building fabric, to utilize free energy (mainly solar and wind), and to reduce heat gain in the cooling season and heat loss in heating season. The third aspect is to further reduce the energy demand through the application of efficient heating, cooling and ventilation service systems. The last aspect is to integrate renewable energy systems in order to decarbonise the remaining energy requirements combined with batteries to provide local energy storage. Passive design strategies can contribute to reducing energy use by modifying the external climate to the benefit of improving the energy performance of the building. In addition, the reduced energy load leads to more options to integrate more efficient building systems with lower capacity. Furthermore, using renewable energy supply and storage systems becomes possible to offset the overall energy consumption.

2   PASSIVE DESIGN STRATEGIES

Passive design emphasizes on site planning, building shapecoefficient (external surface area to internal volume ratio), orientation, building envelope which related to insulation level, solar control, air tightness, thermal mass, daylight and natural ventilation. Passive design strategies have been widely applied to projects around the world, while a large number of associated research has been conducted by various researchers.

Watson and Labs stated that the reduction of building shape coefficient by using compact building geometry to reduce heat losses led to energy efficiency [30]. Depecker et al. pointed out that the energy consumption was inversely propositional to building shape coefficient only in cold,

overcast winters [4]. On the other hand, compact building shapes tend to lead to a deep floor plan which is not desirable for daylighting and natural ventilation, and the increased lighting and cooling load can offset the energy saving from heating load. In order to achieve sufficient daylighting, the plan should less than 18m deep [15], while the floor to ceiling height should be no less than 3m [3]. Design for natural ventilation should have a narrow plan (approximately 15m) [3]. Liddament concluded the maximum depth of the room is 2.5 times of the room height in order to achieve sufficient single sided ventilation, and 5 times of the room height for cross ventilation [16].

This research focuses on building envelope design since there is generally less flexibility associated with site planning, building geometry and orientation. In addition, building envelope has a great impact on energy efficiency and built environment, as well as the overall capital cost since they are an important component in the building structure [19]. Four aspects of envelope design have been explored: 1) insulation, 2) glazing system, 3) air tightness and 4) thermal mass.

2.1   Insulation

For a multi-layer building fabric, thermal properties are expressed by thermal transmittance, or U-value, which is the heat flow that passes through a unit area of a complex component or in homogeneous material due to a temperature gradient equal to 1K; and the unit is W/m2K (ISO 2007).The incidence of the thermal losses through the opaque walls on the whole energy losses of the building represents a large amount, so the use of adequately insulated walls has become essential [19]. The common insulation materials include stone wool, glass wool, expanded polystyrene, extruded polystyrene, phenolic foam, polyurethane, cellulose, cork, wood fibers, hemp, sheep wool, straw bale etc. There are advanced insulation materials with a much lower thermal conductivity, such as Vacuum Insulation Panels (VIPs, 0.004W/m/k), aerogels (0.013W/m/k) and Gas-Filled Panels (GFPs, 0.012W/m/k) [19]. Thin panels of these materials have good thermal performance which can contribute to a lighter weight structure and larger net floor areas; however, they are kept off main market due to high cost and real-life application.

The Building Regulations has set requirement for the minimum thermal performance of building envelope. The 1965 Building Regulations were made under the Public Health Act in 1961 and they mainly focused on health and safety issues [21]. Minimum U-values for the exposed walls and pitched roof were set to reduce the risk of condensation and mould growth in households rather than from the concern of saving energy. However, the range of issues covered by the Building Regulations has expanded considerably beyond the initial concern with the safety of a building's structure and its resistance to fire [12]. The escalating oil price in 1970s alarmed the world and prompted governments to seek secure sources of energy and

reduce dependency on imported fuel, which had a great influence on architectural design. In response to these crises, the emergence of building energy codes took place globally [23]. In the UK, the Building Regulations were reviewed in line with the requirement to decrease the energy consumption in buildings.

Currently, there are 14 parts (from Part A to Part P) in the Building Regulations. 'Part L: Conservation of fuel and power' is the key document addressing carbon reduction and energy efficiency in England and Wales. Table 1 summarizes the updating of the U-values in each revised version of the Building Regulations as an example showing the increased requirements.

Care homes are not dwellings which are defined as a self-contained unit (including a house or a flat) designed to be used separately to accommodate a single household. Therefore, new care home design should comply with guidance provide in Approved Document L2A.

The current mandatory requirement in Wales sets the minimum U-value requirement for roof, external wall, floor as $0.25W/m^2K$, $0.35/m^2K$, and $0.25/m^2K$ respectively.

Year	Categories	Roof	Wall	Floor	Window	Party wall	Air permeability	Compliance method
1965	Dwellings	1.40	1.70	NA	NA	NA	NA	Prescriptive
1976	Dwellings	0.60	1.00	1.00	NA	NA	NA	Prescriptive
1985	Dwellings	0.35	0.60	0.60	5.7	NA	NA	Prescriptive
	Nondwellings	0.6/0.7	0.6/0.7	0.6/0.7	5.7	NA	NA	
1990	Dwellings	0.25	0.45	0.45	5.7	NA	NA	Prescriptive
	Nondwellings	0.45	0.45	0.45	5.7	NA	NA	
1995	Dwellings	0.20/0.25	0.45	0.35/0.45	3.0/3.3	NA	NA	Prescriptive
	Nondwellings	0.25	0.45	0.45	3.3	NA	NA	
2002	L1	0.16-0.25	0.35	0.25	2.2/2.0	NA	10	Prescriptive
	L2	0.16-0.25	0.35	0.25	2.2/2.0	NA	NA	
2006	L1A	0.25	0.35	0.25	2.2	NA	10	Performance
	L1B	0.25	0.35	0.25	2.2	NA	NA	
	L2A	0.25	0.35	0.25	2.2	NA	NA	
	L2B	0.25	0.35	0.25	2.2	NA	NA	
2010	L1A	0.20	0.30	0.25	2.0	0.20	10	Performance
	L1B	0.16/0.18	0.28	0.22	1.6	NA	NA	
	L2A	0.25	0.35	0.25	2.2	NA	10	
	L2B	0.16/0.18	0.28	0.22	1.8	NA	NA	
2013 England	L1A	0.20	0.30	0.25	2.0	0.20	10	Performance
	L1B	0.16/0.18	0.28	0.22	1.6	NA	NA	
	L2A	0.25	0.35	0.25	2.2	NA	10	
	L2B	0.16/0.18	0.28	0.22	1.8	NA	NA	
2014 Wales	L1A	0.15	0.21	0.18	1.6	0.20	10	Performance
	L1B	0.15	0.21	0.18	1.6	NA	NA	
	L2A	0.25	0.35	0.25	2.2	NA	10	
	L2B	0.15	0.21	0.18	1.6/1.8	NA	NA	

Notes:
1. U-values are in $W/m^2.K$
2. The U-values for the existing building categories (L1B and L2B) are the standards for the new thermal elements. More relaxed values are set for the retained components.
3. Air permeability unit is $m^3/(h.m^2)$ @ 50Pa.

Table 1. U-values in the Building Regulations from 1965 to 2014 [13]

2.2 Glazing systems

Glazing system is an important component of building envelope since it determines heat loss from internal space to the external (U-value), solar heat gain into internal space (G-value) and daylighting condition (T-value). Total solar energy transmittance (G-value) is the fraction of solar radiation at near normal incidence that is transferred through the glazing by all means (both short wave and long wave components); where 100% represents 100% (all) solar energy passing through it, while 0% represents an opaque window with no solar energy transmittance (Pilkington,

2010). Light Transmittance (T-value) is the proportion of visible light at near normal incidence that is transmitted through the glass [24].

The current Building Regulations for care homes in Wales sets the minimum requirement for window U-value as 2.2W/m2K (which corresponds to an average performance of a double glazed window). Compared to external wall, windows are regarded as having a much lower thermal performance, therefore, to limit window to wall ratio has been considered to be an efficient passive design strategy. However, there is no prescriptive requirement for window to wall ratio, as well as for minimum G-value and T-value. The G-value of notional building recipe for compliance is 40%.

With the development of technology to assist in tackling fuel poverty and climate change, multi-layer glazing which consists 4, 6 and 7-pane glazing has a low U-value (ranging from 0.13-0.49 W/m2K) to reduce the heat loss and a low G-value (ranging from 10% to 35%) to reduce solar heat gain starts being used [29]. The application of multi-layer glazing can 1) reduce annual heating and cooling consumption to reduce the running cost, 2) reduce peak loads to decrease capacity of service systems and 3) improve thermal comfort due to better internal glazing surface temperature.

2.3 Airtightness

Air leakage is the uncontrolled flow of air through gaps and cracks in building fabric which can lead to unnecessary heat loss and discomfort from cold draughts. Air leakage can account for up to 50% of the total heat loss in a well-insulated house with a poor standard of airtightness[5].

In order to achieve high airtightness, early consideration on an air barrier line is important. A good airtight strategy is to have less penetration between the internal and external environment and all gaps properly sealed [5].

The current Building Regulations for care homes in Wales sets the maximum infiltration rate as $10m^3/s/m^2$@50Pa, while the infiltration rate of notional building recipe for compliance is $3m^3/s/m^2$@50Pa. For Passivhaus standard, the infiltration limit is 0.6ac/h @50Pa. The conversion between the infiltration rate at 50Pa and the infiltration rate at 0Pa is 1) to convert the infiltration rate from X $m^3/s/m^2$ at 50Pa to Y ac/h at 50 Pa, and 2) divide Y ac/h by 20 to get Z ac/h at 0 Pa.

2.4 Thermal mass

Utilization of thermal mass in buildings can be one of the most effective ways of reducing building heating and cooling loads. Thermal mass can convert solar energy into its own molecular motion and allow indoor air to remain cooler during solar gain and return some of absorbed energy to the space through radiation and convection when external temperature drops [1]. In addition, massive building envelope components delay and flatten thermal waves caused by exterior temperature swings. In heating

condition, thermal mass can be installed with a conservatory (such as a trombe wall system) to provide warm air to internal space when the sun is down. In cooling condition, thermal mass incorporated in the building fabric can absorb both solar and internal heat gains during day time to delay internal temperature rise. The heat can be taken away by night time ventilation. This approach is applicable in locations that have large daily temperature variations. However, several studies conducted indicated that passive design strategies with thermal mass was not suitable to cloudy winter climates like the UK [25].

Specific heat capacity[1], density[2] and thermal conductivity[3] are the relevant factors to determine the level of thermal mass of a material [9]. A material of 'high' thermal mass has a high specific heat capacity, high density and a moderate so that the absorption and release of heat synchronizes with the building's heating and cooling cycle. Heavy construction materials such as stone, brick masonry and concrete have a high level of thermal mass. On the other hand, lightweight construction materials tend to have a low level of thermal mass, including insulation materials, timber, steel, plaster board and carpet.

Both heavyweight and lightweight structures are commonly used in the construction of care homes in the UK (Figure 2). For the heavyweight structure, it can brick cavity wall with insulation in the middle, externally insulated brick/ block/ concrete wall, and internally insulated brick/ block/ concrete wall. For the lightweight structure, Structure Insulated Panels (SIPs) are composite building panels, consisting of a layer of platunum insulation (expanded polystyrene)sandwiched between two layers of structure Oriented Strand Board (OSB). SIPs can be used as both wall panels and roof panels. When used wall panels applied with 100mm brickwork with 60 clear cavity, it can provide a U-valude of 0.13W/m^2/K. Other advantages of this system include fast build times, good air tightness, avoiding cold bridging as well as fast erection [28].

3 METHODOLOGY

Building energy simulation is used to investigate the performance of a care home design. This takes account of the complex dynamic thermal interactions between the external environment, building fabric, internal heat gains,

---

[1] Specific heat capacity is the amount of heat needed to raise the temperature of one gram of mass by 1 kelvin, and the unit is J/kg/K.

[2] Density is the mass per unit volume of a material and is measured in kg/m^3.

[3] Thermal conductivity is the ease with which heat can travel through a material, and the unit is W/m.K.

HVAC systems, as well as renewable energy systems and storages if available.

HTB2 (Heat Transfer in Buildings version 2), a dynamic building energy model developed at the Welsh School of Architecture (Cardiff University), is used in this study. Based on weather data, building construction and layout details, shading masks, building services, and occupation profiles for people, equipment and lighting, it can simulate hourly and annual energy consumption and internal thermal conditions. It has undergone extensive testing, validation, including the IEA Annex 1 [8], IEA task 12 [17], the IEA BESTEST [20], and has been applied in many research and design projects [10].

This study is conducted in four steps:

1. A typical design of care home in the UK is selected and located in Cardiff. Design information in relation to internal gains (occupants, lighting and equipment) with consideration of energy efficiency and HVAC systems which meeting the minimum requirements of the Building Regulation Part L2A for Wales

2. A simulation matrix of seven variables in relation to building fabric is established, including 1) external wall u-value, 2) roof u-value, 3) ground floor u-value, 4) glazing u-value, 5) g-value, 6) window to wall ratio and 7) infiltration rate. The matrix consists of three levels of standard:
   - The minimum level represents the minimum limitation from the Building Regulation Part L2A for Wales
   - The average level represents higher requirements of the notional building recipe for compliance from the Building Regulation
   - The good practice level represents advanced practices

3. Building energy simulations are conducted using HTB2 for:
   (i) the energy and thermal performance of the typical care home with three types of structure (heavyweight, medium-weight and lightweight structure) at three levels of fabric efficiency (minimum, average and good practice level)
   (ii) based on the case with medium-weight structure at the minimum level of fabric efficiency, the energy performance of seven cases with one variable in the simulation matrix changed at a time at the average level, and the other seven cases with one variable in the simulation matrix changed at a time at the good practice level

4. The simulation results are compared and discussed, for the impact of each design variable of building fabric on annual energy.

4 BUILDING SIMULATION

4.1 Typical care home design for simulation

A typical care home design (total floor area 2030.5m^2) is selected with a floor to ceiling height 3.0m (Figure 3). Window to wall ration is 45%. The building shape (shape coefficient) is less compact which is considered suitable for daylighting and natural ventilation.
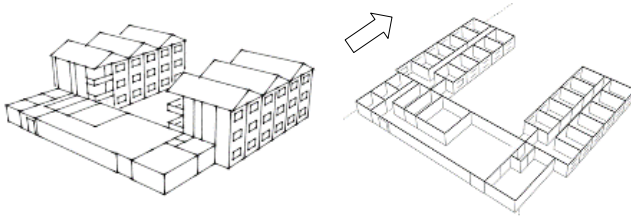
Figure 3. Typical care home design

Design data used in the simulation for the typical care home is summarized in Table 2.

Internal gains and ventilation				
Rooms	**Occupant** (W/m^2)	**Lighting** (W/m^2)	**Plug load** (W/m^2)	**Ventilation** (L/P/s)
Single rooms	3.5	6.3	1.0	10
Corridor	0	4.6	0	10
Reception	5.0	5.6	1.5	10
Office	4.8	11.6	3.0	10
Dinning/ lounge	22.1	4.6	1.0	10
Kitchen	7.8	12.5	30.0	30
Treatment room	14.2	11.6	10.0	10
Systems				
Gas boiler	SEDBUK=88%, Operating 00:00 - 24:00			
1. The building is only used 7 days a week, 24 hours a day; 2. The internal gain profile and lighting and plug load are based on good practice in SIA 2024. The occupant's load is based on the design. 3. The fabric energy efficiency U-value and systems, as well as the infiltration rate are based on British Building Regulation Part L A1. 4. The designed room temperature used in the simulation is 22 (winter) to 24ºC (summer) in accordance to CIBSE Guide A.				

Table 2. Simulation input data for the typical care home

The simulation explores the energy and thermal performance of the same typical care home in heavyweight, medium weight and lightweight structure. In order to make the simulation results comparable, the U-values of the components in these three types of structure are kept the same.

4.2 Simulation matrix with seven variables of building fabric and three levels of standard

The matrix for investigation on various design parameters of building envelop (Table 3).

Design parameters	Minimum level	Average level	Good practice level
External wall U-value (W/m^2/K)	0.35	0.26	0.17
Roof U-value (W/m^2/K)	0.25	0.18	0.11
Ground floor U-value (W/m^2/K)	0.25	0.22	0.19
Glazing U-value (W/m^2/K)	2.2	1.6	1.0
Glazing G-calue	60%	40%	20%
Window to wall ration	45%	35%	25%
Infiltration rate @50Pa	10m^3/h/m^2 (0.16ac/h @0Pa)	3m^3/h/^{2} (0.048ac/h @0Pa)	0.6ac/h (0.03ac/h @0Pa)

Table 3. Simulation variables and levels

The U-values, G-value and infiltration rate for the average case are from the notional building recipe, while the values for the good case are good practice from advanced schemes.

4.3 Simulation cases

Two sets of building simulation have been conducted:

1. With heavyweight, medium weight and lightweight structure to explore the impact of three different structures on energy consumption and thermal performance at three levels of fabric standards: the minimum level, average level and good practice level.
2. Based on the reference case which is the medium weight structure with minimum level of fabric efficiency, to investigate the impact of each of the seven individual parameters with two levels: the average level and good level

5 RESULTS AND DISCUSSIONS

Figure 4 shows the simulation results of three levels of standard with a heavyweight, medium weight and lightweight structure. With the minimum requirements of building fabric from the Building Regulation, the simulation results suggest that for a typical care home with a heavyweight structure, the annual energy consumption is 134.3kWh/m^2/year. When the standard of building fabric improved to the average level, the annual energy consumption can be reduced by 18.2%, to 109.9kWh/m^2/year. With further improvement of the building fabric to advanced design parameters, the total yearly energy consumption can be reduced by 28.1% (compared to the minimum level case), to 96.5kWh/m^2/year.

In addition, the annual energy consumptions of the typical care home with a medium weight and lightweight structure are more or less the same with the one with a heavyweight structure. The annual energy consumptions of a medium weight are 134.7, 111.8 and 98.5kWh/m^2/year for the minimum level case, average level case and good practice case respectively. The annual energy consumptions of a lightweight are 136.6, 112.6 and 98.9kWh/m^2/year for the three cases respectively. The total annual energy consumption of the heavyweight structure is slightly lower than the medium weight structure (ranging from 0.3% to

2.0%) and lightweight structure (ranging from 1.7% to 2.4%) for all three levels of fabric efficiency.

Furthermore, energy savings achieved through improvement of building fabric are similar for the heavyweight, medium weight and lightweight structure. For the medium weight structure, with the design parameters of the notional building recipe, the total energy saving is 17.0%, while with the advanced design parameters, the total energy saving is 26.9%. For the lightweight structure, with the design parameters of the notional building recipe, the total energy saving is 17.6%, while with the advanced design parameters, the total energy saving is 27.6%. If more advanced multi-lay glazing applied (U-value=0.3W/m^2/K, G-value=15%), a total 30.9% and 31.6% energy saving can be achieved for the medium weight and lightweight structure respectively.
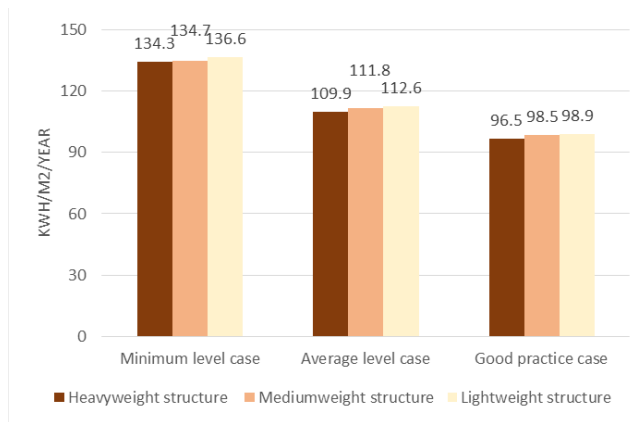
Figure 4. Comparison of annual energy consumption (kWh/m2/year) of three cases between a heavyweight, medium weight and lightweight structure.

The simulation results of thermal comfort indicate there is little difference in thermal comfort among the three types of structure for both summer and winter conditions (Figure 5). During a typical summer day, the internal air temperature can be maintained at 24°C, while the mean radiant temperature (MRT) reaches 2°C higher than the air temperature. During a typical winter day, the MRT remains lower (maximum 1°C) than the air temperature. In addition, the lightweight structure has the highest MRT during summer and lowest MRT during winter.
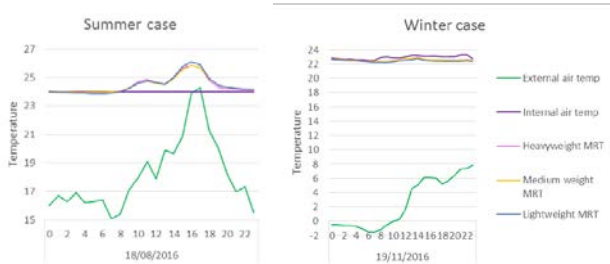
Figure 5. Comparison of thermal comfort of the heavyweight, medium weight and lightweight structure at the average level case

Figure 6 shows the simulation results of the impact of three levels of building fabric (minimum, average and good practice) on the annual energy consumption based on the medium weight structure. The major energy consumption is from heating, then lighting and equipment. Heating is required throughout the year, while cooling is only required from April to August. The annual energy consumption from heating with three levels of building fabric is 81.6kWh/m^2, 61.9kWh/m^2 and 51.5kWh/m^2 respectively. The annual energy consumption from cooling with three levels of building fabric is 9.1kWh/m^2, 5.9kWh/m^2 and 3.0kWh/m^2 respectively. It suggests that the cooling energy consumption can be nearly diminished when good practice level of building fabric applied, while the heating energy consumption can be reduced by 35.2%. In addition, the annual energy consumption from lighting and equipment is 44.0kWh/m^2.

Minimum level

Average level

Good practice level

Figure 6. Comparison of the impact of three levels of building fabric design on annual energy consumption (kWh/m2/year)

Figure 7 shows the simulation results of the impact of each design parameter of building fabric on the annual energy consumption based on the medium weight structure. The simulation results suggest that the design parameter with the most impact on energy consumption is the window u-value. With only the improvement from double glazing with an average performance to with a good performance, and to a triple glazing with good performance, an overall energy saving of 7.1% and 12.6% can be achieved. In addition, the improvement of infiltration rate and window to wall ratio from the reference standard to the good standard can achieve 6.4% and 6.0% energy saving respectively. Furthermore, the u-values for external wall, roof and floor have less impact on energy saving. With the improvement from reference standard to good standard, the total energy saving is 3.5%, 3.1% and 1.5% respectively. However, reducing window g-value has a negative impact on annual energy consumption in the climate of Cardiff with dominant heating requirement. The reason is that more heating is need with the reduction of solar gain to the space.

on annual energy consumption in the climate of Cardiff with dominant heating requirement. The reason is that more heating is need with the reduction of solar gain to the space.
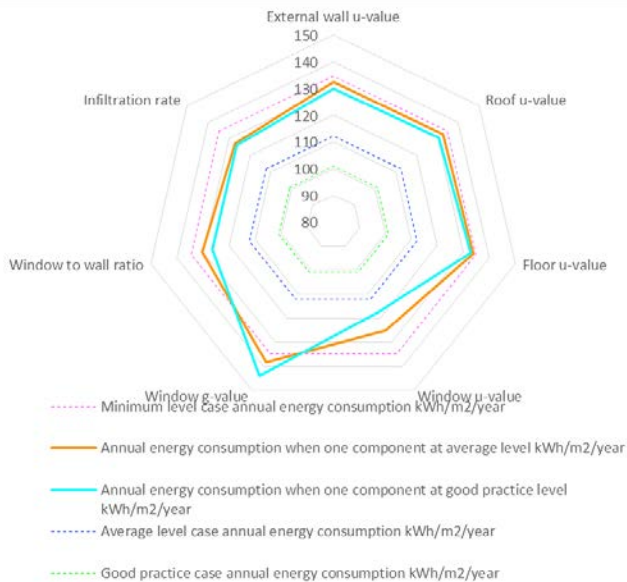
Figure 7. Comparison of the impact of seven variables of building fabric design on annual energy consumption (kWh/m2/year)

Figure 8 shows the energy saving percentage of annual energy consumption from each design parameter of building fabric.

The sum of seven individual percentage of energy saving from the notional building recipe is 16.9%, while the sum of energy saving percentage from the advance practice is 26.3%. Both sums of individual percentage of energy saving matches the overall energy saving percentage (17.0% and 26.9% for the average case and good case) when all the parameters applied.
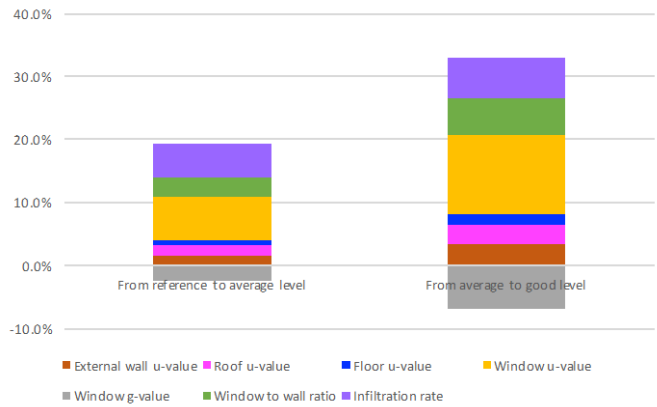
Figure 8. Comparison of annual energy saving (%) of seven variables of building fabric design

6 CONCLUSIONS

This study has explored the impact of passive design strategies on annual energy consumption of a typical care home in Cardiff. This study has identified that passive design has an impact on energy efficient design, which accounts for up to 28% of energy saving.

The most effective design components are window u-value, infiltration rate and window to wall ratio. The u-values of external wall, roof and floor has less impact. However, the reduction of window g-value can increase the overall energy consumption due to the increase of heating demand to offset the reduced solar heat gain.

In addition, in relation to annual energy consumption and thermal performance, there is no significant difference among a heavyweight, medium-weight and lightweight structure in a care home. The main reasons are the high cloud cover in the climate condition, as well as heating and cooling provided to maintain a constant thermal environment during 24/7 operation.

Future study on passive design of care homes can be conducted to explore the impact of passive wall systems on energy efficiency and carbon reduction, such as Trombe Walls, Autoclaved Aerated Concrete Walls, Double Skin Walls, and Green Walls.

Furthermore, passive design is not sufficient to deliver nearly-zero carbon care homes in the UK climate. However, it can bring the annual heating energy consumption to 51.5kWh/m^2, while reduce annual cooling requirement to near zero. It opens options to integrate more energy efficient systems (such as MVHR and heated surface system) and renewable energy supply and storage systems to deliver nearly zero-carbon care homes in Wales.

REFERENCES

1. Alexandra R, Rempel A W, Rempel, K R G, Barbara S. Climate-responsive thermal mass design for Pacific Northwest sunspaces, Renewable Energy, Volume 85 (2016), 981-993, ISSN 0960-1481.

2. BRECUS, ETSU. Energy consumption guide for nursing and residential homes (1996).

3. CIBSE. Natural ventilation in non-domestic buildings. London: CIBSE (1998).

4. Depecker P, Menezo C, Virgone J, Lepers. S. Design of buildings shape and energetic consumption. Building and Environment. 36. (2001), 627–635.

5. Energy saving trust. Enhanced Construction Details: Thermal bridging and airtightness (2009)

6. EPBD-CA. *TOWARDS 2020 – NEARLY ZERO-ENERGY BUILDINGS*. Retrieved from http://www.epbd-ca.eu/themes/nearly-zero-energy. As of 2014.

7. European Commission. Energy Performance of Buildings Directive. Retrieved from http://ec.europa.eu/energy/efficiency/buildings/buildings_en.htm. As of 2010.

8. Faber O. and partner. IEA Annex 1 Computer Modelling of Building Performance: Results and Analyses of Avonbank Simulation. St Albans: Faber O. and partner (1980).

9. Greenspec. Thermal mass. Retrieved from http://www.greenspec.co.uk/building-design/thermal-mass. As of 2016.

10. IEA. Integrated Building Concepts State-of-the-Art Review Report. Accessed online: http://www.ecbcs.org/docs/Annex_44_SotAr. As of 2006.

11. International Organization for Standardization,. ISO6946:2007 Building components and building elements – thermal resistance and thermal transmittance – calculation method (2007).

12. Killip, G. Building fabric and building regulations. Oxford: Environmental Change Institute, University of Oxford (2005).

13. Hou S S, Thomas A, Jones P,. A detailed review of the development of Building Regulations in relation to energy efficiency and carbon reduction in the UK. Urban Flux 12(4) (2015), 9-21.

14. Jones P, Hou S S, Li X. *Towards zero carbon design in offices: integrating smart facades, ventilation, and surface heating and cooling,* Renewable Energy. Vol. 73 (2015), 69–76.

15. Lehner N. Heating, cooling, lighting. Sustainable design methods for architects. Hoboken: John Wiley and Sons (2009).

16. Liddament M. A guide to energy efficient ventilation. Coventry: Air infiltration and ventilation centre (1996).

17. Lomas K J. Empirical validation of detailed thermal programs using test room data, Final report, Vol 1. International Energy Agency (1994).

18. McGhee G. Generic Energy Demand Profile Generation. A thesis submitted in fulfilment for the requirement of the degree Master of Science Sustainable Engineering: Renewable Energy Systems and the Environment. University of Strathclyde (2012).

19. Mendez Echenagucia T, Capozzoli A, Cascone Y, Sassone M. The early design stage of a building envelope. Multi-objective search through heating, cooling and lighting energy performance analysis. Applied Energy. 154 (2015), 577–591.

20. Neymark J, Judkoff R, Alexander D, Felsmann C, Strachan P, Wijsman A. IEA BESTEST Multi-zone non-airflow in-depth diagnostic cases. In: 12th IBPSA. Sydney (2011).

21. NPT. The Building Regulations [Online]. Available at: http://www.npt.gov.uk. As of 2014.

22. Panagiotidou, M, Fuller, R J. Progress in ZEBs: a review of definitions, policies and construction activity. *Energy Policy,* 62 (2013),196-206.

23. Perez-Lombard, L, Ortiz, J, Coronel, J F & Maestre, I R. A review of HVAC systems requirements in building energy regulations. Energy and Buildings, 43 (2011), 255-268.

24. Pilkington. Pilkington Glass Handbook, Retrieved from www.pilkington.com. As of 2010.

25. Porteous C, MacGregor A W K. Latitude myths challenged, in: Solar Architecture in Cool Climates, Earthscan Publishers, London UK (2005), 2-5.

26. Physical Assets, Maintenance, Minor Works, The Energy Team. Warwickshire County Council Annual Energy Review Building Energy Consumption April 2010 to March 2011 (2011).

27. Schiavoni S, D′Alessandro F, Bianchi F, Asdrubali F. Insulation materials for the building sector: A review and comparative analysis, Renewable and Sustainable Energy Reviews, Volume 62 (2016), 988-1011, ISSN 1364-0321.

28. SIPs WALES. SIPs Technology. Accessed on online: http://www.sipswales.co.uk/sips_technology.html. As of 2016.

29. Trimo. World's best-performing glass curtain wall: Qbiss Air (2015).

30. Watson D, FAIA, Labs K. Climatic building design. Energy-efficient building principles and practice. New York: McGraw-Hill (1993).

Design Tool Development

Flexible Parking - a Model for Calculating Parking Norms

Based on Urban Form and Accessibility Factors

Todor Stojanovski[1] and Yusak Susilo[2]

[1]KTH Royal Institute of Technology
Stockholm, Sweden
todor@kth.se

ABSTRACT

Sweden enacted ambitious environmental goals to mitigate climate change and cut carbon emissions. Swedish public agencies have promoted restricting parking norms to decrease automobile travel. This paper proposes a model to calculate flexible parking norms based on urban form and accessibility factors commonly used in architectural, urban design and planning practices. The transportation systems like the automobile need physical preconditions such as parking spaces and motorways that support driving around. The rationale is that parking norms can be lowered to zero if the urban form supports sustainable mobility choices, namely walking, cycling and public transportation. The model aims to inform municipality officials and developers, architects, urban designers and planners about urban form preconditions and integration with walking, cycling and public transportation as possibilities to reduce parking requirements and meet sustainable mobility goals.

Author Keywords

Parking; urban form; model; mobility choices; sustainable mobility.

1   INTRODUCTION

Many Swedish cities in the middle of the 20th century were designed to facilitate undisturbed circulation of private automobiles. Parking norms played important role to create the Swedish car society [1-2]. Many smaller cities and suburban neighborhoods in the larger Swedish cities today are highly dependent on the automobile and imports of oil. Cutting carbon emissions and mitigating oil dependence are most important issues in environmental debates concerning sustainable mobility in Sweden. Swedish public agencies have promoted restricting parking norms to decrease automobile travel [3-6]. Despite the fact that there is an extensive literature on parking [7-16] and empirical research on the effect of urban form (usually as D-variables such as density, diversity, design, etc.) on travel [17-38] there are no model to calculate parking norms based on urban form and accessibility factors that support walking, cycling and public transportation.

Parking problems are conventionally solved by predicting future demand and prescribing minimum parking requirements. The parking norms are typically intertwined with zoning ordinances, land uses and legislated by strict urban codes and building ordinances. This paper describes a new model to assess whether parking norms can be reduced based on available sustainable mobility choices, namely walking, cycling and public transportation. The model is based on analyzing integration of transportation modes as a set of urban form and accessibility factors used in architectural, urban design and planning practices. Some transportation modes such as the private automobile are easily integrated with parking spaces, whereas walking, cycling and public transportation demand density, walkable streets, commercial frontages, no building setback, proximity to transit stops with frequent service, and so on. Many of these urban form factors have also advocated by architects and urban designers to replace conventional land uses and two dimensional zoning with more complex Form-Based Codes (FBCs) and design guidelines in three dimensions [39-44].

2   PARKING NORMS

Municipal officials, urban planners and designers conventionally use parking ratios as part of zoning ordinances and building codes. Parking ratios are minimum or maximum parking requirements that provide number of parking spaces to match travel patterns and motorization trends (Figure 1A). For each car, typically one parking space is provided at home and three to four elsewhere [12, 45]. These numbers were copied between the USA and Sweden in the 1960s [1].

The rationale of predicting parking ratios in relationship to car ownership has not changed until today (Figure 1A). The concerns for managing future motorization and car ownership remains in the parking ratios calculus. Swedish planners predominantly work with zoning, land uses and building codes. The parking ratios are mandatory and are negotiated with the municipalities that usually own the land for development. The parking ratios are conceived as flexible norms (that can be adjusted), maxinorms or mininorms (maximum and minimum parking requirements). Stockholm Municipality [46] recommends green parking ratios of 0.3-0.7 parking spaces per dwelling to facilitate lower motorization rates in the future. Lower

parking ratios of 0.6 parking spaces per dwelling are typically introduced to curb automobile use in new sustainable neighborhoods like Hammarby Sjöstad in Stockholm, Sweden. There are new Swedish sustainable neighborhood experiments with zero parking norm.
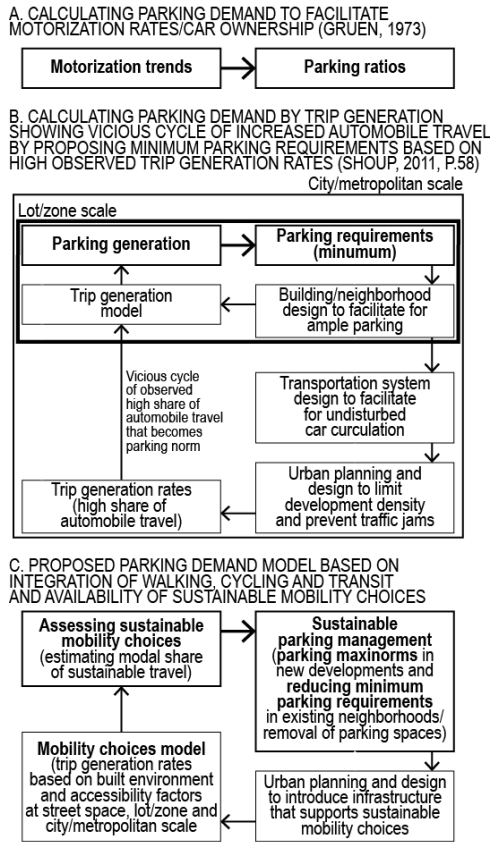
Figure 1. Methodologies to access parking demand and theoretical framework to gradually reduce parking or propose restrictive parking maximums in new developments by availability of sustainable mobility choice

In the USA, the land uses define minimum parking requirements. The Institute of Transportation Engineers (ITE) publishes and continuously updates the *Trip Generation Manual* [47] and *Parking Generation Manual* [48]. The minimum parking requirements calculated by observing and predicting trip and parking generation rates (Figure 1B). The handbooks compile empirical studies on trip and parking generation in different land uses. There is an extensive literature on parking and land use [7-16]. This literature criticizes the empirical evidence in these influential manuals e.g. low significance of the empirical evidence in the observation studies. The observations are often conducted in places where automobile travel dominates the modal split. As such they prescribe exaggerated minimum parking requirements based on high trip generation rates for the automobile. This creates a vicious cycle where high modal shares for automobile and parking generation rates are prescribed as minimums. New research aims to correct and enrich the normative

assumptions of ITE's handbooks by conducting observation in transit-oriented neighborhoods [49-53]

A problem in calculating parking demand is that trip and parking generation are observed on a scale of a lot/zone. The accessibility aspects at city/metropolitan scale that influence travel are neglected. The experiential qualities of the urban environments and street space within viewsheds are also not considered. The methodology in the following section describes the mobility choices model [54-55] that compiles urban form and accessibility factors at various scale to model integration with walking, cycling, public transportation and private car.

3   METHODOLOGY

Transportation professionals typically use land uses or density of development (usually as number of jobs and residents) to predict trip generation and parking rates [47-49]. The mobility choices model [54-55] builds upon theory of environmental perception and sustainable indicators to further develop the existing trip generation models. It also utilizes research on travel budgets for accessibility [56-57]. Figure 1 shows the steps in the mobility choices model.
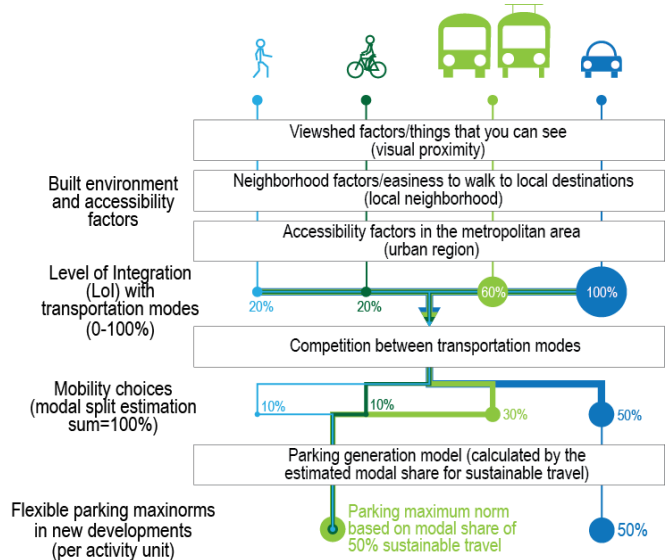
Figure 2. Methodologies to access parking demand and theoretical framework to gradually reduce parking or propose restrictive parking maximums in new developments by availability of sustainable mobility choice

The mobility choices model uses a measure Level of Integration (LoI) for each transportation mode that combines and weighs urban form and accessibility factors at three scales (Table 1). The LoIs illustrate preconditions to travel embedded in the built environment and they vary in complexity. A private automobile needs parking space at the destination and a quick access to an expressway. These two crucial factors give 100% integration. Walking, cycling and public transportation require a very complex combination of urban form and accessibility factors. The integration with public transportation include more than ten

weighed factors (including the factors for walking). Based on the LoIs for the transportation modes (0–100%) as precondition to travel, the modal choices model calculates modal shares proportionally. Better preconditions to travel with particular transportation modes means better integration with the buildings and that would arguably result in higher modal shares for that transportation mode.

Urban form and accessibility factors	Walking	Cycling	Public transport	Private car
Sidewalk design and continuity	(3) 5[1]			
Street segment length/city block width	(7) 15			
Speed limit	(3) 5[1]			
Bike infrastructures (racks, parking and cycling lanes)		(3) 20		
Bus line/busway/ tramway on street			(3) 5	
Transit stop/station exit on street			(3) 5	
Parking				(9) 60
No visible congestion				(3) 10
Building setbacks	(3) 5[1]			
Building height/street width ratio	(3) 5[1]			
Building façade activity/openness	(9) 20[1]			
Lot/block density (residents and jobs)	(9) 40[2]		(3) 5	
Neighborhood topography (slope)		(9) 40		
Access to everyday activities	(9) 20			
Access to event-type activities	(3) 5			
Access to a mix of activities	(9) 20			
Access to a local transit stop			(9) 30	
Access to a regional transit stop			(9) 30	
Access to an expressway				(5) 30
Bikable location		(9) 40		
			Walking (5) 20	
Sum	(49[1]/37[2]) 100	(21) 100	(32) 100	(17) 100

[1] assigned to street space
[2] assigned to city blocks/perimeter within building façades

Table 1. Integration with transportation modes (set of sustainable mobility indicators/urban form and accessibility factors).

In the first step of the method, the urban form and accessibility factors indicators are surveyed in the field or measured, analyzed and geocoded in GIS according to the methods presented in Table 2.

Urban form and accessibility factors	Method
Sidewalk design and continuity	I_1 is surveyed ($I_1 = 100$ is assigned for continuous sidewalks)
Street segment length/city block width	$I_2 = 200 - cbw_x$ where $cbw_x = \sqrt{cba_x}$ cbw_x City block width (width lower than 100 m and minimum 0 points for width over 200 m).
Speed limit	I_3 is surveyed ($I_3 = 100$ if speed limit = 30km/h)
Bike infrastructures	I_4 is surveyed (bicycle parking and cycling lanes on a street give $I_4 = 100$, $I_4 = 50$ if there are only bike racks or cycling lanes on the street)
Bus line/busway/tram way on street	I_5 is surveyed (street segments with bus lines receive $I_5 = 50$, whereas $I_5 = 100$ with busways/tramways on street)
Transit stop/station exit on street	I_6 is surveyed (city blocks with a transit stop/station exit on the streets receives $I_6 = 100$)
Parking	I_7 is surveyed ($I_7 = 100$ is assigned if there is visible parking)
No visible congestion	I_8 is surveyed (if there is no visible congestion $I_9 = 100$)
Building setbacks	I_9 is surveyed (building façade within 0.5m get $I_9 = 100$, between 0.5 and 5m $I_9 = 50$ and $I_9 = 0$ for over 5m)
Building height/street width ratio	I_{10} is surveyed (if the ratio is 1:3 or lower $I_{10} = 100$)
Building façade activity/openness	I_{11} is surveyed (if any part of the building façade is publicly accessible $I_{11} = 100$)
Lot/block density (buildings)	$I_{12} = \frac{q_{rj}}{100}$ q_{rj} residents and jobs per ha (if number of residents and jobs per ha > 100 then $I_{12} = 100$)
Neighborhood topography (slope)	I_{13} is calculated in GIS with raster map algebra method. Two raster maps with cost distance from the central points of the neighborhoods are created to calculate the travel ratio (TTR): 1) without slope; and 2) with slope degree penalty: no penalty was given for 0-0.5 degrees, 50% for 0.5-1, 100% for 1-2, 300% for 2-5, 400 % for 5-10 and beyond 10%-degree slope got 100 times penalty (1000%). By dividing the raster without and with slope penalty it is possible to see how difficult is to reach a destination. A TTR of 1 would mean two points on the map connect without slope obstacles, whereas 2 would mean 0-1% slope. I_{15} is normalized (0-100) with map algebra formula: $I_{13} = -10 \times ttr_x + 110$ (negative values are corrected to 0) ttr_x Travel time ratio in a cell of the raster map
Access to everyday activities	I_{14} is calculated in GIS. O-D matrix network analysis in ArcGIS is used to calculate distances from each supermarket, shop, restaurant, bar and so forth to every building in the neighborhood. Interpolation method (IDW) is used to calculate ranges. $I_{14} = 100$ if building is within 100 m (buffer tool is used), 60 if between 200-400 m, 30 if within 400-800 m network distance.

Access to event-type activities	I_{15} is calculated in GIS with the same method as I_{14}, just destinations included in this case churches, libraries and so forth
Access to a mix of activities	I_{16} is calculated in GIS. Service area network analysis in ArcGIS is used. Service area polygons within 400 m to entries with different land uses (shopping, culture, recreation, bars and restaurants, services, education and public spaces) are created. An overlay in GIS is used to sum up the total number of land uses: The polygons are converted in a raster map with following values: $I_{16} = 0$ (0-1 uses); $I_{16} = 25$ (2-3 uses); $I_{16} = 50$ (4-5 uses); $I_{16} = 100$ (6-7 uses).
Access to a local transit stop	I_{17} is calculated in GIS. O-D matrix network analysis in ArcGIS is used to calculate distances from local transit stops to every building in the neighborhood. Each local transit stop received a Transit Stop Performance Benchmark (TSPB) in respect to the frequency and type of service: $TSPB = \frac{\ln(f_{ts})}{\ln(6000)}$ (values over 6000 are corrected to 6000) f_{ts} Frequency at transit stop (weekly departures multiplied by 2 for commuter rail/subway/regional bus lines, 1.5 for local trunk buses and 1 for standard buses). The reference for the calculus (TSPB = 100) is a bus stop with six lines operating with 10 minutes headway (roughly 1000 departures/week).
Access to a regional transit stop	I_{18} is calculated in GIS is used with the same method as for access to a local transit stop, just for transit stops with regional service.
Access to an expressway	$I_{19} = 100$ if the neighborhood center is within 5 km to an exit to an expressway
Bikable location (regionally)	I_{20} is calculated by the formula: $I_{20} = -20 \times w_{cc} + 200$ w_{cc} Distance to the metropolitan core (in km) (if $w_{cc} > 10$ km then $I_{20} = 0$)

Table 2. Methods used to assess urban form and accessibility

The LoIs for the transportation modes are calculated by the formula (1):

$$LoI_m = \sum_{i=1}^{n}(wBEA_i \times BEA_i), \quad (1)$$

LoI_m Level of integration (0–100) for transportation mode m
$wBEA_i$ Weights for built environment or accessibility factor *I* (the weights are indicated Table 1)
BEA_i Built environment or accessibility factor *i*

In the second step, the modal shares are calculates as proportion of the *LoI* for a specific mode in respect to the sum of the *LoIs* for all modes. The formula is:

$$MS_m = \frac{LoI_m}{\sum_{i=1}^{n}(LoI_N)}, \quad (2)$$

MS_m Modal share for transportation mode m (in percentage)
LoI_m Level of integration (0–100) for transportation mode m
LoI_N Level of integration (sum for all transportation modes N)

Parking depends on the activity and the accessibility *(77)*. Activity denotes number of units such as dwellings, residents, floor space, etc. in the building/lot/zone. Accessibility includes mobility choices. Higher modal shares of sustainable mobility, lower parking maxinorms. To calculate maximum parking norms the formula includes activity and parking modifier with the share of private automobile. The formula is:

$$P_i = (MS_{car} - 30\%) \times p_i \times U_i \quad (3)$$

P_i Number of parking spaces per unit U_i (maximum parking norm if $U_i = 1$)
MS_{car} Estimated modal share of private automobile in a newly proposed development.
p_i Parking modifier to estimate maximum parking norm for one unit U_i
U_i Activity modifier (number of units e.g. if $U_i = 1$ dwelling and if $p_i = 4$ and m_{car} is 90% then the $P_i = 0.6 \times 4 = 2.4$ for 1 dwelling, see Table 3)

Table 3 shows the normative framework for the parking modifier p_i to estimate maximum parking norms as tradeoff of sustainable mobility choices. The private car is in a vicious circle in the modal split with the other sustainable transportation modes. The sum of walking, cycling and transit is always 100% minus the modal shape of private car. Car free development is possible if the estimated modal share of sustainable transportation modes (the sum of walking, cycling and transit) is over 70% (that means that the modal share of private car is 30%). This applies usually for areas that are close to major regional and local transit nodes, but also highly walkable and bikeable locations. To visually inform in the results about maximum parking norm (parking generation by one dwelling), the activity modifier is set to $U_i = 1$ unit (1 dwelling).

Estimated modal share of walking, cycling and transit)	Modal share of private car	Maximum parking norm
70%	below 30%	0 (car free possible)
40%	60%	1
10%	above 90%	2 (car oriented)

Table 3. Normative framework for parking maxinorm based on integration with sustainable transportation modes

4   TESTING THE MODEL IN THE CITY OF LULEÅ

Luleå is a small coastal city in northern Sweden. It is the seat of Luleå Municipality and the capital of Norrbotten County. The city has 75000 inhabitants in a county of 250000. Two neighborhoods are selected for the study, the downtown of Luleå and Kronan (Figure 2). The downtown of Luleå is a typical urban core of a small Swedish city. It has a grid street plan with a rectangular pattern of city blocks with courtyards. Kronan (The Crown) is a proposed neighborhood development roughly 2 km (1.5 mile) eastward from the downtown. The development plan for Kronan includes residential buildings (city blocks with courtyards as in the downtown) and a new square

surrounded by commercial storefronts and public buildings. The street plan is irregular. The new neighborhood will house 7000 inhabitants. Today, roughly 1500 people live in the area. Six residential towers were recently built on the hill westward from the newly planned development.
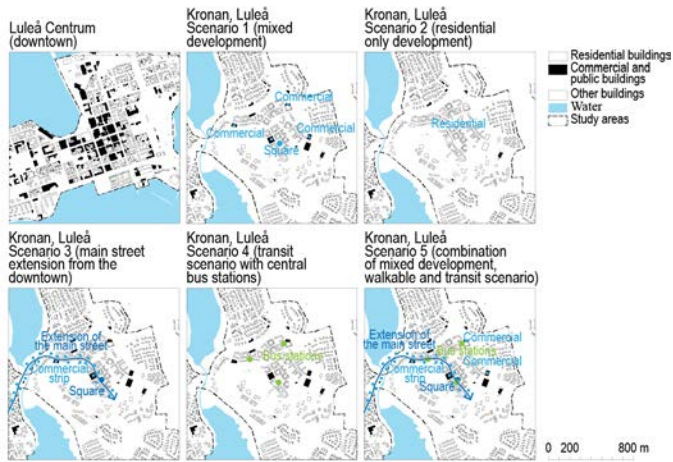
Figure 1. The study areas Luleå Centrum (the downtown) and five neighborhood development scenarios for Kronan

Five scenarios of Kronan are considered in this study. Scenario 1 illustrates the proposed mixed development with residential buildings, new square and commercial buildings. Scenario 2 shows a typical residential only development. In the smaller Swedish cities, the new developments are very dense residential. The demand for commercial in the suburbs is often low. Scenario 3 shows a walkable city scenario where the main street in the downtown extends 2 km (1.5 mile) eastward and connects the new square. It creates a commercial strip induced by the existing main street. Scenario 4 is Transit-Oriented Development (TOD) scenario. Three bus station are introduced with six bus lines operating with 10 minutes headway (with frequency of 1000 weekly journeys). Scenario 5 overlays the walkable city Scenario 3 and TOD scenario 4.

5   RESULTS AND DISCUSSIONS

This paper presents a model to calculate parking maxinorms based on availability of sustainable mobility choices. Figure 3 show heat maps of the analyzed sustainable mobility indicators from Table 1 and Table 2, Levels of Integration (LoIs) and modal shares for walking, cycling, public transportation and private car for two scenarios, Luleå Centrum (downtown) and the most probable Kornan Scenario 2 (residential only scenario). The mobility choices model analyzes a set of urban form and accessibility factors (Figure 3As) to visually inform integration with walking, cycling, public transportation and the private automobile (Figure 3Bs) and estimate modal shares proportionally (Figure 3Cs). The model shows hotspots. Greenest areas have highest integration and red worst. The downtown of Luleå is highly integrated with sustainable transportation modes (particular walking and cycling), in contrast to the dense residential only development in Kronan.
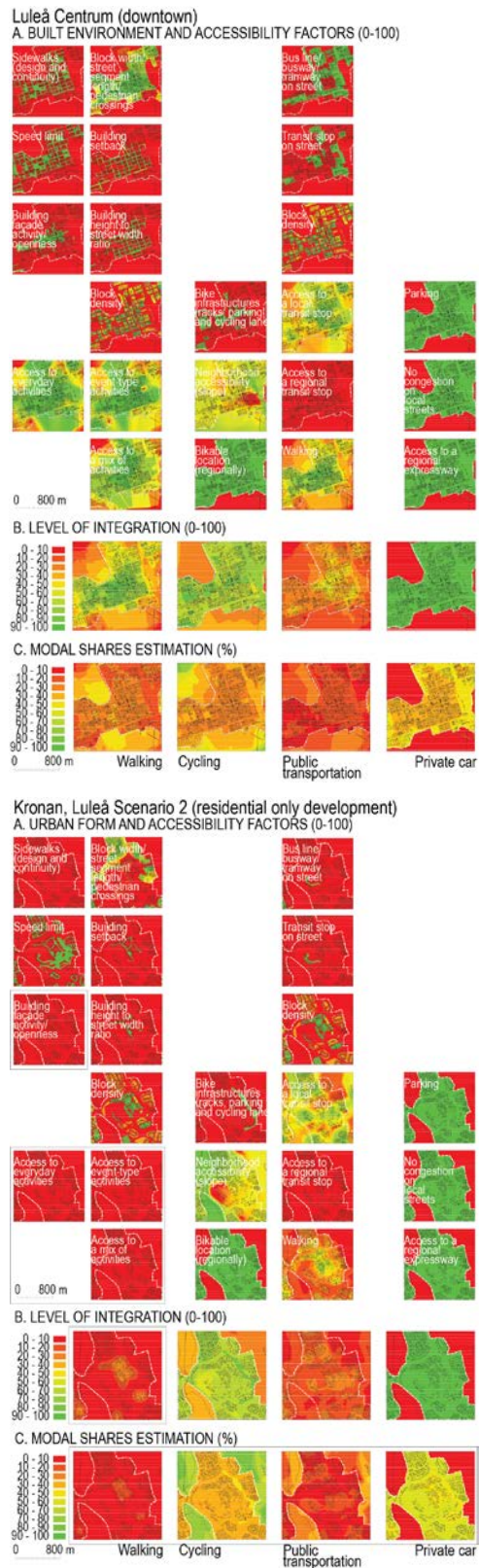
Figure 2. Heat maps of Luleå Centrum (downtown) and Kronan (Scenario 2, most probable) showing the built environment and accessibility factors, Levels of Integration (LoIs) and modal shares estimatio

Figure 4 visualizes the share of sustainable transportation modes (sum of walking, cycling and public transportation) alongside parking maxinorms based on the modal shares estimates for private car. In the proposed mixed development Scenario 1 in Kronan there is a sustainable mobility hotspot around the new square, but it quickly dissipates along the residential buildings. The hill on the east of the new development poses particular difficulties for pedestrians and especially bikers. The integration of sustainable transportation modes (mainly the biking possibilities to the downtown) produces estimates of modal shares between 40-60%. In the typical residential only Scenario 2 (the most probable scenario for small cities like Luleå) the walkability hotspot disappears and the estimates of sustainable modal shares drops by 10% (Figure 4A). The high density residential only development does not help. There is no reason to walk without commercial development, but it is still possible to bike to the downtown. Scenario 3 shows an increase of walkability along the extension of the main street. The commercial strip replicates a downtown and that reflects on the estimates of modal shares. Scenario 4 shows a sustainable mobility hotspot around the new square and bus stations. This area extends even further in Scenario 5 which overlays walkable main street with highly frequent bus stations. The integration of sustainable transportation modes in these sustainable mobility hotspot produces estimates of modal shares between 40-60% (like in Luleå Centrum, the downtown). The recommendations for maximum parking norms vary from 0-0.2 parking spaces/dwelling for the sustainable mobility hotspots to 1 parking space/dwelling for the towers on the hills and the residential only Scenario 2 Based on this estimates and formula (3), Figure 7B calculates and proposes parking maxinorms for new residences.

The model aims to contribute in shaping appropriate parking management strategies by informing municipality officials and developers, urban planners and designers about integration with walking, cycling and public transportation and possibilities to reduce parking requirements. The model captures orientation of buildings to sidewalks, cycling paths and transit facilities (as part of the FCB advocacy [39-44]), but also regional accessibility factors that derive from transportation and land use research. The conventional zoning, trip generation models based on land uses and minimum parking requirements are often simplistic and focus on a lot or zone, neglecting both three-dimensional aspects. The parking maxinorms maps (Figure 4B) show hot spots in the neighborhoods where zero parking can be introduced and under which conditions (very walkable streets, frequent transit stops, etc., see Figure 3 for urban and accessibility factors). Higher integration with walking, cycling and public transportation would allow for lower parking norms and would give more space for sustainable travel alternatives in cities. This can reflect in removal of parking spaces, creating more

walkable streets, orientation of buildings towards transit stops and constructing new cycling infrastructures. In a virtuous cycle for sustainable mobility, less parking would produce more walking, cycling and transit use, contribute to modal shift towards environmentally friendly mobilities and decrease in carbon emissions and oil consumption.
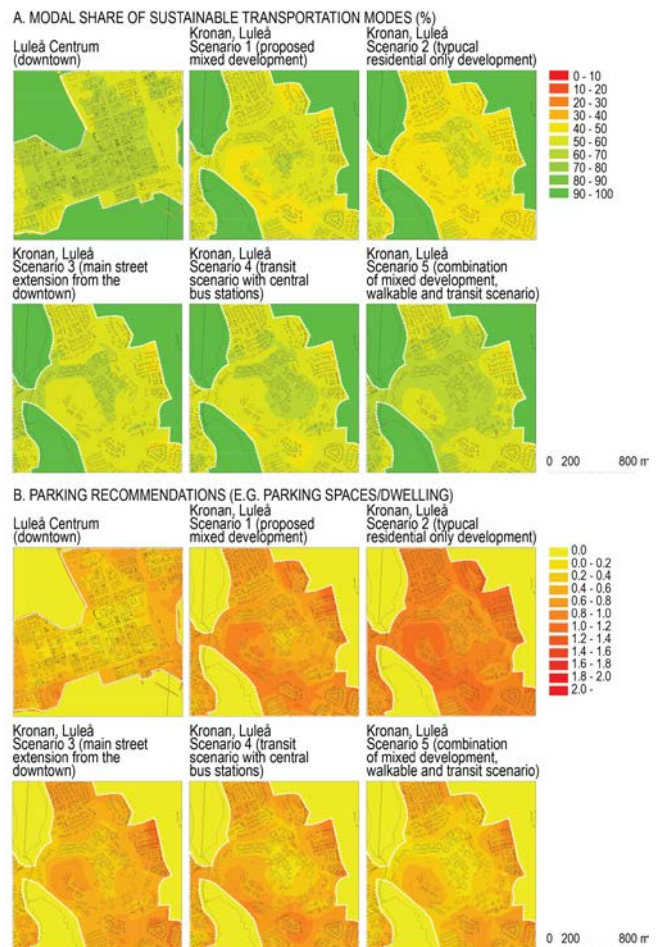
Figure 3. Heat maps of Luleå Centrum (downtown) and Kronan (Scenario 2, most probable) showing the built environment and accessibility factors, Levels of Integration (LoIs) and modal shares estimation

7 CONCLUSIONS

Achieving sustainable mobility is a major challenge in many cities that were designed for the automobile. This paper presents an analytical tool that visualizes integration with different transportation modes, estimation of modal shares and parking maxinorms. The visual information as heat maps aims to inform urban planning and design practitioners, municipal officials and developers about sustainable mobility choices. There are discussions about flexible parking norms in Sweden, but there are no practical models that can be used in the debate. Many attractive developments are located on hills or waterfronts that are not always walkable, bikeable and easily supplied with frequent transit services. Developing these neighborhoods with lack

of sustainable mobility choices contributes to additional automobile travel and oil dependence.

The model informs directly about sustainable mobility in a concise form (modal shares and parking recommendation) and shows a complex set of urban form and accessibility factors as background for the modal shares forecast. Both the concise and complex information are need to reflect upon the complex link between built environment and sustainable mobility both in the planning, development and parking management phases. Sustainable mobility as paradigm implies continuous managerial processes of monitoring and revising the parking norms and the set of sustainable mobility indicators.

8 ACKNOWLEDGMENTS

REFERENCES

1. Lundin, P. American numbers copied! Shaping the Swedish postwar car society. Comparative Technology Transfer and Society, 2(3), (2004), 303-334.

2. Lundin, P. Bilsamhället: ideologi, expertis och regelskapande i efterkrigstidens. Stokholmia forlag, Stockholm, 2008.

3. Trafikverket (author Envall, P). Parkering i täta attraktiva städer, 2013. http://fudinfo.trafikverket.se/fudinfoexternwebb/Publikationer/Publikationer_001701_001800/Publikation_001730/Parkering_i_t%C3%A4ta_attraktiva_st%C3%A4der_100_599_WEBB.pdf

4. Boverket. Parkering – ett effektivt verktyg för hållbar stadsutveckling 2018. https://www.boverket.se/sv/PBL-kunskapsbanken/Allmant-om-PBL/teman/parkering_hallbarhet/

5. Johansson, F., G. Henriksson & J. Åkerman. Parking Benefit Districts – The Transferability of a Measure to Reduce Car Dependency to a European Context. Transportation Research Part D: Transport and Environment, 56, (2017), 129-140.

6. Johansson, F., Henriksson, G., & Envall, P. Moving to Private-Car-Restricted and Mobility-Served Neighborhoods: The Unspectacular Workings of a Progressive Mobility Plan. Sustainability, 11(22), (2019), 6208.

7. Shoup, D.C. & Pickrell, D.H. Problems with parking requirements in zoning ordinances. Traffic quarterly, 32(4), (1978), 545-561.

8. Willson, R.W. & Shoup, D.C. Parking subsidies and travel choices: assessing the evidence. Transportation, 17(2), (1990), 141-157.

9. Shoup, D.C. An opportunity to reduce minimum parking requirements. Journal of the American Planning Association, 61(1), (1995) 14-28.

10. Willson, R.W. Suburban parking requirements: a tacit policy for automobile use and sprawl. Journal of the American planning association, 61(1), (1995), 29-42.

11. Shoup, D.C. The trouble with minimum parking requirements. Transportation Research Part A: Policy and Practice, 33(7-8), (1999), 549-574.

12. Shoup, D. The high cost of free parking. Updated edition (first published 2005). Routledge, New York, 2017.

13. Willson, R. Parking policy for transit-oriented development: lessons for cities, transit agencies, and developers. Journal of Public Transportation, 8(5), (2005), 79-94.

14. Litman T. Parking management best practices. American Planning Association, Chicago, 2006.

15. Cervero, R., Adkins, A. & Sullivan, C. Are suburban TODs over-parked? Journal of Public Transportation, 13(2), (2010), 47-70.

16. Willson, R. W. Parking reform made easy. Island Press, Washington, 2013.

17. Cervero, R. America's suburban centers: the land-use-transportation link. Boston, Unwin Hyman, 1989.

18. Cervero R. & K. Kockelman. Travel demand and the 3Ds: density, diversity, and design, Transportation Research Part D: Transport and Environment, 2(3), (1997), 199-219.

19. Southworth, M. Walkable suburbs?: An evaluation of neotraditional communities at the urban edge. Journal of the American Planning Association, (1997), 63(1):28-44.

20. Boarnet, M., & S. Sarmiento, S. Can land-use policy really affect travel behaviour? A study of the link between non-work travel and land-use characteristics. Urban Studies, 35(7), (1998), 1155–1169.

21. Crane, R., & R. Crepeau, R. Does neighborhood design influence travel?. Transportation Research Part D: Transport and Environment, 3(4), (1998), 225-238.

22. Boarnet, M. G., & R. Crane. Travel by design: The influence of urban form on travel. Oxford, Oxford University Press, 2001.

23. Cervero, R. Built environments and mode choice: toward a normative framework. Transportation Research Part D: Transport and Environment, 7(4), (2002), 265-284.

24. Boarnet, M. G. A broader context for land use and travel behavior, and a research agenda. Journal of the American Planning Association, 77(3), (2011), 197-213.

25. Crane, R. The influence of urban form on travel: an interpretive review. Journal of Planning Literature, 15(1), (2000), 3-23.

26. Ewing, R., & Cervero, R. Travel and the built environment: a synthesis. Transportation Research Record, 1780, (2001), 87-114.

27.Krizek, K. J. Residential relocation and changes in urban travel: Does neighborhood-scale urban form matter?. Journal of the American Planning Association, 69(3), (2003), 265-281.

28.Southworth, M. Designing the walkable city. Journal of urban planning and development. 131(4), (2005), 246-257.

29.Ewing, R., King, M. R., Raudenbush, S., & Clemente, O. J. Turning highways into main streets: Two innovations in planning methodology. Journal of the American Planning Association, 71(3), (2005), 269-282.

30.Naess, P. Urban structure matters: residential location, car dependence and travel behaviour. London, Routledge, 2006.

31.Ewing, R., & Handy, S. Measuring the unmeasurable: urban design qualities related to walkability. Journal of Urban design, 14(1), (2009), 65-84.

32.Krizek, K. J., Handy, S. L., & Forsyth, A. Explaining changes in walking and bicycling behavior: challenges for transportation research. Environment and Planning B: Planning and Design, 36(4), (2009), 725-740.

33.Ewing, R. & R. Cervero. Travel and the Built Environment. Journal of the American Planning Association, 76(3), (2010), 265-294.

34.Naess, P. "New urbanism" or metropolitan-level centralization? A comparison of the influences of metropolitan-level and neighborhood-level urban form characteristics on travel behavior. Journal of Transport and Land Use, 4(1), (2011), 25-44.

35.Næss, P. Urban form and travel behavior: Experience from a Nordic context. Journal of Transport and Land Use, 5(2), (2012), 21-45.

36.Lee S. & Talen E. Measuring Walkability: A Note on Auditing Methods, Journal of Urban Design, 19 (3), (2014), 368-388.

37.Papa, E., & Bertolini, L. Accessibility and transit-oriented development in European metropolitan areas. Journal of Transport Geography, 47, (2015), 70-83.

38.Stojanovski, T. How density, diversity, land use and neighborhood type influences bus mobility. Journal of Transport and Land Use, 11(1), (2018), 769-789.

39.Duany, A., & Talen, E. Transect planning. Journal of the American Planning Association, 68(3), (2002), 245-266.

40.Ben-Joseph E. The code of the city: standards and the hidden language of place making. Cambridge, Mass., MIT Press, 2005.

41.Walters, D. R. Designing Community: charrettes, master plans and form-based codes. Oxford, Architectural Press, 2007.

42.Talen, E. Design by the rules: The historical underpinnings of form-based codes. Journal of the American Planning Association, 75(2), (2009), 144-160.

43.Marshall, S. (ed.). Urban coding and planning. Abingdon, Oxon, Routledge, 2011.

44.Talen, E. Zoning for and against sprawl: the case for form-based codes. Journal of Urban Design, 18(2), (2013), 175-200.

45.Gruen V. Centers for the urban environment: survival of the cities. New York, Van Nostrand Reinhold, 1973.

46.Stockholm City. Riktlinjer för projektspecifika och gröna parkeringstal i Stockholm för bilparkering, Stockholm City, Stockholm, 2015.

47.ITE (Institute of Transportation Engineers). Trip Generation Manual (first published 1976), 10th edition. Washington, D.C., ITE, 2017.

48.ITE. Parking Generation Manual, 5th edition (first published 1985), Washington, D.C., ITE, 2019.

49.Currans, K.M. and Clifton, K.J. Using household travel surveys to adjust ITE trip generation rates. Journal of Transport and Land Use. 8(1), (2019), 85–119.

50.Clifton, K.J., Currans, K.M., Cutter, A.C. & Schneider, R. Household travel surveys in context-based approach for adjusting ITE trip generation rates in urban contexts. Transportation Research Record, 2307(1), (2012), 108-119.

51.Clifton, K.J., Currans, K.M. & Muhs, C.D. Adjusting ITE's Trip Generation Handbook for urban context. Journal of Transport and Land Use 8(1), (2015), 5-29.

52.Ewing, R., Tian, G., Lyons, T. & Terzano, K., Trip and parking generation at transit-oriented developments: Five US case studies. Landscape and Urban Planning, 160, (2017), 69-78.

53.Tian, G., Park, K. & Ewing, R. Trip and parking generation rates for different housing types: Effects of compact development. Urban Studies, 56(8), (2019), 1554-1575

54.Stojanovski, T. Urban Form and Mobility Choices. Informing about Sustainable Travel Alternatives, Carbon Emissions and Energy Use from Transportation Sustainability, 11(2), (2019), 548.

55.Stojanovski, T. Urban Form and Mobility - Analysis and Information to Catalyse Sustainable Development. Doctoral dissertation, KTH Royal Institute of Technology, Stockholm, 2019.

56.Milakis, D., Cervero, R., Van Wee, B., & Maat, K. Do people consider an acceptable travel time? Journal of Transport Geography, 44, (2015), 76-86.

57.Milakis, D., & Van Wee, B. For me it is always like half an hour: Exploring the acceptable travel time concept in the US and European contexts. Transport Policy, 64, (2018), 113-12.

Solving Thermal Bridging Problems for Architectural Applications with OpenFOAM

Patrick Kastner[1], Timur Dogan[1]

[1]Environmental Systems Lab, Cornell University, Ithaca, NY, USA, pk373@cornell.edu

Abstract

Although recent advancements in computational architecture show promising capabilities, it remains difficult for architects to conduct advanced simulations due to the limited software interoperability. For thermal bridging analyses, the architectural community traditionally relies on specific software tools that are not integrated into a CAD environment. To integrate such analyses into the ongoing design process, we implement a software tool to run heat transfer simulations with OpenFOAM from Grasshopper and Rhinoceros. This paper presents an implementation for box-shaped geometries and compares its results to a thermal bridge analysis from a validated simulation engine. We show that OpenFOAM's chtMultiregionFoam solver is capable of accurately predicting temperature distributions in a geometry setup with 13 different regions and 8 different materials. In conclusion, we show that heat transfer studies can be highly automated and integrated into an iterative design process.

Author Keywords

Thermal Bridging, Grasshopper, Rhinoceros, OpenFOAM, Parametric Design, Computational Architecture

Introduction

With the onset of CAD tools, which aimed to reduce repetitive tasks in design, the field of computational architecture evolved. Soon after, parametric software platforms followed to control the dimensions of digital objects while bringing rapid variability into design (Picon et al., 2016). With a growing number of available simulation tools, the problem of CAD (Computer-Aided Design) software interoperability for the export and import of geometries between different software tools increased over time (Pal, 2016). Such interoperability impedes the use of advanced simulation tools during the fast-paced, iterative design process as re-applying geometric changes throughout the project development is often cumbersome and time-consuming, if not error-prone.

In regard to heat transfer analyses, there are many software tools available that range from free options such as *THERM*[1] and an additional GUI for *Ladybug Tools*[2], or *Elmer*[3] to paid options like *Flixo*[4], *HTflux*[5], *HEAT3*[6], *Psi-Therm*[7], *QuickField*[8], *SOLIDO*[9], and *AnTherm*[10]. Yet, most of those tools use individual file formats and CAD implementation which impedes the interoperability between architectural design and analysis. Further, the free options are inherently limited to two-dimensional (2D) analyses which might not be sufficient for some problem statements. To overcome those impediments, this study uses *Rhinoceros* and *Grasshopper*, a software suite that is widely adopted in architectural design, to establish a coupling with *OpenFOAM* to run heat transfer and thermal bridging analyses.

Heat transfer analyses in architecture

Building performance simulations are necessary to obtain green building certifications (ASHRAE, 2007; DIN, 2013). Although the criteria to be eligible for those certifications vary from country to country, most of them require a detailed analysis of the building's envelope to ensure minimized heat loss due to thermal bridging.

Thermal bridges

Thermal bridges are discontinuities in the building envelope that cause an increased overall thermal transmittance and thus unnecessary heat loss. We refer to them as structural thermal bridges if materials with different heat conductivities are conjoining. If this does not apply, we refer to them as

---

[1]https://windows.lbl.gov/software/therm
[2]https://www.ladybug.tools/
[3]https://www.csc.fi/web/elmer
[4]https://www.flixo.com/
[5]https://www.htflux.com/en/
[6]https://buildingphysics.com/heat3-3/
[7]https://www.psitherm.uk/
[8]https://quickfield.com/
[9]http://www.physibel.be/solido.htm
[10]http://www.antherm.at/antherm/EN/

geometrical thermal bridges. Thermal bridges in the building envelope increase the risk of interstitial condensation and mold growth if temperatures drop below the dew point and condensation occurs as a result. Material and insulation degradation, as well as discomfort due to temperature asymmetries, may be implications of interstitial condensation. Thus, it is recommended to avoid thermal bridges or at least understand and constrain their severity whenever possible. The overall quality of a building envelope is ensured by reporting the average thermal transmittance, also referred to as the U-value which for a single layer is defined as:

$$U = \frac{k}{d} = \frac{1}{R} \tag{1}$$

where d is the thickness of the layer, k is the conductivity, and R is the thermal resistance. More generally, for a composite wall with n layers, we are able to calculate the thermal transmittance by summing the reciprocal resistance of each layer and considering internal and external convective heat resistances:

$$U_{wall} = \frac{1}{R_{tot}} = \frac{1}{R_{s,ext.} + \sum\limits_{i=1}^{n} \frac{d_i}{k_i} + R_{s,int.}} \tag{2}$$

The influence of thermal bridges on the overall thermal transmittance of the building envelope can be estimated by evaluating the Ψ-value which is defined as:

$$\Psi = \frac{\Theta}{\Delta T} - \sum_{i=1}^{n} U_i \cdot d_i \tag{3}$$

where ΔT is the temperature difference and d_i is the thickness of the layer in question. The ΔT values used to assess thermal bridges are usually described in engineering standards such as DIN, and they are assessed under steady-state conditions. While the material properties and the temperature in Equation (3) are either given or can easily be chosen, it remains difficult to estimate the crucial heat flux Θ for real-world geometries without conducting detailed simulations. Thus, deducing that heat flux via simulations of the goal of this study. Once the heat flux is known, we are able to report the reduced thermal transmittance as:

$$U_{tot,red} = \frac{\sum_{i=1}^{n} \Psi_i \cdot d_i}{A_{tot}} + U_{wall,ref} \tag{4}$$

Governing equations of heat conduction

The governing equation for heat transfer problems can be derived from the energy balance over a differential control volume in Cartesian coordinates $dx\,dy\,dz$, where $\dot{E}_g$ denoted the generation term and $\dot{E}_{st}$ denotes the stored energy (Bergman et al., 2011):

$$\dot{E}_{in} + \dot{E}_g - \dot{E}_{out} = \dot{E}_{st} \tag{5}$$

Reformulation and simplification yield the "Heat Diffusion Equation" which can be solved both analytically and numerically.

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} + \frac{\dot{q}}{k} = \frac{1}{\alpha} \frac{\partial T}{\partial t}, \text{ with } \alpha = \frac{k}{\rho c_p} \tag{6}$$

Here, x, y, and z are the Cartesian coordinates, α is the thermal diffusivity, $\dot{q}$ is energy generation rate per unit volume, and c_p is the specific heat capacity. While an analytical solution may be feasible for simple 1D and 2D problems, 3D problems with complex geometries and boundary conditions preclude such approaches. In this study, we use the open-source software *OpenFOAM* (Open source Field Operation And Manipulation) to employ a finite volume method approach to numerically solving such heat transfer problems. Although there is a multitude of boundary conditions available in *OpenFOAM*, we exclusively use Dirichlet boundary conditions of the form:

$$T(x, y, z) = T_s \tag{7}$$

to specify the surface temperature on boundaries for this study. These boundary conditions do not account for any convective heat transfer resistance and are thus considered to be a simplified approach[11], see Equation (2).

Reference case & simulation setup

The reference case simulated in this study stems from an example project called "Wärmebrücke" (thermal bridge) that is available in the software tool HTflux[12]. This example exemplifies a structural thermal bridge due to a wall-to-floor junction where the insulation material is partly tapered. The simulation domain consists of 13 adjacent regions with 8 different materials summarized in Table 1. The temperature boundary conditions were assumed to be steady-state both internally and externally.

Methodology

Figure 1 illustrates the simulation steps that have been implemented in the *Grasshopper C#* environment. The steps necessary to run a simulation can be categorized as *pre-processing* (necessary as input by the architect), *simulation*, and *post-processing*. The geometry for this simulation is modeled in *Rhinoceros* and is further detailed in Figure 7 in the appendix.

---

[11]For an analysis according to the German DIN/ENEV, one would need to also specify $R_{s,int.}$ and $R_{s,ext.}$.

[12]HTflux reports to be "validated under the relevant standards EN ISO 10077-2:2007 and EN ISO 10211-2:2012" (Rüdisser, 2018).
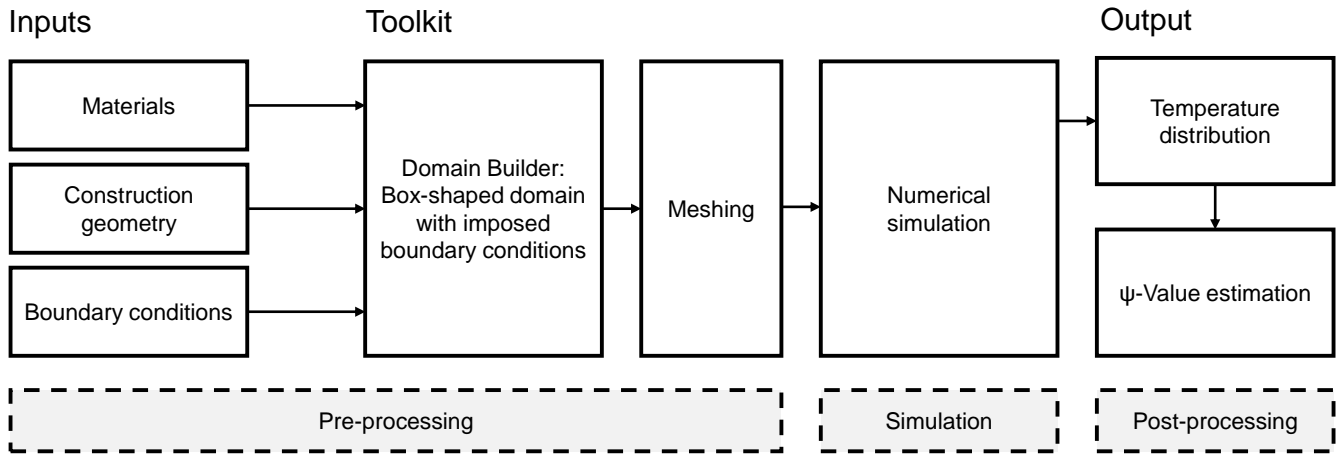
Figure 1: Flow chart of the simulation steps that have been implemented in the Grasshopper C# environment.

Table 1: Construction material properties used for the simulation domain illustrated below. The data was taken from the demo example "Wärmebrücke" (thermal bridge) that ships with the commercial software tool HTflux (Rüdisser, 2018).

	Construction materials	$k \left[\frac{\text{W}}{\text{m K}}\right]$	$c_P \left[\frac{\text{J}}{\text{kg K}}\right]$	$\rho \left[\frac{\text{kg}}{\text{m}^3}\right]$
	Concrete (1 % steel reinf.)	2.30	1000	2300
	Brick	0.08	920	710
	Insulation	0.039	1450	35
	Plaster (inside)	0.7	1000	1400
	Plaster (outside)	0.87	1100	1800
	Perimeter isolation strip	0.87	1100	1800
	Cement screed	1.33	1080	2000
	Impact sound insulation	0.041	1450	15

	Boundary conditions	$T \ [°C]$	BC type
	Inside temperature	20	fixedValue
	Outside temperature	-5	fixedValue
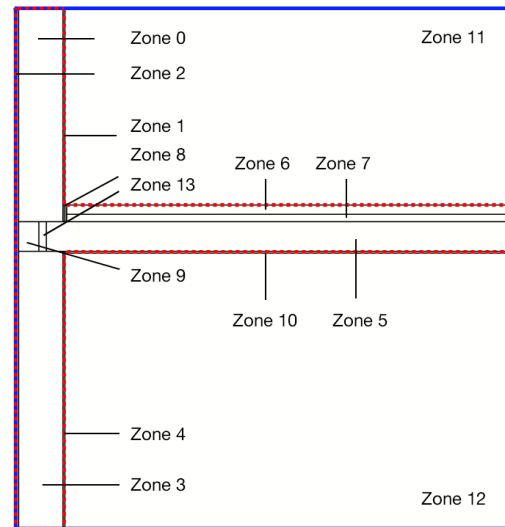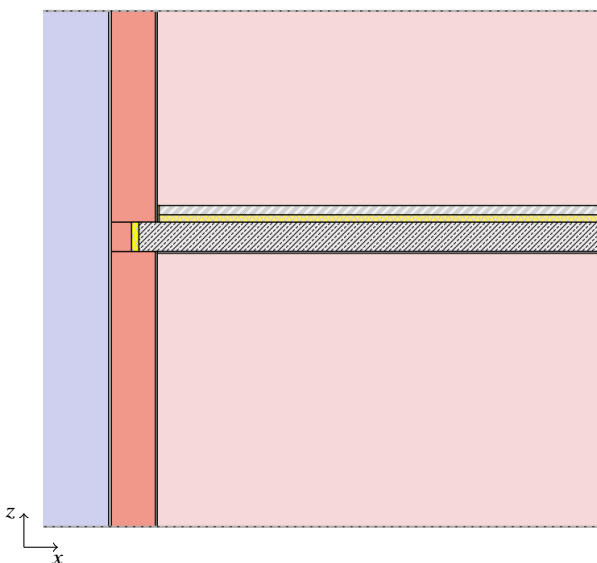
Figure 2: Imposed simulation boundaries for the background mesh (blue) and the internal fixed-temperature boundary conditions (red).

407

Pre-processing

During pre-processing, the materials the corresponding properties from Table 1 are chosen and assigned to each region via user dictionaries. All regions (including air) are modeled as extrusions with conjoining faces parallel to the x, y, and z-axis. In other words, every volume in the simulation domain has to be assigned to either a material with its corresponding boundary conditions or to air. However, it is not necessary to split the regions into sub-regions with congruent conjoining surfaces which means that some surfaces may have multiple adjacent zones (or faces), whereas other zones might exhibit one adjacent zone per surface. For instance, the zone with the material *Plaster (outside)* is adjacent to 3 *Brick* zones along the positive x-direction whereas the brick zone is adjacent to 1 *Plaster (outside)* zone along the negative x-direction, see Table 1. The flexible modeling of multi-adjacent zones simplifies both the modeling of the geometry in question and reduces the number of subregions and faces that need to be created during pre-processing. To create and export this zonal adjacency graph necessary to specify all boundary conditions between those zones, Open-FOAM features the tool *changeDictionary* which is capable of testing and reporting the "connectedness" between each zone. By coupling Rhinoceros and Grasshopper, however, we solve the cross-reference of all internal geometries from which OpenFOAM retrieves the combinations needed for the internal zonal adjacency graph. As a result, there is no need to make use of *changeDictionary*.

A unified boundary representation (Brep) of all 13 regions defines the dimensions for the background mesh. To create the background mesh, we use the *blockMesh* tool to create a voxelized mesh for the highlighted zone in blue in Figure 2. The resulting structured mesh consists of 1.9×10^6 cells and is refined in the x and z-direction. As OpenFOAM inherently operates on 3D simulation domains, this setup is to be seen as a section of a wall with symmetry boundary conditions in the y-direction, and as such representing a quasi-2D case setup. This quasi-2D case setup was chosen to allow for comparison against the validated tool *HTflux*. On all surfaces dashed in red, a "fixedValue" boundary condition was imposed which implies that the convective heat transfer resistance between wall and air on both sides was neglected.

Simulation

For the numerical simulation, we use the *chtMultiRegion-Foam* solver from the latest OpenFOAM 6.0[13] release which is capable of solving "steady or transient fluid flow and solid heat conduction, with conjugate heat transfer between regions, buoyancy effects, turbulence, reactions, and radiation modeling." (Weller et al., 2018). The simulation was run in a steady-state mode for 200 iterations until sufficient convergence was reached.

---

[13]https://github.com/OpenFOAM/OpenFOAM-6

Post-processing

For post-processing purposes, the case was sampled with OpenFOAM's internal *postProcess* utility and a "cellPoint" interpolation and plotted along the x-axis for $z = 1.58$ m and $z = 2.2$ m, see Figure 3. Finally, to calculate the heat fluxes for each region i in Figure 4, Fourier's law of heat conduction was used for which ∇T was sampled from the simulation results.

$$q_i = -k_i \nabla T \qquad (8)$$

Results

Figure 3 shows the temperatures plotted at $z = 1.58$ m and $z = 2.2$ m for simulations conducted with both *OpenFOAM* and *Htflux*. Similarly, the 2D temperature distributions are illustrated in Figure 6. For $z = 2.2$ m, we predict a piecewise linear temperature progression as one would expect for a steady-state case. For $z = 1.58$ m, we predict a piecewise, partly non-linear progression due to the difference in thermal conductivity for floor and ceiling materials. It is evident that OpenFOAM accurately predicts both the 1- and 2D temperature changes compared to the validated data simulated with *HTflux* for $T_{Htflux,h=0}$.

Figure 4 depicts the heat flux through the wall-ceiling junction. This visual representation indicated the effectiveness of the low-conductive insulation material which insulates the concrete from the outside. We also see the asymmetric temperature distribution in the ceiling and floor layers and the resulting asymmetric heat fluxes towards the insulation material. To calculate the final Ψ value, q may be integrated over the outside boundary and normalized by its length in m. The resulting Θ in W m^{-1} may the be used as input for Equation (3).

The temperature plots denoted with T_{HTflux} in Figure 3 depict the temperature distributions for which external and internal heat transfer coefficients h are considered which was not part of this study. The additional heat transfer resistances imposed are $R_{s,ext.} = 0.04$ m^2 K W^{-1} for the external boundary condition and $R_{s,int.} = 0.13$ m^2 K W^{-1} for the internal boundary conditions. For comparison, taking into account both additional resistances leads to a temperature difference of 2 °C at $x = 0.25$ m.

Figure 5 shows the temperature residuals in all zones.

Discussion

The results show that it is possible to use Rhinoceros and Grasshopper to streamline the pre- and post-processing of multi-zone architectural heat transfer studies.

The study focused on a linear thermal bridge, that is represented by a 2D box-shaped problem discretized with *blockMesh*. Only a limited number of real-world problems, however, may be represented with such simplified box-shaped geometries. OpenFOAM comes with meshing tools (*snappyHexMesh/cfMesh*) to produce high-quality, 3D,
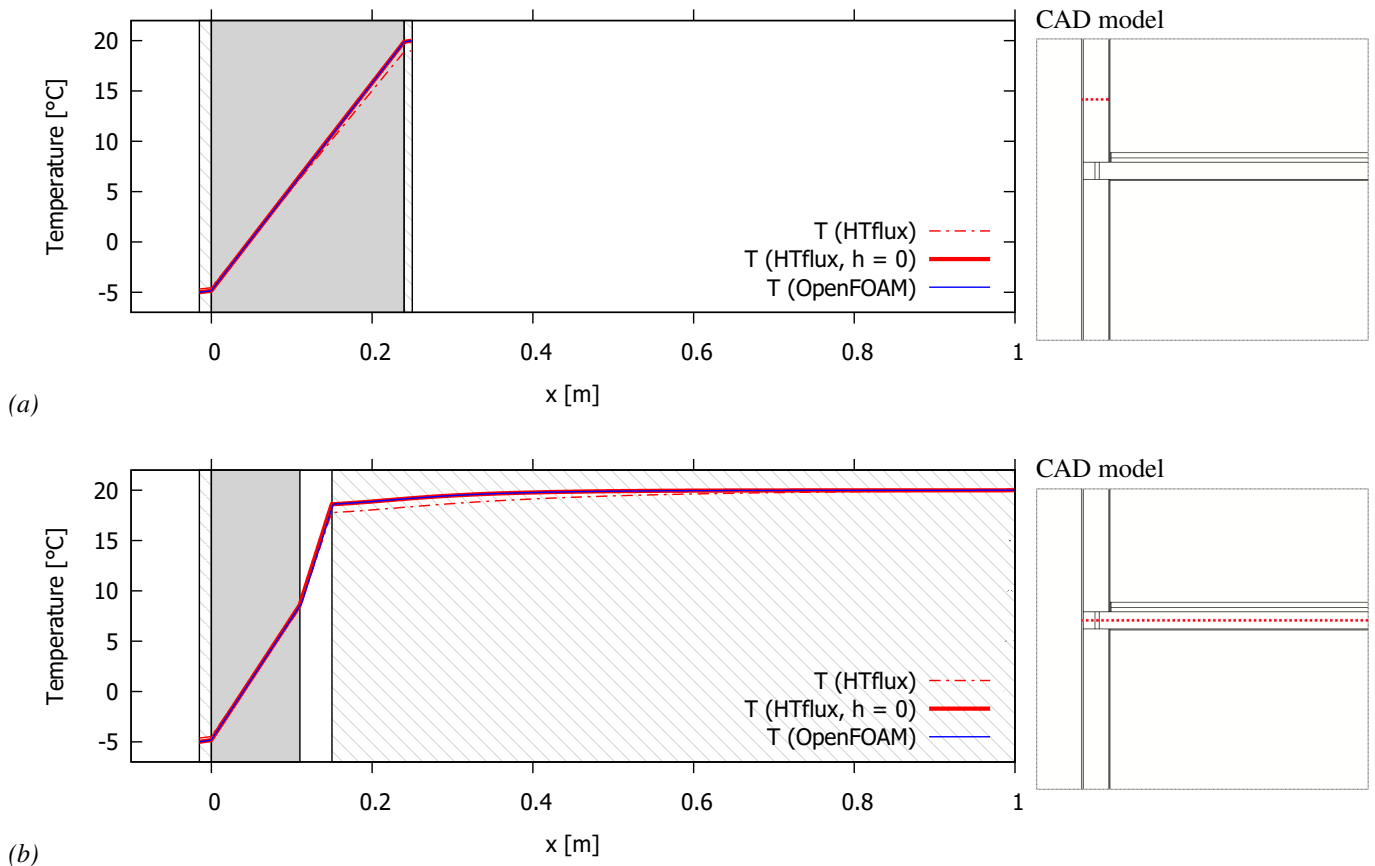
Figure 3: Temperature plot at z = 2.20 m (a) and z = 1.58 m (b). The plots $T_{h=0}$ refer to the variant for which the heat transfer coefficient was neglected. For the T_{HTflux} plots, $R_{s,ext.} = 0.04\ m^2\ K\ W^{-1}$ for the external boundary condition and $R_{s,int.} = 0.13\ m^2\ K\ W^{-1}$ for the internal boundary condition was assumed.

unstructured meshes by snapping arbitrary geometries to the background mesh that has already been used in this study. This may either be achieved by using the internal OpenFOAM tools or by importing meshes from other engines as there is an abundance of conversion algorithms available. While the possibility to produce such meshes exists, it should be noted that the snapping procedure adds a level of complexity. We mentioned that all regions need to constitute extrusions with conjoining faces along the three axes to ensure a robust convergence. As such, the usage of *snappyHexMesh* may likely introduce irregularities at conjoining surfaces of extrusions that in turn might lead to a significant decrease in simulation robustness.

Further, we have shown how the heat fluxes can be calculated by probing ∇T from the simulation results The necessary integration to compute the final Ψ values could be done either with OpenFOAM function objects or by manually probing and integrating Θ during post-processing. This could easily be extended to 3D cases to estimate the thermal transmittance χ of point thermal bridges.

Finally, we showed that imposing an additional convective heat transfer coefficient leads to a temperature difference of

2 °C at the internal wall of the wall-ceiling junction. In colder climates, sub dew point temperature levels may be reached at internal walls with unfortunate material configurations. In such cases, the simplified approach presented would significantly underestimate the temperature distributions. Apart from that, condensation in insulation materials might degrade their thermal resistance over time. A fully transient approach that takes into account moisture transfer should be considered to be the gold standard.

For highly-accurate results, it is, therefore, necessary to consider the full spectrum of heat transfer modes (conductive, convective and radiative) with a transient approach coupled with moisture transport.

Conclusion

In this study, Rhinoceros and Grasshopper were used as a pre-/post-processing interface to use OpenFOAM for heat transfer studies of box-shaped, multi-zone architectural geometries. By comparison with an example case from a well-established and validated software tool, it was shown that OpenFOAM's *chtMultiregionFoam* solver is capable of accurately predicting temperature distributions in a geome-
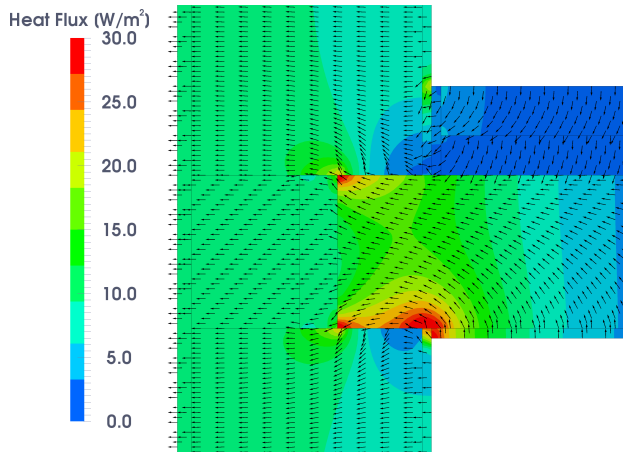
Figure 4: Heat flux through center regions calculated with OpenFOAM. It is evident how the insulation material impedes the heat flux from in the concrete from penetrating the outside layers of the building envelope.
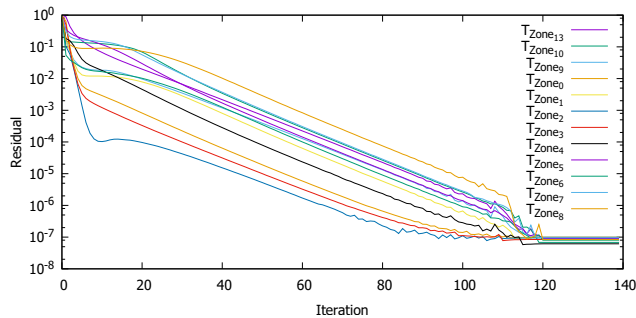
Figure 5: Temperature residuals of all zones.

try setup with 13 different regions and 8 different materials. In conclusion, heat transfer analyses with OpenFOAM can be highly automated and integrated into an iterative architectural design process. Although we were able to show excellent agreement of simulation results for a simplified set of boundary conditions, it is worth noting the limitations of the current implementation: thus far, only box-shaped geometries and fixed-temperature boundary conditions have been presented and compared to validated data which does not necessarily represent the temperature distribution that would occur in presence of convective and radiative heat losses/gains.

The *chtMultiRegionFoam* solver is not only equipped with transient simulation capabilities but also supports conjugate heat transfer. These capabilities should be explored to validate both the transient behavior and the interplay of conjugate heat transfer phenomena. Further, the illustration of heat flux through the wall-ceiling junction in Figure 4 was created manually by applying Fourier's law of heat conduction to every zone in the simulation domain. Since the heat flux is necessary to calculate the Ψ values, more time should be

spent on automatically extracting the heat flux from the simulation domain and the subsequent calculation of Ψ values. Moreover, future work should focus on the implementation of convective and radiative boundary conditions as well as complex 3D, non-orthogonal geometries.

Nomenclature	
α	Thermal diffusivity, $\mathrm{m^2\,s^{-1}}$
c_p	Specific heat capacity, $\mathrm{J\,kg^{-1}\,K^{-1}}$
d	Length, m
E	Energy, J
h	Heat transfer coefficient, $\mathrm{W\,m^{-2}\,K^{-1}}$
k	Thermal conductivity, $\mathrm{W\,m^{-1}\,K^{-1}}$
Ψ	Linear thermal transm. coeff., $\mathrm{W\,m^{-1}\,K^{-1}}$
Θ	Heat flux, $\mathrm{W\,m^{-1}}$
q	Heat flux density, $\mathrm{W\,m^{-2}}$
$\dot{q}$	Energy generation per unit volume, $\mathrm{W\,m^{-3}}$
t	Time, s
R	Thermal resistance, $\mathrm{m^2\,K\,W^{-1}}$
T	Temperature, °C
U	Thermal transmittance, $\mathrm{W\,m^{-2}\,K^{-1}}$
χ	Point thermal transmittance, $\mathrm{W\,K^{-1}}$
x, y, z	Cartesian coordinates, m

References

ASHRAE (2007). Ashrae standard 90.1.

Bergman, T. L., A. S. Lavine, F. P. Incropera, and D. P. Dewitt (2011). Fundamentals of heat and mass transfer.

DIN (2013). Din en 4108-2:2013.

Pal, S. (2016). Is the CAD Interoperability Problem Over? URL: https://www.engineersrule.com/is-the-cad-interoperability-problem-over/. Last visited on 2018/10/12.

Picon, A., A. Menges, and F. Aish (2016). Advancements in Design Computation. URL: https://www.aaschool.ac.uk/VIDEO/lecture.php?ID=3376. Last visited on 2018/10/12.

Rüdisser, D. (2018). HTflux: Hygric and thermal simulation. https://www.htflux.com.

Weller, H., C. Greenshields, and C. de Rouvray (2011-2018). Openfoam - The OpenFOAM Foundation. https://openfoam.org/.

Appendix

(see following pages)

410

Simulated temperature distributions
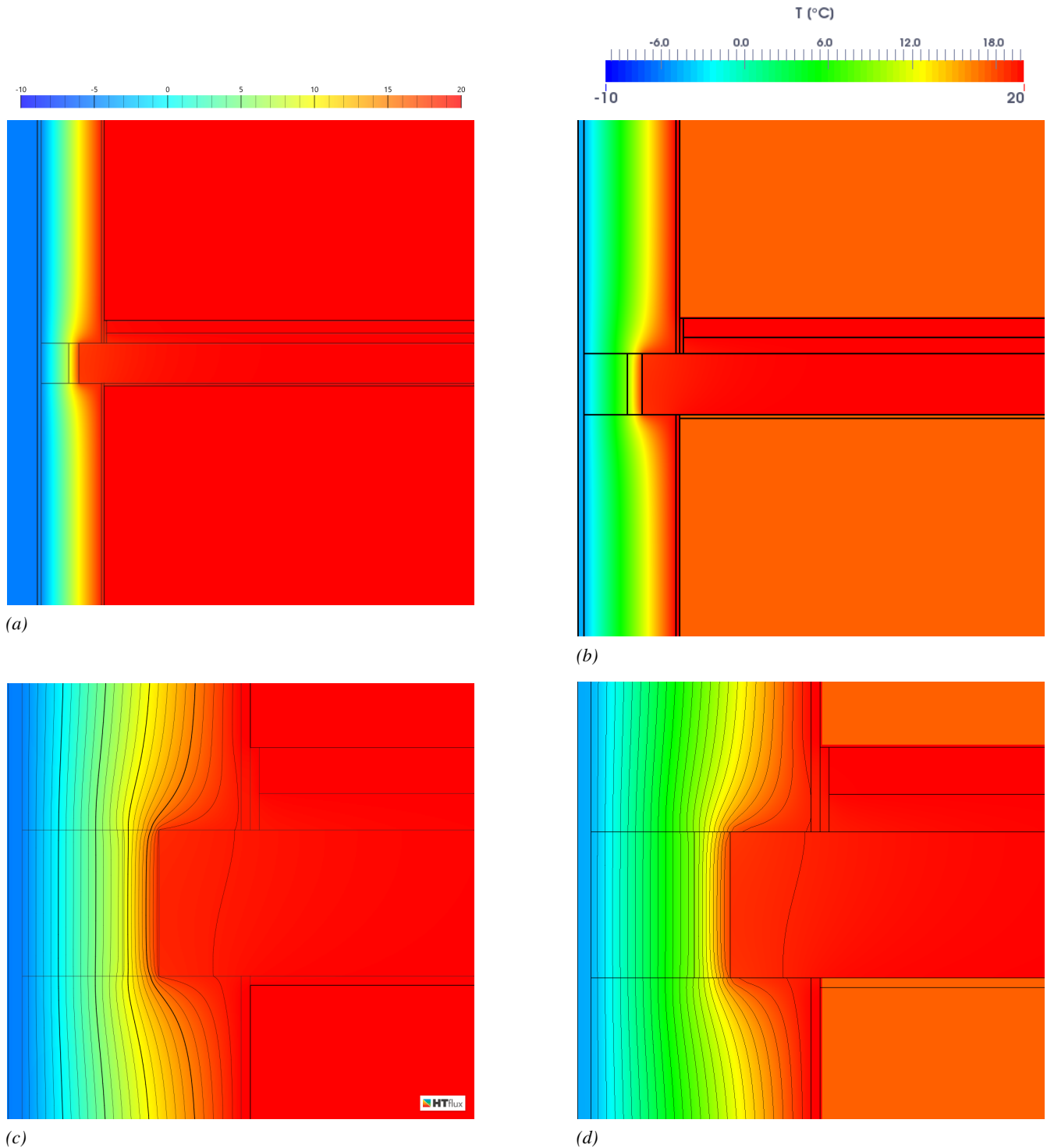
Figure 6: Temperature distributions of the analyzed geometry simulated with both HTflux (left) and OpenFOAM (right): (a) Overview HTflux; (b) Overview (OpenFOAM); (c) Geometric detail with isotherms (HTflux); (d) Geometric detail with isotherms (OpenFOAM). As can be seen by comparing left and right, OpenFOAM is able to accurately predict the reference temperature distribution.

CAD model
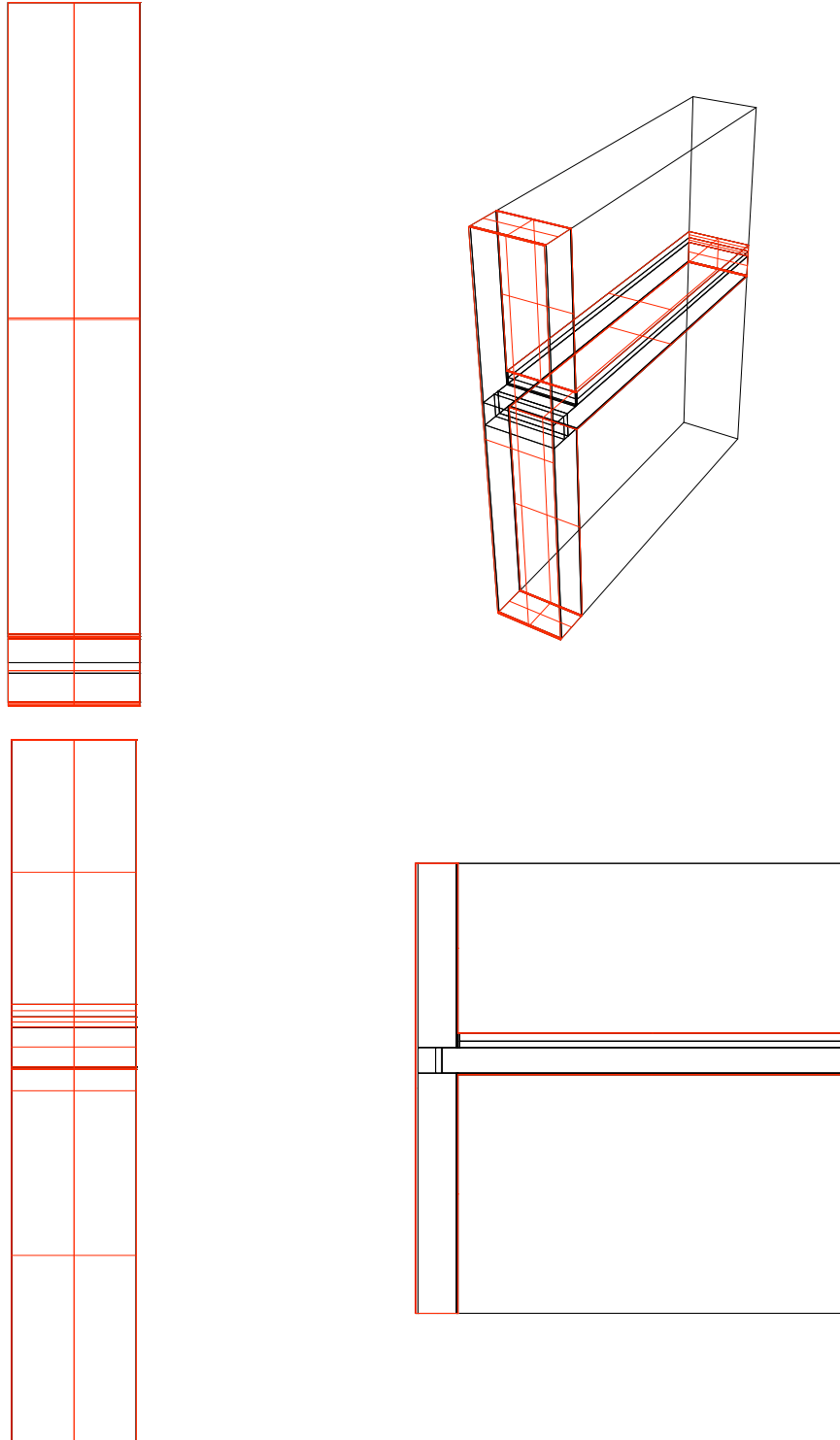
Figure 7: CAD model

SpeckleViz: A Web-based Interactive Activity Network Diagram for AEC

Paul Poinet[1], Dimitrie Stefanescu[1], Eleni Papadonikolaki[1]

[1]University College London
Bartlett School of Construction and Project Management
London, United Kingdom
{p.poinet, d.stefanescu, e.papadonikolaki}@ucl.ac.uk

ABSTRACT

As architectural design and construction projects tend to tackle larger scales and become more complex, the multiple involved disciplines in the Architecture, Engineering and Construction (AEC) sector often need to work globally from different remote locations. This increased complexity impacts digital design up until to fabrication workflows, which become more challenging and discontinuous, as each industry partner involved in the construction of a given project operates on different software environments and needs to access the precise fabrication data of specific design components. Consequently, managing and keeping track of design changes and data flow throughout the whole design process still remains a challenging task. This paper discusses how this particular challenge can be tackled through the development of a web-based interactive Activity Network Diagram (AND) - named SpeckleViz - that continuously maps the data transfers of the design and building processes, enabling the end-user to explore, interact and get a better understanding of the constantly evolving digital design workflow. Through this paper, the authors qualify an "end-user" as an advanced or expert user that performs complex geometry modelling tasks within wider collaborative workflows involving other advanced end-users. SpeckleViz (2020) is an application built upon Speckle (2020), an open-source data platform for the AEC. We illustrate the usefulness of interactive visualization of ANDs in the development of digital design workflows.

Author Keywords

Digital Workflow; Project Management; Data Visualization; User Interface; Activity Network Diagram.

ACM Classification Keywords

Project management techniques; Visualization techniques; Graph theory; User interface design; User centred design; Activity centred design.

1 INTRODUCTION

Design is a key phase across projects' lifecycle and collaboration among design actors is complex and crucial for project success. A recent report released by the Association of Project Management (APM) emphasizes the need for developing custom specific solutions to tackle contemporary large-scale and complex projects (Davies 2019). Indeed, contemporary design to manufacture process of large-scale and geometrically complex architectural projects remains a significant challenge, even though digital literacy keeps improving and computational design knowledge becomes more available.

It is not enough to be able to model complex geometry, but the design process must also be curated, shared and understood in more simple, transparent and intuitive ways than it is currently taking place within the Architecture, Engineering and Construction (AEC) industry. In this regard, previous research has demonstrated that simplifying intricate workflows through custom interactive visualization features enabled the end-users to better understand complex processes taking place across multiple stakeholders or software engineers.

In the AEC realm, current design processes are still segregated, and laborious manual interventions sometimes become a daily routine. In order to tackle this issue, custom management and visualization tools enabling better understanding and curation of complex projects activity networks have been proposed by different firms and individuals. Those proposed solutions converge towards the need for defining low level open-source infrastructures enabling more transparent collaborative workflows.

The present paper reviews both these contemporary solutions, the web-based open-source interoperability framework Speckle (described in section 3) that starts to gain traction in the built environment, and SpeckleViz: an interactive activity network diagram built within Speckle.

This research paper is divided into six sections: the present introduction, a state of the art in managing and visualizing complex digital design workflows (both in practice and through existing standards), an introduction to the Speckle framework, a description of the developed SpeckleViz interface, an illustration of the SpeckleViz interface through the description a selected case study, a discussion section and final concluding remarks with an outline of future works.
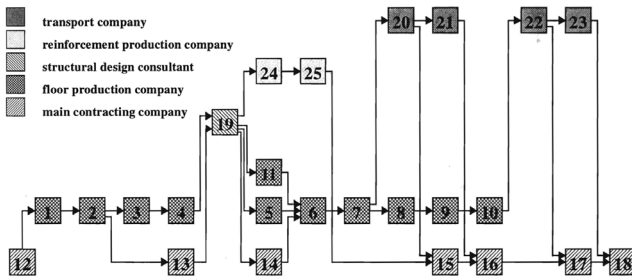
Figure 1. The Activity Network modelled as a PERT (Program Evaluation and Review Technique) network by de Vries (1995).

2 RELATED WORKS IN VISUALISING DESIGN PROCESSES

2.1 De Vries' Activity Network

Bauke de Vries (1995) conceptualized an Activity Network for the AEC industries in "Message Development in the Building Process". The author defined the activity network as follows: *"An activity network shows the information flow between activities. [...] The most important property of the network is that an activity can only start executing if all input channels contain the required information. The squares in figure 1 symbolize the activities. [...] The arrows symbolize the channels through which the information flows from one activity to another."* (de Vries, 1995). Such Activity Network can also be analogically compared with a Social Network, which has been defined by Wasserman and Faust (1994) as a "social structure" of actors (nodes) connected by one or more relations (ties), such as friendship or alliance. In the case of an Activity Network, the nodes represent activities and the ties represent data transfers – input(s)/output(s) – between the activities. Although more than 20 years old, De Vries' Activity Network for the AEC industries anticipated future design workflows and strategies that have been deployed since within practice. Those are illustrated within the next sections.

2.2 Front's Building Information Generation

During the realization of the City of Dreams Casino Hotel in Macau conceived by Zaha Hadid Architects, the consultancy practice Front developed a modelling strategy labelled "Building Information Generation" (Van der Heijden et al., 2015) enabling parallel generation of information and attributes necessary for further fabrication. In order to manage attributes and assign user data to the processed geometrical objects, Front developed an in-house a custom Rhino3D plug-in called "Elefront". The whole modelling process consisted of a strategic alternation between the generation of objects in Grasshopper and their subsequent storage and classification within staged Rhino3D models, within which the geometry is "frozen", thus devoid of any geometrical linkage. From these static, fixed geometrical objects were generated further information through a next iteration of Grasshopper sessions in which parametric linkage was kept. This process repeated itself until the last level of detail was modelled and ready for manufacture, fabrication and assembly (see Figure 2). In the words of the
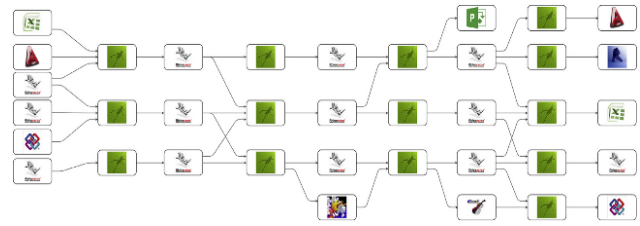
Figure 2. Project ecosystem of staged models and generating logic, across Rhino3D and Grasshopper files.

Computational Design Specialist at Front, this process is called "staging", consisting of a "discretization of logics", where many different smaller files are individually processed, instead of embedding intelligence within one single very large model. Each smaller file can be manually triggered and re-computed when a change occurs up-front, allowing therefore data propagation throughout the whole digital chain. Similarly, the developed SpeckleViz interface enable the end-user to obtain a clear overview on such discretization of logics that is operated within Speckle.

2.3 Woods Bagot's Metagraph

The architecture firm Woods Bagot has been developing internal methods and customized workflows for improving communication and software interoperability. Software platforms are not seen here as limiting the design possibilities but rather as an array of tools from which the architect can pick and choose, serving the project's needs to be designed and delivered.

Based on this approach, Woods Bagot developed Metagraph (see Figure 3), a data visualization tool based on Flux (a former interoperability platform described in the next section) data keys (or data transfer identifiers) which are used to represent all relationships between the different
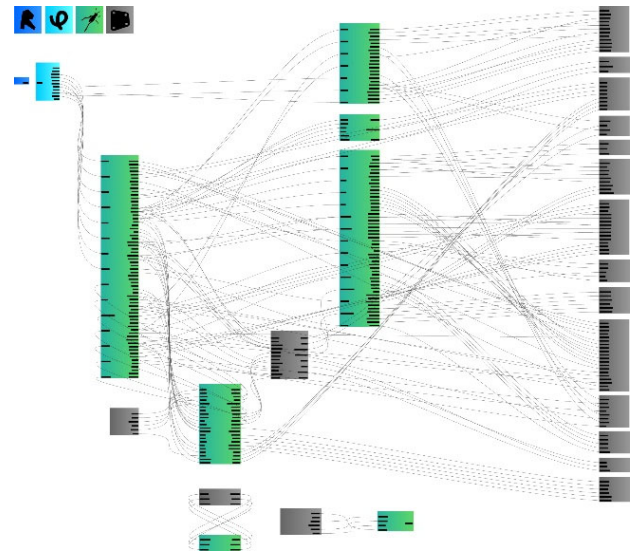
Figure 3. The Metagraph developed by Woods Bagot represents data key relationships among different scripts from different software platforms, using the Grasshopper canvas with its parameters and wire connections. (Ringley, 2017)

414

models, software and programming environments (Rhino, Grasshopper, Revit and Dynamo) constituting the same architectural project. The data exchange between these different software platforms are represented through a decentralized global network of nodes. While being useful for understanding and debugging models at systemic level, it is still being developed to allow for local-level programmatic control. Brian Ringley, former associate and global specialist at the design technology department of Woods Bagot, speculated that the next step of the Metagraph would be to *"set another layer of parametric intelligence where different scripts can adapt and update based on changes of other scripts."* (Ringley, 2017)

2.4 Flux Data Inc.

Flux Data Inc. was an interoperability platform which paved the way to transfer seamlessly building data across different software platforms. It became popular within the AEC community and was largely used by many architects, engineers, and consultants before it ceased software development on the 31st of March 2018. This event was unfortunate for a substantial part of the industry, especially for those who have built up their digital workflows upon this platform. Flux had four main management tools that were accessible from the site: "Community" enabled the users to help each other on an exchange platform, "Data Explorer" helped in assessing the project's workload, "Flow" allowed to visualize the different data flows within the same project, and "Projects" enabled the different disciplines to keep track and see the current status of the projects in progress.

Depending on the client's need, the appropriate applications could be downloaded and installed from the main website, enabling better communication between the specific software environments used by the office. If the company used a software platform that was not supported by the Flux applications, it was still possible to undertake third party software development through the available Software Development Kit (SDK). Flux software was being used by *"more than 6,200 companies in 151 countries and relied upon by Computational Designers, Engineers and sophisticated BIM professionals at Frank Gehry Partners, BIG, SHOP, Arup, BuroHappold Engineering, Thornton Tomasetti and more."* (Flux, 2018).

2.5 Existing Standards in Process Modelling

Paralelly to the practices described above, different standards have been developed by the industry to standardise the formal representation of exchanges and activities hapenning during a business process. For example, the Business Process Modeling Notation (BPMN) has been introduced in 2004 by Stephen A. White, in order *"to provide a notation that is readily understandable by all business users, from the business analysts who create the initial drafts of the processes, to the technical developers responsible for implementing the technology that will perform those processes, and, finally, to the business people who will manage and monitor those processes."* (White, 2004). The author has identify for different categories of elements: Flow Objects (Event, Activity, Gateway), Connecting Objects (Sequence Flow, Message Flow, Association), Swimlanes (Pool and Lane) and Artifacts (Data Object, Group, Annotation).

While the BPMN notation helps all business users to understand business processes more clearly, it reasonates very strongly with the Unified Modelling Language (UML), which focuses instead on modeling software system and acts as general-purpose visual modeling language intedend to specify, visualize and construct the artifacts of a software system. UML was developed in 1994-95 and was originally motivated by the desire to standardize the different notational systems and approaches to software design. UML also comes with its sets of graphical rules which helps in documenting a system model (Booch et al., 2005).

Not specially focused on business or software modelling processes but looking instead at timed event systems in a more general way, the Discrete Event System Specification (DEVS) has been introduced by Bernard P. Zeigler in 1976 to formalise the modeling and analysis of discrete event systems (DESs). A DES can be represented by a step function, and is defined as a non-linear process in which different events can happen in parallel, one event triggering an other in an asynchronous manner. This differs a lot from the more traditional, continuous imulations which could be represented by a continuous function in which optimization tasks are running one after the other (Ziegler et al., 2000).

As building design processes are not linear or continuous but discrete and intricate (as described in the previous subsections), the formalisms introduced by BPMN, UML and DEVS represent useful sources of inspiration to visualize Activities Networks in AEC. The next section introduces Speckle (2020), an open-source data platform for the AEC which enables the deployment of SpeckleViz, the Activity Network Diagram described in section 4.

3 SPECKLE, AN OPEN-SOURCE DATA PLATFORM FOR AEC

Speckle (2020) differentiates itself from commercial web-based interoperability platforms by proposing a complete open-source data framework for architects, designers and engineers. Speckle was originally developed at University College London in 2016 by Dimitrie Stefanescu. Speckle does not enforce a predefined topology of communication patterns, but rather allows for the emergence (and analysis) of meaningful data-driven dialogue amongst the different actors involved in the design process

With regards to schemas, Speckle, in contrast with the existing industry standard IFC (Industry Foundation Classes), promotes composability over completeness and provides a programmatic infrastructure for end-users to define their own, domain-, company-, or even project-specific, object models. Furthermore, Speckle can support pre-existing object models (such as IFC) "out of the box",
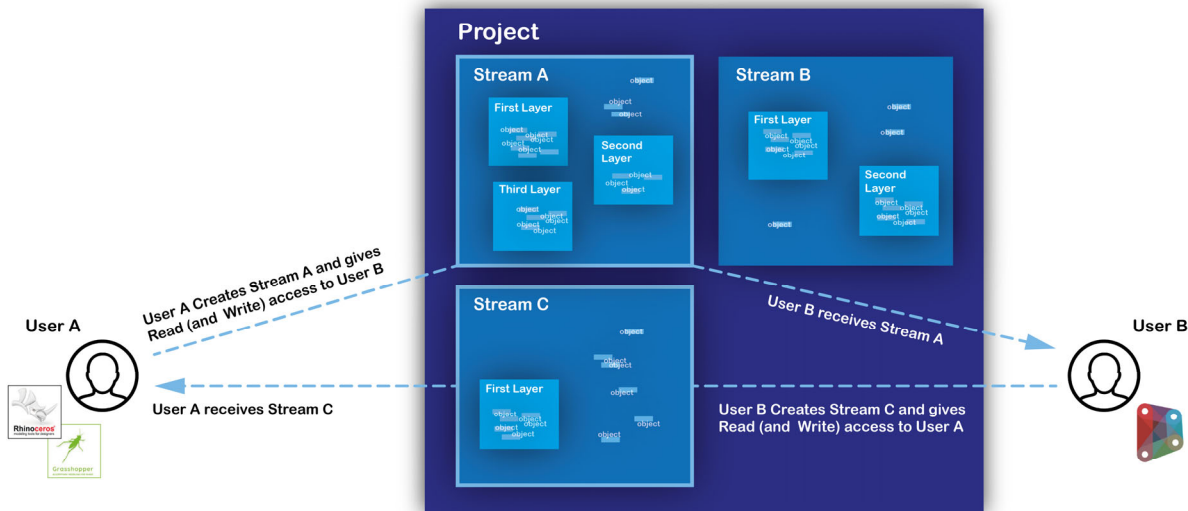
Figure 4. Data exchange protocol between two users.
Streams are created and stored within the Project panel, and contain the project's exchanged data.

provided that they exhibit certain technical characteristics (Speckle, 2020).

Data transfer to and from each end-user is orchestrated by a given Speckle Server, which ensures its availability in the case the original source is offline. Furthermore, the server allows also for efficient updates by leveraging several mechanisms, such as caching, object immutability and partial, differential updates. Speckle implements a discretionary access control model, which gives full control to data authors on how accessible their information is, and with whom. This allows for either fully public or private resources but as well for granular privacy and security settings customised to the roles and needs of each design actor a particular resource is shared with Speckle (2020).

Resources in Speckle are organized in a hierarchical manner as a Work Breakdown Structure (WBS) (see Figure 4) through Objects, Layers (collections of Objects), Streams (collections of Layers and/or Objects) and Projects (collections of Streams). Furthermore, Speckle allows for resources to be enriched with extra metadata such as description, tags, comments, so as to be able to respond to the project's needs and allow for diagonal queries.

Since its inception in 2016, it has been adopted by a large number of progressive AEC companies as a key piece in their digital transformation efforts.

4   SPECKLEVIZ

Via Speckle Streams, users are able to share data from the different existing Speckle clients and plug-ins, which expose a User Interface to both share data (Senders) and receive (Receivers). For example, User A creates a Sender to share Stream A to User B, who creates a Receiver to receive Stream A from User A. As data transfer protocols in Speckle operate in a unidirectional (as opposed to bidirectional) manner, User B would need to create a new Stream (after

working upon the data sent by User A via Stream A) to share new data to User A. This simple yet crucial triple protocol (Sender-Stream-Receiver) defines the basis of the Speckle's Activity Network and is illustrated in Figure 5, along with the aforementioned hierarchical directory structure of Speckle (Object, Layers, Streams, Projects). Although one Stream can be contained within multiple Projects, SpeckleViz only renders the activity network happening within a single Project. In other words, SpeckleViz is a tool for illustrating data flows among the project network, a *"network that gets re-initiated for each project"* (Chinowsky et al., 2008, p. 806, Chinowsky et al., 2010, p. 453).

As an architectural project involves a large number of actors working from different software platforms, the above described procedure scales up to form a larger non-linear workflow that is composed of multiple Sender-Stream-Receiver protocols. In general, Speckle Streams are
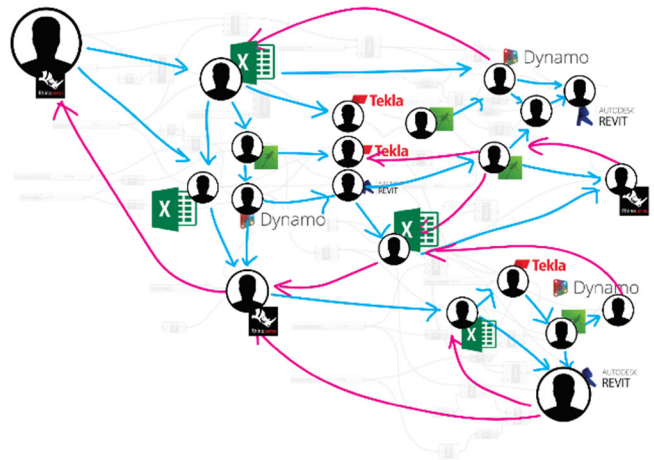
Figure 5. The overall workflow in Speckle is non-linear and contains multiple feedback processes between the different involved trades and software platforms.

416
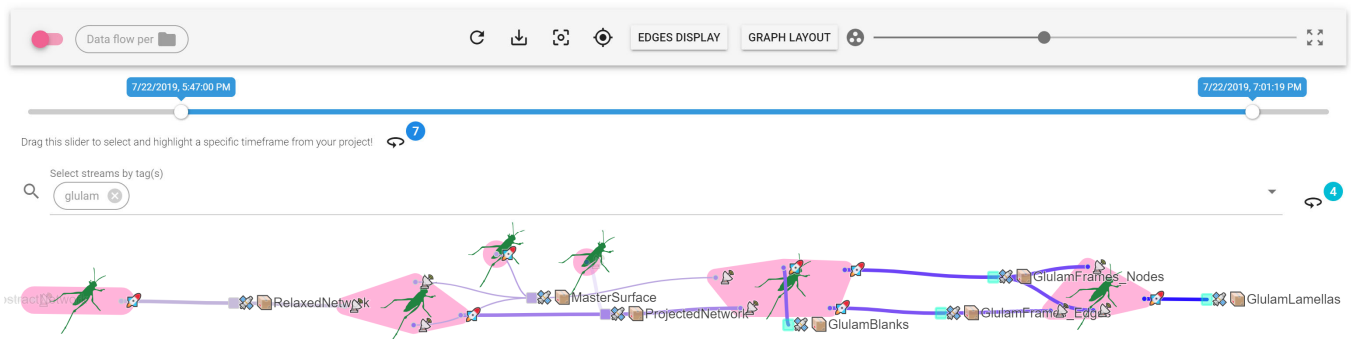
Figure 6. The current SpeckleViz interface (SpeckleViz, 2020).

considered to be single directional flows of data, as this reduces conflicts arising from various sources updating a stream with conflicting information. Although the data flow is deliberately acyclic (as data transfer mechanisms in Speckle are unidirectional) one could interpret the pattern UserA-StreamA-UserB-StreamB-UserA as a data cycle. In other words, whereas the underlying data transfer in Speckle is acyclic (and can therefore be defined as a Directed Acyclic Graph - DAG), its representation can be, in some instances, perceived as cyclical because it starts and ends with 'UserA'. Figure 5 illustrates how data transfers would evolve and look like throughout the design process timeline of a building project, amongst multiple disciplines and across different software platforms. The Speckle's Activity Network described in the next section attempts to visualize these processes by offering an interactive user interface from which the user can query the different created streams.

4.1 Technological Framework

Back-end

On the back-end, the SpeckleViz activity network diagram harvests data through the Application Programming Interface (API) calls using Axios (Axios, 2019). The initial HTTP request takes a Project ID as an input, returning the list of contained streams as a response. New HTTP requests are made to retrieve each stream's corresponding resources, such as: `_id` (Stream's ID), `owner` (Stream's owner), `createdAt` (Stream's creation time), and `updatedAt` (Stream's last update), etc. Finally, last HTTP requests are made to get the corresponding Clients (Sender and/or Receiver) resources per Stream, such as: `_id` (Client's ID), `owner` (Client's owner), `documentGuid` (Document's GUID), `documentName`, `createdAt` (Client's creation time) and `updatedAt` (Client's last update). Although most of the resources can inform the graph, the main ones used to create its nodes and edges are the Client's `_id` and Stream's `_id` properties. The collected resources are formated into a JSON (JavaScript Object Notation) `objects` which will further feed the graph on the front-end.

Front-end

SpeckleViz (Figure 6) is built upon the Speckle web management admin interface, within the Project tab. As the latter has been designed with Vue.js, an open-source JavaScript framework for building user interfaces and single-page applications (Macrae, 2018), a basic layout has been designed with the same framework in order to host the graph itself, which has been rendered using D3.js (also known as D3, short for Data-Driven Documents) - a JavaScript library for producing dynamic, interactive data visualizations in web browsers (M*urray, 2017).*

As Vue.js and D3.js operate on different levels and through different mechanisms, a suitable pattern had to be established to enable the passing of data seamlessly from one framework to the other. In this context, a Vue template has been created to receive the SVG elements from D3.js. For example, the SVG elements `<svg>` `<g>` and `<rect>` elements are added individually rather than through the familiar D3.js method chaining pattern. This allows to dynamically bind these elements to D3.js data within the Vue component, and take advantage of Vue's reactivity. In general, the graph is generated through d3-force, a D3.js module dedicated to force-directed graph layout using velocity verlet integration. In regards to the styling of the toolbar and control panels, SpeckleViz relies on Vuetify.js (Vuetify, 2020)

4.2 Visualization Features

While circle nodes represent Senders (S) and Receivers (R), square nodes represent Streams. Arrows (or graph edges) represent either data that has been shared to a stream by the user (Receiver to Stream) or data that has been retrieved by a user from a stream (Stream to Sender). The edge's thickness is proportional to the number of exchanged geometrical objects. Generally, both nodes and edges are coloured according to their respective timestamp: dark blue for the newest created, and light grey for the oldest.

As the graph is force-directed and rendered dynamically, its overall layout might sometimes become too convoluted and not tidy enough to be grasped as a whole. Therefore, several options have been exposed to the end-user in order to manually adapt the graph representation: while the display mode of the graph edges could be switched between three different modes (straight line, arc or diagonal), the force-directed graph layout could be altered in order to force its alignment along the X or Y axis, taking the shape of a tidier

Directed Acyclic Graph (DAG) (see Figure 7). Furthermore, the end-user could at any time during render stop the force-directed simulation.

Senders and Receivers can be grouped either by identical Document GUID or identical Client's owner ID. While the latter is represented by a blue convex hull, the former is visualized by a pink convex hull. The next section will elaborate on an interaction feature enabling the end-user to switch between two modes of representation: user-centred or document-centred.

4.3 Interaction Features
Multiple interaction features have been implemented on the front-end in order to give the end-user a more granular control over the data exposed by SpeckleViz. These features operate on four different levels:

Drop-down menus
Right-clicking on one the graph nodes would display their related drop-down menu. In the case of a Stream, the user can choose between accessing the Stream's information through the Speckle web management admin interface, viewing the Stream within the Speckle viewer interface or accessing the Stream's data available through the API itself. In the case of a Client, the user can retrieve basic information, such as its `_id`, `createdAt` and `updatedAt` properties.

Time frame selection
As specified above, both Streams and Clients expose the createdAt property informing on when the Stream or Client was created. This data has been brought to the front-end of SpeckleViz by integrating a slider within the application, enabling the end-user to select a specific time frame of the project. When dragging the slider, the graph's nodes and links fade out when they are out of range and fade back in when they are in range. Furthermore, the Streams created within the selected time frame are continuously collated and can be visualized altogether inside the Speckle viewer through a dedicated button.

Tab-based queries
In Speckle, Streams can be tagged by the end-user through the web management admin interface. Input tags are then exposed on the API side through the Stream's property `tags`, which is collected on the back-end before rendering the graph in SpeckleViz. On the front-end side, the Vuetify `v-autocomplete` component (Vuetify, 2020) has been integrated, enabling the end-user to select/deselect the existing tags present within the API. The selection dynamically updates the display of the Stream nodes within the graph by highlighting the ones containing at least one tag present within the current selection. Furthermore, the selected tagged Streams are continuously collated and can be visualized altogether inside the Speckle viewer through a dedicated button.

Adaptive representations
While the activity network in Speckle has been so far illustrated from a user-centred perspective (Figure 4 and 5), the related works referenced earlier highlighted workflows that were instead document-centered (Figure 2), or company-centered (Figure 1). Data flows within the AEC sector can therefore be interpreted (and visualized) from different perspectives. In such a context, SpeckleViz attempts to give the end-user the possibility to visualize and adapt its graph from different points of view: the SpeckleViz toolbar exposes a toggle button enabling the user to choose between the *"Data flow per user"* and *"Data flow per document"* modes. Therefore, SpeckleViz provides the users with illustrations of the data flows among the Social Network, as described by Wasserman and Faust (1994), and the inherent Activity Network, as defined by Bauke de Vries (1995). Switching between these two modes dynamically updates the graph that reorganizes its nodes according to the chosen data flow perspective.

Although the main visualization and interaction features have been described above, every single one has been described in more details on the main documentation page of SpeckleViz (SpeckleViz, 2020).

5 CASE STUDY: SPECKLE WORKSHOP AT SIMAUD 2018
As the Project management panel within the web management admin interface is relatively new, not many practices have used it to organize their exchanged Speckle Streams. Instead, Speckle has been used so far in a more informal manner without too much focus on the Project management interface. As a consequence, there exist today only a very few publicly available data sets on which SpeckleViz could be eventually deployed. Therefore, the authors have exploited an existing digital workflow deployed during a workshop focussing on the Speckle platform, conducted at Delft University of Technology on the 4th of June 2018 in the context of the Symposium on Simulation for Architecture and Urban Design (SimAUD) conference. The Streams created and used during this workshop have been reorganized through the Speckle Project panel, serving here as initial data sets to test and deploy the SpeckleViz graph.

5.1 Case study brief and set-up
After an initial introduction to the Speckle communication platform, a parametric modelling workflow of a free-form timber structure developed within the Grasshopper interface has been segregated into eight different computational *"pipelines"* distributed amongst the eight different participants: (1) Global Network Control, (2) Geometrical Optimization (angle maximization between the members), (3) Radius Control and Maximization, (4) Master Surface Control, (5) Blank Mesh Generation, (6) Volume Mesh Generation, (7) Lamella Mesh Generation, (8) Result Overview. Each participant had control over its own local design space. When satisfied with the local design outcome, the workshop participant could communicate the output of
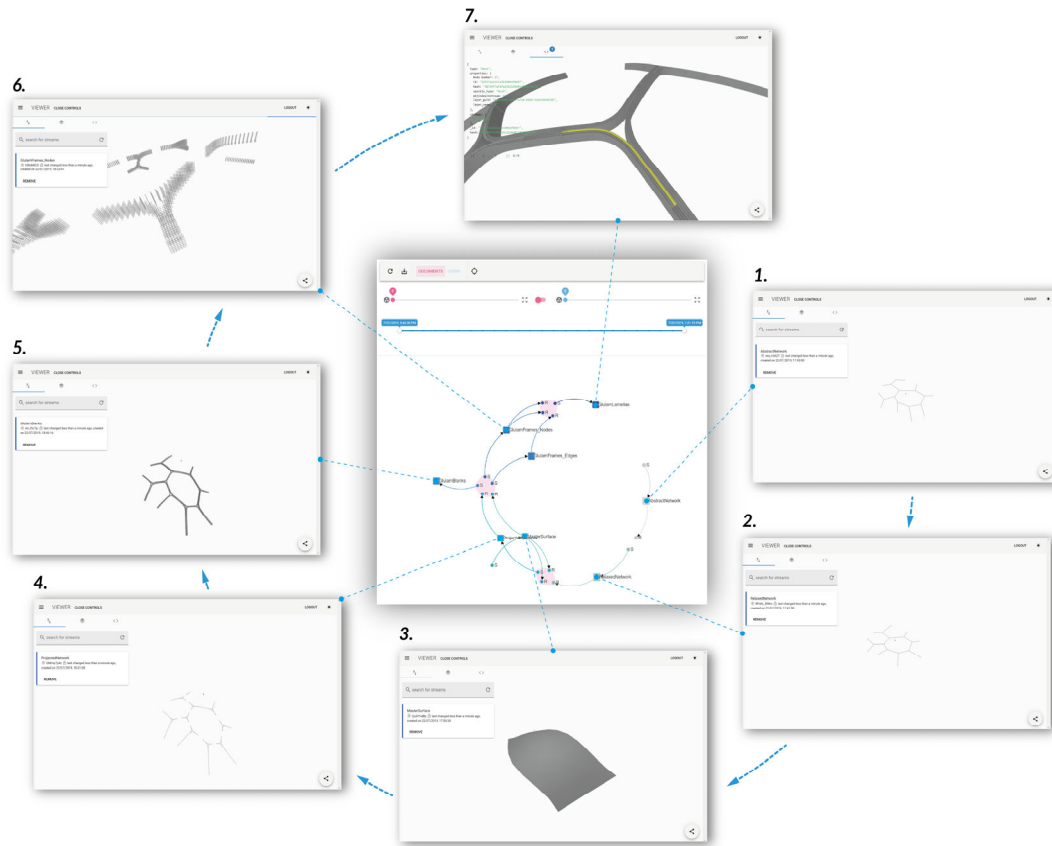
Figure 7. The deployed SpeckleViz interface representing the data flow of the Speckle Workshop at SimAUD 2018

its own computational pipeline to the next design actor, which also had control over its own design space, and so on.

5.2 Case study data collection and analysis

Through the Grasshopper Speckle Client, both Speckle Senders and Receivers were used to seamlessly share design data across all the phases of the design process. The data collected originated directly from the different Rhino-Grasshopper sessions manipulated by the students, and could serve three different purposes:

- **Sharing design ideas:** data could be shared in the sole purpose of exchanging design ideas. This way, students could always log in to the admin interface, explore design possibilities and be inspired by the different models shared by their classmates.

- **Monitoring a Stream's design history:** data could be collected in order to keep track of the design history (enabled by the Speckle platform) of a particular Stream.

- **Keeping track of the project's timeline history:** Finally, data could also be gathered in order to keep track of the chronological evolution of the design process, from the lowest level of detailing to the highest.

The present case study can be seen as an experimental, distributed design chain across all participants, forming the overall design workflow of the workshop. Speckle Senders and Receivers were used to seamlessly share design data

across the pipelines. The overall workflow has been represented through the SpeckleViz interface (Figure 7). The end-user (here, a speculative project manager) is able to visualize all the data exchanged during the workshop and access each Stream within the Speckle viewer environment. The interaction features described above were all operable, such as the tag-based query interface. For example, streams tagged as "fabrication" and/or "global design" could be called, collected and visualized within the Speckle viewer.

6 DISCUSSION AND CONCLUSION

This paper introduced the current challenges in collaborative design workflows within the AEC sector and described the SpeckleViz interface, an experimental and work-in-progress web-based interactive visualization tool which aims at representing the activity network (across users or documents) that operates within the open-source Speckle framework.

The SpeckleViz interface has demonstrated how existing open-source frameworks, such has Vue.js (Macrae, 2018), D3.js (Marray, 2017), Vuetify.js (Vuetify, 2020) and Speckle (Speckle, 2020) can be leveraged and orchestrated altogether in order to build a custom application that answers specific needs from the industry (the AEC sector in the present case). The developed application can then feed back to the open source community, and hopefully inspire future open source contributors to develop custom applications built on Speckle.

6.1 Limitations

Maintenance

Currently, Speckle provides five different application integrations (such as Grasshopper3D and Dynamo), and it is expected that this number increases over time as the Speckle community continuously grows and therefore might gain interest in integrating the open-source platform within other specific software packages. Speckle already requires high maintenance as each software package that integrates a Speckle client could modify its .NET API (A or SDK unexpectedly. Consequently, the contributors to the Speckle platform would need to revisit the affected open-source repositories and rewrite specific object model conversions.

Beyond Visualization

The current SpeckleViz interface can mainly be used to monitor a design process, and its interaction features are mostly limited to the inspection of one or multiple streams based on a time range or tag(s). The authors believe that these existing features could be extended beyond simple visualization or Stream inspection by enabling the end-users to directly manage a Speckle project through the graph itself, by adding, deleting, or tagging Streams. More than a visualization tool, SpeckleViz would then act as a data management platform.

6.2 Future work

Although it has not yet been widely adopted by architectural or engineering practices, future collaboration with different AEC companies (e.g. BuroHappold Engineering and Grimshaw Architects) will look precisely at how SpeckleViz could be deployed within and adapted for data exchanges within practice. Future work will also look at how SpeckleViz could represent the data flows beyond a single project environment. Visualizing the data exchange across multiple projects, companies and/or servers remain open questions that still need to be addressed. Enabling adaptive display strategies to focus at different levels of representation (collapsing the nodes belonging to the same document or same user into one) is another aspect that needs to be tackled. Finally, further research will investigate more in depth the data available within the Speckle API on the back-end to increase analytics and give richer insights to the end-user on the front-end.

REFERENCES

1. Autodesk Revit is a Building Information Modelling (BIM) software for architects, engineers, designers and contractors.

2. Axios, A promise based HTTP client for the browser and node.js. https://github.com/axios/axios. 2020.

3. Booch, G., Rumbaugh, J. & Jacobson, I. *The Unified Modeling Language User Guide.* (2nd ed.) Addison-Wesley Professional, 2005.

4. Chinowsky, P., Diekmann, J. & Galotti, V., Social network model of construction. *Journal of construction engineering and management*, 134 (2008), 804-812.

5. Chinowsky, P.S., Diekmann, J. & O'brien, J., Project organizations as social networks. *Journal of Construction Engineering and Management* (2010), 452-458.

6. Davies, A. *Project management for large, complex projects*. London, United Kingdom: The Bartlett School of Construction and Project Management (2019).

7. De Vries, B. Message Development in the Building Process. *Modeling of Buildings through their Life-Cycle. Proceedings of the CIB w78 Conference.* Standford (1995) 467-479.

8. Dynamo is a graphical programming interface within Revit.

9. Flux, 2018. Retrieved from http://flux.io/ (accessed before the 31st of March 2018).

10. Grasshopper3D (typically abbreviated Grasshopper) is a visual programming language and environment that runs within Rhino3D.

11. Macrae, C. *Vue.js – Up and Running*. O'Reilly, 2018.

12. Murray, S. *Interactive Data Visualization for the Web - an Introduction to Designing with D3*. (2nd ed.) Sebastopol: O'Reilly, 2017.

13. Ringley, B. San Francisco Computational Design User Group, June 2017: *What is the point of using Dynamo?* (https://www.youtube.com/watch?v=y6N1ICoFoyU)

14. Rhinoceros (typically abbreviated Rhino, or Rhino3D) is a commercial 3D computer graphics and computer-aided design (CAD) application software developed by Robert McNeel & Associates. Available from https://www.rhino3d.com/ (accessed on the 17th of March 2020).

15. Speckle, 2020. Retrieved from http://speckle.systems (accessed on the 17th of March 2020).

16. SpeckleViz, 2020. Retrieved from https://speckle.systems/docs/web/speckleviz/ (accessed on the 17th of March 2020).

17. Van Der Heijden, R., E. Levelle, and M. Reise. Parametric Building Information Generation for Design and Construction. In *Proceedings of the 35th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA 2015)*. 417–430.

18. Vuetify, A Material Component Framework for Vue.js. https://github.com/vuetifyjs/vuetify. 2020.

19. Wasserman, S. and Faust, K. *Social network analysis: Methods and applications.* Cambridge, United Kingdom: Cambridge University Press, 1995.

20. White, S. A. *Introduction to BPMN.* BPTrends, July 2004.

21. Ziegler, B. P., Praehofer, H. & Kim, T. G. *Theory of Modeling and Simulation*. (2nd ed.) Academic Press, 2000.

Graph Machine Learning using 3D Topological Models

Wassim Jabi[1], Abdulrahman Alymani[1]

[1]Cardiff University
Cardiff, United Kingdom
{jabiw, alymaniaa}@cardiff.ac.uk

ABSTRACT

Classification of urban and architectural works using machine learning techniques have typically focused on 2D pixel-based image recognition. In this paper we present a novel proof-of-concept workflow that enables a machine learning computer system to learn to classify 3D conceptual models based on topological graphs rather than 2D images. The system leverages two main technologies. The first is a custom designed software library that enhances the representation of 3D models through non-manifold topology and embedded semantic information. The second is an end-to-end deep graph convolutional network (DGCNN) that accepts graphs of an arbitrary structure without the need to first convert them into vectors. The experimental workflow consists of two stages. In the first stage, a generative parametric system was designed to create a large synthetic dataset of an urban block with several typological variations. The geometric models were then automatically labelled and converted into semantically rich topological dual graphs. In the second stage, the dual graphs were imported into the DGCNN for graph classification. Experiments demonstrate that the proposed workflow achieves accuracy results that are highly competitive with DGCNN's classification of benchmark graph datasets.

Author Keywords

Machine Learning, Topological Models, Graphs.

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING

1 INTRODUCTION

Understanding and classifying the typology of urban and architectural forms has been a topic of interest for a long period [19]. By understanding the typology of an urban form, relevant performative information can be extrapolated that helps urban designers and planners to make better decisions. Quantitative and statistical methods aided by computational tools have been shown to be effective in morphological studies [7]. While the impact of these approaches has thus far been limited in the mainstream practice of urban design and planning, there are signs that machine learning technologies (ML) are starting to revolutionize the recognition and classification of urban form. Several challenges still face the adoption of these technologies. First, supervised machine learning requires large datasets for training that are labelled. Second, most machine learning systems rely on 2D pixel-

based image recognition. While this may seem compatible with the representation of the available data – mainly city plans and drawings, this leads to major limitations. The urban environment is three-dimensional, topologically connected, and complex. However, most machine learning systems do not understand the semantics of the recognized image. To an ML algorithm, the data is simply a vector of mostly low resolution RGB values. If the data is 3D in nature and semantically rich, that information must be stripped away, and the data converted into images for an ML algorithm to operate properly. The lack of shareable 3D datasets is a real challenge and while some open-source sets do exist [9], their format, suitability, accessibility and licensing varies.

Even if 3D datasets are made available, there continues to be a challenge in recognizing and classifying them. Some researchers have focused on feature recognition of 3D models [12, 18]. These approaches capture several 2D snapshots of 3D models and match them to an image-based query, thus, still not fully capturing the three-dimensional and topological information embedded in the data. A slightly more sophisticated approach is to extract features from a 3D model and encode them as a vector to be used as input into a neural network [17]. This approach extracts only a portion of the data and must transform it into a standard input vector. It ignores the overall topological information that can be an indicator of the type of object.

A promising approach is to use machine learning on graphs [14, 16, 22]. The limitation of some approaches is that they decompose the graphs into small substructures such as walks or paths and derive the similarity between graphs based on a summation of attributes. The Deep Graph Convolutional Neural Network (DGCNN) avoids the above limitations and provides an end-to-end deep learning system for classifying graph-based data [26]. One of the main advantages of DGCNN is that it accepts arbitrary graphs without the need to first convert them into vectors.

Our aim in this paper is to design a novel proof-of-concept workflow that leverages DGCNN to classify 3D conceptual models based on their topological graphs rather than on their 2D representation. As a domain of investigation, we chose an urban block morphology. The workflow is divided into two stages. In the first stage, we use a software library, developed by one of the authors, that enhances the

representation of 3D models through non-manifold topology and embedded semantic information. The software is used to automatically and generatively create a large synthetic dataset of an urban block with several typological categories. The geometric models are then automatically labelled and converted into semantically rich topological graphs. In the second stage, we use this dataset as input to train DGCNN for graph classification.

The remainder of this paper is organised as follows: Section 2 briefly reviews related work carried out using graphs and machine learning in architecture. Section 3 provides a brief summary of Graphs theory. Section 4 describes the Topologic software library. Section 5 discusses related variations of Graph Convolutional Neural Networks (GCNN). Section 6 describes the Deep Graph Convolutional Neural Network (DGCNN) that was used for this research project. Section 7 details the experimental case study. Section 8 reports on the results of the experiments. Section 9 lists the limitations of this work. Finally, Section 10 provides concluding remarks and lists our plans for future work.

2   RELATED WORK

Beetz has researched the use of graph databases for harmonized distributed concept libraries for building information models [2]. His goal is to create "flexible, granular and cascading concept libraries for the building industry." This work will enable standardization of the input data to graph neural networks. Tamke has researched the use of unsupervised and supervised machine learning approaches to deduce implicit information in building information models [20]. His platform can extract literal values, aggregates and derived values from IFC SPF files. Furthermore, the system possesses geometrical and topological analysis functionality that allows it to detect anomalies and to classify floor plans into two categories based on their geometrical appearance. Derix and Jagganath have researched methods to autonomously recognise spatial typologies through associations between spatial attributes of a layout [5]. They presented cases in which a building floor layout was analysed using space syntax concepts such as isovists, centrality and visual connectivity [10]. Their aim is to use these models to classify types and sequences of user experiences across buildings rather than strict typologies thus enabling an experience-based approach to architectural and urban design. Harding and Derix used a two-stage neural network and spectral graph theory as a spatial pattern recognition tool to develop a plan form of a reconfigurable exhibition space. To recognise and classify plans, they derived a graph using spatial adjacency and then reduced it by finding a "graph spectrum" which then forms a synaptic vector representation in feature space and thus "makes comparisons between graphs much easier to conduct." What is interesting in their research is that they also use this approach for the automatic generation of spatial layouts. In their paper they describe a repulsion algorithm combined with a Voronoi diagram that distributes the graphs evenly

over the boundary plan. Topological connections are simulated as springs to maintain node adjacency.

3   GRAPHS

Explaining graph theory is beyond the scope of this paper, so a brief summary of its main concepts, data structures and metrics is provided here. For more detailed information on graph theory, please consult [23]. Graph theory is a branch of mathematics used to model relations between objects. A simple graph G consists of a set of points called vertices $V(G)$, and the lines that join pairs of points are called edges $E(G)$. The *degree* of a vertex in a graph is the number of edges connected to it. Vertices that are connected by an edge are called *adjacent vertices*. Similarly, edges that share a common vertex are called *adjacent edges*. Any two graphs that have a one-to-one correspondence between the number vertices, the number of edges and the degree of vertices are called *isomorphic* graphs.

4   TOPOLOGIC

Topologic [21] is a 3D modelling software library developed by one of the authors that enhances the representation of space in 3D parametric and generative modelling environments such as Dynamo [6] and Grasshopper [8]. Topologic is based on the concept of non-manifold topology and has been described in [1, 11]. Topologic's classes include: Vertex, Edge, Wire, Face, Shell, Cell, CellComplex, Cluster, Topology, Graph, Aperture, Content, and Context (see Figure 1). A Vertex is a point in 3D space with X, Y, Z coordinates. An Edge connects a start Vertex to an end Vertex. A Wire connects several Edges. A Face is made of a set of closed Wires. A Shell is a set of connected Faces that share Edges. A Cell is made from a closed Shell. A CellComplex is a set of connected Cells that share Faces. A Cluster is a grouping of topologies of any dimensionality. A Graph is a special data structure that is derived from Topologies. An Aperture is a special type of Face that is hosted by another Face. Any Topology can have additional Topologies added to its Contents. In turn, these Content Topologies will have a pointer back to their Context Topologies. This is similar to a parent/children relationship. In addition, any Topology can have a Dictionary that can hold any number of arbitrary key-attribute pairs.
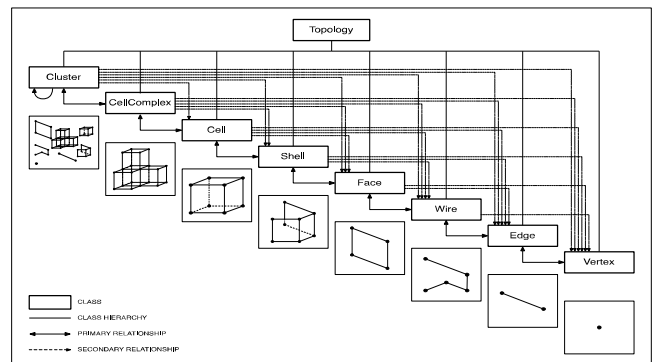

Figure 1. Topologic Core Class Hierarchy.

In this paper, we will focus on two features of Topologic that were essential for the proposed workflow: 1) the automatic derivation of 3D topological dual graphs using the Cell, CellComplex, and Graph classes, and 2) the embedding of semantic information through custom dictionaries.

As mentioned above, In Topologic, a CellComplex is made of enclosed 3D spatial units (Cells) that share Faces. Cells that share Faces are called adjacent Cells (see Figure 2).

The Graph class and associated methods are based on graph theory. A Graph is composed of Vertices and Edges that connect Vertices to each other. A Graph in Topologic accepts as input any Topology with additional optional parameters and outputs a Graph. In its simplest form, the dual graph of a CellComplex is a Graph that connects the centroids of adjacent Cells with a straight Edge (see Figure 3).
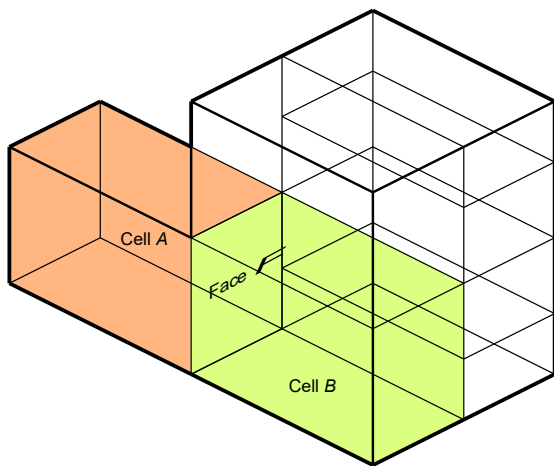
Figure 2. An example CellComplex. Cell *A* and Cell *B* are said to be adjacent because they share Face *F*.
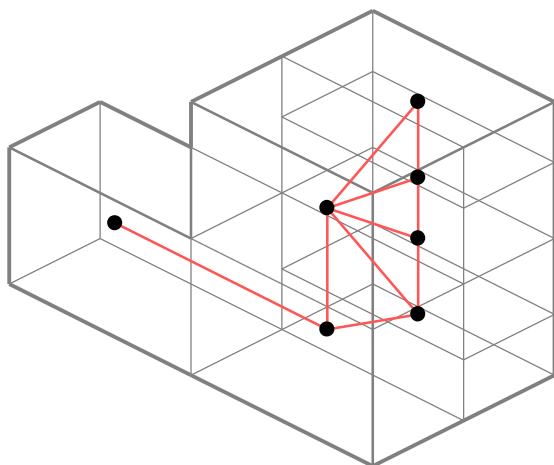
Figure 3. An example Dual Graph of the CellComplex. Each Cell is represented by a Vertex and the Vertices of adjacent Cells are connected by an Edge.

A dictionary is a data structure made of key/value pairs. A key is any identifying string (e.g. "ID", "Type", "Name"). The value of a key can be of any data type (e.g.float, an integer, a string). Topologic enables the embedding of arbitrary dictionaries in any topology. As topologies undergo geometric operations (e.g. slicing a Cell into several smaller Cells thus creating a CellComplex), the dictionaries of operand topologies are transmitted to resulting topologies. Furthermore, when a dual graph is created from a Topology, the dictionaries of the constituent topologies get transferred to their corresponding vertices. We use this capability to label the vertices in the dual graph.

5 GRAPH CONVOLUTIONAL NETWORKS

Graph convolutional neural networks were first introduced by Bruna *et al* [3] where they demonstrated that it is possible to learn convolutional layers with a number of parameters independent of the input size, resulting in efficient deep architectures. In 2018, Xie and Grossman proposed a crystal graph convolutional neural network (CGCNN) that learns the properties of crystal atoms [24]. CGCNN offered highly accurate predictions of eight different properties of crystals. Chai *et al* proposed in 2018 a multi-graph convolutional neural network that can predict the bike flow at station-level in a bike sharing system [4]. Their model can outperform state-of-the-art prediction models by reducing the prediction error by up to 25.1%. Li *et al* proposed a Diffusion Convolutional Recurrent Neural Network (DCRNN), a deep learning framework for traffic forecasting that incorporates both spatial and temporal dependency in the flow of traffic [15]. Yu *et al* proposed graph convolutional networks to predict traffic speed in road systems that consistently outperforms state-of-the-art baselines on various real-world traffic datasets [25]. In 2019, Kipf and Welling introduced a scalable approach for learning on graph-structured data [13]. Their model scales linearly and can encode both local graph structures and features of nodes. They showed that their approach outperforms related methods by a significant margin on datasets in the domain of citation networks.

6 DEEP GRAPH CONVOLUTIONAL NETWORKS

In 2018, Zhang *et al* introduced an end-to-end deep-graph convolutional neural network (DGCNN) that accepts arbitrary graphs without the need to first convert them into tensors [26]. DGCNN accomplishes this by first passing the inputted graph through multiple graph convolution layers where node information is propagated between neighbours. Then a second layer sorts the graph vertices in a consistent order which are then inputted into a traditional convolutional neural network (see Figure 4). By sorting the vertex features rather than summing them up, DGCNN keeps far more information thus allowing it to learn from global graph topology. Furthermore, Zhang *et al* provide a theoretical proof that in DGCNN, if two graphs are isomorphic (i.e. have an identical structure), their graph representation after sorting the vertices is the same. This avoids the need to run additional costly algorithms to canonize the graph. Compared to state-of-the-art graph kernels, DGCNN has

achieved highly competitive accuracy results on benchmark graph classification datasets.
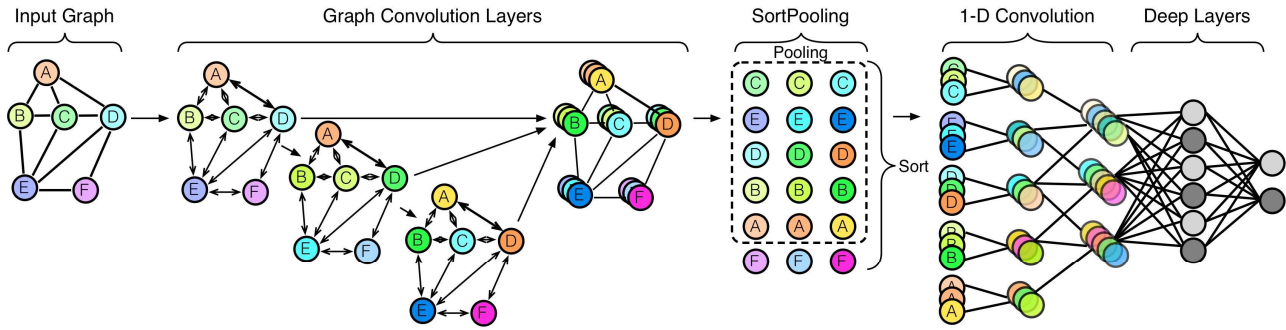
Figure 4. The general structure of DGCNN (after Zhang *et al* [26]).

7   EXPERIMENTAL CASE STUDY

For the experimental case study, we created a workflow using Dynamo and Topologic that was designed to generate many 3D parametric models, and their associated topological dual graphs, of an urban block with a ground plate, an optional base that sits directly on top of the ground plate, and one or more tower blocks that sit either directly on the ground plate or directly on top of the base (see Figure 8).
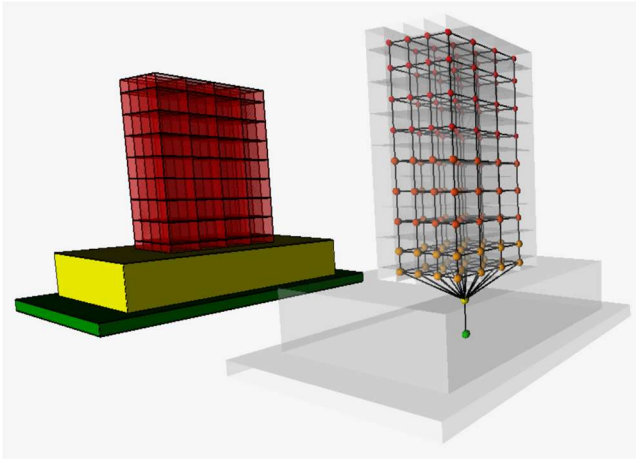
Figure 5. The 3D model and dual graph in Autodesk Dynamo.

The ground plate can vary in size. The base is then dimensioned to be a certain percentage of the ground plate with equal offsets. One or several tower geometries are then placed with appropriate offsets and spacing. The height of the tower blocks is varied, but all tower blocks maintain the same height. Finally, the tower geometries are subdivided internally into a grid of Cells. This creates a 3D lattice structure. It is important to note that the internal sub-division of the tower block is to aid the neural network in distinguishing structures of different heights rather than to provide room-level detail.

To create the dataset, we needed to accomplish three tasks. The first task is to label the overall graph. We chose to limit

the classification of the graph to four categories: 1) Tower, 2) Slab, 3) Tower on Base, and 4) Slab on Base. The first classification (Tower) occurs when the base is not created, and the height of the tower is larger than its width and length. Similarly, a Slab classification occurs when the base is not created, but the height of the tower is less than or equal to its width and height. The last two categories repeat the rules of the first two but occur when a base is introduced.

The second task was to label the vertices. In addition to labelling the overall graph, DGCNN requires that each vertex in the graph is labelled. To experiment with the effect of labelling vertices, we created two separate datasets that contain the same graph topologies but with a different vertex labelling scheme. In the first dataset, called 3 labelled, we labelled vertices according to three categories: 1) Ground, 2) Base, and 3) Tower Cell. For the second dataset, called 5 labelled, we subdivided the tower cells into three categories based on their Z height value: 1) Low, 2) Medium, and 3) High. This yields five total labels: 1) Ground, 2) Base, 3) Low Tower Cell, 4) Medium Tower Cell, 5) High Tower Cell (see Figure 6).
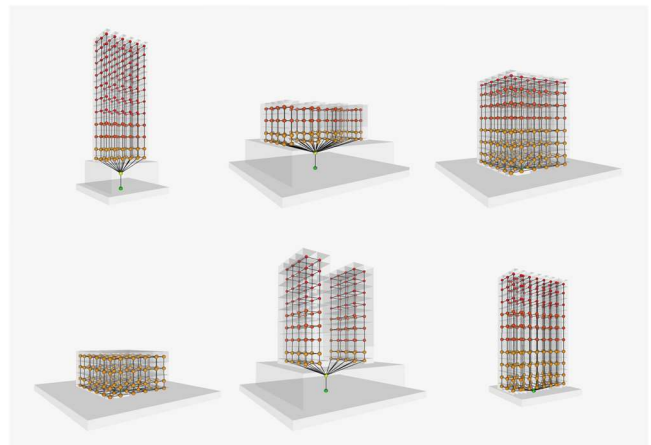
Figure 6. Examples of auto-generated urban block configurations with associated dual graph.

424

The last task was to integrate the visual dataflow definition with a custom python script to convert the 3D dual graph created by Topologic into a text file according to the format required by DGCNN (see Figure 7). The first line of the text file contains the total number of graphs (*g*). This is followed by *g* blocks of graphs where each block starts with a line that contains the number of vertices (*n*) followed by a number that indicates the classification (*p*) of that graph. This is then followed by a block of *n* vertices where each line starts with the label of the vertex (*vl*) followed by the indices of its adjacent vertices. The index of a vertex is implied by its line position using a zero-based numbering system.

```
g
n p
vl1 a b c d …
vl1 a e f g h …
vl2 m n o …
vl2 x y a …
 …
```


Figure 7. The general dataset format required for DGCNN.

Since the dataset in synthetically produced through an iterative *for loop*, the resulting list is implicitly cyclical in complexity and height (see Figure 8). To avoid biased training or the testing of only a specific level of complexity, the final list of graphs is reordered randomly.

The produced dataset totalled 1000 graphs as follows:

- 411 tower graphs
- 80 slab graphs
- 420 tower-on-base graphs
- 89 slab-on-base graphs
- The total number of vertices is 292,570
- The average number of vertices per graph is 292.57
- The minimum number of vertices in a graph is 10
- The maximum number of vertices in a graph is 1585
- Tower graphs have an average of 399 vertices, minimum of 37 vertices, and a maximum of 1,585 vertices
- Slab graphs have an average of 134 vertices, a minimum of 19 vertices, and a maximum of 361 vertices
- Tower-on-base graphs have an average of 259 vertices, a minimum of 14 vertices, and a maximum of 706 vertices
- Slab-on-base graphs have an average of 101 vertices, a minimum of 10 vertices, and a maximum of 433 vertices

The reason towers-on-base and slabs-on-base have fewer vertices than those without a base is because they have a smaller footprint when placed on a base.

All experiments were implemented on an ordinary laptop computer running macOS Catalina 10.15 operating system with the following configuration: An Intel Core i7 Quad-Core CPU running at 2.7GHz with 16 GB of memory. DGCNN was deployed using the *pytorch* python environment.
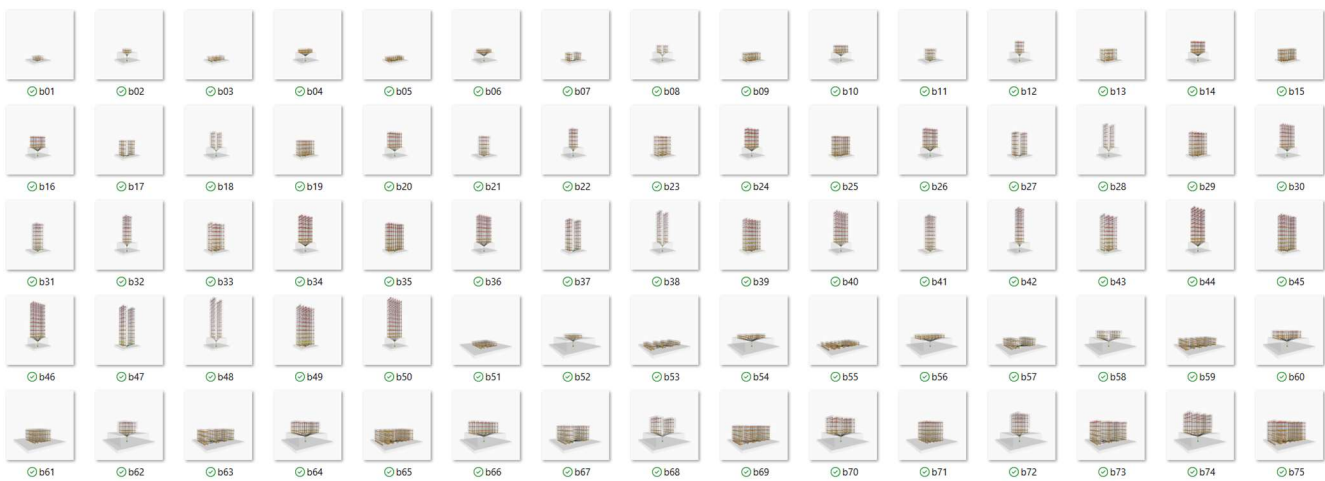
Figure 8. Sample of automatically generated urban block typologies.

For our experiments, we maintained most of the default values used by DGCNN. For details on these, please consult [26]. Below is a brief summary of these unmodified parameters:

- Size of graphlets: 3.
- Decay parameter: the largest power of 10 that is smaller than the reciprocal of the squared maximum node degree.
- SortingPool k: Set such that 60% of the graphs have nodes more than k.
- Two 1-D convolutional layers. The first layer has 16 output channels with a filter size of 2 and a step size of 2. The second 1-D convolutional layer has 32 output channels with a filter size of 5 and a step size of 1.
- The dense layer has 128 hidden units followed by a softmax output layer.
- A dropout layer is added at the end with a dropout rate of 0.5.
- DGCNN uses a nonlinear hyperbolic function (*tanh*) in the graph convolution layers and a rectified linear units (ReLU) in the other layers. DGCNN does not use validation set labels for the training.

8 EXPERIMENTAL RESULTS

As detailed in the experimental results below, we varied only the following hyper-parameters: the training/testing ratio, the learning rate, the number of epochs, and the batch size.

8.1 Training and testing division ratio

The 1000 graphs in the two datasets were divided into training and testing data. We experimented with using 20%, 30% and 40% of the total graphs as testing data and then documented the resulting graph classification prediction accuracy (see Table 1). The 5 labelled dataset achieved consistently higher accuracy than the 3 labelled dataset. The highest accuracy (78.66%) was achieved for the 5 labelled dataset with a 30% testing ratio. Therefore, in subsequent experiments, we used this testing ratio to continue improving the accuracy of the results.

Training data	Testing data	% of testing	3 Labelled	5 Labelled
800 graphs	200 graphs	20%	65.50%	69.00%
700 graphs	300 graphs	30%	64.45%	78.66%
600 graphs	400 graphs	40%	71.17%	72.00%

Table 1. Accuracy results using various training and testing ratios.

8.2 Learning rate

In convolutional neural networks, the learning rate is the amount by which the weights of nodes are updated during training. Varying the learning rate can dramatically affect the accuracy of the results. We experimented with four learning rates (1e-5, 1e-4, 1e-3, and 1e-2) and documented the results (see Table 2). A steep drop in the classification result was reported when using 1e-2 for both 3 labelled and 5 labelled datasets which was less than the acceptable rate. The highest

prediction accuracy result in this stage (79.66%) was achieved through a learning rate of 1e-4 on the 5 labelled dataset using a 30% testing ratio. This improved on the results achieved in the first round of testing. Therefore, in subsequent experiments, we used this testing ratio and learning rate to continue improving the accuracy of the results.

Learning Rate	% of Testing	3 Labelled	5 Labelled
1e-5	30%	64.45%	78.66%
1e-4	30%	67.00%	79.66%
1e-3	30%	64.66%	67.00%
1e-2	30%	31.66%	34.66%

Table 2. Accuracy results using various learning rates.

8.3 Number of epochs

The number of epochs is the number of complete iterations through the training dataset. We experimented with various numbers of epochs while maintaining a testing rate of 1e-4 and a testing ratio of 30% (see Table 3). The best classification accuracy result was reported using 800 epochs for the 5 labelled dataset (84.33%). Exceeding that value (e.g. 1000 epochs) for the 5 labelled dataset resulted in a significant decrease in the classification accuracy which indicates that the model became over-fitted. The variation of the number of epochs for the 3 labelled dataset reported insignificant variation and was consistently worse than the accuracy reported for the 5 labelled dataset.

Number of epochs	3 Labelled	5 Labelled
500	67.00%	79.66%
800	67.33%	84.33%
1000	67.66%	78.66%

Table 3. Comparison of accuracy results using various epochs.

The line graph below charts the graphs' classification accuracy for an average of 20 epochs (see Figure 9). The chart illustrates an elevated result for the 5 labelled dataset, with a moderate classification result for the 3 labelled dataset.
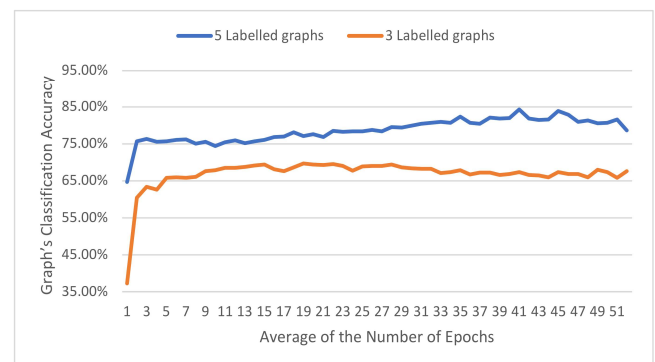

Figure 9. Learning rate and number of epochs for 3 labelled and 5 labelled datasets.

The 3 labelled model started learning and improved dramatically from the first few epochs. However, the accuracy level tapered after reaching approximately 220 epochs which means that the model stopped learning after that point. The 5 labelled model started learning well from the initial few epochs which began at approximately 65%. It then continued improving to reach a peak of 84.33%.

8.4 Batch size

The batch size of gradient descent in convolutional neural networks controls the number of training samples to iterate through before the model's internal parameters are updated. For this last experiment, we used the 1000 5 labelled dataset and maintained 30% testing ratio, 1e-4 learning rate and 800 epochs. We experimented with a diverse range of batch sizes (1, 10, 20, 30, 50 and 100). The best classification accuracy was reported for a batch size of 1 (see Table 4). However, a significant reduction in processing time (approximately 76%) can be achieved by increasing the batch size from 1 to 20 with a relatively modest loss in accuracy (approximately 5.5%). The associated line graph indicates the effect of the batch size on the classification accuracy (see Figure 10).

Batch size	Accuracy	Total Processing Time
1	84.33%	02:35:57
10	79.33%	00:36:30
20	79.66%	00:31:09
30	77.00%	00:31:22
50	76.33%	00:41:37
100	75.33%	01:09:50

Table 4. Comparison of accuracy results and total processing time using various batch sizes.
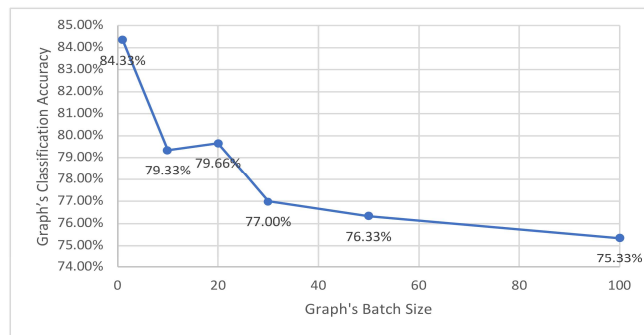
Figure 10. The effect of increasing the batch size on the classification accuracy

9   LIMITATIONS

Due to the lack of 'real' datasets, we resorted to generating synthetic datasets based on parametric variation. Even if we were to obtain real datasets, they may need intervention and translation work to make them amenable for dual graph extraction. Furthermore, this paper focused on the domain of an urban block tower. Its applicability to other typologies remains to be tested in future work. We compared graphs with 3 labels and to ones with 5 labels and found the latter to

be more effective for machine learning. However, we do not know the effect of further increasing the number of labels, inventing a different labelling scheme, or using more complex topologies that represent further internal space division. By using a computer with moderate power, we were limited in the amount of training and testing that we could conduct. Finally, our workflow was tested and fine-tuned independently and was not compared to other approaches which may or may not be more effective.

10   CONCLUSION

In this paper we aimed to find out if we can classify urban form through a novel workflow that uses machine learning on three-dimensional graphs rather than on two-dimensional images. We leveraged a sophisticated topology-based 3D modelling environment to derive dual graphs from 3D models and label them automatically. We then fed those graphs to a state-of-the-art deep learning graph convolutional neural network. To discover the best accuracy rates, we experimented with the vertex labelling scheme, testing ratios, learning rates, number of epochs and number of batches.

At the conclusion of our experiments we found that the 5 labelled dataset of 1000 graphs with a testing ratio of 30%, a learning rate of 1e-4, using 800 epochs, and a batch size of 1 gave us the best prediction accuracy (**84.33%**). This result is highly competitive with accuracy results on benchmark datasets. Our approach shows strong promise for recognising urban and architectural forms using more semantically relevant and structured data. Planned future work will experiment with other datasets, labelling schemes, and will compare this novel workflow to other approaches.

We have identified several new areas of research based on our findings. First, we are planning to investigate node classification rather than just overall graph classification. Second, we are planning a system that may recognise the topological relationships the designer is building in near real-time and suggest precedents from a visual database. Other future planned work includes the use of this technique as a fitness function within an evolutionary algorithm to generate and evaluate urban block forms that fit within a context based on user's preferences.

REFERENCES

1. Aish, R. et al. 2018. Topologic : Tools to explore architectural topology. *Advances in Architectural Geometry 2018* (2018), 316–341.

2. Beetz, J. 2014. A Scalable Network of Concept Libraries Using Distributed Graph Databases. *Computing in Civil and Building Engineering*. (2014), 955–1865.

3. Bruna, J. et al. 2014. Spectral networks and deep locally connected networks on graphs. *2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings* (2014).

4. Chai, D. et al. 2018. Bike flow prediction with multi-graph convolutional networks. *GIS: Proceedings of the ACM International Symposium on Advances in Geographic Information Systems*. (2018), 397–400. DOI:https://doi.org/10.1145/3274895.3274896.

5. Derix, C. and Jagannath, P. 2014. Digital intuition – Autonomous classifiers for spatial analysis and empirical design. *The Journal of Space Syntax*. 5, 2 (2014), 189–215.

6. Dynamo: *http://dynamobim.org*. Accessed: 2020-03-11.

7. Gil, J. et al. 2012. On the discovery of urban typologies: Data mining the many dimensions of urban form. *Urban Morphology*. 16, 1 (2012), 27–40.

8. Grasshopper 3D: *http://grasshopper3d.com*. Accessed: 2020-03-11.

9. Gröger, G. and Plümer, L. 2012. CityGML - Interoperable semantic 3D city models. *ISPRS Journal of Photogrammetry and Remote Sensing*.

10. Hillier, B. and Hanson, J. 1984. *The Social Logic of Space*. Cambridge University Press.

11. Jabi, W. et al. 2018. Topologic: A toolkit for spatial and topological modelling. *Computing for a Better Tomorrow, Proceedings of the 36th eCAADe conference* (Lodz, Poland, 2018).

12. Kasaei, H. 2019. OrthographicNet: A Deep Learning Approach for 3D Object Recognition in Open-Ended Domains. (2019).

13. Kipf, T.N. and Welling, M. 2019. Semi-supervised classification with graph convolutional networks. *5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings*. (2019), 1–14.

14. Kriege, N. and Mutzel, P. 2012. Subgraph matching kernels for attributed graphs. *Proceedings of the 29th International Conference on Machine Learning, ICML 2012*. 2, (2012), 1015–1022.

15. Li, Y. et al. 2018. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings*. (2018), 1–16.

16. Orsini, F. et al. 2015. Graph invariant kernels. *IJCAI International Joint Conference on Artificial Intelligence*. 2015-Janua, Ijcai (2015), 3756–3762.

17. Qin, F. et al. 2014. A deep learning approach to the classification of. 15, 2 (2014), 91–106. DOI:https://doi.org/10.1631/jzus.C1300185.

18. Sarkar, K. et al. 2017. Trained 3d models for CNN based object recognition. *VISIGRAPP 2017 - Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. 5, (2017), 130–137. DOI:https://doi.org/10.5220/0006272901300137.

19. Steadman, P. et al. 2000. A classification of built forms. *Environment and Planning B: Planning and Design*. (2000). DOI:https://doi.org/10.1068/bst7.

20. Tamke, M. 2015. Assessing implicit knowledge in BIM models with machine learning. *Modelling Behaviour*. December (2015). DOI:https://doi.org/10.1007/978-3-319-24208-8.

21. Topologic: 2019. *https://topologic.app*. Accessed: 2020-03-14.

22. Vishwanathan, S.V.N. et al. 2010. Graph kernels. *Journal of Machine Learning Research*. 11, (2010), 1201–1242. DOI:https://doi.org/10.1007/978-0-387-30164-8_349.

23. Voloshin, V.I. 2009. *Introduction to graph theory*.

24. Xie, T. and Grossman, J.C. 2018. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Physical Review Letters*. 120, 14 (2018). DOI:https://doi.org/10.1103/PhysRevLett.120.145301.

25. Yu, B. et al. 2018. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *IJCAI International Joint Conference on Artificial Intelligence*. 2018-July, (2018), 3634–3640. DOI:https://doi.org/10.24963/ijcai.2018/505.

26. Zhang, M. et al. 2018. An end-to-end deep learning architecture for graph classification. *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*. (2018), 4438–4445.

How Can Bio-Mapping the Foraging Excursions of Ants Translate Into a Prototype for Human Living on Mars?

Farah El-Hakim[1], Wassim Jabi[2]

[1]Cardiff University
Cardiff, UK
El-HakimF@cardiff.ac.uk

[2]Cardiff University
Cardiff, UK
JabiW@cardiff.ac.uk

ABSTRACT

This research project converges myrmecology, computer vision and architecture together to bio-map the locomotion of ants in a spatial-temporal dynamic setup. What can we learn from the de-centralised, flexible and optimal way of living from ants? Despite no journey plan and no limit on the size of areas they explore, ants never find themselves in traffic jams [1]. A number of simple rules give rise to a much more complex system, allowing ants to overcome many environmental obstacles with minimum energy expenditure and assessed risk. Through a series of stigmergic experiments, whereby a Kinect V2 is used for tracking and recording movement, this study looks to translate a spatial template of algorithmic foraging excursions into a resource-poor prototype for human living in an Extra-Terrestrial (ET) environment.

1 INTRODUCTION

When I think about the many dystopian futures that paint negative visions of human society, a repetitive occurrence lies within them. They tend to draw on the selfish behaviour of humans. This is particularly evident in J.G. Ballard's 'High-Rise' fictional book. Set in a residential block that offers all services necessary for living, a lack of empathy and altruism reveals itself and the society sees a complete breakdown of social organisation. Chaos and anarchy ensue, transcending into barbarism and eventually death. How is it that the most intelligent species living on Earth should find themselves essentially killing off their own when faced with adversities of "fear, hunger and isolation" [2], instead of adapting and developing resilience? For way of comparison, we can look at the contradicting and holistically altruistic behaviour and socio-psychology of ants. How do they go about necessary life activities and hostilities, surviving and staying highly organised? What can we learn from them, and how can we use it to influence new environments that may then alter the way humans respond in them?

The queen in an ant colony has no power over others: instead they work together, with each individual knowing what to do for the colony to succeed. They are a "fundamentally democratic" [2] society, whereby career changes evolve from: nursing for young, to cleaners, to foraging (and defence). They are organised into working groups that segregate spatially – giving us reason to explore how we may promote optimal living for humans by optimising spatial arrangement. Whilst we are unlikely to copy an ant's lifestyle, we can look to them for inspiration. Combining this with the aforementioned theme of hostile factors, this project envisions sitting in an ET environment, such as Mars. How, in fact, would these foraging excursions over time be translated to reflect both the environmental setting, as well as shape human living, on Mars?

As Warren M. Brodey said, "evolution now must include evolving environments which evolve man, so that he in turn can evolve more propitious environments in an ever quickening cycle" [4]. This "evolutionary dialogue" [3] considers organism with its environment, understanding how they may influence one another. This project looks to perceive spatial configuration resulting from bio-mapping ants over time with a Kinect V2, building a computational translation of "the functional space of an organism" [4]. It starts with a qualitative study on how the world is perceived through an ant's sensory capabilities, then goes on to extract data from a series of experiments with ants. From this, a prototype for human living in ET/extreme environments will be realised. This project is broken down into three main tasks: capturing activity, data mapping their 2D/3D co-ordinates over time, and addressing the rationale and implications of this. For the purpose of this study, harvester ant species were used. The aim is to understand why certain behaviour/form is the outcome.

2 BACKGROUND

Humans have been living in complex societies for only several thousands of years, whereas ants have been doing so for over 100 million years [5]. Despite the same daily travel challenges as people, ants don't get stuck in traffic jams. Crowding, bottlenecks and slow movers are all things experienced by ants too in their foraging activities. When ants walk past one another, the ant carrying something heavy (food item) is given priority, and the other ant will move to its side [1].

Speaking with leading myrmecologist, Dr. Nigel Franks [6] from Bristol University, we discussed the interdisciplinary connection between architecture and biology. Quoting Winston Churchill, he said "we shape our buildings, and

afterwards, our buildings shape us." Here, he refers to the effect that an ant's social nature of ants has on their environment and alterations of it [7]. They function collectively as though they are one entity. We can consider this on many levels from political to social, for example - through a person's desire to move to a city because of how they either feel there, or based on other's experience in such place [8].

It is interesting to note how both humans and social insects utilise space as a means of encoding social information; for instance, ants have separate areas for foraging, brooding, nest building, keeping debris and corpses aside [9], and humans too for private and public uses. This spatial organisation we see in these social insects is promoted by a "stigmergic nature" [10], whereby a trace left in the environment stimulates a feedback, or reaction, from another member. This concept will fuel the methodology in this study.

Basic needs for a human society to thrive include: constructing shelter, defence, gathering food and raising children. This requires some level of energy dissipation and organisation within a labour workforce. What if we then took away the centralised organisation and instead relied on a biological-based algorithm? [11] Harvester ants are chosen in this study because they are a species best known for their foraging capabilities, hunting predominantly for seeds, and occasionally being lured by honey/sugary sources [12]. We can build a perception of how a resource-poor living in an extreme environment (e.g. Mars) could work, both spatially and temporally.

3 RELATED WORKS
Bill Hillier, who wrote about strategies of effective design policies, explained that the mapping between "human behaviour and its spatial containment" [13] can be abstracted to "psycho-physiology and the environmental filter" [13]. From a practical point of view, mapping structures that communicate this relationship are used as "autonomic devices" to resolve challenges. From a research view though, the purpose of using a mapping device is to understand and improve them [13]. So in the same way, through plotting ant behaviour in a confined space, we create a mapping device that can be developed into an autonomic device, or prototype, that solves problems and optimises with time.

3.1 The Purpose Of Tracking
The mechanisms behind the biological algorithms present in the foraging excursions of harvester ants can solve real-world problems. These include: traffic organisation [14] (including the 'traveling salesman' dilemma [15]), economising energy, and minimising energy output for maximum food return – e.g. pheromones indicate the near-optimal path for ants to reach food.

Deborah Gordon, ecologist from Stanford University, has found that harvester ants, local to hot climates like deserts or Mediterranean countries, are biologically programmed to store food for several months and only hunt when it is worth the energy expenditure and loss of water. This is dependent on the rate of foragers returning to the nest in a very short amount of time [16]. Such a study serves the purpose of inspiring transport networks and managing traffic jams [17]; the use of sand and gel environments aims to shed light on this. Biologist Guy Theraulaz explained that "[ants] get together to form complex yet fully functional and reactive structures" [18], and this study hopes to apply this thinking to a prototype representative of simple rules leading to something more complex overall. Ecological success of ants lies in their ability to switch between one algorithm, or task/role, to another as needed. Moreover, they are able to find the near-optimal paths by following pheromones (static) as well as follow a backup plan in case of a glitch/break in trail in a dynamic environment [15]. This defines the flexibility and optimisation that such a system offers, both lucrative qualities for a prototype on Mars.

Inspired by Neri Oxman, whose research investigated biological behaviour stimulated by environmental factors, this study aims to address the gap of 3D tracking of insects. Oxman has used her bio-computational design research to synthesize between nature and man-made – this is a result of her investigation between an organism and its immediate surroundings. Neri and her team translated the "motion-capture data" [19] of silkworms with a 3D printer to create a 'Silk Pavillion' [20]. There are many ways to interpret the tracking of biological creatures, and feeding off this, my research looks to translate the tracking of ants to develop a spatial-temporal prototype for human living in ET/extreme climates, based on optimised paths depicted in an ant's biological algorithm.

The work of computational architect John Frazer examined forms and process that emerged from natural evolution. Common ground between this study and his reveals itself in the application of a biological system's optimisation capabilities. Frazer's work focuses on genetic algorithms (GA)/computational methods, whilst my research places more emphasis on social biological algorithms. Additionally, cybernetician Gordon Pask, highlights the relation between machine and organism and the conversation between them that starts to construct an intelligent environment [21]. This raises an interest in architecture adapting to positive feedback [22], much like the prototypes that arise later in this project do in accordance to ant trails.

3.2 Use of RFID Tags
Entomologists have used RFID tags to track animal behaviour: each individual member's change in position over time is captured. One key case study, based at the University of Lausanne in Switzerland [23] recorded each ant's position twice per second; this allowed them to recognise that ants make 'career moves' with time – progressing from nursing the larvae (nurses), to cleaning rubbish (cleaners), to foraging for the colony (hunters) [23].

Another study by Dr. Nigel Franks sees the use of RFID tags to "get each individual ant to identify itself" [24] and find out worker's tasks, methods of communication and decision-making for nest-finding. This led to the understanding of how tandem running displays a mentoring system amongst ants [25], similar to the release of pheromones when something good has been found [26]. Unfortunately I could not use this tracking system myself due to both access/financial cost of RFID tags and the difficulty in personally attaching a tag to each ant member, especially as an amateur ant keeper. Dr. Frank's studies did, however, raise the question about spacing of food. A scattered food situation would likely represent a scenario closer to that they would experience in nature, whereas a clustered one makes it easier for them: so how would the two scenarios affect their behaviour? Similarly, what effect would obstacles have on their journey/decision-making routes? These are much like the considerations regarding the use of complex transport systems in human societies to relocate goods where they are required [27].

3.3 Cameras and Sensors
3D ant tracking has also been achieved using an "image acquisition system" [28] that studies the moving kinematics of ants with 3D values. However, consisting of five very high definition cameras, also calibrated together with several infrared lights, it is also a very costly system. To adapt this to my project however, I could use optical cameras with either an IR sensor or Raspberry Pi, all calibrated and connected to one central system.

3.4 Kinect V2 vs. Wii Motion Plus
Both the Kinect and Wii devices utilise cameras and infra-red (IR) sensors, making them capable of tracking movement and objects. Whilst very comparable, the Kinect V2 has overall been reviewed a little higher, with its use even extending as far as to the military [29]. Moreover, there are more python image libraries and Java Processing libraries available for use with the Kinect, and generating point cloud data, making it a more suitable choice. Whilst there is a newer version of the Kinect boasting higher accuracy – the Microsoft Azure device – it is unfortunately out of reach with sales only in the US at the time being.

3.5 Video and Real-time Tracking
The video-tracking software named 'trackR' [30] went some way to inspiring the methodology used in this study; it performs multi-object tracking from videos while retaining individual tracker identities. To complement such methodology, mine will incorporate real-time tracking, which also allows for a 3D dimensional aspect. Biologist Simon Garnier, from Swarm Lab, has produced research with computational architect Dr. Tim Ireland on the passing of information through chemical cues between ants in a colony, and suggested that construction arises from the organisation of "individuals, their activities and their environment" [9]. This forms the basis upon which both the methodology will be built on, as well the analysis of the spatial templates that result.

4 METHODOLOGY

4.1 Part A: Animate Object Exploration in Spatial-temporal Dynamism
This section looks to explore the spatial-temporal dynamism of ants in their environment to develop a qualitative understanding of how ants distinguish their environment. This philosophical approach is inspired by Rosenbleuth, Weiner and Bigelow's cybernetics-rooted paper entitled "Behaviour, Purpose and Teleology", as well as Biologist Jakob von Uexküll's concept of 'umwelt'. The term is used to describe how organisms perceive the world using sensor (touch with antennae) and motor organs (smell and laying of pheromone trails), linking it to their subsequent behaviour [31]. The "intrinsically purposeful" performance of an organism is defined by a constant feedback loop. By mass communicating (quality and quantity) of food sources, ants work on a positive feedback loop [32] to produce a pheromone trail and recruit others in doing so [33]. This links to the aforementioned idea of 'stigmergy' and producing an 'active' region. Ants do not expand unnecessary energy [34]. Furthermore, pheromones not only indicate food sources, danger, and the queen nearing the end of her fertility, but also inform orientation. Their searching efficiency is maximised by these individual interactions. The mandible action from their jaws, defensive movement, and overall fluid movement come together to form a series of reactions to their environment. Ants perceive and navigate the world attaining towards a goal, where "survival and purpose intermingle" [35]. I imagine their world to appear like a monotone pixelated grid from their perspective, where each pixel is representative of a chemical cue or not, and then they make a binary decision (see Figure 1). This denotes how the 'social engineering' experiments are created for them.
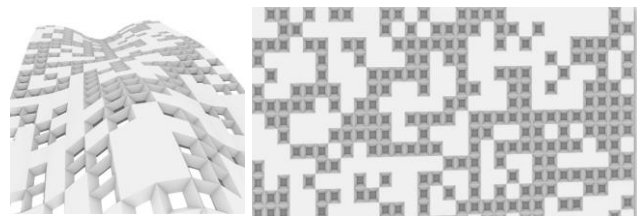
Figure 1. An ant's perception of the world.

An Organism's Perception and Response to Chemical Cues in the Environment
Organisms "detect and react" [36] to molecules in their surrounding environment, such as (hydrocarbon) pheromones, chemical cues warning of germs or predators, or pointing to food and shelter. Olfactory receptors allow ants to read the "biological barcodes" [36] which are the hydrocarbon molecules on the bodies of their nestmates. Therefore, this begs the question: how does an ant's ability to detect and discriminate between wanted (e.g. food) or unwanted (e.g. a repellent scent like cinnamon) generate different behaviour that can be tracked? More importantly too, could an architectural language be subsequently

interpreted from this? This will be integrated into the methodology for the experiments in the next section.
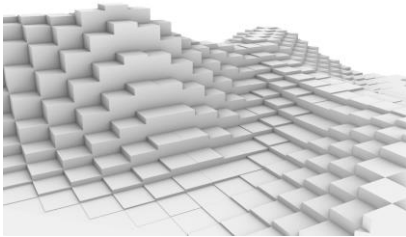
Figure 2. How individual decisions, based on chemical cues and interactions, add up to a fluid and collective movement.

4.2 Part B: Experimental Study

This part of the methodology looks to employ 11 harvester (Messor barbarous species) worker ants and a queen for the purpose of capturing their food foraging excursions with a Kinect V2, while influencing their paths (social engineering) over time. This is made up of the physical aspect of the organisms-in-environment, and the digital that allows them to be tracked and lead to a human prototype for living.

Computational Methodology steps:
1. Deep Learning: I used a single object image annotation tool (VoTT) to prepare/annotate 1000 images for ant detection and build in Yolo (darkflow) software (Real-Time Object detection) [37], using tiny-yolo weights.
2. Python with darkflow and Kinect V2: Now that the ants were recognisable by the system, the next step was to specifically track the movements of individual ants between frames/over time. Python libraries (openCV, darkflow, numpy and PyKinect2) were used to bring together the Yolo detection and the Kinect. This was done by getting the Kinect to return the depth information, then the centre of mass for all ants, and then the depth information for every ant. Finally, this data was recorded both in video format and as .csv format, storing each ant's ID and 3D co-ordinates over frames.
3. 3D prototype model: The final digital step sees the interpretation of the spatial-temporal ant trackings as prototypes for human living. This is done using the Wasp plugin for Grasshopper/Rhino, aggregating discrete modules that are essentially 3d modelled/printed along the ant traces.

Physical Methodology steps:
This part details the 'social engineering' that these experiments perform on the harvester ant colony. It is hoped that form and performance would be synthesized. The same ant colony is used per environmental setting.

Spacing of Food		Home (no. of nests)		Obstacle		Environment/setting		
Scattered (K1)	Clustered (K2)	1 (K0)	2 (K3)	Cinnamon (K4)	Perspex (K5)	Sand (K7)	Gel (K8)	Plain ant farm (K0, K1-5)

K6 represents the combinations from K2-5.

Table 1. Summary of experiments designed.
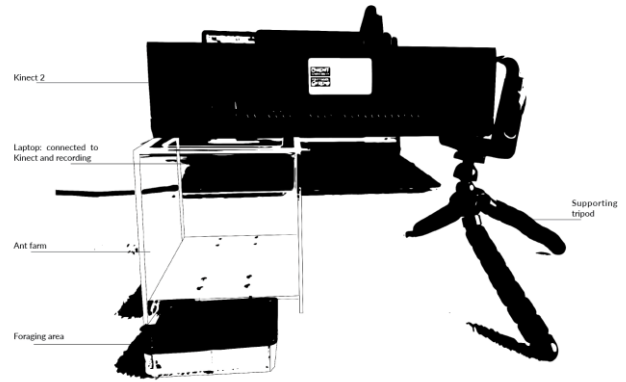
Figure 2. Setup using laptop, Kinect V2, and main ant farm. Experiments: Decision-making relating to foraging and nest-making/shelter, as effected bye environmental setting, obstacles and nest options.

Strategy
My interventions in their environment that aimed to encourage nest migration (for K3) included: tracing sugar trails towards the new nest, moving a couple of workers near the opening of the new nest, moving a couple of workers near the opening of the nest, adding food supplies and a red cover/lid to the new nest to immediately appear attractive.
1. Scattered: Food is spread around the main arena.
2. Clustered: Seeds are clumped up in two corners of the main arena.
3. Number of nests: This involves adding a connector that leads to a new nest. However this one is smaller than the first. Will they move, and if so how long will it take them to decide?
4. Environmental settings: Plain ant, sand or gel. White sand loam is used and it requires water to be mixed in until it visibly penetrates all the sand. Food is kept in the petri dish that is connected via tube. How does each affect behaviour or ease of tracking them?
5. Obstacles: One is a physical object – Perspex – cut out in such a way that it mimics ant journey paths and method of building, with traces of honey along it as a climbing lure. Cinnamon represents a chemical odour they find undesirable. This was placed along the entrance to the original nest, to see if it can deter/persuade them to relocate.

5 FINDINGS
Increasing complexity through experimental stages has given more opportunity to capture the transfer of information and resources by workers in their 'network analysis' [38].

Caring for the Ants
Initial difficulties found in caring appropriately for the ants include: the queen dying in the first colony (which shortened survival rate for the whole colony), a colony discretely escaping, and another colony not surviving in sand loam. This could be due to internal air temperature increasing

beyond what it should have (due to strong sun and heat in the month of July), mould growth (toxic for a colony), and lack of sufficient moisture in their environment. To improve conditions, a humidity sponge was added for later colonies. The advantage of using gel is that it provides an environment where ants can both feed from and dig into, like they would earth, whilst revealing their 'underground' travels within the gel; this was the easiest to take care of. The ants would also pile up dead bodies and waste to the side – making cleaning easier.

Social Engineering on the Ants

I have been leaving hints or traces in the environment of the ants (stigmergy (Grassé 1995)), manipulating their opportunities and challenges [34]. Change of behaviour was recorded; at least one ant felt the need to explore the quality/quantity of a food source.

It was found that changing scenarios in the ant farm also led the ants themselves to change between clustered and scattered groups. It is evident that individuals stick to their role (foraging, defence, nest maintenance or brooding) unless temporarily recruited to fill other shortages [34]. Splitting up the colony temporarily in the sand environment, between the connected petri dish, saw the colony first appear lost but to self-organise within 12h. Most food was stored at the bottom of dug tunnels in the gel and sand loam.

Hardware
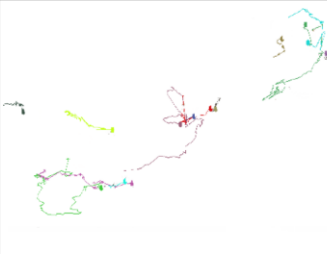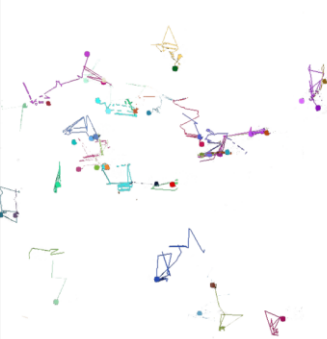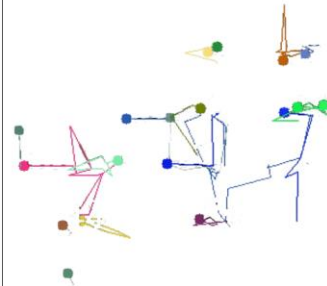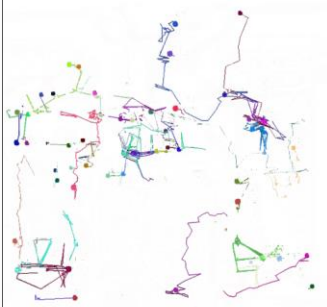
Limitations faced with the Kinect include camera quality, and size of pixel/object it can capture. Whilst this was partially overcome through YOLO/object detection, perhaps extended model training could have been put in place.
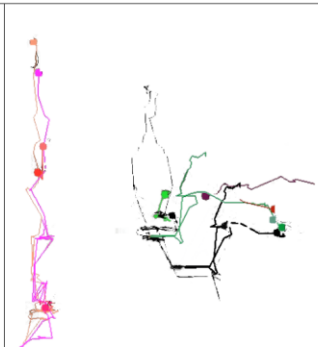
6 REFLECTION AND DISCUSSION

Increasing complexity through experimental stages has given more opportunity to capture the transfer of information and resources by workers in their 'network analysis' [38].

Due to the dispersion and vaporisation of pheromones, many of the results depict journey paths that "emerge, converge, fluctuate, and expire" [10] – which makes it hard to analyse the optimal route of an organism moving through a dynamic environment.

Aggregation Pheromones describes insects being attracted from a distance to cluster around a pheromone source. It is a natural evolutionary response to an organism's environment, as it may protect from desiccation and predators [32]. Just as their aggregation structure can be affected by environmental factors, so too can the architectural language and prototype that develops.

Exp	Spatial Template	Extractable Logic	Architectural Language
K0		Separation More individual work Prototype: describing arrival in a new environment	If we use the overall curve of the movement of the ants in this scenario as a boundary field, parts can follow and aggregate more heavily along the curve turns. This means more modules will cluster on those bends, as if to highlight their change in direction/decision.
K1		High interaction rate – lot of communication Geometric intersections Networking/ collective workings. Trying out lots of new routes until perhaps the decide on the most useful ones to take.	Just as the scattered food causes small cluster formations amongst the ant colony, here we see voxels aggregating along the field points driven by the ant networking trail. The prototype expresses the fluctuations in interactions over time too.
		Module: describing 'hot' foraging activity Repetitive modules: indicates this journey path works best for them	
K2		Individuals venturing to find best way (i.e. simplest) to food. Perhaps less successful than the other experiments, there is less of a specific language here.	This could be because it is likely their least challenging environment out of all the experiments: no over-excitement over many seeds everywhere, and no obstacles. Therefore no urgency to go out foraging perhaps. The overall movement can work as a field boundary line.
K3		Various tasks going on: foraging, nursing, storing food, cleaning. Upon feeling exposed, they might be communicating about the potential of finding another home, as there seems to be a lot of	Module extracted to express their repetitiveness and interaction: In line with their need to be organised for their survival and make moves worth the energy expenditure, the prototype here

433

		interaction and movement. Short routes to find new food - stochastic behaviour.	expresses a more orderly aggregation.
K4		Tandem running likely to have taken place here – a type of mentoring system. Not all ants returned the same way that they went out, indicating that some optimised their paths with time.	There are two overlaying languages here: one is a more strict geometric/90 degree turns in this spatial template than some other ones, and the other follows slightly more erratic lines. Therefore, the two modules also have their own rules, or 'connection grammars'. These come together in one prototype through aggregating the ends, long sides, and short sides of one module to the other.
K5		Either avoid obstacles and create their own direct corridor, or work around them as closely knit as possible within, maintaining close proximity to fellow foragers. Repetitive journey paths. Disjointed	Modules derived from movement patterns (as a result of tracing over their most repeated geometric routes): In line with their geometric pattern-making, we can translate their rules are directional, connecting the modules together through their centre. These parts can then be aggregated additively, joining together various faces of one module to another. Just like the ants navigate their environment, different chemical cues lead their decision making.
K6a (45)		Minimum energy expenditure – living as though resource-poor. Ideal for a living prototype in an ET/extreme climate.	-
K6b (234)		Optimised paths – short and near. Strong pheromones trails motivate the same journey paths.	-
K6c (245)		A kind of optimisation where one destination is more popular than another and so the route there is easier because there are chemical cues leading to it.	Finding the path of least resistance/difficulty towards a richer food source, can be expressed through a resembling sine wave field distribution, which is then modulated against the 3D co-ordinate points of their locomotion over time.
K7		Repetitive journey paths from many individuals – indicates strong pheromone routes and communicates colony's need to build. Most efficient way to dig through a solid ground to get there?	The ants building algorithm. The language described here expresses the consistent and continuous movement of aggregating one sand particle atop another – much in the same way, a simple module is aggregated atop the other along their movement lines. How we can create habitats that are dug into a solid
K8			environment, that are as efficient as possible, but create: e.g. an assembly hall, etc.
		-	-
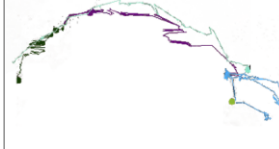
Table 3. Detailed analysis and extractable architectural language from each experiment where possible. The different colours denote different ant identity.

Aggregation Pheromones

Just as ants use a combination of 'recruitment systems, with signals and pheromones from two potential glands [32], spatial templates like K1 use more than one type of module that aggregates in stochastic ways. Furthermore, ants also use 'tandem running' recruitment, where one ant is led to a "new nest of food source" by a "returning scout" – this seems likely for quite a few of the spatial templates, such as K2, K3 and K5, where we see either longer routes taken or an ant following a returning one. This can be recognised as field lines in some of the prototypes to follow. Foraging communication methods also allow for the colony to "retain a memory of previously rewarding locations" [32] and to be able to select between them. This would explain the findings in experiments like K6C, where I identified that there seemed to be a preferred food source, and also explicates how the ants knew to re-visit the one with a richer amount of seeds. Spatial templates, like K1 and K4 in particular, which show a lot of ants on the scene indicates a mixture of 'group recruitment' and mass communication took place – which entails recruiting a lot of their nestmates for foraging [32]. The journeys mapped are themselves a "communication network" [10] based on pheromone trails. It is also interesting to note that there is a lack of symmetry in all the spatial templates. This indicates a different route outbound than inbound over their foraging excursions.

In order to relocate, ants would need to see new benefits being offered, such as improved nesting. Whilst I did try to

reduce attractiveness of their current home by removing the red lid (thus making them feel exposed) [24], moving away some of the food they had gathered and to stop adding as much moisture to their nest – the ants found it simpler to just re-gather food. They never moved and I realised this must be because relocation also has its costs: energy expenditure [39], potential loss of workers [40], reduced foraging time, and desiccation [41]. However, this is part of their biological efficiency and natural defence system.

Experiment Limitations

Some limitations were observed over the course of the study. Firstly, the ant farm setting is contained and so with limited travel distances, ants likely act in a different way than they would in nature. Environmental challenges such as finding a suitable nest site, establishing territories, defending against enemies, and traveling far to find food sources are all reduced [34]. Furthermore, temperature and humidity were not accurately monitored – this would greatly affect rate of evaporation of pheromones and in turn the way journeys are taken. This was not used as a variable though because it could kill the ants if not done with suitable sensitivity. Colony sizes used were relatively small (up to 11 workers) to ensure they could be tracked with ease, especially when near one another. Perhaps in the future, this could be expanded upon.

7 A PROTOTYPE FOR HUMAN LIVING

This section is a synthesis of the analysis so far and experimental results, where the ant paths are translated through fractalated voxels.

An Aggregation of Pixels

We know now that "ants don't waste their energy and share clustered resources" [42] - they only go out foraging as a result of positive interactions. Ants also know their limits and don't exhaust themselves. Imagine the system we could build for humans if we followed these simple rules. This chapter will therefore apply the outcomes of the experiments to a prototype for human living on Mars, an 'unreliable environment', where we assume life there would start out depending on resource-poor construction.

The models produced in this study are data-driven, fusing both theoretical/qualitative work as well as experimental. This will culminate in the translation of these "efficient collective decisions in unreliable environments" [43]. Ant behaviour is based on adaptability: according to the number of ants and changes in the environment, their architecture changes – much in the same way, a living building would need to alter according to environmental changes [18] and the symbiotic relationship between organism and environment.

Brodey described that "an aggregation of simple machines grows only into a complicate machine decomposable into simple elements" [3], much in the same way that ants follow simple rules to create a collectively more complex system

[44]. The social algorithm of ants can be broken down into a series of binary decision points. At each decision point, an individual colony member either proceeds down one route or another until the next decision needs to be made, or the sequence ends. This is makes up the basis of their self-organised system. Thus, the aggregation of modules gives rise to a more complicated prototype. In the same way that some of the spatial templates display stochastic behaviour [32], the aggregation modules in these prototypes are too stochastically aggregated – i.e. different rules are applied in the additive process. Additionally, in the previous chapter we see the modules start revealing themselves in the spatial templates – what if these elements become fractalized modular pieces, aggregating together, translating each pixel of ant locomotion as a voxel? We can now consider how discrete modules could be used to realise this into a prototype that is customised to evolve space for specific needs.

Discrete Architecture

A computational design theory that approaches abstraction in architecture, known as 'Discrete' or as 'computational mereology' [45], loses focus on the holistic and instead prioritises its parts. Similar to the brushstrokes in Rothko's abstract art that are concerned with organisation through parts, architects of this movement concern themselves with "organisation of material parts or particles" [46]. Mario Carpo, architectural historian, has also described computation as an abstract and discrete process [47]. Also of interest, is that the post-capitalist technology employed works to "democratise and decentralise production" [48] with its automation, which in itself is also reminiscent of the democratic and decentralised organisation of the ant society. With this in mind, the architectural prototype that emerges from abstract parts can even be configured from "extreme discretisation of statistical pixel values" [45]. This is similar to how ants perceive the world in bits/parts, how they make their binary decisions, and how their trails are made up of each pixel/ant over space and time. This method of aggregational architecture is also fully reflective of both the process and findings produced in this study – the spatial templates are in themselves pixels of "coarse discreteness" [49].

Furthermore, in our consideration of material organisation that is indicative of big data organisation, we also lean towards a type of discrete assemblage and fabrication. This allows parts to be disassembled, and then aggregated with differentiation [50]. Much like how the ants avoid entropy through ascending adversities over time, this kind of fabrication too holds "low-entropy heterogeneity" by being feasible and cheap due to 'serial repetition' [51]. This makes it especially suitable for a hostile environment like Mars.

As such, the data collected from the experiments has been input to a multi-object behaviour library, effectively creating

scenarios for human living in an ET setting. Some of the outcomes can be seen below.

Figure 4. K4 3D prototype made up aggregated parts representing ants following chemical cues in their environment in a spatial-temporal sense.
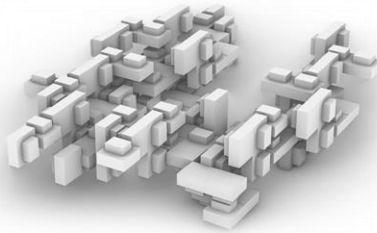
Figure 5. K5 3D prototype.

8 CONCLUSION

This interdisciplinary project has undertaken experimental research in the social organisation of ants, showcased through observation, bio-mapping their locomotion as affected by their decision-making in varying situations with increasing complexities over time. This study has extracted their biological behaviour to apply to spatial organisation in architecture, expressing their activities as a "dialogue between the subject and its environment" [10]. Every trail has a purpose and intent, which defines organisation, and results in a "spatial intelligence" that articulates the actions of these algorithm-guided individuals as a pattern [52]. These spatial templates are in themselves a "social product" that has unfolded as a result of interaction [53]. However, just as ants run on algorithms and know what to do, an intelligent architectural environment for humans wouldn't think like us. Rather, it would be capable of "rational autonomous action" that optimises the surrounding functions [21].

In this concluding chapter, I come back to the fictional story of 'High-Rise' to finish the comparison between man and ant, with themes revolving around resilience and optimisation. In the face of growing adversity over time in a fixed environment, the ants behaved in the polar opposite way to people. Much like how entropy deduces that chaos increases with time, as was the case in J.G. Ballard's story, the ants actually became more organised and worked better together. They were more wary of energy expenditure, organisation, colony survival, and teamwork (which comes in the form of group recruitment, mass communication, and tandem running). By following algorithms, worker ants avoid chaos, while still acting in a timely fashion [34].

Moreover, if an animate object and machine could self-regulate, then the chance of entropy are reduced through information return [21], as it adds control to the situation.

Ultimately, the concluding prototypes are a type of nest morphology: it shows how an organism has reacted to its environment, and the communication that has occurred between workers. The voxelated formations can be considered as frozen moments in time as the ants trace through their environment. Basing this process on a set of general rules that extracted from their biological algorithm acknowledges the balance of objectives against efficiency. The implications of this is the concept of an intelligent environment [3] – entailing "complexity, self-organisation… [and] responsive" as it perceives and reacts to surrounding stimuli [21]. The evolution in ants has also taught them to avoid disease and produce high productivity workers [54] – which hints ability to help even in medical fields.

Social engineering on the ants allows us to gain an understanding of how ants problem-solve. This is a key aspect of space syntax. Each pixel of ant movement over time in space is interpreted as a part, where they aggregate as a result of stochastic patterns, boundary field lines, and rule grammars regarding connecting faces and edges. Indeed, the olfaction of ants has created a new (3D) architectural language for responding to the environment. As we see from the prototypes, there are a lot of combinatorial possibilities and no finite answer. This, along with their possibilities of being easily disassembled, offers greater flexibility in construction and spatial arrangement – making it especially adaptable. For an environment like Mars, resilience would be another key factor. With its difficult atmospheric conditions, the situation lends itself to the creation of habitats that are dug into a solid environment as efficiently as possible; these studies have revealed some economical possibilities of digging through. This in itself starts to answer how the setting of Mars would alter human behaviour and living. Just as our buying and selling habits set market prices, and frequently visited websites contribute high rankings on Google searches [7], we can understand how we shape something bigger and more complex than our individual self.

The proposed prototypes are an offspring from the new knowledge gained about ant's behavioural relevance to architecture processes and products in creating a resource-poor living that would be feasible in an ET setting. A step further would be for the system to self-repair materially too, and to consider a more compact living arrangement for thermal and economic efficiency in an extreme environment. Therefore, further research could see these architectural aggregations optimized for extreme thermal environments.

REFERENCES

1. Hanson, Dr. Joe. *Why Don't Ants Get Stuck In Traffic?* Pbs Studios, 2016.

2. TEDxZurich, [writ.]. *Fascinating ants -- lessons for humans?* [perf.] Danielle Mersch. TEDxTalks, 2014.

3. *The design of intelligent environments, soft architecture.* Brodey, Warren M. 1, Cambridge, MA : s.n., 1967, Landscape, Vol. 17, pp. 8-12.

4. *Umwelt-theory and pragmatism.* Sharov, Alexei. January 2001, Semiotica, Vol. 134, pp. 211-228.

5. Abboud, Rebecca Elias, [perf.]. *Ants help us learn more about human society.* University of Pennsylvania, 2015.

6. Franks, Dr. Nigel. *'Tracking ants'.* [interv.] May 2019.

7. TEDxLausanne. *Ant societies and what we can learn from them.* [perf.] Laurent Keller. TEDx Talks, 2012.

8. Adams, Jane, [perf.]. *Emergence (or: How Ants Find Your Picnic Basket).* [prod.] TEDxGallatin. TEDx Talks, 2013.

9. *Architecture, space and information in constructions built by humans and social insects: A conceptual review.* Ireland, Tim and Garnier, Simon. 2018, Philosophical Transactions of The Royal Society B Biological Sciences.

10. *Stigmergic Planning.* Ireland, Tim. New York : ACADIA 10: LIFE in:formation, 2010. 30th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA).

11. Gordon, Deborah M., [perf.]. *Inside the ant colony.* TED-Ed, 2014.

12. Reynolds, Matt. The natural genius of ants is helping us build better algorithms. *Wired.* [Online] 01 June 2019. [Cited: 01 July 2019] https://www.wired.co.uk/article/ant-algorithms-nature?utm_medium=applenews&utm_source=applenews.

13. *How is design possible?* Hillier, Bill and Leaman, Adrian. 1, January 1974, Journal of architectural and planning research, Vol. 3.

14. Kubota, Taylor. Stanford researchers found an algorithm that explains how ants create and repair trail networks. *Stanford News.* [Online] 2 October 2017. [Cited: 2 September 2019.] https://news.stanford.edu/2017/10/02/algorithm-ants-create-trail-networks/.

15. *Swarm Smarts.* Bonabeau, Eric and Theraulaz, Guy. 3, 2000, Scientific American, Vol. 282, pp. 72-9.

16. *Collective Wisdom of Ants.* Gordon, Deborah M. February 2016, Scientific American.

17. Latty, Tanya. Building better transport systems: lessons from ants and fungi. *Cosmos: The Science of Everything.* [Online] 09 June 2018. [Cited: 09 September 2019.] https://cosmosmagazine.com/biology/building-better-transport-systems-lessons-from-ants-and-fungi.

18. Singer, Emily. The remarkable self-organization of ants. *The Guardian.* [Online] 11 April 2014. [Cited: 10 September 2019.] https://www.theguardian.com/science/2014/apr/11/ants-self-organization-quanta.

19. Slane, Kevin. 5 things to know about Neri Oxman, the MIT professor linked to Brad Pitt. *Boston.* [Online] 10 April 2018. [Cited: 10 September 2019.] https://www.boston.com/culture/entertainment/2018/04/10/5-things-to-know-about-neri-oxman-the-mit-professor-linked-to-brad-pitt.

20. *Silk Pavilion: A Case Study in Fiber-based Digital Fabrication.* Oxman, Neri, et al. 2014. Farbicate: Negotiating Design & Making. pp. 248-255.

21. Yiannoudes, Socrates. *Architecture and Adaptation: From Cybernetics to Tangible Computing.* s.l. : Routledge, 2016.

22. *Adaptive-Conditional Architecture.* Eastman, Charles. Manchester, UK : London Academy Editions, 1971. Design Participation: Proceedings of the Design Research Society's Conference. pp. 51-57.

23. Yong, Ed. Tracking whole colonies shows ants make career moves. *Nature.* [Online] 2013. [Cited: 02 September 2019.] https://www.nature.com/news/tracking-whole-colonies-shows-ants-make-career-moves-1.12833.

24. Franks, Dr. Nigel, [perf.]. *Antzilla - Tracking Ants / Wild Showcase.* National Geographic Wild, 2012.

25. *Teaching in tandem-running ants.* Franks, Nigel R and Richardson, Thomas. 7073, 2006, Nature, Vol. 439.

26. Franks, Nigel and Sudd, John. *The Behavioural Ecology of Ants.* New York : Chapman & Hall, 1987.

27. *The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles.* Guimerà , R, et al. 22, 2005, Proc Natl Acad Sci U S A., Vol. 102.

28. Vision, R&D. R&D Vision. [Online] 2017. [Cited: 14 September 2019.] http://www.rd-vision.com/items-en/3d-tracking-of-ants#contact.

29. Fitzpatrick, Alex. 3 Ways the Military Wants to Use Microsoft Kinec. *Mashable UK.* [Online] 19 June 2012. [Cited: 19 July 2019.] https://mashable.com/2012/06/19/military-kinect/?europe=true.

30. SwarmLab. *trackR.* Berkley, CA, US : University of California at Berkeley, 2018.

31. Uexküll, Jakob Johann von. *A Foray Into the Worlds of Animals and Humans: With A Theory of Meaning.* s.l. : U of Minnesota Press, 2013.

32. Wyatt, Tristram D. *Pheromones and Animal Behaviour.* 2nd. Oxford : Cambridge University Press, 2014.

33. Gordon, Deborah M., [perf.]. *Inside the ant colony.* TED-Ed, 2014.

34. Wilson, Edward and Hlldobler, Bert. *The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies.* s.l. : Norton & Company, 2008.

35. *Behavior, Purpose and Teleology.* Rosenblueth, Arturo , Wiener, Norbert and Bigelow, Julian. 1943, Philosophy of Science, Vol. 10, pp. 18-24.

36. Riverside, University of California -. Olfactory receptors that enable ants to smell and recognize workers, males, and their queen identified. *ScienceDaily.* [Online] 2017. [Cited: 15 September 2019.] https://www.sciencedaily.com/releases/2017/08/170817122 127.htm.

37. Redmond, Joseph. *YOLO: Real-Time Object Detection.* Seattle, USA : Darknet, 2016.

38. *Resource redistribution in polydomous ant nest networks: local or global?* Ellis, Samuel, Franks, Daniel W. and Robinson , Elva J.H. 5, 2014, Behavioral Ecology, Vol. 25, pp. 1183–1191.

39. *Speed versus accuracy in collective decision making.* Franks, Nigel R, et al. 2003. Biological Sciences. pp. 2457–2463.

40. *Nest Relocation and Excavation in the Florida Harvester Ant, Pogonomyrmex badius.* Tschinkel, Walter R. 19 November 2014, Plos One.

41. *Variation in nest relocation of harvester ants is affected by population density and food abundance.* Pinter-Wollman, Noa and Brown, Mark J.F. 6, 23 July 2015, Behavioral Ecology, Vol. 26, pp. 1569–1576.

42. Blackman, T.J. Life Lessons We Can Learn From Ants. [Online] 9 July 2015. [Cited: 13 September 2019.] https://learn.eartheasy.com/articles/life-lessons-we-can-learn-from-ants/.

43. NJIT. Garnier, Simon J. *New Jersey Federated Department of Biological Sciences.* [Online] 2012. https://biology.njit.edu/faculty/garnier#research.

44. Colchester, Joss. Simple Rules – Complex Behavior. *Systems Innovation.* [Online] 29 June 2019. [Cited: 15 July 2019.] https://systemsinnovation.io/simple-rules-complex-behavior/.

45. Foster+Partners. Mars Habitat. *Foster+Partners.* [Online] 2015. [Cited: 2019 July 01.] https://www.fosterandpartners.com/projects/mars-habitat/.

46. *Autonomous Additive Construction on Mars.* Wilkinson, Samuel, et al. Florida, USA : s.n., 2016. Earth & Space 2016 - ASCE International Conference on Engineering, Science, Construction and Operations in Challenging Environments.

47. *Discrete Architecture in the Age of Automation.* Retsin, Gilles. 2, 01 March 2019, Architectural Design, Vol. 89.

48. *In Part Whole: The Aesthetics of the Discrete.* Retsin, Gilles. 5, 05 September 2019, Architectural Design, Vol. 89.

49. Carpo, Mario. Breaking the Curve: Big Data and Design. *Artforum.* February 2014.

50. Mason, Paul. *Post Capitalism: A Guide to Our Future.* London : Allen Lane, 2015.

51. Carpo, Mario. *The Second Digital Turn: Design Beyond Intelligence.* Cambridge, MA : The MIT Press, 2017.

52. *Post-capitalist design: Design in the Age of Access.* Sanchez, Jose. 2014. Paradigms of Computing, ACADIA 2014.

53. *Discrete Assemblage as design and fabrication strategy.* Retsin, Gilles. 2015. TxA Emerging Design + Technology.

54. *Sniffing space II: The use of artificial ant colonies to generate circulation patterns in buildings.* Ireland, Tim. Montreal : s.n., 2009. CAAD Futures: Joining Languages, Culture and Visions. pp. 214-227.

55. *The Spatiality of Being.* Ireland, Tim. January 2015, Biosemiotics.

56. Pennsylvania, University of. *Ants help us learn more about human society.* [perf.] Rebecca Elias Abboud. 2015.

57. Hall, Chris. Why JG Ballard's High-Rise takes dystopian science fiction to a new level. *The Guardian.* [Online] October 2015. [Cited: 12 September 2019.] https://www.theguardian.com/books/2015/oct/03/jg-ballards-high-rise-takes-dystopian-science-fiction-to-a-new-level.

58. *Architecture, space and information in constructions built by humans and social insects: A conceptual review.* Ireland, Tim and Garnier, Simon. 1753, August 2018, Philosophical Transactions of The Royal Society B Biological Sciences, Vol. 373.

59. *Decoding ants' olfactory system sheds light on the evolution of social communication.* d'Ettorre, Patrizia, Deisig, Nina and Sandoz, Jean-Christophe. 34, 2017, Proceeding of the National Academy of Sciences of the United States of America (PNAS), Vol. 114, pp. 8911-8913.

A Three-Tier Architecture Visual-Programming Platform for Building-Lifecycle Data Management

Mahmoud M. Abdelrahman[1], Sicheng Zhan[1], Adrian Chong[1]

[1]**Department of Building, School of Design and Environment,**
The National University of Singapore, 4 Architecture Drive, Singapore 117566.

ABSTRACT

In this paper, we present a platform that integrates three main aspects in the building industry: 1) Building data from both IoT devices and Building Management System (BMS), 2) Building Energy modeling and simulation engine, and 3)Data analysis and optimization libraries. All of which are combined in a three-tier architecture cloud platform. The platform aims to provide useful representations of the data to different stakeholders. We defined three main types of stakeholders. Each stakeholder uses the platform differently: (1) End-use programmers, who use visual programming and textual programming interfaces to perform computational tasks, (2) Dashboard viewers who are interested in viewing insightful real-time data about performance in the form of charts and diagrams, and (3) Data feedback inputters such as occupants to give feedback or fill questionnaires. The three-tier architecture enables the spatial and physical separation of the databases, the computational engines, and the user-interfaces. This separation resulted in some advantages such as flexibility, scalability (horizontal and vertical), reusability, and latency reduction. Currently, the platform is in the final stage of the alpha release development and will be released as open-source platform. The following stage includes community testing and user experience enhancement.

Author Keywords

three-tier architecture; n-tier architecture; Building Energy Modelling; BEM; building lifecycle; actors; presentation layer; application layer; data layer

ACM Classification Keywords

I.6.1 SIMULATION AND MODELING: Computing methodologies Model development and analysis - Software and its engineering Visual languages

1 INTRODUCTION

The building sector has witnessed immense development recently in the way by which building systems are managed [45]. This development aimed at alleviating the significant environmental impact of this sector (30% of the world energy consumption and a third of the associated CO2 emissions [20]). Decreasing this impact could be achieved by better controlling the resources [5]; providing sustainable, and more efficient solutions [46]; developing a better understanding of different deterministic and stochastic aspects of the built environment [10, 36, 43, 11]. In addition, making better decisions based on mining ground-truth data (black-box approach)[31], physics-based simulation models (white-box approach)[41, 31] or both of them (gray-box approach)[46].

The emergence of the Internet of Things (IoT) devices has enabled a massive amount of data, which posed some challenges. The data collected during the last two decades exceeded that which has ever been collected in history[34]. Two significant challenges, other than the availability of data, are to be addressed. On the one hand, managing this big data: transferring, storing, preprocessing, wrangling and mining, optimization, and control in a robust cyber-infrastructure. This lead to cloud computing or ubiquitous computing, that is, computing data in place, without paying much effort in transferring data to local storage/processing machines using scalable storage and computational power on demand maintained by professionals [28, 7, 14]. On the other hand, delivering useful information to different stakeholders based on their use is yet another challenge.

"Data is not information; data must be presented in a usable form before it becomes information" [35, p.134]. Raw data from the built environment varies in its degree of usefulness to different stakeholders (actors). "Usability" is different for different stakeholders. For instance, A Building Energy Modeller (BEM) requires data such as Green Building eXtensible Markup Language (gbXML) [19], schedules of operations, set-points, number of occupants,etc.. At the same time, a Facility Manager (FM) would be more interested in reducing maintenance costs by fault detection and diagnostics (FDD)[47] algorithms and dashboards. This difference in uses requires users with proper domain knowledge alongside with programming or procedural thinking skills for automating, prototyping, analyzing, building work-flows using the big data and IoT sensors. At the same time, they do not need to be professional programmers, but rather End-use programmers.
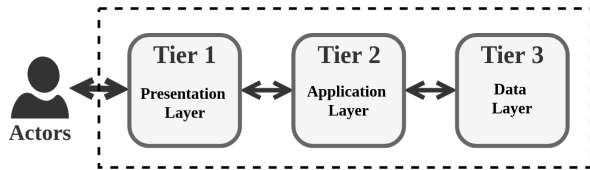
The term *End-use programmer* (or Novice developer) first introduced by [32] is used to describe non-professionals without programming and code structures knowledge [**?**]. Most programs today are written by novice programmers [37] who use spreadsheets, writing add-ons/plugins to support their work or add functionality to the software, running MATLAB simulations [26], writing python codes, or IPython notebooks [39]. The main difference between the end-user developer and the professional developer is the goal of the development. The former writes a program focusing on getting tasks done without paying attention to debugging, unit-testing, or re-usability. However, the latter develops, maintain, debug, and test a robust software for others to use[25]. End-use development mostly involves modifying or extending a software artifact (known as programming by example, or programming by demonstration). For example, *Extended annotation or parameterization*: whereas a user can define a new functionality by selecting a readily developed set of functions by other users and stored in shared repositories [24]. To make this task handy to end-use programmers, a proper programming notation should be utilized.

Programming notation is the way by which a programming language is used to represent the state of the world[17]). The selection of the proper programming notation (e.g., textual, diagrammatic, or gesture-based) to describe the problem is a trade-off between several evaluation criteria[17]. This trade-off should lead non-specialist to find the programming environment easy-to-use and easy-to-understand.

A large body of studies have addressed some difficulties related to the everyday programming tasks such as, but not limited to, code navigation [13], and hidden dependencies. A study conducted by Bragdon et al. [9] that uses visual programming 2d spatially connected blocks of codes where it is possible to track each function in the form of bubble showed that visual programming significantly improve code understanding with less code navigation time.

Visual Programming Language (VPL), also known as box-and-wire programming or data-flow programming utilizes user-friendly graphical elements to display the flow of data between different components (Figure 1). To date, there has been little agreement on the superiority of visual programming languages (VPLs) over the textual programming languages (TPLs). Green et.al. [18] and Moher et.al. [30] adopted this stance. However, [29] argued that VPLs could be significantly useful in motivating beginners to learn to program as they are informative in the sense of spatial reasoning and ideas generation. However, he argued that those advantages are not generalized and should not be isolated from the task being studied. Also, TPL is more productive in terms of large-scale and complex software development tasks. There is no conclusive evidence of the superiority of one over the other. Thus, a hybrid VPL/TPL is proven to be suitable amongst users. Many software enables this type of scripting which gained wide popularity such as Rhinoceros VPL Grasshopper[6, 27] whose user-community extended its limits by contributing to developing TPL based functions amongst others for building and urban simulation

Figure 1. A screenshot from our platform indicating a parametric simulation process on the cloud using (BuildSimHub). The figure is an example of Visual programming Language (also known as box-and-wire programming notation, and data-flow programming) where blocks of code functions are encapsulated in the boxes (known as components), while data flows from one or more component output variable to other component input variable. A component can hold as many inputs as needed, while the code behind each component follows a specific schema. Some components are used only as input components such as "file, numerical, panel" and some components work as output visualizations such as "panel, 3d visualization canvas, plotting components .."

[38, 33, 22], daylight analysis [21, 23], Machine learning[2, 1, 3].

This research addresses two primary problems:

1. Managing buildings' big data requires flexibility and scalability.

2. Different users use data differently.

We introduce a Three-tier architecture web platform as a resolution to these problems. The objectives of this research are: (1) Making IoT data from the built environment useful to different stakeholders, (2) An Open platform for developing and sharing different models amongst end-use programmers. (3) Provide flexibility and scalability by separating the platform into three distributed layers. The rest of the paper is structured as follows: In section 2, we explain the three-tier architecture. The components of the three-tier architecture are explained in details in the subsequent sections 3 (Actors), 4 (Presentation Layer), 4.2 (Application Layer), and 6 (Data Layer).

2 THREE-TIER ARCHITECTURE

Three-tier architecture is a client-server software architecture pattern that has several potentials in distribution applications [8]. The key aspect of this architecture is that it enables the reusability of program building blocks (also referred to as components) such as sharing software components at run time, replicating the same component, controlled adaptation of code, or independent adaptation of components [42, 44]. This pattern consists of three functionally separated layers [40] (Figure 2), those are: **1) Presentation layer** (the client-side), **2) Application layer** (also known as business logic, controllers and process management layer), and **3) Data layer** (also known as data storage). The presentation layer constitutes the user-interface side of the application. All human-computer interactions (HCI) happen on this layer. On

Figure 2. Three-tier architecture structure of the platform

the other side, the Business logic layer consists of the core engine where component functions that carry out the workload by receiving requests from the presentation layer get the relevant data from the Data layer and process this data then, sends back the response to the presentation layer. Finally, the Data layer consists of data querying (acquire from a data source), storage (stored in the relational database on a local server), and accessing (feed to the user). Each of these three tires is discussed in detail in the subsequent sections. A summarized workflow is illustrated in figure 3.

Advantages of using three-tier architecture [16]:

1. Reusability: The use of a modular component-function system enables many users/developers to contribute to the development. This contribution is shared in a public repository to be reusable by other users with small or no change

2. Scalability (both horizontally and vertically) is another essential feature of using three-tier architecture because of using distributed servers as well as because of its separation nature. Horizontal Scalability is achieved by adding more nodes of the same types where required. While the Vertical Scalability is achieved by adding more resources to a host node on demand. On the other hand, separating different tiers allows scaling each one independently depending on the needs at any given time.

3. Flexibility: As server nodes that are used in the application layer could be developed, configured and tested separately without affecting any of the other layers, and then it could be added to the system on demand. Readily PaaS (Platform as a Service) is used to maintain consistency and Scalability.

4. Latency reduction: As the nodes are distributed on data centers that are spatially close to the end-user as possible. This is basically handled by adopting open PaaS (Platform as a Service) where the infrastructure complexity of servers are maintained and optimized by professionals.

5. Anytime data processing regardless of the user's connection speed. The connection between the application layer and the data layer is separated from the presentation layer, reducing the bandwidth load on the user's side.

3   ACTORS

We use the term "Actor" to refer to different users/stakeholders within the built environment. "Usability" is the critical factor in obtaining "information" from data. Data obtained from the built environment does not hold equal usefulness for all stakeholders. Furthermore, in its raw format, IoT data consists of many problems, including noise, missing data,

non-valid data, ill-labeled data points, duplicates, and non-standardized IoT representation. These problems should be dealt with before delivering this data to non-expert users. We define some actors or stakeholders who are involved in the building's life cycle energy modeling and data analysis and their corresponding data usage based on a systematic review conducted by Zou et.al [48] of 10 years publications till 2017 and using Natural language processing (NLP) [4].

Then, we identified three main data-related actions, namely, **End-use programming**, **Data monitoring**, and **Data input**. Different stakeholders are clustered based on these three activities (Figure 4). Then, these activities are reflected in the presentation layer *(explained in section 4)* as three major components: 1) Hybrid VPL/TPL interface, 2) Interactive, shareable, embedded dashboards, and 3) Shareable, embedded forms. The interface of the platform is shown in figure 5.

4   PRESENTATION LAYER

The presentation layer acts as the interface for the users. Each new project has a unique global unique identification (GUID), which is stored on the data layer and could be called using its GUID. Three main functions used in the presentation layer: (1) Visual/Textual programming interface, (2) dashboards, and (3) forms are used.

4.1   VPL/TPL interface

Both the visual and the textual programming interface use two main languages: Python and JavaScript. However, the selection of which is based on the complexity of the task and the response time required. For example: if the component function requires a real-time reaction, then it is more suitable to use a JavaScript component/script. We refer to the component in this case as *"Shallow Function"*. While if the function requires heavy calculation such as Machine Learning (ML), or Energy Simulation (ES), then Python-based components/scripts are used and are referred to as *"Deep Functions"*. Examples are shown in figure 6.

Each component must follow specific abstraction schema that defines its structure and relationships as well as its types. Mainly, it consists of inputs, the function body, and outputs. However, other information must be provided when defining a new component e.g., The `dataflowType`: either shallow or deep, the component `type`: numeric, panel, optionList, ListItem, Plot, Generic, etc., the `category` where this component falls into, and other supplementary features such as the color, documentation, and license.

Each component has a unique GUID on a project level. This GUID is used to trigger the component and call its inputs and outputs using the TPL panels. For example: the deep component (Figure 6) has a GUID : `'c56ad1cf-47ea-4fe3-805b-84d104ecff8b'` could be triggered using the Python API . This allows the end-use programmer to extend the functionality of the components by running blocks of codes either before running the component (i.e. doing preprocessing to the inputs) or after running the component (i.e. doing post-processing to
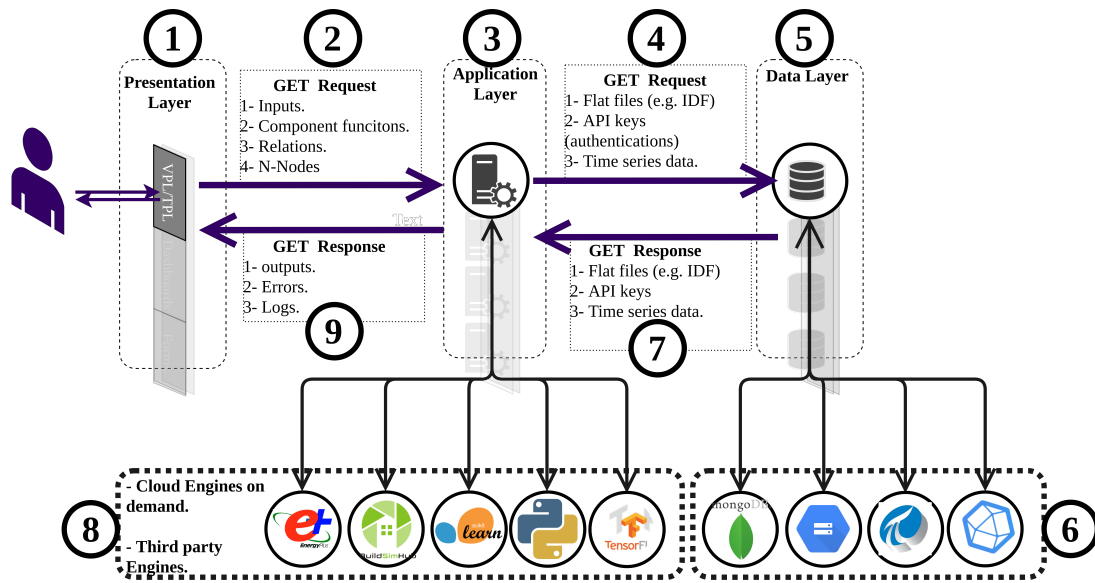
Figure 3. This figure shows the flow of data within the three tiers: (1) The actor interact with the presentation layer three ways: programming, viewing dashboards, or giving feedback; (2)The user interaction results in sending requests to the application layer to be processed; (3)The application layer starts processing inputs and relevant functions; (4) If there is any required data from the databases, a QUERY is sent to the data layer; (5) then, the data layer respond with the corresponding data from different databases (6,7); (8) The application layer then runs the required components using the distributed engines; (9) The output of step (7) is responded back to the presentation layer along with the logs/errors in JSON format. higher resolution image can be downloaded here
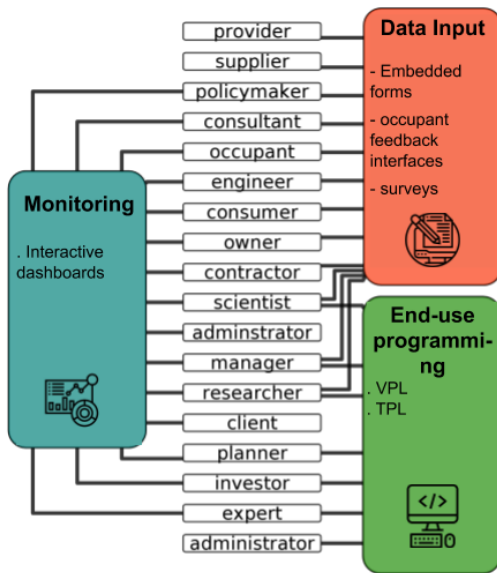
Figure 4. Actors are functionally clustered into three categories: Data monitoring, Data input, and End-use programming. Those categories constitute the Presentation layer.

the outputs) using the functions `before(callback)` and `after(callback)` as shown in listing 1 and figure 8.

```
1  import buildFit as bf
2  import json
3
4  component = bf.getComponent.by_guid("c56ad1cf-47ea
      -4fe3-805b-84d104ecff8b")
5  if component.inputs[0] == None:
6      compoent.inputs[0] = "http://google.com"
```
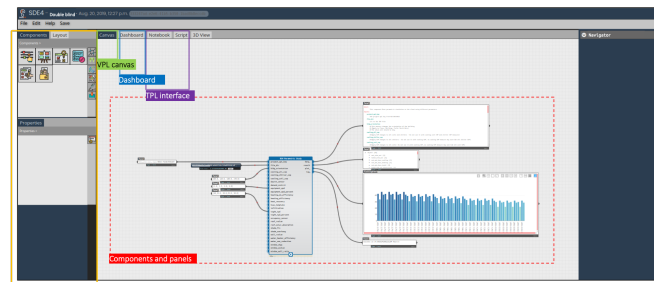

Figure 5. The main interface functions – a higher resolution image can be viewed here.

```
7
8  def preprocessInputs(input1):
9      '''This function is applied to the input
        before running the component'''
10     return input1
11
12 def postprocessOutputs(output1):
13     '''This function is applied to the output
        after running the component'''
14     return json.parse(output1)
15
16 inp1 = component.before(preprocessInputs,
        component.inputs[0])
17 out1 = component.after(postprocessOutputs,
        component.outputs[0])
```

Listing 1. Triggering a VPL component using TPL Python interface

4.2 Dashboard and forms

The dashboard and forms are located in the same tab called 'Dashboard'. It is an information management tool directed to stakeholders who are not involved directly in the development of the project i.e., to track performance, metrics, and
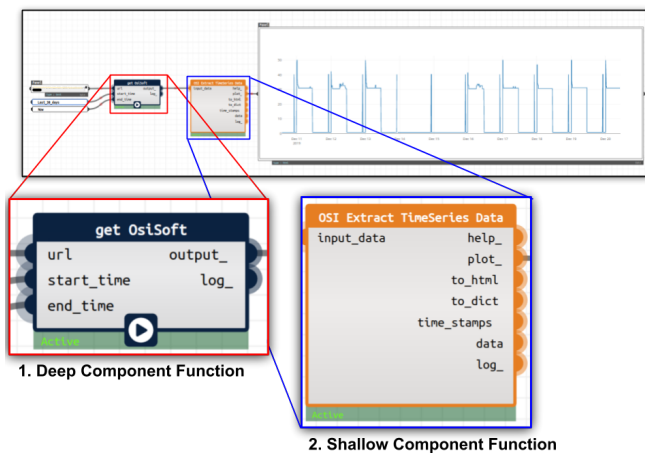
442

Figure 6. Two examples of Shallow and Deep functions. 1. The deep function is Python-based, and hosted on the application layer, the "play button" on the bottom of the component must be clicked to run the deep function. 2. The shallow function is JavaScript based and is hosted on the front-end. However, all the shallow functions are stored in the data layer and follow specific schema. The shallow functions are used to implement real-time actions, such as arithmetic operations, and JavaScript Object Notation (JSON) parsing. In this figure, the deep function on the left uses OSISoft Python API to request time-series data from the data layer and outputs it as a JSON object. Then, the shallow component on the right parses the JSON object and converts it into different formats, including plotting.
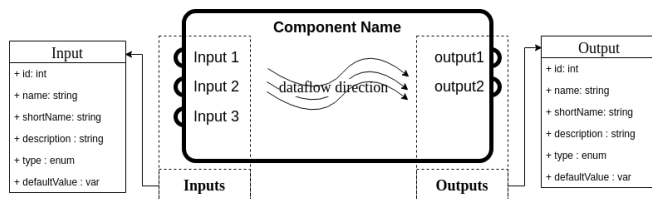
Figure 7. The data flows within components from left to the right. Each input of the component either carries a default value or receive value from other component's output, this value is stored in the input until the component runs, then the input variables are processed by calling the corresponding function. After the function is called, the return output of the function is stored back into the output of the component unless this output is connected to other component's input, it runs the other component automatically.

other key data points; or to get feedback and other user inputs (questionnaires or thermal comfort feedback). The project owner does the design of the dashboard, using the input components (e.g., numeric slider, optionList, listView, RadioButtons, and Panels) and the output components (e.g., Panels, Plots, 3D views, 2D plans, images, videos etc.). Then, the dashboard could be shared with public or specific people to view or interact. It can also be embedded within other websites.

5   APPLICATION LAYER

The application layer receives the deep function requests from the presentation layer in the form of a JSON object following specific schema. On the one hand, If there are data-related processes, it sends a request to the data layer (explained in detail in section 6. On the other hand, if there are no database-related processes, the JSON object consisted of the inputs and the component function callback. The function then starts to operate in three cases:
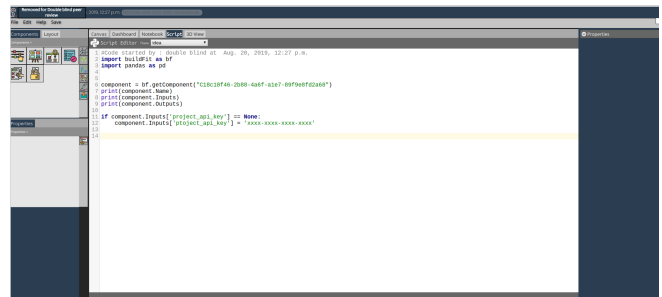

Figure 8. TPL interface using Python script. The python script enables reading, writing inputs and outputs of different components, as well as adding pre/post processing functions to the inputs/outputs – a higher resolution image can be viewed here

1. It runs the function directly from the node.

2. It starts other cloud engines to run the functions.

3. It uses third party engines (e.g. BuildSimHub) to perform the operation.

Consequently, the function response back to the presentation layer as a JSON object with the outputs (if any), Errors, and logs. The response JSON object also follows a predefined schema as shown in figure 7. The whole process is illustrated in figure 3 (steps: 1, 2, 3, 8, 9).

6   DATA LAYER

There are four major types of data involved in this platform: metadata of the users and projects, metadata of the buildings, building models, and building operational data. The data layer has three main functions: **Data querying**: to acquire data from multiple data sources, **Data storage**: to store different types of data, and **Data feed**: to respond to the requests from the application layer.

The metadata and models are either input by the users through the presentation layer or assigned by the application layer. These data are mainly text. Once passed to the data layer, they are static and stored in the relational databases (Figure 9). As for the building operational data, the time series data comes from servers of different Building Management Systems. BMS in different buildings are various, in terms of data structure, sampling rate, communication protocol, etc., making data querying a troublesome task. The platform dealt with this by deploying the PI system from OSIsoft, which queries data from different types of servers, such as BACnet (Building Automation and Control networks) and OPC (Open Platform Communications), and stores the compressed data on the local server.

Data exchange between the application layer and the data layer is done through RESTful API and mainly in JSON format. For example, if a user wants to see the energy consumption trend of a building, he/she will select the data point and define the time period in VPL canvas. The deep functions in the application layer will get the information and accordingly send the request. The API will retrieve data from the server and send it back as a JSON file, which will then be plotted in the canvas.
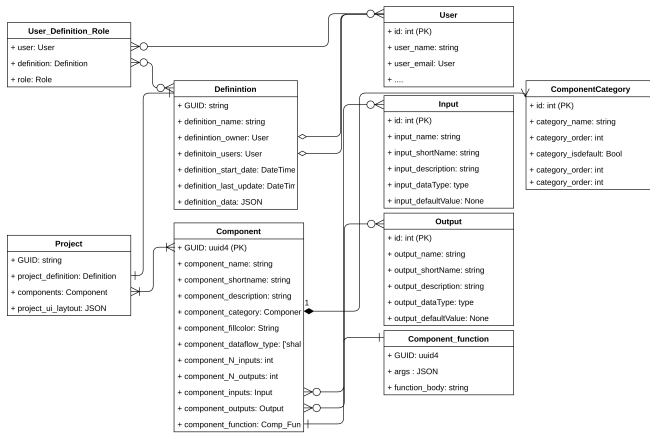
Figure 9. Relational database: A higher resolution image can be downloaded here.

7 USE CASE SCENARIO

In this section, we introduce a use case based on Cockburn's template [12] to illustrate how the platform works. The use case is a parametric energy simulation of a small office building from ANSI/ASHRAE/IES Standard 90.1 commercial reference buildings [15] shown in figure 10.
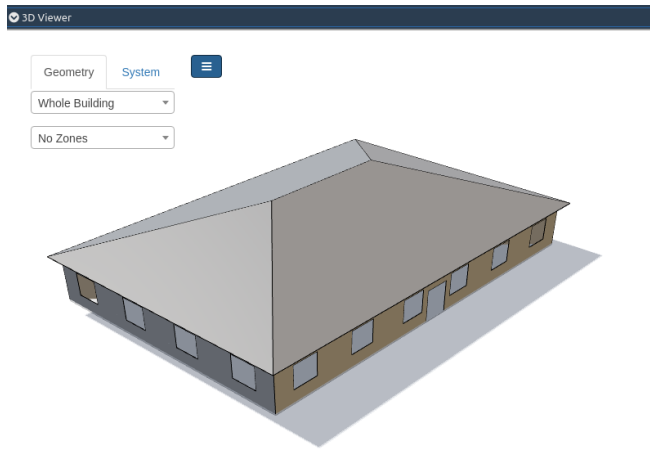

Figure 10. One of the reference buildings developed by the U.S. Department of Energy (DOE). The building represents a small office with 5,500 square feet area and one floor. The image is a screenshot taken from the embedded 3d viewer in our platform, and provided by BuildSimHub (in the application layer). an animated GIF of the 3d-viewer could be viewed here

USE CASE:1 Conduct a Parametric Energy Simulation
- -
CHARACTERISTIC INFORMATION
Context of use: EnergyModeller conducts a parametric building energy simulation of a small office building using the three-tier architecture platform.
Scope: Platform
Level: Summary
Preconditions: The user must have a valid authentication to the platform and a supported web browser (e.g. Google Chrome).
Actor(s): EnergyModeller, BuildSimHub (cloud simulation engine).

Trigger: Firstly, the EnergyModeller should start a new project. Secondly, IDF file of the EnergyPlus model of version greater than 8.0.

Description:

1. EnergyModeller starts a new definition – (Tier1 - presentation layer).

2. In the presentation layer, the EnergyModeller uploads the IDF file to the data layer (Tier 3) which is a scalable google cloud bucket using a file upload component (Figure 11). This step takes place through the application layer (Tier2) which contains a Google Cloud Storage application programming interface (API) and other file validation and security checks – an animated GIF image could be viewed here.

3. The IDF file link and global unique id (GUID) are retrieved.

4. EnergyModeller loads a "buildSimHub Parametric Study" component (Figure 12).

5. After running the component, a request is sent to the simulation engine in the application layer including the inputs (IDF file, project api key, and the simulation parameters).

6. The output of the simulation comes in a form of JavaScript Object Notation (JSON) as well as a chart (Figure 13) – a full view of the project can be downloaded here.
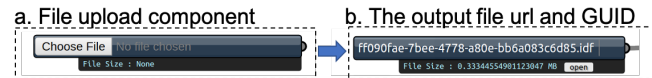

Figure 11. File upload component in the presentation layer: a) The user is prompted to select a file. The file is uploaded to a google cloud bucket; b) Then, the file could be triggered by its global unique id

8 CONCLUSION

In this paper, we explained an approach for applying Client/Server based architecture called 3-tier architecture (summarized in figure 3). This approach is used for managing data from the built environment using a cloud-based user-friendly visual programming interface. This approach depends on separating the client-side (called the presentation layer) from the back-end engines (the application layer) and the databases (the data layer). This separation eased the distribution of computation power over many nodes without compromising the efficiency of the other layers. Furthermore, it overcame some problems related to latency and connection speed. Moreover, flexibility and scalability are two key features of this type of architecture.

Currently, the platform still in the final stages of developing and debugging the alpha release. The platform will be released as an open source and contributions are welcome from the community. Contribution instructions will be available on the following GitHub repository:

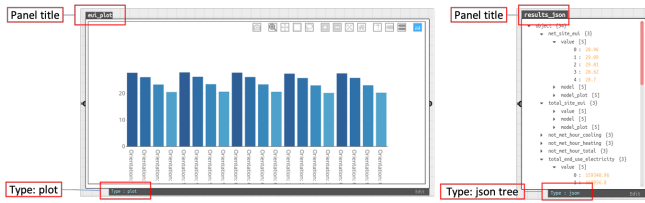https://github.com/ideas-lab-nus/paper-SIMAUD2020-three-tier-architecture-platform

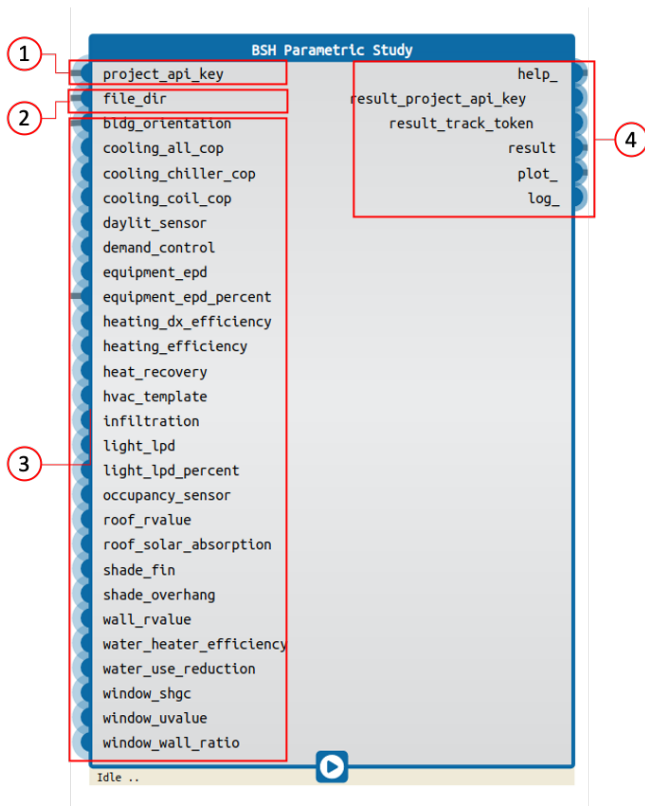**Figure 13.** Display results as a chart and/or JSON tree in the presentation layer.

Figure 12. Parametric simulation study: this component runs a cloud simulation engine in the application layer. The inputs are: (1) The project_api_key, i.e. a unique id for each simulation project, and is used to track the project and run different simulation models' progress; (2) file_dir: which is the IDF file stored in the data layer (tier 3) in the form of a flat file; (3) The simulation parameters. After running the component, the outputs (4) come in the form of JavaScript Object Notation (JSON) object and a plot which can be displayed in a panel (Figure 13) and embedded in a website.

Future development includes enabling real-time teamwork, version controls, besides continuous testing and user-experience improvements.

REFERENCES
1. AbdelRahman, M. GH_CPython: CPython plugin for grasshopper, 2017.

2. Abdelrahman, M. M. Enhancing Computational Design with Python high performance scientific libraries : Integration of Grasshopper and CPython language. 2–3.

3. Abdelrahman, M. M., and Toutou, A. M. Y. [ANT]: A Machine Learning Approach for Building Performance Simulation: Methods and Development. *The Academic Research Community publication 3*, 1 (2019), 205.

4. Abdelrahman, M. M., Zhan, S., and Chong, A. Building Life-Cycle Usability Data Segmentation: A NLP-based review. *Unpublished work* (2020).

5. Allcott, H., and Mullainathan, S. Behavior and energy policy, 3 2010.

6. Bachman, D. *Grasshopper: Visual Scripting for Rhinoceros 3D*. 2017.

7. Bhardwaj, S., Jain, L., and Jain, S. Cloud Computing : a Study of Infrastructure As a Service (Iaas). *International Journal of Engineering 2*, 1 (2010), 60–63.

8. Booch, G., Maksimchuk, R. A., Engle, M. W., Young, B. J., Connallen, J., and Houston, K. A. Object-oriented analysis and design with applications, third edition. *ACM SIGSOFT Software Engineering Notes 33*, 5 (2008), 29.

9. Bragdon, A., Zeleznik, R., Reiss, S. P., Karumuri, S., Cheung, W., Kaplan, J., Coleman, C., Adeputra, F., and Laviola, J. J. Code bubbles: A working set-based interface for code understanding and maintenance. In *Conference on Human Factors in Computing Systems - Proceedings*, vol. 4 (2010), 2503–2512.

10. Brohus, H. ., Frier, C. ., Heiselberg, P., and Haghighat, F. Quantification of Uncertainty in Predicting Building Energy Consumption: a stochastic approach. *Energy and Buildings 55* (2012), 127–140.

11. Chong, A., Xu, W., and Lam, K. P. Uncertainty analysis in building energy simulation: A practical approach. *14th International Conference of IBPSA - Building Simulation 2015, BS 2015, Conference Proceedings* (2015), 2796–2803.

12. Cockburn, A., and Cockburn, A. Use Case Template Basic Use Case Template. Tech. rep., 1998.

13. De Line, R., Czerwinski, M., Meyers, B., Venolia, G., Drucker, S., and Robertson, G. Code Thumbnails: Using spatial memory to navigate source code. In *Proceedings - IEEE Symposium on Visual Languages and Human-Centric Computing, VL/HCC 2006* (2006), 11–18.

14. Dillon, T., Wu, C., and Chang, E. Cloud Computing: Issues and Challenges.

15. Field, K., Deru, M., and Studer, D. Using DOE commercial reference buildings for simulation studies.

16. Furht, B., and Escalante, A. Handbook of Cloud Computing. Tech. rep., 2010.

17. Green, T. R. G. Instructions and descriptions. In *Proceedings of the working conference on Advanced visual interfaces* (2004), 21–28.

18. Green, T. R. G., Petre, M., and Bellamy, R. K. E. Comprehensibility of Visual and Textual Programs: A Test of Superlativism Against the 'Match-Mismatch' Conjecture. *Proceedings of the Fourth Annual Workshop on Empirical Studies of Programmers*, January (1991), 121–146.

19. Green Building XML (gbXML) Schema. gbXML Green Building - Current Schema, 2019.

20. Iea. Tracking Clean Energy Progress 2016: IEA Input to the Clean Energy Ministerial. *International Energy Agency (IEA) Directorat* (2013), 148.

21. Jakubiec, J. A., and Reinhart, C. F. DIVA 2.0: Integrating daylight and thermal simulations using rhinoceros 3D, DAYSIM and EnergyPlus. In *Proceedings of Building Simulation 2011: 12th Conference of International Building Performance Simulation Association* (2011), 2202–2209.

22. Koltsova, A., Zurich, E., Schmitt, G., Schneider1, C., Koltsova2, A., Schmitt3, G., and Strasse, W. P. Components for parametric urban design in Grasshopper from street network to building geometry. Components for Parametric Urban Design in Grasshopper. From Street Network to Building Geometry. Tech. rep., 2011.

23. Lagios, K., Niemasz, J., and Reinhart F, C. Animated Building Performance Simulation (Abps) – Linking Rhinoceros / Grasshopper With Radiance / Daysim. *conference proceedings of SimBuild 2010*, August (2010), 7.

24. Lieberman, H., Paternò, F., Klann, M., and Wulf, V. End-User Development: An Emerging Paradigm. In *End User Development*. Springer Netherlands, 10 2006, 1–8.

25. Lieberman, H., Paternò, F., and Wulf, V. End-user development. Tech. rep.

26. Mathworks. MATLAB - Mathworks - MATLAB & Simulink, 2016.

27. McNeel, R. Grasshopper generative modeling for Rhino. *Computer software (2011b), http://www. grasshopper3d. com* (2010).

28. Mell, P., and Grance, T. The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology. Tech. rep., 2011.

29. MENZIES, T. EVALUATION ISSUES FOR VISUAL PROGRAMMING LANGUAGES. 5 2002, 93–101.

30. Moher, T. G., Mak, D. K. H., Blumenthal, B. B., and Levanthal, L. M. Comparing the comprehensibility of textual and graphical programs, 1993.

31. Molina-Solana, M., Ros, M., Ruiz, M. D., Gómez-Romero, J., and Martin-Bautista, M. J. Data science for building energy management: A review, 2017.

32. Nardi, B. *A small matter of programming: perspectives on end user computing*. 1993.

33. Peronato, G., Kämpf, J. H., Rey, E., and Andersen, M. Integrating urban energy simulation in a parametric environment: a Grasshopper interface for CitySim. Tech. rep.

34. Ramaswamy, S., and Tripathi, R. Internet of Things (IoT): A Literature Review. *Journal of Computer and Communications 3* (2015), 164–173.

35. Reen, T. R. G. G., and Etre, M. P. Usability Analysis of Visual Programming Environments : A 'Cognitive Dimensions' Framework. Tech. rep., 1996.

36. Reviews, N. S. R., Energy, S., and 2012, u. Stochastic techniques used for optimization in solar systems: A review. *Elsevier*.

37. Rothermel, G., Shaw, M., and Wiedenbeck, S. The state of the art in end-user software engineering. *ACM Comput. Surv 43* (2011), 44.

38. Roudsari, M., Pak, M., Lyon, A. S. I. c. h. i., France, u., and 2013, u. Ladybug: a parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. *ibpsa.org*.

39. Scaffidi, C., Shaw, M., and Myers, B. Estimating the Numbers of End Users and End User Programmers. In *2005 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC'05)*, 207–214.

40. Schuldt, H. Multi-tier Architecture. In *Encyclopedia of Database Systems*. Springer New York, 2017, 1–3.

41. Tardioli, G., Kerrigan, R., Oates, M., James, O., Procedia, D. F. E., and 2015, u. Data driven approaches for prediction of building energy consumption at urban level. *Elsevier*.

42. Taylor, P. The Three-Tier Architecture: An Object-Oriented Perspective. 1998, 79–94.

43. US Department of Energy. Technology for Building Systems Integration and Optimization – Landscape Report.

44. Wijegunaratne, I., and Fernandez, G. *Distributed Applications Engineering*. Practitioner Series. Springer London, London, 1998.

45. Wong, J., Li, H., construction, S. W. A. i., and 2005, u. Intelligent building research: a review. *Elsevier*.

46. Zhang, Z., Chong, A., Pan, Y., Zhang, C., and Lam, K. P. Whole building energy model for HVAC optimal control: A practical framework based on deep reinforcement learning. *Energy and Buildings 199* (2019), 472–490.

47. Zibion, D., Singh, D., Braun, M., and Yalcinkaya, D. Development of a BIM-enabled Software Tool for Facility Management using In-teractive Floor Plans, Graph-based Data Management and Granular Information.

48. Zou, P. X., Xu, X., Sanjayan, J., and Wang, J. Review of 10 years research on building energy performance gap: Life-cycle and stakeholder perspectives, 11 2018.

Data-driven Design Based on the Outdoor Thermal Comfort

Navid Hatefnia[1], Marjan Ghobad[2] and Paul Carew[3]

[1]TUM Ph.D. Candidate
Berlin, Germany
Navid.Hatefnia@tum.de

[2]Independent Researcher
Berlin, Germany
Marjan.Ghobad@gmail.com

[3]Associate Director
Berlin, Germany
Paul.Carew@burohappold

ABSTRACT

In recent years the subject of the outdoor thermal comfort analysis has been discussed extensively as an essential part of activating the urban areas. The significance of this subject relates to the efficiency of the analytical method to extract meaningful information from the thermal comfort simulations. The outcome of the analysis should have the potential to establish a series of design strategies base on them, especially to construct the design basis in the initial stages. This paper proposes a methodology to analyze the derived results from the outdoor thermal comfort simulations. This methodology can reveal the hidden pattern of the effectivity of the environmental parameters in order to thermally improve the exterior spaces. Consequently, it provides efficient design strategies through its potential to prioritize the environmental parameters based on their beneficence.

Author Keywords

Data mining; Thermal Comfort; MRT; Outdoor Comfort; Data Gathering.

ACM Classification Keywords

I.5 PATTERN RECOGNITION

1 INTRODUCTION

This paper proposes a methodology to improve and guide the design of the open urban spaces with consideration of the thermal comfort in the initial stages. Since understanding the thermal comfort deals with many different parameters, the proposed method shows how to extract meaningful data from the executed analyses through data mining. It indicates how the derived data could create a guideline for the design of the open urban spaces to thermally improve the comfort level. It suggests how to prioritize the effective environmental parameters in a site of interest. Also, it proposes how to classify the site based on the efficiency of the parameters. Moreover, it suggests what strategies would be more beneficial for each classified area.

In recent decades, the approach to design the open urban spaces have improved by developments in the fields of the computer simulations. This process of design relies on running the abundant number of simulations to consider thermal comfort, especially in the open-air spaces. The

evaluation of the thermal comfort through simulation is a very time-consuming process, as it deals with multi parameters. Commonly, the analysis of the outdoor thermal comfort starts with building the 3D model of the context geometry within the computer software. After assigning the proper material into the surrounding environment of the context, simulations run. After each simulation, the model updates based on the results. The continuous operation of the simulations creates multi choices for decision making in the design. Then the comparison between all the executed simulations leads to the development of the design. Besides the time-consuming problem, the main disadvantage of this approach is to gain only one conclusive number that indicates the comfort level. This number is commonly achieved by placement of the associated amount of each variable in the UTCI (Universal Thermal Climate Index) formula. Subsequently, the individual evaluation of each parameter would not be possible because the assigned value to the comfort level is the sum of the variables in the UTCI equation.

On the contrary, the presented method in this paper has many potentials. First of all, it has the potential to analyze the impact of each effective parameter separately. Since this method is the developed version of the earlier published paper by the authors, [1] the collected data relating to the UTCI variables have been stored in separate matrices. These matrices create a database that has the potential to track the emerging changes in the site by changing the assigned value to each parameter after making the predictive model. Secondly, the created database allows prioritizing the parameters based on their level of effectiveness. Thirdly, it is beneficial to be evaluated for a specific period. Lastly, it should be emphasized that the simulation is not enough by itself to improve the design. Therefore, the collected data derived from the analyses should have the potential to be analyzed in order to extract meaningful information. The collected database in the presented method provides the opportunity to comprehensively analyze all the involved parameters.

Moreover, it allows applying the data mining methods to visualize the hidden patterns behind them to create the design guidelines directly on the 3D model. The classification of the simulation results in the separate matrices allows post-

processing data mining. Then this data mining facilitates a better understanding of the effective parameters that the simulation, by itself, is not capable of it. Also, the execution of such a process with the available software is an arduous task. Because the calculation of the parameters proceeds in a black box, and then the results would be presented in absolute numbers. However, to analyze the different parameters in detail, it is required to have access to all the calculations, the intermediate variables, and the computational methodologies to develop a predictive model for the post-processing step.

2 METHODOLOGY

In this section, it is tried to briefly explain all the stages of the method. Then, the analytical process has been clarified with details in the sub-sections.

The process starts with collecting Typical Meteorological Year (TMY) data from the closest weather station to the studied site. [2,3] The collected data should be analyzed and validated accurately [1]. Thus, the recorded TMY data in the weather station should be aligned with the urban morphology, urban geometry, and surface materials of the site. In the two presented examples of this paper, the Urban Weather Generator (UWG) has been used for this purpose [4]. The UWG alters the TMY files of the airport data based on the specifications of the site location into Actual Meteorological Year (AMY) data. To do so, the typical urban canyon of the context is evaluated in the simulation of the energy balance of the heat fluxes from it. Also, the effect of the surrounding building geometries and materials, program, traffic, vegetation, and ambient heat transfer are assessed. [4] Afterward, the annual analysis of the outdoor thermal comfort has been executed on two different scales. First, the 'Open-Field' analysis, which focuses on the essential principles and second, the 'Contextual Simulation,' which deals with specified features of the particular site.

To carry out this process, the related data from all the effective parameters for the initial climate analysis and then for all the plotted points on the selected site should be calculated and recorded. In a way that for every hour in a year, the UTCI parameters; Air Temperature (Ta), Relative Humidity (RH), Air Velocity (Va) and Mean Radiant Temperature (MRT) and the parameters that lead to evaluate MRT [5]; Atmospheric long-wave radiation, Surrounding Surface Longwave Radiation, and total Solar radiation should be assessed and sorted. After running the simulation, the derived results are classified and recorded in the separate data matrices. Then, post-processing and data mining of the derived results from the simulation in both scales provides the opportunity to visualize the hidden patterns of them. The visualizing of the hidden patterns determines the probability of occurring each condition. Thereafter, it highlights the most effective parameters and the feasibility of applying the different strategies before any further simulations. This process of data mining proceeds to answer the "What-If" scenarios on each scale.

It should be taken into consideration that for the outdoor spaces, it is not enough to only run some simulations for a few specific hours or extreme conditions and generalize the results to the whole year. In the thermal comfort analysis of the exterior spaces, the probability of the occurring conditions in each specific spot defines what strategy should be prioritized for that spot, which requires a comprehensive simulation. The probability for the creation of the different conditions should be considered for the specified periods, and it should not be generalized for all the hours in a year. The examined domain may be considered annually, or it may be limited to the specific periods that deal with more users. For instance, in the studied examples of this paper, which are located in the middle-east, it is expected that the majority of users would be in the open urban spaces after the sunset. In this regard, it can be assumed that the urban circadian rhythm defines which periods should be considered to prioritize the design concerns.

Since the process of the proposed method has been applied to evaluate the two cases in Abu Dhabi, the applied analyses and the provided figures are related to the Abu Dhabi context to describe the method. However, it should be taken into account that the method can be applied for diverse contexts. Since the main aim of the method is to highlight the most effective parameters on the comfort level, it has the potential to be used for different contexts. Clearly, the results would indicate the different parameters as the effective ones, and consequently, the design process would follow the different "What-If" scenarios.

2.1 Open-Field Analysis

The main goal of the Open-Field analysis is to assess the outdoor thermal comfort in the context that is limited to the essential principles. Therefore, this analysis is proposed with the evaluation of the comfort condition of a person who is located on a ground surface without any surrounding surfaces. Therefore, the ground is the only geometrical surface that has a direct impact on the person's body. In this rudimentary condition, the human body is exposed to Solar Radiation and the ground surface. Thus, the ground surface bonuses a certain amount of radiation to the human body based on the Sun's angle. In addition, even during the night, with consideration of the Sky View Factor (SVF) (which is equal to one in this case), the human body is in a long-wave heat exchange with the sky and the ground. Since the Open-Field analysis deals with the essential principles, it can show the direct impact of the environmental parameters in the simplified context. Therefore, this analysis provides a better understanding of the environmental potentials such as night sky cooling and parameters like daily solar radiation and the ground cover material on the thermal comfort condition.

At this stage, the simulation for the assessment of the thermal comfort should be executed. Based on the UTCI criteria [6,7], the assigned value to each of the individual parameter, Ta, RH, Va, MRT [5,8] and the incorporated parameters of the MRT; Atmospheric Radiation, Solar Radiation, Infrared

Radiation and Reflected Short-wave Radiation from the ground should be recorded in the separate matrices [1]. For further information, the method to create these matrices have been explained in detail in the earlier work of the authors, "*Analytical method to assess outdoor comfort based on universal thermal climate index (UTCI).*" [1] The formation of such matrices provides the database from the Open-Field Analysis. This database has the possibility to change the assigned value to the initial parameters and simultaneously compare the derived results with the initial outcomes from the predictive model. An initial predictive model has to be made in order to assess the probability of occurring similar situations by changing each variable that is likely to influence the outcomes. Changing variables individually in the data series might lead to the creation of the new series that have not been existed in the data. Hence, this new condition cannot be accepted. It is important to consider the variables as a cluster and assess the probability of occurring each one.

In general, it should be taken into account that before addressing the details in the design, it is required to identify the effective strategies regarding the improvement of the thermal comfort conditions in the outdoor spaces. Since the gathered database from the Open-Field simulation has all the required data in a classified way in itself, it provides the potential to develop a predictive model to change the input variables and, consequently, observe the outcome instantly. The assumptions of changing the variables present the extremums from the performance of the different design strategies. These extremums seem to be undoable in reality. However, it is required to first analyze the level of the effectiveness of a strategy and assess the feasibility with the highest performance. If the result shows a high rate of improvement, then the analyzed strategy requires more investigation. Otherwise, from the beginning, the analyzed strategy would be eliminated from the design priority concerns. Subsequently, there is no need to run several time-consuming simulations with lots of details.

In order to clarify the priorities of the effective parameters, the different parameters should be analyzed without consideration of the design strategies at this stage. To do so, the four primary parameters, based on the UTCI criteria [6,7], Ta, RH, Va, MRT should be evaluated separately in order to elucidate the impact of the applied changes in each of them on the thermal comfort level. To implement this process, the easiest way is to apply a change into one of the parameters by neglecting the design strategies and observe what would happen on the desired point. For instance, what would happen by reduction of 3° Celsius in the air temperature? Or what if to have 3 m/s wind velocity at all times? Or what if the ground cover would have a material feature to reflect more and absorb less heat?

The applying of the changes individually in each of the parameters and the comparison of them with the derived outcomes from the Open-Field analysis leads to define the

priorities and to recognize the most effective parameters. Furthermore, it clarifies that the change of a parameter how much and in what time of the year would be effective. One example that can indicate how the changes in one of the parameters could be a practical approach to specify the appropriate design strategies is to provide a canopy. To better understand this design strategy, the shade, the assumption is to replace the calculated MRT with the Air Temperature. The produced data from this replacement illustrates the imaginary condition from the ideal design for the canopy. The canopy that could have reflected the solar radiation to the sky without any absorption and its lower layers could have had the temperature equal to the ambient temperature. Clearly, the design of such canopy has poor performance and inferior functionality in reality. However, in this example, it has been tried to use the collected data from the database in order to achieve the rapid and straightforward perception of the environmental conditions and to observe the role of the design strategies instead of running another simulation. This example highlights the hidden pattern of the pros and cons of using shading strategy. Figure 1 illustrates the impact rate of the canopy placement on the UTCI level. As this figure shows, using the shade is not always beneficial, and in the absence of solar radiation, it does have a negative impact since it mitigates the sky cooling effect for the underneath area.
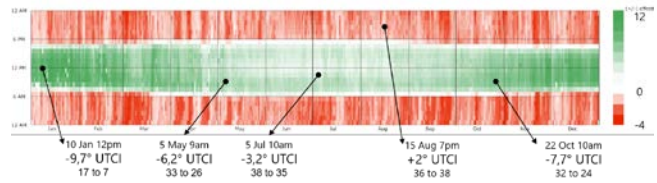

Figure 1. The pros and cons pattern of using a canopy.

After individually recognition and analysis of each parameter, the changes in the combination of the parameters play an essential role in determining the design strategies. It is mainly because the combining of the strategy, such as evaporative cooling, can create more effective changes in the shade rather than applying it alone. Therefore, the next stage is to define the "What-If" scenarios to better comprehend the functionality of the different strategies in the design. The example that can show the effectivity of applying changes in the combination of the parameters is to create an imaginary situation with the existence of the shade and to have 4 m/s wind velocity at all times on the selected point. The design strategies can be extended with similar assumptions, and the different combinations of the changes in the parameters can be analyzed in a basic approach. For instance, the combination to design a canopy with providing the evaporative cooling can determine the performance of this combination in the extreme climate of the context (Figure 2). While there are considerable differences in the UTCI level in the daytime, the canopy needs to be folded after sunset. For instance, the providing of the shade on 22 October at 10 a.m. decreases the UTCI level by 6 degrees while the Air

449

Temperature is 28°C and the Relative Humidity is 45% with 4 m/s airflow.

The pattern of the assessment result can be visualized based on the UTCI improvement or comfort level condition, and these two should be considered separately since improving the UTCI condition does not necessarily lead to thermal comfort conditions.
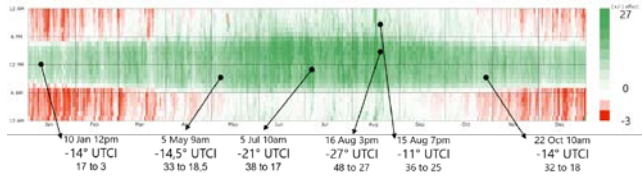

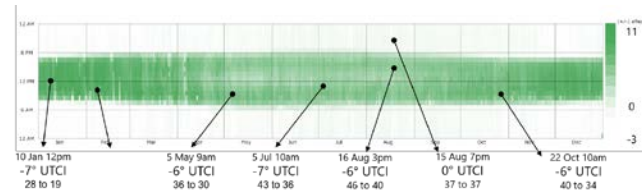**Figure 2.** The pros and cons pattern of the combination of shading with evaporative cooling.

Figure 3. The advantages of providing a moveable shade in Abu Dhabi during the whole year.

2.2 Contextual Simulation

While the Open-Field analysis deals with analyzing only one point in the open space without the presence of any surrounding urban surfaces, on the contrary, the Contextual Simulation considers detailed features of the circumambient surfaces of the site and provides the distributed points within site. [1] The accurately scaled model of the site requires to be built in computer software via including the details related to the geometry, material properties that indicate the emissivity and albedo of the surrounding environment. In addition, the model should be built in a way to be used for further analysis, such as CFD (Computational Fluid Dynamic) and MRT simulations. In this regard, the assigned values of the environmental parameters as the initial parameters are collected directly from the AMY. (In the previous stage, the collected TMY data has been converted by UWG into AMY data). At this point, a comprehensive simulation is executed to assess the outdoor thermal comfort based on all the considered details. At the same time, the simulation results are recorded in the separate matrices for each point in the site for every hour throughout the year, which creates a database. As a result, this process not only records the final result of the thermal comfort level based on the UTCI criteria [6] for each point in every hour but also it stores the details related to the assessed variables like MRT [8] that is formed by environmental aspects and geometries on the human body.

Recordings of these variables provide the opportunity to analyze the role of each parameter individually in the next stages, and also, it gives the potential to visualize the effect of the different design strategies directly on the studied site.

As an example, before considering any solution to improve the design, it is possible to provide a better understanding from a parameter like the air temperature and its effect on the site, only by the reduction of the air temperature by 3°C and observes its effect directly on the site. It is clear that because of the different morphologies and variable geometries for each point in the site, the level of the effectivity from the temperature reduction would relatively differentiate for the various points. The unequal situation in this example indicates which part of the site has more potential to use the air temperature reduction strategy and which parts have less potential to use this strategy.

In this regard, the analysis of the priorities to employ the different design strategies in the various points in the studied site based on the effectivity of the different parameters leads to an efficient design with the maximum impact. This process can be done for each parameter individually or for the combination of the parameters in a way to prioritize the strategies such as usage or blockage of the airflow, canopy, misting, cooling surfaces, Passive Downdraft Evaporative Cooling (PDEC), different albedo and emissivity of the surfaces. (Figure 4)
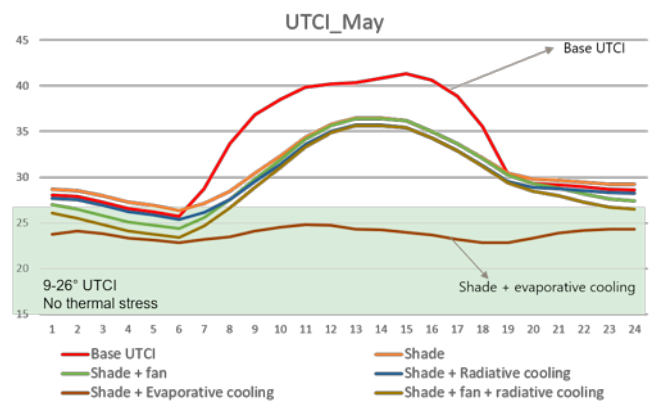

Figure 4. The comparison of the various cooling strategies highlights the efficiency of each strategy for the Abu Dhabi climate in May.

3   EXAMPLES

In the following section, two examples from several ongoing projects in Abu Dhabi have been presented. Abu Dhabi is a coastal city with an extreme climate by high daily temperature, humidity, and high solar radiation. Usually, there is a gentle wind from the South and South-East in the mornings and the North and North-West in the evenings from the sea. The very-strong heat stress and extreme heat stress happen between April and October, which cause difficulty in the use of the outdoor spaces due to uncomfortable thermal condition. This extreme condition would be relatively harsh in the urban context. This is not only because of the absorbed heat from the urban surfaces, but also it is due to the shortcoming to emit the absorbed long-wave radiation during the night to the sky regards to the high levels of humidity and the dense urban geometry. However, the night sky cooling can be highlighted as one of

450

the limited potentials for the cooling purpose in Abu Dhabi. Besides, the produced heat from the mechanical air-conditioning systems should be added to this critical condition. Also, air pollution and dust level should be considered in the design.

All these concerns have caused the urban designers and planners to consider ameliorating the design of the outdoor public spaces, especially with regards to improvement of the thermal comfort condition to create active and livable open spaces in the neighborhood areas. To do so, many different areas have been planned as urban plazas or the urban parks in different scales. In the following section, the process of the presented methodology has been stated in detail in two examples of the pocket parks.

The analysis process begins with the collection of the TMY data. The raw data from the Bateen airport weather station has been used in the presented projects from the two existing sources to collect Abu Dhabi weather data. The collected raw data should correspond with the urban morphology of each site. UWG has been used [4] to convert TMY data into the morphed weather file that captures the Urban Heat Island effect. Then the AMY is achieved and becomes the base to run the open-field simulation.

The Abu Dhabi Open-Field analysis shows about 35% of comfort conditions during a year. However, this includes the night-time domain, and during the day hours, this percentage is much lower. As an instance, the assessed comfort condition by the basic simulation from 6 a.m. to 10 a.m. shows just 18% of the total. Different strategies, from a simple shade to more complexed ones, the combination of the strategies, are assessed to determine the impact on the thermal comfort during a year, in the morning and the afternoon separately. Although the strategies have less impact individually, the combination of them seems to be more promising in this climate. A combination of a well-designed shade with providing evaporative cooling seems to be efficient enough to achieve 9-26°C UTCI on a daily average. The diagram in figure 2 shows a 20°C UTCI reduction for 3 p.m. on average in May.
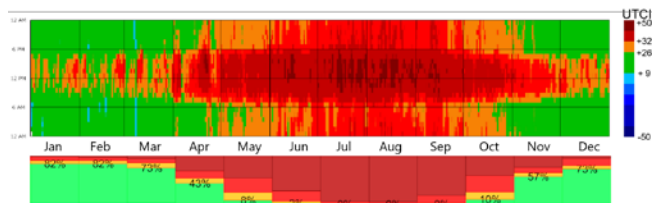

Figure 5. The illustration of the UTCI after running the Open-Field simulation for Abu Dhabi.

Figure 5 shows the UTCI results from the open-field simulation. The figure illustrates more heat stress and discomfort conditions from May to November, including the summer months and more thermal comfort conditions from November to May.

The recorded variables in the Open-Field analysis and the obtained database provide the opportunity to analyze the different conditions based on the heat stress level and to clarify the essential parameters that affect the discomfort conditions.

The understanding of the existing range in the collected data has significant importance. Providing minor changes in the existing condition can shift the discomfort into a comfortable condition. To understand the most effective change to modify the discomfort state into a comfortable condition, the data mining methods have been applied. As an instant, the Dynamic Time Warping (DTW) [9,10] method has been used to find the closest comfort state to the discomfort condition to highlight the most advantageous parameter. The probability of each of these beneficial parameters in each condition is considered in a specific period to computationally calculate the prioritizing the effective aspects in support of the decision-making process. To use this method, the collected data in the separate matrices should be compared with each other to unfold a pattern among them. To do so, the environmental parameters should be calculated for every hour of the year for each point of the site. As previously mentioned, the collected data is stored in separate matrices. The stored data should be evaluated based on the UTCI criterion, and it reveals whether they are in comfort or discomfort condition. In order to analyze the environmental parameters, the parameters that have led to the comfort condition should be compared with those that have led to discomfort. Therefore, by using the DTW method, the closest state from the comfort condition to the discomfort condition can be exposed. This comparison shows the contribution of different parameters to create discomfort status. By expansion of the DTW method to all the matrices, the level of effectiveness of the environmental variables for each point in the site and for each interested period can be measured.

Since in the presented method, it has been suggested to limit the examined period to maximize the efficiency of the comfort evaluation; the analysis of the comfort condition in the two sites have been focused on the period that the majority of people use the outdoor open spaces instead of analyzing for every hour of the day. In order to clarify the most effective portion of the time to limit the analysis process, the existing patterns that local people use the outdoor spaces should be considered. The urban daily rhythm shows that the majority of people intend to spend a more significant portion of time in the outdoor areas during the evening and the late-night, mostly from 6 p.m. until midnight. The existing pattern of the urban circadian rhythms indicates how the local people have adapted and planned their daily routine based on the outdoor comfort condition. In this regard, the analysis period for the projects has been focused on the evening hours to maximize the efficiency of the design for the majority of people.

After recording and classifying data during the computational process of the open-field simulation and analyzing the role of each parameter on the comfort level, then the design scenarios should be considered. As it has been formerly stated, the effectivity of each scenario can be assessed with the change of the collected data and the variables in the obtained database. To do so, the created predictive model form the database considers a series of assumptions for each design scenario and clarifies the probability of occurrence of similar phenomena to what exists in the database.

The design strategies can be evaluated with the derived outcomes from the Open-Field analysis. Since the sites are located in the middle-east, the examined strategies are focused more on cooling purposes such as shading, radiant cooling, evaporating cooling, and misting. One example that can show the usage of the derived results from the Open-Field analysis is to provide a canopy to see whether it can create a comfortable status or not. This is a simple example to illustrate the performance of the predictive model without the need to repeat the simulations. In this situation, it has been assumed that the provided canopy can reflect the solar radiation entirely, and the temperature of the bottom layer of the canopy is the same as the ambient temperature. Thus, this canopy has a direct impact on the absorption of radiant energy on the human body.

Moreover, as a result, it mitigates the MRT rate during the day. However, as it has been illustrated, it has a reverse impact before sunrise, and after sunset also, it prevents sky cooling. It is required to repeat that this situation is the most idealistic condition to practically understand the collected data, and this canopy would have less efficiency in its performance in reality. A similar process can be replicated for other scenarios such as evaporative cooling, airflow, radiant cooling, and even for the combination of the two or more scenarios together. The provided graph in figure 6 elucidates the impact of the different scenarios on the comfort level in the mornings, evenings, and annually.
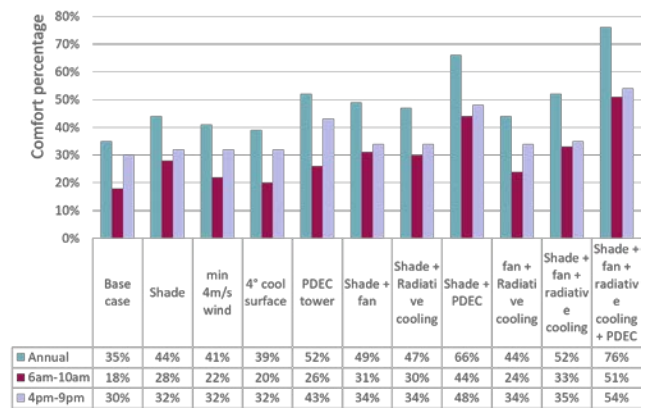
create shade. And in the absence of solar radiation, there is no need for the shade. In this regard, a movable canopy has been recommended to be used in the projects for the area which is exposed to the sky.

In the next sub-sections, two sites have been evaluated in order to visualize the application process of the method. In this process, it is explained that how to analyze the collected data, how to prioritize the role of the parameters based on their level of effectiveness, how to classify the areas of the site based on the efficiency of the environmental parameters and how to discover the practical strategies for the design.

3.1 Site One

The first site is surrounded by various buildings with different opening ratios and with variety of façade materials such as stone, aluminum, glass, and concrete. This site is oriented with a 40° angle towards the South-East and North-West. In order to improve the outdoor comfort condition, the 3D model of the site with the accurate orientation and with the precise alignment of the property material to the surrounding surfaces has been simulated annually. Some of the simulation outcomes, such as sky view factor are constant numbers that are not a function of time and are specified for this particular geometry. However, the other parameters are varied based on the weather conditions and at different times of the day, such as MRT. The higher sky view factor means more night cooling after sunset, and it is preferred to keep them as an open-air area in this climate.

The obtained results for the different points from the analyzed area can be directly visualized. As it has been shown in figure 7, the considered period of the design should be selected precisely to have better perception form the existing condition in that specific period. After overlaying of all data, it is required to develop the predictive model and apply the data mining process accurately. For example, it is not recommended to apply the 'average' function in some of the parameters like UTCI. Instead, it is better to use the probability equation to demonstrate the comfort rate in a percentage format after omitting the deviations. Since all data, including the input, output, and contingent variables data are recorded, it is possible to visualize the annual heat map of the specified parameter based on the extractable data for each point that has been previously simulated.
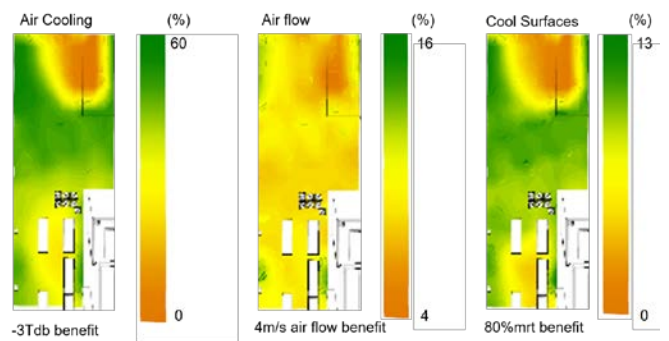

	Base case	Shade	min 4m/s wind	4" cool surface	PDEC tower	Shade + fan	Shade + Radiative cooling	Shade + PDEC	fan + Radiative cooling	Shade + fan + radiative cooling	Shade + fan + radiative cooling + PDEC
Annual	35%	44%	41%	39%	52%	49%	47%	66%	44%	52%	76%
6am-10am	18%	28%	22%	20%	26%	31%	30%	44%	24%	33%	51%
4pm-9pm	30%	32%	32%	32%	43%	34%	34%	48%	34%	35%	54%

Figure 6. The comparison of the different cooling strategies for annual, morning, and evenings.

It is clear that to construct an efficient design and to use the environmental parameters during the day, it is required to

Figure 7. The visualization of the different parameters' assessments.

Same as the open field analysis, in the contextual analysis it is possible to analyze the effectivity rate of applying changes to assess the design strategies. Figure 7 illustrates the potential of improving thermal comfort. Cooling air improves thermal comfort from 7% to 60% on different points of the area. Applying the cooling strategies are more practical in the area with the more potential percentage.

The effect of considering different strategies can be directly mapped on the site to highlight the area to consider the relative strategy. As it has been shown in figure 8, environmental parameters are assessed to find the most feasible area to apply the relevant design strategy. This figure illustrates in what area of the site it is more effective to reduce Ta, increase airflow, use sky cooling with higher SVF, utilize the low thermal mass and use less surface temperature in order to consider the beneficial design strategies.
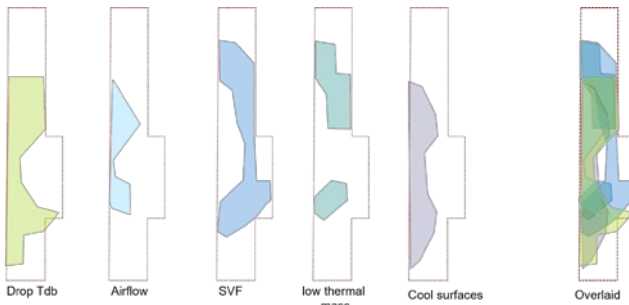
Figure 8. The illustration of the most effective parts of the site to apply the relevant strategies.

3.2 Site Two

The second site is an urban plaza that is surrounded by a series of low-rise buildings in a residential context, in front of a local mosque. Similar to the first site, the surrounding buildings have been covered with various façade materials such as stone, aluminum, glass, and concrete with the different glazing ratios. The same analytical methodology is applied for this site to evaluate the role of the environmental parameters, to determine the most effective strategies, and to indicate the relevant scenarios that should be applied for specified areas. The visualization of SVF analysis can highlight the potential of gaining benefit from the sky cooling. It has been suggested to use the movable canopy for the area that has higher SVF. Thus, this canopy has the potential to not only reduce the solar radiation during the day but also to benefit from the advantage of the sky cooling effect after sunset.

The annual outdoor thermal comfort simulation is visualized in terms of comfort percentage, heat stress percentage, and slightly heat stress percentage in figure 9 in October in the evenings from 6:00 p.m. to midnight. This visualization highlights the importance of the surfaces that are being exposed to solar radiation and the reflectance of the materials. Less heat stress is shown in front of the North faced facades, and more heat stress is displayed in front of

the reflective facades of the Mosque. The predictive model emphasizes the role of the environmental aspects of this context. Figure 10 illustrates the effect of 3°C Ta reduction, the airflow increase and decrement of the surfaces heat flux directly on the distributed points of the site in October during the evenings.
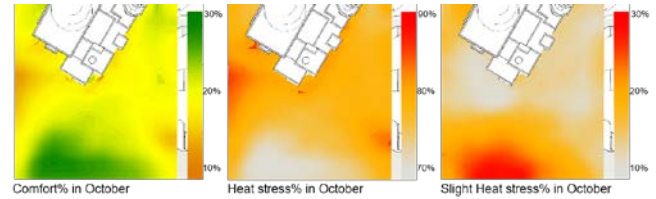

Figure 9. The thermal comfort percentage in site 2 in October evenings (18:00-24:00).
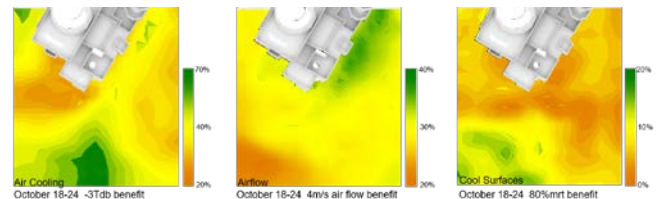

Figure 10. The effectivity of the Ta reduction, Va increase, and the surface heat flux decrease in October.

The outcome of the applied methodology not only priorities the role of each parameter but also highlights the most effective area that the practical strategy needs to be considered. The efficiency analysis of each environmental parameter forms a corresponding pattern of the site that displays in which area has the potential to be used more efficiently to enhance the outdoor comfort condition. The visualization of this analysis on all the site's points leads to the emergence of a specific hidden pattern for the site. The illustration of these patterns for all the environmental parameters leads to construct a guideline map for the entire site. As a result, this map provides a platform for the designer to define the practical design strategy in each area of the site that correlates with the derived pattern. In other words, the subsequent design decisions can be made in a way to maximize the beneficence of the created pattern of the effectivity of the environmental parameters.

4 CONCLUSION

To conclude, this paper aims to elucidate the methodology to improve the outdoor comfort condition without the necessity to execute the numerous simulations. The presented method demonstrates the simplified process to define practical design solutions. The applied process clarifies how the derived outputs from the method can prevent the multiple replications of the required simulations in order to achieve the same result. This advantage is caused by the specific method of storage and the classification of the gained data from each point of the site in a database format. The achieved database provides the opportunity to precisely understand the impact of the different environmental parameters (Ta, RH, AT, and Va) on the site. In a way that by applying the desired

changes in the environmental parameters, it shows which one has the maximum efficiency to improve comfort.

In addition, the presented method has the advantage to be applied in a specific period, which has the most beneficial impact on the improvement of the comfort situation. To do so, by clarification of the urban daily rhythms, it is possible to find the pattern that people use the outdoor spaces and focus on analyzing that period with the utmost effectiveness.

Lastly, the advantage of this methodology can be summarized by highlighting the fact that it is not only the application of one simple simulation. However, it intends to collect, classify, and analyze the environmental data in order to reveal the hidden patterns behind them that lead to finding practical design strategies. In other words, it is not required to enter into the try-and-error process after the initial simulation; because the collected database gives the opportunity to analyze the environmental parameters and to prioritize the design strategies to improve the comfort condition in each desired point.

REFERENCES
1. Hatefnia, N., & Ghobad, M. (2015). Analytical method to assess outdoor comfort based on universal thermal climate index (UTCI). *Proceedings of 31th International PLEA Conference*.

2. B.Crawley, D. (1998). TRANSACTIONS 104 Part 2 - TO-98-2-2 Which Weather Data Should You Use for Energy Simulations of Commercial Buildings? *ASHRAE Journal*.

3. Pappenberger, F., Jendritzky, G., Staiger, H., Dutra, E., Di Giuseppe, F., Richardson, D., & Cloke, H. (2014). Global forecasting of thermal health hazards: the skill of probabilistic predictions of the UTCI. *Int J Biometeorol, doi: 10.1007, 56*, 311-23.

4. Bueno, B., Norford, L., Hidalgo, J., & Pigeon, G. (2012). The urban weather generator. *Journal of Building Performance Simulation, 6(4)*, 269–81.

5. Huang, J., Guillermo Cedeño-Laurent, J., & D.Spengler, J. (2014). CityComfort+: A simulation-based method for predicting mean radiant temperature in dense urban areas. *Building and Environment, 80*, 84-95.

6. Bröde, P., Fiala, D., Blazejczyk, K., Holmér, I., Jendritzky, G., Kampmann, B., Havenith, G. (2012). Deriving the operational procedure for the Universal Thermal Climate Index (UTCI). *Int J Biometeorol, 56*, 481–94.

7. Jendritzky, G., de Dear, R., & Havenith, G. (2012). UTCI-Why another thermal index? *Int J Biometeorol, 56*, 421-8.

8. Weihs, P., Staiger, H., Tinz, B., Batchvarova, E., Rieder, H., Vuilleumier, L, Jendritzky, G. (2012). The uncertainty of UTCI due to uncertainties in the determination of radiation fluxes derived from measured and observed meteorological data. *Int JBiometeorol, 56*.

9. Sakoe, H., Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. IEEE Trans. Acoustic Speech and Signal Processing, v26, pp. 43-49.

10. Souza, C.F.S., Pantoja, C.E.P, Souza, F.C.M. Verificação de assinaturas offline utilizando Dynamic Time Warping. Proceedings of IX Brazilian Congress on Neural Networks, v1, pp. 25-28. 2009.

An Algorithm for Efficient Urban Building Energy Modeling and Simulation

Orçun Koral İşeri[1], İpek Gürsel Dino[2]

[1] Ph.D. Student, METU
Ankara, Turkey
koral.iseri@metu.edu.tr

[2] Assoc.Prof.Dr., METU
Ankara, Turkey
ipekg@metu.edu.tr

ABSTRACT

The urban population increases continuously since the industrial revolution, and the residential buildings have the primary responsibility for the total energy demand. There is a need for the analysis of the residential building stock for energy efficiency and sustainable planning. However, energy modeling and simulation in urban scale is expensive in computational complexity and time, due to various building geometries and occupancy types. This research proposes a method to increase the efficiency of the simulation process by reorganizing the building geometries with functional clustering and radiation analysis scaling. In order to accelerate the urban building energy modeling (UBEM) process, the building geometries are modified based on energy simulation standards, then, clustering is determined based on radiation analysis and outside boundary conditions. The candidates are selected according to the selection percentage that has been identified before the process to simulate in building energy software. Three different simulation types are compared to validate the performance of the proposed algorithm with complete model simulations in terms of the error rate of the objectives and the simulation runtime.

Author Keywords

Urban energy simulation; Building energy flows; Energy management; Automation; Residential building stock

ACM Classification Keywords

I.6.4 [Simulation and Modeling]: Model Validation and Analysis; I.6.5 [Simulation and Modeling]: Model Development – Modeling methodologies.

1 INTRODUCTION

Sustainable, energy-efficient solutions are a priority for cities towards decreasing carbon emissions and increasing user comfort [24]. 70% of the global CO2 emissions are attributed to the urban areas, due to their population density, high rates of economic activity, and associated energy and resource consumption. As 92% of the total population of Turkey lives in cities, and the residential urban areas are responsible for 41% of total energy demand, it is crucial to understand the energy consumption profiles of residential buildings [6,20,32]. However, energy simulation of a wide variety of residential units requires comprehensive models with many parameters, i.e., building volumes, user types, and layout. Due to close interactions with the environment of the residential building stock, the analyzed models are organized as a build-up from the neighborhood level [18,24]. Therefore, simulating an urban region is expensive in terms of time and computing power. This study proposes a bottom-up method for energy modeling of the residential building stock to address this problem. The method aims to decrease the total simulation time for new design projects or retrofit analysis of the urban building model.

1.1 Literature Review

Urban building energy models (UBEM) have the potential to support energy policy decision-making processes for cities to form effective design strategies for building sustainable urban environments [5]. City-scale building data sets are essential for UBEM, which demand different level of building properties for modeling, e.g., footprint, floor area, date of construction, space conditioning type, heating and lighting load, internal load [12]. There are two classes of modeling methods to analyze and estimate overall building stock energy performance, such as top-down, bottom-up approaches [15].

1.1.1. Top-down Approach

The top-down approach evaluates urban models while ignoring household energy demand. Generally, top-down models use collected historical energy data instead of using physical features of individual units and predict end-use energy demand of the building stock by top-level variables, e.g., energy cost, climate effect, macroeconomic indicators such as inflation, gross domestic value [13,24]. The main goal is to find a correlation between macro patterns of the past and the future.

1.1.2. Bottom-up Approach

Bottom-up models consist of the building geometry compositions that are defined as 'archetypes' [19,20]. Comprehensive models enhance the description of each building unit in terms of how the unit energy efficiency can be improved and CO_2 emissions can be reduced. These models present cost-effective options for energy demand estimations and CO_2 reduction strategies [22]. This study

adopts a bottom-up approach by processing information from the smallest unit to the top neighborhood level.

1.1.3. Residential Sector and Energy Demand

As the residential sector is one of the leading sectors for energy usage and the environmental impact, residential building stock should be analyzed in detail. However, energy usage profiles of residential types differ widely compared to industrial or commercial buildings due to the variety of building types, areas, or materials. Even, the neighborhood planning composes of a holistic planning strategy; individuals could change their units from years to years for different purposes. Different occupant behaviors and schedules complicate the metering of household energy demands [16]. The situation becomes complex in the aspect of energy demand types of residential units because there are multiple types, e.g., heating load (Q_H), cooling load (Q_C), domestic hot water (DHW), appliances, and lighting (Q_L) [25]. Each demand types is interacting with others from season to season based on user behavior and interior unit layout. Therefore, building energy modeling demands a comprehensive study to reach accurate simulation results for residential buildings.

1.2 BEM and Neighborhood Models

UBEM requires a wide range of information on buildings, such as geometric and non-geometric characteristics (constructions, appliance systems, schedules), and meteorological features of the environment [4]. For geometrical data, Geographic Information Systems (GIS) databases can supply valuable information that contains building age, user type, shape; however, when GIS data is missing, local municipality databases can present reasonable solutions [5,11]. Nevertheless, the computing cost of the model does not only increase depending on the geometry of the structure, the surface heat transfer (u-value, airtightness) and infiltration rates also contribute to the cost [9]. Therefore, there is a need to simplify the modeling process.

Building energy simulation (BES) is an informative model for building energy performance analysis in terms of presenting accurate performance indicators. Energy simulations could serve as feedback at the early design stage, in the way of comparing design alternatives, analyzing the problematic area for the evaluation of the architectural design. The method developed in this study can calculate the energy demand patterns in the neighborhood, and it can sustain valuable information for decision-making on the neighborhood level in terms of energy efficiency and sustainability to architects, planners, or policy-makers.

1.3 Neighborhood Models and Occupant Behavior

Mostly, building energy performance accuracy in simulations decreases when variances in occupant behavior is ignored as a model parameter [14]. However, the occupants have a significant influence on energy demand [26]. The lack of a realistic model that captures occupant behavior creates a demand gap due to the difference between estimated energy performance and actual energy demand

influenced by occupant behavior, e.g., daily user schedules, interaction with lighting, and appliances [10,29]. As an important feature, a methodology that can be associated with occupant types can contribute to reducing the variance between calculated and actual energy demand levels [3,23]. In this study, occupant types are generated based on government statistical data repository that represent the actual occupancy profiles, instead of using standard libraries for simulations [27].

2 MATERIALS

The study area is the Kültür neighborhood in Izmir, Turkey (Figure 1). This neighborhood mostly contains retail units on the ground level and residential units on the upper floors. Based on the GIS information, the neighborhood contains 726 residential units with approximately 76.344 m^2 floor area. Approximately, 200 buildings were eliminated as they were atypical in terms of building footprint area for residential function; consequently, 525 residential units were simulated in the study area. The threshold value of the floor area can be changed according to the district. The information on buildings' total number of floors was derived from the Turkish Statistical Institute (TUIK) [27]. On the other hand, the floor height values were not specified in the in the *Open Street Map* file (.osm). Therefore, researchers determined the height of the building and units measurements based on in-situ observations. Accordingly, the maximum height of the buildings is set to approximately 24 meters, corresponding to 6-7 floors. However, if the height information can be accessed, the model should be constructed based on these values.

Figure 1. Selected Urban Area in Izmir, Turkey – Red Border Area (2019)

The .osm urban models contain different data types, such as roads and connection points. This information is derived from the GIS, and the dataset converted from 2 dimensional to 3-dimensional model. All layout curves are transformed into four-edged convex geometry for energy simulation tool (i.e. EnergyPlus [17]), in accordance with the modeling restrictions. Except for building geometries and ground surfaces, all other elements in the urban model are ignored, such as roads, urban street elements.

Several parameters are taken from TUIK to generate the model precisely, such as occupant types by ages, the ratio of

space conditioning types [27], while others are estimated, e.g., window to wall ratio, building height, number of floors, residential unit zone division. In addition, the residential block is located in a dense urban area, which can eliminate possibilities of natural daylighting and ventilation. Objects found outside the analysis area are introduced as context geometries as shade-making elements. The context geometry is defined by the ray-casting process, which includes only visible surfaces for simulations. This method decreases total simulation runtime by using only visible elements instead of complete, intricate geometries [8].

3   METHODOLOGY

This chapter presents the method developed for bottom-up neighborhood energy and occupant comfort modeling workflow. The method consists of five steps; data input (as explained previously in Section 2), radiation analysis, model development, simulation, comparative analysis (Figure 2).
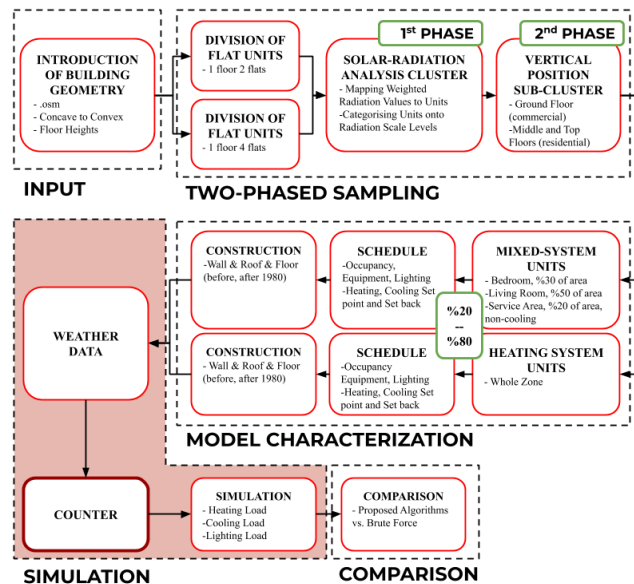
Figure 2. Flowchart of the proposed method

3.1 Two-Phased Sampling

This phase involves the statistical sampling technique that helps reduce the computing cost of simulations due to the high number of thermal zones in the urban model. A two-phase stratified sampling method is used, which aims to reduce the number of unit zones to a smaller set that is representative of the whole population by dividing the whole number of members into strata (subgroups) that are have similar thermal characteristics. Stratified sampling works in two steps. Firstly, the algorithm splits the whole data set into different groups with similar characteristic elements according to radiation analysis results, then to distribute the units under the subgroups according to the floor based clustering, which is the division based on the outside boundary conditions. Two-phase grouping leads to an increase in the possibility to select the right distribution for the samples. Finally, randomly sampling within each strata by selecting representative members according to the

selection ratio [31]. This process is based on the total solar radiance (SR) incident on envelope surfaces for each unit, which is an adaptation of Dogan and Reinhart's method [8].
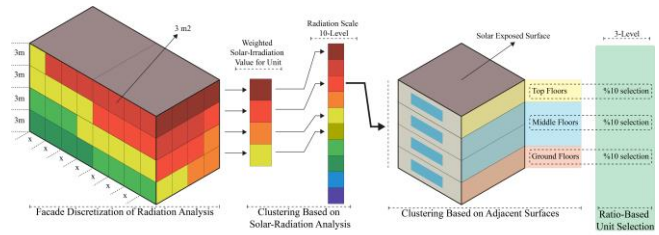

Figure 3. Solar Radiation-Based Clustering and Vertical Position Sub-Clustering

3.1.1.   Surface Discretization

All vertical building surfaces in the urban model are discretized in vertical and horizontal (each floor) directions (approx. 3m). To form an equal radiation surface area for solar radiation analysis, the division numbers are proportioned based on the façade dimensions of the building, i.e., width, length (Figure 3). The roof surfaces are excluded.

3.1.2.   Solar-Radiation Analysis

After discretization, solar radiation analysis (SRA) is conducted for all surfaces. SRA aims to systematically sort residential and retail units according to their incident solar radiation. The radiation results are assigned to the discretized surfaces. The radiation value of a unit is calculated based on the average value of the radiation analysis surfaces found adjacent to the unit. Based on the façade dimensions, each unit could have different number of radiation analysis surfaces. Then, units are sorted according to the weighted radiation values (Figure 3). For this study, a 10-level radiation scale is set.   Residential and retail units are categorized by the level of solar radiation.

3.1.3.   Clustering based on Vertical Position of Unit

SRA values help the sampling of the units into ten different clusters. Units are sorted based on their weighted radiation values; then, each cluster is divided into three sub-clusters based on the vertical position of each unit in the building (Chapter 3.1). This is because the ground temperature or exposed roof surfaces have different levels of heat transfer compared to adjacent horizontal surfaces in the middle floors. Therefore, three groups are formed based on solar exposure surface type, e.g., top floors, middle floors, ground floors, under the radiation analysis clustering as sub-clusters. The 5%, 10%, 20% sample size are applied for each sub-cluster to execute a uniform selection. This second sampling step is for the equally-distributed selection of the units based on similar thermal characteristics.

3.1.4.   Selection of Units for Energy Simulations

Yearly solar-radiation simulations were carried out on the 3D urban model, and the results were sorted into ten different radiation level groups. Then, the units are divided into sub-groups with floor-based clustering. The units for energy simulations are extracted from these clusters. In terms of

efficiency, the 5% sample size was 95 minutes, the 10% sample size was 149 minutes, and the 20% sample size was 258 minutes, while the full model lasted 1134 minutes. These simulation types are generated based on the selection ratio in Table 1.

Parameter	Property	Value
Non-Residential elimination	Based on the footprint area	Multiple values, {top, bottom}
Radiation level	Radiation scale	Single value, 10
Extracting candidates for simulation	Division of total number of units	Single value, 10
Number of floors	Based on building height limit [30]	Multiple values, {a,b,c}
Floor height	Based on building type	Multiple values, {a,b,c}
Space Conditioning Type	Zone Conditioning	Selection (heat, cool, mixed)

Table 1. Parameters of the UBEM Model

3.2 Model description

This chapter describes the space conditioning type and indoor thermal characteristics of each unit. The model consists of a cooling space conditioning system for some units based on the usage ratio of the total number of units, if the climate of the region demands it [27]. For this purpose, residential units are divided into two different clusters, that are mixed-mode (cooling and heating both exist, %20 of all units) and heating-only mode (only heating exists, %80 of all units). This ratio is parameterized in the model, and can be changed with users' preferences (Table 2). For mixed-mode residential units, there are different zones; living room and bedroom with cooling, and service areas without cooling. For heating-only residential units, the whole unit is considered as one single zone (Figure 4). All retail units are considered as heating-only.
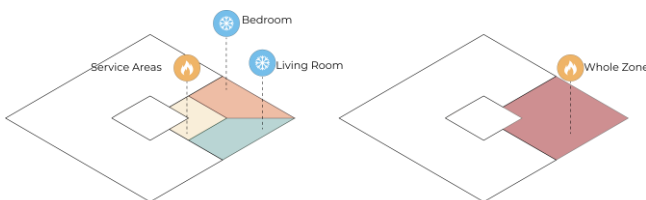

Figure 4. Mixed-mode (left) and Heating-only (right) Zone Unit Division

The space conditioning division, occupant characterization based on age and number, construction definition based on building construction period are identified from the government statistics bureau dataset [27]. The building constructions and schedules is selected according to TS-825 Turkish Standards and ASHRAE standards [1,2,28] (Table 2).

Surface Heat Transfer	Value
U-value, Wall (before, after 1980)	0.60, 1.88 W/m²-K
U-value, Roof (before, after 1980)	1.88, 3.12 W/m²-K
U-value, Floor (before, after 1980)	0.93, 1.92 W/m²-K
U-value, Window (before, after 1980)	5.1, 2.1 W/m²-K
Cooling Set Point (Mixed-mode)	25.0 ºC
Heating SetPoint	20.0 ºC
Heating Set Back	10.0 ºC
Natural VentilationType	One-sided
Infiltration Rate Per Area	0.0003 m³/s-m²
The fraction of Glazing Area	0.25
Natural Ventilation Limits	21.0 – 24.0 ºC
Number of People Per Area	0.0078 – 0.0394 ppl/m2

Table 2. Object Properties Used in the Model

Each unit zone contains different surface types and it could change due to outside boundary conditions, e.g., ceiling or roof surface. Therefore, each unit zone is distributed under different groups based on boundary condition properties on a vertical scale. Besides, the window openings and thermal heat transfers from the surfaces are organized based on the same methodology (Figure 5). The adjacent surfaces were determined as adiabatic surfaces to increase the simulation performance in terms of efficiency.
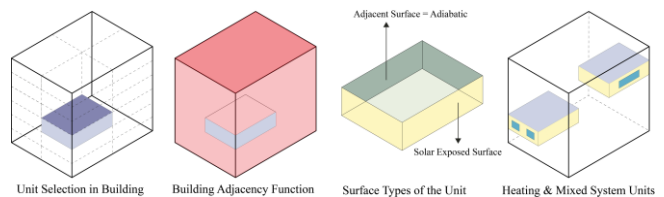

Figure 5. Unit Selection and Surface Adjacency

In the second development phase, internal loads are defined. Both population characteristics and occupancy habits are generated under nine occupant types based on the dataset of the TUIK [27]. These types are clustered according to occupant age that represents the population of the urban model, i.e., 0-65 adults from 0-65 age range and adolescent from 0-65 age range, 65+, in accordance with the national statistical datasets [27].

According to occupation types, schedules are generated to represent daily occupancy profiles, e.g., number of people, lighting, equipment and ventilation schedules. Additionally, natural ventilation is introduced in the model to reduce the cooling load in summer when the zone is occupied.

3.3 Simulation

For this chapter, the simulation process for the selected residential and retail units are described. Each unit in the buildings are modeled as described above, and is simulated separately. Annual energy simulations are performed using EnergyPlus. For radiation analysis, RADIANCE is selected [30]. Both simulation tools were operationalized using in the Grasshopper visual coding platform that contains both geometry formation, and energy calculation possibility with extra plug-ins, i.e., Honeybee, Ladybug [21].

3.4 Comparative Analysis

A second UBEM was built for a systematic comparative analysis and validation. For this, all zones (11972 units in 545 buildings) are constructed and simulated. This approach is commonly known as exhaustive search. The comparison metrics of energy demand types (i.e., heating, cooling, lighting) are arithmetic mean ($\overline{x}$), standard deviation (σ), mean absolute error (MAE). The success rates of the simulation types with the proposed algorithm are compared based on the 5% confidence interval statistical metric. A confidence interval is the range of elements in a group that evaluates under a degree of confidence [7].

4   RESULTS

This chapter presents the results of the simulation types with the proposed algorithm compared to the complete-model. Four urban building energy models were generated, the first three models are the proposed algorithm and the last one is the complete model of the whole neighborhood.

4.1 Modeling and simulation with proposed algorithm

The proposed algorithm consists of three different simulation types that are differentiated based on unit selection ratio 5%, 10%, 20% sample size. Each simulation process starts with the radiation analysis of the whole unit zones and distribution of the zone units for grouping the zone units. Then, zone units are selected from these groups randomly based on the selection ratio parameter. The proposed algorithm aims to provide efficiency for the total simulation process and reliability for the success rate of the process.

4.1.1.   Radiation Analysis Results

The radiation analysis step is only included in the proposed algorithm simulation types. It has resulted in the sampling of the residential units took approximately 40 minutes for each simulation. 162275 surfaces and 11972 units were analyzed for their solar radiation value on vertical surfaces (Figure 6). According to the radiation results, the selected units were sorted for their radiation values. Following, each radiation scale level was divided in terms of floor-based clustering as sub-groups, i.e., top floors, middle floors, ground floors. In total, 30 different groups were formed for each simulation type with the proposed algorithm. Based on the radiation scale parameter or the floor-based clustering parameters, the total group number could increase or decrease to increase the reliability.
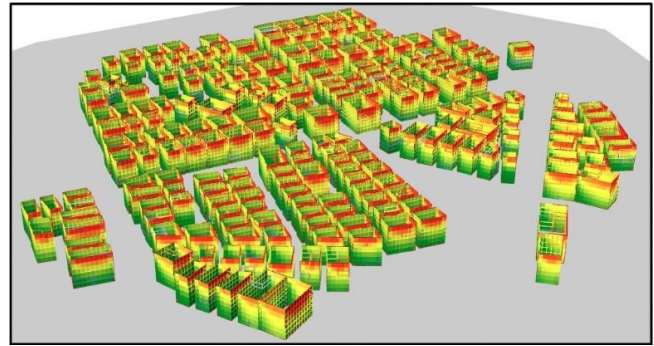

Figure 6. View of Radiation Results on the Façade of Buildings

Table 3 presents the ten radiation analysis groups with their average radiation value (kWh/m2) and three sub-clusters for each division. Based on different radiation values, the radiation groups were formed to execute better sample distribution of sample which is close to a total number of unit zones.

Radiation Scale	Average (kWh/m2)	Ground Floor	Middle Floor	Top Floor
0	90.5	812	1346	0
1	212.9	585	2492	2
2	352.6	252	2102	64
3	492.4	103	1330	145
4	633.2	41	520	493
5	774.9	24	317	397
6	913.1	12	123	335
7	1056.5	4	47	248
8	1197.9	4	19	105
9	1323.2	1	0	49
Total Number		1838	8296	1838

Table 3. Weighted Radiation Clustering Level

4.2 Modeling and simulation of complete model

Unlike the proposed algorithm, the complete-model simulation does not include radiation analysis and floor-based clustering processes. All unit zones are simulated individually with the brute-force method to provide validation for measuring the performance of the simulation types with different sample sizes. For this reason, although the duration of radiation analysis has been extracted from the simulation time, the total time is considerably longer than the proposed algorithms since the all unit zones are put into the structure energy simulation individually.

4.3 Comparative analysis between the proposed model and the complete model

Three different selection ratio numbers were simulated for testing the performance of the proposed algorithm (Table 1). For the Kültür Neighborhood of Izmir case, the ratio of space conditioning type were determined as 20% of a mixed-mode,

80% of the heating-only. The comparative analysis process is shown as:

- 20% sample size; heating-only (residential, retail), mixed-mode (residential)

- 10% sample size; heating- only (residential, retail), mixed-mode (residential)

- 5% sample size; heating- only (residential, retail), mixed-mode (residential)

- Complete model; heating-only (residential, retail), mixed-mode (residential)

(kWh)	5% sample	10% sample	20% sample	Complete model
$\overline{x}_{QH}$	35.7	36.6	38.9	38.7
σ_{QH}	16.1	16.9	15.7	17.0
$\overline{x}_{QL}$	17.2	17.3	17.3	17.1
σ_{QL}	1.4	1.4	1.4	1.4

Table 4. Heating-only units (%80 of total simulations) simulation comparison

The heating-only and the mixed-mode units are simulated separately to realize comparison based on space conditioning type. Table 4 presents the objective comparison of heating-only units for Q_H and Q_L in terms of average values and standard deviation to point out the reliability of the simulation types compared to complete-model results. From 5% to 20% sample sizes, the $\overline{x}_{QH}$ and $\overline{x}_{QL}$ values are improving as the number of sampling size is increased, however, the proposed algorithm results are highly close to the complete-model. There is a similar trend for σ_{QH} and σ_{QL}, and especially, σ_{QL} values are similar for all simulation types including complete-model. Lastly, the $\overline{x}_{QH}$ values varied more than $\overline{x}_{QL}$ values because the space heating is related to many thermal characteristics of the units, such as window directions, floor number, and occupant schedules. Lastly, the objective of the equipment load is not crucial for residential units; hence, the objective was not included in the comparison.

kWh	5% sample	10% sample	20% sample	Complete-model
$\overline{x}_{QH}$	30.8	33.6	33.7	34.8
σ_{QH}	9.4	12.8	12.9	12.7
$\overline{x}_{QC}$	29.2	30.8	31.7	32.4
σ_{QC}	12.8	13.3	13.5	14.1
$\overline{x}_{QL}$	8.2	8.1	8	8.0
σ_{QL}	1.8	1.8	1.8	1.7

Table 5. Mixed-mode units (%20 of total simulations) Simulation Comparison

In Table 5, the comparison of the units with mixed-mode simulation results is presented in the aspect of Heating Load (Q_H), cooling load (Q_C), and lighting load (Q_L) with average and standard deviation. The reliability for mixed-mode units is lower, due to their lower sample size as compared to the heating-only units. For $\overline{x}_{QH}$, the 10% and 20% sample size results are similar to each other as 33.6 and 33.7 kWh/ m^2, but the 5% sample size is lower than the other types as 30.8 kWh/m^2. In parallel, a similar trend is observed in the standard deviation values of QH. From 5% to 20%, as the selection rate increases, the reliability ratio also increases for $\overline{x}_{QC}$ and σ_{QC}. On the other hand, $\overline{x}_{QL}$ and σ_{QL} values result in approximately the same as 8.0 to 8.2 kWh/ m^2 and 1.8. Lastly, although some comparisons had similar results with other model types, the 5% sample size performed worse than the 10% and 20% sample sizes in all comparisons.
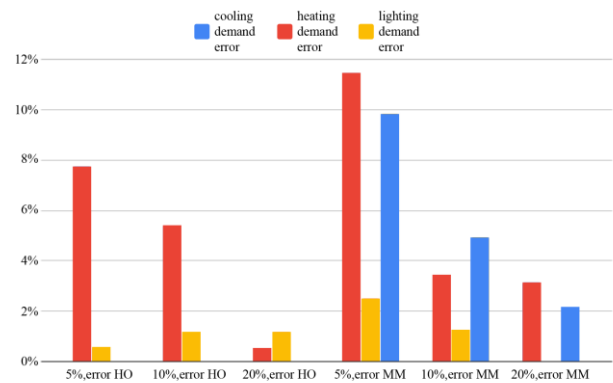

Figure 7. Error Ratio for Heating-Only (HO) and Mixed-Mode (MM) units

Figure 7 shows the error ratio of average values for zone heating-only; Q_H, Q_L, and mixed-mode; Q_H, Q_C, Q_L. For 10% and 20% sample sizes, the error ratios are close or lower than 5% confidence value. It is evident that as the sample size decreases, the error ratio increases for both heating-only and mixed-mode units. For the heating-only units, the highest error is below 8% with the heating demand of heating-only zone units. Therefore, to increase the reliability, the calculation accuracy should be increased when the total number of sample sizes decreases in turn. This can be possible with the parameters in Table 1. On the other hand, error ratios of mixed-mode units are higher, as the sample size for all simulation types is four times lower than the heating-only units. The highest error ratio is 12% for heating demand in the mixed-mode zone. Similarly, the cooling demand for the 5% sample size mixed-mode zone has the highest ratio between cooling demands of mixed-mode units. On the other hand, Q_L values are under a 5% error rate for all simulation types including heating-only and mixed-mode.

Minutes, Runtime	5% sample	10% sample	20% sample	Complete-model
Heating-only, %80	95	149	258	1134
Mixed-mode, %20	51	62	85	420
Radiation Analysis	40	40	40	-

Table 6. Runtime Comparison; Proposed Algorithm vs. Complete Model

The aim of the algorithm is to increase efficiency in terms of simulation period by means of a sampling method that implements radiation analysis and floor-based clustering for candidate selection. Table 6 presents the time comparison between consecutive simulation processes. As all simulations with proposed algorithm include radiation analysis for the whole urban geometry model, the radiation analysis is also added in the simulation time. Floor-based clustering process does not contains any analysis. Therefore, there is no time addition for the simulation types due to floor-based clustering.

In conclusion, the proposed algorithm performed successfully for the %10 and %20 sample sizes in terms of Q_H, Q_C, Q_L based on 5% confidence interval comparison. The error ratio proportionally increased when the sample size was lowered. On the other hand, the simulation runtime was significantly lower than the complete model simulation process. If the number of buildings to be simulated increases, the time difference between the proposed algorithm and complete model simulation in expected to increase.

5 DISCUSSION

In this study, a bottom-up approach for UBEM is presented. There are some comparative steps for different selection levels of the units for the energy simulation to calculate the success rate of the algorithm. The 10% and 20% sample size simulation types were successful, resulting in lower than the 5% error rate for the selected objectives. However, the %5 sample size simulations did not show complete success for heating and cooling loads. More reliable simulations should be conducted by increasing radiation scale levels for unit extraction for energy simulations or increasing the sample size. Moreover, the simulations were executed in the urban context, but for this process, there was no context element differentiation as a constraint between three different levels of simulation groups. The context geometries are highly effective for solar gain and natural ventilation objectives, which have an impact on both heating and cooling demands. Similarly, user types provided important input values for energy simulations by organizing the occupancy, heating/cooling set point, and setback schedules. As future work, additional parameters can be added related to context geometries and occupant types to increase the precision of the simulation results.

6 CONCLUSION

This study proposed an algorithm that decreases the computing cost of the UBEM and its simulations with a bottom-up approach. The proposed framework categorizes the building geometries as residential and non-residential geometries before energy simulation. By implementing solar-radiation analysis at the beginning, the units of the buildings are sorted based on weighted radiation values on the façade. Then, clustered units are categorized according to their positions in the building under three different sub-clusters such as top, middle, and ground floor clustering. Finally, the selected units are categorized according to their space conditioning. Based on the unit selection rates, three simulation levels were compared with the complete model simulations. 10% and 20% sample sizes error ratio was lower than 5% for all objectives. The 5% sample size resulted in increased errors due to the insufficient number of sample sizes. For future work, it is necessary to develop an inclusive UBEM framework with more precision for objectives, both to facilitate the simulation process and increase the content of the model in terms of other urban context components such as landscape elements (i.e. greenery, water) and urban infrastructure.

REFERENCES

1. ASHRAE. *ASHRAE Standard 55-2004 -- Thermal Comfort*. 2004.

2. ASHRAE. *ASHRAE Standard 90.1-2013 -- Energy Standard For Buildings Except Low-rise Residential Buildings*. 2013.

3. Azar, E. and C. Menassa, C. *Agent-Based Modeling of Occupants and Their Impact on Energy Use in Commercial Buildings*. 2012.

4. Bueno, B., Norford, L., Pigeon, G., and Britter, R. A resistance-capacitance network model for the analysis of the interactions between the energy performance of buildings and the urban climate. *Building and Environment*, *54*, (2012), 116–125.

5. Buffat, R., Froemelt, A., Heeren, N., Raubal, M., and Hellweg, S. Big data GIS analysis for novel approaches in building stock modelling. *Applied Energy*, *208*, October (2017), 277–290.

6. Davila, C.C. and Reinhart, C. *Urban energy lifecycle: An analytical framework to evaluate the embodied energy use of urban developments*. 2013.

7. Devore, J. *A Modern Introduction to Probability and Statistics: Understanding Why and How*. 2006.

8. Dogan, T. and Reinhart, C. Shoeboxer: An algorithm

for abstracted rapid multi-zone urban building energy model generation and simulation. *Energy and Buildings*, *140*, (2017), 140–153.

9. Engvall, K., Norrby, C., and Sandstedt, E. *The Stockholm Indoor Environment Questionnaire: A sociologically based tool for the assessment of indoor environment and health in dwellings*. 2004.

10. Groot-Marcus, J.P., Terpstra, P.M.J., Steenbekkers, L.P.A., Butijin, C.A.A. User Behaviour and Technology Development. In *Technology and Household Activities*. 2006.

11. Groppi, D., de Santoli, L., Cumo, F., and Astiaso Garcia, D. A GIS-based model to assess buildings energy consumption and usable solar energy potential in urban areas. *Sustainable Cities and Society*, *40*, May (2018), 546–558.

12. Hong, T., Chen, Y., Luo, X., Luo, N., and Lee, S.H. Ten questions on urban building energy modeling. *Building and Environment*, *168*, (2020), 106508.

13. Howard, B., Parshall, L., Thompson, J., Hammer, S., Dickinson, J., and Modi, V. Spatial distribution of urban building energy consumption by end use. *Energy and Buildings*, *45*, (2012), 141–151.

14. ISO. *ISO 9241-210:2010- Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems*. 2010.

15. Kavgic, M., Mavrogianni, A., Mumovic, D., Summerfield, A., Stevanovic, Z., and Djurovic-petrovic, M. A review of bottom-up building stock models for energy consumption in the residential sector. *Building and Environment*, *45*, 7 (2010), 1683–1697.

16. Kelly Seryak, J. and Kissock, K. *Occupancy and behavioral affects on residential energy use*. 2003.

17. Lawrence Berkeley National Laboratory. Input Output Reference. EnergyPlus. 2009.

18. MURE, O.-. *Monitoring of EU and national energy efficiency targets*. 2012.

19. Oliveira Panão, M.J.N. and Brito, M.C. Modelling aggregate hourly electricity consumption based on bottom-up building stock. *Energy and Buildings*, *170*, (2018), 170–182.

20. Pachauri, S. An analysis of cross-sectional variations in total household energy requirements in India using micro survey data. *Energy Policy*, *32*, 15 (2004), 1723–1735.

21. Roudsari, M.S. and Pak, M. Ladybug: A parametric

environmental plugin for grasshopper to help designers create an environmentally-conscious design. *Proceedings of BS 2013: 13th Conference of the International Building Performance Simulation Association*, (2013), 3128–3135.

22. Schuler, A., Weber, C., and Fahl, U. Energy consumption for space heating of West-German households: empirical evidence, scenario projections and policy implications. *Energy Policy*, *28*, 12 (2000), 877–894.

23. Soebarto, V.I. and Williamson, T.J. Multi-criteria assessment of building performance: theory and implementation. *Building and Environment*, *36*, 6 (2001), 681–690.

24. Swan, L.G. and Ugursal, V.I. Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, *13*, 8 (2009), 1819–1835.

25. Swan, L.G., Ugursal, V.I., and Beausoleil-Morrison, I. Occupant related household energy consumption in Canada: Estimation using a bottom-up neural-network technique. *Energy and Buildings*, *43*, 2–3 (2011), 326–337.

26. Tanimoto, J., Hagishima, A., and Sagara, H. A methodology for peak energy requirement considering actual variation of occupants' behavior schedules. *Building and Environment*, *43*, 4 (2008), 610–619.

27. TÜİK. *TÜRKİYE İSTATİSTİK KURUMU Turkish Statistical Institute*. 2010.

28. Turk Standartlari Enstitusu. TS825 Binalarda Isi Yalitim Kurallari. *825*, (2009), 80.

29. Vine, E. and Barnes, B.K. Monitored indoor temperatures and reported thermostat settings: How different are they? *Energy*, *14*, 5 (1989), 299–308.

30. Ward, G.J. The RADIANCE lighting simulation and rendering system. *Proceedings of the 21st annual conference on Computer graphics and interactive techniques - SIGGRAPH '94*, (1994), 459–472.

31. Ye, Y., Wu, Q., Huang, J.Z., Ng, M.K., and Li, X. Stratified sampling for feature subspace selection in random forests for high dimensional data. *Pattern Recognition*, *46*, 3 (2013), 769–787.

32. Yılmaz, M. Changes of Rural Population in Turkey and its Distribution by Provinces (1980-2012). *Eastern Geographical Review*, *33*, 20 (2015).

Conditional Generative Adversarial Networks for Pedestrian Wind Flow Approximation

Sarah Mokhtar[1], Aleksandra Sojka[2], and Carlos Cerezo Davila[2]

[1]Kohn Pedersen Fox Associates
London, United Kingdom
smokhtar@kpf.com

[2]Kohn Pedersen Fox Associates
New York, United States of America
{asojka, ccerezodavila}@kpf.com

ABSTRACT

Despite the vital role the wind microclimate plays in urban design, the time, cost and computational expense of wind simulations make them ineffective during early design stages. Surrogate models have been explored as an alternative to the traditional time-consuming computational fluid dynamics (CFD) methods for wind flow approximation in urban contexts. Current approaches, however, have limitations that hinder their applicability and integration in early design pedestrian wind microclimate studies. The high-volume of data generated from CFD simulations provides a practical opportunity for exploring data-driven surrogate models. This paper proposes and evaluates a conditional generative adversarial network surrogate model for the design contribution of urban morphologies to pedestrian wind flow conditions. The validity of the approach is demonstrated at a fraction of the time that would be required to perform the equivalent conventional simulation. Variations in dataset encoding techniques, image resolutions and geometric diversity of the training set are explored to identify the key parameters affecting model's accuracy and suitability for the intended application. Larger urban configuration datasets are recommended to further reduce errors and instill confidence in model predictions for urban contexts quite far from the training set.

Author Keywords

Deep Learning; Performance Simulation; Computational Fluid Dynamics; Generative Adversarial Networks; Surrogate Modeling; Pedestrian Wind Comfort.

1 INTRODUCTION

The ongoing rapid urbanization, with more than 68% of the world's population expected to live in cities by 2050 [1], presents architects, urban designers and city planners with an additional layer of complexity to create livable, resilient and enjoyable cities. In the face of density, cities will face higher vulnerability to climate change, and building professionals will have to design to mitigate the anticipated higher urban heat island effects as well as higher urban wind speeds from high-rise buildings [2]. Wind microclimate studies [3] have extensively explored the relationship between urban morphology and the pedestrian wind conditions, including impacts of building form, building height variations, building arrangements, building porosity, street canyon aspect ratio and orientation. Shaping the built environment to promote favorable wind microclimate conditions can contribute positively to the quality and usability of public spaces through improvements in wind comfort, thermal comfort and air quality through city ventilation and pollution dispersal [2]. These in turn define the activity types and walkability levels which are key to the functioning, vitality and essence of the city.

With design projects increasing in scale from one building to the planning of an entire city, architects and engineers are faced with problems of a complexity that intuition, experience and generic studies insights alone cannot address and where simulation becomes an essential tool to inform design decisions [4]. In response to this growing necessity, several cities have issued and enforced industry standards for the study of a design-contribution to pedestrian wind comfort, including the Dutch NEN 8100 and NPR 6097 in 2006, the EU-RTD COST guidelines in 2007, the Japanese AIJ guidelines in 2008, with the most recent being the wind microclimate guidelines published by the City of London in early 2019. The later recommends the use of either computational fluid dynamics (CFD) or wind tunnel testing or both depending on the complexity of the urban context [5]. Both approaches come with steep time and cost challenges, while CFD has the advantage of providing full flow field data.

Despite the vital role the wind microclimate plays in urban design and its integration in planning regulations, the time, cost and computational expense of wind simulations make them ineffective during early design stages. In practice, the design process is generally characterized by uncertainties in problem definition as well as rapid and frequent changes, both of which necessitate an adaptive, fast, robust [6] and seamlessly integrated simulation workflow [7]. The focus of building performance simulation developments in the past decades, particularly in CFD, has been improving the accuracy rather than speed, a trade-off driven by the low error tolerance in aerospace engineering applications [8]. In contrast to other wind engineering applications, pedestrian wind level (PLW) is "one of the few topics […] where nature is kind to us concerning turbulent flows" [9]. At less constrained comfort ranges and more relaxed error

tolerances, enabling CFD for the exploration and optimization of a large design space through faster methods can provide an understanding of performance variability that bring insights in early design stages that one high-accuracy solution cannot.

In order to address the limited integration of slow feedback conventional CFD simulations in early design optimization workflows [7], a number of approaches have been explored to produce fast approximations models (further expanded in the following section). The high-volume of data generated from CFD simulations provides a practical framework for exploring data-driven surrogate models. End-to-end feature learning through deep learning approaches was shown to be effective in multiple applications in the recent years [10]. This study draws on the opportunity presented to design practitioners to use the CFD simulation data that is produced for multiple projects to train a machine learning model to predict pedestrian urban wind flows during early design explorations. It proposes and evaluates a deep learning methodology, and explores multiple training dataset parameters and encoding strategies as they relate to acceptable error thresholds suitable for the intended application.

2 RELATED WORK

The following review covers different approaches to tackling the high-computational expense of traditional CFD methodologies through both solver approximation and data-driven workflows. It focuses on identifying the approaches' suitability for wind microclimate studies and potential integration in early design architectural and computational process.

2.1 Solver Approximations for CFD

In order to address the time-intensive nature of CFD simulations, a general research focus has been on developing solver approximations. This was achieved through "simplified meshes, use of lower-order equations or the treatment of turbulence through modelling" [8]. For wind microclimate studies, there has been a general focus on the use of steady Reynolds-averaged Navier–Stokes (RANS) simulations. In Toparlar et al.'s review of 183 CFD studies from 1998 to 2015, 96% used RANS despite large eddy simulations (LES)'s potential to provide higher accuracy [3]. At one to two orders of computational time magnitude larger than RANS and a larger modeling complexity [11], the use of LES on practical applications becomes rather limited. This, in turn, reinforces the industry's need for valid but simpler and faster methods for evaluating outdoor urban conditions.

Advanced numerical methods coupled with increased computing capabilities have been investigated to provide from moderate acceleration of conventional approaches to potential real-time CFD feedback. Those can be classified as mesh-based, mesh-free methods and hybrid methods [12]. One approach, which was highly explored in architectural

early wind-focused design explorations, is the fast fluid dynamics solver (FFD). Originally developed for real-time fluid visualization in the gaming industry, it gained traction due its notable speed in comparison to RANS [13], about 500-1500 faster on a GPU [14]. In the scope of design exploration, a low Reynolds number model was validated as suitable for simulating air flow inside buildings for natural ventilation applications [15]. It was, however, shown to be less accurate in capturing wind flows between buildings particularly downstream [15], a discrepancy attributed to its lack of turbulence model.

Other GPU-enabled methods are being explored in the context of external urban wind flows. Those include the Smoothed-Particle Hydrodynamics (SPH), a mesh-free approach based on solving particles movement and fluid properties within a certain distance, and the Lattice Boltzman Method (LBM) based on solving the Boltzman equations for particles on a lattice [16]. While those are underway to be developed to suit urban scale applications, they have not yet been demonstrated to provide 'reliable and repeated accuracy' for pedestrian wind comfort applications [16]. Additionally, they rely on high GPU computing requirements and cannot yet be integrated in architectural workflows in a way that can support iterative early design explorations.

2.2 Data-driven CFD approximation models

Machine learning approaches have been proposed as an alternative to approximate wind flow through learning a relationship between geometry represented as input feature vectors and corresponding measured fields [8]. Neural network implementations investigating the urban wind interference impacts are not recent, with early experimentations dating to more than two decades ago [17], [18]. These focused on using wind test data to estimate wind interference factors based on two building separation distances both the along-wind and across-wind directions. More recent studies focused on capturing wind fields from large datasets of full CFD simulations. Artificial neural networks, random forest algorithms and other reduced order models were investigated for approximating wind pressures of buildings with and without urban wind interference impacts [8], [17]–[20]. While effective in relating shape features to interference factors or wind fields, the studies are limited in geometrical parameters and do not provide full field wind flow data for the spaces around buildings. Those limitations, although not directly inferable to urban scale applications, are being challenged in other CFD domains. An example of that is [21] combined a PolyCube novel 3D geometrical representation with a Gaussian method for real-time interactive full field wind speed and pressure for car shape optimization.

2.3 Deep Neural Networks and CFD

The deep neural networks' (DNN) remarkable capability to learn complex, high-dimensional, non-linear dynamic systems, that was shed to light in the last few years, gained

traction in modelling turbulence flows, a high dimensional (spatial and temporal) physical phenomenon [22]. Some of these approaches focused on augmenting physics-based models with their uncertainty quantification and aimed at improving accuracy predictions [23]. Other studies explored applications of DNN models as surrogates for enabling CFD-based design optimization through faster predictions. Deep generative model implementations including convolutional neural networks (CNN) started being particularly explored in this context due to their ability to better represent data in non-linear input and output functions. For instance, Kim et al. [24] constructed, using a CNN, a model that can estimate fluid velocity fields based on data from Eulerian fluid simulations at up to 700 times faster speed than a CPU-CFD. Another example is Guo et al. [10] CNN models that were capable of predicting two and three-dimensional non-laminar flow velocity fields twice as fast as the equivalent GPU-accelerated CFD simulation.

Generative adversarial networks (GAN), which consist of a generator and a discriminator model competing against each other, have been explored as well in predicting fluid behavior due to their additional ability to learn implicit data distributions without the need to define an explicit loss function. An implementation of that is the conditional generative adversarial networks (cGAN) that were used by Farimani et al. [25] for direct predictions of steady state heat conduction and incompressible fluid flow. High accuracy predictions, based entirely on learning from observations, were achieved with a mean absolute error percentage of 1% for different boundary conditions and domains. The deep learning model implementations show the potential of the approach to provide high accuracy inference at state-of-art computational performance in the fluid. Those methodologies can be used to predict steady state flow without any knowledge of fluid flow governing equations and only requiring the encoding of boundary conditions. These results, although promising, are specific to their applications and focused on simple geometries, and thus not generalizable to complex urban systems.

2.4 Limitations and Opportunities

The aspiration for integrating CFD in the early design process has pushed research towards looking into novel methods for fast wind flow approximations. Despite the range of explorations, the current approaches have limitations that prohibit them from being actively used in the urban and architectural design process in the context of pedestrian wind comfort. Advanced numerical methods that are GPU-enabled, such as SPH, LBM and methods, are promising but necessitate high hardware computing requirements, are not integrated in design workflows and have not yet been validated as suitable for predicting external wind flow around buildings. Data-driven workflows, particularly deep learning approaches in CFD, have shown their effectiveness in predicting flow without any knowledge of fluid flow governing equations. The capacity to infer flow fields solely on a set of observations is remarkable and its

agnostic nature to the source of the data makes it accessible to a wider audience (architects, engineers, designers …), not limited to any particular software, platform or simulation method and capable of continuously expanding accuracy as more data becomes available. As with the research into accelerated numerical methods, the current workflows tested are specific to the certain applications they were designed for, and results cannot be directly used to infer wind flow in complex urban contexts, neither has their suitability for pedestrian wind comfort applications been explored.

The approach presented in this paper draws on the opportunities that deep learning generative models open for surrogate modeling of physics phenomena and how that can allow for wind-influenced insights into the early design process. The study applies a cGAN to obtain a surrogate model for the design contribution of urban morphologies to pedestrian wind flow conditions. The main contribution of this research is the exploration of the suitability and limitations of this method for the pedestrian wind comfort applications in early design stages, and the evaluation of different training set encoding strategies to identify what simplifications are acceptable for this purpose.

3   METHODOLOGY

The development of a pedestrian wind flow surrogate model for early design stages is proposed based on a cGAN model. This adopted process, represented in Figure 1, starts with the generation of a synthetic dataset based on CFD simulations of procedurally generated urban configurations. An investigation of different geometry representation and dataset encoding techniques are then explored. This is followed by deep learning training, network architecture optimization and finally model evaluation.
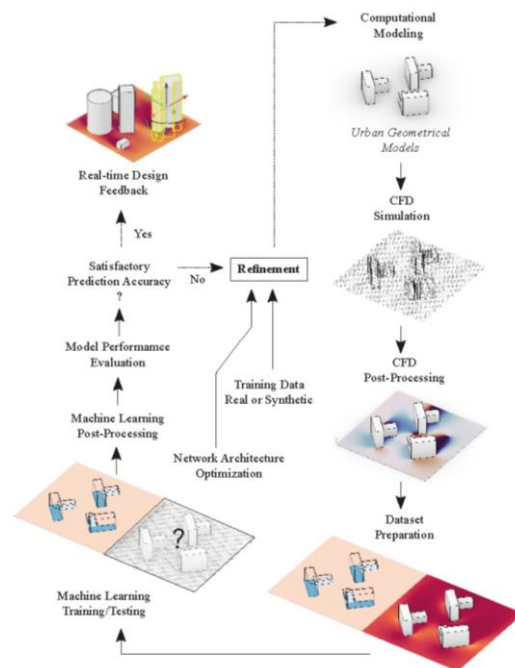

Figure 1. Workflow

465

A cGAN model architecture for image-to-image translation applications has been adapted for this study based on a PyTorch implementation [26]. This model is an extension to the generative adversarial network (GAN) [27] deep learning model which is composed of two competing networks: a generator, which maps a sample to a desired distribution and a discriminator which determines if a sample belongs to this distribution or has been generated by the generator. The conditional model relies on conditioning an output to an input. In this context, the model inputs and outputs are RGB images with constant pixel count. Abstract and high dimensional data from the geometrical representation of the urban environment is extracted in the encoding layers of the network. These are then mapped, through multiple decoding layers, into wind speed values.

3.1 Urban Geometrical Models
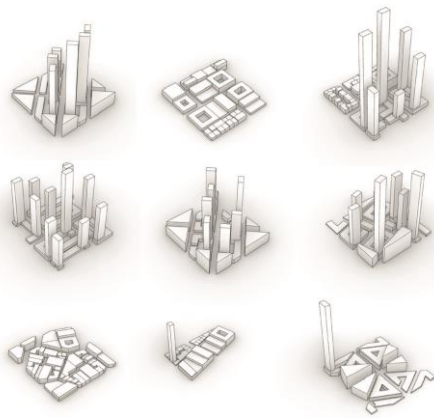
Figure 2. Sample of Urban Geometrical Configurations

Two sets of geometrical configurations were created in order to provide a larger design representation potential: the first set focused on simple configurations of two or three buildings, while the second set provided a larger fidelity to the complexities and rules governing the urban grids and morphologies. These were generated through procedural modelling in Rhinoceros 3D CAD environment using Grasshopper its visual programming environment, while balancing between representability of the urban context and design space dimensionality. The first set included 800 primitive geometries varying in base shape, rotation, relative location within the boundary and height, set within a boundary of 350 m. The second set included 225 urban configurations, varying in grid typology and orientation, density, building morphologies and heights, set within a boundary of 550 m. These were procedurally generated based on typical dimensions and densities of urban contexts.

3.2 CFD Simulation

The *OpenFOAM* solver was used to solve the governing equations of steady-state RANS with a turbulence model of realizable k-ε, following the support in the literature and industry standards for the robustness of this CFD approach for PLW applications [5], [11], [16]. Due to the varying complexity and scales of geometries tested, the time each simulation took to reach convergence ranged from 20 to 120

minutes on a 2.9 GHz i7, running in parallel on Ubuntu Linux distribution. The threshold for convergence was defined as residuals falling below the recommended 10^{-6} [5], which typically meant a termination at around 800 iterations for simple geometries and 2000 for more complex urban configurations. A rectangular domain was used for the analysis, defined as 5 times the maximum building height on the windward, top and sides, and 15 times on the leeward side. The urban configurations tested varied in maximum building height, ranging from 40m to 250m. Cartesian-based 3D structured hexahedral meshing was used at a cell-to-cell expansion ratio of 1.1. A minimum cell size of 3m was selected for this study has been identified as an adequate balance between computational feasibility and a low average error through a sensitivity analysis. A flat terrain was assumed for all cases and a landscape roughness of a dense urban environment was used. At the upstream inlet, a reference wind speed of 6 m/s at a 10m reference height was used for all simulations.

3.3 Geometry Representation and Dataset Preparation

To convert a 3D geometry representation into a 2D matrix, a height map on a Cartesian grid was created for each, where zero level is defined as the wind speed height of interest. In areas where buildings extend beyond the height of interest, a positive distance was assigned, and in other cases, a negative distance to building roofs or ground plane was assigned.
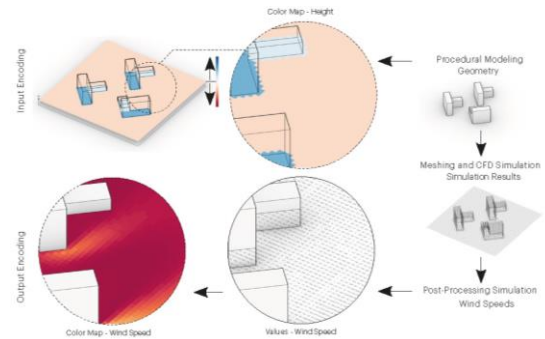

Figure 3. Dataset Encoding

For each simulation case, an automation process was developed through Paraview Python scripting to extract mesh geometries and their corresponding speed values at assigned heights of interest which were selected to range from 1.5m above ground to 36m, at 1.5m intervals. Different data encoding strategies were explored, as listed in Table 1.

Image Sizes	Color Scheme	Colormap Ranges	Training Set Size	Geometric Diversity
256px	Turbo	0 – 8 m/s	5,400	Primitive
512px	Spectral	0 – 16 m/s	13,400	Complex urban
1024px	Inferno	0 – 32 m/s		Hybrid

Table 1. Dataset Encoding and Training Parameters Explored

For the height map matrix and its corresponding speed values, values have been converted to RGB through colormap functions with either 256 or 766 discrete states. To

466

ensure consistency between input and output images, both were extracted from the same platform. The image pairs were used to create the training and testing sets for the cGAN network. For the first set of urban geometries, 13,800 image pairs were generated, and 5,400 image pairs for the second set of complex urban configurations. For each case, in order to investigate the impacts of image size, colormap functions and colormap ranges on training, multiple sets of each case were created. Three image resolutions were extracted at 256, 512 and 1024 pixels, two high-contrast colormaps were extracted using matplotlib's Spectral and Google AI's Turbo colormap functions and a 766-unique-colors linear colormap function as well in addition to three maximum colormap ranges of 32, 16 and 8 m/s. Image pairs were visually inspected for errors and the dataset was cleared accordingly with a loss of about 15% of the set.

3.4 Network Architecture and Model Training

A Pytorch implementation of a Pix2Pix network architecture [26] was selected. The Pix2Pix model's generator is built with Unet Skip Connection architecture, a variant of encoder-decoder network where the series of convolutional layers of the encoder are 'connected' to deconvolutional layers of the same size of the decoder. The down-sampling (encoder) blocks consist of Convolution-BatchNorm-LeakyReLU layers and the up-sampling (decoder) blocks consist of Deconvolution-BatchNorm-Dropout-ReLU layers. For our experiments, Batch-Normalization was replaced with Instance-Normalization layer for both up-sampling and down-sampling blocks and additional skip connection blocks were added to scale the network up and allow us to train on larger input image sizes, up to 1024x1024 pixels. The model's discriminator follows PatchGAN architecture, a deep convolutional neural network that is designed to classify NxN patches of an input image (generators output) as a real or fake, and then averaging the results, instead of classifying the entire image once. The discriminator model is trained to minimize the negative log likelihood of identifying real and fake images. The generator model is trained using both the adversarial loss for the discriminator model and the L1 loss, the mean absolute difference between the generated translation image and the input image. The generator model had 79.577 M parameters and the discriminator 6.963 M. The datasets were trained for 400 epochs, using kaiming weight initialization, batch size of 32, initial learning rate of 0.0002 with scheduled linear learning rate decay after 200 epochs. After every 5 epochs, the trained weights were saved in order to monitor the progress of the training. Depending on the size of the dataset, the training varied from 500 sec/epoch (4000 x 1024px x 1024px images) to 2500 sec/epoch (12000 x 1024 px x 1024 px images). The training was conducted using several hardware configurations: Windows with two NVIDIA GV100 Volta GPUs with 32MB memory and 128GB RAM; Linux with NVIDIA RTX8000 GPU with 48MB and 128GB RAM; Windows with eight NVIDIA Quadro P6000 GPUs with 24MB memory.

4 RESULTS

To evaluate the accuracy of the proposed framework, the generated wind speed predictions from the cGAN models are compared with the solutions from the RANS CFD simulation. The mean absolute errors (MAE) and 90th percentile errors in m/s are computed for a testing set of 22 geometries unseen to the training network with 528 corresponding image sets. For each image representation, the MAE is defined as the sum of the absolute differences between the predicted and simulated wind speeds for every pixel representation in the image divided by the total number of pixels. Since the interest is in predicting wind flow at pedestrian level, in addition to the cumulative set values, the error was calculated separately for testing image sets below 7.5m height. In addition to a quantitative assessment, a visual inspection of the testing sets further provide an understanding of the models' ability to represent wind behaviors and patterns around buildings and identify zones with highest wind speed amplification.
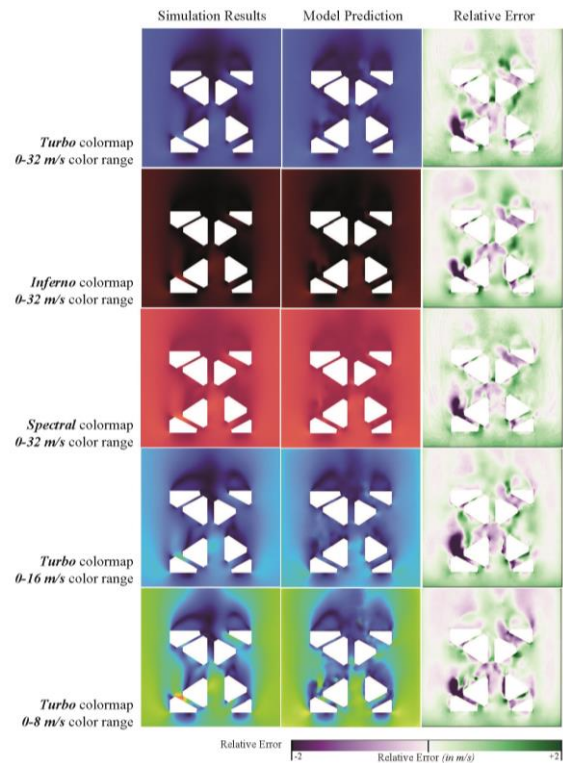

Figure 4. Data Encoding Techniques and Performance

This evaluation method was used to compare different data encoding techniques and training parameters against their corresponding model performance, as illustrated in Figure 4 and 5. Figure 4 illustrates, for the same simulation case, the different data encoding methods used for wind speeds and their respective impact on model prediction errors. The columns from left to right represent the visual representation of simulation results, the corresponding model prediction and the relative error between both. Figure 5 represents a box plot summarizing the MAE and 90th percentile errors for the different encoding strategies and datasets. From left to right,

the represented sets are colormap representations, ranges, building mask color and image size. A distinction is made as well between results of wind speed errors for all trained heights versus pedestrian heights which represent the focus of this work. The selection of data encoding technique for the wind speed was shown to have an impact on the accuracy of the results and can be listed in descending order of impact as: the building mask color, the color range and the color map.
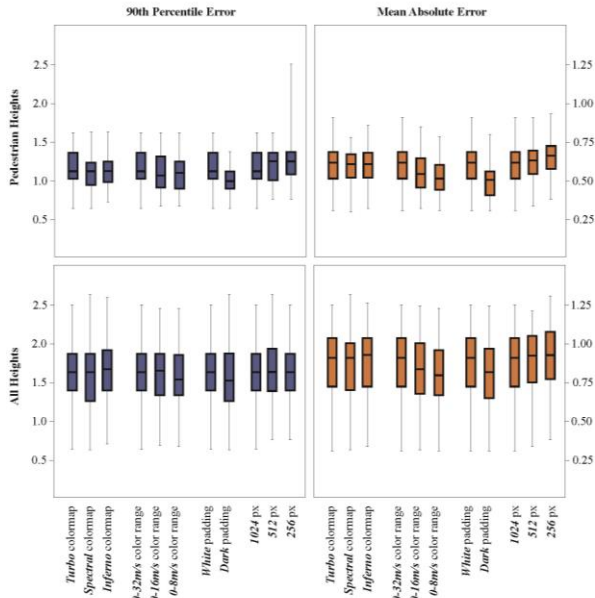

Figure 5. Models' Performance

Firstly, despite being useful for dataset visual inspection, assigning a white color for buildings comes with a disadvantage. A near-grayscale layer of pixels at building edges is observed in many cases and results in the need for a less straightforward building/non-building pixel separation post-process before a conversion to a wind speed matrix is possible. As a by-product, an increase in MAE is observed at an average of 0.1 m/s.

Secondly, a shorter color gradient range representation achieves lower overall MAE of about 0.05 m/s on average as well as more visually legible maps due to the larger number of visually distinct colors. The rare instances in the training set at which high wind speed values reach 16 m/s or more that would be capped by the 0-8 and 0-16 m/s color ranges do not seem to have a significant impact in the overall MAE.

Thirdly, the improvements from the high-contrast smooth visualization of the 'Turbo' and 'Spectral' maps outweighs the three-times larger number of unique colors in the 'Inferno' linear color map. Multiple input training image sizes were as well tested: 1024px, 512px, and 256px. As clear in Figure 4, there is a general improvement observed when training with larger inputs ranging from 0.03 to 0.1 m/s average error difference. This rise in accuracy, however, comes at a steep computational cost, where a 4000-images

set training on Windows with 2 GPU/GV100, 32MB memory, 128 GB RAM, would take about 48hrs to complete 400 epochs for a 1024px set compared to 3hrs for 256px. The geometry complexity and diversity of both training and testing sets have the largest noticeable influence on model improvement and prediction accuracies. The first set of training solely included urban geometries with about 4000 training images compared to the same set augmented with simple geometry configurations with about 15000 training images achieved minimally lower accuracy predictions. The larger set, on the other hand, expanded the capacity of the model to infer the morphology contribution to wind speed of both simple building arrangements and dense urban configurations. The predictions accuracy varied as well depending on the deviation of tested instances from the training set geometries. Tested geometries that deviated from the continuous street grids and consistent building morphologies that were generally maintained in the training set corresponded to the highest sources of speed differences.

5 DISCUSSION

Putting the proposed model results in perspective with its intended application to predict pedestrian wind flow conditions around buildings is essential to evaluating its suitability. In contrast to the average MAE of 0.7m/s and the 1.5m/s 90th percentile error for the full testing set, the pedestrian level heights results were capped at an average MAE and 90th error of 0.5m/s and 1.2m/s respectively, decreasing to 0.3m/s and 0.7m/s for urban configurations following the consistent street grids and morphologies it was trained on. Among the comprehensive criteria for pedestrian wind comfort is Lawson criteria which classifies wind speeds in 2m/s interval bands to which is assigned a corresponding percentage of time exceedance. Given the 2m/s-wide band, the proposed model's 0.5m/s MAE reflects its potential suitability for early design pedestrian wind comfort predictions when augmented with a larger repertoire of urban morphologies and further optimizations of the model. The proposed approach presents some challenges and limitations.

Firstly, as with other DNN implementation for physics-informed applications [22], machine learned models are prone to overfitting and there is no direct methodology to ensure that this can be avoided that is applicable to all morphologies that are quite far from the trained dataset, it is essential to cross-validate the model against overfitting. As was observed in the results, although the model was able to infer wind speeds on urban configurations outside the training set, a larger disparity between simulation and model prediction was observed when the continuous street grids and building shapes that the model was trained on was disturbed. This suggests that a larger repertoire of urban morphologies is desirable to further improve its usability, which should be weighed against the computational expense of simulating those with CFD.
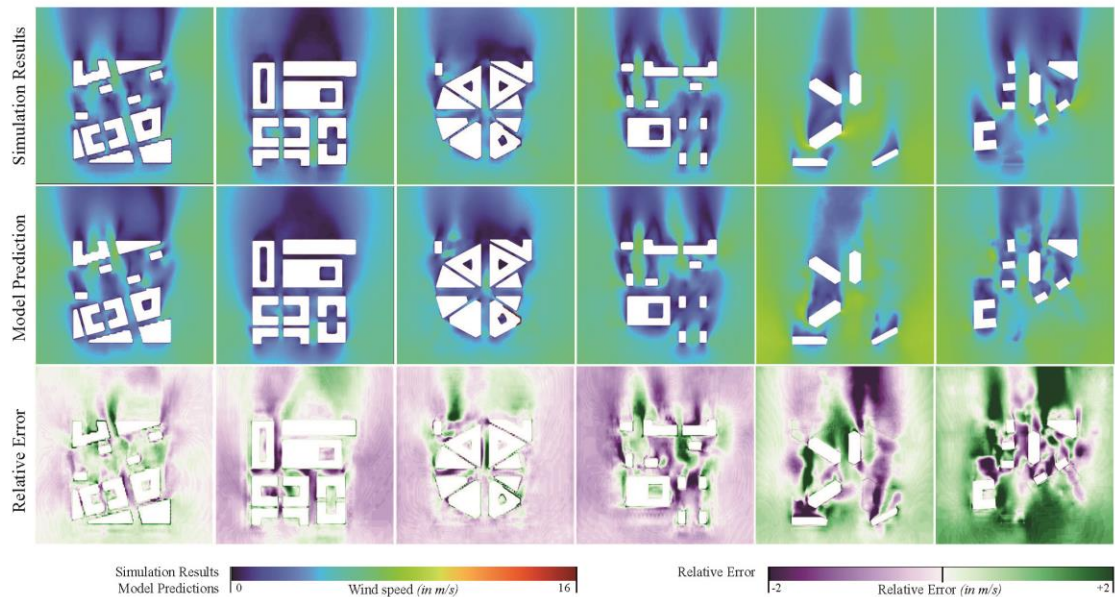
Figure 6. Sample of Model Predictions from Lowest MAE to Highest MAE

Secondly, the current geometry encoding strategy presented can represent a large variety of urban morphologies through an adapted version of a height map. It is, on the other hand, limited to buildings that are continuously solid throughout and cannot yet represent elements such as tunnels, bridges, carved terraces, etc. Other representations of the 3D geometries can be further explored to expand the model capabilities through either augmenting the image encoding for voids using additional color channels or exploring adapted mesh or voxel data structures in a different model architecture.

Thirdly, although applicable to other initial wind speeds, the proposed workflow has only been explored for an initial wind speed at the inlet of 6m/s at a 10m reference height. Different initial wind directions are, on the other hand, possible through rotating the geometry relative to the base south wind direction. In order to overcome the initial wind speed boundary condition limitation, a number of strategies could be further explored. An additional encoding layer to the input representation can be integrated to represent input speed. Another method could be to base the machine learning training on wind factors and use those to infer resulting wind speeds based on the inlet speed.

Fourthly, the lack of structured data for CFD simulations of urban morphologies necessitate the creation of synthetic training sets which require the computationally expensive RANS CFD simulations and thus limit the capability to expand considerably training sets. The challenge here extends beyond solely the computational expense of those simulations, but also the time expense associated with all the automation processes that are needed to convert simulation data into structured datasets suitable for machine learning workflows, as well as the time associated with data and simulation results validation and quality control.

Lastly, as GPU-enabled computational approaches become increasing adopted in the CFD simulation domains and particularly for wind flow in urban contexts, they will present solutions that are significantly more cost effective than CPU-based models. If validated as suitable for pedestrian wind flow studies, they can potentially become viable in the future for faster wind flow approximation, and would substitute the need for data-driven surrogate workflows.

6 CONCLUSION

The paper presented a conditional generative adversarial network approach to developing a surrogate model for the inference of the design contribution of urban morphologies to pedestrian wind flow conditions for any given urban configuration. The developed model is capable of generating pedestrian wind flow approximations for a wide range of urban configurations at an accuracy of about 0.3m/s within seconds, a fraction of what would be needed for a an equivalent CPU-based CFD simulation. For wind microclimate studies, where comfort band ranges are about 2m, the current accuracy results reflect its potential suitability for early design pedestrian wind comfort predictions. The selection of dataset encoding technique, training set parameters and geometrical diversity was shown to have an impact on the accuracy of the results. Those averaged to about 0.1m/s change for each strategy to an accumulated combined 0.3m/s error difference. While some of those improved encoding strategies come with no additional computational cost, the benefit from increased image size and training data size should be evaluated against the corresponding computational expense. The ability of the cGAN to infer wind flow physics from observation was demonstrated and its potential to provide high fidelity results through expanding the training set and further optimizing the

469

model network was shown to be possible at state-of-the-art computational performance.

REFERENCES

[1] United Nations, "World Urbanization Prospects: the 2018 Revision," 2018.

[2] P. Moonen, T. Defraeye, V. Dorer, B. Blocken, and J. Carmeliet, "Urban Physics: Effect of the Micro-climate on Comfort, Health and Energy Demand," *Front. Archit. Res.*, vol. 1, pp. 197–228, 2012.

[3] Y. Toparlar, B. Blocken, B. Maiheu, and G. J. F. van Heijst, "A Review on the CFD Analysis of Urban Microclimate," *Renew. Sustain. Energy Rev.*, vol. 80, pp. 1613–1640, 2017.

[4] S. Hanna, L. Hesselgren, and V. G. Ignacio Vargas, "Beyond Simulation: Designing for Uncertainty and Robust Solutions," in *Proceedings of the 2010 Spring Simulation Multiconference*, 2010.

[5] City of London and RWDI, "Wind Microclimate Guidelines for Developments in the City of London," 2019.

[6] S. Hanna, L. Hesselgren, and V. Gonzalez, "Beyond Simulation: Designing for Uncertainty and Robust Solutions," in *Proceedings of the 2010 Spring Simulation Multiconference*, 2010.

[7] A. Chronis, A. Dubor, E. Cabay, and M. S. Roudsari, "Integration of CFD in Computational Design - An evaluation of the current state of the art," in *Proceedings of the 35th eCAADe Conference*, 2017, pp. 601–610.

[8] S. Wilkinson, G. Bradbury, and S. Hanna, "Reduced-Order Urban Wind Interference."

[9] J. H. Ferziger, "Approaches to Turbulent Flow Computation: Applications to Flow Over Obstacles," *J. Wind Eng. Ind. Aerodyn.*, vol. 35, pp. 1–19, 1990.

[10] X. Guo, W. Li, and F. Iorio, "Convolutional Neural Networks for Steady Flow Approximation | Autodesk Research," in *ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2016.

[11] B. Blocken, "LES over RANS in Building Simulation for Outdoor and Indoor Applications: A Foregone Conclusion?," *Build. Simul.*, vol. 11, pp. 821–870, 2018.

[12] M. L. Hosain and R. Bel Fdhila, "Literature Review of Accelerated CFD Simulation Methods towards Online Application," in *Proceesings of the 7th International Conference on Applied Energy (ICAE2015)*, 2015, pp. 3307–3314.

[13] A. Chronis, A. Turner, and M. Tsigkari, "Generative fluid dynamics: integration of fast fluid dynamics and genetic algorithms for wind loading optimization of a free form surface.," 2011, pp. 29–36.

[14] W. Zuo and Q. Chen, "Fast and Informative Flow Simulations in a Building by Using Fast Fluid Dynamics Model on Graphics Processing Unit," *Build. Environ.*, vol. 45, no. 3, pp. 747–757, 2010.

[15] J. Mingang, W. Zuo, and C. Qingyan, "Simulating Natural Ventilation in and around Buildings by Fast Fluid Dynamics," *Numer. Heat Transf. Part A Appl.*, vol. 64, no. 4, pp. 273–289, 2013.

[16] D. Philips and M. Soligo, "Will CFD Ever Replace Wind Tunnels for Building Wind Simulations?," *Int. J. High-Rise Build.*, vol. 8, no. 2, pp. 107–116, 2019.

[17] A. C. Khanduri, C. Bedard, and T. Stathopoulos, "Modelling Wind-Induced Interference Effects using Backpropagation Neural Networks," *J. Wind Eng. Ind. Aerodyn.*, pp. 71–79, 1997.

[18] E. C. English and F. R. Fricke, "The Interference Index and its Prediction using a Neural Network Analysis of Wind-Tunnel Data," *J. Wind Eng. Ind. Aerodyn.*, vol. 83, pp. 567–575, 1999.

[19] S. Wilkinson, S. Hanna, L. Hesselgren, and V. Mueller, "Inductive aerodynamics," 2013.

[20] A. Zhang and L. Zhang, "RBF Neural Networks for the Prediction of Building Interference Effects," *Comput. Struct.*, vol. 83, pp. 2333–2339, 2004.

[21] N. Umetani and B. Bickel, "Learning Three-Dimensional Flow for Interactive Aerodynamic Design," *ACM Trans. Graph.*, vol. 37, pp. 1–10, 2018.

[22] J. Nathan Kutz, "Deep learning in fluid dynamics," 2018.

[23] K. Duraisamy, G. Iaccarino, and H. Xiao, "Turbulence Modeling in the Age of Data," *Annu. Rev. Fluid Mech.*, vol. 51, pp. 357–377, 2019.

[24] B. Kim, V. C. Azevedo, N. Thuerey, T. Kim, M. Gross, and B. Solenthaler, "Deep Fluids: A Generative Network for Parameterized Fluid Simulations," in *Proceedings of EUROGRAPHICS 2019*, 2019, pp. 59–70.

[25] A. B. Farimani, J. Gomes, and V. S. Pande, "Deep Learning the Physics of Transport Phenomena," 2017.

[26] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[27] I. J. Goodfellow *et al.*, "Generative Adversarial Nets," in *Proceedings of the Advances in Neural Information Processing Systems*, 2014.

Suggesting Design Directions: Early Examples of Simulation-Based Guidance for Common Model Types

Nathan C. Brown

Pennsylvania State University
University Park, PA, USA
ncb5048@psu.edu

ABSTRACT
Simulation data can inform early design, but for many design spaces, there is a need for distilling large amounts of performance data into guidance that can support creative, interactive design. Automated computational processes could help support this activity. This paper demonstrates a procedure for finding suggested design directions based on automatic simulations of design options for three common model types. The workflow involves reading in a static geometry and basic simulation information, automatically generating dummy variables, running a series of simulations, applying data analysis to find a direction for improvement, and then returning that direction to the user. This process is demonstrated on a 2D truss for reducing weight, a 3D surface structure for increasing stiffness, and an urban neighborhood concept for increasing PV potential. These simulations reveal one possible foundation for a future design system that intelligently suggests areas of performance improvement for an initially fixed truss, shell, or urban form.

Author Keywords
Parametric design; design suggestion; improved performance; canonical correlation analysis; trusses; surface structures; urban massing

ACM Classification Keywords
I.6 SIMULATION AND MODELING; I.2 ARTIFICIAL INTELLIGENCE; J.6.1 Computer-aided design (CAD)

1 INTRODUCTION
Early stage architectural design involves creative processes in which geometry and other design choices are proposed, discussed, and evaluated. Often, this is done collaboratively among people with different preferences, intentions, expertise, and goals. Such a dialogue can produce useful exchanges in which a solution is proposed, and experts with specific concerns and knowledge suggest a way to improve the given design. This guidance might involve modifications to a system, a set of components, how they are to be arranged, or other decisions that must be made. For certain buildings and related structures, discussion can be primarily about geometry, which has the potential to move or morph continuously within a chosen system. For example, an experienced structural engineer might suggest a more efficient shape than the initial concept, or a building scientist might propose a shifted massing or orientation that reduces energy loads or increases access to daylight.

While these conversations often depend on the experience, knowledge, and creativity of those involved, there is potential for computer simulations to play a role in making suggestions. To start, simulations can quantify the impact of recommended design modifications. Beyond this, computers themselves might find areas of improvement that are novel or counterintuitive for conceptually complex geometries. In other areas of life, numerous recommendation systems exist that operate in a related way, suggesting new songs or products a user might like to purchase. Yet these systems operate on extensive datasets involving many people and their actions. Researchers in architecture and urban design are considering ways to build up similar datasets in the design field [15] and finding other methods for computationally responding to preferred design ideas [4]. Both approaches could have significant future benefits.

However, this paper takes a targeted approach to finding an immediate direction for design improvement. The strategy starts with an automatic parameterization and simulation-based exploration of nearby designs, and then finds trends or patterns in a dataset generated based on these designs. Ideally, a fully formed artificial intelligence system for building design could act like a general suggestion engine— the designer likes "this" geometry; would she or he consider trying "that" geometry, which may improve performance? Or at least try moving in that direction? While not a full suggestion system, this paper considers ways in which three typical geometric types might be evaluated to generate a performance-based direction for design improvement. Through three case studies, it provides examples of meaningful suggested geometric modifications that arise from analyzing a dataset originating in the design itself.

2 BACKGROUND
Today, design firms are increasingly using parametric simulation to generate possible outcomes, determine their performance, weigh their merits and drawbacks, and ultimately make decisions [9, 13, 18]. The idea of performance here refers to quantifiable, desirable aspects of buildings, often in the domain of structure, energy, daylighting, and acoustics. Many common simulation

engines are connected directly to generative design software, making the digital feedback loop between geometry and performance more efficient than ever. As this connection is refined, researchers have developed techniques for generating catalogues of designs, visualizing complex design spaces, finding design space trends, and conducting interactive optimization [3, 12, 19, 20].

Despite the increased prevalence of parametric simulation in architectural practice, there remain unresolved issues. One is that designers must often put forth considerable effort to code a parametric logic, shape grammars, or other strategy to generate possibilities. This task is frequently reserved for specialists and can become time-intensive and tedious if done repeatedly for varied design concepts and projects. To relieve burdens related to coding parametric options, researchers have attempted to create parametric models that generate themselves [10], or domain-specific models that can gradually learn how to predict performance from underlying structures without extensive new simulation [23].

Specific to geometry, some have also created generalized workflows for automatically processing geometry and connecting it to performance outcomes, including for building energy [6], and for views and other urban criteria [7]. Disciplines outside architecture have done considerable work in shape morphing, as it is fundamental to computer graphics and increasingly useful in engineering design [8, 16, 17]. Much of this work contains sophisticated geometric manipulations that improve the design process, but there are still opportunities for removing tedium in computational workflows through targeted automation. This paper introduces one such application, which is suggesting design directions based on performance without the need for coding a parametric model specific to the problem being studied.

Second, the presence of a large design space does not necessarily provide guidance [12]—it is up to the user to interpret the possibilities and make decisions. In many cases, optimization can fill this role of guidance, finding a high-performance design within a user-coded parametric design space [21]. Yet, there are plenty of instances in practice in which a design team generates a conceptual geometry but wants a computer to answer a simple question—from a design that is already preferred, what is a direction that could improve the design in terms of performance? How should the geometry be morphed, and by how much would this adjustment improve a given metric? This paper offers a computational strategy for addressing these questions using automated parameterization, simulation, and analysis, which is especially useful for multi-objective design scenarios or when certain objectives are difficult to quantify.

3 GOALS AND METHODOLOGY

The goal in this paper is to demonstrate a framework for suggesting design improvement on three geometric model types that arise frequently in conceptual parametric design—trussed structures, surface structures, and urban form. The emphasis is on how geometry can be automatically

parameterized on granular level, and yet pattern recognition can find smooth, meaningful, performance-based suggested design directions for these geometric typologies. This paper considers structural material quantity, strain energy, and PV potential, but this approach could be used with other quantitative design objectives, as well as with other data science techniques that use an initial dataset to map between geometric variables and a global simulation response.

The general methodology for this workflow is indicated in Figure 1. For each model type, a user must first input a geometry along with basic information necessary for performance evaluation to be implemented. The input geometry can be a preferred design already under consideration for architectural reasons, or an atypical shape for which designers want to know how to improve the performance. The additional input is specific to technical domains—in the examples in this paper, require information includes support locations and loading for the structural case studies. Ideally the input information required beyond geometry is minimal, such that most of the process is automated and users can concentrate on rapidly iterating input geometries or other aspects of the design. However, there are tradeoffs between the specificity of a given evaluation type and its generalizability.
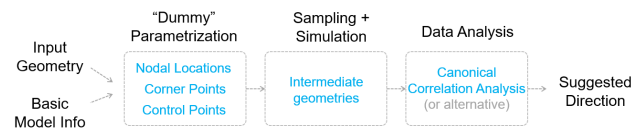

Figure 1. Basic workflow for finding performance-based suggested design directions.

Next, the provided geometry is automatically parameterized with "dummy" variables based on its type: nodal locations for trussed designs, control points for surfaces, and building corner points for urban massing. Since these selections are specific to the model type, they are not generalized further. However, a future "user" could have a library of scripts to choose from and apply the appropriate base parameterization as needed. Some assumptions must be made at this point—for example, how many degrees of freedom should the trusses and control points have, and how far should they be allowed to move? Must building faces stay orthogonal, or should corners be allowed to move independently? In this paper, assumptions are made about each of these questions related to scale, dimensionality, and design freedom. Although in practice such decisions might need frequent updating, certain assumptions could be written once and held constant for, as an example, a firm that frequently works on gridshell roofs of a typical scale.

A series of design samples are then generated, and their performance is simulated. These intermediate design samples represent slightly perturbed versions of the original. As shown in the case studies, the intermediate samples are often not useful designs themselves—they may be wrinkled shells or nonsensical trusses. However, trends can still be

found by considering them together as a coherent dataset. To analyze this design space and find a direction for improvement, canonical correlation analysis (CCA) [11] is used. This analysis provides a set of linear coefficients for input variables that maximize correlation with another dataset. Although the primary application of this technique is comparing multidimensional sets of data, when used to find coefficients for dummy design variables that maximize linear correlation with performance, it can produce architecturally compelling directions for improvement. These directions are then returned to the user by way of their coefficients. A 2D visualization of such an outcome is given in Figure 2, showing how two initial parametric variables could be remapped such that a user is moving in a direction that should improve performance.
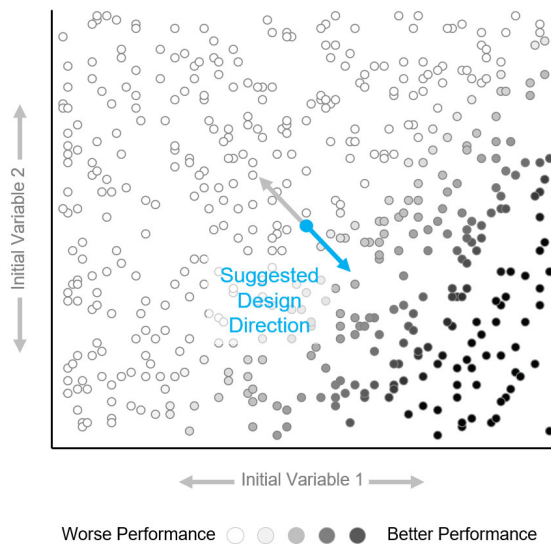
Figure 2. A visualization of how data analysis might find a better design direction than original variables, for a 2D design space.

Although the application of CCA to interactive design has been described before in [2], this paper extends the method to further geometric typologies. Across all model domains, the potential benefits of this approach include finding directions for improvement of a complex geometry that would be difficult to reveal otherwise; finding directions that are counterintuitive; focusing on specific modifications that seem to matter most; and synthesizing desired outcomes through combined analysis of multi-objective guidance.

The current implementation of this workflow uses native Grasshopper components and custom scripts for the automatic parameterization, and the plug-in Design Space Exploration [1] for generating data and manipulating the design. Performance simulations rely on additional plug-ins mentioned in the next section. Once the analysis is completed, designers have access to a slider that morphs the design along the suggested direction. As it is directly on the Grasshopper canvas, this method of suggestion inherits the typical interface and visualization of the parametric design software itself. Although there are currently a few manual

steps in the procedure (such as triggering the sampling and data analysis), these could be automated in the future. Although the design manipulation and visualization occur directly in Grasshopper, the functionality for suggesting directions could also be added to a separate interface.

At present, this methodology requires many intermediate simulations to generate suggested directions. Although the sampling and simulation is automated and could be completed while a user is executing different tasks, the time is still significant. Furthermore, this strategy does not provide every way to improve a design, instead giving a single direction per desired quantitative objective (although in a multi-objective scenario, various combinations of priorities could give many directions). Nevertheless, this approach is worthwhile for complicated geometries in which it is difficult to extract a single, cohesive suggested direction without using data. Its relevance would increase for future design environments that run faster simulations and project future directions for multiple design goals simultaneously. The following three case studies demonstrate initial success in finding architecturally meaningful directions of design improvement for basic geometries.

4 CASE STUDIES

4.1 Trusses

In this first example, a script automatically parameterizes a basic truss and then finds suggested directions for improving its performance. Based on the general methodology and interface described in Section 3, this workflow incorporates: 1) reading in a static truss geometry (Figure 3), identifying nodes and creating dummy design variables corresponding to the location of each node; 2) acquiring design information, which includes supports, loads, and boundary conditions; 3)
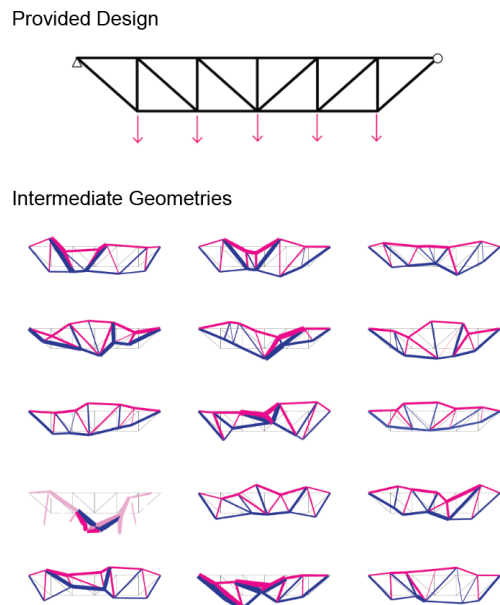

Figure 3. Initial input geometry and a selection of generated intermediate designs for the truss case study. Compressive members are visualized in pink, and tension members in blue.

sampling the initial design space; 4) analyzing the resulting dataset using canonical correlation analysis; and 5) mapping the coefficients back into the original variables to provide a new direction for morphing the structure.

The truss is simply supported and loaded vertically at its lower nodes. For every geometry, Karamba is used to [14] apply loads, calculate internal forces, chose an adequate member size, and return the sized member. Example intermediate designs for the truss are also shown in Figure 3. Very few sample designs are viable solutions, due to the complete freedom of each node to move independently both vertically and horizontally. However, suggested directions for improving the performance of the truss (i.e. lowering its structural material quantity and subsequent weight) seem realistic, as shown in Figure 4. This figure shows the corresponding suggested geometries along with simulation results evaluating the structural weight and deflection in these geometric directions. The three columns represent different amounts of data, which gives a sense of how many simulations are required to generate a coherent result.

encouraging in themselves—no symmetry, constraints, or reasonable bounds were imposed on the design initially, since the raw nodal locations were used directly as trial variables. More information about this approach and case study can be found in [2].

4.2 Surface structures

In the next example, the input geometry is the surface structure shown in Figure 5, and the quantitative design goal is to minimize strain energy. In early design, simulation of this design objective can assess the efficiency of the shape, which has implications for required structural material. The chosen geometry is a loft through a series of curves at different heights and orientations, leading to complex double curvature. Due to BIM and digital fabrication, such geometry is increasingly common for grid shell roofs and other structures around the world. Furthermore, the selection of curves with different heights and curvatures will demonstrate how canonical correlation analysis can highlight specific regions of the surface that require urgent attention for improving performance.
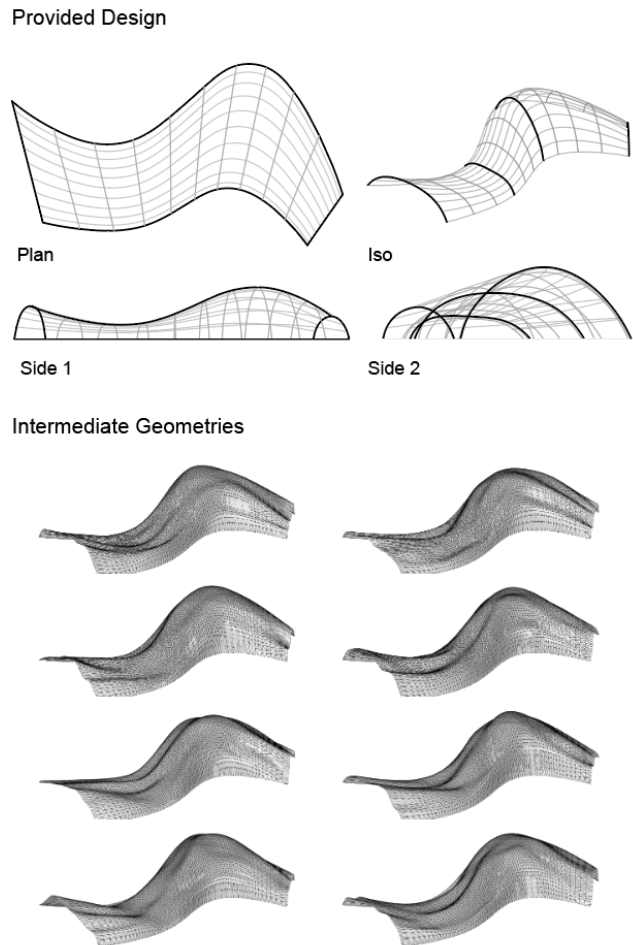

Figure 4. Suggested directions for structural improvement from a flat truss, indicating that increased depth can lower weight.

When asked to find a direction that lowers structural weight, modifications were returned that generally corresponded to truss depth, which tends to control truss efficiency. A reasonably smooth transformation occurred for both the 2,000 and 5,000 sample datasets. This suggests that a fairly large amount of data is needed to extract a discernable pattern, but there are diminishing returns in the smoothness that can be eventually created. Yet these truss directions are

Figure 5. The initial geometry and corresponding samples of intermediate, automatically generated surface structures. The intermediate structures are not smooth, yet they lead to a rational suggested improvement to the original structure.
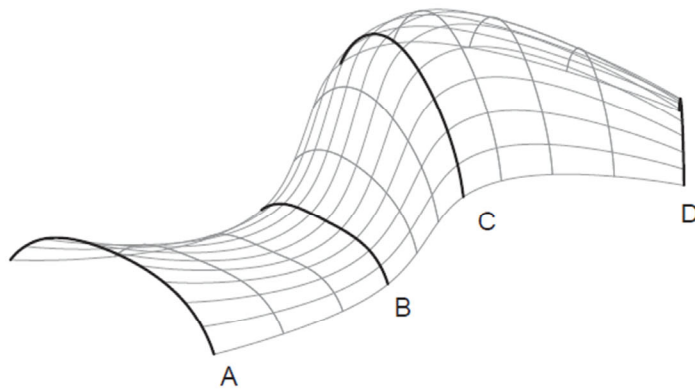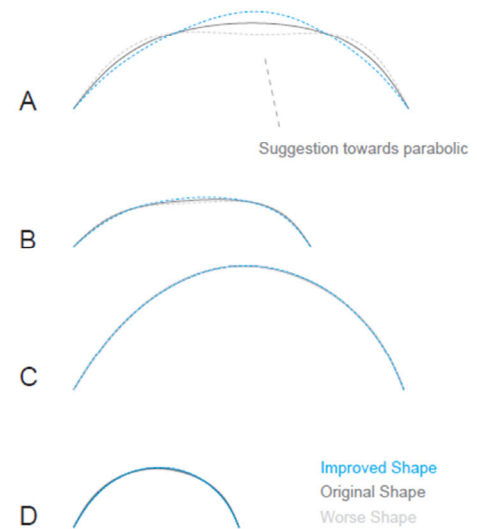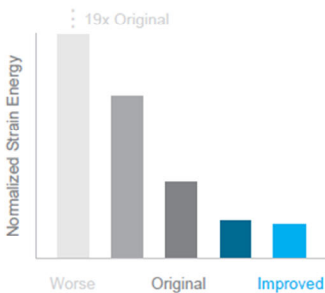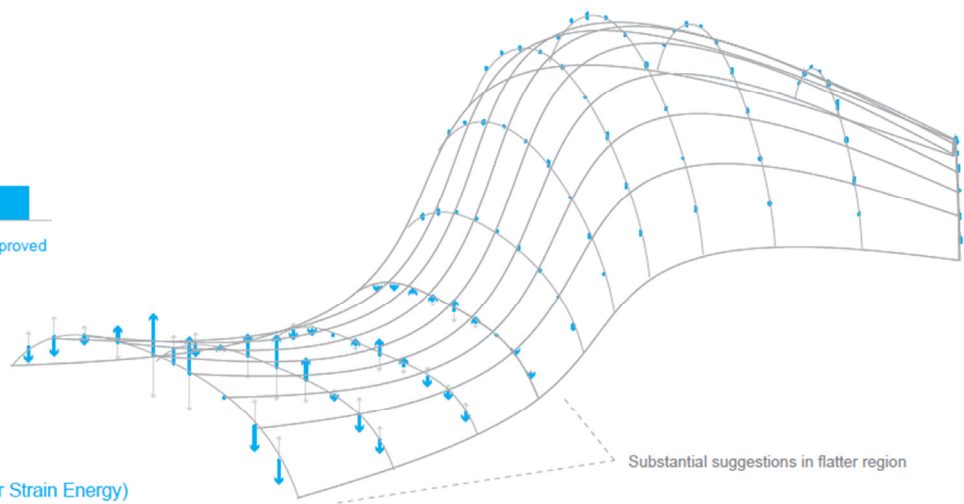
474

Figure 6. The original surface structure geometry, suggested improvements, and change in performance for these improvements.

The initial shape is first fed into a script that automatically locates control points for manipulation. In this case, 121 control points were located along the surface, and were given a degree of freedom in the z-direction. Next, 10,000 samples were generated based on this new parameterization, and a simulation determined their strain energy. Karamba was again used for the structural simulation, with a uniform load in the negative z-direction. Any edge of the surface touching the bottom plane was assumed to be translationally fixed. The script created for this paper is generalizable to any input surface for which a vertical load and supported edges can be assumed and the software can adequately find control points. However, unlike the truss, there are no set nodal locations for the control points. Thus, for future testing, the resolution of control points might need to be modified to generate meaningful results. The structure is approximately 50 m long, with a maximum span of ~20 m between supports.

Next, a canonical correlation analysis was conducted to determine coefficients that map a direction to improve the performance of the structure. This direction (and its opposite) are visualized in Figure 6, in which the length of blue arrows shows how the structure should be modified to reduce its strain energy, making it stiffer against the vertical load case. By observation, this technique concentrates on the shallower part of the roof, where there is a structurally problematic flat zone. The suggested modification is to morph the curvature of certain sections towards a parabola, which is more efficient for the load case. Although this modification is visually subtle, moving to the structure visualized in Figure 6 cuts the strain energy in half for the entire design, as shown in the normalized graph. It is again notable that although the intermediate samples taken represent wrinkled, unrealistic structure, the overall analysis yields a relatively smooth suggested transition.

It must be acknowledged that when working with surface structures, there are sophisticated formfinding methods to directly arrive at a more efficient structure. In many applications, the force density method, dynamic relaxation, or another technique could find a better shape from the beginning. Nevertheless, there are situations in which an architectural designer might begin with a complex initial form based on their own design goals, which may relate to constructability, competing performance objectives, or another desired quality. In these cases, a suggested direction like this example might show an avenue to improve performance while still staying true to designer intent, or initiate a conversation about tradeoffs between structural performance and preferred geometric outcomes.

4.3 Urban massing

The next case study demonstrates the potential of this automated procedure at the urban scale. In this case, the designer provides the computer with an initial massing for an urban building complex with varying floorplates and heights. The goal is to find a direction for modifying the geometry that would increase the PV potential of the roofs, assuming they would be covered with solar panels. This methodology could be used for urban daylighting or energy simulations, although at present a PV calculation runs considerably faster. A similar procedure involving dummy parameterization, sampling, and analysis is used to determine geometric transformations that correlate with increased PV potential. The initial variables automatically assigned are the corner points of buildings, with the script assuming that buildings can grow taller but remain rooted on the ground. The simulations for PV potential were conducted for the Boston climate using Archsim [5]. The resulting direction is given in Figure 7, which shows changes to the design that affect the overall appearance and spatial sequence of the buildings.

Some aspects of this geometric transformation show that the urban massing workflow is the least developed, as it is a more complex problem requiring additional assumptions in the script if designers require a clean final output. A more refined artificial intelligence process will eventually include constraints on what is feasible in a real urban setting, and build smart rules and capabilities into the design process, such as the ability to pick which walls should stay frozen and recognize that buildings should be combined when they overlap.
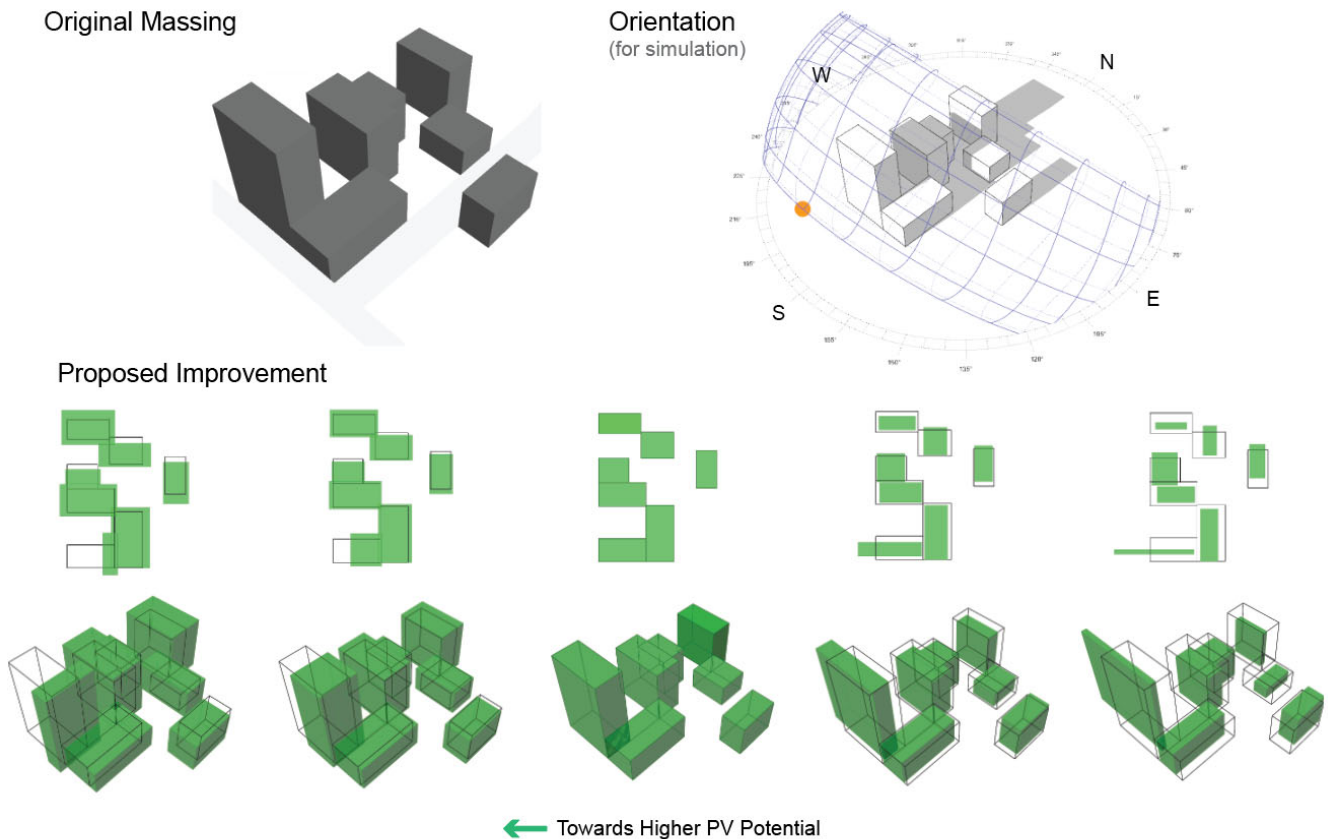

Figure 7. The original massing, simulation orientation, and proposed improvement for increasing PV potential across all building roofs.

Still, the results indicate potentially compelling directions for design exploration. For example, the direction for increasing PV potential seems to show that the tall, southernmost building, which significantly blocks the others in this orientation, should be peeled away, while most other buildings should spread out to maximize surface area. At the urban scale, such suggestions may be more useful than building-by-building adjustment due to the geometric complexity of separate buildings. Consequently, they can provide a starting point for urban designers for where to move next beyond simulation feedback. Such guidance would likely be generated for multiple competing interests in the design, providing a rich exploration of both the design and objective spaces.

5 DISCUSSION

Figure 8 gives a visual a summary of the parallel approaches to providing a design improvement path for the three geometry types. Though these methods show initial success in generating meaningful suggestions for design, a discussion of applications is provided here. It should be mentioned that this automated suggestion workflow should not be the only step in design exploration. As Wortmann [22] notes, simplifying a proeblem into a single (or few) directions for improvement may miss alternatives that are higher performing. A design space might need to be expanded, contracted, or modified throughout the process, and design space exploration or optimization-based workflows are often suitable for this task.

Thus, this workflow is not a replacement for the approaches to data-driven design that have been previously established. However, for applications in which designers want to specifically consider desirable geometries and have a sense for how they might be improved, or control the tradeoff between affinity for the original geometry and potential performance gains, such an automated process may have a role. These situations could include time-pressured processes in which there is not time for systematic optimization for a single solution, when strong qualitative goals are driving the initial geometry generation, or when it is suspected that a computer might discover a compelling direction for improvement that would be difficult to find intuitively due to model complexity or other factors. Although the examples here are preliminary, the ability to extract smooth transformations from datasets full of automatically generated designs that are not individually useful is encouraging. Future experimentation with advanced machine learning methods might yield even more effective results.

Although currently implemented with problem-specific parametric design scripts, this research imagines a future interface with faster simulation and cumulative information from previous designs, in which a variety of metrics could be queried automatically as need. For example, the designers of a complex gridshell shape could provide a desired geometry and ask how to improve stiffness, reduce structural weight, reduce incident solar radiation, or reduce overall energy consumption, while having control over subtle geometric manipulations that achieve each of these desired outcomes. This interaction might occur before or after a more systematic optimization procedure and could augment conversations around preferred designs and worthwhile changes, whenever these conversations occur in the design process. The preliminary case studies in this paper indicate that automatic parameterization combined with data science can find compelling design suggestions for such situations.
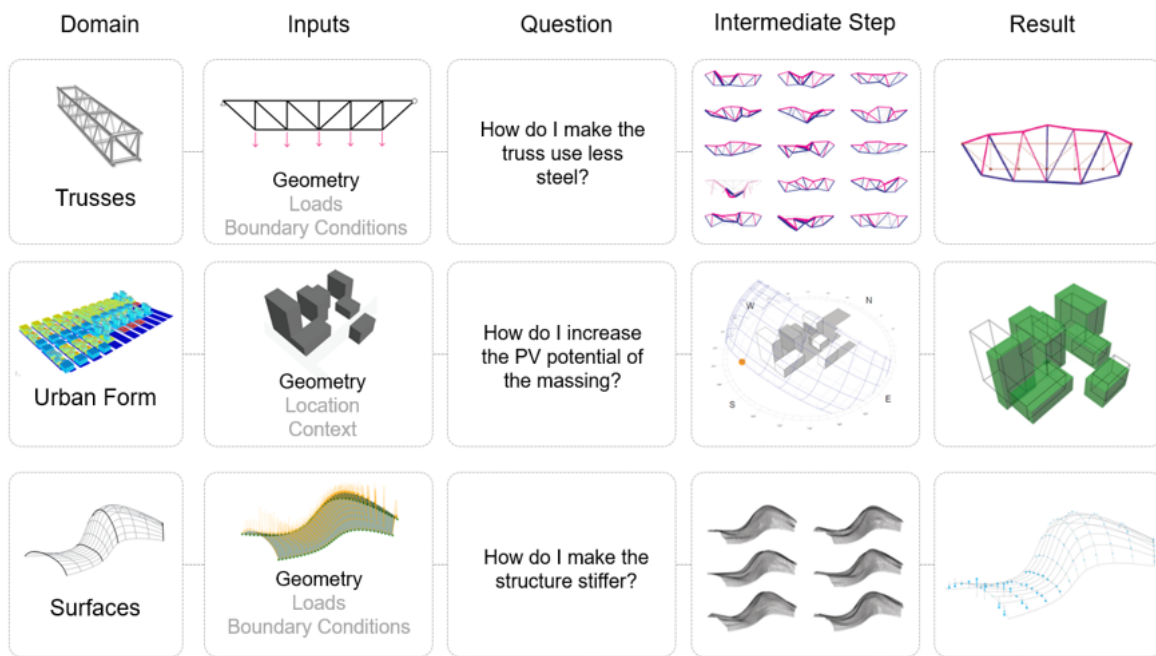

Figure 8. Summary of workflows and suggested directions for the three case studies.

6   CONCLUSION

This paper has demonstrated a computational procedure for generating suggested design directions. This approach includes automatic parameterization, performance simulation, and data analysis, and it was applied to trusses, surface structures, and rectangular buildings. As this workflow is still early, there are many areas for future work. First, rather than using a combination of scripted components and custom code, this could be implemented in its own design interface. A new interface could make additional assumptions for each geometric type, as mentioned for the case studies, to increase its generalizability for input designs. The workflow should also be tested for more complex geometry, additional performance metrics, and other data analysis techniques that go beyond linear correlation. Such a process could also be combined with ongoing research on automatically translating a 2D sketch into 3D models, which would allow designers to literally sketch ideas and get feedback rather than drawing first in CAD. Nevertheless, the initial results in this paper are a step towards intelligent design suggestion based on performance simulation.

ACKNOWLEDGMENTS

REFERENCES

1. Brown, N. and Mueller, C. Design Space Exploration, 2017. www.food4rhino.com/app/design-space-exploration.

2. Brown, N. and Mueller, C. Gradient-based guidance for controlling performance in early design exploration. *Proceedings of IASS*, (2018), Cambridge, MA.

3. Brown, N. and Mueller, C. Design variable analysis and generation for performance-based parametric modeling in architecture. *International Journal of Architectural Computing 17*, 1 (2019), 36–52.

4. Chatzikonstantinou, I. and Sariyildiz, I.S. Addressing design preferences via auto-associative connectionist models: Application in sustain architectural Façade design. *Automation in Construction 83*, (2017), 108–120.

5. Dogan, T. 2014. ARCHSIM. www.solemma.com.

6. Dogan, T. and Reinhart, C. Shoeboxer: An algorithm for abstracted rapid multi-zone urban building energy model generation and simulation. *Energy and Buildings 140*, (2017), 140–153.

7. Doraiswamy, H. et al. Topology-based catalogue exploration framework for identifying view-enhanced tower designs. *ACM Transactions on Graphics 34,* 6 (2015), 230.

8. Gao, L. et al. A data-driven approach to realistic shape morphing. *Computer Graphics Forum 32*, 2.4 (2013), 449–457.

9. Gerber, D.J. et al. Design optioneering: multi-disciplinary design optimization through parameterization, domain integration and automation of a genetic algorithm. *Proceedings of SimAUD,* (2012), 1–8.

10. Harding, J. and Shepherd, P. Meta-parametric design. *Design Studies 52*, (2016), 73–95.

11. Hardoon, D.R. et al. Canonical Correlation Analysis: an overview with application to learning methods. *Neural Computation 16*, 12 (2004), 2639–2664.

12. Mueller, C.T. and Ochsendorf, J.A. Combining structural performance and designer preferences in evolutionary design space exploration. *Automation in Construction 52*, (2015), 70–82.

13. Oxman, R. Performance-based design: current practices and research issues. *International Journal of Architectural Computing 6,* 1 (2008), 1–17.

14. Preisinger, C. 2014. Karamba. www.karamba3d.com.

15. Reynolds, D. et al. Structural design taste test. *Proceedings of IASS*, (2018), Cambridge, MA.

16. Schulz, A. et al. Retrieval on parametric shape collections. *ACM Transactions on Graphics 36,* 1 (2017), 1–14.

17. Sederberg, T.W. and Parry, S.R. Free-form deformation of solid geometric models. *SIGGRAPH '86 20*, 4 (1986), 151–160.

18. Shea, K. et al. Towards integrated performance-driven generative design tools. *Automation in Construction 14*, 2 (2005), 253–264.

19. Turrin, M. et al. Design explorations of performance driven geometry in architectural design using parametric modeling and genetic algorithms. *Advanced Engineering Informatics 25*, 4 (2011), 656–675.

20. Wortmann, T. 2017. Surveying design spaces with performance maps: a multivariate visualization method for parametric design and architectural design optimization. *International Journal of Architectural Computing 15*, 1 (2017), 38–53.

21. Wortmann, T. and Nannicini, G. 2017. Introduction to Architectural Design Optimization. In *City Networks – Collab. And Plan. for Health and Sustainability,* (2017), 259-273.

22. Wortmann, T. and Schroepfer, T. 2019. From Optimization to Performance-Informed Design. *Proceedings of SimAUD,* (2019), 261–268.

23. Xuereb Conti, Z. and Kaijima, S. A flexible simulation metamodel for exploring multiple design spaces. *Proceedings of IASS*, (2018), Cambridge, MA.

Performance-Based Facade Framework

Automated and Multi-Objective Simulation and Optimization

Mahsa Minaei[1], Ajla Aksamija[2],

[1]University of Massachusetts
Amherst, U.S.
mahsa@umass.edu

[2]University of Massachusetts
Amherst, U.S.

ABSTRACT

Buildings have a considerable impact on the environment, and it is crucial to consider environmental and energy performance in building design. Buildings account for about 40% of the global energy consumption and contribute over 30% of the CO_2 emissions. A large proportion of this energy is used for meeting occupants' thermal comfort in buildings, followed by lighting. The building facade forms a barrier between the exterior and interior environments, therefore it has a crucial role in improving energy efficiency and building performance.

In this regard, decision-makers are required to establish an optimal solution, considering multi-objective problems that are usually competitive and nonlinear, such as energy consumption, financial costs, environmental performance, occupant comfort, etc. Sustainable building design requires considerations of a large number of design variables and multiple, often conflicting objectives, such as the initial construction cost, energy cost, energy consumption and occupant satisfaction. One approach to address these issues is the use of building performance simulations and optimization methods.

This paper presents a novel method for improving building facade performance, taking into consideration occupant comfort, energy consumption and energy costs. The paper discusses development of a framework, which is based on multi-objective optimization and uses a genetic algorithm in combination with building performance simulations. The framework utilizes EnergyPlus simulation engine and Python programming to implement optimization algorithm analysis and decision support. The framework enhances the process of performance-based facade design, couples simulation and optimization packages, and provides flexible and fast supplement in facade design process by rapid generation of design alternatives.

Author Keywords

Performance-based facade design; building performance simulation; simulation-based optimization; energy efficiency; occupant comfort; Genetic Algorithm; Python script.

ACM Classification Keywords

B.5.2 Design Aids (Simulation & Optimization)

1 INTRODUCTION

The buildings and buildings construction sectors combined are responsible for 36% of global final energy consumption and nearly 40% of total direct and indirect CO^2 emissions (IEA 2019). Energy demand from buildings and buildings construction continues to rise, driven by improved access to energy in developing countries, greater ownership and use of energy-consuming devices, and rapid growth in global buildings' floor area, at nearly 3% per year (IEA 2019). A large proportion of this energy is used for meeting occupants' thermal comfort in buildings, followed by lighting. The building facade forms a barrier between the exterior and interior environments, and has a crucial role in improving energy efficiency and building performance. Therefore, this research focuses on performance-based facade design, appropriate simulation and optimization tools and methods for design analysis and support.

Building performance simulation (BPS) provides relevant design information by indicating potential (quantifiable) directions for design solutions. BPS tools and applications facilitate the process of design decision-making by providing quantifiable data about building performance. BPS tools are an integral part of the design process for energy efficient and high-performance buildings, since they help in investigating design options and assess the environmental and energy impacts of design decisions (Attia 2013). The important aspect is that simulation does not generate design solutions, instead, it supports designers by providing feedback on performance results of design scenarios.

Optimization is a method for finding a best scenario with highest achievable performance under certain constraints and variables. There are different methods for optimization, requiring use of computational simulation to achieve optimal solution, or sometimes requiring analysis or experimental methods to optimize building performance without

performing mathematical optimization. But in BPS context, the term optimization generally indicates an automated process that is entirely based on numerical simulation and mathematical optimization (Nguyen 2014). Integrating BPS and optimization methods can form a process for selecting optimal solutions from a set of available alternatives for a given design problem, according to a set of performance criteria.

2 BACKGROUND

Previous related research regarding building envelope and simulation-based optimization are discussed below. Application of computational optimization methods in sustainable building envelope design with focus on residential retrofits was reviewed in detail by Envis (2013). Simulation-based optimization methods applied to buildings with focus on handling discontinuous multi-model optimization problem was discussed by Nguyen (2014). Building design optimization techniques for real-world design challenges, as well as advances and obstacles were described in this study (Nguyen 2014). Optimization algorithms in building design was studied by Machairas (2014). The study reviewed methods and tools used for building design optimization, exploring reasoning behind the selection process, abilities and key characteristics (Machairas 2014). Energy efficient building design optimization from the architectural perspective was reviewed by Shi (2016). Analyzed subjects includes general procedures, origin and development, classification, design objectives and variables, energy simulation engines, optimization algorithm and the applications (Shi 2016). Literature review regarding optimization for building envelope design over the last 30 years was reviewed by Huang and Niu (2016). Numerous studies on optimization of building envelope design were assembled and reviewed, and targeted objectives were collected and summarized (Huang and Niu 2016). Building simulation and computational techniques were discussed by Wang and Zhai (2016). This study focused specially on six different topics, including ventilation performance prediction, building energy and thermal load simulation, lighting and daylight modeling, building information modeling, indoor acoustic simulation, and life cycle analysis of buildings. Major advances in each are highlighted, as well as trends for development and applications (Wang and Zhai 2016). Ostergard reviewed building simulation methods in early building design, discussing developments in academia and commercial software industry that focus on time-consuming modeling, rapid change of the design, conflicting requirements, input uncertainties and large design variability (2016). Building energy modeling for control and operation was reviewed by Li and Wen (2016). An effort was aimed to develop a zero-energy building design support tool (ZEBO) that facilitates the use of building performance simulation in the early design stage in hot climate (Attia 2013). However, there are not that many studies that focus on integration of simulation and optimization methods, specifically for building envelope design. This was the driving factor behind the research discussed in this paper.

2.1 Current Gaps in Research and Literature

A limited number of studies have focused on the performance-based design process for building facades which integrate simulation and optimization methods. There is a lack of workable framework that implements both simulation analysis and optimization methods for facade design, taking into account performance criteria specific to this building system. High-performance facades require integration of building performance simulations during the design process, because it is essential to quantify the effects of different design decisions on building envelope performance (Aksamija 2013). The future performance-based design approaches and simulation tools for facades should increase effectiveness, speed, quality, assurance and users' productivity.

Energy modeling and simulations in design process are usually limited to analysis of few different scenarios. It is not possible to simulate and analyze all possible design scenarios because of time constraints. Therefore, this research focused on developing a framework that couples simulation and optimization processes and allows multiple design scenarios to be tested rapidly. The framework was implemented by coupling Python scripting with EnergyPlus simulation engine, enabling users to consider more variables during the design process.

2.2 Benefits of Data-Driven Framework

The basic characteristics that differentiate the developed framework and improve decision-making process can be summarized as:

- Automation and speed: The framework enables users to automatically send the design scenarios to simulator and gather the outputs, and then screens out and sorts these outputs to find optimized results. The advantages of this automated process are efficient testing methodology, consistency, reliability and increase in the number of possible design scenarios. Also, by implementing this framework, simulation time will be decreased for thousands of design scenarios.

- Variety of variables (multi-objective variables): This framework enables users to test multiple variables at the same time during the design process.

- Modularity: The framework is designed in multiple modules, which work independently. The key benefits of modularity in this framework are distinct functionality and manageability. Each module provides a distinct function and can be combined to provide entirely new collective function. The separate modules make it easier to test and implement this framework in design process or detect the errors.

3 METHODOLOGY: DEVELOPMENT OF FRAMEWORK FOR PERFORMANCE-BASED FACADE DESIGN

The new framework for performance-based design approach, aiming to minimize building energy consumption and energy cost while considering occupants' comfort level, was developed as part of this research. This is a modular framework, consisting of independent scripts that represent modules, steps and functions of application under test. The modules are used in a hierarchical fashion to apply the framework, consisting of four steps:

1) Defining goals, performance criteria, facade design variables and their properties, acceptable ranges for high-performance facade design

2) Generating the database that includes all possible design scenarios based on the variables with permutation in Python and selected outputs after simulation in EnergyPlus. This is module 1.

3) Coupling Python script with simulation engine (EnergyPlus) to automatically perform simulations for scenarios from database (measurements methods) to quantify variables and generate the needed outputs. This is module 2.

4) Filtering and narrowing down the results by implementing Python script, genetic algorithm (GA) and machine learning to evaluate outputs and find the optimal scenarios. This is module 3.

The next sections discuss the components of the framework and its implementation in detail.

3.1 Step1: Defining Goals, Performance Criteria and Facade Variables

Figure 1 shows the components of the framework. Performance-based facade design requires a holistic approach, considering performance indicators, such as energy performance, occupant comfort and energy cost. These performance requirements (variables) must be quantified. The goals (objective function) for this framework are to aid the design decision making process, where energy consumption and cost are minimized, and occupant comfort (thermal and visual) is maximized. The energy requirements for heating, cooling, and lighting of buildings are strongly driven by the performance of the facade, especially glazing parts. The objectives for reducing energy consumption are to reduce heating, cooling and lighting loads. Performance requirements (variables) to meet this objective are window to wall ratio (WWR), wall assembly, insulation, solar control, and glazing system. Performance-based facade
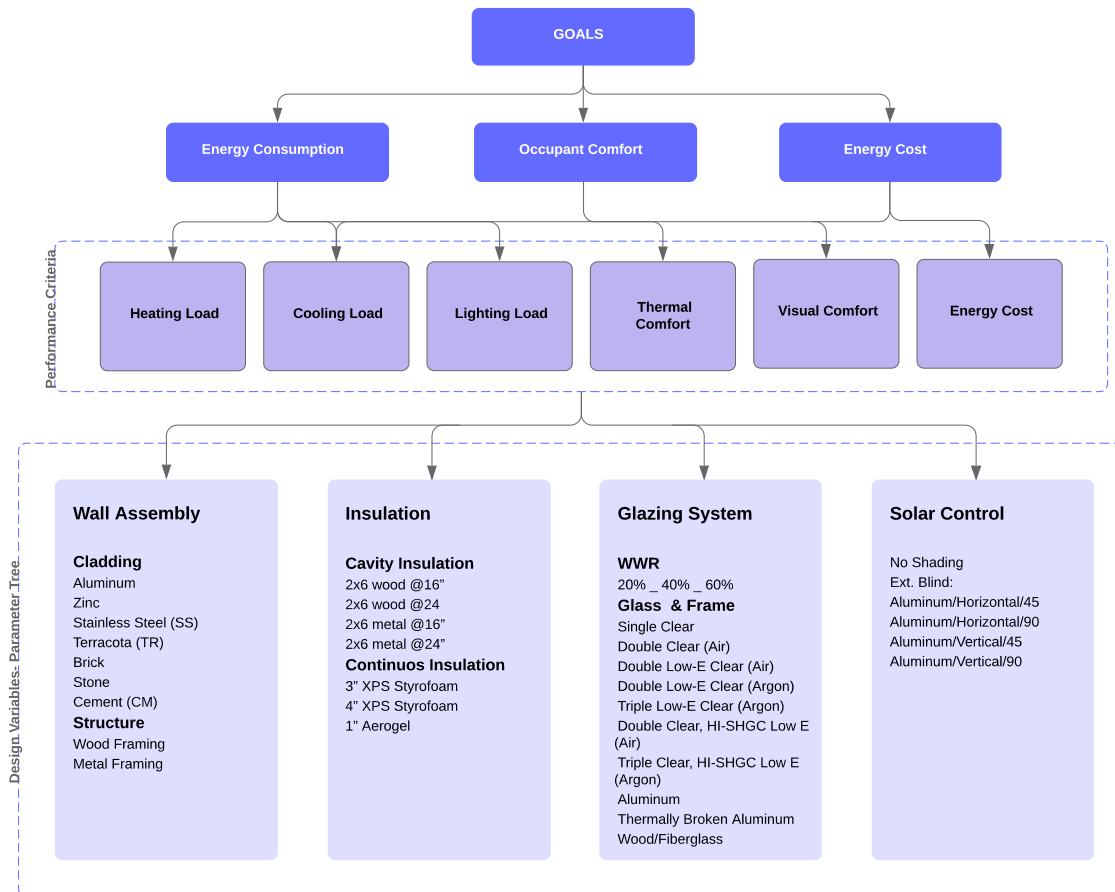
Figure 1. Conceptual diagram, showing components of the framework design objectives that are related to human factors and contribute to occupant comfort and satisfaction in buildings

481

include thermal comfort and visual comfort. The variables that relate to facade design include: air temperature, mean radiant temperature, air movement, relative humidity, clothing levels and activity levels. The predictive mean vote (PMV) suggested by Fanger (1970) predicts the effects of these six factors on thermal comfort. Predicted Percentage of Dissatisfied (PPD) persons predicts the percentage of people who would feel discomfort with certain thermal conditions.

3.2 Creating the Database

After setting variables and parameters for facade design, all possible scenarios are generated using Python programming. With permutation in Python script, design scenarios are generated and added to database with specific scenario ID. In this study, we have 38,400 scenarios to investigate for the test cell, described in the next section. After running simulation in EnergyPlus, all outputs in step 3 are populated in this database with identical scenario ID. EnergyPlus provides wide range of outputs, but for this purpose, the following results are obtained: cooling, heating and lighting loads, Energy Use Intensity (EUI) for electricity and gas, PMV and PPD, and total energy costs for electricity and gas. Module 1 is responsible for generating all scenarios with defined variable and populating these scenarios in database. Module 2 is responsible for sending automatically these scenarios to simulation engine and for populating the selected outputs in the database. Data Flow Diagram (DFD) in Figure 2 shows the overview of the framework system that represent the flow of data through this process.
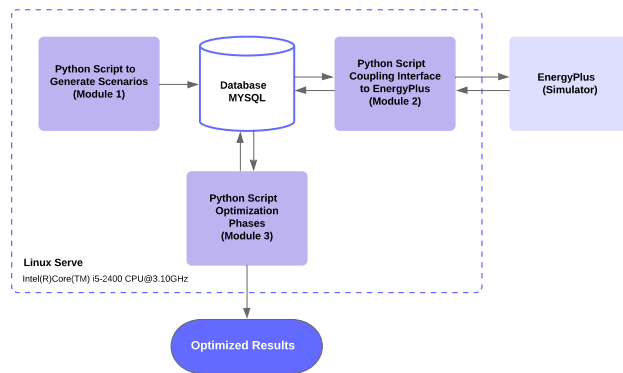
Figure. 2. Data Flow Diagram for the framework.

3.3 Step 3: Coupling Python Script with Simulation Engine

EneryPlus 8.5 is used in this research as an energy modeling engine. EnergyPlus has been chosen as BPS tool for two main reasons: (a) this program allows reliable modeling of both building and HVAC systems, and, (b) it works with text-based inputs and outputs, and these facilitate the interaction with Python scripts. EnergyPlus can investigate discussed variables as inputs and simulate envelope related outputs in the study. Thermal comfort is calculated based on PMV and PPD. The formulas for both PMV and PPD are

built into EnergyPlus and their values can be obtained directly from the simulation output file.

Initial simulation test cell considered a single office space (40'x40'x10'), located in Atlanta, Georgia. The south-facing facade was used to develop different design scenarios, varying WWR, cladding materials, Insulation materials, glazing system and shading control. Defining related parameters as inputs and setting data needed for outputs are the primary method for connecting design scenarios in the database with the simulation engine. Python script works as an interface to call scenarios from database and to send them to simulator. Each parameter must identify a well-defined relation with discussed variables, which reveals facade behavior in relation to performance aspects being analyzed.

3.4 Step 4: Filtering and Narrowing Down the Results by Implementing Python Script, GA and Machine Learning

In building optimization studies using GA generally reduce the computational time by two methods: use very simplified model instead of complete simulation. This simplification can cause inaccurate modeling of building. The other method is select very small size for GA populations or relatively small number of generation. But one efficient solution to reduce the computational time associate with GA algorithm is use of machine learning techniques to reduce time and increase the accuracy in the results. The machine learning used in this research is a combination of Batch normalization which is an Artificial Neural Network (ANN) technique and flood fill algorithm. ANNs are effective methods that imitate the complex relationship of the network to solve multi objectives and non-linear problems. ANNs resemble the biological neural system, composed of layers of parallel neurons and weighted links. They learn the relationship between the input and output variables by studying previously-recorded data.

The optimization method in this study is a combination of GA and machine learning. The GA in combination with flood fill algorithm and batch normalization create a new technique to find a relation between the outputs, to assign weights and dynamically adjust the target position. For this framework, three indices are defined for consumption, comfort and cost as indicators. Indicators are combined values that are used to measure performance, achievement or the impact of changes.

Batch normalization technique is used for the first phase of optimization. This is a technique for improving speed, performance and stability of artificial neural network (ANN) by adjusting and scaling the activations. The batch normalization was introduced in 2015 (Loffe and Szegedy 2015). The intention behind batch normalization is to optimize network training. Several benefits of this methods are: faster training, higher learning rates, reduced sensitivity, easier methods to initialize and produce better results (Hinton 2012). This technique, combined with flood field

algorithm, facilitates the optimization by sorting the highest indicators and decides which scenarios have to be simulated.

The flood field algorithm takes three parameters: start node, target and replacement, and determines the area connected to our target. This algorithm facilitates the optimization by sorting the highest indicators and decides which scenarios have to be simulated, based on the specific scenario ID. Using this algorithm decreases the process time, because it is not necessary to simulate all scenarios—rather, only scenarios that are closer to the target. The comparison is based on the assigned indicator value. In dynamic system, it is necessary to scale indicators to represent the impact of the indicators, so as to configure following tasks, and converge the results to the goal based on these scores. Figure 3 shows a sample for scoring total EUI electricity, EUI gas, PMV and energy cost indicators.

These indicators work as fitness functions in genetic algorithms which is particular type of objective function to summaries and guide the simulations towards optimal design solutions. These indicators or fitness functions must correlate closely to the goals and must be computed quickly. Because it needs to be iterated many times in order to produce usable results. The initial population is generated randomly, based on the range of possible design scenarios. It is sent to the simulator to run the initial calculations, and then results are returned to the database to compare with the goals and standards. Then, design scenarios that have results closer to the goals are kept, and others are removed. In this framework, goal is summation of three indicators, for energy consumption, comfort and cost.
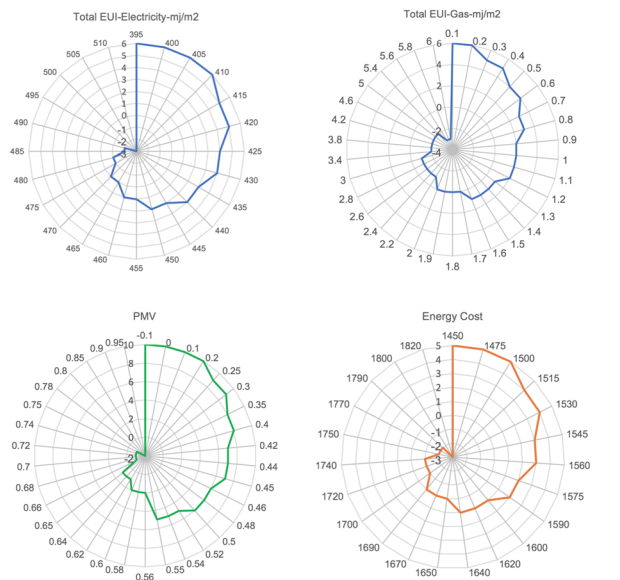
Figure 3. Total EUI-Electricity (MJ/m^2), EUI-Gas, PMV and Energy Cost indicator scores.

The indicators are dynamically updated based on the range of results. The top left chart in Figure 3 shows an example,

where indicators from 6 to -3 are used for the initial test cell energy consumption results. Occasionally, the solutions may be "seeded" in areas where optimal solutions are likely to be found. Individual solutions are selected through a fitness-based process, where fitter solutions (as measured by a fitness function) are typically more likely to be selected. This method accelerates the simulation process and the results give us clusters of optimized scenarios for analysis in next phase of optimization. The whole optimization process is illustrated in Figure 4.
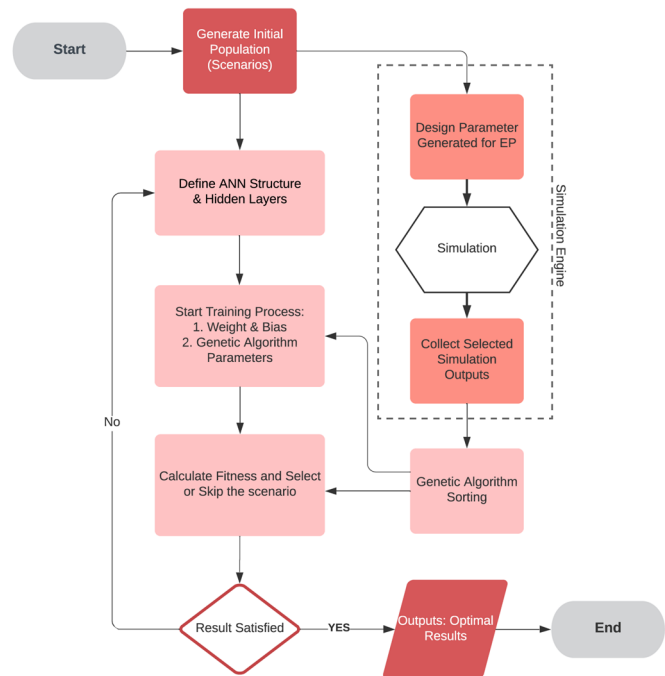

Figure 4. Optimization Flow chart Illustrating the Application of GA and ANN.

Figures 5 and 6 show how optimization algorithm selects and sorts the fitted results for this framework. Figure 5 shows the results before applying optimization for 2,061 scenarios and Figure 6 shows the result of 18,103 scenarios with assigning the first step of optimization. In this case, we have 1,627 scenarios that scored 20 and more than 20 (1,591 scenarios at 20 and 36 more than 20). Using this process decreases the process time, because it is not necessary to simulate all scenarios-rather, only scenarios that are closer to the target. After running all scenarios (38400 ID) with applying batch normalization technique, 3,164 scenarios are selected. Next step focuses on comparison of results.
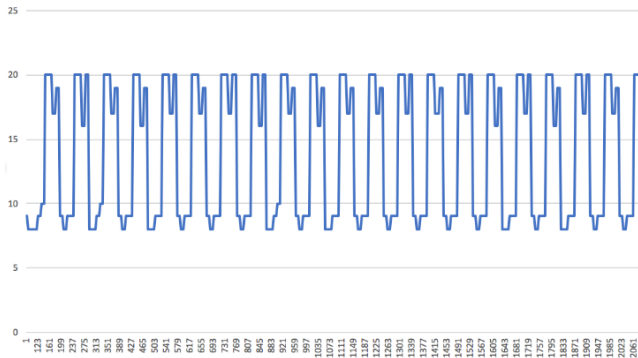
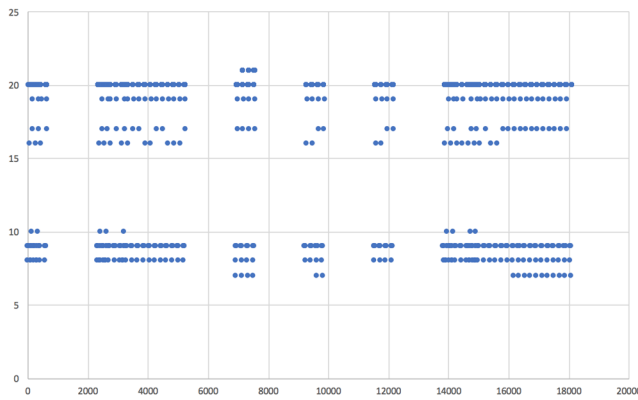**Figure 5:** Total Indicators vs. Scenario ID (for 2,061 scenarios)

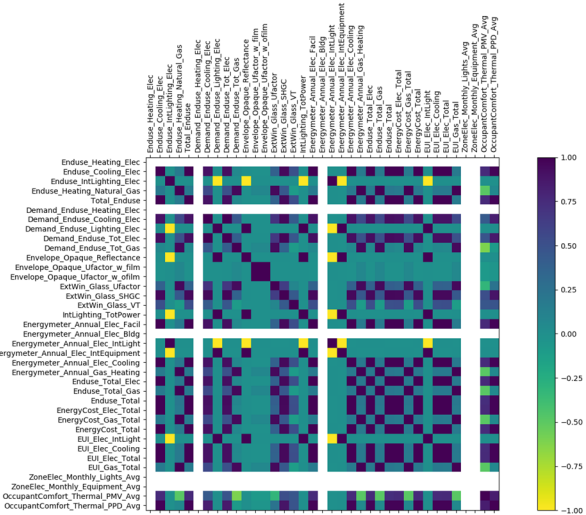**Figure 6:** Total Indicators vs. Scenario ID (for 18,103 scenarios

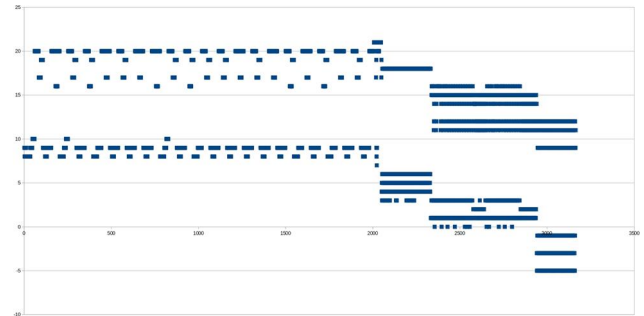**Figure 7:** Correlation Matrix

Figure 8. Results for all scenarios with applying batch normalization and flood field algorithm (Phase 1 optimization)
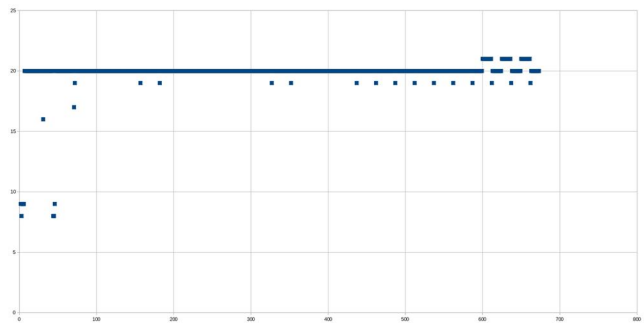

Figure 9. Results for all scenarios with applying correlation matrix and cluster eliminating (Phase 2 optimization)

Next step of optimization is applying integrated correlation matrix clustering as dropout technique, then comparing the results. This dropout is a regularization technique for reducing overfitting in neural networks by preventing complex co-adaptations on training data. Dropout refers to dropping out units (hidden or visible) in neural network. In machine learning, correlation clustering provides a method for clustering a set of objects into optimum number of clusters based on the similarity. So, in correlation matrix, the relationship between the objects (variables) are known instead of the actual representations of the objects. Figure 7 shows the correlation matrix based on output data, integrated with optimization method to sort the results. Figures 7 and 8 show the final results of all scenarios with both techniques implemented. Figure 8 represent the first phase of optimization, and Figure 9 shows second phase after applying correlation matrix. Results show that process time, performance and accuracy are improved by using this method.

4. CONCLUSION

This paper discussed the role of simulations and optimization in design decision-making process. Then, a novel performance-based facade design framework was described, where different performance criteria and variables have been defined for achieving energy efficiency, occupant comfort and cost optimality. The framework has been implemented

by coupling EnergyPlus as a simulation engine, and custom scripts using Python programing language. The paper describes the components and functionality of this framework in detail, as well as two-step optimization technique. A case study for a test cell was presented, illustrating how the framework is used to test a variety of design possibilities. Future research will focus on testing and validation of this framework, as well as its application for facade design in different climates and developing the user interface. In addition to developing the user interface and web application this product will develop to take any IDF files then users are enabled to choose their variable for optimization and the rest of the process will be automated and results will be represented. Other important development for this research after release the open source application, is collecting the data for implement deep learning on these data and outputs form different iterations so enable the correlation matrix to generate automatically.

It is important to note that the inputs can be all design and environmental related variables with regards to their significance in design decision-making. In real-life situations, however, architects may assign more weight to some design inputs or environmental impacts because of design priorities, or in order to overcome the conflicting effects of inputs or impacts, so future methodologies and tools should follow this research capability to assign weight to factors in their decision-making support process.

A future direction of the research in this area could be to develop a software tool that can help conduct the entire methodology within one tool. Such tools need to include all design variables and construction inputs beyond what is studied in this research. The variables such as building systems and other component of the building in design phase also environmental inputs can be evaluate it this tool and researches for decision making.

REFERENCES

1. Aksamija, Ajla. *Sustainable facades: Design methods for high-performance building envelopes,* John Wiley and Sons, 2013.

2. Attia, Shady, Mohamed Hamdy, William O'Brien, and Salvatore Carlucci. "Assessing gaps and needs for integrating building performance optimization tools in net zero energy buildings design." *Energy and Buildings* 60 (2013): 110-124.

3. Clarke, Joseph Andrew, and J. L. M. Hensen. "Integrated building performance simulation: Progress, prospects and requirements." *Building and Environment* 91 (2015): 294-306.

4. Crawley, Drury B., Jon W. Hand, Michael Kummert, and Brent T. Griffith. "Contrasting the capabilities of building energy performance simulation programs." *Building and Environment*
43, no. 4 (2008): 661-673.

5. Evins, Ralph, Philip Pointer, and Stuart Burgess. "Multi-objective optimization of a modular building for different climate types." In *First Building Simulation and Optimization Conference* (2012)*,* Loughborough, pp. 173-180.

6. Evins, Ralph. "A review of computational optimisation methods applied to sustainable building design." *Renewable and Sustainable Energy Reviews* 22 (2013): 230-245.

7. Fanger, Poul O. "Thermal comfort. Analysis and applications in environmental engineering." *Thermal comfort: Analysis and applications in environmental engineering.* (1970).

8. Gero, John S., Neville D'Cruz, and Antony D. Radford. "Energy in context: A multicriteria model for building design." *Building and Environment* 18, no. 3 (1983): 99-107.

9. Gosavi, Abhijit. *Simulation-based optimization*. Berlin: Springer, 2015.

10. Hensen, Jan LM, and Roberto Lamberts, eds. *Building performance simulation for design and operation*. Routledge, 2012.

11. Hinton, Geoffrey E., Nitish Srivastava, Alex Krizhevsky, Ilya Sutskever, and Ruslan R. Salakhutdinov. "Improving neural networks by preventing co-adaptation of feature detectors." *arXiv preprint arXiv:1207.0580* (2012).

12. Huang, Yu, and Jian-lei Niu. "Optimal building envelope design based on simulated performance: History, current status and new potentials." *Energy and Buildings* 117 (2016): 387-398.

13. Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." *arXiv preprint arXiv:1502.03167* (2015).

14. Li, Xiwang, Jin Wen, and Er-Wei Bai. "Developing a whole building cooling energy forecasting model for on-line operation optimization using proactive system identification." *Applied Energy* 164 (2016): 69-88.

15. Machairas, Vasileios, Aris Tsangrassoulis, and Kleo Axarli. "Algorithms for optimization of building design: A review." *Renewable and Sustainable Energy Reviews* 31 (2014): 101-112.

16. Nguyen, Anh-Tuan, Sigrid Reiter, and Philippe Rigo. "A review on simulation-based optimization methods applied to building performance analysis." *Applied Energy* 113 (2014): 1043-1058.

17. Ostergard, Torben, Rasmus L. Jensen, and Steffen E. Maagaard. "Building simulations supporting decision making in early design–A review." *Renewable and Sustainable Energy Reviews* 61 (2016): 187-201.

18. Shi, Xing, Zhichao Tian, Wenqiang Chen, Binghui Si, and Xing Jin. "A review on building energy efficient design optimization rom the perspective of architects." *Renewable and Sustainable Energy Reviews* 65 (2016): 872-884.

19. Wang, Haidong, and Zhiqiang John Zhai. "Advances in building simulation and computational techniques: A review between 1987 and 2014." *Energy and Buildings* 128 (2016): 319-335.

20. Zhang, Yi. "Parallel EnergyPlus and the development of a parametric analysis tool." In *11th Conference of International Building Performance Association (IBPSA),* Glasgow, UK, July, pp. 27-30. 2009.

Modeling and Simulation of a Municipal Solid Waste Management System based on Discrete Event System Specification

Chang-Hyun Lyoo[1], Jinho Jung[1], Changbeom Choi[1] and Eun-Young Kim[2]

[1]Handong Global University, Pohang, Rep. of Korea, {21300259, 21300704, cbchoi}@handong.edu
[2]Pohang TechnoPark, Pohang, Rep. of Korea, hellosally@daum.net

ABSTRACT

The cleanliness of residential areas is a critical consideration for urban planning. In particular, municipal waste management is a vital service that affects the satisfaction level of residents. Since the satisfaction level may be affected by the amount of garbage accumulated at the time residents put out their garbage, an urban planner may consider residents' living patterns to develop an efficient waste collection strategy. As the composition of residents in a residential area becomes more complicated, estimating the accumulation of garbage based on the living patterns of residents so as to achieve a high level of residents' satisfaction with public services becomes an increasingly difficult problem for urban planners. This research focuses on the representation of the temporal behavior patterns of residents in a given area using a discrete event system. Additionally, this research introduces a simulation environment that utilizes the discrete event system specification formalism to simulate the living patterns of residents and to analyze waste generation and the municipal waste management system. The simulation results show that modeling the living patterns of residents based on their occupations may yield a satisfactory representation of the dynamics of a waste management system.

Author Keywords

Agent-based Simulation; Discrete Event System Formalism; Municipal Solid Waste Management System.

ACM Classification Keywords

I.6.5 : SIMULATION AND MODELING (Model Development).

1 INTRODUCTION

Today, almost 55% of the world's population lives in urban areas, and the urban population is continuing to increase [1]. Among the many issues that arise in an urban environment, the problem of garbage is most closely related to residents' satisfaction [2]. As the population increases, the total amount of garbage generated in a local area may increase. Consequently, the satisfaction level of residents may decrease when the government does not collect the garbage on time. Moreover, unsatisfied citizens may file civil complaints, leading to an increase in the workload of government employees and higher social costs. To address these issues, urban planners or employees of the local government analyze residential areas to estimate the corresponding amounts of waste generation and ensure the provision of appropriate public services, such as utilizing garbage trucks to collect waste. Despite such efforts, however, it is difficult to estimate waste disposal patterns because the behavior patterns of the residents are more important factors than the geographic or demographic characteristics of the residential area. Notably, residents' living patterns may differ depending on the cultural and occupational characteristics of different residents; consequently, capturing these living patterns is a difficult problem.

This research proposes a modeling method focusing on human characteristics and behavior patterns to help optimize municipal waste management systems. Since the temporal behaviors of residents may differ based on their living habits and occupations, this research utilizes the discrete event system specification (DEVS) formalism to model residents' behavior. The DEVS formalism is a set-theoretic framework with modular characteristics [3, 4, 5]. This allows users to develop their own simulation models to serve as components of an urban system or to reuse simulation models built by other practitioners. To maximize the modular capabilities of the DEVS formalism, this research proposes an object-oriented discrete event simulator based on the Python programming language. The simulator captures the common behavior of the residents of an urban area using the DEVS formalism and specializes in analyzing residents based on their occupations. Therefore, practitioners may reuse this DEVS simulation model to simulate the temporal behavior of residents while differentiating residents by their occupations.

2 RELATED WORK AND BACKGROUND

This section introduces and analyzes previous related studies. Additionally, it provides background knowledge to help understand the proposed method and the simulation environment.

2.1 Related Work: Waste Management

Many studies have pursued the optimization of waste management through the application of both analytical solutions and simulation techniques. For example, Dao-Tuan et al. (2017) used integer linear programming and mixed-integer linear programming to optimize the routes of garbage trucks considering emissions control [6].

Some work has been done to optimize the routes of garbage trucks to reduce waste collection costs using agent-based simulations and geographic information system (GIS) techniques. Chalkias et al. (2009) constructed a spatial geodatabase by integrating GIS and municipality databases [7]. They proposed a method of deriving optimal paths for garbage trucks using Dijkstra's algorithm. Das et al. (2015) solved the waste transportation problem by optimizing waste collection and transportation paths with the aid of the traveling salesman problem to develop a time- and cost-effective waste management system design [8]. Nambiar et al. (2013) compared the static and dynamic vehicle routing problems, with a focus on increasing the capacity utilization of each participating vehicle [9]. Nguyen et al. (2013) proposed a model to optimize the collection and transportation of municipal solid waste in terms of fuel consumption and pollutant emissions [10]. Shi et al. combined hybrid agent-discrete modeling techniques with a Voronoi-based genetic algorithm and a kth nearest neighbor search algorithm to build an optimization model [11]. Additionally, they recommended vehicle routing and resource allocation strategies for solid waste management systems. Karadimas et al. (2006) suggested the loose coupling of GIS and simulation software in the domain of municipal waste management [12]. Hua et al. (2016) proposed that applying GIS data and real-time waste data from smart devices to solve routing problems would reduce the travel distance of the collection vehicles [13]. System dynamics approaches have also been used in forecasting waste generation and simulating an efficient waste management system [14, 15, 16]. Dyson et al. (2005) predicted waste generation using 5 different models [15]. Guariso et al. (2009) predicted the limit on the performance of a waste transfer station in Australia [14]. Wang developed a waste prediction and response model based on system dynamics [16].

However, to the best of the authors' knowledge, there has been little discussion of the modeling of human behavioral patterns or cultural characteristics in the context of waste management. This study proposes a modeling method to capture the behaviors of residents using the DEVS formalism.

2.2 Background: Discrete Event System Specification (DEVS)

The DEVS formalism is a set-theoretic formalism developed for specifying discrete event systems. The DEVS formalism comprises two types of models to represent discrete event systems: *atomic models* and *coupled models*. There are two ways to build a model using the DEVS formalism. First, the user may specify the behavior of some indivisible component of the discrete event system using an atomic model. On the other hand, the user may compose a system represented by

a coupled model by assembling multiple atomic models or other coupled models.

In this research, an atomic model specification is adopted to represent the behavior of the simulated entities in an urban simulation to capture the dynamics of the waste management system and garbage disposal patterns. Since the proposed urban simulation environment manages various simulation models and the connections between them, this simulation environment is a coupled model. However, the user need not consider the specification of the coupled model because the simulation environment manages the insertion of and connections between different entities in the urban simulation. The atomic model specification is defined as follows:

$$AM = <X, Y, S, \delta_{ext}, \delta_{int}, \lambda, ta>$$

where

X: a set of external input event types;
Y: an output set;
S: a sequential state set;
δ_{ext}: $Q \times X \to S$, an external transition function,
  where Q is the total state set of
  $M = \{(s, e) | s \in S \text{ and } 0 \leq e \leq ta(s)\}$;
δ_{int}: $S \to S$, an internal transition function;
λ: $S \to Y$, an output function;
ta: $S \to \mathcal{R}_{0,\infty}^+$, a time advance function,
  where $\mathcal{R}_{0,\infty}^+$ is the set of non-negative real
  numbers up to ∞.

An atomic model AM is a model that is affected by external input events X and generates output events Y based on the model state. The state set S represents the unique description of the model. The internal transition function δ_{int} and the external transition function δ_{ext} are used to compute the next state of the model. If an external event arrives after an elapsed time e that is less than or equal to the value $ta(s)$ specified by the time advance function ta, then the new state s' is computed using the external transition function δ_{ext}. Subsequently, a new $ta(s')$ is computed, and the elapsed time e is set to zero. Otherwise, when an internal event arrives at $ta(s)$, the new state s' is computed using the internal transition function δ_{int}. In the case of an internal event, the output specified by the output function λ is produced based on the state s, which means that the output function is processed before the internal transition function. Then, as before, a new $ta(s')$ is computed, and the elapsed time e is set to zero.
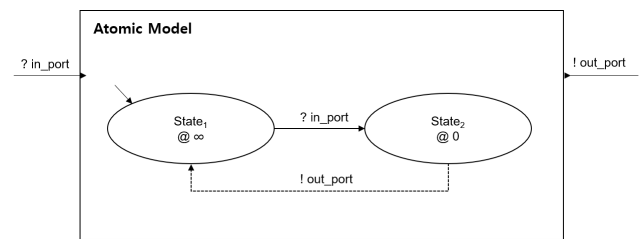

Figure 1. Example of an Atomic Model

Figure 1 shows a diagram of an atomic model. The label starting with a question mark denotes the input port, and the

label starting with an exclamation mark denotes the output port. The ellipses represent the states of the model, and the ellipse indicated with an arrow denotes the initial state of the model. Each state is associated with a state name and a positive real number that represents the occupancy time. When the elapsed time is equal to the occupancy time, the simulation algorithm will sequentially trigger the output function and the internal transition function.

The proposed simulator partially relies on the notation of the DEVS formalism and its simulation algorithm. Listing 1 shows the Python code for the example atomic model. As shown in Listing 1, the proposed simulator provides a programming interface that has a one-to-one correspondence with the atomic model specification of the DEVS formalism.

```
1  class AtomicModel(BehaviorModelExecutor):
2    def __init__(self,instance_time,
       destruct_time,name,engine_name):
3      BehaviorModelExecutor.__init__(self,
       instance_time,destruct_time,name,
       engine_name)
4
5      self.init_state("State1")
6      self.insert_state("State1",Infinite)
7      self.insert_state("State2",0)
8
9      self.insert_input_port("in_port")
10     self.insert_output_port("out_port")
11
12   def ext_trans(self,port,msg):
13     if port == "in_port":
14       self._cur_state = "State2"
15
16   def output(self):
17     name = self.get_name()
18     msg = SysMessage(name,"out_port")
19     return msg
20
21   def int_trans(self):
22     if self._cur_state == "State2":
23       self._cur_state = "State1"
```

Listing 1. Example Code for an Atomic Model

3 PROPOSED METHOD AND ENVIRONMENTS

This section introduces the modeling results for the simulated entities of the municipal solid waste management system and its simulation environment. In this research, we propose two methods for modeling a municipal waste management system. First, we separate the domain-related part and the simulation-related part. Second, we adopt an object-oriented concept to allow practitioners to select entities to be simulated and synthesize the desired simulation based on the target situation. In this way, experts with various levels of understanding may participate in simulation development and experimentation.

To capture the dynamic behavior of the residents, we classify the residents into different types and model them using the DEVS formalism. The parameters required for simulation are modeled by defining a corresponding class in the Python programming language. Additionally, simulated entities that are modeled using an atomic model inherit the properties of

the base class for discrete event simulation. Each class instance of each resident type and the discrete event simulation model are synthesized during the initialization phase of the simulation process.

As an example, Listing 2 shows the code for the resident-type class for a student. The code returns information on the resident's behavior, such as the wake-up time, sleep time, and amount of waste generated. Each piece of information is based on a survey titled "What time did you wake up today?" conducted by Gallup Korea in 2013 [17].

```
1  class Student(HumanType):
2    def __init__(self,_id):
3      HumanType.__init__(self,_id)
4
5    def get_type(self):
6      return "Student"
7
8    def get_wakeup(self):
9      return TimeStruct(7,58,Statistic(0,0,1))
10
11   def get_sleep(self):
12     return TimeStruct(24,51,Statistic(0,0,1))
13
14   def get_out(self):
15     return self.get_wakeup() + 1
16
17   def get_in(self):
18     return TimeStruct(21,00,Statistic(0,0,1))
19
20   def get_trash(self):
21     return 0.3
22
23   def get_satisfaction_func(self,trash):
24     if trash >= 0.8:
25       return -10
26     elif trash <= 0:
27       return 20
28     elif trash < 0.8:
29       return 10
```

Listing 2. Example Code for the Resident-Type Class for a Student

Figure 2 shows the overall framework of the proposed simulation environment. The solid rectangles represent real models of the behavior of simulated entities instantiated in the simulation environment. Each real model implements the atomic behavior of the corresponding simulated entity. On this basis, an urban planner may compose various simulation models of interest. For example, some less sensitive residents may not file complaints to the government. Accordingly, the urban planner may use only the basic resident model to model such a less sensitive resident rather than modifying the parameters of the simulated entity. The dotted rectangles represent conceptual models. Each conceptual model may comprise several real models. When a practitioner inserts real models into the simulation environment, the simulation environment connects the ports of the newly introduced entities to those of other entities.

Figure 3 shows a diagram of the resident model used to model a resident of the residential area. The resident model can calculate the departure time of a resident based on the wake-up time of the corresponding resident-type instance, as shown
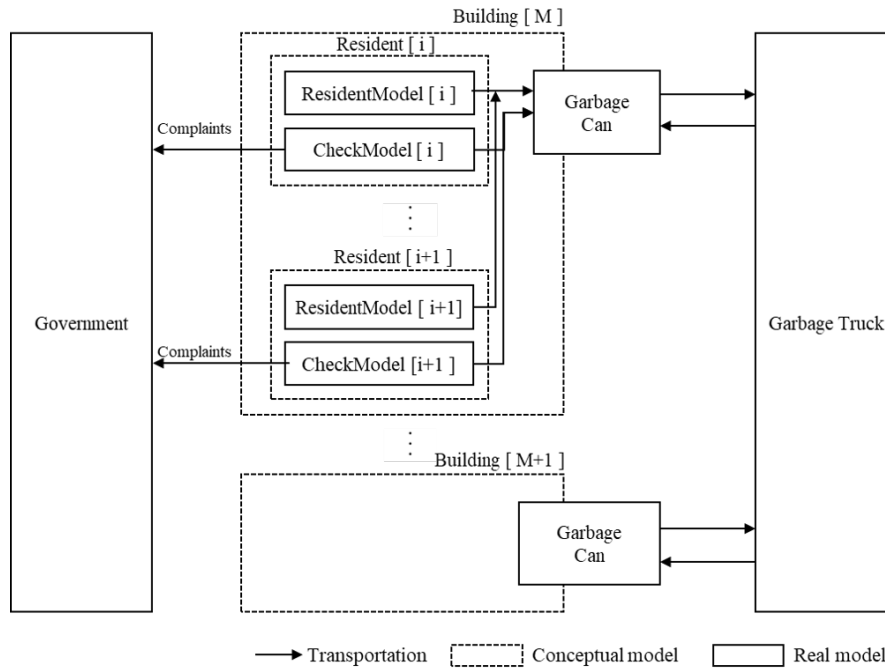
Figure 2. Atomic Model for Calculating the Satisfaction Level of a Resident

in Listing 2. The practitioner may reuse this resident model for different resident-type object instances. Notably, this research assumes a normal distribution for the departure time to support stochastic simulation. During the simulation, the resident model may generate a trash event with a given amount of waste generation from a particular resident-type instance and send it to the corresponding instances of the family model and check model.
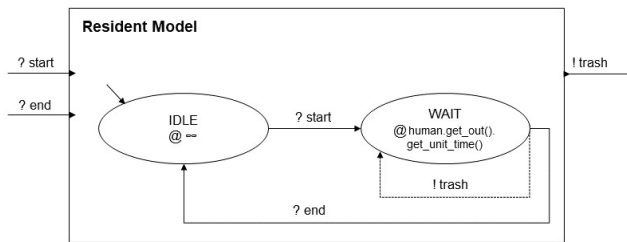
Figure 3. Atomic Model of a Resident
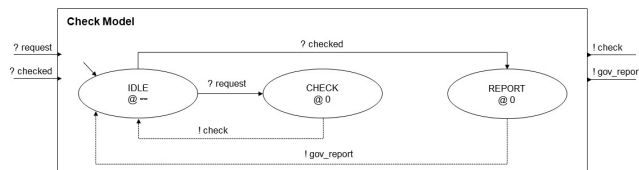
Figure 4. Atomic Model for Checking the Garbage

Figure 4 shows a diagram of the check model. This is the model used to measure the satisfaction level of a resident in accordance with the capacity of the garbage can. Since most

residents do not check the garbage during the daytime, the check model checks the garbage can at the departure time of the resident. Therefore, the check model checks the garbage can periodically upon departure events. Since the resident model generates garbage events based on the departure time, the check model utilizes these events to handle departure events. As a result, the level of satisfaction of a resident with the cleanliness of the garbage collection site is measured every departure time. When the check model receives a request message from the resident model, a check message is sent to the garbage can model to check the status of the garbage can. The check model then passes the value received from the garbage can through the check port to the resident-type instance. Each resident-type class has a function that calculates the satisfaction level based on the accumulation ratio of the garbage can. After checking the satisfaction level, the check model may trigger a claim message to the government model through the gov-report port if the satisfaction level is below a set threshold.
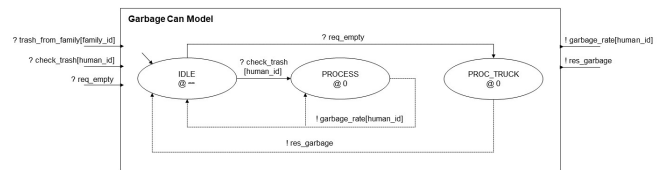

Figure 5. Atomic Model for Calculating the Garbage Accumulation of a Building

Figure 5 shows a diagram of the garbage can model. The garbage can model models the temporary garbage collection area for each building. Each instance of the garbage can model may have a different amount of free space during the

simulation. If residents dispose of a large amount of garbage in the garbage can or a garbage truck does not collect sufficient garbage from the garbage can, some garbage may remain in the garbage can. The remaining garbage may affect the satisfaction level of a resident. Therefore, the practitioner may observe changes in residents' satisfaction depending on the garbage can capacity and estimate a suitable capacity according to the residents. When the garbage can model receives an event from the $trash\_from\_family$ port, the garbage can is filled with an amount of garbage equal to the given value. When the garbage can model receives an event from the check model, the garbage can model calculates the accumulation ratio of the garbage can and sends the result back to the check model. When the garbage can model receives a message from the garbage truck model, the accumulated amount of garbage is sent to the garbage truck model.
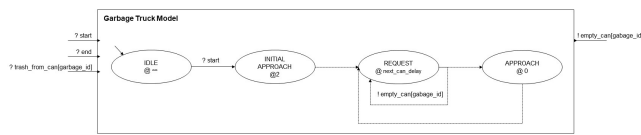
Figure 6. Atomic Model for Collecting the Garbage at a Building

Figure 6 shows a diagram of the garbage truck model. This model is designed to retrieve garbage from buildings at regular intervals. The parameters of the garbage truck model are the visit order of the buildings, the collection time for each building, the capacity of the garbage truck, and the collection cycle. Each parameter can be adjusted to create an optimized waste management system. When the garbage truck model visits a building, garbage is taken from the garbage can and added to the truck. The truck has a fixed storage size, which the amount of garbage that the truck is carrying cannot exceed. Therefore, the garbage truck may not collect garbage from a garbage can when the truck does not have sufficient storage space available. The garbage truck model sends a message to receive the accumulated garbage from a garbage can. This process is repeated until the truck has visited all buildings. After the truck has visited all buildings, the amount of garbage in its storage space is set to zero, and another round of garbage collection is scheduled.
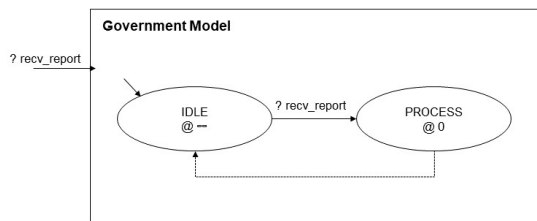

Figure 7. Atomic Model for Receiving a Report from a Resident

Figure 7 shows a diagram of the government model. The government model collects complaints from residents. To assist the practitioner, the model records the complaint time and satisfaction level of each resident.

4 CASE STUDY

This section illustrates a case study of the proposed method and simulation environment. Based on the population composition of an administrative district in Pohang City, Republic of Korea, experiments were conducted by modeling the living patterns of construction workers, students, and homemakers, who represent the majority of the occupants of the administrative districts. The living pattern for each occupation was compiled based on the analysis report for a public opinion survey targeting the national population [17].

The following assumptions were adopted when modeling the residential area and the waste management system:

- Each building has a garbage can to hold garbage temporarily.

- Each resident departs on his or her own schedule.

- Each resident takes out the garbage every day when he or she departs.

- The satisfaction level of a resident may change at the moment when the resident takes out the garbage.

- A resident may complain to the government when the garbage accumulation is above a certain level.

4.1 Experimental Design

Because the actual operations of garbage trucks may be different for each garbage disposal company, in this research, an abstract model was built for the garbage collection strategy. The garbage truck collects trash from the garbage can of each building based on the current capacity of the truck. Additionally, single-person households were assumed to analyze the simulation results with regard to the following patterns: 1) a fixed sequence, 2) a palindrome sequence, and 3) a random sequence. A fixed sequence is a sequence in which the garbage truck visits the buildings in a predefined order. For example, if there are three buildings, A, B, and C, the garbage truck may visit the buildings in alphabetical order. Accordingly, the sequence of the schedule would be $A \rightarrow B \rightarrow C \rightarrow A \rightarrow B \rightarrow C \rightarrow \ldots$ A palindrome sequence is a sequence in which the order is reversed after the completion of each cycle of collection. In the palindrome sequence for the example given above, the garbage truck would visit the last building first in the next cycle, and the building visited first in the previous cycle would be visited last, as follows: $A \rightarrow B \rightarrow C \rightarrow C \rightarrow B \rightarrow A \rightarrow \ldots$ Finally, a random sequence is a sequence in which every building is visited in a random order.

The experiments were conducted using a system with a 4.20 GHz Intel(R) Core(TM) i7-7700K CPU with 4 physical cores and 64 GB of RAM running 64-bit Windows 10 Pro, build 18362. Ubuntu 18.04.4 LTS (x86_64 GNU/Linux 4.19.76-linuxkit) and Python 3.6.9 were mounted using Docker. Each experiment was conducted 30 times to measure the average number of complaints from residents concerning each waste collection strategy: fixed, palindrome, and random. On average, each simulation required 333 seconds to simulate 31 days.

Table 1. Description of Simulation Parameters

	Construction Worker	Student	Homemaker
Departure Time	6:22:00	7:58:00	13:00:00
Amount of Waste Generated	0.9	0.9	0.9
Satisfaction Function	$Sat(x) = \begin{cases} +10 & \text{if } x < 0.8, \\ 20 & \text{if } x = 0, \\ -10 & \text{if } x > 0.8 \end{cases}$	$Sat(x) = \begin{cases} +10 & \text{if } x < 0.8, \\ 20 & \text{if } x = 0, \\ -10 & \text{if } x > 0.8 \end{cases}$	$Sat(x) = \begin{cases} +10 & \text{if } x < 0.8, \\ 20 & \text{if } x = 0, \\ -10 & \text{if } x > 0.8 \end{cases}$

Table 1 shows the simulation parameters deduced from the assumptions. $Departure\ Time$ represents the interval at which a resident leaves his or her home. We modified the $Departure\ Time$ based on the average wake-up time data from the behavior survey [17]. $Amount\ of\ Waste\ Generated$ represents the amount of waste generated between departure times. To ensure a realistic simulation, the amount of waste generated by residents was drawn from a distribution corresponding to the average waste generation per capita [18]. The $Satisfaction\ Function$ represents a resident's satisfaction level depending on the level of garbage accumulation. This function was defined heuristically based on the experiences of several students.

4.2 Simulations and Lessons Learned
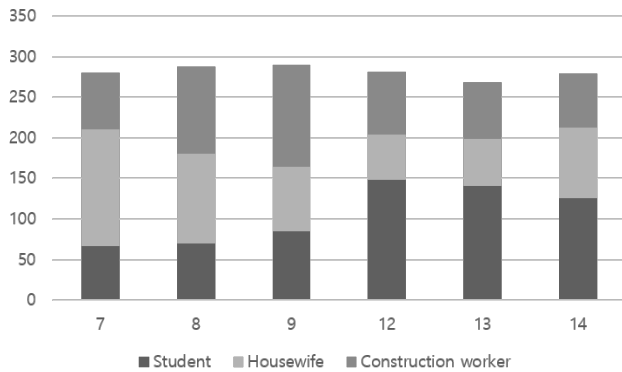
Figure 8. Fixed Sequence

Figure 8 shows the simulated numbers of complaints filed by a population with an equal composition of students, homemakers, and construction workers (33:33:33) served by a garbage truck following a fixed sequence. At 7:00, the number of complaints from students was the lowest, and the number of complaints from homemakers was the highest. As the collection time was delayed from 7:00 to 12:00, the number of complaints from students continuously increased, and the number of complaints from homemakers decreased. As the collection time was further shifted from 13:00 to 14:00, the number of complaints from students decreased, and the number of complaints from homemakers increased. In the case of the construction workers, the number of complaints increased gradually from 7:00 to reach a maximum at 9:00 and then continuously decreased until 14:00. In terms of total hourly complaints, the lowest number of complaints occurred at 13:00, and the highest number of complaints occurred at 9:00. The average number of complaints across all times was 280.87.
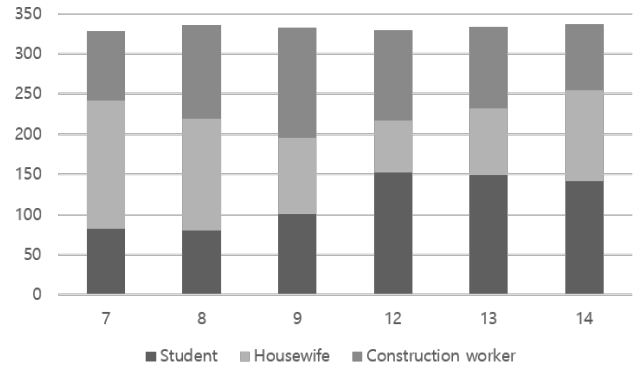

Figure 9. Palindrome Sequence

Figure 9 shows the simulated numbers of complaints filed by a population with an equal composition of students, homemakers, and construction workers served by a garbage truck following a palindrome sequence. The number of complaints from students continuously increased as the collection time was delayed from 7:00 to 12:00, while the number of complaints from homemakers decreased. This trend reversed at 12:00, with the number of complaints from students decreasing from 13:00 to 14:00, while the number of complaints from homemakers increased. In the case of the construction workers, the number of complaints increased from 7:00 to 9:00 and decreased beginning at 12:00. Overall, the lowest number of complaints was at 7:00, and the highest number of complaints was at 14:00. The average number of complaints across all times was 332.52.
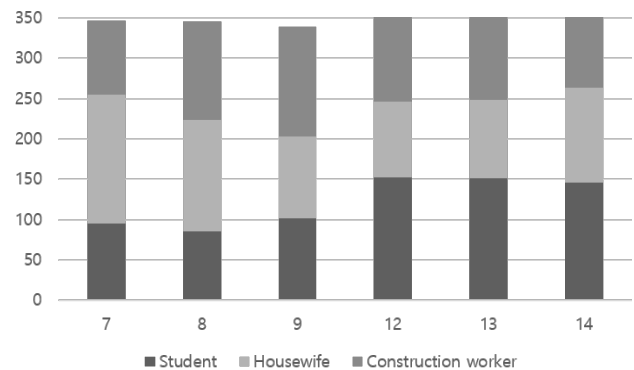

Figure 10. Random Sequence

Figure 10 shows the simulated numbers of complaints filed by a population with an equal composition of students, homemakers, and construction workers served by a garbage truck

492

following a random sequence. The number of complaints from homemakers was the highest when the garbage collection time was at 7:00, while the students filed the highest number of complaints at 12:00. Construction workers made the most complaints at 9:00. Across all times, the total number of complaints was lowest when garbage was collected at 9:00 and the highest at 12:00. The average number of complaints across all times was 353.14. Among the three garbage collection strategies, the fixed sequence resulted in an average number of civil complaints that was 18 to 25% lower than those in the cases of the palindrome and random sequences in these experiments. The actual waste collection company operating in the study area follows a fixed collection schedule, that is, a fixed sequence, which results in the lowest average number of complaints.
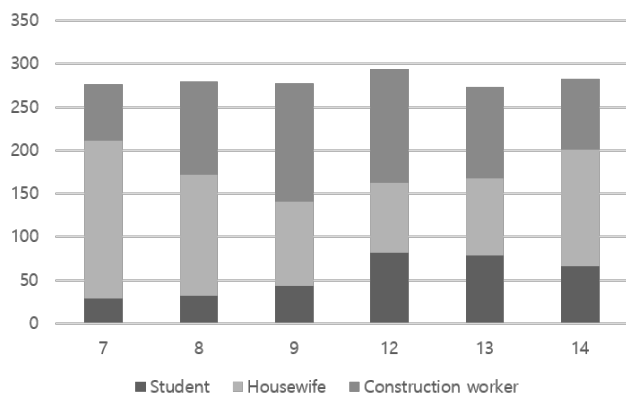
Figure 11. Fixed Sequence with an Imbalanced Population Composition

Figure 11 shows the simulated numbers of complaints filed by a population with a composition of 18:33:48 (students:homemakers:construction workers) served by a garbage truck following a fixed sequence. At 7:00, the number of complaints from students was the lowest, and the number of complaints from homemakers was the highest. At 12:00, students and construction workers filed the most complaints. Overall, the total number of complaints was the highest at 12:00 and the lowest at 13:00. The patterns of increase and decrease in the numbers of complaints from residents with each occupation were similar to those in the case of equal composition rates. The average number of complaints across all times was 280.69, almost the same as in the equal-composition simulation. From these experiments, it was found that the results vary depending on the population composition. However, with fixed-sequence collection, the average number of complaints is the lowest for collection at 13:00, even for different population compositions. It can be concluded that changing the garbage collection time to 13:00 should reduce the average number of complaints.

5 CONCLUSION

As the complexity of urban areas increases, estimating waste generation and providing alternative solutions for waste collection are essential for urban planners and employees of the local government to increase the satisfaction level of residents. In this research, the living patterns of residents with different occupations were modeled to estimate waste generation and collection to compute the residents' approximate satisfaction level. In the proposed modeling approach, the simulation environment was modeled by considering residents, buildings, garbage trucks and different operation strategies based on the DEVS formalism. In this environment, various assumptions and operation strategies were established and tested.

The experimental results showed that the number of complaints may vary depending on the population composition; however, the total number of complaints was the lowest when garbage was collected at 13:00, independent of the population composition, for the case of fixed-sequence collection in the given experimental environment. This study has demonstrated that it is possible to perform modeling based on the population composition of a residential area in an actual city and thus determine the optimal collection time on the basis of the existing collection pattern. Alternatively, the proposed modeling approach and simulation environment can be used in combination with alternative derivation methods to derive the optimal collection pattern. In future research, an experiment can be conducted to investigate family-unit waste disposal patterns, ranging from single-person households to multi-person households. Additionally, an experiment to validate the effectiveness of installing smart garbage cans can be performed by adopting dynamic vehicle routing in the collection schedule. Implementing a graphical user interface to increase the usability of the model for urban planners or local government employees can also be considered.

REFERENCES

1. United Nations. 2018 revision of world urbanization prospects, 2018.

2. Mohammad Abdul Mohit, Mansor Ibrahim, and Yong Razidah Rashid. Assessment of residential satisfaction in newly designed public low-cost housing in kuala lumpur, malaysia. *Habitat international*, 34(1):18–27, 2010.

3. Bernard P. Zeigler. *Multifaceted Modeling and Discrete Event Simulation*. Academic Press, 1984.

4. Bernard P. Zeigler. *Object-Oriented Simulation with Hierarchical, Modular Models*. Academic Press, 1990.

5. Bernard P. Zeigler, Tag Gon Kim, and Herbert Praehofer. *Theory of Modeling and Simulation*. Academic Press, Inc., Orlando, FL, USA, 2000.

6. Anh Dao-Tuan, Anh Nguyen-Thi-Ngoc, Khanh Nguyen-Trong, Anh Bui-Tuan, and Van Dinh-Thi-Hai. Optimizing vehicle routing with path and carbon dioxide emission for municipal solid waste collection in ha giang, vietnam. In *International Conference on Industrial Networks and Intelligent Systems*, pages 212–227. Springer, 2017.

7. Christos Chalkias and Katia Lasaridi. A gis based model for the optimisation of municipal solid waste collection: the case study of nikea, athens, greece. *technology*, 1:11–15, 2009.

8. Swapan Das and Bidyut Kr Bhattacharyya. Optimization of municipal solid waste collection and transportation routes. *Waste Management*, 43:9–18, 2015.

9. Sindhya K Nambiar and Sumam Mary Idicula. A multi-agent vehicle routing system for garbage collection. In *2013 Fifth International Conference on Advanced Computing (ICoAC)*, pages 72–76. IEEE, 2013.

10. Khanh Nguyen-Trong, Van Dinh-Thi-Hai, Anh Nguyen-Thi-Ngoc, and Doanh Nguyen-Ngoc. Emission control and route optimization in municipal solid waste collection and transportation using agent-based model. *International Journal of Innovative Technology and Exploring Engineering*, 7(2):23–28, 2017.

11. Xiaoran Shi, Aristotelis E Thanos, and Eric D Antmann. Multi-objective agent-based modeling and optimization of single stream recycling programs. Technical report, University of Florida, 2013.

12. Nikolaos V Karadimas, George Rigopoulos, and Nikolaos Bardis. Coupling multiagent simulation and gis- an application in waste management. *WSEAS Transactions on Systems*, 5(10):2367–2371, 2006.

13. Tuyen Manh Hua, Trong Khanh Nguyen, Hai Van Dinh Thi, and Ngoc Anh Nguyen Thi. Towards a decision support system for municipal waste collection by integrating geographical information system map, smart devices and agent-based model. In *Proceedings of the Seventh Symposium on Information and Communication Technology*, pages 139–146. ACM, 2016.

14. Giorgio Guariso, F Michetti, F Porta, and S Moore. Modelling the upgrade of an urban waste disposal system. *Environmental Modelling & Software*, 24(11):1314–1322, 2009.

15. Brian Dyson and Ni-Bin Chang. Forecasting municipal solid waste generation in a fast-growing urban region with system dynamics modeling. *Waste management*, 25(7):669–679, 2005.

16. Chen Wang. System dynamics simulation on municipal solid waste in hong kong. In *Proceedings of the Institution of Civil Engineers-Municipal Engineer*, volume 171, pages 185–195. Thomas Telford Ltd, 2017.

17. Gallup Korea. Korean life time-wake up / sleep / sleep time. `http://www.gallup.co.kr/gallupdb/reportContent.asp?seqNo=516`. Accessed:2020-01-15.

18. Korea Resource Recirculation Information System. National waste generation and disposal status (2018). `https://www.recycling-info.or.kr/rrs/stat/envStatDetail.do?menuNo=M13020302&pageIndex=1&bbsId=BBSMSTR_000000000002&s_nttSj=KEC011&nttId=867&searchBgnDe=&searchEndDe=`. Accessed: 2020-01-15.

PANDO: Parametric Tool for Simulating Soil-Plant-Atmosphere of Tree Canopies in Grasshopper

Ata Chokhachian[1], Marion Hiller[2]

[1]Technical University of
Munich
Munich, Germany
ata.chokhachian@tum.de

[2]TRANSSOLAR
Energietechnik GmbH
Stuttgart, Germany
Hiller@transsolar.com

ABSTRACT

PANDO is a numerical process-based simulation tool that enables the user to model individual trees and canopies in Rhino Grasshopper interface. This plugin simulates radiation fluxes in three different wavebands of photosynthetic (PAR), near infrared (NIR) and thermal. It also counts for photosynthesis and transpiration by tree canopies. PANDO can be used also to estimate root water uptake, drainage, canopy interception and cooling potential of trees by coupling with outdoor thermal comfort models. The tool enables the user to plan and design urban greenery more efficiently in cities considering the benefits and future paybacks.

Author Keywords

Tree canopy simulation; urban greenery; annual outdoor comfort; grasshopper; MAESPA.

CCS Concept

• Computing methodologies ~ Modeling and simulation ~ Simulation support systems ~ Simulation tools

1 INTRODUCTION

Trees are one of the critical components for cities to define livability and health measures. Studies show that healthy and well planted trees can mitigate the impacts associated with the built environment [1]. Over the last 30 years the primary purpose of street trees has changed from an aesthetic role of beautification and ornamentation to one that also includes the provision of services such as storm water reduction, energy conservation and improved air quality [2, 3]. As an example a recent report from Britain's Office of National Statistics estimates that trees saved the capital more than 5 billion pounds from 2014 to 2018 through air cooling alone. Additionally, by keeping summer temperatures bearable for workers, trees prevented productivity losses of almost 11 billion pounds as well [4]. Beside economic benefits, trees improve urban life and microclimate of cities to a more enjoyable place to live, work, and play, while mitigating the city's environmental impact [5].

There is an ample literature proving that trees in urban areas make significant economic, environmental, social, cultural and spiritual contributions to the well-being of people [6-10].

Starting from 2006 projected until 2040, several cities over the world committed to the "milliontrees" initiative. Metropolitans like; Los Angeles, Denver, New York City, Shanghai, London, Ontario and Pune started planting one million trees aiming at increasing urban forest pointing at reduction of carbon dioxide in the air to reduce the effects of global warming [11]. Moreover, It has been already proven that there is a significant relationship between amount of green space and tree in the public spaces and how people occupy and choose walking routes in daily commutes [12, 13]. As an example, Larsen, Gilliland [14] investigated the influence of the physical environment on mode of travel finding that land use mix and presence of street trees were the only significant physical environment variables. The study highlights that the likelihood of walking or biking to school was positively associated with both increased land use mix and greater number of street trees. Although land use mix and population density should also be considered with regard to increasing rates of non-motorized travel, the study recommends that planting trees may be the most efficient and cost-effective environmental intervention to encourage walking to school.

Nevertheless, the amount of investment on tree planting some may question the need for additional investigation on benefits and future paybacks of such a service. For that reason, further empirical research is necessary in order to efficiently guide the design and planning of urban green space, and specifically to investigate the importance of the abundance, distribution and type of greening [15]. Accounting for these effects requires proper modeling of vegetation, including vegetation physiology to estimate accurate latent energy fluxes and the soil vegetation-atmosphere. Bowler, Buyung-Ali [15] suggests that the current research body does not determine exactly how urban greenery should be employed in terms of abundance, type, and distribution. Therefore a suitable modeling tool is needed to provide solid quantitative assessments of the effectiveness of tree canopies and water in terms of heat mitigation for pedestrian outdoor comfort in urban environments, and therefore provide guidance on how to most effectively use trees in cities [16].

At the other end of the complexity spectrum, there are number of models and tools, mainly based on computational fluid dynamics, are able to resolve at a micro-scale and account for many of the influences of vegetation [17-20]. However, their complexity and computational intensity puts them out of reach for less specialised users, plus they lack providing annual simulation results which are fundamental for informing design and urban planning.

Nice, Coutts [16] argues that there is a need for an approach that fits in between the two levels of complexity, a new innovative method to balance detail, accuracy, and complexity with efficiency and usability, while also being able to consider a wide variety of vegetation types and arrangements. To fill the gap, we developed a tool based on MAESPA [21, 22] a process based model simulating fluxes of energy, water, and carbon at the tree and stand-scale levels.

2 SOIL-PLANT-ATMOSPHERE MODEL

Numerical Process-based Simulation Models (PBMs) are one of the useful methods to incorporate the fundamental physical and eco-physiological processes (e.g., photosynthesis and transpiration) to estimate fluxes of energy, water, and carbon in the ecosystem, as a function of climate, soil, and plant characteristics. Also they can be used as an effective tool to identify interactions between environmental drivers, plant and canopy structure, leaf physiology and soil water availability and their combined effects on water use and carbon uptake [23].

There are different approaches to model trees in terms of complexity and scale. The two-dimensional multi-layer models [24] are limited since they consider the interception of light from only a single direction which ends up underestimating the transpiration and photosynthesis rates [25]. Therefore, three-dimensional tree-scale models are more accurate and appropriate to simulate the radiation transfer and soil-plant atmosphere [26]. One model that uses voxels to make the required trade-off between computation time and precision is MAESPA [21], which is a recent coupling of the tree-scale light interception and ecophysiology model MAESTRA [22, 27] and the soil and ecosystem water and energy balance SPA [28]. MAESPA occupies a very interesting niche in the PBM complexity continuum, between the complex, detailed 3-D models [17] and the less-detailed multilayer models [29]. Thus, MAESPA is a relevant candidate for addressing effects of climate change upon horizontally heterogeneous tree canopies.

3 MAESPA MODEL EVOLUTION

MAESPA has a long history, going back to the work of John Norman and Paul Jarvis in the 1970's and 80's and later continued by Wang [30] called MAESTRO. In 2004 Medlyn [27] revised the model, with the purposes of modularising the code to make the program easier to understand and modify; and incorporating standard formulations of leaf gas exchange models called MAESTRA. In 2008, Duursma and

Medlyn [21] developed an individual tree based model named MAESPA largely based on combining the MAESTRA and SPA [31] ecosystem models. The collective model includes a hydraulically-based model of stomatal conductance, root water uptake routines, drainage, infiltration, runoff and canopy interception, as well as detailed radiation interception and leaf physiology routines from the MAESTRA model.

In 2018, Vezy, Christina [32] modified the MAESPA model to add calculation of foliage surface water evaporation at the voxel scale, computation of an average within-canopy air temperature and vapour pressure, and use of iterative calculations of soil and leaf temperatures to close ecosystem-level energy balances. In the modified version, each tree in the canopy is described individually, and can have different sets of physiological and structural parameters; for instance, according to each tree's species, age, or size. MAESPA simulates the foliage light absorption, photosynthesis, soil evaporation, transpiration, and balances of water and energy. The coupling with SPA model allows a precise computation of soil water balance and plant hydraulics, so that stomatal conductance can respond to leaf water potential.

4 MAESPA RADIATION MODEL

In MAESPA the foliage energy balance is computed at the voxel scale for each tree. Each voxel contains a given amount of leaf area within a tree, which has a set of homogeneous properties such as leaf inclination angle distribution, optical properties and photosynthetic parameters (Figure 1).
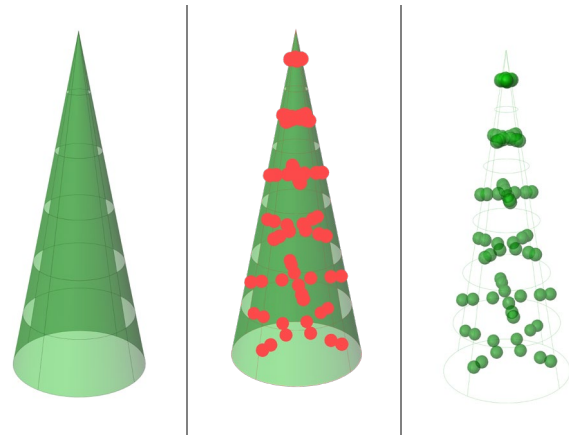

Figure 1. 3D representation of tree geometry in MAESPA showing the voxels and sensors points.

The net radiation of each voxel is computed from the light interception sub-modules in three spectral domains which are: Photosynthetically Active Radiation (PAR), Near Infrared (NIR) and the thermal domains (Figure 2). The light interception submodule takes into account the 3D representation of the stand, in which each tree is located according to its x,y,z coordinates, and characterized by its height, crown length, radius, shape and total leaf area [32].

The initial radiation routines are described in detail by Wang and Jarvis [22]. The canopy consists of individual tree

crowns, which are described by a basic shape, length, height to crown base, and width. Radiation calculations can be performed for the entire canopy or only for a set of target crowns specified by the user. The points in the voxels are used to calculate radiation domains based on shading within the crown and by neighboring trees, the location of the sun, and whether radiation is direct or diffuse. Scattering of radiation is approximated following the method described in detail by Norman [33]. Leaf area within crowns is assumed to be distributed randomly, or to follow a beta distribution in horizontal and/or vertical directions [34]. At each grid point, leaf area is separated into sunlit and shaded leaf area [35], and the coupled stomatal conductance and photosynthesis model is run separately for each fraction.
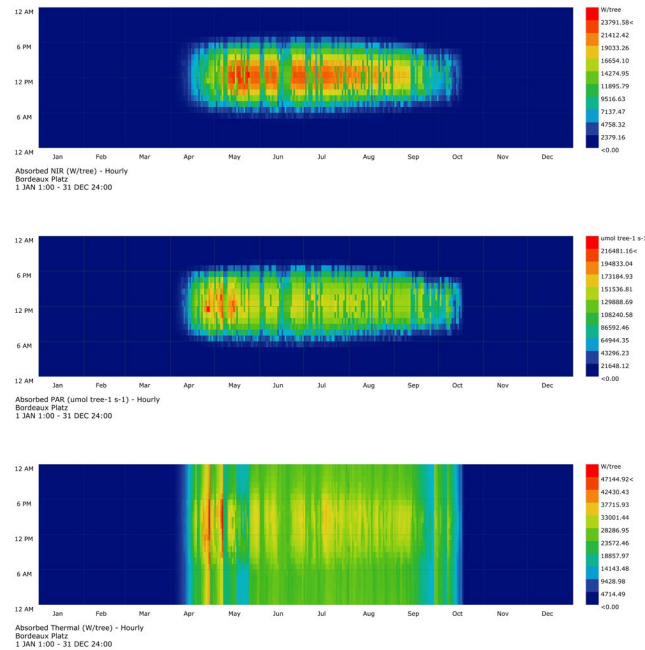
Figure 2. Annual radiation spectrums (NIR, PAR and Thermal) with dynamic Leaf Area Index.

5   PANDO DEVELOPMENT

PANDO is a parametric tool scripted in python for visual programming interface (Grasshopper). The interface consists of components to prepare files directly from 3D modeling program (Rhinoceros) for MAESPA. The workflow consists of components to generate tree geometries, weather data, water balance and simulation setup (Figure 3). After running the simulation PANDO reads the results back to visualize wide range of hourly or daily outputs including: absorbed, transmitted and reflected radiation fluxes; transpiration; sensible heat fluxes; leaf photosynthesis rates; water potential; CO_2 concentration and Leaf temperatures all in single tree or canopy scales. The workflow of the tool including the brief definition and parameters of the input files, different sub-models of simulation and output parameters are illustrated in Figure 5.
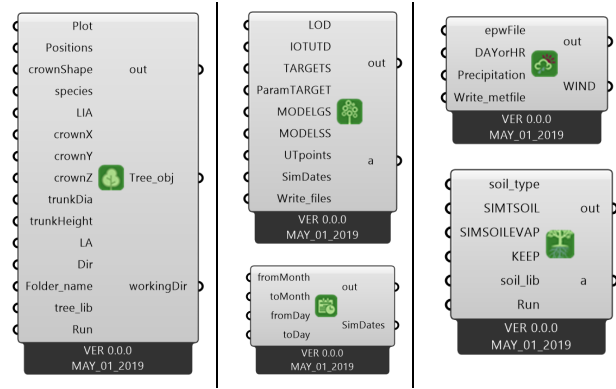

Figure 3. Grasshopper components of PANDO.

5.1 Geometry and Structure Settings

The geometry component enables the user to customize the tree canopy in terms of shape, size, species and leaf area parameters. Each of these settings can be assigned separately for each tree geometry or uniform for the entire canopy. The component transforms the geometric inputs from Rhino Grasshopper interface into data files including the location coordinates, crown and trunk size and lead area parameters. In the structure settings the user can choose between different predefined shapes including: conical, half ellipsoidal, paraboloid, full ellipsoid, cylinder and box for the crown shape (Figure 4). For the leaf angle distribution which is an input for the radiative transfer calculations, it can be uniform for the entire canopy or a ratio of horizontal and vertical distribution.
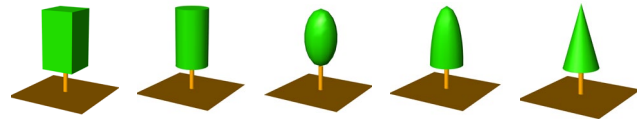

Figure 4. Optional tree geometries for canopies.

5.2 Weather Data

This file contains meteorological observations that drive MAESPA including location information, atmospheric pressure, temperature, radiation and humidity. PANDO converts the commonly used TMY file formats into meteorological files readable by MAESPA. The input file can be in hourly or daily time steps. The precipitation rates can also be separately included.

5.3 Phenology Model

One of the main challenges of modeling tree canopies is to estimate seasonal behavior of trees. This was crucial since one of the main objectives of this study is to enable the user to simulate and estimate benefits of urban greenery in annual cycles. Thanks to MAESPA's phenology model [36] the user can define the flushing date and the number of days from the flush date to the end of the first flush, end of the second flush, start of leaf senescence, and end of leaf senescence to create a dynamic leaf area index for the canopy.
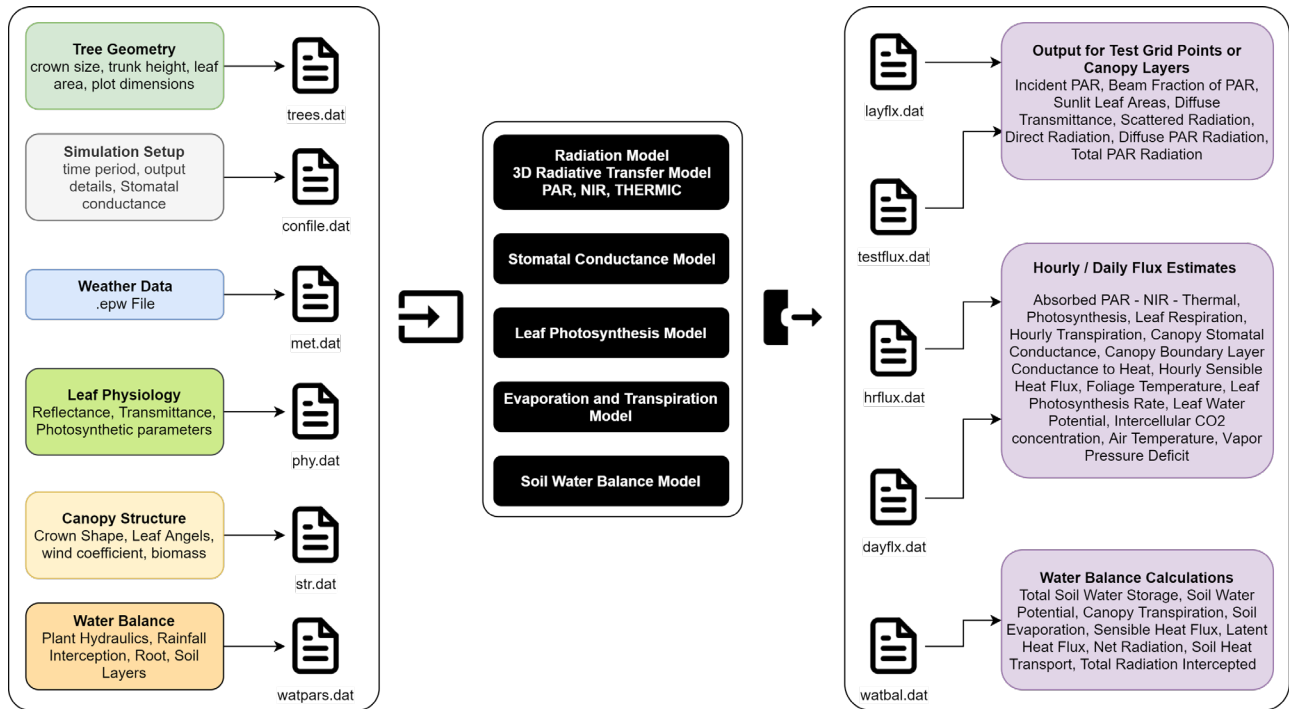
Figure 5. The workflow of input and output files in PANDO.

Also the model needs the time taken for a single leaf to fully expand in days as an input. Using the provided inputs, the model will generate annual variation of leaf area (Figure 6) which is used to estimate the radiation, evaporation, transpiration and photosynthesis fluxes.
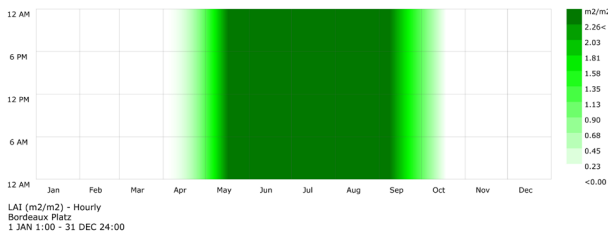

Figure 6. Annual Leaf Area Index (LAI) variations based on the phenology model.

5.4 PANDO Tree Templates

The parametrization of MAESPA model for a non-expert user can be fairly complicated. To foster this problem, PANDO comes with tree and soil templates in json file format. The user based on the desired plant species can import from the library. The plant template files include information about leaf physiology, leaf transmittance and reflectance, stomatal conductance, Quantum yield of electron transport and day respiration rates. The user can modify and update the parameters and generate custom templates as well.

6 MODEL VALIDATION

Radiation is one of the main drivers for having reliable energy balance in process based models. For this reason we confronted the simulated PAR values of PANDO with onsite measurements obtained from previously published research by Rahman, Moser [37]. The experiment was done in city of Munich in Bordeaux Platz as an open green square with avenue canyon running from North to South and from South to North. The site has four rows of T. cordata trees with approximately same age. Two rows each were planted on the two sides of a pedestrian walk-way on each side of the square. Table 1 shows the morphological parameters of the trees in Bordeaux Platz used for parametrization of PANDO simulation model [38].

Parameter	Unit	Value
Age	Years	40±20
Crown Projection Area (CPA)	m^2	62.2±0.4
Crown radius	m	4.46±0.03
Crown surface	m^2	670±24
Crown volume	m^3	290±26
Diameter at Breast Height (DBH)	cm	28.7±0.76
Height	m	14.9±0.29
Leaf Area Index (LIA)	m^2 / m^2	2.3±0.28

Table 1. PANDO parametrization for tree geometry.

The measured PAR values are collected using PQS1 sensor (Kipp & Zonen, Delft, The Netherlands) installed on top of a 3.3 m street lamp post by a 3.5 m cross arm, 2 m outward from the lamp to avoid influence of lamp and shade of the nearby trees and buildings. the data were recorded continuously at a 15-min resolution from August 6th to October 13th, 2015. Figure 7 shows the

confrontation of measured and simulated PAR values with 15 minutes intervals from 5th to 15th of August 2015. The results show high correlation between simulated and measures Photosynthetically Active Radiation with squared correlation coefficient of 0.99 over the study period.
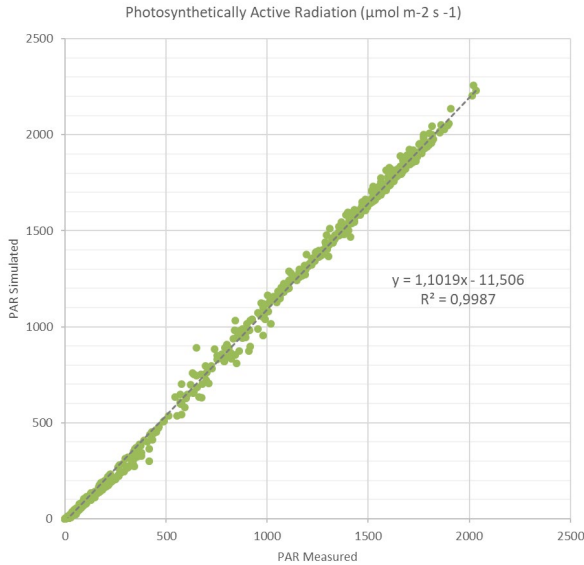

Figure 7. Confronting measure and simulated PAR values from August 5th to 15th.

Comparing the hourly values in Figure 8, shows that the simulation model overestimated the daily peaks up to 230 μmol/m^2/s around noon. The PANDO weather file input should normally have direct and diffused radiation fluxes separately, however in our measurements we only had the global radiation. This input is separated by the radiation model into direct and diffused automatically and this simplification overestimates the output PAR values.
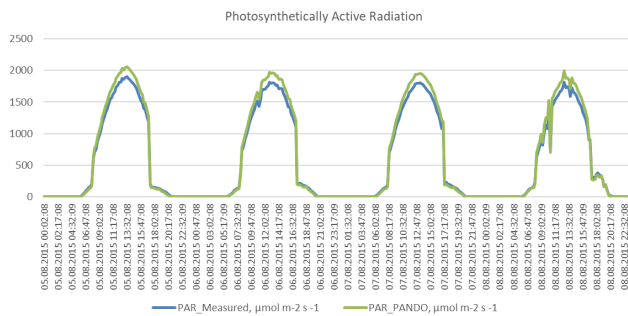
Figure 8. Line graph of measured and simulated PAR between 5th to 8th of August (15 minutes interval).

7   MODEL APPLICATION

As highlighted before, trees are one of the critical component of urban planning to enhance livability and health. However, quantifying the benefits of trees in term of annual thermal comfort stays the main challenge due to the complexity and dynamic behavior of tree canopies.

Using the PANDO workflow for modeling the tree canopies, we simulated annual perceived temperatures by Universal Thermal Climate Index (UTCI) index [39] for two scenarios of a person fully exposed and shaded under the tree. One of the main drivers and most effectives parameters to have reliable UTCI assessment is a detailed and proper model for the radiant environment. The Mean Radiant Temperature (MRT) modeling integrates exposure to shortwave and longwave radiation in a three-dimensional environment assuming that the radiant heat transfer from the human body is equal to the radiant heat transfer in the actual non-uniform enclosure [40]. To calculate MRT, the process is divided into shortwave fluxes (direct, diffuse and reflected solar radiation) and long-wave radiation (urban surfaces, trees and sky) using the weather data for city of Munich.

To be accurately represented, direct and diffused short-wave solar radiation is simulated using radiance raytracing with Daysim's hourly irradiation method. For trees, the radiation intercept differently compared to solid objects in the urban environment. They prevent incident shortwave and longwave radiation arising from the context reaching pedestrians and nearby surfaces. As incident radiation reaches the tree canopy, a certain amount of it is blocked until the radiation passes out. The ratio of intercepted energy is calculated using [41]:

$$Tree\ intercepted\ fraction = 1 - exp\left(-K_{bs}\Omega LAI\frac{w}{x}\right)$$

Where K_{bs} is the extinction coefficient of the leaves (assumed 0.5 for this study) and Ω is the clumping factor ranging from 0 (clumping leaves) to 1 (random distribution) [42]. LAI is dynamic leaf area index, w and x are distance between tree and crown size respectively.

For the longwave fluxes, the surface temperatures are estimated based on air temperature, wind velocity, net all-wave radiation and Bowen ratio [43]. The sky temperature is calculated based on the method developed by Martin and Berdahl [44] taking into account the emissivity model introduced by Duffie, Beckman [45]. The average leaf temperature of the tree canopy is extracted from PANDO model for every hour. Having radiant fluxes from the adjacent surfaces, based on the view factors from the body position, the MRT component is calculated for each hour of the year based on the Stefan-Boltzman law [46].

In the next step, using the annual hourly MRT values with local wind speed, air temperature and humidity, the perceived temperatures (UTCI) are estimated for the two cases of fully exposed and under the tree (Figure 9). the results show significant improvement by decreasing the strong heat stress from 16.9 % to 3.9 % over the year due to shading effect and cooling potential of leaves caused by lower surface temperatures. For climate of Munich, we also see considerable improvement (8 %) in the comfort band of UTCI ranging between 9 to 26 °C (Figure 10).
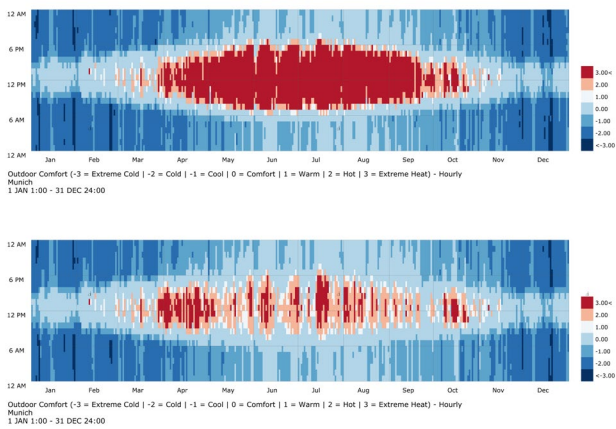
Figure 9. Comparison of the annual perceived temperatures (UTCI) exposed (top) and under the tree (bottom).
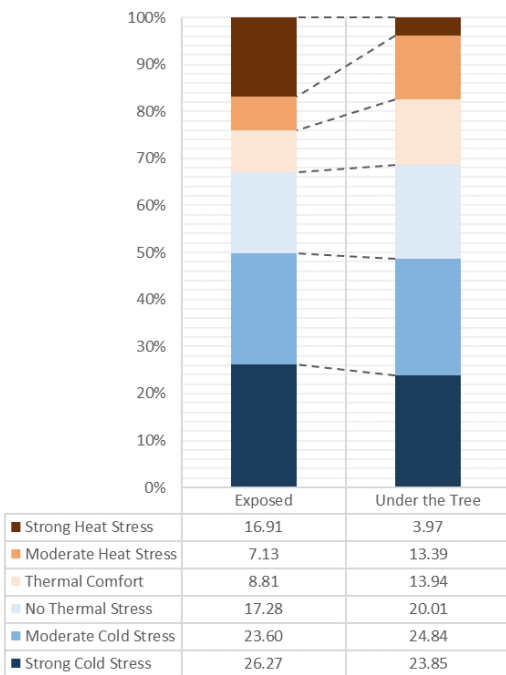

	Exposed	Under the Tree
■ Strong Heat Stress	16.91	3.97
■ Moderate Heat Stress	7.13	13.39
▫ Thermal Comfort	8.81	13.94
▫ No Thermal Stress	17.28	20.01
▪ Moderate Cold Stress	23.60	24.84
▪ Strong Cold Stress	26.27	23.85

Figure 10. Thermal comfort bands for exposed and shaded scenarios (annual condition of person).

8 CONCLUSION

This study presents the development, validation and application scenarios of parametric tool for modeling tree canopies in grasshopper. Currently, there is no sufficiently detailed PBM application to model trees in urban environments to output annual statistics on their effects on outdoor thermal comfort. The developed tool can be used as standalone program to simulate behavior of tree canopies or can be coupled with building and/or urban simulation programs to model urban canyons including vegetation. Furthermore, PANDO has the potential to be used with the daylight modeling workflows to properly estimate the effect of trees on daylight availability and energy consumption thanks to the phenology model that gives dynamic LAD or LAI. The application scenarios of the tool can go from modeling a standalone tree or single row of tree canopy in a street canyon to modeling a park in the middle of a residential neighborhood (Figure 11).
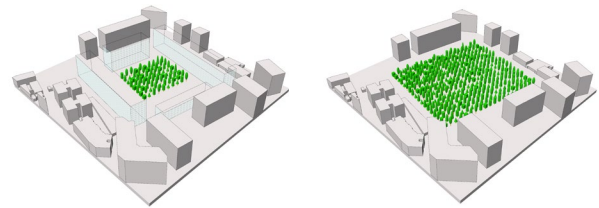

Figure 11. Conceptual modeling scenario of a park using PANDO.

REFERENCES

1. Zhao, Q., D.J. Sailor, and E.A. Wentz, *Impact of tree locations and arrangements on outdoor microclimates and human thermal comfort in an urban residential environment.* Urban Forestry & Urban Greening, 2018. **32**: p. 81-91.

2. Silvera Seamans, G., *Mainstreaming the environmental benefits of street trees.* Urban Forestry & Urban Greening, 2013. **12**(1): p. 2-11.

3. Salmond, J.A., et al., *Health and climate related ecosystem services provided by street trees in the urban environment.* Environmental Health, 2016. **15**(1): p. S36.

4. Statistics, O.f.N., *Woodland natural capital accounts, UK: 2020.* 2020: UK. p. 47.

5. Peper, P.J., et al., *New York City, New York municipal forest resource analysis.* 2007, Center for Urban Forest Research, USDA Forest Service, Pacific Southwest Research Station, University of California: Davis, CA.

6. Elmendorf, W., *The importance of trees and nature in community: A review of the relative literature.* Arboriculture and Urban Forestry, 2008. **34**(3): p. 152.

7. Dwyer, J.F., H.W. Schroeder, and P.H. Gobster, *The significance of urban trees and forests: toward a deeper understanding of values.* Journal of Arboriculture 17 (10): 276-284, 1991. **17**(10).

8. Greene, C.S., A.A. Millward, and B. Ceh, *Who is likely to plant a tree? The use of public socio-demographic data to characterize client participants in a private*

urban forestation program. Urban Forestry & Urban Greening, 2011. **10**(1): p. 29-38.

9. Mullaney, J., T. Lucke, and S.J. Trueman, *A review of benefits and challenges in growing street trees in paved urban environments.* Landscape and Urban Planning, 2015. **134**: p. 157-166.

10. Fernandes, C.O., et al., *Between tree lovers and tree haters. Drivers of public perception regarding street trees and its implications on the urban green infrastructure planning.* Urban Forestry & Urban Greening, 2019. **37**: p. 97-108.

11. McPherson, E.G., et al., *Million trees Los Angeles canopy cover and benefit assessment.* Landscape and Urban Planning, 2011. **99**(1): p. 40-50.

12. Reinhart, C.F., J. Dhariwal, and K. Gero, *Biometeorological indices explain outside dwelling patterns based on Wi-Fi data in support of sustainable urban planning.* Building and Environment, 2017. **126**: p. 422-430.

13. Dhariwal, J., et al., *Evaluating the effectiveness of outdoor evaporative cooling in a hot, arid climate.* Building and Environment, 2019. **150**: p. 281-288.

14. Larsen, K., et al., *The Influence of the Physical Environment and Sociodemographic Characteristics on Children's Mode of Travel to and From School.* American Journal of Public Health, 2009. **99**(3): p. 520-526.

15. Bowler, D.E., et al., *Urban greening to cool towns and cities: A systematic review of the empirical evidence.* Landscape and Urban Planning, 2010. **97**(3): p. 147-155.

16. Nice, K.A., A.M. Coutts, and N.J. Tapper, *Development of the VTUF-3D v1.0 urban micro-climate model to support assessment of urban vegetation influences on human thermal comfort.* Urban Climate, 2018. **24**: p. 1052-1076.

17. Bailey, B.N., et al., *A new three-dimensional energy balance model for complex plant canopy geometries: Model development and improved validation strategies.* Agricultural and Forest Meteorology, 2016. **218-219**: p. 146-160.

18. Kunz, R., I. Khatib, and N. Moussiopoulos, *Coupling of mesoscale and microscale models—an approach to simulate scale interaction.* Environmental Modelling & Software, 2000. **15**(6): p. 597-602.

19. Schlünzen, K.H., et al., *Joint modelling of obstacle induced and mesoscale changes—Current limits and challenges.* Journal of Wind Engineering and Industrial Aerodynamics, 2011. **99**(4): p. 217-225.

20. Huttner, S. and M. Bruse. *Numerical modeling of the urban climate–a preview on ENVI-met 4.0.* in *7th international conference on urban climate ICUC-7, Yokohama, Japan.* 2009.

21. Duursma, R.A. and B.E. Medlyn, *MAESPA: a model to study interactions between water limitation, environmental drivers and vegetation function at tree and stand levels, with an example application to [CO_2] × drought interactions.* Geosci. Model Dev., 2012. **5**(4): p. 919-940.

22. Wang, Y.P. and P.G. Jarvis, *Description and validation of an array model — MAESTRO.* Agricultural and Forest Meteorology, 1990. **51**(3): p. 257-280.

23. Morin, X., C. Augspurger, and I. Chuine, *PROCESS-BASED MODELING OF SPECIES' DISTRIBUTIONS: WHAT LIMITS TEMPERATE TREE SPECIES' RANGE BOUNDARIES?* Ecology, 2007. **88**(9): p. 2280-2291.

24. Braden, H., *The model AMBETI: A detailed description of a soil-plant-atmosphere model.* 1995: Selbstverl. des Deutschen Wetterdienstes.

25. Pearcy, R.W. and W. Yang, *A three-dimensional crown architecture model for assessment of light capture and carbon gain by understory plants.* Oecologia, 1996. **108**(1): p. 1-12.

26. Simioni, G., et al., *Treegrass: a 3D, process-based model for simulating plant interactions in tree–grass ecosystems.* Ecological Modelling, 2000. **131**(1): p. 47-63.

27. Medlyn, B., *A maestro retrospective.* Forests at the land-atmosphere interface, 2004: p. 105-122.

28. Williams, M., B.J. Bond, and M.G. Ryan, *Evaluating different soil and plant hydraulic constraints on tree function using a model and sap flow data from ponderosa pine.* Plant, Cell & Environment, 2001. **24**(7): p. 679-690.

29. Hanson, P.J., et al., *OAK FOREST CARBON AND WATER SIMULATIONS: MODEL INTERCOMPARISONS AND EVALUATIONS AGAINST INDEPENDENT DATA.* Ecological Monographs, 2004. **74**(3): p. 443-489.

30. Wang, Y., *Crown structure, radiation absorption, photosynthesis and transpiration.* 1988.

31. WILLIAMS, M., et al., *Modelling the soil-plant-atmosphere continuum in a Quercus–Acer stand at Harvard Forest: the regulation of stomatal conductance by light, nitrogen and soil/plant hydraulic properties.* Plant, Cell & Environment, 1996. **19**(8): p. 911-927.

32. Vezy, R., et al., *Measuring and modelling energy partitioning in canopies of varying complexity using MAESPA model.* Agricultural and Forest Meteorology, 2018. **253-254**: p. 203-217.

33. Norman, J., *Modeling the complete crop canopy. ln: Barfield, G.* Modification of the Aerial Environment of Crops. American Society of Agricultural Engineers, 1979. **249**: p. 280.

34. Wang, Y. and P. Jarvis, *Mean leaf angles for the ellipsoidal inclination angle distribution.* Agricultural and Forest Meteorology, 1988. **43**(3-4): p. 319-321.

35. Norman, J.M., *Scaling processed between leaf and canopy levels.* Scaling physiological processes: leaf to globe, 1993: p. 41-76.

36. Wang, E. and T. Engel, *Simulation of phenological development of wheat crops.* Agricultural Systems, 1998. **58**(1): p. 1-24.

37. Rahman, M.A., et al., *Microclimatic differences and their influence on transpirational cooling of Tilia cordata in two contrasting street canyons in Munich, Germany.* Agricultural and Forest Meteorology, 2017. **232**: p. 443-456.

38. Rahman, M.A., et al., *Within canopy temperature differences and cooling ability of Tilia cordata trees grown in urban conditions.* Building and Environment, 2017. **114**: p. 118-128.

39. Błażejczyk, K., et al., *Principles of the New Universal Thermal Climate Index (UTCI) and its Application to Bioclimatic Research in European Scale.* 2010. **14**(1): p. 91.

40. Thorsson, S., et al., *Different methods for estimating the mean radiant temperature in an outdoor urban setting.* International Journal of Climatology, 2007. **27**(14): p. 1983-1993.

41. Park, C.Y., et al., *Variations in pedestrian mean radiant temperature based on the spacing and size of street trees.* Sustainable Cities and Society, 2019. **48**: p. 101521.

42. Chen, Q., et al., *Modeling radiation and photosynthesis of a heterogeneous savanna woodland landscape with a hierarchy of model complexities.* Agricultural and Forest Meteorology, 2008. **148**(6): p. 1005-1020.

43. Oke, T.R., *Boundary layer climates*. 2002: Routledge.

44. Martin, M. and P. Berdahl, *Characteristics of infrared sky radiation in the United States.* Solar Energy, 1984. **33**(3): p. 321-336.

45. Duffie, J.A., W.A. Beckman, and W. Worek, *Solar engineering of thermal processes*. Vol. 3. 2013: Wiley Online Library.

46. Matzarakis, A., F. Rutz, and H. Mayer, *Modelling radiation fluxes in simple and complex environments: basics of the RayMan model.* International Journal of Biometeorology, 2010. **54**(2): p. 131-139.

502

Optimizing Urban Systems:
Integrated Optimization of Spatial Configurations

Serjoscha Duering[1], Angelos Chronis[2], Reinhard Koenig[3]

[1] AIT Austrian Institute of
Technology
Vienna, Austria
Serjoscha.duering@ait.ac.at

[2]AIT
Vienna, Austria
Angelos.chronis@ait.ac.at

[3]AIT, Bauhaus University
Weimar
Munich, Germany
Reinhard.koenig@ait.ac.at

ABSTRACT

The importance of performance feedback and simulation in early design stages has been argued for numerous times. Therefore, the employment of generative design solutions becomes more and more an indispensable part in the production pipeline of architecture and space. However various barriers prevent the integration of such tools in early design stages, including the lack of design practice studies, domain expertise and commonly time-consuming and disintegrated simulation engines. In our study we demonstrate a method and workflow on how the interoperability and computational cost barriers can be reduced, using an integrated simulation and analysis approach based on machine learning and generative design. We illustrate our approach by optimizing the spatial configuration of a hypothetical, medium sized urban design project in the city of Vienna for a diverse set of fitness objectives. We conclude by showing how gained results can also support manual design processes.

Author Keywords
Simulation; Deep Learning; Urban Design; Optimization.

ACM Classification Keywords
I.6.1 SIMULATION AND MODELING.

1. INTRODUCTION

Despite ever-growing computing power, calculation time is yet a significant bottleneck in the field of data-driven design and the optimization of spatial configurations of buildings - and even more so of cities. Nevertheless, the inclusion of a diverse array of performance indicators is crucial in the quest to capture the factors defining the performance of urban systems holistically.

Looking at the bigger picture, global challenges like the climate crisis and peaking rates of urbanization in many developing countries call for the integration of both climatic and social aspects when drafting plans and designing cities. At the same time, the complexity of a design task grows exponentially with the number of factors considered, creating new challenges for urban designers and planners.

Thus, the employment of generative design solutions becomes more and more an indispensable part in the production pipeline of architecture and space. Recent advances in generative design, simulation, and computational optimization opened up possibilities in employing data driven workflows to achieving design solutions of unprecedented performance. Notably, the adoption of novel paradigms such as deep-learning appears to have the potential to become another game-changer.

This paper showcases an attempt to optimize the spatial configuration of urban systems through the integration of multiple simulation engines into one framework. Most importantly, we combine deep-learning enabled real-time estimations of microclimate (thermal comfort and wind speed) with graph-based mobility and accessibility models. This setup permits previously unfeasible and more comprehensive options when defining fitness objectives used during an evolutionary optimization process (EO) thus allowing a more holistic understanding of the performance potential inherent in the design space of urban design projects.

2. BACKGROUND

The importance of performance feedback and simulation in early design stages has been argued for numerous times. Despite their significance in optimizing the performance of architecture and urban design projects, simulation and analysis tools are still not meeting the expectation of design practice [1]. Various barriers prevent the integration of such tools in early design stages, including the lack of design practice studies, domain expertise and commonly time-consuming simulation engines. Another important barrier for the integration of diverse sets of key performance indicators in early design stages is the lack of interoperability in simulation and performance feedback tools [2]. In our study we demonstrate a method on how the interoperability and computational cost barriers can be reduced, using an

integrated simulation and analysis approach based on machine learning and generative design.

The integration of different key environmental performance indicators for urban design in an integrated and interoperable software environment has been shown to have a significant advantage as a digital workflow to achieve higher environmental performance [3]. Generative design has also been used widely for optimizing urban form scenarios, using a variety of environmental simulations as performance indicators [4]. In our study the aim is to include a variable set of KPIs, including not only environmental performance, but also accessibility and other spatial metrics.

Studies that integrate other urban metrics, such as combining financial and energy goals show the benefit of a more holistic performance driven approach [5]. In our case, the focus is both on spatial metrics which are easily computable but also on environmental metrics, using a deep learning simulation prediction framework that makes their integration feasible. The description of the deep learning framework used is out of the scope of the paper and it is based on previous research by the authors [6]. The focus on this paper is mostly on the coupling of the environmental performance prediction with a number of other spatial performance indicators and within a generative design framework in order to evaluate the potential of the integration of these diverse performance goals and metrics.

3. METHODOLOGY

3.1. Overview
The approach followed here can be described by the concept of *Cognitive Design Computing* [7]. which emphasizes the role of human designers within a computer-augmented design process. For instance, in our case, the designer guides the optimization and design space exploration by both defining design-altering parameters as well as the fitness objectives, each generated solution is evaluated upon. The automatized generation of thousands of solutions provides an overview of what is possible for a project area - in terms of its spatial configuration and performance. Next, the computed solutions can both be used for inspiration and as a benchmark of design proposals manually created by the designers. Due to its high level of flexibility and wide distribution among practitioners, Rhino/Grasshopper was chosen as the central platform for the case-study. Its visual programming interface, along with the large developer community, makes it easy to adapt and extend models by additional analytical modules. However, next to these benefits, Grasshopper comes with the disadvantage of being computationally relatively inefficient and slow, which is especially problematic for applications of multi-criteria EO. Therefore, we implemented several parts of the framework in python scripts, some of which run as external processes parallel to grasshopper.

Below, the framework used for our case-study is presented step by step. We start by briefly describing the generative module and it's parameters, followed by an introduction of our analytical models and the performance metrics employed for the optimization module.

3.2. Framework
Our design pipeline consists of three basic modules: the parametric generation of a design solution, its quantitative evaluation, and a genetic optimization engine (see Fig. 1). All parts are connected through a recursive loop, in which the optimization algorithm systematically adjusts the parameters of the generator intending to maximize the performance of a design. Below each module is briefly described.

Figure 1. Structure of the proposed workflow

3.3. Generation
The generative part of the model was designed to be able to generate street-networks, blocks and basic building volumes resembling an open perimeter block typology. For the street network and block generation, the plugin DecodingSpaces was used [8].

3.4. Evaluation
Accessibility: The model for accessibility-related analysis is based on a weighted, multi-modal graph network that is combined with an extended version of the Huff-model (a gravitational model, see [9] and [10]), which stochastically distributes trips. It enables to simulate interaction effects between land-uses, associated sociodemographic properties, and points of interest through a transportation network. This set-up allows computing performance metrics, especially
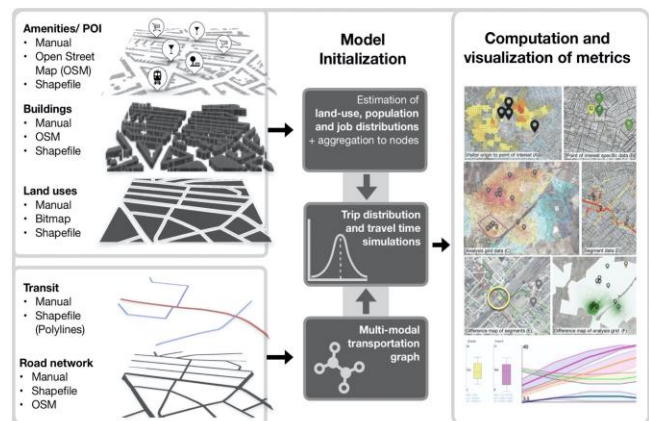

Figure 2. The accessibility and mobility model. Data is structured and aggregated to nodes of a multi-model transportation graph. A gravitational model stochastically distributes trips between defined origins and destinations.

504

regarding mobility and accessibility-related measures. More details on this module can be found in [11]. Fig 2 provides an overview of the structure.

Microclimate: For analyzing solar radiation and wind flows, we employ a novel, deep learning (DL) based method developed by [6], and based on [12]. It allows reducing computation time by the order of several magnitudes, thus making it feasible to include microclimate related indicators into iterative planning and optimization processes. In the first step a DL model is trained with thousands of image pairs of building heightmaps (Fig. 3, A) and analysis results gained through traditional simulations (B), in this case conducted throughout six months using the grasshopper plugin Ladybug [13] and openFOAM [14]. A trained model can then be employed to predict simulation results based on building heightmaps. While the preparation of training data took months, and the training a day, a pre-trained model makes predictions in milliseconds.
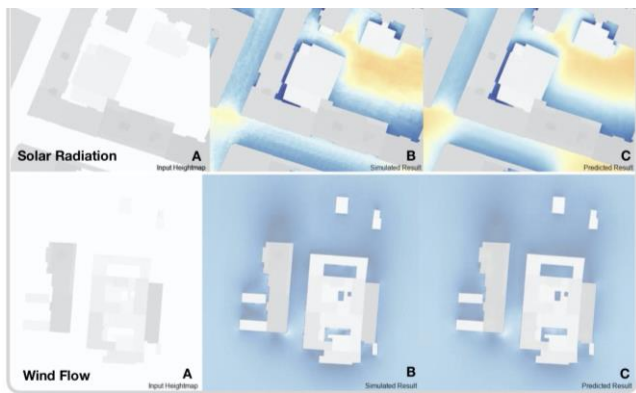
Figure 3. In and outputs of the deep learning based prediction of solar radiation (top) and wind flows (bottom). Based on the building height map (A) the model estimates the results for solar radiation and wind speeds (C). B shows results of the physical simulation using ladybug and openfoam.

3.5. Optimization
For the optimization process, we use the Octopus EO engine [15], a plugin available for grasshopper. It allows us to take multiple fitness objectives into account and provides a basic Interface to analyze results. Especially for the class of complex problems, urban design is typically part of, where different goals are not directly comparable with one another, multi-criteria optimization appears to be the preferable approach to take. Instead of weighting and subsuming all indicators into one aggregated fitness function, multi-criteria optimizers are useful in finding solutions that potentially work well as compromises between all fitness objectives. This implies that the algorithm cannot return a single best solution but collections of relatively well-performing solutions a designer hast to choose from or can be inspired by in the end.

3.6. Case Study
To try our framework, we test an urban design intervention in Vienna. The hypothetical project covers roughly the size of six typical blocks. However, the area of analysis is much larger, covering more than four square kilometers of the surrounding context. In this way, interaction with, and impact on the whole area can be included in the analysis. Fig. 4 gives an overview of the site. The newly generated blocks are highlighted in black while the metric reachable people within ten min. walking is visualized on the surrounding building geometries on a gradient from red (low) to green (high). Next to building volumes, a street network (street center lines) is modeled for the context.

In order to define both parameter and evaluation criteria of the parametric model, we first need a scenario to base our decisions on. The new quarter should:

- Provide around 600 residential units

- Promote walkability of the whole area, act as a new local center and be well integrated in its surrounding quarters

Along with the construction of the quarter itself, an additional pedestrian railway overpass shall be added to improve connectivity between areas on both sides and to the new center. As we are in an early design stage, it is sufficient to generate a basic street network, block arrangement, building volumes, and switch between several possible locations for the bridge
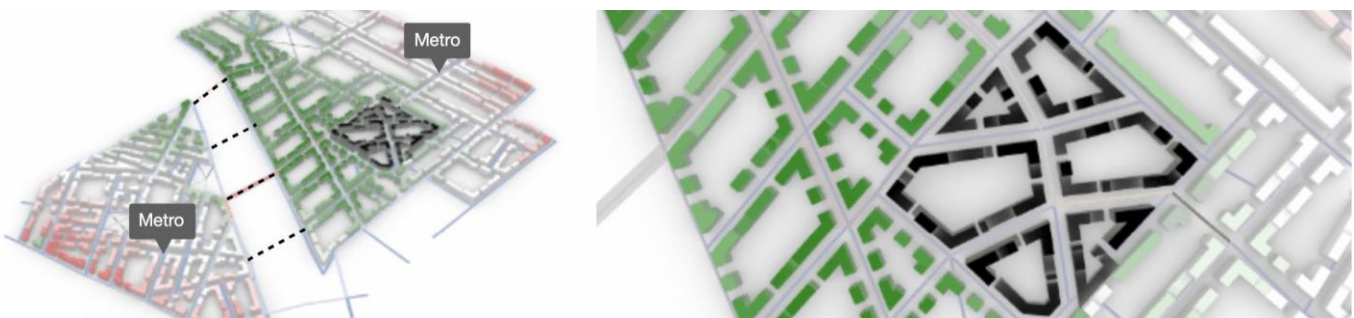

Figure 4. The Case-study area with context. Generated buildings are highlighted in black.

3.7. The Parametric Model

Foremost, we must set up the parametric model that shall help us to explore the design space within the project. It is important to find a compromise between the number of Based on the scenario, our generative module consists of three parts that are described below and illustrated in Fig. 5

Figure 5. Parameter of the generative model. From left to right: Angle of central street segment, bridge location, building volumes.

Street network and Blocks: For the parametric generation of the streets and blocks, we employ components from the DecodingSpaces toolbox. The only parameter is the rotation angle of a guiding street segment in the middle of our project site. The network generation starts from this segment and tries to connect to incoming surrounding streets (see the first image in Fig. 5). We chose this approach as the orientation of the street network has an impact on both pedestrian flows as well as thermal comfort (see below). For the sake of computational efficiency, the number possible angles to choose from is reduced to forty (40 parameter).

Bridge: Four possible locations for another railway overpass are preselected (4 parameter).

Buildings: We chose an open perimeter block typology for the building generation. A custom definition creates multiple buildings along the block perimeter. Each building can vary in size and height while maintaining the same overall GFA. Each Block has five buildings on average and five height configurations to choose from (150 parameter).

3.8. Evaluation Criteria and Fitness Objectives

The next step is to translate the stated quantitative and normative goals into a strategy we can act upon. measurable performance metrics the computational model can work with. Additionally, we also need to think about which parameters of the generative module do impact the proposed

Figure 6. From a design vision to measurable performance metrics.

indicators. Fig. 6 gives an overview of this process for the presented scenario.

Based on these indicators, we started to define our final fitness objectives for the EO. The results are shown in Fig. 7. It is noteworthy that we combined both simulation modules employed in this example, for instance, by weighting the value computed for thermal comfort in streets with the estimated pedestrian frequency.
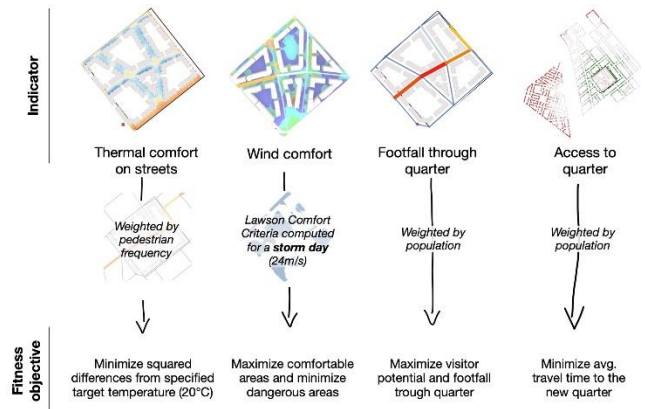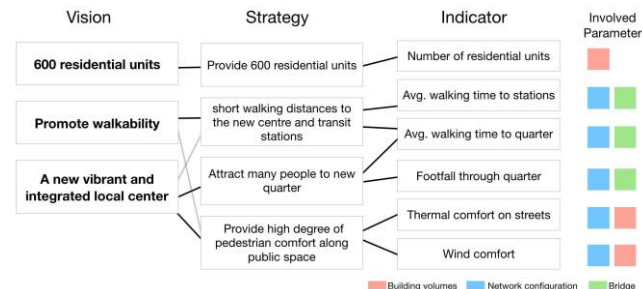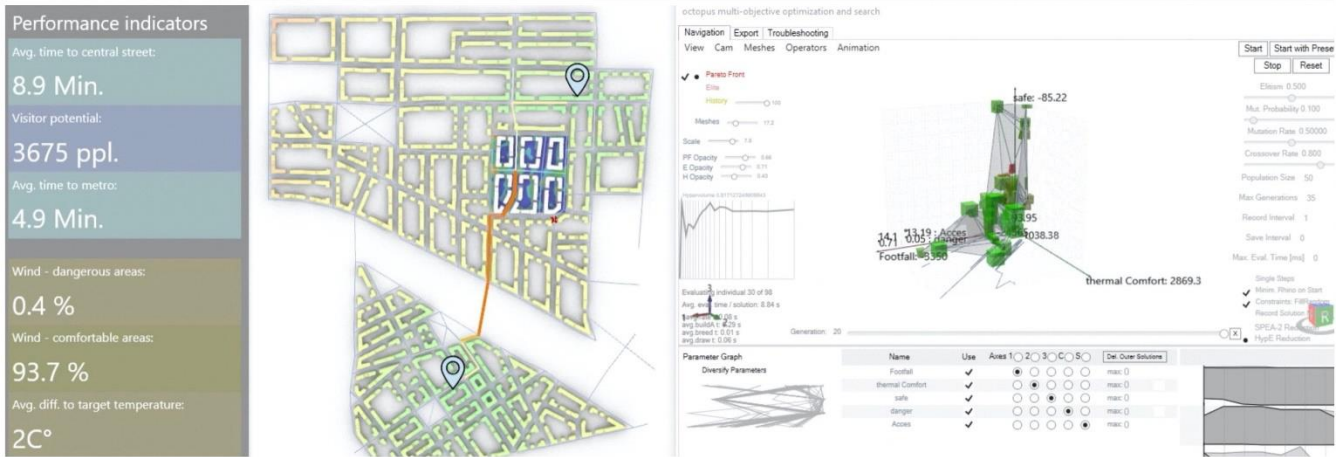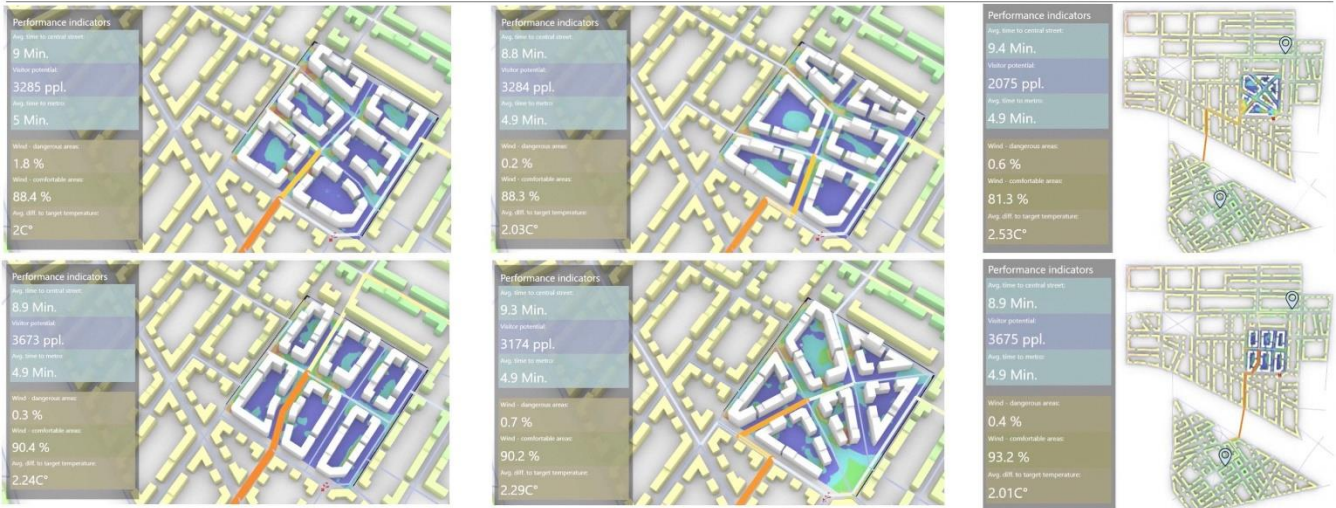

Figure 7. Defining fitness objectives.

3.9. Optimization

Having everything setup, the optimization process can be started. As mentioned above, we use the Octopus multi-criteria solver. Figure 8 shows snapshots taken during the optimization process. One complete iteration, including the generation of a solution and its evaluation through the presented indicator system, took around three and a half seconds. This allows to quickly gain an overview of the solution and performance space of a design problem.

A. Performance Metrics and Octopus Interface

B. Several solutions

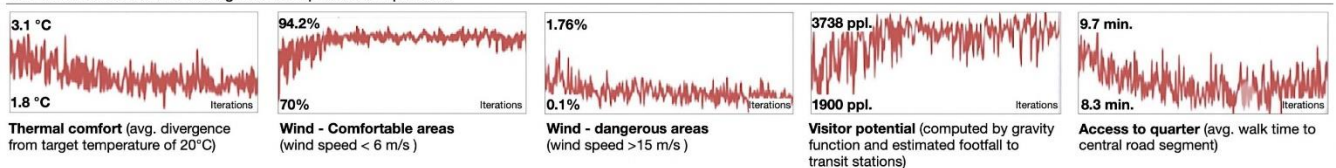**C. Performance metrics throughout the optimization process**

Thermal comfort (avg. divergence from target temperature of 20°C)	**Wind - Comfortable areas** (wind speed < 6 m/s)	**Wind - dangerous areas** (wind speed >15 m/s)	**Visitor potential** (computed by gravity function and estimated footfall to transit stations)	**Access to quarter** (avg. walk time to central road segment)

Figure 8. Results of the optimization process. Interface (A), several generated solutions (B) and trajectory of the fitness objectives throughout the optimization process (C).

4. RESULTS

We stopped the optimization process after two hours and approx. 2000 iterations, as we saw clear patterns of convergence. Figure 8 provides an overview of the Octopus interface (A), several solutions computed during the process (B) and the trajectory of fitness objectives throughout the optimization process (C). Regarding the explored solution and performance space, all selected fitness objectives showed significant variance (Fig. 8, C) indicating their relevance. Especially visitor *potential to the new quarter* showed a high variance (computed by the addition of the indicator *footfall through quarter* and *visitor potential to quarter*) with values ranging from 1900 to 3738. The observed values in *avg. difference to target temperature* varied between 1.8 and 3.1 °C, while the share of *dangerous areas* measured by the Lawson Wind Comfort Criteria varied between 1.76 and 0.1%. All five FO's showed clear patterns of convergence. For Instance, better performing solutions ended up choosing option C for the pedestrian overpass (see Fig. 5) and a street orientation connecting most directly to it.

Besides the ability of exploring the generated designs for factors and parameters that can be attributed to either well or poorly performing solutions, the generated data also allows to put new variants into perspective. For instance, manually designed ones (see Fig. 9). Having the design space explored,
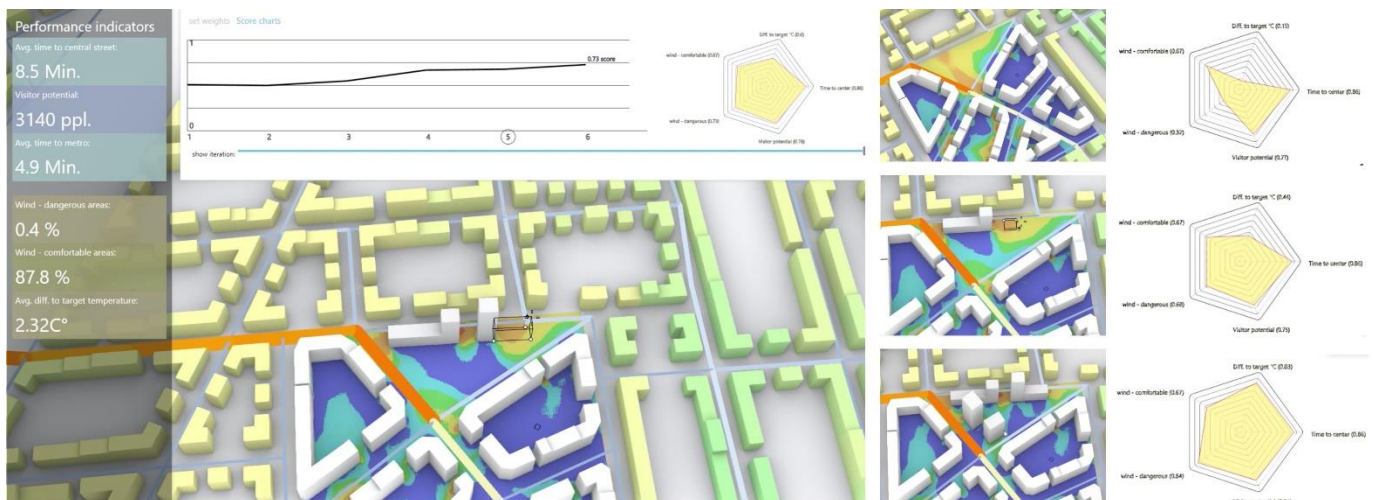
Figure 9. Data-driving manual design based on the automatically explored design space. Left: Interface and trajectory of summed overall score tracked throughout a manual design process. Right: three interventions and resultant performances of the five indicators plotted on a radar-chart. Lawson Wind Comfort criteria is displayed on the map (calculated for a storm day).

it is now possible to relate any new design variant to the previously observed performances potentials. Thus, allowing for a visualization and comparison on normalized scales, which greatly increases readability. Since, the re-computation of all metrics takes about 3.5 seconds, a manual, yet data-driven and iterative design process is made possible.

5. CONCLUSION

In this paper, we showcased the coupling of multiple simulation engines into one integrated optimization framework for the spatial configurations of urban systems. We combined traditional simulation techniques and spatial performance indicators (here a graph-based mobility model) with a novel, deep learning-based approach to simulate the environmental aspect of spatial configurations (here thermal comfort, wind speeds). The main goal was to evaluate the potential of the combination of this diverse set of performance metrics with generative modeling of a medium-sized quarter and multi-criteria evolutionary optimization.

First of all, the greatly reduced computation time of microclimate related metrics made their inclusion possible in the first place: Instead of minutes, the full calculation of one design variant (including generation and evaluation) took 4 seconds on average. A short computing time allows us to quickly explore the solution and performance space of a design problem. Thus, increasing the potential of similar workflows to be integrated into practice. Secondly, the different KPI's where successfully combined to define more meaningful fitness objectives compared to an separated employment of different simulation engines. For instance, the integration allowed us to weight the indicator thermal comfort on the street level with the estimated flow of pedestrians through a street.

Even though a diverse set of fitness objectives was used in this example, their number and individual design still resemble a normative and incomplete selection of possible and important criteria in the quest of measuring and optimizing the performance of urban morphologies.

As a too large number of fitness objectives render the application of evolutionary optimization algorithms as ineffective, the most important FO for a design problem should be identified and supplied to the EO engine. Alongside, additional performance indicators can, nevertheless, be computed for use later on. In this sense, the optimization is not the end of the design process. Rather the results gained can be regarded as an exploration of the design space and a pre-selection of solutions that need be further analyzed in detail.

REFERENCES

1. Purup, P. B., & Petersen, S. (2020). Research framework for development of building performance simulation tools for early design stages. Automation in Construction, 109.

2. Bazjanac, V. (2004). Building energy performance simulation as part of interoperable software environments. In Building and Environment (Vol. 39, pp. 879–883)9.

3. Naboni, E., Natanian, J., Brizzi, G., Florio, P., Chokhachian, A., Galanos, T., & Rastogi, P. (2019). A digital workflow to quantify regenerative urban design in the context of a changing climate. Renewable and Sustainable Energy Reviews, 113.

4. Chang, S., Saha, N., Castro-Lacouture, D., & Yang, P. P. J. (2019). Generative design and performance modeling for relationships between urban built forms, sky opening, solar radiation and energy. In Energy Procedia (Vol. 158, pp. 3994–4002).

5. Nagy, D., Villaggi, L., & Benjamin, D. (2018). Generative urban design: Integrating financial and energy goals for automated neighborhood layout. In Simulation Series (Vol. 50, pp. 190–197). The Society for Modeling and Simulation International

6. Galanos, T. Chronis, A. Vesely, O. Aichinger, A. & Koenig, R. (2019) Best of both worlds – using computational design and deep learning for real-time urban performance evaluation. 1st International Conference on Optimization Driven Architectural Design

7. Koenig R., Schmitt G., & Standfest M. (2018). Cognitive Computing for Urban Design. In: The Virtual and the Real in Planning and Urban Design: Perspectives, Practices and Applications (eds. C Yamu, A Poplin, O Devisch and G De Roo), pp. 93–111. New York, NY: Routledge

8. .www.decodingspaces-toolbox.org

9. Huff D. (1963). *A probabilistic analysis of shopping center trade areas*. In Land Economics (Vol. 39, pp. 81-90).

10. Sevtsuk A., Kalov R. (2017): Patronage of urban commercial clusters: A network based extension of the Huff model for balancing location and size. Environment and Planning: Urban Analytics and City Science (Vol. 45(3), pp. 508–528).

11. Düring S., Sluka A., Vesely A., König R. (2019): Applied Spatial Accessibility Analysis for Urban Design: An integrated graph-gravity model implemented in Grasshopper. Proceedings of the 37th eCAADe and 23rd SIGraDi Conference (Vol. 3, pp. 333–342).

12. Isola, P., Zhu, J., Zhou, T., & Efros, A.A. (2016). Image-to-Image Translation with Conditional Adversarial Networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5967-5976.

13. www.ladybug.tools

14. www.openfoam.com

15. www.food4rhino.com/app/octopus

Material Studies

Wood-based Responsive Systems:
A Workflow for Simulating, Predicting and Steering Material Performance in Architectural Design

Vasiliki Fragkia[1], Isak Worre Foged[2]

[1,2] Department of Architecture, Design and Media
Technology, Aalborg University
Aalborg, Denmark
[1,2] {vafr, iwfo}@create.aau.dk

ABSTRACT

The presented paper address the complex condition of analyzing and simulating anisotropic responsive material systems in architecture, focused on thermally-activated wood envelopes. The objective is to define, develop and present a design integrated workflow for simulating, predicting and steering material performance, as a key factor for the development of adaptive material systems, developable into adaptive building envelopes. Using methods from associated disciplines of engineering, material science and computer graphics, the studies employ adopted and adapted techniques for advanced material analysis and simulation. From this approach, we illustrate a new cyclical design method for high predictability and simulation of anisotropic material systems, with low scale specification data. With a special focus on hybrid responsive systems and functionally graded materials (FGM), we ask how to integrate material performance, engage with high degrees of interdependency and allow the emergence of design feedback between the multiple scales of architectural construction. With a parallel equally concerned approach on material specification and data processing methods, the experimental computational and prototyping studies conclude with a series of critical factors for high precision steering of material performance in an architectural scale.

Author Keywords

Design simulation integration, active material simulation, material specification, computer vision

ACM Classification Keywords

Design, Simulation, Performance

1   INTRODUCTION

Materials were for years conceived as passive receivers of a previously generated form. For that reason, the range of materials used for construction exclude anisotropy and are mostly restricted to homogeneity of pre-fabricated defined shapes (plates, cubes, pipes), for optimizing performance within a construction environment [1]. However, there is a great variety of heterogeneous materials in nature which have to go through several industrial processes and, oppositely to any sustainable logic, result in much material waste before becoming suitable for a standard fabrication environment.

As contemporary building culture is finding itself in an era of resource scarcity on materials needed for construction, [2] this challenges profoundly the perception of materiality into architectural design, as well as the established relationship between the processes of design and physical making. This challenge calls for development of new material practices and design workflows that optimize resource allocation and reduce waste [3].

Moving from the paradigms of standardization and mass production, these new practices position the making of functionally graded materials (FGM) for site and use specific application, as central means of building smarter with less.

This paper discusses new methods by which functionally graded materials can be specified and produced, using robotic fabrication. The aim of this paper is the identification, development and implementation of a simulation method and model for active anisotropic materials, focused on application for material driven building active envelopes.

Through a case study, it explores an integrated workflow for simulating and predicting highly anisotropic material composites, interfacing high fidelity with low scale specification data. The experimental case study examines how simulation of material heterogeneous performance can be integrated into architectural design and fabrication, as an aspect of sustainability and speculates on future fabrication paradigms for lighter, adaptive, material-driven architecture.

2 SIMULATION IN DESIGN

The current forming of the architectural design practice is facing increasing demands in terms of predictability and performance of its outcomes. From design to construction and operation, today different means of simulation cover a wide range of architectural concerns, including the simulation of a design's environmental, structural and material performances, its fabrication and assembly.

As we deal with more and more complex design tasks, employing different material systems and different scales of material interaction, the emergence of new simulation methods and workflows, that interrupt the otherwise linear progression from design to analysis, seems to be critical. This becomes even more essential when designing for and with material performance of highly anisotropic responsive material systems. Current studies provide a wide range of simulation methods, varying from Geometric Representation Models (GRM) [4], Spring –based Simulation Models (SSM) [5], use of Finite Element Method (FEM) [6] and Isogeometric Analysis (IGA) [7, 8].

Representing various levels of resolution and accuracy, all aforementioned methods provide extensive information on the system's behavior. However the development of a responsive architectural system based on anisotropic material properties of wood, presents methodological and technical challenges that need to be considered and further implemented, within the design chain. One such aspect involves the homogeneity between samples. Current designs rely on a homogeneous, linear grain pattern for all elements and, despite the fact that there are wood species that provide very linear grain topology, all samples still require fairly rigorous quality control to ensure consistent grain curvature [9].

Given the organic and extensive parameters of wood as a vascular tissue, the excessive or uneven distribution of heat across the grain can cause uneven distribution of strain in the system, which may consequently result in microstructural damage (plastic deformation) and an overall reduction of responsiveness over time.

Moreover, the geometric arrangement of the responsive elements can lead to areas that are over exposed to UV radiation, over extended periods of time. This increases material fatigue [10] and can affect the operational life of the system. Thus, it is a performative aspect that cannot be excluded from the simulation methods used in the design process.

Under this scope, these challenges seem to address the importance of a new material design practice, where simulation is understood as an integrated part of the whole chain from design to construction, introducing constant feedback and cyclical thinking between the physical and the digital design environments, as well as between the different stages and scales of architectural thinking.

In other words, the implementation of adaptive material systems in architecture and construction entails the integration of an adaptive simulation workflow within architectural design process.
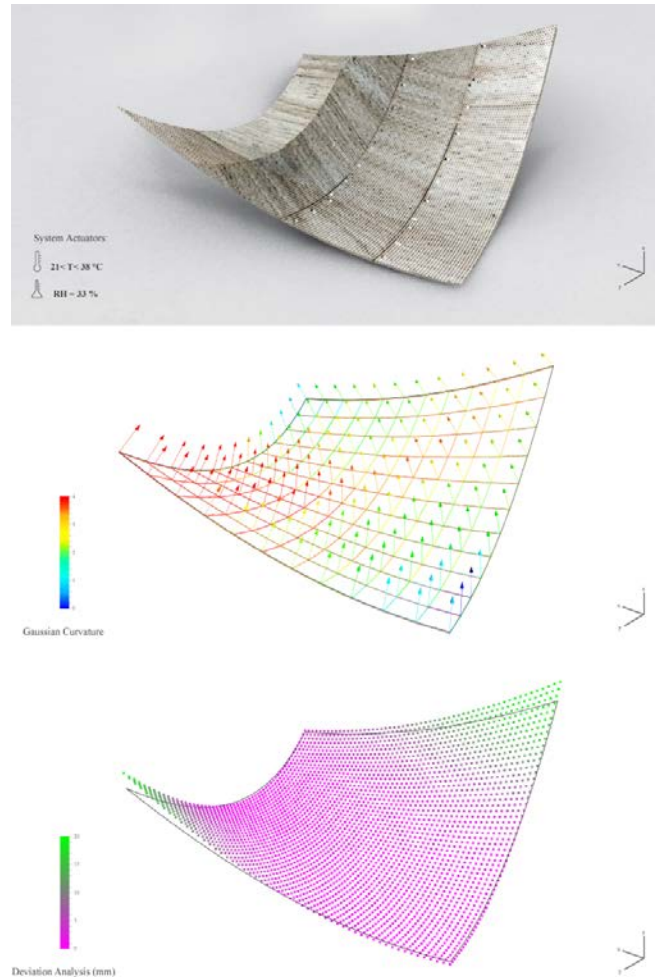

Figure 1. From top to bottom: 3D scanning of activated multi-grain timber composite (38 °C), Dynamic Simulation displaying Gaussian curvature and Deviation Analysis between physical and simulated material performance.

3 LIGHTWEIGHT WOOD-BASED RESPONSIVE COMPOSITES

The integration of active materials with the capacity to compute behavioral response, based on environmental conditions, represents a growing field in architecture and construction, in which digitally controlled mechanisms are replaced by material systems that are based on the design and organization of the material itself [5]. Understanding in depth the characteristics of wood as a natural material can give us a good insight into the possibility to re-conceptualize the material use in architectural design, in order to fulfill formalized design requirements. However, programming a system's response in the material level requires the development of material-driven computation

strategies that incorporate a range of design, fabrication and actuation parameters.

With the intention to study correlations between a material exergy-based system [11], which utilizes energy that is available in the environment in order to actuate responsive behaviors, and the environmental dynamics of temperature, a series of critical influencing parameters were derived:

a. Wood type and sample selection. Species that show high thermal behavior increase responsive performance. Given the environmental exposure of the material in continuous actuation cycles, a number of other performance parameters, such as density and Modulus of Elasticity, were taken into account.

b. Homogeneity in fiber directionality. This allows for high precision control and steering of the system's performance.

c. Material layout. Width/Length ratio, thickness and shape variables determine the level of performance, as well as the timeframe within responsiveness occurs.

d. Dimensional limitation of samples. Quarter-cut veneer ranges from 40-120 mm in width, which directly constraints the working width and the scale of the design elements.

e. Homogeneity of samples. Even if oak is considered homogeneous wood type, it is still required a detection method of wood directionality for informing both the simulation and fabrication process, so as to ensure consistent behavior.

Using temperature as the environmental stimuli of the responsive system, the selection of the right material is approached through an intensive research and understanding of the microstructural principles that facilitate the thermal actuation of wood [12]. Based on its high thermal expansion coefficient, tangentially to growth rings (quarter-cut veneer), the presented case study uses oak composite veneer elements to transfer thermally activated dimensional changes into responsive bending, in a wood-based architectural envelope. The ability of the single discrete wood elements to implement reactive, shape-changing behaviors, triggered by environmental stimuli, gives the opportunity of large scale responsive architectural assemblies.

Under this scope, several material compositions were tested in relation to their bending performance, from single layer veneer elements to bi-layer (combining two material layers) and multi-layer (combining more than two material layers) composites. Specifically, the material study presented explores the case of gradual material layering deposition of various grain directionality, allowing for local control of bending stiffness and direction. Within these investigations, the hierarchy and the anisotropic characteristics of the functionally graded composite is proved to be the driving force of the programmed deformation and the basis for developing a simulation model for anisotropic responsive composites.

Figure 2. Material studies to identify, understand, develop and test novel composites for exergy based adaptive behavior.

4 THE DESIGN SETUP: A WORKFLOW FOR INFORMED DESIGN ITERATION

The interdisciplinary frameworks that architecture needs to adopt in order to address the current advances and challenges in building practice, have resulted to an inevitable coexistence of different levels of precision, methods and calculation tools, which leads to a larger research enquiry of how simulation of different levels of fidelity can interface along the design process [13].

The presented case study devises an experimental workflow for wood-based responsive systems, equally concerned with design at the scale of the material, the element and the structure, in which an adaptive iterative design process is formed. While existing solutions have implemented a sequential approach to simulation studies from low to high precision and from material to structure, this workflow explores a parallel approach [14] where simulation, prediction and steering coexist in a non-hierarchical framework, allowing for bi-directional information flows, within the design chain. More specifically, physical material experiments on various wood fiber deposition configurations define the threshold of the bending performance of the responsive elements, as well as the basis for the assembly topology.

Using Computer Vision methods, a series of qualitative data from the responsive elements are being collected over several activation cycles of the system. Alongside, these data interface with the spring-based simulation model, providing us with quantitative understandings of the system's adaptive and anisotropic behavior.

Finally, with the use of evolutionary algorithms, these two computational frameworks are reconciled in overall fitness functions, for addressing specific design requirements, providing us with geometric data that inform robotic fabrication.
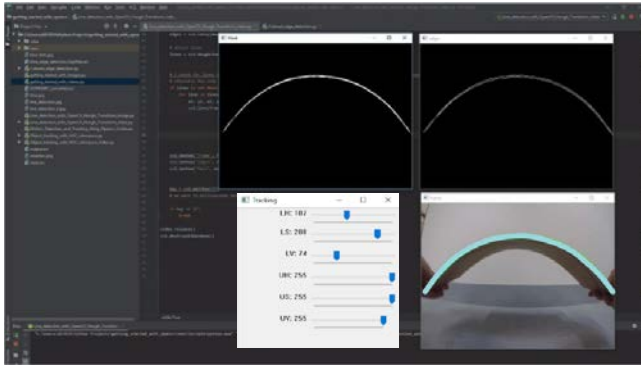
Figure 3. Object Tracking and Bending Ratio Detection using Linear Hough Transform (LHT) method in OpenCV.
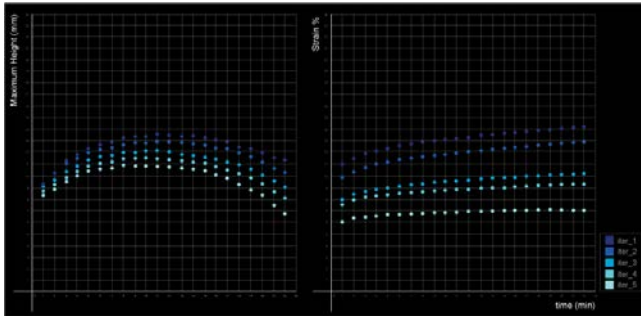

Figure 4. Exporting quantitative data of maximum bending ratio and strain of oak responsive veneer element, after 5 activation cycles.

4.1 Simulation

The project takes advantage of the bending behavior of multi-directional oak veneer composites in a reciprocal open topology, in order to form a thermal adaptive membrane. The presented integrated workflow is being applied for the simulation, prediction and steering of a 1 x 1 meter reactive specimen behavior. While single-directional responsive curvature can be estimated mathematically

through Geometric Representation Models (GRM) [4], predicting the interaction between multiple parts can be a computationally intensive and complex task. For that reason the form-finding and overall structural behavior of the design system, in terms of deformation and bending stresses, is simulated with spring-based physics engine, using K2 solver in Grasshopper, Rhinoceros. The form-finding of the reactive elements occurs in two levels. In a material level, the bending behavior of a multi-layered and multi-fiber directional composite is being computationally represented by assigning the grain orientation of each fiber into weighted values for hinges, along the internal edges of a triangulated mesh. The resulted vertex map of graded weighted values is being informed by the elastic limitations of the material and recalibrated based on measurements from the physical experiments.

The accuracy of the simulation at this level can be up or down sampled by changing the resolution of the mesh topology. The resulted form-found mesh is then reduced into a low resolution mesh and simulated within a system of reactive interdependent elements. This second level of simulation ensures that the system performs within the limits of the material properties and orchestrates the design of a more complex multi-part system behavior.

K2 constitutes a fast and accurate method that provides and permits direct interaction of the designer at any point of the simulation process, allowing for visual evaluation and constant feedback on the design. Although this simulation method can provide extensive information on the system's behavior, its integration into a design model requires increased computational resources. In addition, this method becomes challenging in providing accurate geometries at intermediate simulation stages, (not fully curled) [15] which can be critical in calibrating an environmental responsive highly anisotropic system. To address the aforementioned challenges, the case study shifts towards a more agile adaptive workflow, by introducing Computer Vision algorithms, parallel to the main K2 simulation model.
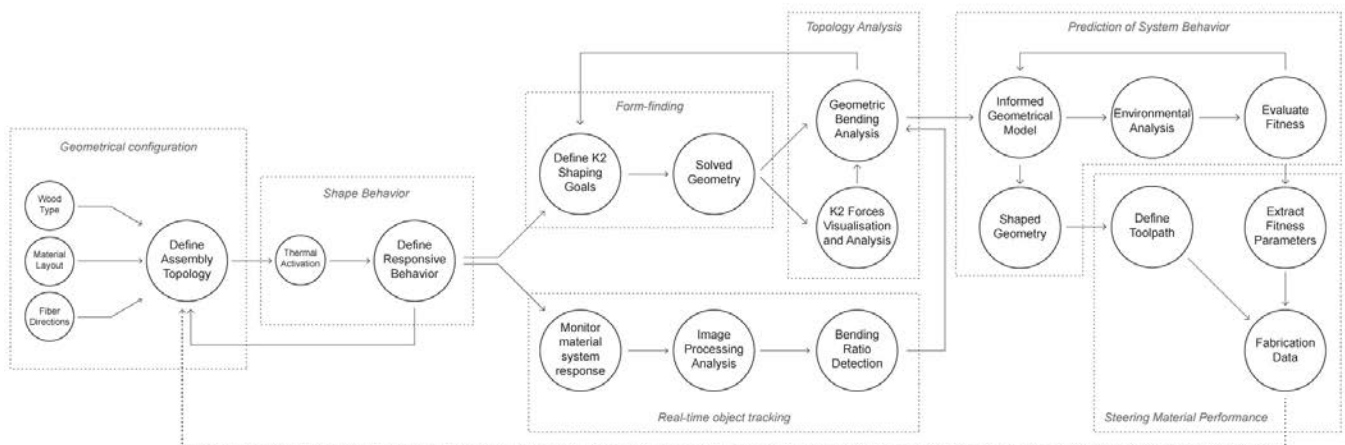

Figure 5. Flowchart of the design integrated simulation workflow for thermally activated wood composites.

516

Computer vision is an interdisciplinary field, including methods for acquiring, processing, analyzing and understanding digital images, and extraction of high-dimensional data from the real world, in order to produce numerical or symbolic information in the forms of decisions [16]. The presented project uses OpenCV programing library and, specifically applies Linear Hough Transform (LHT) methods for real time object tracking, as well as grain detection of the wood-based responsive composites. LHT is a voting process aimed to detect the existence of a line within a certain class of shapes, in the pixel space of an image [17], by assigning to each pixel the parametric equation that describes a line:

$$r = x\cos\theta + y\sin\theta$$

The collected data are in the form of arrays of 3-dimensional points, where each array represents an analysed video frame and can be easily imported to any design simulation environment.

Along with the K2 simulation engine, this tracking method provides us with extensive information regarding the responsive behavior in the level of the material, while allows for real time bi-directional data flow between the physical and the digital environmnets of the design chain.

4.2 Prediction

When predicting and steering material performance as a key factor for the development of adaptive material systems, developable into adaptive building envelopes, multi-parameter optimization can act as an instrumental tool to mediate conflicting goals of environmental control, performative limitations of the material and changing user requirements[18]. For the prediction of the specified geometrical configuration of the responsive wood-based composites, the enquiry employs Galapagos Evolutionary Solver for Grasshopper, Rhinoceros, developed by David Rutten [19]. Along with the technical setup of the evolutionary engine, there are a series of critical parameters that define the accuracy, resolution and speed of its operation performance, such as describing the fitness function and altering the variables of the population and mutation rate. The evolutionary algorithm derives its fitness parameters by an environmental benchmark model, created in Ladybug plugin for Grasshopper, Rhinoceros [20]. The model showcases a parallel evaluation study of how the various bending states of the responsive membrane affect the visual, self-shading, radiation and energy analysis, providing the designer with a direct relation between material behavior and performance. Under the presented case study, the parameter of visual analysis has been selected to describe the fitness function within the evolutionary search. Through continuous adjustment of the bending responsive elements in various configurations as input parameters, this method provides an accurate prediction of the optimized result that best meets the goal outcome. By enabling a continuous and interactive search

loop, where resolution is continuously redefined, the evolutionary engine generates fit candidates disengaging the designer from compromising between extensively complex coexisting design and performance demands.
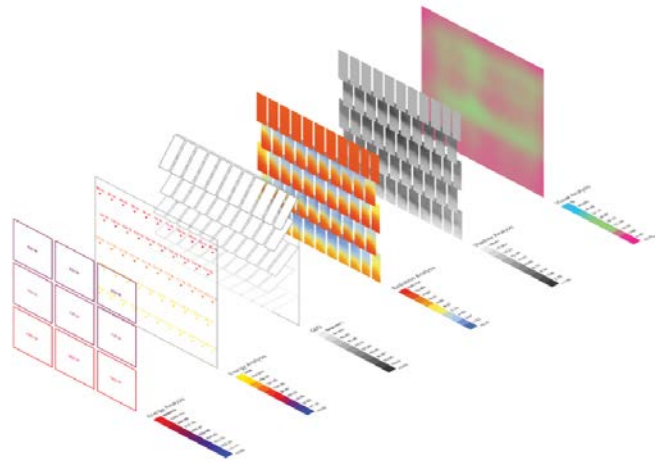

Figure 6. Benchmark model for Environmental Analysis of the Adaptive Envelope design. From left to right: Energy Analysis (per m2), Energy Analysis (per reactive element), Geometric Bending, Radiation, Shading and Visual Analysis.

4.3 Steering

Along with the simulation and prediction methods, a customized robotic fabrication process has been developed for steering of the wood-based responsive material composites behavior. This process is formed around and for high resolution timber surfaces, using gradual material layering deposition of various grain directionality, allowing for local control of bending stiffness and direction. For these studies, the material used for the fiber deposition is quarter-cut oak veneer 0.6 mm. The robotic fabrication developed in the presented case study takes place in two levels –in a material and in a system level- and is being specified by three essential parameters- Material Production, Robotic Tool Development, and Robotic Path Planning. In a material level, the high resolution form-found mesh of each individual responsive composite element, resulted from the K2 simulation environment and evaluated through the Galapagos evolutionary solver, is being down sampled and deconstructed into the individual wood elements of different grain directionality. This process is achieved by remapping the grain directionality data acquired through the Open CV algorithm into the form-found mesh. As these methods are interconnected and inform each other in a parallel set-up, the behavior accuracy of the specified functionally graded wood- based composites are highly interdependent to the resolution of the form-found mesh and the fidelity within which K2 goals where defined, including the geometric constraints, the material elastic and mechanical properties, applied loads and other energies [21].This method assigns directly 3d mesh topologies to 2d fabrication data and creates an

important link between the various simulation models and the robotic fabrication process. In the system level, steering a responsive membrane for addressing formalized design requirements is achieved by the assembly of the functionally graded wood composites, produced from the aforementioned material-based fabrication process. Once again, the classification, clustering and distribution of these composites, within a large scale architectural assembly, is linked to the optimized geometric solution of the evolutionary engine. Based on the above fabrication requirements and challenges, a custom robotic end effector was developed using pneumatic suction caps, with an adaptive operational area for elements between 10-110 cm long and with the ability to transform from linear to rectangular configuration, with maximum operational area of 40 x 40 cm, serving various scales and material layouts. The robotic movement of a single "pick and place" movement is defined by the operational area of the assembly which is directly linked to the scale of the intended responsive membrane. The robotic paths are planned as a repetitive trajectory between the material stock (picking) and the target point of the final assembly (placing) within the topology of the greater architectural surface. This allows for an optimized fabrication interface and facilitates a rather complex compositional overlaying of various elements. Moreover, there has been carried out extensive research on different attachment methods for binding wood fibers together, ranging from synthetic adhesives (acrylics, liquid seals, lockers, structural glues) to sewing. In the presented prototypical method, binding is not yet integrated in a single end effector development, but occurs as a post fabrication process, in the form of local customized PLA snap joints, allowing for fast assembly and disassembly of the composites.
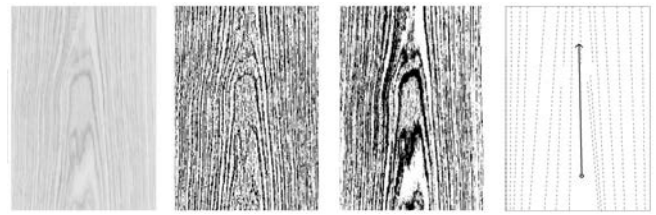

Figure 7. Grain detection using Linear Hough Transform (LHT), OpenCV. From left to right: Original black and white image, Binary image, Line creation following grain directionality, Grain Direction vector.
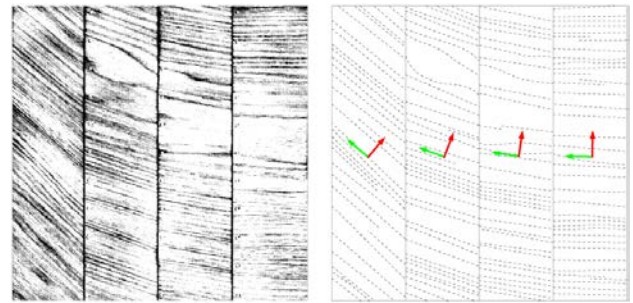

Figure 8. Extraction of robotic path coordinates for clustering and assembly of FGM multi-grain oak composites.

5   RESULTS

Moving from a binary interpretation of simulation, expressed as a linear progression from design to evaluation, this prototypical method allows us to re- conceptualize simulation practice as an assemblage of distributed events of various resolutions and scales, along the design chain. Within this process, the role of the designer lies on the interfacing of those parallel simulation and design models, as well as the data processing between the various resolution design agencies.
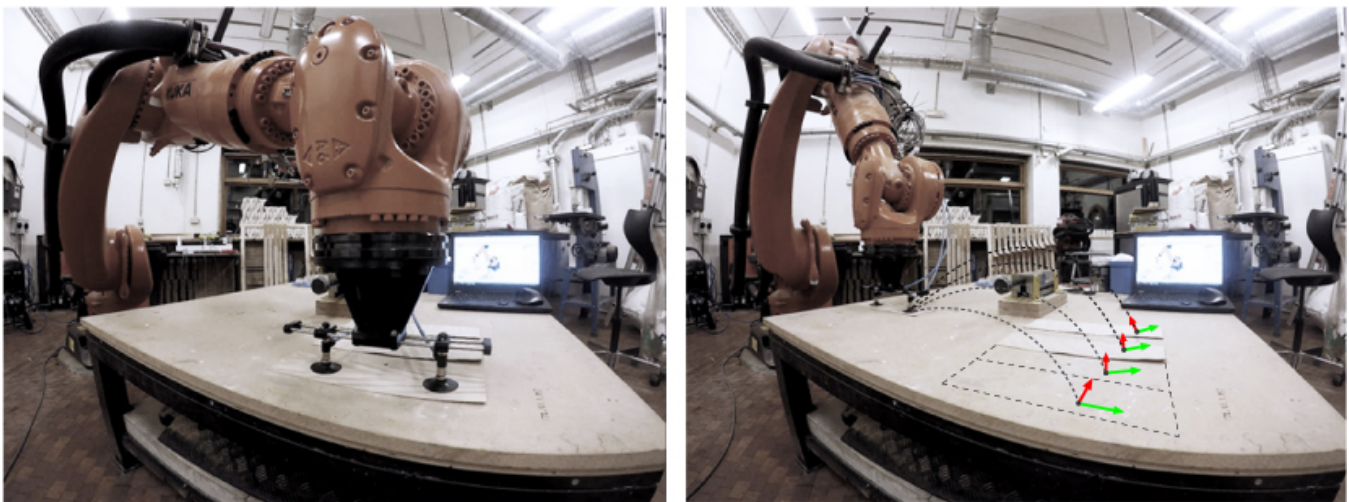

Figure 9. Robotic assembly of various grain direction reactive elements for an exergy-based thermally active wood envelope.

The presented case study forms an experimental workflow that can simulate, predict and steer non-linear anisotropic behaviors, by interfacing highly specified behavioral descriptions, with low specification data. The case study proves that the integration of a multi–scalar adaptive workflow from design to production, can inform the design chain with a series of critical results, allowing for high precision performance steering of responsive systems:

a. Tracking real time the material behavior and capturing all in between stages of the material performance, allows for qualitative and quantitative evaluation of the material decay over the responsive operational cycles.

b. Benchmarking bending ratio in relation to temperature conditions and time. This can provide us with a high resolution 3d mapping of the wood responsive composite's thermotropism [22], where each timeframe can be assigned with a unique material state and a surface temperature value.

c. Grain detection can up sample the K2 simulation accuracy by creating a higher resolution differentiated mesh for dynamic relaxation in the scale of the material, providing high control of bending stiffness and curvature.

d. Informed Design Iteration by identifying infidelities or mismatches between the simulated and the physical material behavior. This could be useful to simply re-inform and re-parametrize specific features of the integrated workflow in accordance to physical world data.

e. Additionally, this evaluation could allow for fabrication optimization. Aiming for a highly accurate pre-programmed responsive material system behavior, this feedback can be used to detect and exclude specific material elements from the fabrication process, for not fulfilling the desired performance requirements.

6   DISCUSSION

Further development needs to occur in the interfacing between the various resolution simulation platforms and the fabrication process. This careful classification of data exchange will also allow for applying the presented methods in a bigger architectural scale, maintaining accuracy. The binding method applied on the wood-based composites affects the overall performance and needs to be integrated into one fabrication workflow. Moreover, bio-based adhesives such as cellulose need to be investigated. Further investigation on how the robotic fabrication workflow can be refined needs to be addressed. Feedback processes during the fabrication process could allow for adaptive tool paths and iterative informed material layer deposition. Finally, along with the multi-grain material deposition that is being investigated in the presented study, the shift towards a

multi-material layering could introduce new hybrid properties in the responsive architectural system.

7   REFERENCES

1. DeLanda, M. The New Materiality. Architectural Design (2015), Vol. 85(5) 16–21

2. Valero, A. Physical geonomics:Combining the exergy and Hubbert peak analysis for predicting mineral resources depletion. Resources, Conservation and Recycling (2010), 1074-1083.

3. Ramsgaard Thomsen, M., Nicholas, P., Tamke, M., and Svilans, T. A new Material Vision. In Diniz, N.(eds), Data and Matter, Routledge (2019).

4. Foged, I., Pasold, A. An oak composite thermal dynamic envelope. In: Structures and Architecture - Proceedings of the 3rd International Conference on Structures and Architecture, ICSA (2016).

5. Wood, D., Correa, D., Krieg, O., Menges, A. Material computation-4D timber construction: Towards building-scale hygroscopic actuated, self-constructing timber surfaces. International Journal of Architectural Computing (2016), Vol. 14(1) 49–62.

6. Deleuran, A. H., Pauly, M., Tamke, M., Friis Tinning, I, and Ramsgaard Thomsen, M. Exploratory Topology Modelling of Form-Active Hybrid Structures. TensiNet (2016).

7. P. Längst, A. M. Bauer, A. Michalski, and J. Lienhard. The Potentials of Isogeometric Analysis Methods in Integrated Design Processes. In Proceedings of the IASS Annual Symposium 2017 Interfaces: Architecture. Engineering. Science (2017).

8. Baseta, E., Bolinger, K. Construction system for reversible self-formation of gridshells: Correspondence between physical and digital form. ACADIA 2018 Re/ calibration: on imprecision and infidelity (2018).

9. Menges A, Reichert S. Material Capacity: Embedded Responsiveness. Architectural Design (2012) Vol. 82 52-59.

10. Smith, I., Landis, E., Gong, M. Fracture and Fatigue in Wood. Wiley (2003), 123-141.

11. Shukuya, M. Exergy concept and its application to the built environment. Build. Environ (2009), 44, 1545–1550.

12. Foged, I., Pasold, A. Material Studies for Thermal Responsive Composite Envelopes. Proceedings of the 37th eCAADe Conference (2019), Vol.1 207-214.

13. Ramsgaard Thomsen, M., Tamke, M., Nicholas, P., Deleuran, A. H., Ayres, P., La Magna, R., Gengnagel, C. Simulation in Complex Modelling. In Proceedings of SimAUD (2017)

14. Winsberg, E. Science in the Age of Computer Simulation. University of Chicago Press (2010)

15. Reichert, S., Menges, A., Correa, D. Meteorosensitive Architecture: Biomimetic building skins based on materially embedded and hygroscopically enabled responsiveness. Computer Aided Design (2015), Vol.60 50-69.

16. Learning OpenCV: Computer Vision with the OpenCV Library by Gary Bradski, Adrian Kaehler 2011

17. Richard O. Duda and Peter E. Hart (April 1971). Use of the Hough Transformation to Detect Lines and Curves in Pictures. Artificial Intelligence Center.

18. Sastry, K, Goldberg, D 2005, Genetic Algorithms. Search Methodologies pp 97-124. Springer.

19. http://www.grasshopper3d.com/group/galapagos

20. Roudsari, M.S., Pak, M., Smith, A. and Gordon Gill Architecture. Ladybug: A Parametric Environmental plugin for Grasshopper to help Designers create an environmentally-conscious design. Proceedings in Conference of International Building Performance Simulation Association

21. Bauer, A.M., Langst, P., La Magna, R., Lienhard, J., Piker, D., Quinn, G., Gengnagel, C., Bletzinger, K.U. Exploring Software Approaches for the Design and simulation of Bending Active Systems. Proceedings of the IASS Symposium (2018)

22. Robertson McClung, C.; Davis, Seth J. Ambient Thermometers in Plants: From Physiological Outputs towards Mechanisms of Thermal Sensing. Current Biology, Vol. 20, no. 24

Theory, Views & Visions

Indoor Localization and Building Occupancy Count Estimation using LTE-A Ultra-Dense Networks

Ala'a Al-Habashna[1], Vinu S. Rajus[1], Nicolas Arellano Risopatron[2], Cristina Ruiz-Martin[1],

Stephen Fai[2], Liam O'Brien[3], Gabriel Wainer[1]

[1]Systems and Computer Engineering   [2]Carleton Immersive Media Studio   [3]Civil and Environmental Engineering

Carleton University   Carleton University   Carleton University

Ottawa, Canada   Ottawa, Canada   Ottawa, Canada

{alaaalhabashna, cristinaruizmartin, gwainer}@sce.carleton.ca

{vinu.rajus, nicolasarellanorisop, stephen.fai, liam.obrien}@Carleton.ca

ABSTRACT

Recent studies have shown that awareness of occupants' presence, location, and count can be used for optimizing building operations and management. We present innovative ideas on how to improve such building sustainability reducing CO_2 emissions and energy consumption, through occupants' localization and tracking, and building occupancy count estimation. We propose to use Long Term Evolution-Advanced Ultra-Dense-Networks to locate users and to estimate the occupancy count. Furthermore, we discuss how awareness of occupants' location and count will be integrated into other parts of our project, namely, Building Information Modeling (BIM), building simulation, design, retrofitting, and studying occupant's behavior.

Author Keywords

LTE-A; Occupancy count estimation; Indoor localization

ACM Classification Keywords

Applied computing → Architecture (buildings) → Computer-aided design.

1 INTRODUCTION

There is an urgent need to improve processes for sustainably managing buildings over its entire lifecycle. Sustainability needs to be considered in building design, construction, and operation. According to [3], buildings consume approximately 40% of the total primary energy use in the U.S. and Europe and 27.3% in China. Total building energy end-use is dominated by *space* and *water heating*. This is translated in approximately 40% of total direct and indirect CO_2 emissions [8]. There is potential to improve energy efficiency although the total floor area is expected to grow by 60% by 2040 (according to the Efficient World Scenario, by 2040, buildings could be around 40% more energy-efficient than today [8]).

To achieve these improvements, buildings should be designed, and run near-optimally to maximize performance and user comfort. Improving building control, operations, and

management is low-cost and non-invasive, it can address inefficiencies and improve energy usage. For example, Natural Resources Canada's Office of Energy Efficiency initiatives for existing homes saved 388,000 tons of greenhouse gas emissions and 3.991 PJ of energy [14]. Our research aims to improve building sustainability, investigating new methods that use Building Information Modeling (BIM), exploiting data obtained from different sources such as sensors, network operator elements (such as Base-Stations) and user deployed elements (such as Wi-Fi Access Points). The architecture of the proposed effort is presented in Figure 1.
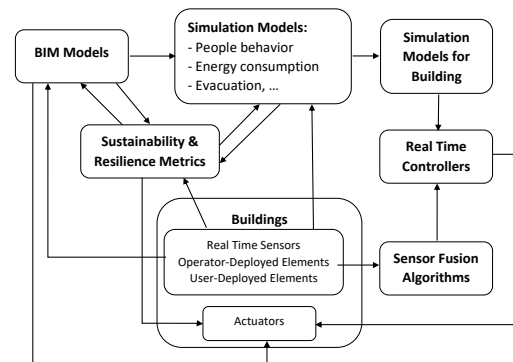

Figure 1. Software Architecture.

As a part of this research, we are investigating standard metrics for evaluating the sustainability and resilience of buildings. These metrics will be used on all the phases of the project: (1) in the design phase, as criteria to select the designs that meet the requirements; (2) in the optimization phase to evaluate the results of building simulations; and (3) during operation to adjust the building controllers.

We use Generative Design to explore building designs (as BIM models) that satisfy all the design requirements. Once we have a set of designs, we run simulations to evaluate energy consumption, evacuation time, etc., including different user's behavior (i.e. movement patterns, preferences, etc.).

These simulation results are used to select designs or to suggest a reiteration in the design process. They are also used to build the control systems using the Discrete Event Methodology for Embedded Systems (DEMES) [21], which allows transforming models into the actual controllers used in buildings. The raw data from sensors, operator-deployed elements, and user-deployed elements are combined with sensor fusion algorithms to supply fault tolerance. Data extracted from the building is also used to generate Data-Driven Simulation and to populate the BIM models for building management.

We present a part of this software architecture, namely, the set of methods proposed for occupant localization and tracking, and occupancy estimation. Recent studies showed that this is an important aspect for optimizing building operations and management [18], as awareness of these aspects can help to deliver building services (e.g., lighting) when and where needed. We propose using Long Term Evolution-Advanced (LTE-A) Ultra-Dense-Networks (UDNs) to locate and track occupants and to estimate their count. The advantage of this method over approaches based on sensor data (e.g. measurement of CO_2, camera data, humidity, etc.) is that it does not need to install and set up sensors in the building: LTE-A is already available for cellular communications. Our approach is expected to be accurate for occupancy estimation because LTE-A covers areas where Wi-Fi or Bluetooth devices cannot. Additionally, the geographical area covered by cellular networks can provide valuable data. For example, we can track occupants to understand individuals' movement and extract emergent behavioral patterns.

We also detail how accurate localization and occupancy estimation can benefit other aspects of our software architecture, such as the development of models and simulations of user behavior. We expect that having advanced building controllers will increase comfort while reducing energy consumption and CO_2 emissions. These new methods for localization will also provide accurate data for building retrofitting and additional data for BIM models used for building management. For example, the data about the time slots when the corridors of a business building have minimum occupancy can help to schedule maintenance and cleaning operations. Movement patterns from the students on a University Campus may suggest that the location of the food court is creating traffic on an area initially designed for study (i.e. an area that should not have traffic or noise). This information suggests a change in the location of the study area or the food court, which could be used when retrofitting operations are needed in that area of the Campus.

The data provided by these new localization methods are integrated into BIM models for visualization purposes. BIM models provide Digital Storytelling, i.e., digital techniques to create narratives that transform data into information for end-users. If we want to present our findings to non-specialized users or Architecture, Engineering, Construction, and Operations (AECO) professionals, we need to use a common language to all of them. There are visualization features provided by some BIM platforms that allow users to disseminate data interactively and intuitively. This includes diagrams, adaptive geometry, and interactive parameters, among others. This way of disseminating the data allows to understand and interact with the results of the study.

The rest of the paper is organized as follows. Section 2 discusses indoor localization and building count estimation. Section 3 describes how to use LTE-A UDNs for occupant count estimation, user localization and tracking. Section 4 discusses user behavior on building performance, and occupant count estimation, localization and tracking for the development of user behavior models. Section 5 presents how BIM models benefit with this estimation and vice versa.

2 BACKGROUND AND RELATED WORK

Indoor localization is becoming important for location-based services, for instance, various mobile applications require accurate location of running smart devices, which shows the importance of indoor localization. We propose to use indoor localization for building occupancy count estimation to provide energy efficient buildings. Reducing this significant portion of the world's energy consumption [20] would help with energy shortage and reduce carbon footprints of buildings. In order to deliver building services to occupants in an energy-efficient manner, such services need to be provided in the correct time, location, and amount [4, 18]. This applies for a number of services such as lighting as well as heating, ventilating, and air conditioning (HVAC).

Much work in the literature has been conducted on building occupancy estimation [4, 18]. Most of the existing work is based on data that is extracted from sensors deployed in buildings. This includes passive infrared (PIR) sensors, CO2 sensors, temperature sensors, humidity sensors, pressure sensors, RFID tagging, camera data, keyboard and mouse activities. Other localization and occupancy estimation methods are based on Wi-Fi and Bluetooth signal sniffing. Such methods either use data extracted from Wi-Fi access points or use designated devices to sniff Wi-Fi and Bluetooth signal from surrounding devices to estimate the position of these devices or estimate the number of devices in a building [24].

A cellular network is one where the last link to the end user takes place over a wireless radio link. The coverage area of the network is divided into smaller areas referred to as cells. Each cell is covered by a stationary transceiver that is called the evolved Node B (eNB) [2]. Voice and data communication between the network and User Equipment (UE) take place over radio frequency links between the covering eNB and the UEs. The part of the network that includes the eNB, UEs, and connecting frequency links is called the Radio Access Network (RAN). eNBs are usually connected via a high-speed wired network called the backhaul.

LTE-A is standard for the 4th generation (4G) mobile networks introduced by the 3rd generation partnership project (3GPP) to satisfy mobile broadband services with higher data rates and Quality of Service [2]. Cellular and mobile networks witnessed an increasing demand for higher data rates and continuous growth of data traffic and number of subscribers. Furthermore, the number of devices to be connected will continue increasing exponentially due to Internet of Things (IoT) applications that can be deployed over cellular networks (smart cities, autonomous vehicles, etc.). Network densification is a key technology to satisfy these demands; it is achieved by increasing the density of elements in the RAN. This includes operator-deployed and user-deployed elements to increase coverage, frequency reuse, and achieved data rates [10]. Ultra-Dense Networks (UDN) are expected to be widely adopted in the future, to the point where each UE might have its own serving element.

Recent research considered employing the radio signals transmitted by LTE-A cellular networks for localization. The infrastructure of such systems is available for cellular communications. Furthermore, it can provide accurate results due to the wide spread of mobile devices and the ability to detect them, which can provide accurate estimation of occupants' headcount. The advantage of a cellular-based system is its wide availability and ability to cover areas where Wi-Fi access points or Bluetooth devices are not available. The geographical area covered by cellular networks can provide valuable data. With cellular-based localization, occupants can be tracked over the area of interest (e.g., at the building or university campus). This can allow analyzing occupants' movement and understand individual as well as emergent behavioral patterns. For example, the movement of students on campus can be analyzed to find the locations and times to reduce traffic jams. A large number of students on campus might all have to go through a certain corridor to get to a theater. Providing another way or entrance to the theater might improve the situation. As another example, analyzing such movement patterns of occupants might reveal that many occupants must move for a long distance during the day to get to a certain service (e.g., coffeeshop). These findings might help resolving such issue by introducing minor changes to building design.

Several localization systems based on measurements from LTE signals have been proposed. In [16], a localization system that employs Channel State Information (CSI) extracted from LTE signals was proposed. The system uses CSI measurements for signal fingerprinting localization. Experiments in indoor and outdoor environments show that localization based on CSI from LTE signals can be used for both indoor and outdoor localization. The authors in [23] proposed a fingerprinting approach for localization of UE in LTE-A networks mapping multiple radio channel parameters formulated as a fingerprint vector and a geographical location. They employ a feature-extraction algorithm to identify unique channel parameters and use a neural network to build a fingerprinting database of channel parameters and UE locations. Results show that by using only one LTE eNB, the proposed technique provides a median error distance of 6 and 75 meters in indoor and outdoor environments, respectively.

The authors in [15] also considered employing the CSI from LTE signals for fingerprinting-based indoor localization. The authors propose a technique where the fingerprint contains a vector that serves as the shape of the channel frequency response instead of the CSI. The approach uses eNB signaling messages and does not need designated communication between the eNB and the UEs. The approach reduces computation complexity and memory requirements.

In [11], the authors evaluate the accuracy of localization based on radio fingerprinting of LTE signals on 800 MHz, 1800 MHz and 2600 MHz frequency bands. Field measurements are conducted to collect training data that consist of UE locations and the corresponding received signal strength radio measurements from several base stations. Collected data are used to provide a fingerprint of the radio conditions at a specific location. The performance of two systems composed of LTE and LTE+WLAN grid-based RF fingerprint measurements utilizing partial fingerprint matching were studied and compared. Obtained results show that partial fingerprints that consist of LTE and WLAN radio measurements improves localization accuracy by at least a factor of 3.5x while keeping the percentage of discarded samples low.

The work in [6] used Cell-Specific Reference signal measurements from LTE signals for indoor localization to complement outdoors localization systems such as Global Navigation Satellite System. Two algorithms were used for localization. The first one is a Time-Of-Arrival approach called the Threshold-to-Noise Ratio algorithm. The second one is an estimator that is more complex but also robust against multipath fading; it provides more accurate, robust and smooth results indoors, at the cost of increased complexity.

All the research above considers LTE-A networks with macro cells, where a macro eNBs with high power provides the coverage for a wide geographical area and high number of users. In our work, we will studying the performance of localization over LTE-A UDNs. The availability of high number of elements such as femtocells and picocells are supposed to increase the accuracy of localization over mobile networks for indoors environments.

3 LTE-A UDN FOR OCCUPANCY ESTIMATION

3.1 LTE-A UDN Scenarios
New network architectures such as UDNs and Ultra-Dense Heterogeneous Networks (UDHetNets) are enabling technologies to meet increasing demands and achieve the required performance of 5G cellular networks [10]. With UDNs, the density of the operator-deployed and user-deployed elements is reduced, improving coverage, frequency reuse, and achieving higher data rates. In UDHetNets, several types of wireless access nodes are employed, and hence, macrocells are overlaid with low-power nodes such as Remote Radio Head, Pico eNB (PeNB) and Home eNB. These

smaller cells can be used to offload traffic, which improves the network coverage at the cell edge and increase data rates.

The LTE-A Pro standard [1] proposes different scenarios for implementation of UDNs and UDHetNets. These include scenarios for UDNs where similar elements are employed, such as PeNBs, as well as heterogeneous scenarios where distinct types of cells coexist such as eNBs and PeNBs.
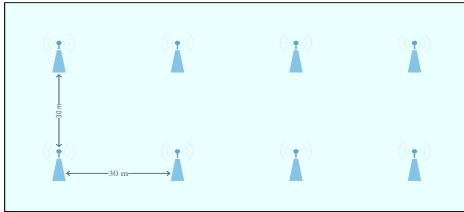

Figure 2. LTE-A UDN scenario A.

Following, we list the possible scenarios considered in the LTE-A Pro standard [1]:

- Scenario A-Indoor small cell deployment: this scenario consists of a single layer of small cells in an indoor environment. This scenario is shown in Figure 2.

- Scenario B-macro cell deployment: this scenario consists of a single layer of macro cells.

- Scenarios C and D-Heterogeneous network of urban macro and outdoor small cell deployment: these contain macro cells coexisting with small cells. The two differ in the method of channel allocation for the two layers.

Table 1. Transmission parameters for scenario A.

Parameters	Scenario A
Type	Indoor Hotspot (Figure 2)
Layout	Single layer Indoor TP: Number of TPs: N=8, N=12 (optional) per 120m x 50m
ISD (inter-site distance)	20m, 30m
Carrier frequency	3.5GHz
Coordination cluster size for ideal backhaul	All sites
System Bandwidth	10MHz (50RBs)
Channel model	Channel model available in document TR 36.814

We are interested in scenario A because our work is focused on indoor localization and building occupancy count estimation. The transmission parameters for such scenario is provide by the LTE-A standard and presented in Table 1. These parameters will be adopted in our study.

3.2 Methodology and localization approach

We propose using the Channel Quality Indicator (CQI) or Received Signal Strength Indicator (RSSI) values sent from the UEs to the eNBs for localization. A fingerprinting-based method will be used where a database of pairs of locations inside the building and corresponding CQI or RSSI values will be first built. During the localization phase, we will estimate location from the built database.

We propose simulating indoor LTE-A UDN scenarios. First, we will run simulation scenarios involving simple prototype floor plans. Afterwards, real building designs will be obtained from existing BIMs to generate simulation scenarios automatically. Various BIMs developed to produce a digital campus at Carleton University contain various attributes (spatial, areas, volumes, and uses of the rooms), which can be used to create real simulation scenarios.

As discussed in the next section, we will include occupants' behavior and their personas. This can result in developing models for the movement of occupants. Such realistic movement models can be used to create more precise simulation scenarios to study localization and count estimation. For instance, a study of occupants may reveal that many occupants tend to sit close to windows during the summer. In such case, occupants will stay close to the edge of the building, which means that an accurate localization algorithm would be needed to produce an accurate count estimation.

From these simulations, we will extract various data sets for the UEs locations and corresponding RSSI values as per the approach presented in the previous section. From the collected data, we will build a fingerprinting database, and evaluate the performance of indoor fingerprinting localization in LTE-A UDNs. The localization accuracy as well as building occupancy estimation accuracy will be considered as the performance metrics. Additionally, the performance of such scenarios will be studied and compared to those achieved in the case of a macro cell architecture.

4   LOCALIZATION AND BUILDING DESIGN

4.1 Occupant Behavior and Building Operation

Occupants are one of the leading causes for the difference in predicted and actual energy usage in buildings [5, 22]. The knowledge of how occupants behave and interact within a context is not available to them. Such behavior is usually more complex than the assumptions made by the modelers. The modeling and simulation (M&S) community does not have access to the 'lived experience' of the people, therefore they need to assume the possible set of occupants' behavior in the buildings. Furthermore, factors like socioeconomic conditions, available technology, environmental conditions, and temporal adaptations influence the occupant's behavior. The context decides the interaction possibilities. Privacy limits the possibility of understanding user behavior in buildings. Hence, measured data in buildings plays a crucial role in understanding occupancy behavior though it lacks qualitative interpretation. Various technologies are integrated in new buildings (e.g., sensors), and this makes various data

(e.g., CO2 levels) available to designers. IoT allows us to have access to certain information without invading occupants' privacy (e.g., their current activities). Though there is a limitation in assumptions for occupants' behavior from measured data (like their perception or personal comfort), it could be used to improve design decisions. Our proposal focuses on creating and generating personas from measured data and build a model that optimizes the design based on the criteria [12]. The research goal is to use personas at the design stage and during building operation and automation. Using the personas at building operations will help the buildings to adapt sustainably. Likewise, the automation system could suggest efficient interactions based on the occupants' location. Additionally, it enables the building systems (like blinds, thermostat) to make dynamic changes to improve occupancy comfort and building performance.

4.2 The use of personas

During the design stage, building simulation can be used to analyze lux levels or temperature in a room, and the behavior of occupants for those conditions. However, during building operation more precise parameters can be defined through LTE-A. A qualitative questionnaire or inputs can collect the occupants' preferences. The measurable and qualitative information can enable automation to fine-tune comfort level during building operation. Persona gives flexibility: the same personas used at the initial stages of design could be used with finer granularity in real-time to improve building performance. The attributes are individual characteristics (age, clothing, activity, and role), comfort preferences (thermal comfort, visual comfort, and views), interactive behavior (blind, door, windows, equipment states), eco-behavior (active or passive decisions made to save energy), and social behavior (socio-economic conditions and group dynamics).

During operation, the occupants' location enables us to make the spatial relationship with other measurable data like available interactions, room temperature, and lux levels. The automation can use the information to make effective dynamic changes. Further, occupancy count and tracking enable us to understand group dynamics like preferred location, most likely used space as a group and as an individual. All these pieces of information may help in refining building operation. The following section explains how the personas help in understanding the user behavior at the design stage.

4.3 Personas for Occupancy

Figure 3 illustrates the integration of personas to the model of a typical generative design (we limit the discussion to the parameters of the personas and not the whole generative design components). We propose defining the geometry using Dynamo [7] and Autodesk Refinery for Generative Design. The model runs energy simulation with the initial geometry, and then evaluates the output for occupant behavior.

The personas are generated randomly using two main parameters: the number of occupants and the building type (see Figure 4). It creates different persona types for simulation

and evaluation. Once the simulation is completed, the system runs fitness criteria for sustainable behavior goals.

Fitness evaluation is performed using Discrete Event Simulation. Once the performance data is updated, the model runs a fitness check on how the different personas will behave on those conditions. If their interactions with building elements and systems (like windows or thermostats) lead to more energy usage than the defined goal, the system modifies the geometry and reruns the process to produce an optimal design. Personas may be a solution to minimize the discrepancy between predictive and actual scenarios of use for energy use and comfort. Understanding the behavior of the occupants will affect the design decisions. Personas could be used in automated building to control thermostat, blinds, or lighting.
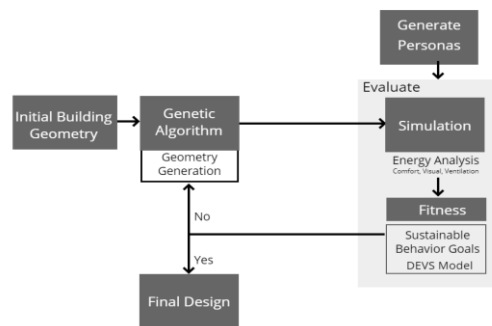

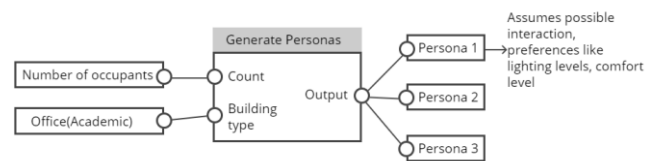**Figure 3.** Integration of personas to a typical generative design model using Revit/Dynamo/Refinery/DEVS.

Figure 4. Conceptual model for generating personas

4.4 Personas and localization

The novelty of the idea is to use the personas with occupants' location, count, and tracking for building operation (see Section 3). In this section, we discuss the use of persona with the localization concept at a small geographical and quantitative scale (i.e., room or building, and individuals). Figure 5 shows the parameters considered to develop personas for automation purposes.
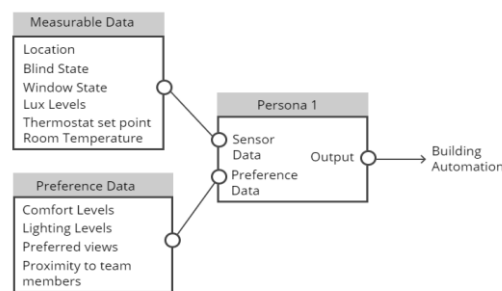

Figure 5. Personas for building automation

There are two categories for the collected parameters. First, the measured data (like blind state, room temperature) and second occupant's preference data (like preferred light settings). Based on occupant location, the automation system collects other corresponding data like room temperature, lux levels, and blind state. It compares the information with the preferences of the occupants to automate the blinds or thermostat or any automated system. The automation can consider the decisions at the individual level or collective level. It depends on the influence on its surrounding. For example, at a collective level, a person sitting close to windows may have more daylight compared to the person sitting on the far end. Hence, a person sitting at the far end may need more artificial light than the person sitting closest to the windows. Likewise, a person sitting closest to the window may feel colder during winter compared to the person sitting at the far end. Hence, considering the location and corresponding measurable data will help in refining occupant comfort.

The comfort of an occupant needs to be considered at an individual level and a group level.

5 BIM AND BUILDING OCCUPANCY SIMULATION

We use BIM models as a host where to produce simulations, to apply building retrofitting concepts (using data extracted from the occupancy count estimation as a parameter in the model to inform future designs); and as a tool for digitally assisted storytelling (which refers to graphical dissemination of data and visual communication of the simulation results).

Figure 6. Render of InfraWorks of the Digital Campus model.

We built a BIM model of Carleton University campus, consisting of a federated digital assembly of more than 50 buildings, roads, tunnels, landscape, etc. (Figure 6). The model includes many different layers of data received from various parties, some of which are anticipated to be beneficial for the three applications mentioned above [19].

5.1 Using BIM models for prediction

To produce simulations, an architectural setting is needed; however, in some cases, it is not necessary to have a model that reflects an actual physical place. If one has an accurate virtual representation of a real building that contains all the attributes required for occupancy count estimation and tracking, one can replace the need for a physical one. For this reason, a digital model was a more feasible alternative. The digital campus has all the essential elements required to run the simulations and has the potential of holding more parameters if it is required in the future. The campus BIM model contains rooms with parameters for spatial attributes, such as their area, volume, uses, etc. It also contains walls, ceilings,

floors and all the architectural elements needed to understand the space, as well as location and all the attributes of the eNBs. With these components, it is possible to simulate scenarios as accurate and as close to reality as possible. The virtual representation of physical spaces and architectural elements creates a good environment for hosting both the simulation and the personas.

The model can be used at different scales. At building scale, it is possible to understand the characteristics of the indoor environment. For example, one can simulate the impact on new buildings over existing buildings to predict the consequences this new relationship is going to have over the occupants of the existing space. Additionally, at campus scale, it is possible to visualize occupancy in relation to groups of buildings, circulation, services, shared areas, landscape, etc. This could be beneficial to larger scale planning strategies by producing a better understanding of campus use.

5.2 Using collected data to inform design

The second BIM application for this study refers to the use of data collected from the occupancy count estimation, as well as the behavior of people and their location in space to inform future designs. For example, every few years, the university produces a campus master plans to set the parameters, policies and directions for the physical development of its campus. This master plan aim to set the basis for future developments, guiding them to be in harmony with the university's principles. Having a better understanding of the occupancy, behavior and location of people on campus could help designers to generate better and more accurate master plans. For instance, the campus that we used to run the simulation in this study has tunnels to connect different buildings during the winter months. Should the collected occupancy data reveal one tunnel having more intensive use than another, the designers can respond to this information, potentially widening tunnels, reducing others, or even building new ones to reduce congestion in future master plans. Another application could be to define the dimensions of new spaces. For example, if a space demonstrates a greater occupancy than expected, a similar typology in a new building could be designed taking into consideration the results of the simulation. Additionally, it could help to make decisions regarding materiality: designers could pick stronger tiles for a floor that proves to be used more intensely than another do, or to reduce the dimensions of beams supporting a space that is not as frequently occupied. In both cases, the process may result in using materials and elements that are better fitted for their use, thus making them more durable and cost-effective in the construction of new buildings. The collected occupancy data, treated as a parameter in the BIM model, could become additional information for HVAC system designers (see Figure 7). Here, optimal systems and equipment for the ventilation of a new building can be developed, taking into consideration the potential use of its spaces. Finally, the data could be used through the BIM model for operation and maintenance (O&M). It is becoming increasingly common for Facility

Managers (FMs) to use BIM models to operate buildings. Integrating real time collection of occupancy data such as location and behavior into a BIM model, could help FMs better understand the use of different spaces, enabling a higher efficiency of O&M management [17].
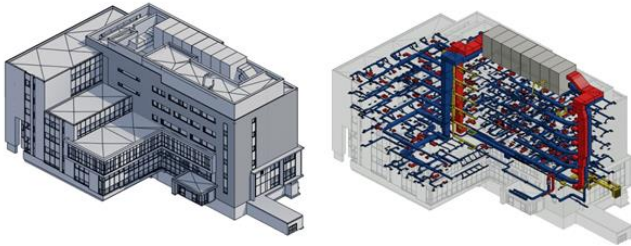

Figure 7. Mechanical, Electrical and Plumbing model (MEP) of the Health Science Building (digital campus model).

5.3 Digital Assisted Storytelling

Digital Assisted Storytelling refers to the use of digital techniques to create narratives that disseminate information and ideas. BIM platforms, such as Revit, provide powerful visualization features that allow users, through two or tri-dimensional geometry, to display data in intuitive and interactive ways. This includes diagrams, adaptive geometry and interactive parameters, among others. This way of disseminating the data obtained through simulation could help non-specialized users or those without any AECO background to understand and interact with the results of the study in real-time.
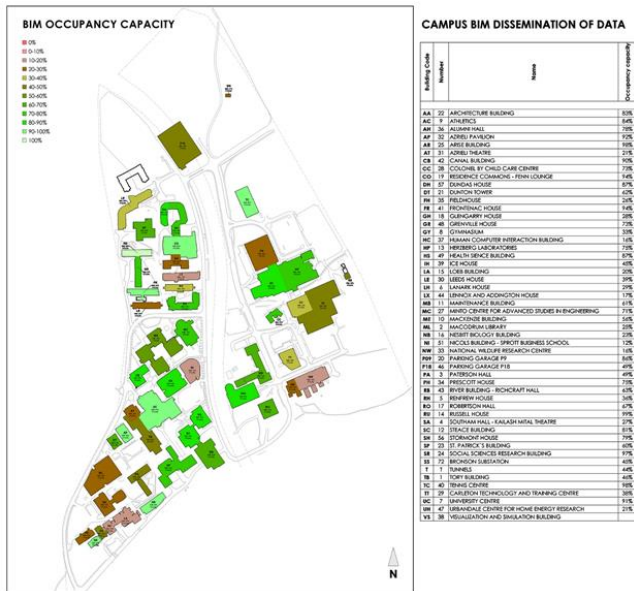
Figure 8. Simulated occupancy diagram.

BIM helps to graphically disseminate the results of a simulation in multiple ways; since the BIM model contains all the architectural and geometrical information of each building in the campus, the data could be displayed as floor plans, sections, elevations, axonometric views, renderings, etc. In Revit, through Visibility Graphic Display, it is possible to pre-set the expected visualization to update according to the changes in the data (Figure 8). One can also create View Templates that can adjust the properties of multiple views at the same time, making the process more efficient while also providing visual consistency. Furthermore, through Revit, one can use Dynamo, a visual programming tool, to manipulate large amounts of data and complex geometry with great precision. With Dynamo, it is possible to manipulate the data and produce real time diagrams and graphs. This way of visualizing data helps us to better communicate the results of our simulation, making the information available to more people, and allowing better feedback and collaboration.

Using a BIM model allows us to generate a bridge between complex technical language—which was required for the simulation—and the common designer or user. This communication enriches the process of Storytelling. As the architect and scholar, Marco Frascari wrote: "Buildings are not experienced as data 'fed to passive spectators' but, instead, are experienced culturally through the stories found embodied in buildings and retold by architects… the real architectural craftsmanship is the crafting of a good story" [13]. Data, on its own, has no real power to generate an impact on people, or to have a deep implication towards their experience; the real power is in the story that one is able to tell or to graphically display. We are facilitating the understanding of complex processes in a simplistic and didactic way.

When working on multidisciplinary projects, it is important to maximize the diverse capabilities of each team member. Indeed, for a group composed by professionals and academics with diverse backgrounds, BIM becomes an ideal tool to congregate technical analysis and qualitative results with a rich and intuitive visualization. "BIM has the untapped potential to unhinge the link between instrumentality and architectural representation. For example, the capacity to simultaneously incorporate large and diverse sources and types of information, represent it in multiple formats, and react to input in real time present an opportunity to develop modes of architectural representation that are in flux and responsive to the people, history, materials, and environment that contribute to the making of architecture." [9]. BIM promotes collaboration and communication among professionals of the AECO industry and it has the capability to transmit the produced information massively and in elemental ways.

6 CONCLUSION

We propose the use of LTE-A UDNs to provide estimates for occupancy count and user location and tracking in buildings. We specifically proposed to use CQI or RSSI values from the UEs to the eNBs. To use this data, we are building a database of pairs of locations inside the building and corresponding CQI or RSSI values. We will use different algorithms for the localization phase. We will also evaluate the performance of those algorithms and provide a comparison. We will evaluate them on different simulation scenarios created from the information stored in BIM models.

The advantage of using LTE-A UDNs over other approaches based on data collected over a network of sensors (e.g. CO_2 levels, camera data, humidity, etc.) is that it eliminates the need to install and set up sensors in the building. We reuse the LTE-A infrastructure that is already deployed for cellular communications. Because we do not need to deploy specific sensors for occupancy and user location, we expect that this approach will reduce the cost of the building equipment and its maintenance. Additionally, the LTE-A UDNs are maintained by the cellular service provider. We also expect that this method will provide better results than other methods that also use already deployed infrastructure for communications (e.g. Wi-Fi AP). As future work, we will validate this assumption comparing our proposed method with others based on Wi-Fi APs and Bluetooth.

Having accurate occupancy count estimation, and user location and tracking can have an impact in optimizing building Operation and Management. Employing occupancy detection and localization (using LTE-A UDNs) is being investigated. We propose using occupancy data to optimize the operation of the building through actuators and controllers (e.g., controlling HVAC systems). We also propose using both sensor data as well as occupants' locations and count estimation to study occupants' behavior and generate personas (e.g., movement patterns, preferred locations within offices, etc.) that are also used in the process of analysis and design of the building and its controllers through simulation. Furthermore, occupants tracking, and count estimation will be included as parameters in BIM models and visualized. Such data will be used to design future buildings with the same purpose as the current building, or to generate designs during expansion or retrofitting of the same building.

REFERENCES

[1] 3GPP 2017. *Technical Specification Group Radio Access Network; Study on further enhancements to Coordinated Multi-Point (CoMP) Operation for LTE.*

[2] Akyildiz, I.F., Gutierrez-Estevez, D.M. and Reyes, E.C. 2010. The evolution to 4G cellular systems: LTE-Advanced. *Physical Communication*. 3, 4 (Dec. 2010), 217–244.

[3] Cao, X., Dai, X. and Liu, J. 2016. Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy and Buildings*. 128, (Sep. 2016), 198–213.

[4] Chen, Z., Jiang, C. and Xie, L. 2018. Building occupancy estimation and detection: A review. *Energy and Buildings*. 169, (Jun. 2018), 260–270.

[5] Clevenger, C.M. and Haymaker, J. 2006. The impact of the building occupant on energy modeling simulations. *Joint International Conference on Computing and Decision Making in Civil and Building Engineering* (Montreal, 2006), 1–10.

[6] Driusso, M., Marshall, C., Sabathy, M., Knutti, F., Mathis, H. and Babich, F. 2016. Indoor positioning using LTE signals. *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (Oct. 2016), 1–8.

[7] Dynamo: *https://dynamobim.org/*.

[8] Energy Efficiency: Buildings The global exchange for energy efficiency policies, data and analysis: *https://www.iea.org/topics/energyefficiency/buildings/*. Accessed: 2019-07-22.

[9] Gallant, M. and Fai, S. 2018. BIM: The virtual capriccio: New paradigm of the modern tool in the heritage world. *Proceedings of the 2017 23rd International Conference on Virtual Systems and Multimedia, VSMM 2017* (Apr. 2018), 1–8.

[10] Gotsis, A., Stefanatos, S. and Alexiou, A. 2016. UltraDense Networks: The New Wireless Frontier for Enabling 5G Access. *IEEE Vehicular Technology Magazine*. 11, 2 (Jun. 2016), 71–78.

[11] Hiltunen, T., Turkka, J., Mondal, R. and Ristaniemi, T. 2015. Performance evaluation of LTE radio fingerprint positioning with timing advancing. *2015 10th International Conference on Information, Communications and Signal Processing (ICICS)* (Dec. 2015), 1–5.

[12] Lene Nielsen 2019. *Personas - User Focused Design*. Springer.

[13] Marco Frascari 2012. An architectural good-life can be built, explained and taught only through storytelling. *Reading Architecture and Culture Researching Buildings, Spaces and Documents*. Adam Sharr, ed. Routledge. 272.

[14] Natural Resources Canada 2018. *ENERGY EFFICIENCY IN CANADA REPORT TO PARLIAMENT UNDER THE ENERGY EFFICIENCY ACT 2016–2017.*

[15] Pecoraro, G., Di Domenico, S., Cianca, E. and De Sanctis, M. 2018. CSI-based fingerprinting for indoor localization using LTE Signals. *EURASIP Journal on Advances in Signal Processing*. 2018, 1 (Dec. 2018), 49.

[16] Pecoraro, G., Di Domenico, S., Cianca, E. and De Sanctis, M. 2017. LTE signal fingerprinting localization based on CSI. *2017 IEEE 13th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)* (Oct. 2017), 1–8.

[17] Rafael Sacks, Chuck Eastman, Ghang Lee, P.T. 2018. *BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors, and Facility Managers*. Wiley.

[18] Shen, W., Newsham, G. and Gunay, B. 2017. Leveraging existing occupancy-related data for optimal control of commercial office buildings: A review. *Advanced Engineering Informatics*. 33, (Aug. 2017), 230–242.

[19] Shi, Z., Abdelalim, A., O'brien, W., Attar, R., Akiki, P., Graham, K., Van Waarden, B., Fai, S., Tessier, A., Khan, A., Canada, C. and Author, C. *Digital Campus Innovation Project: Integration of Building Information Modelling with Building Performance Simulation and Building Diagnostics.*

[20] Tianzhen Hong, O., Langevin, J. and Hong, T. 2017. *The Human Dimensions of Energy Use in Buildings: A Review.*

[21] Wainer, G. 2013. Applying modelling and simulation for development embedded systems. (Sep. 2013), 1–2.

[22] De Wilde, P. 2014. The gap between predicted and measured energy performance of buildings: A framework for investigation. *Automation in Construction*. 41, (2014), 40–49.

[23] Ye, X., Yin, X., Cai, X., Perez Yuste, A. and Xu, H. 2017. Neural-Network-Assisted UE Localization Using Radio-Channel Fingerprints in LTE Networks. *IEEE Access*. 5, (2017), 12071–12087.

[24] Zafari, F., Gkelias, A. and Leung, K. 2017. A Survey of Indoor Localization Systems and Technologies. (Sep. 2017).

Simulation-Powered Smart Buildings Management enabled by Visible Light Communication

Yehuda E. Kalay [1,2], Haripriya Sathyanarayanan [1], Davide Schaumann [3], Albert Wang [4], Gang Chen [4] and Ramdas G. Pai [4]

[1] University of California Berkeley, US
{kalay, haripriya_snarayanan}
@berkeley.edu

[2]Technion, Israel Institute of Technology
Haifa, Israel
kalay@technion.ac.il

[3]Jacobs Technion-Cornell Institute
New York, US
davide.schaumann
@cornell.edu

[4] University of California
Riverside, US
{aw, gachen}@ece.ucr.edu
ramdas.pai@medsch.ucr.edu

ABSTRACT

Throughout history, buildings have been considered passive containers in which occupants' activities take place. New sensing technologies enable buildings to detect people presence and behavior. At present, this information is mostly used to trigger reactive responses, such as heating and cooling operations. We argue that truly smart environments can leverage sensed information to proactively engage with the occupants and inform decision making processes with respect to which activities to execute, by whom and where. Such ability will transform buildings from passive to active partners in the daily lives of their inhabitants. It stems from the omniscience of sensor-equipped buildings that will "know" all that is happening everywhere within (and around) them at any given moment and can predict, through simulation, the expected consequences of alternative operational decisions. Such ability is mostly relevant for hospitals and other complex buildings, where actions taken in one part of the building may affect activities in other parts of the building. We are developing a simulation-powered building management system that resolves space, actor and activity-based conflicts while harnessing data collected via visible light communication. We demonstrate this approach in a case study in the catheterization lab of a major hospital.

Author Keywords

Smart Environments; Human Behavior Simulation; Space Utilization; Hospital Environments; Visible Light Communication.

ACM Classification Keywords

I.6.3 SIMULATION AND MODELING: Applications; I.6.5 SIMULATION AND MODELING: Model Development; J.6 COMPUTER-AIDED ENGINEERING.

1   INTRODUCTION

Sensing technologies that enable buildings to detect people's presence have been in use for the past few decades, mostly to trigger reactive responses to people's presence (e.g., heating/ventilating, lighting, security, etc.). We argue that truly smart environments can leverage sensed information about the locations and activities of their inhabitants to proactively engage with the occupants and inform their decision-making processes with respect to which activities to execute, by whom and where (in addition to autonomously activating building resources).

To help assess the potential impact of "smart" buildings on their occupants, we are developing a simulation-powered building management system that can sense the location and activities of human and building assets, extrapolate patterns of utilization, simulate what-if scenarios and suggest future user activities and resource allocation to maximize specific Key Performance Indicators (KPIs). Different from existing approaches, our system is able to evaluate the implications of potential conflict resolution strategies using a multi-agent simulation system that accounts for individual and collaborative activities.

While the approach being developed is agnostic of the sensing and communication technology used, we have chosen Visible Light Communication (VLC) technology, which is embedded in a building's LED lighting system [10].

VLC has been chosen because unlike radio frequency it is highly localized, and it does not interfere with the building's other sensitive instruments, which is critical in the case of hospitals (our chosen case study). In addition, because it is embedded in the building's LED lighting system, it requires little additional infrastructure compared to other technologies.

Information derived from the VLC system is combined with models of actors' activity schedules, profiles, and space affordances to understand what happens in each space at any

given time. This data on the building's current state of the occupancy and utilization is used to simulate alternative possible future actions for each actor and to resolve possible conflicts that may occur. The simulation and decision-making process are driven by a previously developed narrative-based modeling system to simulate human behavior in buildings [12,13]. It produces alternative future states, revealing the consequences of enacting different resource allocation strategies. A priority function is used to evaluate and compare the alternative futures and choose the one that maximizes a utility function. Once the decision is made, the system uses VLC to communicate the information to the relevant actors who can enact them. We demonstrate our approach by applying it to the Catheterization Lab in a hospital.

It is our contention that smart environments of this kind hold promise to enhance the decision-making capabilities of buildings and their inhabitants, thus enable building management strategies that support human needs and efficiency requirements, especially in mission-critical facilities, such as hospitals.

2   ADAPTIVE BUILDINGS

We have identified three levels of building automation that help us understand and assess the potential impact of adaptive buildings on their occupants, namely buildings that can change their performance dynamically in response to the changing environmental conditions and the needs of their inhabitants [5].

Feedback regulated adaptability is based on the concept of *feedback loop*, where the output of a machine is linked to its input and compared against some desired performance measures. Departure from the desired condition triggers adjustments in the performance of the machine, hence its output. The ubiquitous thermostat demonstrates this principle: as the HVAC system heats (or cools) the air inside a building, the thermostat monitors the air's temperature. When that temperature reaches the thermostat's set point (the desired temperature), it sends an electrical signal to the HVAC plant, shutting it off. When the air cools below the set point (or, conversely, heats up beyond it), the thermostat sends a signal which turns the HVAC plant on, and so on.

Enabling the building to sense and respond to changing needs is a relatively simple reactive kind of automation, which has been implemented in areas of control, regulation and supervision of electrical, mechanical and climatic control equipment.

Adding a *functional model* to networked building systems and appliances allows for a proactive adaptability approach to automation, which we call **model-based adaptability**. It helps to regulate the environment in *expectation* of events, rather than in response to them. A functional model of a building is one where the occupants' behavior patterns are programmed in advance, based on learning their typical

preferences, so the building can anticipate and position itself to support recurring events, not only to respond to them.

Model-based adaptability has been demonstrated by the University of Colorado's ACHE (Adaptive Control of Home Environments) project [9]. Using the model, the house could anticipate the inhabitants' preferences and adjust the operation of devices, accordingly, freeing them from the chores of manually controlling of their environment.

Total-environmental adaptability will be reached when the building not only responds—reactively or proactively—to its inhabitants' behavior, but will actively engage, even manage them. Such active management requires much more information than the locations of the inhabitants and prevailing environmental conditions: it must include information about spatial conditions, activities, and the inhabitants themselves.

Such information comprises of three components: (1) *Space* information, which includes the configuration of the building (rooms and the connections between them), the intended purpose of each room (e.g., a hospital patient room, a nurse station, etc.), the environmental conditions prevailing in each space (light, temperature, noise, etc.), and current location of each inhabitant. (2) *Activities* information, which includes each inhabitant current, past and future activities. It also includes information about customary scheduled activity sequences, and what to do in case of unexpected activities (e.g., 'Code Blue' in a hospital). (3) *Inhabitants'* (which we call 'actors') information, which includes the identity of each actor, his/her profile (role in the organization—doctor, nurse, patient, visitor, etc.), abilities, degree of fatigue, and more.

Once the building management system has access to all this information, it can form an image of the current state of the whole building and its inhabitants. Using simulation, it can predict alternative future states, which can be evaluated according to some Key Performance Indicators (KPI), so that the most suitable future state can be chosen.

3   OVERVIEW OF THE PROPOSED APPROACH: THE POWER OF SEEING THE WHOLE

We call this ability "the power of seeing the whole." It is what air traffic controllers use to manage planes in the air, and GPS-based systems like Waze use to help drivers choose the fastest route to their destination.

The ability to see the whole provides an overview of some situation, not visible from the individual actor's point of view. Much like an air traffic controller can direct airplanes without risking midair collisions, a building management system could direct assets (people, equipment) to where they are needed at any given time, alert security personnel in case of disturbances, and more. Furthermore, as evident from Figure 1, this ability extends from the present to the past: it is possible to trace previous locations of individuals and equipment at prior points in time.

532

It is our contention that this ability can also be extended into the future by way of simulation, which will allow the building management system to predict the future locations and activities of the inhabitants. It could, therefore, consider alternative "futures" and help choose the one most desired (according to some predefined criteria).
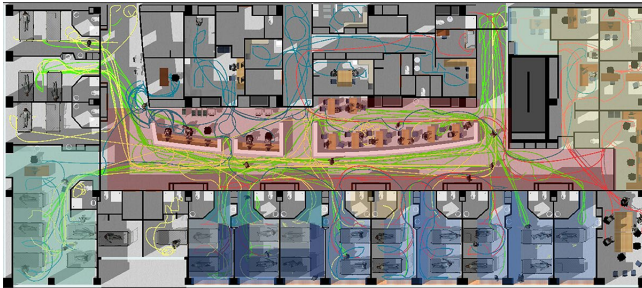

Figure 1. Locating and tracking people in a hospital ward.

4   METHOD

We propose a simulation-powered Building Management System aimed at Total Environmental Adaptability that leverages the power of "seeing the whole." It will sense the presence and location of humans and building assets, extrapolate patterns of behavior and utilization, simulate what-if scenarios and suggest modifications to user activities and building operations to maximize specific KPIs. The system is composed of three main components: (1) sensing, (2) decision-making, and (3) acting.

4.1 Sensing

A building can "sense" various aspect of the environment, including the presence of people, through a variety of embedded sensors. If the data collected by each sensor is to be made available to the building as a whole, these sensors must be networked. While wireless connectivity is preferred, in hospitals—which are the domain and case study for our research—sensors based on radio-frequency (RF) wireless technologies have many major limitations, such as interference with medical devices that may potentially put patient safety at risk [1], and compliance with extremely high security and regulatory standards to ensure patient privacy, which cannot be ensured because RF signals are publicly open, and their security is only guarded by certain encryption methods.

An alternative to RF is communication based on visible light. Traditional incandescent and fluorescent lamps are being replaced by energy-efficient solid-state LEDs [7]. Other than being more energy efficient, LEDs can be switched ON/OFF at a speed of tens of MHz without flickering visible to the eye, enabling Visible Light Communication (VLC) at high data rate by modulating LED light [3]. Wireless VLC has many advantages over traditional RF technology: (a) the optical spectrum is unlicensed, unrestricted and orders of magnitude wider (300THz) than the crowded RF spectrum, making wireless streaming of big data possible for large

number of users; (b) it allows more emission power for higher data rates and better quality of service (QoS) without risking human health; (c) unable to penetrate walls, it ensures high security and privacy; (d) being interference-free, it can co-exist with RF technologies; (e) VLC devices are cheaper than RF components.

Ubiquitous VLC wireless systems will consist of modulated LEDs (lamps) for broadcasting and user terminals (smartphones with embedded photodetector as transceivers) to realize full-duplex optical wireless streaming anywhere, anytime for anyone. In a sense, VLC wireless is "free" because it is built on existing LED lighting infrastructure, providing VLC wireless streaming at beyond-Gbps speeds. LED also allows visible light real-time communication and positioning, making it possible to share data, locate personnel and equipment instantly and securely.

4.2 Decision making

The system combines the sensed data with other data, to help make decisions about future actions. Since these decisions involve human activities, which are dynamic and depend on many factors such as spatial, occupational, and personal conditions, the method chosen to help depict future situations is *simulation*.

Simulation approaches have been used to analyze the dynamic relationship between human activities and the surrounding environments in both existing and not-yet-built environments. In particular, recent work on narrative-based modeling [12,13] demonstrated a viable method to simulating day-to-day occupancy scenarios in complex facilities, like hospitals. The approach is centered on *narratives*, which are rule-based scripts that coordinate the collaborative behaviors of heterogeneous actors (e.g., doctors, nurses, patients) who perform a structured sequence of activities (e.g., checking a patient) that unfold in semantically rich spaces (patient rooms, clinics, etc.). Different from other simulation approaches that mostly focus on linear, straightforward pedestrian movement or evacuation scenarios, the narrative-based model uses a combination of a top-down coordination mechanism to enforces the performing of structured sets of tasks, while allowing for bottom-up adaptations to dynamic social and spatial conditions, such as the emergence of unplanned narratives (staff-visitors interactions) that can potentially delay planned narratives (checking a patient).

Our narrative-based decision-making system consists of three components: (a) a library of spaces, actors, activities and narratives that represent the spaces that people inhabit, the actors that populate the spaces, the activities they perform, and the narratives they are involved in; (b) a simulation engine that calculates the behavior of the entities over time; and (c) an evaluation module that calculates and visualizes a list of KPI so that the simulated outcomes of different future narratives can be compared to predefined KPI.

4.2.1 Library of spaces, actors, activities and narratives

- *Space* entities comprise a model of the building, typically generated using CAD or BIM tools, including both physical (walls, floors, doors, furniture, etc.) and non-physical components (rooms, corridors, and open areas). Both physical and non-physical building components also store semantic information that indicate how they can be used. A 'clinic room' zone, for instance, indicates that the space can be used for clinical activities, such as treating a patient. Such zones can, for instance, record the presence and activities of the occupants within their boundaries.

- *Actor* entities include a profile (e.g., patient, nurse, companion, doctor) that determines the type of narratives the actor can be associated with, stores static information, such as the names of patients that a nurse is responsible for treating, and updates dynamic information about the current activity the actor is performing as well as other actor properties, such as tiredness.

- *Activities* represent the possible interactions that actors have with other actors or with spaces. In this work, we are concerned with abstracted activity descriptions, their spatial location, the identities of the participating actors, and their duration. In this way, we can limit the number of activities modeled and focus on their spatial/social implications in real-world clinical situations. We also model activities in a modular fashion, so they can be reused multiple times within a narrative or across narratives.

- *Narratives* are the heart of the simulation. They use the aforementioned components (spaces, actors, and activities) and combine them into scripts that direct actors' behavior by associating them with specific activities performed at a given time and space(s), while accounting for possible context-dependent adaptations to unforeseen situations.

4.2.2 Simulating

A narrative manager coordinates the unfolding narratives over time. In addition to the execution of planned narratives, the narrative manager is also responsible for triggering unplanned narratives when the necessary preconditions are satisfied (e.g., an impromptu conversation that takes place when a staff member and a visitor meet in a corridor). The simulation is powered by Unity 3D, a popular game engine.

4.2.3 Evaluating

The simulations result in measurable performance indicators, which can be compared to predefined threshold measures or relatively to one another. They may include hard and soft criteria. Hard criteria are quantitative, measurable performances, such as patients and staff walking paths and distances, patients' length of stay, overall throughput, congestion, staff or space utilization. Soft criteria are typically qualitative, based on subjective perceptions, such as social, psychological, and organizational policies.

In many cases, the same performance results may be valued differently by different stakeholders. To create a building-wide management system, it is therefore necessary to create a shared world view that incorporates the relative merits of each action from different points of view and reconciles the differences among them in light of shared, higher-level objectives [2]. A tradeoff mechanism balances competing needs. It may choose to optimize one performance characteristic over others or strike a balance in the degree to which any performance criterion is achieved, assuring that overall performance is maximized [6].

4.2.4 Implementation details

In addition to using Unity 3D as the simulation engine, which was chosen because it features advanced physics and AI libraries to model collision avoidance and pathfinding, we use Microsoft C# to represent the properties of actors, activities, narratives and for the narrative manager. Spaces have been modeled using McNeel Rhinoceros 3D software, and then imported into Unity 3D.

4.3 Acting

Once the comparative evaluations are completed, it is then possible to recommend enacting the most desired—or least disruptive—action. This action is communicated to the relevant stakeholders via the building's two-way communication system, which as mentioned earlier in our case is by means of VLC. If the preferred action involves building systems, such as HVAC, lighting, etc., the preferred action may be communicated directly to the assets involved.

Like other "recommender" systems, such as GPS-based driving instructions, the actors may accept or ignore the recommended action. Either way, their action will be sensed by the building, and become input for the next round of simulation/evacuation/action.

5   CASE STUDY

To demonstrate the proposed system, we have chosen to implement it hypothetically on a Cardiac Catheterization Laboratory (CCL). A CCL is a suite of rooms in a hospital with diagnostic imaging equipment used to visualize the arteries and the chambers of the heart and treat any stenosis or abnormality found. It performs diagnostic, interventional, and electro physiology procedures, serving outpatients, inpatients, and emergency cases. Typically, 20-25 procedures are planned for each workday. In addition, the CCL handles 1-2 unplanned emergency cases every day. The procedure rooms are staffed by 15 staff members and operate from 7.30am to 5pm every day and may run overtime depending on the procedures and other emergencies.

While the CCL is only one unit in the hospital, it impacts, and is impacted by other units. The challenge is to evaluate those impacts to avoid conflicts and maximize the utility of the overall hospital. The simulation process necessarily requires abstraction of a complex system into a simplified

model and experimenting iteratively on it to test the relationship among many variables interacting in complex and often unpredictable ways [14].

The case study CCL has five Cath Labs: three Cardiac Catheterization (CC) labs, one Electro Physiology (EP), lab, and one Hybrid Cath Lab (HCL). A diagnostic procedure involves a team of three staff members (a cardiologist and two nurses) and lasts 20-30 minutes. An interventional procedure involves a six-member team (cardiologist, anesthesiologist, three nurses and a technician), along with a nurse in the observation area, and lasts 45-90 minutes. The labs interact with a 15-bed Cardiac Acute Care Unit (CACU) where patients are prepared for the procedure and recover from it.

Figure 2 depicts the typical (planned activities) workflow of the CCL operations, and includes the actors, activities, space and average duration for each activity. The case study focuses on the workflow that begins after pre-procedure preparation of the patients, either at the CACU for outpatients, the hospital nursing wards for inpatients, or the Emergency Department (ED), depending on the type of patient. The key activities include the patient transfers to and from the procedure room, patient preparation for the procedure at the procedure room and the actual procedure, along with the actors involved in the activities.

To demonstrate the proposed system, we look at the impact of an *unplanned event* on the CCL and the overall hospital. An unplanned event consists of emergency cases known as STEMI (ST-Elevation Myocardial Infarction). That is a very serious type of heart attack during which one of the heart's major arteries (one of the arteries that supplies oxygen and nutrient-rich blood to the heart muscle) is blocked. For STEMI patients, access to a facility with percutaneous

coronary intervention (PCI) capabilities is time-critical: Door-to-Balloon (D2B) time must be less than 90 minutes. The STEMI may arise within the hospital, at the inpatient ward and ED, or arrive from outside the hospital. While the typical protocol for a STEMI is going to the ED first, the D2B can be reduced by pre-activating the CCL for the STEMI thereby improving patient outcomes [15].

Although the frequency of STEMI cases is typically 1-2 a day, they can be considered 'unplanned' events that disrupt planned events at the CCL. A STEMI protocol requires an immediate activation of a suitable Cath lab and medical team to prevent delays in care [8].

This presents a suite of challenges within the CCL, as it may disrupt planned schedules for patients, medical teams, and planned allotment of CLs for the different procedures. Furthermore, not all the Cath Labs are suitable to treat a STEMI. Suitable CLs might be occupied with ongoing procedures at different stages of completion, and cardiologists may be occupied or have a scheduling conflict if assigned the STEMI patient. Furthermore, these challenges go beyond the CCL, as changes to the planned schedules can have a negative effect on other outpatients and inpatients: patients that were scheduled to undergo treatment may be bumped, requiring rescheduling (of outpatients) and longer stays (for inpatients). Hence, an action that may seem optimal for the CCL may adversely affect other units of the hospital, and thus be less optimal overall.

It is the goal of the system described here to critically evaluate all the options facing the CCL in case of a STEMI, and to recommend the overall most suitable plan of action for the hospital as a whole (subject, of course, to the constraints of the STEMI protocol and others).

Figure 2. Typical Cath Lab workflow for the planned procedures.
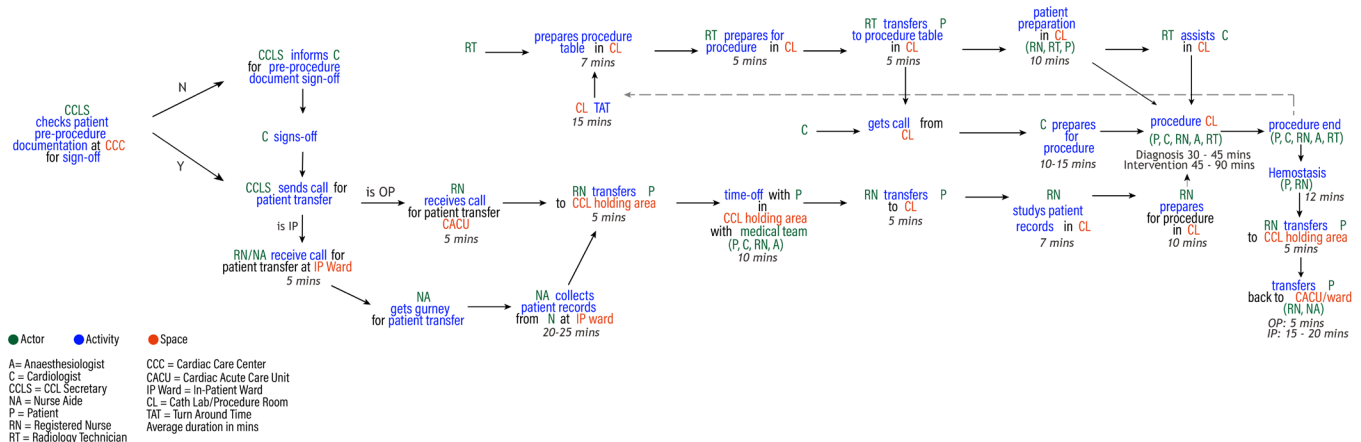
5.1 Sensing

To determine the overall best course of action we must start by detecting the state of the CCL when a STEMI protocol is declared. This is done through the VLC system and associated data profiles for the spaces, actors, and activities. Table 1 shows the hypothetical state of the five Cath Labs in terms of the types of patients and medical teams involved, and the type of ongoing procedures, expected duration and possibility for interruption of the ongoing procedures by the unplanned STEMI event.

Space	Type	Procedure	Patient	Duration	Time left
CL1	CC	Intervention	IP1	80 mins	40 mins
CL2	CC	Intervention	IP2	45 mins	45 mins
CL3	CC	Diagnosis	OP1	30 mins	30 mins
CL4	HCL	Diagnosis	OP2	30 mins	10 mins
CL5	EP	EP Study	OP3	35 mins	30 mins

Table 1. Narratives of the planned procedures in the CCL.
IP = In-Patient; OP= Out-Patient; CC=Cardiac Catheterization; HCL=Hybrid Cath Lab; EP = Electro Physiology

From Table 1 it can be determined that:

- The procedure in CL1 is half-way through.
- The procedure in CL2 has not yet begun.
- The procedure in CL3 has not yet begun.
- The procedure in CL4 is 2/3 complete.
- CL5 is an EP lab, therefore it is not suitable to treat the STEMI.

5.2 Decision-Making

To determine which one of the available labs to choose for treating the STEMI (CL1, CL2, CL3, or CL4), we need to simulate the consequences of choosing each one of the labs and evaluating their relative merits.

5.2.1 Simulation

Using the Event-Based simulation described earlier, we find:

- The procedure in CL1 cannot be interrupted, therefore that lab is not available to treat the STEMI.

- If CL2 is chosen to treat the STEMI, Patient IP2 (an inpatient) who was scheduled to be treated in that lab, will be bumped. The patient will be taken back to the inpatient ward, where he will stay at least another day before he will be treated (we assume his condition allows such postponement of the treatment). Consequently, patient IP2 will not be discharged as planned, and will continue to occupy a bed in the cardiac in-patient ward.

- The continued hospitalization of IP2 will prevent

admission of an incoming patient from the Emergency Department, who was scheduled to be hospitalized in the cardiac in-patient ward. Instead, she will have to stay in the ED for another 24 hours, at a great inconvenience to her and the ED staff.

- If CL3 was chosen to treat the STEMI, Patient OP1 (an outpatient) who was scheduled to be treated in that lab, will be delayed. She will be taken back to the CACU, delaying treatment of other outpatients scheduled for the day. Since the policy of the CCL is to treat all out-patients that were scheduled for the day rather than sending them back home to be treated another day, the CCL clinical staff will have to stay for a longer shift.

- If CL4 is chosen to treat the incoming STEMI, it will take 10 minutes to complete the ongoing procedure, and another 15 minutes to turn the CL around and make it ready for the incoming STEMI. This will result in 25 minutes delay in treating the STEMI.

5.2.2 Evaluation

The results of the simulations are evaluated comparatively to a list of Key Performance Indicators partially drawn from the literature [4,11] and discussed with an expert/lead-cardiologist at the hospital's CCL. Sixteen relevant and feasible KPIs were selected for the CCL, inpatient department and the emergency department. The KPIs were structured hierarchically and clustered into categories. For user specific KPIs that address 'patient satisfaction,' proxies such as patient wait times and staff load schedules were used. Inter-departmental relations were accounted for based on the goals or executive KPIs for the hospital overall to ensure there were no undesired trade-offs where processes within the CCL interact with processes outside the CCL.

The KPIs were grouped into three categories: 'operational,' 'user related' and 'space related,' affecting the operational efficiencies, user experience and space utilization. The KPIs in each category were ranked and prioritized based on relative importance and impact on outcomes. This was also done for the other departments (inpatient and emergency departments) with the assigned priority weights shown in Table 2.

The process included: (1) identifying the relevant KPIs under the categories, (2) evaluating the KPIs from the simulation results, (3) normalizing the results against the benchmarks and goals set by the individual departments and the hospital, (4) arriving at an overall score for each CL.

Benchmarks are used for normalization, with the values based on the organizational goals, performance targets, experts' experience, evidence-based design, policies, etc. The results are normalized and given a score based on the assigned weights for the KPI.

The scores obtained for each category within the department helps understand tradeoffs between the categories. For example, between the operational and user related categories, 'average patient wait times', 'staff load schedule' which impact patient and staff satisfaction could have a higher priority to the 'average LOS', where LOS is the length of stay of the patient in the hospital (a critical KPI). In case of conflicts, the organization's policies and preferences are obtained for the recommended action.

Similarly, the scores obtained between the departments help understand conflicting needs and consequences of actions on the KPIs. The simulation results and ranking based on the scores obtained within the department and between the departments is shown in Table 3.

Category	Key Performance Indicator (KPI)	Unit	Exec	CCL	IP	ED
	Average LOS	days	20	-	10	-
	Bed turnover rate	#	10	-	25	25
	Cancelled Procedures	%	10	5	20	-
	Average LOS for ED	hours	10	-	-	10
	First case on-time starts	%	-	10	-	-
	Average procedural time for procedure	mins	-	5	-	-
Operations	Turnaround time between cases (TAT)	mins	-	10	-	-
	Average time on pre & post procedure holding area by procedure	mins	-	5	-	-
	STEMI patients with D2B ≥ 90 minutes	#	-	15	-	-
	Average LOS post procedure	hours	-	5	-	-
	Overall Patient throughput	hours	-	5	-	-
User	Average Patient Wait Times	hours	15	15	25	25
	ED Waiting Time	mins	20	-	-	20
	Staff Load Schedule	hours	-	5	-	-
Space	Bed Occupancy	%	15	10	20	20
	Room / Asset utilization	%	-	10	-	-

High Priority   Low Priority   - Not Applicable

Table 2. Key Performance Indicators (KPI) and priority weights. Exec=Executive KPI; CCL=Cardiac Cath Lab KPI; IP=Inpatient Department KPI; ED=Emergency Department KPI; LOS=Length of stay.

	Exec		CCL		IP		ED		Score	Rank
CL2	75.6	3.8	84.2	58.9	65.9	6.6	97.3	14.6	83.9	3
CL3	98.3	4.9	97.0	67.9	98.3	9.8	97.3	14.6	97.2	1
CL4	91.2	4.6	98.6	69.0	92.6	9.3	91.8	13.8	96.6	2

Intra department and Inter department scores

Table 3. Simulation results and ranking. Intra=Within or inside the department; Inter=Between the departments.

5.3 Acting
Based on the comparative evaluations, the most desired—or least disruptive—action is communicated to the relevant stakeholders. This is done via the building's two-way communication system, which as mentioned earlier in our case is by means of the VLC. Like other "recommender" systems, such as GPS-based driving instructions, the actors

may accept or ignore the recommended action. Either way, their action will be sensed by the building, and become input for the next round of simulation/evacuation/action.

6. CONCLUSION AND DISCUSSION

We argue that total environmental adaptability—namely, buildings that can dynamically respond to and interact with their occupants—will be achieved when the building's omniscience is harnessed in the service of its inhabitants. Such omniscience implies that the building, unlike its inhabitants, "knows" all that is happening within and around it at any given moment. When coupled with operational procedures and occupant profiles, such knowledge can be leveraged to predict and evaluate future events and recommend choosing the most beneficial one overall.

Towards this end, we presented a simulation-powered building management system that can sense human and building assets based on Visible Light Communication (VLC) technology; simulate alternative future building occupancy scenarios; evaluate them according to specific KPIs for the purpose of choosing one that will minimize/maximize users' welfare and resource allocation.

We demonstrated such abilities in the case of a hospital's Cardiac Catheterization Lab. Results indicate that it holds promise to enhance decision-making capabilities of building inhabitants, thus enabling building management strategies that support human needs and efficiency requirements, especially in mission-critical facilities.

The work reported here is on-going: the VLC system is being developed and tested in lab settings, with commitment for deployment at St. Bernardine Medical Center in California. Data is being gathered, through observations and interviews, of on-going CCL procedures, and event-based simulation software developed earlier is being adapted to whole-building scenarios.

REFERENCES

1. Berger, H. Stephen, and H. Mark Gibson. Managing your hospital RF spectrum. *Biomedical instrumentation & technology* 47.3 (2013), 193-197.

2. Capolongo, S. *Architecture for flexibility in healthcare*, Franco Angeli, 2012.

3. Conti, J.P. What you see is what you send, *Engineering & Technology* 3.19 (2008), 66-68,

4. De Pourcq, K., Gemmel, P., and Trybou, J. Measuring Performance in Hospitals: The Development of an Operational Dashboard to Coordinate and Optimize Patient, Material and Information Flows. In H. Albach, H. Meffert, A. Pinkwart, R. Reichwald, & W. von Eiff (Eds.), Boundaryless Hospital: Rethink and Redefine Health Care Management, (2016), 159–181.

5. Kalay, Y.E. *Architecture's New media: principles, theories and methods of computer-aided design*. MIT Press, Cambridge, MA, 2004.

6. Kalay, Y. E. Evaluating and predicting design performance. Wiley, 2016.

7. Kim, J.K. and Schubert, E.F. Transcending the replacement paradigm of solid-state lighting. *Optics Express* 16, 26, (2008), 21835-21842.

8. Martel, T. J. Activation of the Cardiac Catheterization Lab for STEMI Patients. *Emergency Medicine* 49,6 (2017), 259–262.

9. Mozer, M.C. The Neural Network House: An Environment that Adapts to Its Inhabitants. *AAAI Spring Symp. On Intelligent Environments,* (1988), 110–114.

10. Pan, Z. Lang, T. Li, C. Di, M. Chen, G. Kalay, Y.E. Pai, R. and Wang, A. Visible Light Communication Cyber-Physical Systems-on- Chip for Smart Cities. *Journal of Communications*, 14, 12, 2019.

11. Rahimi, H., Kavosi, Z., Shojaei, P., and Kharazmi, E. (2016). Key performance indicators in hospital based on balanced scorecard model. *Journal of Health Management & Informatics* 4.1 (2017), 17-24.

12. Schaumann, D., Date, K., and Kalay, Y.E. An event modeling language (EML) to simulate use patterns in built environments." SimAUD, (2017), 21-28.

13. Schaumann, D., Pilosof, N.P., Sopher, H., Yahav, J., and Kalay, Y.E. Simulating multi-agent narratives for pre-occupancy evaluation of architectural designs." *Automation in Construction*, 106, 102896, (2019).

14. Shannon, R. E. Introduction to the art and science of simulation. Proceedings of the Winter Simulation Conference. (1998).

15. Martel T.J., Nambudiri V., Kirkman D., Martel J., Pappas L., and Jehle D. Activation of the Cardiac Catheterization Lab for STEMI Patients. *Emergency Medicine*. 2017 June; 49(6):259-262.

Evaluating Temporal and Spatial Light Exposure Profiles for Typical Building Occupants

Megan Danell[1], María L. Ámundadóttir [2] and Siobhan Rockcastle[1]

[1]Baker Lighting Lab
University of Oregon
Eugene, USA
{mdanell, srockcas}@domain

[2]OCULIGHT analytics
Reykjavik, Iceland
maria@oculightdynamics.com

ABSTRACT

Many of the daylight metrics that we use today rely on grid-based illuminance data to evaluate the spatial distribution of daylight at a building scale, but we lack methods to evaluate how an occupant's temporal and spatial behavior impacts our assessment of performance. This is particularly problematic when we consider the effects of light on human health. Light-exposure at eye-level drives responses in both the visual and non-visual (circadian) systems. This paper adapts a simulation workflow to evaluate the non-visual effects of light using existing models like Equivalent Melanopic Lux (EML) and non-visual Direct-Response (nvRD) at an occupant scale, accounting for typical user profiles, sky condition, and time of year. Our paper considers four typical user profiles within a side-lit office environment to query location and view directions from a multi-point, multi view-direction, multi time-step simulation of eye-level illuminance and return individual light-exposure profiles. This approach allows us to compare each profile as a product of the occupant's spatial and temporal behavior and begin to consider the lighting performance of our case study building *through* the performance of its users.

Author Keywords

Daylight performance; health; non-visual; circadian; lighting simulation; human-centric lighting; occupant behavior; visualization

ACM Classification Keywords

•Applied computing~Arts and humanities~Architecture (buildings)•Computing methodologies~Modeling and simulation

1 INTRODUCTION

As we continue to learn more about the circadian and direct effects of light on human health, new models and tools have emerged to quantify the effect of light exposure on the health potential of building occupants [2, 16, 18, 19, 22]. While organizations like the International WELL Building Institute have recommended a standard to support circadian health by setting a minimum threshold, we lack any grounded methods to simulate the eye-level illuminance that a user receives *over space and time* [24].

As daylight is most often analyzed from a horizontal task-plane, a shift towards the consideration of eye-level light exposure from a *dynamic* user perspective would create a radical shift in our understanding of lighting performance at the building scale. Rather than quantifying illuminance levels over time across the floor area of a building, we would need to consider the spatial and temporal behavior of building occupants and compute the aggregated performance of a building's users.

As building occupants move throughout a space over the course of any given day, they are exposed to variations in brightness and spectrum. Changes in sky condition, time of year, time of day, user location, and view-direction creates a vast matrix of possible light exposure profiles for any given user in a space. Depending on where a building occupant spends time, their accumulated light exposure profile will change throughout the day, week, month, or year, impacting alertness, sleep quality, and overall health. To provide a more robust assessment at the occupant level, we must account for both spatial and temporal human behavior within a given building.

This paper proposes a novel method to evaluate and compare dynamic user light-exposure profiles over time and throughout space. To exemplify this method, we've created four typical user profiles for hypothetical building occupants in a side-lit office space in Portland, OR, USA. We queried illuminance values for the exact location and view direction of each user profile and used it as input for both the WELL Building Standard and the non-visual Direct-Response (nvRD) model [2, 24]. By doing this, we are able to illustrate the impact of occupant behavior on daylight performance and compare the performance prediction of existing metrics, which offer divergent performance narratives through the way they 'count' light towards a daily dose. While these four user profiles cannot give a holistic overview of daylight performance for the entire occupied building, this method could be used to evaluate the performance of all building users and begin to create an aggregated narrative about healthy light exposure from a dynamic occupant perspective. Using this approach, we are able to determine if and when various building

occupants would reach a recommended daily light dose and discuss the different outcomes of WELL Building Standard and nvRD models in influencing building design decisions.

2 BACKGROUND

In order to begin evaluating the effects of eye-level exposure on human health, a handful of studies have suggested new evaluation methods to transition from a horizontal task-plane (i.e. illuminance measurements across a 2D surface) to a vertical view-based approach.

Using circadian equivalent threshold values, Pechacek et al. (2008) evaluated vertical illuminance at the eye-level [17], which was later extended to a multi-point, multi-view direction approach by randomly sampling point values for two locations and eight view directions from a uniform distribution [6].

A 2017 study by Amundadottir et al. proposed a method to simulate eye-level daylight performance from an occupant's perspective using rendered 360-degree images that could be unrolled and analyzed across multiple view directions [2]. From these unrolled view-directions, the nvRD model was applied to predict direct non-visual responses in addition to other view-based performance models. While that method used a single point in space, later publications by Rockcastle et al. (2018,2019) applied the same method to an array of view directions, resulting in the OCUVIS web-based visualizer that illustrates the frequency of views that exceed a given threshold for both visual (glare, task brightness, and visual interest) and non-visual light responses (nvRD) [19, 20].

Analyzing light exposure patterns from a fixed sensor point might, however, be too reductive as we move our eyes and head continually by gaze direction and position in space. As a preliminary attempt to test this concept, four different strategies were applied to generate light exposure patterns based on occupants' spatial behavior [5]. These light exposure patterns where used as an input for a preliminary version of the nvRD model [4]. The results confirmed that the use of a space is an important factor, which strongly relates to the temporal pattern property of light.

Additional models have been developed to simulate non-visual or "circadian" responses to light exposure. These models, like Circadian Stimulus (CS) [18] and Equivalent Melanopic Lux (EML) [16, 24] have been integrated into international standards/reports [8, 22] and simulation tools like Adaptive Lighting for Alertness (ALFA) [22].

A 2019 study by Saiedlue et al. used ALFA to quantify how various light sources and glazing types affect non visual light exposure within a space [21]. Acosta et al. considered the effect of view direction by simulating both horizontal and vertical illuminance to represent a healthcare patient either lying down or sitting up [1]. In this study, 3D models were simulated in two latitudes to test high and low room reflectance values and window to wall ratios (10%-80%). A 2017 study by Konis used a grid of view positions and

view directions to map the strength and frequency of the minimum acceptable stimulus frequency threshold on a daily basis within a space [14].

While multi view-direction simulations have become more common, so to have methods of multi-spectral computation using HDR images. Inanici et al. (2015) calculated circadian values with both Rea et al. (2005) circadian spectral sensitivity curve [18] and Lucas et al. (2014) melanopsin spectral sensitivity curve [16]. A north-facing conference room in Seattle, WA was used to test how red and blue partition walls would impact instantaneous photopic and melanopic illuminance using a false color analysis of the space under clear, intermediate, and overcast skies [13].

Despite the uptick in simulation workflows that include non-visual/circadian models, there is still a lack of methods for applying those models to a dynamic user profile rather than to a fixed position in space. Both Amundadottir et al. and Figueiro discuss the importance of considering light exposure history to determine an occupant's total light dose [2, 12]. History implies the knowledge of what a user has been exposed to over time.

A 2018 field study by Konis used similar methods to his simulation-based workflow [14, 15], but collected physical light exposure data in a variety of existing dementia care facilities. The results from this experiment emphasize the importance of location and view direction in receiving an effective circadian light exposure.

Figueiro et al. (2019) conducted an experiment using tunable luminaires at a user's desk to provide an optimal light spectrum throughout the day. Participants wore Daysimeters [7] to calculate their non-visual light exposure and smart watches that logged the user's activity levels. While they should have received sufficient light from the luminaires, the time not spent at their desk and the movement of data loggers throughout the study were both found to impact the participant's circadian light exposure [10, 11].

Of the field and simulation studies mentioned here, none have integrated exposure history, timing, user location, view direction and spatial behavior in a single study. Being able to account for the spatial and temporal behavior of building occupants in a simulation workflow would allow us to predict the circadian performance of each user and compare factors that impact that performance. From the architectural design of the building to the choice of seating location and scheduled activities, we could offer a more holistic assessment of daylight performance on human health.

3 METHODOLOGY

In this paper, the authors adapted a simulation-based framework developed by Amundadottir et al. (2013) to investigate a dynamic occupants' daily light exposure to daylight and its impact on their health and well-being in an

office space [5]. An array of sensor points was simulated across a side-lit building to reproduce all possible locations that an occupant may sit or stand for more than 15 minutes at a time. Each sensor point produced an array of 8 vertical view-directions. Four occupant profiles were then used to query illuminance data for each position, view direction, date, sky, and time-step to evaluate the occupant's light exposure profile across the day. Two sets of performance criteria were used to compare the different light exposure profiles and testing visual representation.

3.1 Daylighting Simulation

Our selected case study consists of a two-story side-lit office building with double-height glazing along the South-West facing facade. This building is located in Portland, OR at 45.5 N, 122.67 W (Figure 1). The building was recently remodeled by SRG Partnership and houses their Portland design offices.

Three variables were simulated in this study: sky condition (clear and overcast), date (March 20th, June 21st and December 21st), and time (9am-6pm in 15-minute intervals. Gensky was used to create CIE clear and overcast sky conditions for these three annual instances.

Illuminance levels are computed vertically at eye-level to simulate the light entering the eye. Photosensors were placed at a height of 1.14m above the floor for standard chairs, and 1.5m above the floor for high stools at workstations and at other locations. Eight view directions were simulated at each position and were aligned to building geometry and desk orientation to account for the slight west-facing angle of the south facade.

Default rtrace RADIANCE v5.2 parameters were used except for the following adjustments to: -dt 0.05, -dc 0.5, -ds 0.15, -dr 3, -ab 3, -aa 0.15, -ar 32, -ms 0.066, -lr 8, -lw 0.002.

No blind controls or devices were used for consideration within this study to minimize variables, but further research should address the use of dynamic facade controls
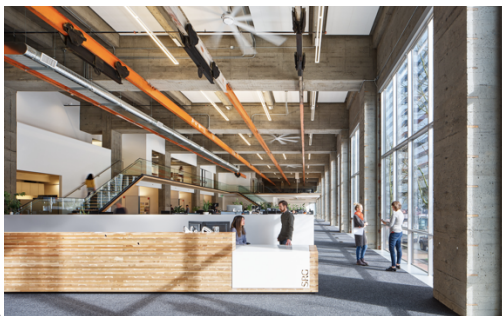
.
Figure 1 A photo of the selected case-study in Portland, OR. Image courtesy of SRG Partnership.

3.2 Occupant Profiles

The workspace was assumed to be occupied from 9am to 6pm. During these hours occupants generally engage in different activities and move around for meetings, lunch, and breaks (Figure 2). Because the time-step used in this study is 15-minutes, none of the circulation zones were simulated as people would typically only spend a second in each as they pass through.

Figure 2a shows schedules for four typical office workers and Figure 2b shows their path in the office space. These profiles were created manually based on typical functions that each user may be expected to perform.

- Occupant 1: The Senior Manager profile consisted of desk work and intermittent meetings, with a majority of time spent receiving phone calls and answering emails at their desk. Their position was located directly adjacent to the windows, with optimal views.

- Occupant 2: The Junior Manager profile was scheduled for more intermittent desk work, with meetings and trips to the material lab that regularly interrupted desk time. Their desk was located farther from the window on the ground-level where they could easily manage their team.

- Occupant 3: As an entry-level position, the Intern profile spent most of their time between their desk and the material lab, with their desk located on the mezzanine, farther from the window in a less desired zone of the space.

- Occupant 4: The fourth employee profile is represented as the IT Manager. Their desk was located in the back of the space adjacent to the company servers. They were scheduled to make frequent visits into the common space on the first level to assist other employees with computational issues.

3.3 Non-visual light-response evaluation

In order to evaluate the effects of light on the non-visual system, RGB values obtained from RADIANCE simulations were weighted and summed into melanopic illuminance, which is based on the spectral sensitivity of the ipRGCs. The weights are obtained from [13]. The melanopic illuminance values are used as inputs to both the WELL Building Standard criteria for circadian lighting and the nvRD model [3, 24]. The nvRD model was then translated into a dose measure (i.e. cumulative response). Unlike the implementation of WELL Building Standard, the nvRD model is influenced by continuous light exposure throughout the day (not only 9am – 1pm) and also considers variations in light intensity, wavelength, duration, pattern, and history. This makes it more sensitive to intermittent light exposure or short periods of bright light exposure.

3.4 Performance criteria for visual representation

The WELL Building Standard requires a certain threshold of melanopic illuminance which is usually converted into photopic illuminance based on the spectral distribution of individual light sources using the term Equivalent Melanopic Lux (EML).
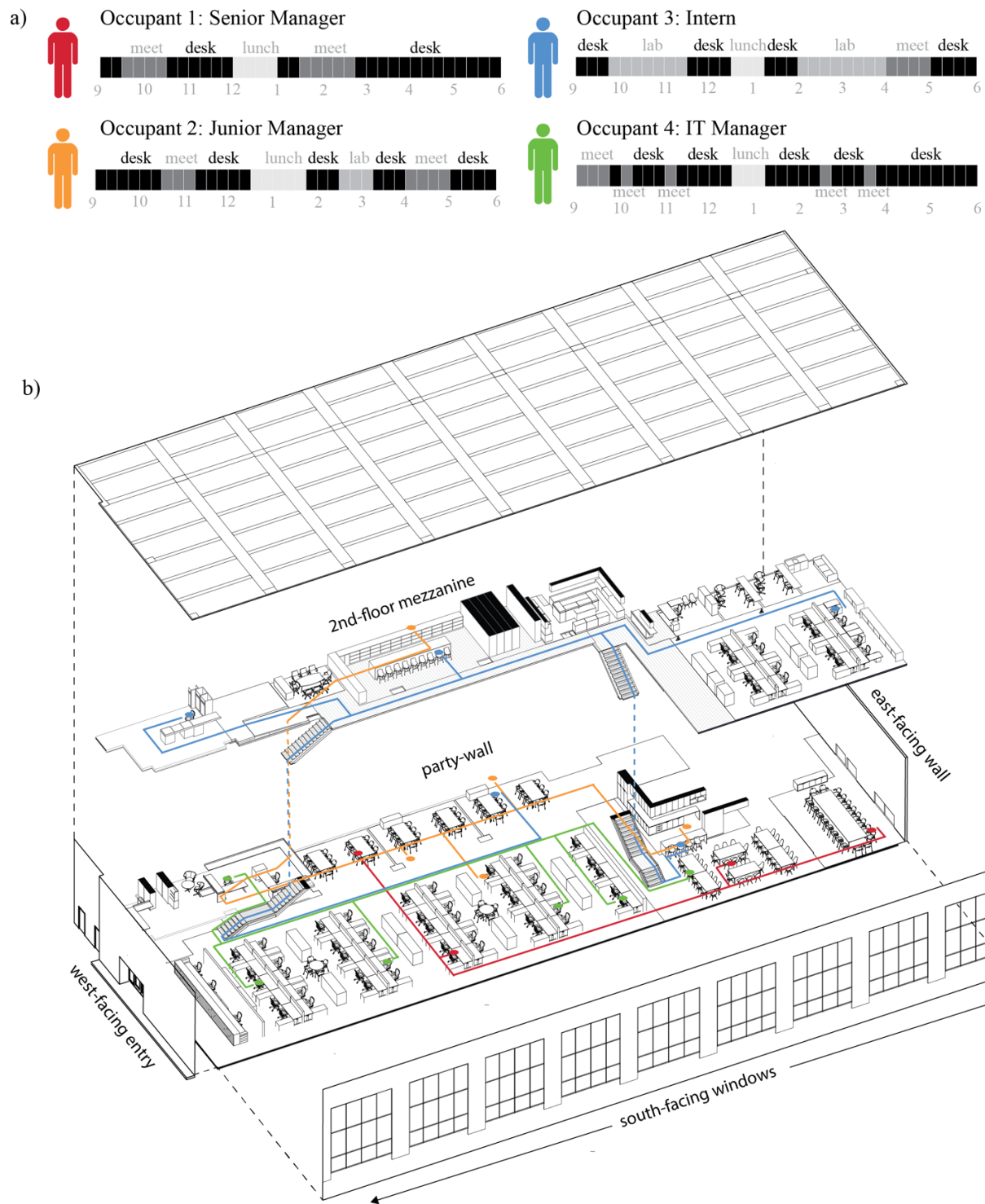
Figure 2 a) Occupant profiles and their daily schedules. B) Spatial distribution of occupant profiles in the side-lit office space.

The WELL Q4 2019 version Feature 54: Circadian Lighting Design [24], implements a minimum threshold of 200 lx of equal-energy light source which is equivalent to 182 lx of daylight illuminant D65. This threshold applies to work areas measured at the vertical plane, which must be achieved (between 9am to 1pm) for every day of the year. While the authors acknowledge that WELL Building Standard assumes the use of electric lighting to supplement daylight during occupied hours, this paper attempts to provide a means of comparing occupant profiles in relative terms.

For the nvRD model, the non-visual response results are accumulated over the day to predict the *daily light dose*. The threshold value of 4.2 was established in [2] as a reasonable criterion for achieving beneficial effects of light. For the purpose of this paper it is lowered to 4, so it is comparable to the period length applied in WELL Building Standard criteria. These tentative thresholds correspond to the number of "vital" hours in a day, when adequate daylighting is desired.

4 RESULTS

The results of this study represent a substantial shift in the way we quantify and visualize lighting performance at a building scale. While existing annual daylight performance metrics account for time and space, which already requires a significant abstraction of information to achieve a compact number, our approach adds yet another dimension with the consideration of occupant activity. To adequately cover the impact this has on our daylight performance narrative, we have organized the results into three sections.

4.1 Daily Performance Overview

This section presents a compact overview of results for our four hypothetical building occupants on June 21st and December 21st. Table 1 shows the daily average of vertical illuminance (E_v), WELL Building Standard criteria (percentage of time when $E_v \geq 182$ lx), and daily nvRD cumulative response. Yellow cells represent when $E_v \geq 182$ lx is achieved between 9am and 1pm. Green cells indicate when a recommended nvRD dose of 4 is achieved.

Table 1. Daily performance overview comparing results for a) June 21st and b) December 21st for all occupants.

a)	clear			overcast		
Occ. id	Ev	WELL	nvRD	Ev	WELL	nvRD
1	2391	84%	7.2	729	76%	6.9
2	982	73%	7.0	335	62%	5.8
3	259	8%	1.5	70	8%	1.1
4	714	22%	4.5	186	19%	3.6

b)	clear			overcast		
Occ. id	Ev	WELL	nvRD	Ev	WELL	nvRD
1	3483	86%	7.1	286	62%	5.2
2	706	73%	6.7	126	14%	3.5
3	199	8%	2.3	23	0%	0.8
4	533	32%	4.6	92	22%	2.8

This table can be used to provide a high-level comparison between our four building occupants, revealing the impact of day (June 21st vs. December 21st) and sky condition (clear vs. overcast) on the percentage of time when E_v exceeds the WELL threshold as well as the cumulative

response nvRD achieved during occupied hours. Based on the performance criteria, the light exposure for Occupants 3 and 4 is the same between the two days, where Occupants 1 and 2 show minor variations.

While this table is helpful in comparing multiple users, it does not reveal the temporal granularity of light exposure that underlies either the WELL Building Standard or the nvRD model and is not super useful in revealing which activities and locations may be preventing the occupants from receiving the recommended light exposure.

4.2 Understanding Temporal Light Exposure

This section explains two different visual representations of the temporal light exposure for a single user across two days (June 21st and December 21st) and two sky conditions (clear and overcast). Figure 3 shows the daily profile for Occupant 2: Junior Manager with the time and activity they were performing indicated along the x-axis. Clear sky results are indicated in grey, overcast in black, and the $E_v \geq 182$ lx threshold is shown as a dotted line.

If we refer to Figure 2b (which shows the occupant paths in axonometric), this occupant's desk position was located several seats from the window, with a head position facing East. As shown in Figure 3 for both days and sky conditions, the occupant received the highest light exposure levels while at lunch and secondarily while at their desk. As they periodically got up from this position to attend meetings and visit other locations within the office, this profile varies as a result of the occupant's shifting position in space and the shifting position of the sun.

Assuming the same exact schedule as on June 21st, Figure 3b shows the exposure profile for the same occupant on December 21st. While the impact of day (June 21st vs. December 21st) and sky condition (clear vs. overcast) may seem obvious, the impact on eye-level exposure may be less intuitive when shifting solar altitude angles change the depth of sun penetration through a building envelope. As a result, some occupants may be exposed to brighter pulses of light during the winter months.

Figure 3 clearly illustrates the dynamic nature of light exposure for a typical building occupant, where scheduled activities determine the available light received at eye-level. The impact of sky condition is clearly shown when comparing the line graph in Figure 3a and 3b, however this type of graph is not well suited for viewing multiple days. Thus, this information can be translated into a heatmap as shown below the line graph. The heatmap maps the $E_{v,ipRGC}$ values on a logscale to four colors, ranging from dark red to bright yellow. The $E_{v,ipRGC}$ values that exceed 200 EML or 182 lx are labeled with the number 1 to indicate those time periods that would meet the recommended light exposure under the WELL Building Standard. Representing the results in a more compact way, will allow us to view the results on a temporal map for the full year.
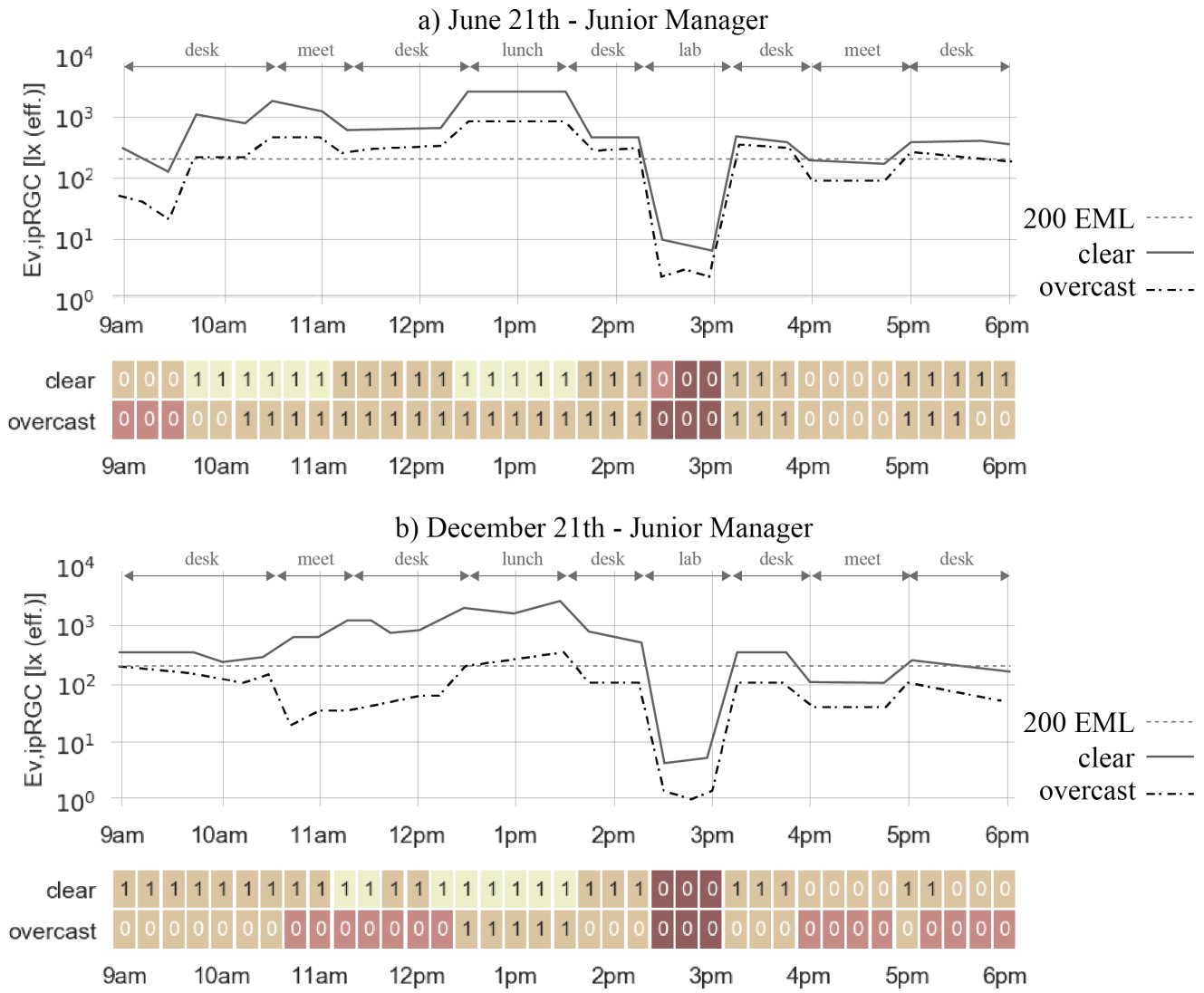
Figure 3 The effective ipRGC illuminance $E_{v,ipRGC}$ (also called melanopic illuminance) as a function of time-of-day for the Junior Manager on a) June 21st and b) December 21st. The upper graph shows the $E_{v,ipRCG}$ on a log-scale for clear and overcast skies. The dashed line shows the threshold of 200 EML. The lower graph shows the same data as the upper graph. The color represents the log-scale $E_{v,ipRGC}$ divided into 4 bins by the gridlines in the graph above ranging from low illuminance (dark red) to higher illuminance (bright yellow). 0 and 1 indicate whether the value is below or above the 200 EML threshold, respectively.

4.3 Comparing Occupants & Metrics Over Time

To better understand the impacts of time and schedule on the performance of our four building occupants, Figure 4 shows a) the light exposure between 9am and 6pm on March 20th in a plain line graph, b) the derived analysis by highlighting periods (or timesteps) when the E_v that exceeds the 182 lx threshold between 9am-1pm as outlined in the WELL Building Standard criteria, and (c) the cumulative nvRD received during occupied hours from 9am - 6pm. This compact comparison reveals the disconnect in performance narrative between the WELL Building Standard criteria and the nvRD model. As WELL

only takes into consideration the light exposure that is received between 9am - 1pm, some occupants will never achieve WELL using daylight alone.

The nvRD model is more sensitive to periodic light exposure profiles, like the one experienced by Occupant 4 (the IT Manager). While the WELL Building Standard criteria would not award credit to this type of light exposure profile due to the lack of continuous exposure over the $E_v \geq 182$ lx threshold, the nvRD model shows that this occupant was able to achieve the recommended dose by the late afternoon.

544
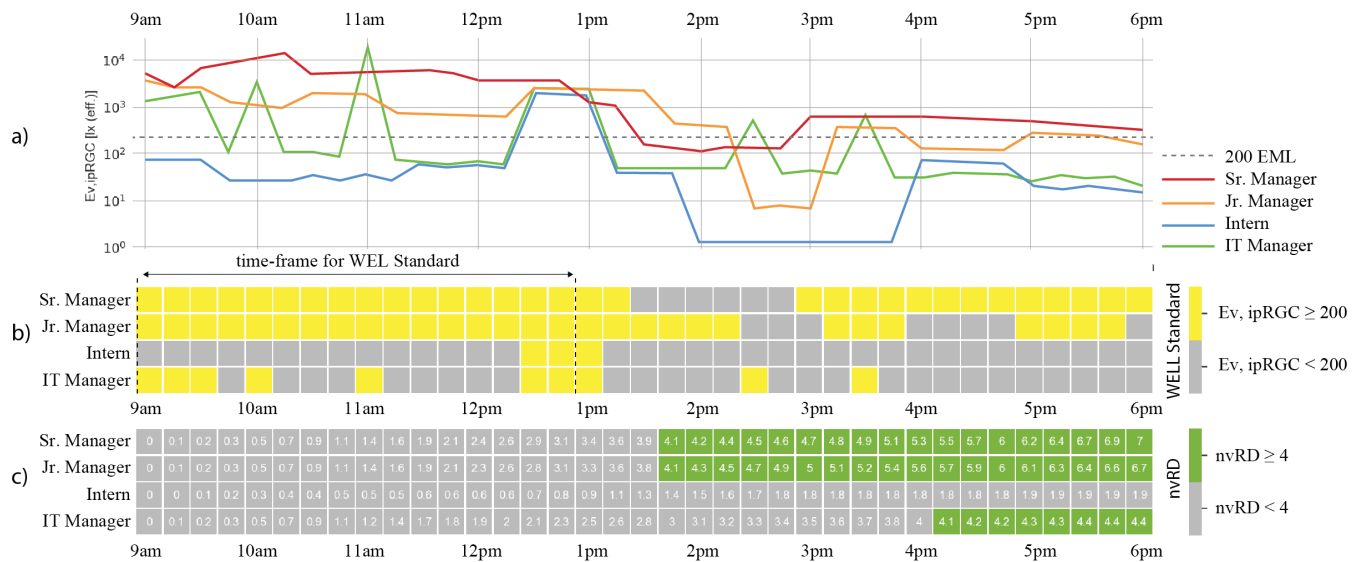
Figure 4 The temporal light exposure for the 4 occupant profiles on March 20th under a clear sky a) The effective ipRGC illuminance $E_{v,ipRGC}$ as a function of time-of-day. b) A heatmap showing when a minimum threshold of 200 EML is achieved or not. c) A heatmap showing when the daily recommended light dose (nvRD ≥ 4) is achieved or not and the cumulative response value at every timestep.

In architectural spaces where light exposure may be continuous, but occupation is intermittent and/or more dynamic, continuous threshold-driven methods like the WELL Building Standard may provide an incomplete narrative about healthy light exposure. The nvRD model accounts for a wider set of variables that influence the non-visual system. As such, nvRD may offer a more robust way to evaluate the impact of intermittent light exposure, which is a reality experienced by occupants who move throughout an office over the course of a day. Accounting for this dynamic behavior may encourage us to think about the impact of programming as well as seating location. A well-lit break room or coffee station frequented by many occupants could help enhance the intermittent exposure of those who are more deprived in their current seating location.

The comparison between occupant profiles underlines the importance of considering spatial and temporal behavior when predicting the health potential of a space for its occupants. The comparison between the WELL Building Standard and nvRD model reveals the need for rigorous discussion about how these models are implemented in an architectural context and how they may be integrated into predictive simulation-based workflows.

5   CONCLUSION

This paper introduced a method for evaluating the circadian performance of daylight for four building occupants based on their spatial and temporal behavior in a side-lit case study. With an increased understanding of dynamic lighting environments, human behavior, and the effects of light on health and well-being, we can produce better predictions regarding the benefits of daylight and eventually, their integration within the design process. As a next step for this research, we would like to extend our evaluation to include the full population of building occupants and develop an aggregated metric that relies on a percentage of people that achieve performance rather than a percentage of space.

By shifting the narrative of daylight performance from the space of the building to the performance of its occupants, we can place more emphasis on the impact of architecture on health in the built environment. In order to do this, we need robust methods of predicting occupant behavior through space and over time. Future work in this area could consider agent-based modelling to generate user profiles in the absence of recorded field data or in the design-phase before occupant profiles have emerged. Building on work by Breslav et al, 2014 and Schaumann et al., 2015, this research could use behavioral narratives to predict granular narratives that acccount for a fine-grain temporal resolution [8, 23]. To provide a holistic evaluation of health in our buildings, the future of predictive daylight modeling must consider human behavior or we may continue to miss a large part of the performance narrative.

REFERENCES

1. Acosta, I. et al.: Analysis of circadian stimulus allowed by daylighting in hospital rooms. Light. Res. Technol. 49, 1, 49–61 (2017).
2. Amundadottir, M.L. et al.: A human-centric approach to assess daylight in buildings for non-visual health potential, visual interest and gaze behavior. Build. Environ. 113, 5–21 (2017).
3. Amundadottir, M.L.: Light-driven model for identifying indicators of non-visual health potential in the built environment. Ecole polytechnique fédérale de Lausanne (2016).
4. Amundadottir, M.L. et al.: Modeling non-visual responses to light: unifying spectral sensitivity and temporal characteristics in a single model structure. In: CIE Centenary Conference "Towards a New Century of Light." pp. 101–110 Commission Internationale de L'Eclairage, Paris, France (2013).
5. Amundadottir, M.L. et al.: Simulation-based evaluation of non-visual responses to daylight: proof-of-concept study of healthcare re-design. In: BS2013. pp. 2757–2764 , Chambery, France (2013).
6. Andersen, M. et al.: Modelling 'non-visual' effects of daylighting in a residential environment. Build. Environ. 70, 138–149 (2013).
7. Bierman, A. et al.: The Daysimeter: a device for measuring optical radiation as a stimulus for the human circadian system. Meas. Sci. Technol. 16, 11, 2292 (2005).
8. Breslav, S., et al.: Towards Visualization of Simulated Occupants and their Interactions with Buildings at Multiple Time Scales. SimAUD 2014, Tampa (2014).
9. CIE: Report on the First International Workshop on Circadian and Neurophysiological Photometry, 2013. CIE Central Bureau, Vienna, Austria (2015).
10. Figueiro, M. et al.: Circadian-effective light and its impact on alertness in office workers. Light. Res. Technol. 51, 2, 171–183 (2019).
11. Figueiro, M. et al.: Light, entrainment and alertness: A case study in offices. Light. Res. Technol. (2019).
12. Figueiro, M.G.: Disruption of circadian rhythms by light during day and night. Curr. Sleep Med. Rep. 3, 2, 76–84 (2017).
13. Inanici, M. et al.: Spectral daylighting simulations: computing circadian light. In: Proceedings of BS2015: 14th Conference of International Building Performance Simulation Association. , Hyderabad, India (2015).
14. Konis, K.: A novel circadian daylight metric for building design and evaluation. Build. Environ. 113, 22–38 (2017).
15. Konis, K.: Field evaluation of the circadian stimulus potential of daylit and non-daylit spaces in dementia care facilities. Build. Environ. 135, 112–123 (2018).
16. Lucas, R.J. et al.: Measuring and using light in the melanopsin age. Trends Neurosci. 37, 1, 1–9 (2014).
17. Pechacek, C.S. et al.: Preliminary method for prospective analysis of the circadian efficacy of (day)light with applications to healthcare architecture. J. Illum. Eng. Soc. N. Am. 5, 1, 1–26 (2008).
18. Rea, M.S. et al.: Circadian light. J. Circadian Rhythms. 8, 1, 1–10 (2010).
19. Rockcastle, S. et al.: OCUVIS: A Web-Based Visualizer for Simulated Daylight Performance. Presented at the SimAUD , Delft, The Netherlands June 5 (2018).
20. Rockcastle, S.F. et al.: A Simulation-Based Workflow to Assess Human-Centric Daylight Performance. In: Proceedings of the 8th Symposium on Simulation for Architecture and Urban Design. , Toronto, Canada (2017).
21. Saiedlue, S. et al.: Assessing circadian stimulus potential of lighting systems in office buildings by simulations. In: ARCC Conference Repository. (2019).
22. Solemma.com: ALFA: Adaptive Lighting for Alertness, https://www.solemma.com/Alfa.html, last accessed 2020/01/10.
23. Schaumann, D. et al, D. (2015). Simulating human behavior in not-yet built environments by means of event-based narratives. SimAUD 2015, Washington D.C. (2015).
24. WELL: The WELL Building Standard - Feature 54: Circadian Lighting Design. International WELL Building Standard (2019).

Fabrication

Implicit Representation for 3D-Printing Aware Shape Design

Efthymia Douroudi

University College London,
London, UK,
efthymia.douroudi.18@alumni.ucl.ac.uk

ABSTRACT

3D printers create objects directly from digital models by stratifying material layer upon layer. A problem that could arise is that, in case an object has surfaces without support material underneath, the print might fail. One can add support structure to a predefined geometry, or segment it, in printable parts, to avoid such failure. However, this paper suggests that design can change such that there are no unsupported areas. The modification of the input design geometry is carried out under the constraints of minimising overhang areas and ensuring balance during printing. Furthermore, changing shape in small increments keeps the difference of the new shape compared to the initial one as small as possible. Implicit geometry -geometry represented using implicit representation- is employed throughout the research, due to its efficiency in topological changes and accuracy compared to mesh or surface. A tool developed to modify shapes minimising the support needed for them to be printed, and the underlying computational framework are presented. This tool not only reduces material and makes manufacturing process faster, but also assists the user in creating variable fabrication aware geometries enhancing creativity.

Author Keywords

3D printing; implicit geometry; function representation; shape processing;

1. INTRODUCTION

The motivation of this research was the rapid growth in desktop and large-scale additive manufacturing and the lack of exploratory design tools thereof [3, 7].

1.1. Additive Manufacturing Technology
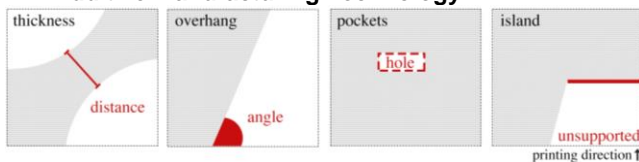
Figure 1. 3D printing constraints (Douroudi and Tsomokou 2019).

Layer-by-layer additive manufacturing -3D printing (3DP)- is a popular technique that a lot of designers and architects use during the ideation and fabrication process. Desktop 3D printers are accurate and fast enough to produce customised products, causing a revolution in industries including: product [2, 12], fashion [18, 24] and sculptures [15, 17]. Also, large-scale application using industrial robots [1, 21], although still in its infancy, is bringing into architecture, engineering and construction the automation technology and control of processes found in the automotive and aerospace industries [8, 10, 23].

However, there are still some limitations that designers have to be aware of during the design process. The most common geometric constraints are thickness, pockets, overhang and island (Figure 1), which usually depend on the 3DP technique [3, 19].

This paper focuses on the limitations of overhang and island in local deposition technique -fused deposition modelling (FDM). When FDM is employed, the solid object is created layer by layer, starting from the lowest one. As a consequence, each layer can only be deposited on top of an existing surface or it can protrude over the lower layer within a certain limit, otherwise the print material falls and solidifies out of the geometry.

In case the overhang exceeds the hardware limit there are three possible solutions to the problem: adding a support structure underneath the non-printable areas, segmenting geometry in printable parts, or changing the geometry in order to avoid such areas. The topic of generation and optimisation of additional support is already rich in algorithms [11, 29], as well as geometry segmentation [20, 28, 31]. Note that the support structures are meant to be removed manually at the end of the process, thus they represent wasted material, energy, as well as time employed for their construction and removal. Segmented geometries also need additional time to be assembled after printing, increasing the fabrication cost. Therefore this approach focuses on making changes onto the design geometry.

1.2. Objectives and Contributions

While there is considerable research in the field regarding adding and optimising support and segmentation, the generation of a geometry that follows printability aspects, is a research gap with very little documentation. The primary objective here is the integration of fabrication-informed means for geometry modification into an early design work-flow.

This paper explores the design-related aspects that have hitherto been side-lined and aims to assist designers to control geometry from an early stage and to create fabrication aware geometries for 3DP. It is focussed on geometries represented using implicit representation; field-based design is deployed, as it is a good way to generate and analyse geometries for 3DP [3, 6, 27]. The shape modification process aims to fulfil three objectives: minimise overhang, create a stable geometry and reach a solution as close to the input geometry as possible. As a contribution, the form-finding tool (Grasshopper plug-in) is proposed along with the underlying computational framework. Implicit modelling techniques are adapted to assist designers to pre-rationalise their geometries in a way so that they avoid printing inefficiencies and difficulties.

2. BACKGROUND

2.1. Design Paradigm Aspects

A common method to print strong overhangs or unstable shapes is to print an additional support structure at the same time. While in small-scale 3D prints this could be an efficient solution, there are cases like large-scale concrete printing where it cannot apply. In these cases, designers need to incorporate those constraints in their design process. Bhooshan, S. et al. [6] are using conservative layer height and overhangs at their design to ensure the shapes are guaranteed to be stable during printing. Efthimiou E. and Grasser G. [14] integrated an optimisation step within the morphogenetic scheme of their design, so that the object never exceeds a maximum overhang angle value at any single point. Also, they implemented an additional force on their geometry as it was growing, for any centre of mass correction necessary. Drawing inspiration from such cases, the author aims to develop a tool that automatically changes the design in order to keep the need for additional support to the minimum.

2.2. Computational Paradigm Aspects

As 3D printing has become a popular manufacturing technique, a lot of people are engaged with the development of tools to handle its constraints.

Overhang
There is continuing investigation into the overhang problem in 3DP. In most cases a support structure fills the free space encountered when a part is projected downward in its build orientation. There are multiple methods that vary in how much material is used and the type of material which can be used [11]. Commercially available software widely used to generate support for meshes includes Ultimaker Cura and Autodesk Meshmixer, whilst there is no application to automatically fix geometry so that additional support structure is not needed.

Cacace et al. [11] propose a level set based method to create object-dependent support structures. In most cases, they manage to reduce both the amount of additional material and the printing time. In particular, they suggest a way to avoid the creation of pillars which touch the build plate by means of optimally shaped chamfers, suitably placed below hanging parts. This method could be deployed to create a new shape, close to the initial one, which does not require additional support material and time to be cleaned.

Balance
The use of the support structure is not just for overhangs, it also keeps the whole model from falling. Consequently, another issue that has to be addressed is the overall stability of the geometry, both as a final object and during the printing procedure. Therefore, an issue that arises is how to change the geometry such that it stands during printing.

Prévost, R. et al. [26] suggest a method to deform a given volume in order to make it balance. They formulate geometry optimisation as an energy minimisation problem through an interactive editing process, aiming to move the centre of mass above the base of the model. They manipulate the volume in two ways: by deforming the original surface, and carving the model interior to create inner voids in order to minimise the deviation from the original shape. The developed tool assists designers to create objects that stand after printing and support removal, but they do not examine stability at each layer.

Interaction and Different Solutions
Optimisation algorithms are widely used in architecture and engineering to maximise the performance of a structure and to minimise the weight or the cost of it, while at the same time they can produce innovative designs. Although multi-objective optimisation methods can produce multiple solutions by applying different weights on objective responses, existing single-objective optimisation algorithms usually produce a single optimal solution. However, most designers are looking for multiple possible solutions to choose from, as ''optimal design'' may deviate significantly from their desired design or functional requirements.

Yang, K. et al. [32] present five strategies for achieving such diverse and competitive structural designs that can be easily implemented in different topology optimisation techniques. With these strategies, the designers can make their own choice based on the multiple solutions, whilst gaining more control in the form-finding process. Ameba Grasshopper plug-in [30] incorporates the aforementioned ways for different solutions. According to the design requirements, the user can apply various loading and boundary conditions to the design domain, which evolves into multiple shapes.

This research aims to create an interactive application that offers ways to produce different solutions for each criterion, overhang and balance, both individually as well as for the combination of the two.

2.3. Implicit Representation

Implicit representation is finding renewed interest in the design domain, with the development of new modelling applications and researches.

In implicit representations, space is described as a scalar field and shapes are as the result of a function -function representation (FRep)- that converts any point coordinate into a real

value, $R3 \rightarrow R$. For instance many geometric primitives (sphere, torus, cone, cylinder, box, etc.) can be described as a function $v = f(x, y, z)$. All points assigned with a negative value (v) lie inside the object, while those with positive value lie outside of it. All the point locations where this function evaluates to $v = 0$ form the boundary of the shape. So, technically, the surface of the object is the zero iso-surface of the scalar field. [9]

Advantages of implicitly represented shapes include: efficient check for whether a given point is inside or outside the object, easy-to-make boolean operations and topological changes, as well as efficient blending between shapes, etc. [6, 3]. Therefore, this representation scheme is increasingly being adopted for shape design in 3DP.

Recently developed 3D modelling applications based on implicit representations include Axolotl [5], Monolith [22], Ameba [16] and nTop [27]. In addition, FRep has been deployed in research on 3DP. For instance, Bhooshan, S. et al. [6] use FRep to design shapes for robotic concrete 3DP. Bernhard, M. et al. [4] also suggest FRep as an alternative representation that allows building elements to be described throughout their entire volume. They present applications of volumetric modelling in the context of 3DP for architecture.

Implicit representation has recently been implemented in the context of form-generation. While design tools have been developed, to the best of the author's knowledge, a tool that optimises implicit geometry for overhang and balance by modifying its 3D scalar field is not commercially available to this day. Therefore, existing knowledge is used to develop an application for function-represented shape processing for 3DP as a form-finding tool.

2.4. Relevant Prior Work
Douroudi and Tsomokou [13] developed a plug-in for Rhino Grasshopper, for volumetric modelling, based on FRep. The plug-in includes tools to make shapes 3D printable, like segmentation, thickness and overhang filter. In detail, overhang filter subtracts the unsupported voxels of the geometry or adds voxels below them, according to the angle with the nearest neighbour below. The limitation of this method is that the geometry changes a lot.

3. COMPUTATIONAL FRAMEWORK
As already mentioned, the aim of this approach is to explore different methods to modify an implicit geometry in order to become 3D printable without any additional support structure. This section examines how to evaluate and how to change geometry such that it meets the three different criteria: minimisation of the support volume required, distance between the centre of mass and the geometry's base and deviation of the new shape compared to the input.

3.1. Implicit Measures
This research is focused on implicit geometries, created by assigning values in a continuous 3D scalar field according to mathematical functions. This field is temporarily discretised into voxels in order to become computationally manageable.

The measures of overhang, balance and deviation are based on these voxels and their values. The computational methods update the scalar field, such that the iso-surface at level 0 represents a 3D printable geometry.

Overhang
The overhang measure values, measured at the boundary voxels, are categorised into three subsets, on the basis of their printability. To this end, for each voxel the angle (θ) between the gravity vector and the gradient vector (always perpendicular to the iso-surface, as it equals to the normal of the surface at that point), is measured as shown in Figure 2. Moreover, a voxel is considered as:

- unsupported, if $\theta \in [0, \alpha)$,
- modifiable, if $\theta \in [\alpha, \pi/2)$,
- safe, if $\theta \in [\pi/2, \pi]$,

where α is a given limit angle. So, unsupported voxels represent areas that cannot be printed, opposite to the safe points. Modifiable voxels are printable since the overhang is small, but they could be modified in order to provide support to the unsupported ones. The overhang solvers presented below aim to reduce unsupported voxels.
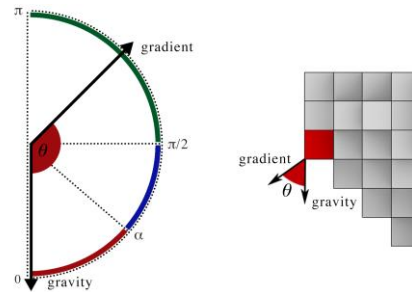

Figure 2. Left: Unprintable (red), modifiable (blue), safe (green) points with respect to the smallest angle between normal and gradient. Right: 2D diagram of overhang measure on voxels.

Balance
The geometry is considered stable during 3DP without external support if it is balanced at every step. Hence, balance is measured for each layer of the geometry in z direction separately. A layer is said to be stable when its centroid is located above the convex hull of the shape's base. For instance in Figure 3 the red circle with centre C1 is unstable while the blue one with centre C2 inside the convex hull is stable. Furthermore balance can be measured as the distance (d) of the centroid to the convex hull curve, which is considered as negative if it is inside and positive if it is outside (Figure 3).
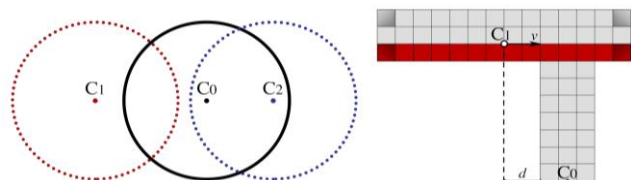

Figure 3. Left: Base convex hull (black), unstable(red) and stable (blue) layers. Right: 2D diagram of balance measure on voxels.

Deviation

The modification methods aim to find a new, fabrication aware geometry as close as possible to the input. As a consequence the deviation has to be measured. One way suggested is boolean operations between the input and result geometry. In this way the number of different voxels increase deviation while the voxels that intersect decrease it (Figure 4). Also, the difference between voxel values is another way to measure if the geometry has changed a lot. The higher the difference of values, higher is the change in geometry.
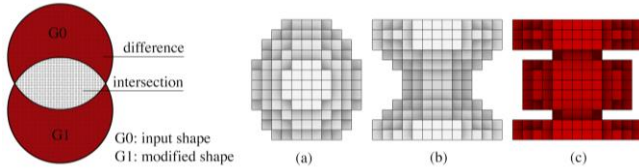
Figure 4. Left: Deviation measure. Right: Voxel representation of two input geometries (a, b) and deviation between them (c).

3.2. Overhang Reduction

Printability can be ensured by the absence of strong overhangs. Two methods deployed to achieve that by minimising the number of unsupported voxels of a given geometry, as described in 3.1, are the Gradient Descent Method (GD) and the Level Set Method (LSM).

Gradient Descent Method for Overhang Reduction

A GD algorithm is deployed to change the geometry in order to fix overhang problems. The algorithm aims to minimise the number of unsupported voxels by changing scalar field values and, as a consequence, the field's gradient.
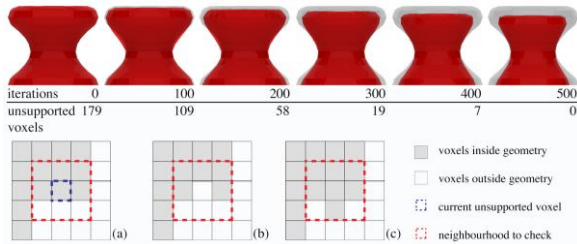
Figure 6. Above: Overhang elimination using GD. Below: 2D diagram of the changes that GD makes in the field. Input field (a), modified field in case of increasing the value of the unsupported voxel (b) or decreasing the value of the voxel below(c).

The process of GD begins by defining an input geometry in scalar field format and the maximum overhang angle allowed. The unsupported voxels are calculated using the field's gradient and the input angle. Next, a learning rate factor (λ) is defined along with the number of iterations.

The algorithm navigates through the unsupported voxels at every iteration and searches the best change between three options:

- increase value by adding λ to delete unsupported voxel,
- decrease value by subtracting λ to a voxel below to add it,
- or do not make any change.

The evaluation of each option is calculated locally. The cost of each one equals to the number of unsupported points in the neighborhood of the current voxel after each change. The choice is made based on that cost. In detail, if the cost is the same at each option, no changes occur. Else, the option with the less unsupported voxels is chosen, while deleting is preferred. Voxels are stored in lists according to the best option and field updates once, after all unsupported voxels are classified. (Figure 6)
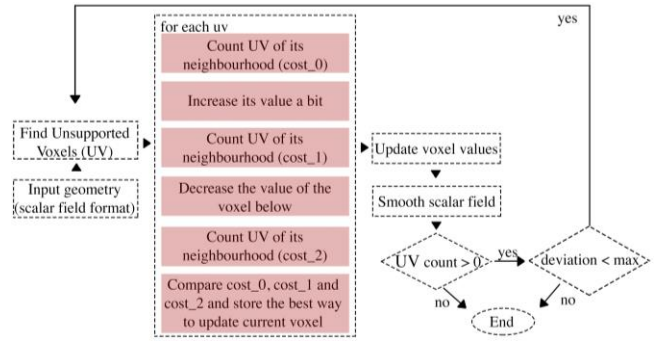
Figure 7. Gradient Descent method to reduce overhang workflow.

After each field update, discrete laplacian smooth is carried out as it was noted the gradient does not align with the geometry. It is important to have this alignment for the next iteration since gradient is used for the calculation of the cost. This process is repeated until the output geometry has no unsupported voxels or the deviation has reached the maximum limit set by the user. The algorithm is also described as a pseudocode in Figure 7.

Level Set Method for Overhang Reduction

A different way to achieve printability, in terms of overhang, is to use the LSM to perturb geometry [11]. Here, how the values of the modifiable voxels in the input field can change so that these voxels move towards the gradient direction (perpendicular to the surface) is examined, aiming to provide support to the unsupported voxels (Figure 8).
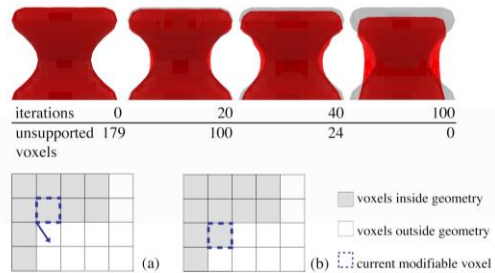
Figure 8. Above: Shape processing deploying LSM to minimise overhang. Below: 2D diagram of the changes of modifiable points using LSM. Input field (a) and modified field (b).

The inputs to the algorithm are again, a geometry in scalar field format, the maximum overhang angle, the number of iterations and the learning rate (λ). In this case modifiable voxels of the geometry are found based on the field and the

angle. For each modifiable voxel a moving vector is calculated, which represents velocity (v), to update its position using the linear movement equation:

$$x = x0 + v\Delta t,$$

where x is the new position, $x0$ the starting position of the voxel and Δt the time between each iteration.

The direction of the velocity is the same as the gradient of the field at that point. The speed depends on λ, which controls the deviation in every step and it is denoted by the user. In addition, increasing the speed of lower points with respect to higher ones, allows them to offer support to voxels above saving material and preventing unprintable regions from reaching the built plate "as it is". For this reason, the term ($zmax - z$) is introduced, where $zmax$ is the height of the object and z the height of the current modifiable point. Furthermore, the speed is set to be higher when θ, the angle between the gradient and the gravity vector, is close to 0 and the voxel tends to be unsupported. Consequently, for $\Delta t = 1$, the moving equation used to move values is updated to:

$$x = x0 + \lambda(zmax - z)/\theta$$

A small learning rate makes the transaction smoother, although it may increase the number of iterations needed to reach an optimal solution. This process is also described in the workflow diagram below (Figure 9).
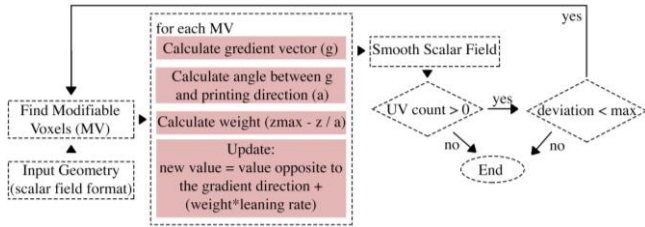
Figure 9. Level set method workflow.

3.3. Balance Improvement
The objective in terms of balance is formed as a layer by layer deformation of the geometry. In order to ensure that it remains stable during printing, each layer at z direction moves towards a stable position (Figure 10b). At first the centroid of each layer is computed to check whether it is stable, according to the measure described above. Then, in case the layer is not stable, a moving vector is calculated to move the voxel values of that layer. The moving vector starts from the current centroid and points to the base projected at the same height, while its magnitude is defined by the user and represents the distance that the layer moves at every iteration. (Figure 11)

An additional option given to the user is to enlarge the base. In this case, the base gets wider at every iteration; consequently, the rest of the layers change less and balance can be reached in fewer iterations. This way, an alternative solution can be found (Figure 10c). Another solution can be reached by moving stable layers towards the centre of the mass of the whole shape at the same time. This may increase deviation

but can also decrease the number of iterations needed to reach balance equilibrium.
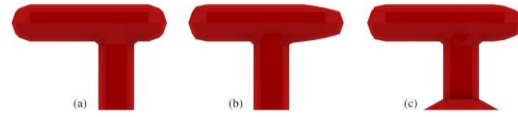
Figure 10. Input geometry (a), balanced geometry with moved layers (b) and enlarged base (c).
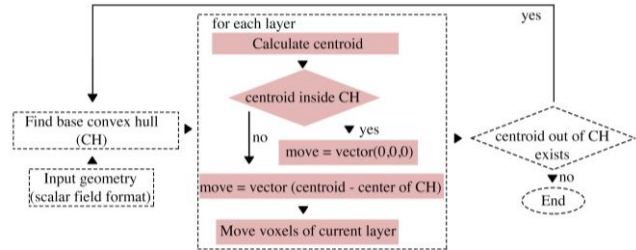
Figure 11. Balance improvement workflow.

3.4. Solver
In order to ensure that a 3D geometry can be printed without additional support structure, two parameters have to be considered: overhang and balance. Modification methods based on those parameters were examined separately, but also a solver which combines the two was developed. Therefore, the shape processing for 3DP is approached with an iterative process that alternates between overhang and balance criteria. Every iteration has two phases. In the first phase, GD is deployed to reduce unsupported voxels. In the second phase, geometry is deformed a bit to improve stability. Finally, deviation should be kept as small as possible and it is recommended to use a small learning rate both for balance and for overhang targeted modification. (Figure 12)
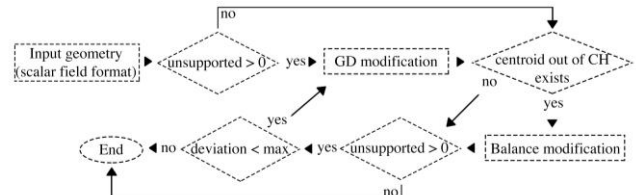

Figure 12. Combinatorial method workflow.

4.   RESULTS
The tools developed managed, as anticipated, to reduce additional support and guarantee balance during printing for several experimental designs. In most cases the time and material need for printing were also minimised too. Shapes with increasing number of complexity and their modified outcomes are presented, discussed and demonstrated below (Figures 13-17). These shapes were generated by the application described in 2.4, the modification occurred in their scalar field and iso-surface contouring was used for the representation. As a proof of concept, 10cm height models of them were digitally tested with Ultimaker Cura v4.2.1 software. Cura creates fabrication simulation and calculates the estimating time and material needed for 3DP. Time and material are the parameters being compared between the input

geometries and the modified outcome. For the following tests settings for generic PLA 0.4mm deposition with Ultimaker 2+ and the default Fine Profile were used. Then, physical models were printed with the use of the aforementioned settings, material and printer.

Digital Tests

Figure 13 illustrates two simple geometries (a), with negative (A) and positive (B) curvature. Also the outcomes of modification towards overhang minimisation deploying gradient descend (b) and level set method (c) are exhibited. In these first case, the tool developed achieved to modify geometry so that there is no need for additional support structure, which was proved by Cura digital tests. The input shape needed 3 hours and 55 minutes to be printed, with 46% of the time being required for the support. Also, the total of required material for the print was 31 grams, from which 8 grams consisted support. The output geometries do not need any support and they require less material and time to be printed, as shown in the Table. In detail, by using the GD to modify the shape, the fabrication time was reduced to 1 hour and 44 minutes and the required material to 19 grams. The LSM achieved to reduce time to 1 hour and 52 minutes and the required material to 21 grams. Respectively, both time and material were reduced for the second shape (Table). Those shapes were symmetrical and stable, so they did not need balance modifications.
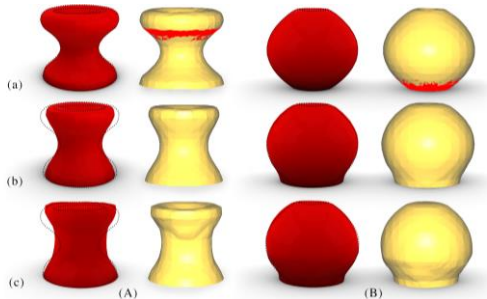
Figure 13. Input geometries(a) with negative(A) and positive(B) curvature, optimised geometry using GD (b) and LSM (c). 3D models(red) and diagrams by Cura software (yellow).
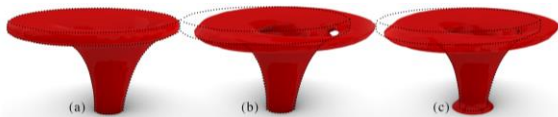
Figure 14. Input geometry (a), balanced geometry generated by moving unstable layers (b) and widening the base while moving unstable layers (c).

Next, another shape was tested, which required both overhang and balance modifications. Figure 14 exhibits two options for balance optimisation: solely moving unstable layers (b) or also widening the base at the same time (c). Based on the deviation criteria the second option is better in this case, with 479 different from the input voxels instead of the 594 of the first option. However, the designer has the opportunity to choose between the two.

That shape was also modified towards overhang minimisation (Figure 15). It was inserted to Cura for digital testing and the initial estimated support required 50 out of 81 grams and 77% of the printing time. The developed tools achieved to deform the geometry (b,c) so that less support was required. In both cases, total material and printing time were reduced, too. The combinatorial tool also managed to modify the geometry so that it demands less support and it is balanced during printing (d).(Table)
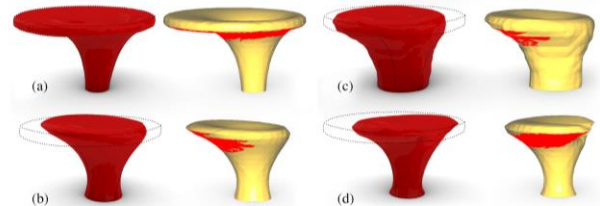
Figure 15. Geometry, Cura diagram where red is the unsupported areas of input (a), modified shape using GD (b), LSM (c) and the solver (d).
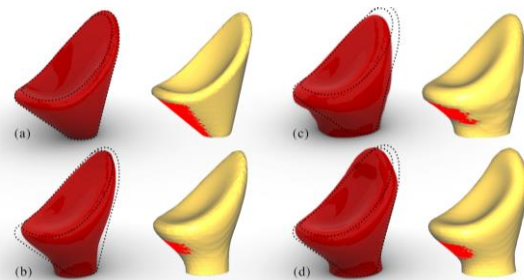

Figure 16. Chair geometry, Cura diagram where red is the unsupported areas of input (a), modified shape using GD(b), LSM (c) and the solver (d).
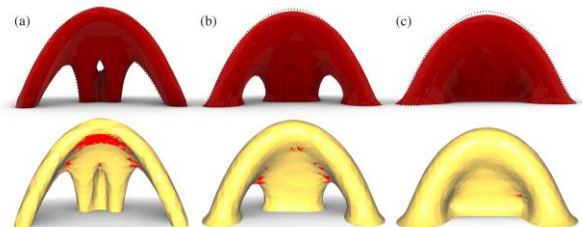

Figure 17. Geometry with multiple touch points: Input (a), overhang reduction by LSM and 50 (b) or 100 (c) iterations.

In addition, a chair design was modified using the methods developed and several iterations were generated (Figure 16). The developed tools achieved to reduce the support required compared to that of the input shape (a) in this case as well. However, as shown in Table, only the GD method achieves to minimise both total time and material (b). The LSM (c) and the solver (d) enlarged the object, so that it needed less support, but both total material and time needed for printing increased.

Furthermore, the suggested tools were tested on another shape, with multiple touch points (Figure 17). In this example both GD and LSM had the same result, the input geometry (a) was enlarged and the overhang was reduced. The LSM

presented in Figure 17 illustrates the outcome of 50 iterations (b), where the support has been reduced from 7 to 3 grams. After 50 more iterations (c) the estimated support becomes 2 grams. However, the output shape deviates a lot from the input one, and significant topological changes are observed, which might reduce functionality.
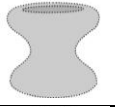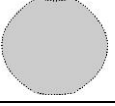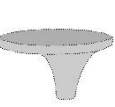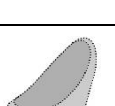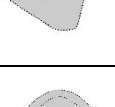
object		material	time	support
	Input	31g	3h 55min	46%
	GD	19g	1h 44min	-
	LSM	21g	1h 51min	-
	Input	40g	4h 47min	34%
	GD	38g	3h 53min	-
	LSM	26g	2h 37min	-
	Input	81g	13h 55min	77%
	GD	38g	5h 46min	65%
	LSM	45g	6h 12min	56%
	Solver	40g	6h 18min	68%
	Input	11g	2h 27min	14%
	GD	10g	1h 53min	8%
	LSM	14g	2h 38min	8%
	Solver	12g	2h 13min	8%
	Input	31g	6h 11min	21%
	GD	35g	6h 27min	10%
	LSM	42g	7h 19min	3%

Table. Estimated time and material for 3D printing.

Carrying out these digital tests, differences between the two methods for overhang reduction were observed. The GD method was slow compared to the LSM and needed more iterations to reach the same amount of unsupported voxels. However, the deviation was smaller and smoother. The LSM, although fast, was usually enlarging the object a lot and the modifications were not smooth. In some cases an additional smooth function was required, due to the voxel operations that did not allow very small steps.

Figure 18. 3D printed models. From right to left: Input shape, GD, LSM and Solver modified shapes.

Physical Tests
Physical tests were carried out using Ultimaker 2+ and PLA material, with the same settings as the digital tests, in order to confirm that the generated shapes can be fabricated. Shapes of Figure 13 were printed successfully without support. Shapes of Figure 18 corresponding to the shapes of Figure 16 were printed with the support generated by Cura software (Figure 18). These objects include the initial design, the outcome of the GD and LSM algorithms for overhang reduction, as well as the model modified by the combinatorial

solver. The alternative solutions provided by the developed toolset can be observed here.

5. FUTURE WORKS
An easy-to-use tool, for implicitly represented shape processing, able to enhance creativity and assist users to design 3D printable objects has been developed. Its limitations along with suggestions for further exploration are presented below.

Gradient Vectors
The proposed methods require the input geometries to be on scalar field format and processes are based on the gradient vectors of it. As a consequence, the gradient of the field needs to be updated after every modification. The difficulty of gradient update is partially overcome by the use of discrete laplacian smooth. This way, however, the geometry shrinks. Currently, this limitation is being addressed by manual scaling, but an additional function to prevent geometry from shrinking would be useful.

Contour based calculations
A continuous scalar field, temporary discretised into voxels, was used to control implicit geometries. Voxel-based operations to measure and update the field are efficient and fast, but sometimes transactions are not that smooth and measures not that accurate. A different way to compute measures and change implicit geometries could be through their iso-contours. Some experimentation has already been done by the author, but it was deemed computationally expensive. It is suggested as a probably more efficient way to be explored in future.

Expansion to robotic fabrication and several materials
Overall, this paper establishes a general approach to achieving diverse and competitive 3D printable forms, which holds great potential for practical applications in architecture and engineering. The proposed method follows a modelling paradigm where material and hardware constraints are abstracted into geometric constraints. Consequently, it can apply to large-scale robotic printing and/or different materials.

6. CONCLUSION
Computational methods capable of producing geometries, aware of the 3DP geometrical constraints of overhang and balance were demonstrated in this paper. Using the developed toolset, a designer is able to control shape from an early stage of the design workflow developing fabrication-aware-shapes, amenable to contemporary additive manufacturing technologies. In conclusion, the suggested methods for implicitly represented shape processing are able to enhance creativity and innovation, as well as to assist users to design 3D printable objects across scales and disciplines.

REFERENCES

1. AiBuild (2016) Making of Daedalus Pavilion, https://www.youtube.com/watch? v=rAbB_AZvCT4 (Accessed 20 August 2019).

2. Assa, A. (n.d.) 3D printed Titanium Pen collection, https://assastudio.com/work/3d-printed- titanium-pen-collection/ (Accessed 20 August 2019).

3. Attene,M. et al.(2018) Design, Representations, and Processing for Additive Manufacturing, Synthesis Lectures on Visual Computing Computer Graphics, Animation, Computational Photography and Imaging, Barsky, B. A. (ed.), University of California, Berkeley, Morgan and Claypool Publishes.

4. Bernhard, M., Hansmeyer, M., & Dillenburger, B., Volumetric modelling for 3D printed architecture, L. Hesselgren et al. (eds.): AAG 2018, pp. 392–415.

5. Bernhard, M. (n.d.) Axolotl, http://dbt.arch.ethz.ch/project/axolotl/ (Accessed 20 August 2019).

6. Bhooshan, S., Ladinig, J., Van Mele, T. & Block, P., Function Representation for Robotic 3D Printed Concrete, J. Willmann et al.(eds.): ROBARCH 2018, pp. 98–109.

7. Bhooshan, S. et al. (2018) Equilibrium-Aware Shape Design for Concrete Printing, De Rycke K. et al.(eds.), Humanizing Digital Reality. Springer, Singapore.

8. Bhooshan, V., Louth, H.D., Bhooshan, S. and Schumacher, P. (2018) Design Workflow for additive manufacturing: a comparative study, Int. J. Rapid Manufacturing, Vol. 7, Nos. 2/3, pp.240–276.

9. Bloomenthal, J. (1997) The Geometry of Implicit Surfaces. Introduction to Implicit Surfaces, 3–51.

10. Buswell, R.A., Soar, R., Gibb, A.G. and Thorpe, A. (2007) Freeform construction: mega-scale rapid manufacturing for construction, Automation in construction, Vol. 16, No. 2, pp.224–231.

11. Cacace, S., Cristiani, E. & Rocchi, L. (2017) A Level Set Based Method for Fixing Overhangs in 3D Printing, Appl. Math. Model. 44, pp. 446-455.

12. Chua, C., Teh, S. & Gay, R. (1999) Rapid prototyping versus virtual prototyping in product design and manufacturing, Int J Adv Manuf Technol 15: 597–603.

13. Douroudi, E. & Tsomokou, E. (2019) FRep for 3D printing, MSc Architectural Computation Studio, Bartlett School of Architecture, London.

14. Efthimiou E. & Grasser G., Liquid rock – Agent based modeling for concrete printing, L. Hesselgren, et al. (eds.): AAG 2018, pp. 236–255, Klein Publishing GmbH.

15. Ervinck, N. (2016) Wolfkiam, http://nick-ervinck.com/en/works/detail-2/wolfkiam (Accessed 20 August 2019).

16. Han, Y.S., Xu, B., Zhao,L. & Xie, Y. M. (2018) Topology optimization of continuum structures under hybrid additive-subtractive manufacturing constraints

17. Hansmeyer, M. & Dillenburger, B. (2013) Digital Grotesque, http://digital-grotesque.com (Accessed 20 August 2019).

18. Herpen, I.V. (2011) Capriole, http://www.irisvanherpen.com/ (Accessed 20 August, 2019).

19. Lefebvre, S., Turning 3D models into 3D prints, SGP2018 Graduate School, Paris, France.

20. Luo, L., Baran, I., Rusinkiewicz, S., & Matusik, W. (2012) Chopper: Partitioning Models into 3D-Printable Parts, ACM Transactions on Graphics, 31(6), 1.

21. Materials (2014) Chinese Company 3D Prints 10 Recycled Concrete Houses in 24 Hours, http://www.designboom.com/technology/3d-printed-houses-in-24-hours-04-24-2014/ (Accessed 28 August 2019).

22. Michalatos, P. & Andrew O. P. , Working with Multi-Scale Material Distributions, Acadia 2013: Adaptive Architecture, 43–50.

23. nTopology (2019) nTop Platform, https://ntopology.com/ntop-platform/ (Accessed 28 August 2019).

24. Peleg, D.(n.d.) The birth of Venus, https://danit-peleg.com/the-birth-of-venus/ (Accessed 20 August, 2019).

25. Perez, B. (2019) nTopology Platform for Additive Manufacturing, https://ntopology.com/ blog/2019/03/21/implicit-modeling-for-additive-manufacturing/ (Accessed 28 August 2019).

26. Prévost, R., Whiting, E., Lefebvre, S. & Sorkine-Hornung, O. (2013) Make It Stand: Balancing Shapes for 3D Fabrication, ACM Trans. Graph. 32, 4, Article 81.

27. Reitz, A. (2019) Computational Modeling with nTop Platform, https://ntopology.com/blog/ 2019/04/18/computational-modeling-with-ntop-platform/ (Accessed 28 August 2019).

28. Shamir, A. (2016) Computational Tools for 3D Printing Part 4 : Modeling and Analysis Modeling 3D Objects.

29. Strano, G., Hao, L., Everson, R. M. & Evans, K. E. (2013) A new approach to the design and optimisation of support structures in additive manufacturing, Int J Adv Manuf Technol 66:1247–1254.

30. Xia, L., Xia, Q., Huang, X., & Xie, Y. M. (2018) Bi-directional Evolutionary Structural Optimization on Advanced Structures and Materials, Archives of Computational Methods in Engineering, 25(2), 437–478.

31. Wu, K. (2017) A Review on Mesh Segmentation Techniques.

32. Yang, K. et al. (2019) Simple and effective strategies for achieving diverse and competitive structural designs. Extreme Mechanics Letters, 100481.

Bridging the Gap Between Traditional Japanese Fabrication and Advanced Digital Tools

Mohammed Makki[3], Masaaki Matsuoka[2], Ana Ilic[1], Lorenzo Franceschini[1], Jorge Beneitez[1]

[1]Architectural Association
London, United Kingdom
franceschini@aaschool.ac.uk

[2]Takenaka Corporation
Osaka, Japan
matsuoka.masaakia@takenaka.co.jp

[3]University of Technology Sydney
Sydney Ausralia
mohammed.makki@uts.edu.au

ABSTRACT

Traditional Japanese woodworking techniques have been inherited between generations in the past millennia through the Daiku, the master carpenter that teaches apprentices the precise and valuable skills of carpentry through methodical and time-intensive processes. Throughout the 21st Century, with the development of advanced construction methods, coupled with younger generations exhibiting little interest in following the Daiku, this valuable cultural artform is becoming less prevalent in Japanese culture. However, the development of advanced digital tools offers an avenue through which the knowledge and skills of older generations can be both transferred and developed by younger generations. In this context, the paper examines the relationship between traditional Japanese woodworking and advanced computational tools by bringing experts from both disciplines for the design and construction of a Japanese Pagoda, in which the significance of bridging the gap between both domains is highlighted across the design, fabrication and assembly of the project.

Author Keywords

Computation; Japanese Woodworking; Daiku; Evolutionary Computation; Pagoda

ACM Classification Keywords

I.6 I.2.8 Problem Solving, Control Methods, and Search (Heuristic methods); J.5 Arts and Humanities (Architecture); J.6 Computer-aided design;

1 INTRODUCTION

The historical development and evolution of traditional Japanese woodworking techniques spans many centuries. Primarily rooted in Japan's religious and geographic locale, the impact of climatic and environmental conditions on the territory's flora, and the deep respect of the Shinto belief of treating wood as a living organism, have played a primary and significant role in the use of wood (and the development of the intricate methods associated with handling the material) in Japanese culture [5]. Ingrained in this is the transfer of knowledge between generations, primarily through a master-apprentice relationship, in which not only were the methods and techniques taught (such as selection, cultivation, distribution, treatment, design and assembly), but also the respect and moral bond between the carpenter

and the tree. The importance of the *Daiku* – master carpenter – went beyond simple construction, acknowledging and contributing to mental and emotional aspects of the trade, such as leadership, compassion and spiritual preparation [1]. Following the decline of the Edo period in which Western methods were actively pursued in Japanese construction, traditional methods remained heavily prevalent in ensuring the techniques and skills developed over centuries remained pivotal in Japanese woodworking approaches; more importantly, in ensuring the inheritance of this knowledge by future generations did not stop [10] (Figure 1).

Figure 1. The Horyu-Ji Pagoda is one of the oldest wooden structures in existence [9] and serves as a prime example of the woodworking knowledge inherited by Japanese master carpenters.

In more recent times, the transfer of this knowledge has showed signs of deterioration [4]. Attributed to a range of different causes, such as there being fewer *Miyadaiku* – carpenters who handle traditional Japanese structures through restoration by means of disassembly and re-assembly, less interest by younger generations in learning the trade, the development of advanced carpentry tools that have replaced traditional techniques and the use of digital tools for the purpose of expedited design and construction methods.

While traditional knowledge continues to be used for shrines, temples, teahouses, and Japanese-style homes, they have ceased to become mainstream architectural techniques, and so the passing down and preservation of these methods has become at risk.

Advanced digital technologies developed in the last decade offer respite and an avenue for the preservation of traditional Japanese woodworking techniques. While the Daiku is primary in the transfer of knowledge, the use of advanced digital tools across a range of domains and disciplines allows for a streamlined sharing of knowledge through a feedback mechanism that is both relatable to younger generations as well as in line with the technological advancements of the 21st century, allowing greater opportunities to explore new formats and frameworks that effectively combine traditional techniques with contemporary methods.

In this context, the presented research examines the relationship between old and new by bringing experts in traditional Japanese woodworking techniques and experts in advanced digital technologies for the design and fabrication of a wooden Japanese Pagoda at the Takenaka Carpentry Tools Museum in Kobe, Japan. Additionally, the project is conducted with 28 participants from 16 different countries, in which the transfer of knowledge (both traditional and digital) played a primary role in ensuring the research examined the above relationship through varying skillsets and backgrounds.

2   BACKGROUND

The future of traditional Japanese woodworking techniques is at a crossroads. At one end, there is the preservation of "cultural heritage for the benefit of future generations", while on the other, the necessary acceptance and adaptation "to the inevitable future change" brought on by technological advancements [6]. In either case, the declining number of Daiku, coupled with a generation that is less interested in inheriting the techniques of previous generations, highlights the significance of bridging the gap between advanced digital design tools, which are in continuous evolution and change, and the artistry of traditional woodworking techniques. Most importantly however is the importance of demonstrating that the former does not diminish the role of the latter, as well as highlight that the latter is easily accessible through the former.

The project's aim is to design a traditional Japanese Pagoda through advanced digital tools; and utilise the expertise of a Daiku - Master Carpenter Akinori Abo - for the refinement of the design as well as the fabrication and construction of the final product. Most importantly, the project aimed to avoid linearity in its development, ensuring a feedback loop is developed between the application of digital tools for the design's development and the intricate fabrication processes and techniques inherent to Japanese woodworking. Through the use of Rhino 3D and Grasshopper 3D, in which structural analysis tools through Karamba [7], environmental analysis tools through Ladybug [8] and multi-objective evolutionary

optimisation tools through Wallacei [3], the design of the Pagoda aimed to preserve its spiritual and historical significance and ensure it allowed for the application of the unique woodworking skills and techniques of the Daiku.

3   DESIGN ITERATIONS

Multiple design options were developed for the Pagoda, where several key characteristics played a significant role in the design process. First, the global morphological traits of a traditional Pagoda had to be respected, i.e. repetitive patterns, multiple levels and formal hierarchy. Second, the spirituality associated with the Pagoda is represented through its form. Third, wood is the primary and only material used in the structure's construction, more precisely, *Sugi* (Japanese Cedar) and Hinoki (Japanese Cypress) (Figure 2). Fourth, no fastening tools are used (i.e. screws, nails, etc..). Finally, to ensure an intricate and attentive use of joints between the wooden elements is considered; in which *Shiguchi* (joints between two members of different direction and function) and/or *Tsugite* (joints between two members in the same direction) [11] construction methods were utilised.

Figure 2. Sugi and Hinoki (Japanese Cedar/Cypress) is native to Japan and known for the wood's strength and durability. The wood used for the Pagoda was sourced from Mount Yoshino.

In this context, participants developed 5 different schemes that aimed to incorporate the above characteristics through a digitally driven design process. Utilising different approaches, the 5 schemes presented varying interpretations of the Japanese Pagoda, and the use of varying digital tools and their impact on the design of the Pagoda. Some of the schemes adopted a multi-objective evolutionary approach to their designs, where the pagoda was optimised for multiple conflicting objectives through the use of the evolutionary engine *Wallacei* (running the NSGA-2 multi-objective evolutionary algorithm developed by Deb et. al. [2]). In doing so, these schemes generated a population of different solutions with varying morphological characteristics driven by design objectives identified to be key in the traditional woodworking techniques used in a traditional Pagoda (Figure 3). Such objectives included the complexity of joints, the quantity of joints, the viewpoints of the different levels of the Pagoda, and the cross sections of the wood used in the Pagoda's construction. Through this process, the evolutionary model would derive the design objectives and morphological variables running the algorithm from the Daiku's expertise, allowing for the digital tools to be used within the framework of the traditional knowledge of the master carpenter.
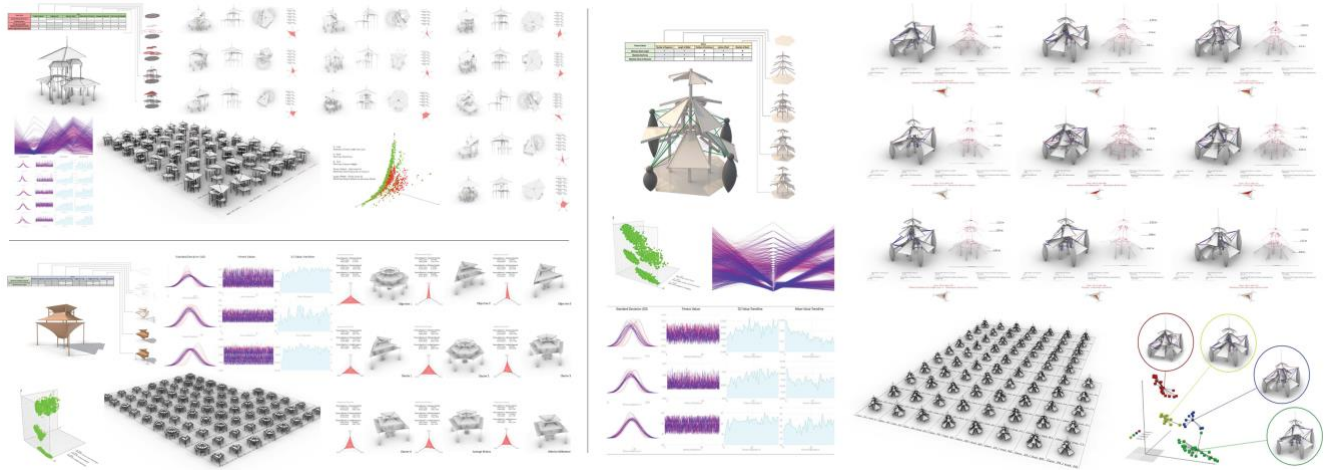
Figure 3. The application of multi-objective evolutionary optimization methods for the design of different Pagoda options.

In presenting the final designs to the Daiku, highlighting each scheme's development and underlying concepts, and the methods behind the application of the advanced digital tools for the design's development (Figure 4), one scheme was chosen by the Daiku as best representing a Pagoda that is both formally and structurally intriguing. In doing so, the design of the selected Pagoda was further refined in preparation for fabrication and construction.

Figure 4. The different schemes developed of the Pagoda (the selected scheme is presented in the following section).

4 THE RECIPROCAL PAGODA

The selected scheme was a 3 level Pagoda structure comprised from a reciprocal pentagonal base through which each level recedes inwards (a feature found in traditional Japanese Pagodas). The structure's boundary sat within a 3m x 3m x 4m bounding box; in which the structural frame - comprised from columns, rafters and tension/compression rings - supported the linear roof slats located on each of the Pagoda's three levels.

4.1 Design Development

The primary aim of the design was to signify the central column of a Pagoda, one that holds significant spiritual significance for Japanese culture. However, rather than incorporating the column within the structure, it is rather representative as a continuous and uninterrupted void throughout the centre of the Pagoda. In doing so, the design is driven by 6 'rings' connected through 3 primary sets of rafters. The 3 larger and outer rings act in tension and join each set of rafters to one another; the 3 smaller and inner rings act in compression and join the centres of each set of rafters. Finally, the base columns lift the structure above ground level, in line with the design of traditional Pagoda structures.

The design of the roof followed the reciprocal morphology of the primary structure; where the objective was for the structure to be primarily fabricated using hand tools, the roof utilised digital fabrication methods – primarily the laser cutter. As such, the design was driven by the size of the laser cutter bed (60cm x 40cm) and developed through ruled surfaces that were independent from the primary structure; thus allowing for roof to be added onto the final built structure rather than incorporating it alongside its construction. In this context, the roof was designed as a gridded waffle structure comprised from primary (longer) slats and secondary (shorter) slats (Figure 5).
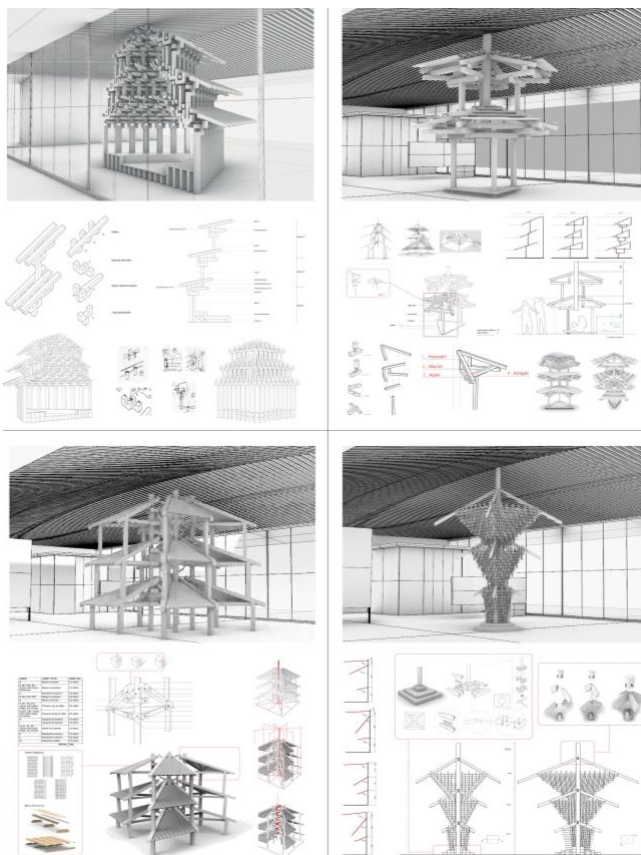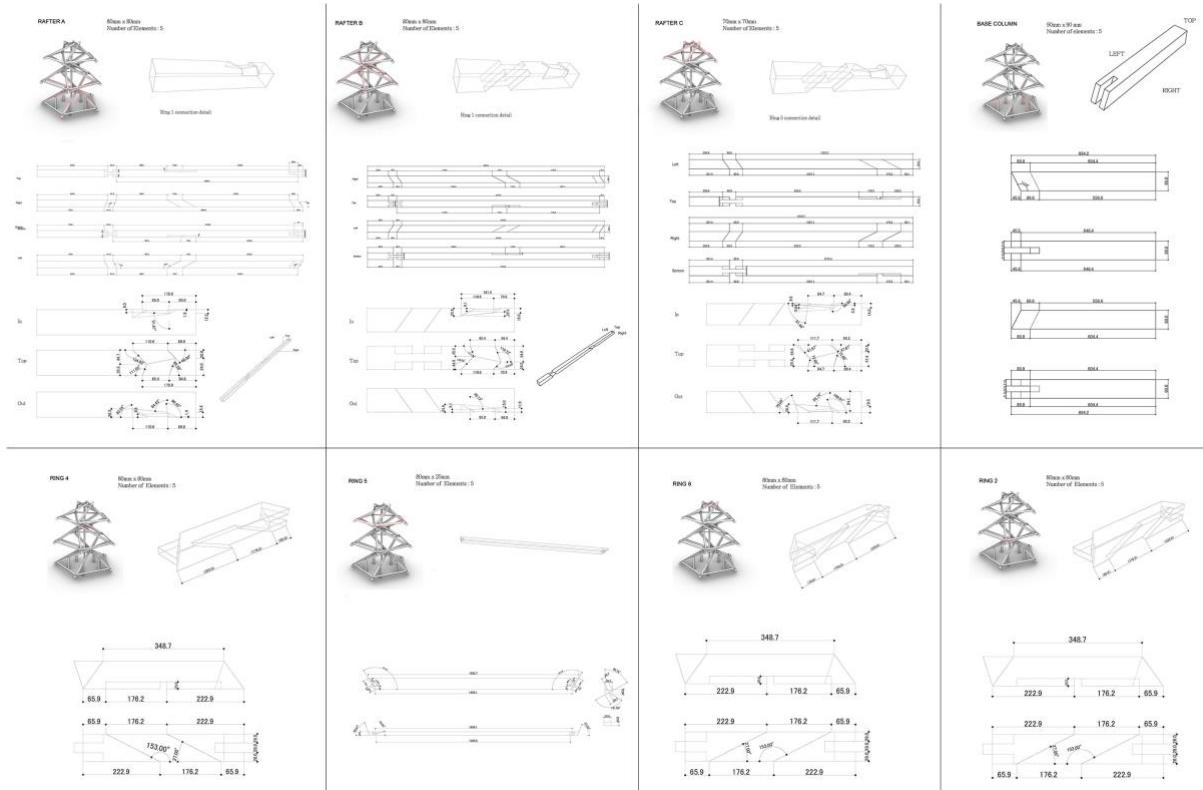
559

Figure 5. Design of the Pagoda selected for further development and fabrication

4.2 Optimisation

The design of the Pagoda was further developed using multi-objective optimisation methods through the use of the grasshopper plugin *Wallacei*. Where most recent applications of evolutionary computation methods have been applied for formal exploration, its application in this context was for the morphological refinement of the structure. The formulated design problem optimised for 3 primary objectives; First, to minimise the overall structural displacement of the Pagoda through using the structural analysis plugin *Karamba*. Second, to minimise the structure's weight by means of reducing the structure's overall size. Third, and in response to the impact of the first two objectives in decreasing the size of the structure, maximise the structure's overall volume. To optimise for the above, the variables inputted into the evolutionary engine were limited to the size and position (in the z-axis) of the compression and tension rings.

The results of the evolutionary simulation, presented through the fitness charts in Figure 6, demonstrate a successful optimisation run for objectives 1 and 2 (displacement and weight), however, seeing as how these objectives conflicted with the third fitness objective (volume), the simulation was less successful in optimising for the latter. However, through clustering the last generation of the simulation using the K-means clustering algorithm (function built into the Wallacei plugin), and analysing the morphologies of the associated cluster centres, the results demonstrate that the fittest solutions for objective 1 (displacement) did not retain the receded feature of the three levels of the Pagoda, instead favouring a linearly extruded form. This highlights the necessity of coupling the statistical analysis of the algorithmic output *with* the morphological analysis of the simulation's results; where the former may present a 'fit' solution according to its numeric attributes, the latter may

highlight alternative results that were not defined through numeric analysis.

Figure 6. Results and analysis of the multi-objective evolutionary algorithm's simulation output (environmental analysis (ladybug) was also conducted on the simulation's output).

560

Figure 7. The complexity of the joints and the varying angles of connections demanded the production of highly detailed cutting files.

4.3 Fabrication

Throughout the design stage of the Pagoda, the influence of the Daiku was in translating the methods associated with preparing the wood for fabrication (i.e. selecting the wood, understanding its characteristics and properties, etc.) as well as understanding the methods and types of Japanese woodworking techniques, highlighting the importance of not using fasteners and relying on the friction between the joints to provide stability (thus signifying the type of joint and its relevance in the overall structure). However, in preparation for fabrication, the Daiku's influence grew exponentially, across multiple fronts which included preparing the wood, marking the wood, reviewing and informing the design of the joints, cutting the joints and finally, assembling the final form. It was at this stage that the gap between digital and master carpenter presented itself, where the simple act of marking the wood for cutting proved to be a challenging task for the digital experts. The precision and attention to detail demanded by the Daiku resulted in not only highly detailed cutting files (Figure 7), but also the understanding that the fabrication of very complex joints is achievable by non-experts through a careful translation from the digital to the physical (Figure 8).

The Pagoda was fabricated over a timeline that spanned only 4 days; however, the actual assembly of the fabricated pieces spanned one afternoon. The primary challenge was cutting the joints between the rings and the rafters, where three timber pieces connected at 3 varying angles across three planes (Figure 9). While in the digital model the joints

Figure 8. The expertise of the Daiku was pivotal in realizing complex cuts and joints.

561

presented themselves as complex, translating them to the physical proved even more so. During this time, revisions were made to the digital model, primarily directed by the master carpenter, that revised the angles in which the rafters intersected the outer rings, with the primary intent to minimise the variation between joints and thus allow for repetition of similar joints at different locations in the Pagoda's design; and although this proved essential, the challenge to fabricate these complex joints continued to prove a difficult task (however imposing repetition did allow for inexperienced users to improve their skills). Although it was clearly evident that the Daiku's experience was vital to the success of realising the final built form (as was expected), what also emerged was the Daiku's influence on the development of the digital model and its translation to the physical, in which it served as a medium of communication between the two.

Figure 9. The assembly of the Pagoda exemplified the elegance (and challenge) of joint connections that avoided the use of fasteners between the timber elements.

5   CHALLENGES

The precision of the digital model was, mostly, successfully translated to the physical structure. However, due to the nature of the project (time limitation, lack of fabrication experience by the participants, absence of auxiliary joints,) and the tools used to fabricate different parts of the structure, there were several challenges in the Pagoda's fabrication. Two such challenges are highlighted below.

Firstly, the complexity of the connection joints between the tension rings and the rafters, primarily due to the deep timber cross section (25mm x 80mm), as well as the positioning and orientation of the timber elements, prevented the joints from securing a perfectly fit and tight connection. Although this was permissible in the scale that the Pagoda was being built at (considering it was a temporary and lightweight structure), the impact of this at a larger scale would have caused the connection joints to not withstand the occurring tension. This could have been avoided however should there have been another round of revision to the digital model, in which the orientation and cross section of the timber elements were revised to achieve a more secure and resolved joint (Figure 10).

Figure 10. Tolerances drastically increased during fabrication of complex joints.

Secondly, the fabrication of the roof plates using the laser cutter (Figure 11) (limiting the cut to one vertical cutting plane) resulted in 'linear contact' points to the multi-directional structural frame, as opposed to the 'surface contact' in the digital model. Although the thickness of the material (5mm) permitted an acceptable tolerance, thicker material sections would have resulted in a much larger tolerance (Figure 12). As before, further time in refining the digital model, so that it is more in line with the fabrication and woodworking techniques of the master carpenter would have allowed for the applied methods to function more efficiently at a larger scale. More importantly, engaging with the Daiku during the design phase in which his expertise in traditional fabrication methods inform the design's development, would allow for these tolerances to be addressed during the digital phase rather than the physical. The above highlights the significance of ensuring that the gap between advanced digital tools and traditional Japanese woodworking methods is based primarily on a feedback mechanism between the two artforms, ensuring a continuous transfer of knowledge between the two.

Figure 11. The roof of the Pagoda took full advantage of the laser cutter, in which all pieces were cut and labelled for pre-assembly before connecting to the structure.
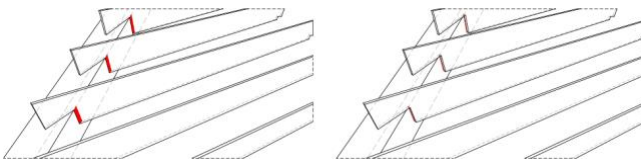
Figure 12. Comparison between surface contact points of the digital model (left), to the linear contact points in the final built form (right)

6 CONCLUSIONS

Traditional Japanese woodworking techniques have been gradually declining in recent years; their use has been limited primarily for the restoration and preservation of historic structures and in the construction of temples and shrines. The traditional craftsmen's manual labour continues to be eclipsed by the sheer volume and speed of machine production, with the presumption that the craftsmanship of the old is no longer applicable to the technological advancements demanded in the modern day. However, advanced tools in computational design and digital fabrication have allowed for the construction of non-identical, unique and complex architectural and structural morphologies, ones inherent to traditional Japanese architecture, easier and cheaper to conduct. Coupled with the younger generations' interest in digital tools and their consequential integration in the development and application of traditional Japanese techniques, facilitates the cultivation of craftsmen's apprentices as successors to the Daiku, preserving the knowledge gained over many generations through a symbiotic relationship with technological advancements of the 21st Century.

To achieve this relationship, the significance of the master carpenter's knowledge and its integration within the digital process is vital. Most evident was the Daiku's relationship with the wood, where his experience and knowledge of the wood's properties, - and most importantly, limitations - proved to be most beneficial, and in turn, the most difficult for the digital translation of the structure to replicate. Thus, highlighting the 'reciprocal' relationship that can be achieved between the master carpenter (and the knowledge of generations past) and the use of advanced digital tools. A relationship that has multiple advantages, most of which is the understanding that rather than advanced digital technologies acting as a hindrance to the transfer of traditional woodworking knowledge, it facilitates a mechanism of support and further evolution of the developed methods and techniques.

The project faced numerous challenges (Section 5), and although these challenges will be addressed in future work, the primary challenge remains in demonstrating that the use of advanced digital tools has a benefit to the preservation of traditional Japanese craftsmanship. Future work will further explore this in applying similar concepts to a larger structure that involves the master craftsman more heavily in the early stages of the design process, in which the contribution of digital tools to Japanese craftsmanship is examined more vigorously prior to fabrication rather than during it.

Figure 13. The final built Pagoda

563

REFERENCES

1. Brown, A., *The Genius of Japanese Carpentry: Secrets of an Ancient Craft*, Revised,Hardcover with Jacket edition, (Tuttle Publishing, 2014).

2. Deb, K., Agrawal, S., Pratap, A., and Meyarivan, T., A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II, in *Int. Conf. Parallel Probl. Solving Nat.*, (Springer, Paris, France, 2000), pp. 849–858.

3. Makki, M., Showkatbakhsh, M., and Song, Y., Wallacei: An evolutionary and Analytic Engine for Grasshopper 3D, *Wallacei* (2018).

4. Matsuura, S., "Miyadaiku" Carpenter Laments Loss of Traditional Knowledge, (2001).

5. Mertz, M., *Wood and Traditional Woodworking in Japan*, (Kaiseisha Press, Otsu City, 2011).

6. Mulligan, M., Craftsmanship For Our Time, in *AU Think. Hand – Takenaka Corp. Takenaka Carpentry Tools Mus.*, (Shinkenchiku-Sha Co., Ltd, 2019).

7. Preisinger, C., Linking Structure and Parametric Geometry, *Archit. Des.* 83 (2013) 110–113.

8. Sadeghipour Roudsari, M., and Pak, M., Ladybug: A parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design, *Proc. BS 2013 13th Conf. Int. Build. Perform. Simul. Assoc.* (2013) 3128–3135

9. UNESCO World Heritage Centre, Buddhist Monuments in the Horyu-ji Area, *UNESCO World Herit. Cent.* (n.d.).

10. Wendelken, C., The Tectonics of Japanese Style: Architect and Carpenter in the Late Meiji Period, *Art J.* 55 (1996) 28–37.

11. Zwerger, K., *Wood and Wood Joints: Building Traditions of Europe, Japan and China*, 2nd edition, (Birkhauser Architecture, Basel, 2012).

Integrated Parametric Design and Production Planning: A Luminaire Design Case Study

Peter Ferschin[1], Georg Suter[2]

[1]TU Wien
Digital Architecture Group
Vienna, Austria
peter.ferschin@tuwien.ac.at

[2]TU Wien
Design Computing Group
Vienna, Austria
georg.suter@tuwien.ac.at

ABSTRACT

The paper describes how to extend the parametric design process by including feedback from the production process and iterating a design towards a realizable object, considering real world constraints such as production costs and manufacturing time. In this paper, we report on our experiences with integrated design and production workflows from teaching a Master level digital design and production class, in which we give students a task to design and produce a luminaire based on a given production infrastructure (3d printers, laser cutters, 3-axis milling machines). Production processes are simulated and and provide feedback to designers about the feasibility of design alternatives regarding material use, cost, machine time. We describe the didactic and technical concepts and conclude with a discussion of open issues.

Author Keywords

Parametric design; digital production; 3d printing; laser cutting; milling.

ACM Classification Keywords

•Applied computing~Arts and humanities~Architecture (buildings)~Computer-aided design

•Computing methodologies~Modeling and simulation~Simulation support systems~Simulation tools

1 INTRODUCTION

Advanced digital design and fabrication technologies provide significant opportunities for designers to realize innovative buildings and architectural products with improved qualitative and quantitative properties regarding shape, material, cost, usability, comfort, or sustainability. In this paper, we investigate the challenge of integrating diverse design and fabrication technologies in order to achieve workflows that let designers rapidly and seamlessly explore the feasibility of design alternatives with respect to given fabrication constraints. Specifically, we explore the linking of state-of-the-art visual scripting based parametric design tools and production planning simulation tools for milling, laser cutting and 3d printing. We describe the necessary knowledge and skills that designers must acquire in order to use these tools productively. In order to mimic realistic conditions for architectural designers, we use the design of a luminaire as a case study.

Our work contributes to the areas of parametric design and digital fabrication by focusing on skills and knowledge, and by considering the integration of production planning simulation for multiple production methods in the design exploration process. Previous related work investigates design knowledge (Woodbury, 2010; Oxman, 2017) and pedagogical aspects of parametric design, and the role of visual programming and scripting languages (Aish & Hanna, 2017; Celani & Vaz, 2012). Further related is work on digital workflows (Wortmann & Tunçer, 2017), design for manufacture and assembly (Austern, et al., 2018), and multi-criteria design (Imbert, et al., 2013)

2 TEACHING DIGITAL DESIGN AND PRODUCTION

We teach integrated parametric modeling and digital production in a set of related, Master level courses (or 'module') to develop and deepen knowledge of topics and skills for integrated parametric design and production. Knowledge of geometry optimization and discretization concepts and techniques is essential for shape designs that are computationally efficient as well as material and production friendly. Such designs may be created with parametric modeling and programming functionality in geometry modeling and CAAD software. Students learn basic data processing concepts and develop programming skills in computational geometry exercises. They learn concepts of digital production methods, including additive, subtractive, and formative fabrication and are aware of workflows, material and production parameters. Acquired knowledge and skills are applied in a small digital production project which we describe in detail in the following sections.

3 PREPARATORY EXCERCISES

To teach the necessary skills, required for an integrated design and production process, we defined various exercises that expose the students to important concepts like surface and solid modeling, visual and text-based scripting as well as digital production methods. We emphasize the main idea of parametric thinking, that allows the generation of numerous design variants, which could be evaluated either by human expertise or by computer simulations (Oxman, 2017).

Important is the concept of identifying key design parameters as independent and dependent variables.

3.1 Surface Modeling

Non-uniform rational B-splines (NURBS) modeling is a quasi-industry standard geometric modeling method for curves and surfaces which is widely used in the development of innovative designs, including industrial design, arts and crafts, and architectural design. Its popularity is due to ease of use in rapidly creating and modifying complex free-form shapes based on intuitive parameters, such as control points and weights. Continuity (or smoothness) is a key quality criterion for free-form shapes. Analysis methods that are provided by geometric design tools include curvature graphs which let designers assess the smoothness of transitions between curve or surface segments. We use the NURBS modeling functionality in Rhino to study surface modeling.

3.2 Visual Scripting

Visual scripting interfaces facilitate graphical rather than text-based program definition. Thus they appear useful for designers who tend not to be expert programmers. Examples of visual scripting environments for computer-aided geometric and architectural design include Rhino Grasshopper and Revit Dynamo. Visual scripting is usually based on the dataflow programming method. A program consists of a directed, acyclic graph where the nodes process data and edges represent data flow between nodes. Each node accepts input data, processes the data, and sends output to related nodes. The graph is automatically evaluated when input data are modified. This is beneficial for design tasks where it is necessary to rapidly modify an evolving design by varying input parameters. We use the Rhino Grasshopper environment for visual scripting.

3.3 Python Scripting

We additionally expose students to text-based scripting, in particular with Python scripting in Rhino as well as in Grasshopper. We choose Python as a language, as it is the most common scripting environment in 3D design applications. We introduce a small subset of the language Python, geometry creation in Rhino and how scripts can be integrated into the Grasshopper environment to extend the predefined functionalities. Each lecture is accompanied by small exercises to verify that the provided scripts of the lectures can be applied with small adaptations. The first lecture is focused on parametric two-dimensional curves, simple 3d solids and extruded geometry from curves. To teach first steps in programming, we implemented a modified version of a Turtlegraphics Python library for Rhino. The exercise for the students consists of writing their first name as a parametric function, additionally supporting different graphical styles (Figure 1). The implementation should define a Grasshopper python component, with a list of polylines as output parameter that can be manipulated further (e.g. by extrusion).
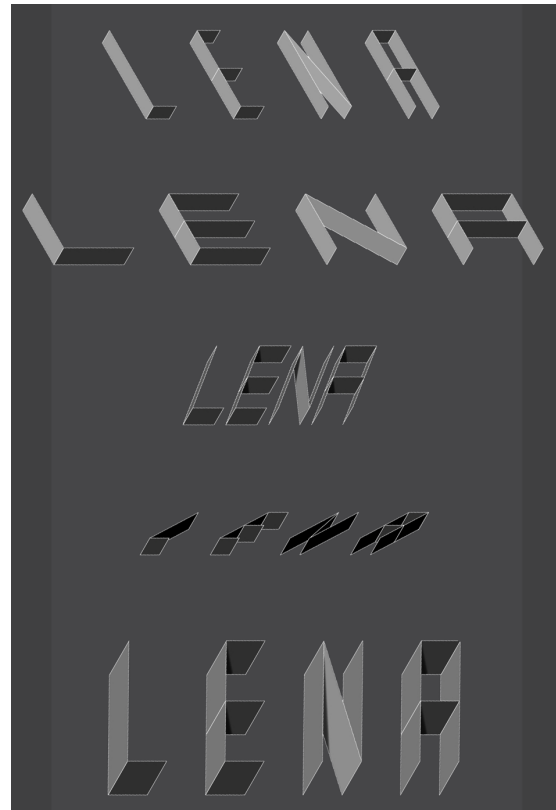
Figure 1. Parametric first name – Lena Roth, 2019

The second lecture focuses on constructing NURBS surfaces from NURBS curves and by defining parametric mathematical functions that specify design parameters and implement surface meshes with adaptable resolutions. As an exercise, students have to define suitable mathematical formulas and design parameters to create parametric dishes with design variations as mesh structures implemented as Python components in Grasshopper.
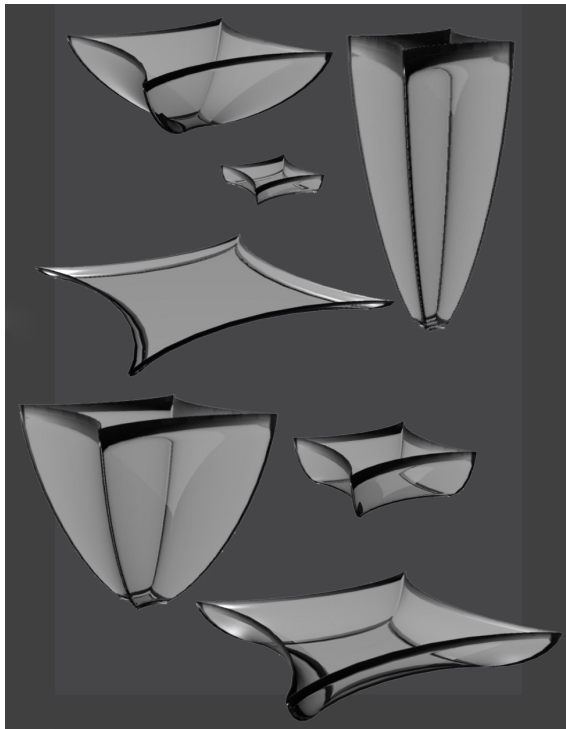
Figure 2. Parametric dishes – Lena Roth, 2019

The third lecture explains how parametrically defined meshes can be converted into buildable structures by generating solid elements at vertices, edges or faces of a mesh. The parametric dish functions of the previous exercise have to be converted into buildable structures.
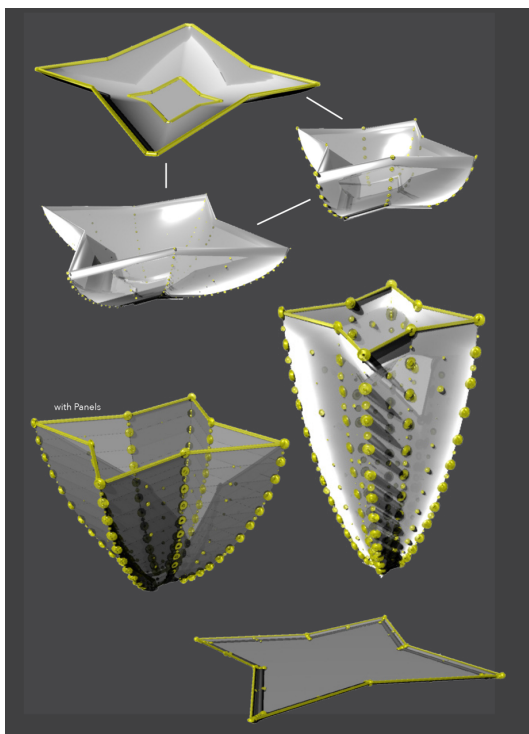

Figure 3. Parametric dish constructions – Lena Roth, 2019

3.4 3D Printing

In a separate lecture ("Digital Production") we explain the principles and possibilities of 3d printing and what requirements for geometric models need to be considered (e.g. solid geometry, closed volumes, consistent normal vectors). As an exercise, students have to prepare a terrain model for printing. For the simulation of the printing process Ultimaker Cura is utilized, parameters that influence the speed and quality of the printing process are explained.
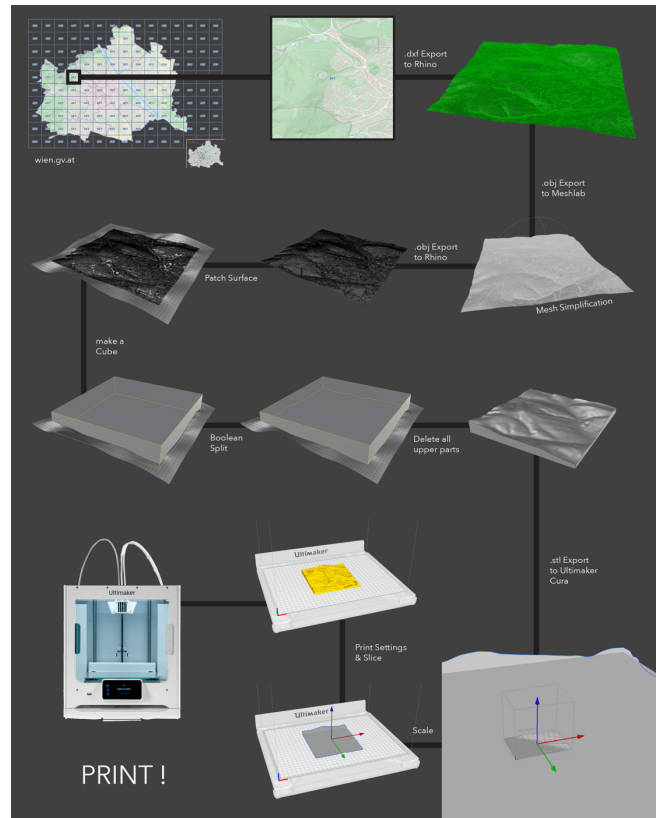
Figure 4. Preparing a terrain for 3d printing – Lena Roth, 2019

3.5 CNC Cutting

In the next step, various CNC cutting strategies are introduced (e.g. laser cutting, waterjet cutting). As an exercise, the same terrain needs to be prepared to be generated as layers on a laser cutter. An additional exercise introduces the task of designing a parametric Kirigami (the art of paper cutting and folding) layout. To simulate the cutting process the tool VisiCut is used to check the layout for the laser cutter as well as to estimate the cutting time.

567
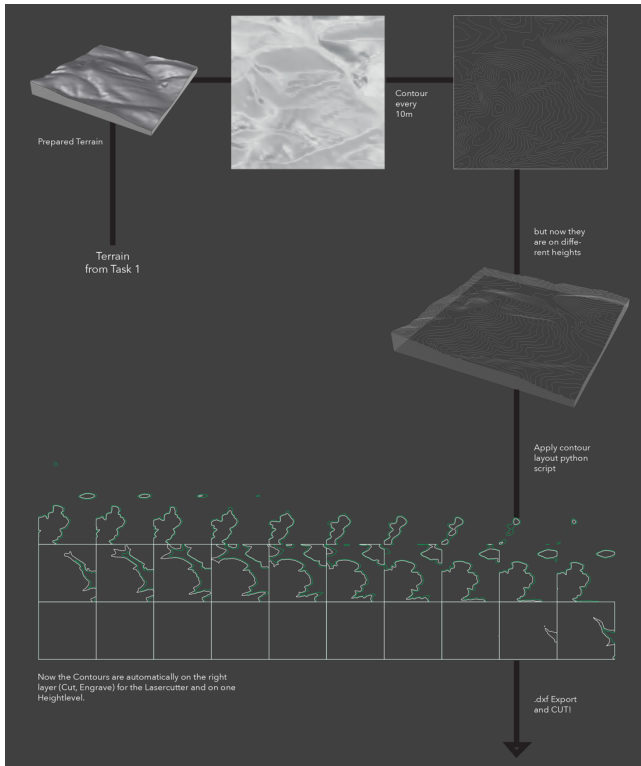
Figure 5. Preparing a terrain for laser cutting – Lena Roth, 2019
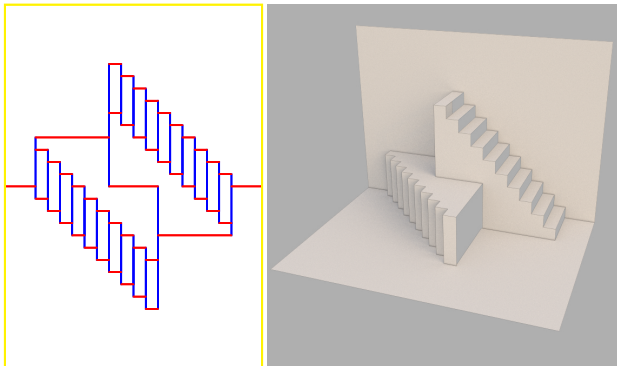
Figure 6. Parametric Kirigami design – Mihael Barada, 2019

3.6 CNC Milling

Next, the principles of CNC milling are introduced and a Python library for generating CNC code is provided to the students. The terrain has to be prepared for a 3-axis milling machine. An additional exercise – to generate parametric milling patterns – has to be implemented as a Python component in Grasshopper. For the simulation of the 3-axis milling process the software Camotics is used. As the milling paths are generated by the students with Python, it is important to check for possible errors that could damage their work piece or even the milling machine. This is done by evaluating the milling paths in Camotics and verifying that no collisions with the milling machine will happen. Furthermore the plausibility of the milling paths in relation

to the stock material is visually inspected. Additionally, the time required for milling is evaluated.
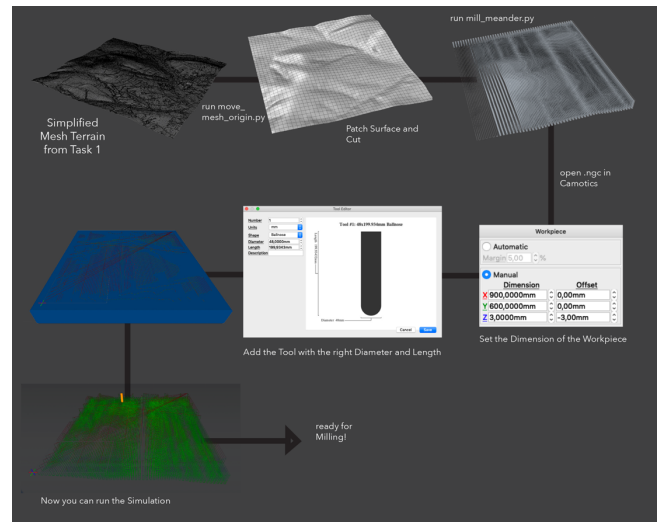
Figure 7. Preparing a terrain for 3-axis milling – Lena Roth, 2019
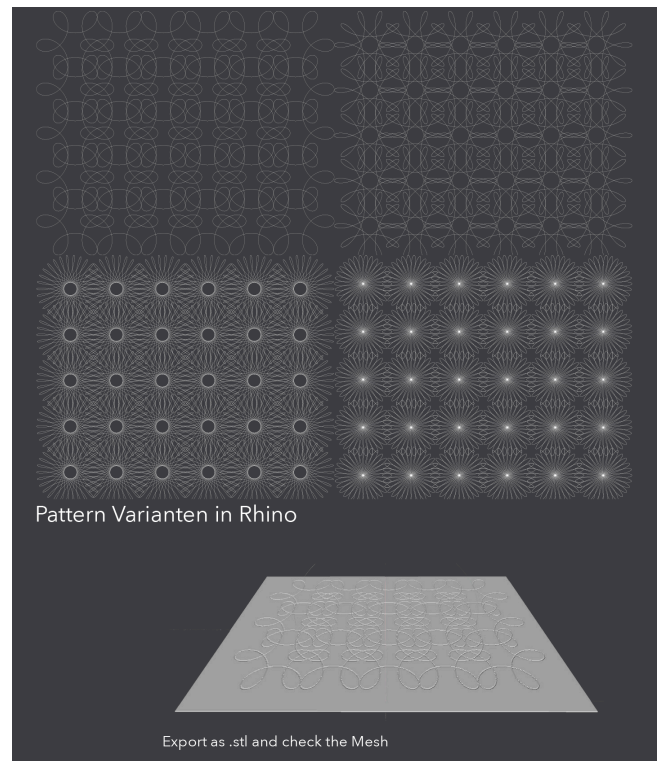
Figure 8. Parametric milling patterns– Lena Roth, 2019

4 PARAMETRIC LUMINAIRE DESIGN AND PRODUCTION

We illustrate the parametric design and production workflows outlined above with an example project.

4.1 Design Brief

The design task involves the design of a luminaire for a specific activity, such as dining or reading, and for a specific room, such as dining room or office. To ensure diversity in luminaire designs, these functional requirements were

568

defined by the students. Designs must further meet specific requirements regarding lamps, screens, and connectors. For example, lamps should be low cost, screens made of acrylic glass, and connectors of synthetic material. Moreover, milling, laser cutting, and 3d printing methods are to be used in production. Production constraints must be considered. For example, the production of luminaire parts by 3d printing should not exceed 4 hours, and the total available volume is at most 11cm x 11cm x 11cm.

4.2 Schedule

Projects need to be completed over a period of 9 weeks. After an initial concept design phase (Week 1), subsequent iterations focus on the development of the parametric model (Week 3), production planning (Week 5), production (Week 7), assembly (Week 8), and final presentation (Week 9). Feedback is provided in review sessions after Week 1, Week 3, Week 5 and Week 9.
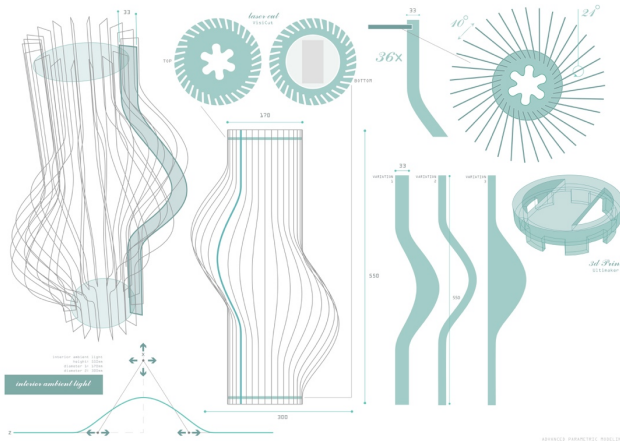
Figure 9. Parametric design, Luminaire "Flora" – Robert Zanona, Christoph Rössler, Erdinc Özcaliskan, 2019

5 SIMULATION AT THE DESIGN AND PRODUCTION PROCESS

Luminaire designs are generated by a parametric model, which is implemented as a combination of manual modeling steps, visual scripting (Grasshopper) and text scripting components (Python). An important part of the design strategy is to decide which components of the luminaire are realized with what kind of production technology. Students need to determine design parameters that influence production time. Through evaluation of design variants, the best parameter sets need to be identified that achieve the intended design while fulfilling production constraints. To facilitate data sharing between parametric design and production planning simulation, files in CNC data formats are automatically created by parametric models.

5.1 Simulating the 3D Printing process

For the parts that need to be 3d printed, Ultimaker Cura is used to verify the geometric consistency and the required printing time. The time constraint was set to 4h printing time per project.
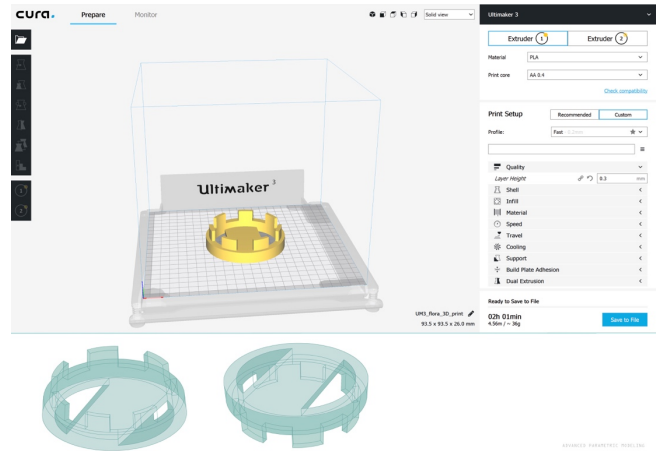
Figure 10. Printing simulation (2:01 hours), Luminaire "Flora" – Robert Zanona, Christoph Rössler, Erdinc Özcaliskan, 2019

5.2 Simulating the Laser Cutting process

Using VisiCut, the layout of the cutting parts is verified, additionally an estimated cutting time is calculated. The time for working on the laser cutter is limited to 1h per group, so a cutting time less than 45min should be achieved. 15min is reserved for machine setup.
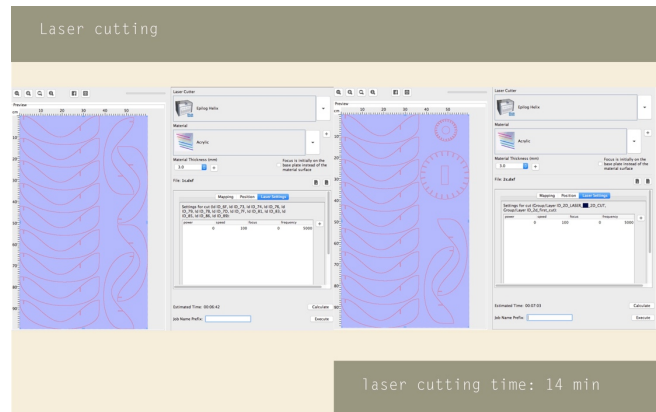
Figure 11. Laser cutting simulation, Luminaire "Onion" – Igor Milekic, Fabijan Farkas, Annabella Lintl, 2019

5.3 Simulating the Milling process

Using Camotics, the simulation of the milling path is used to ensure that no collision with the tool and the milling machine occurs and the milling path is inside the stock dimensions (59cm x 99cm). The tool is fixed with a 6mm ballpoint cutter and tool changes are not allowed. Additionally, the calculation of the milling time is verified with the proposed constraints (45min milling, 15 min preparation).
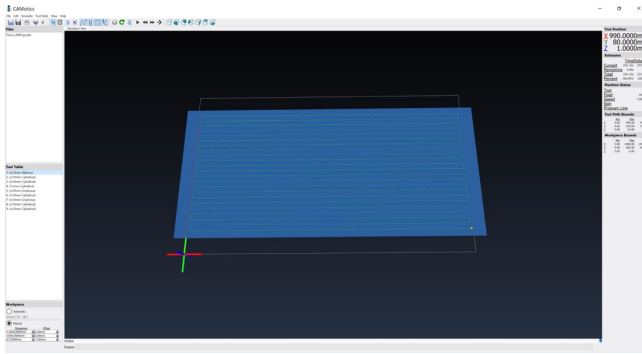
Figure 12. Milling simulation, Luminaire "Flora" – Robert Zanona, Christoph Rössler, Erdinc Özcaliskan, 2019

5.4 Light Simulation of the Lamp Design

To evaluate visual design performance, variants are rendered photo-realistically to identify suitable design parameters for the intended lighting effect.
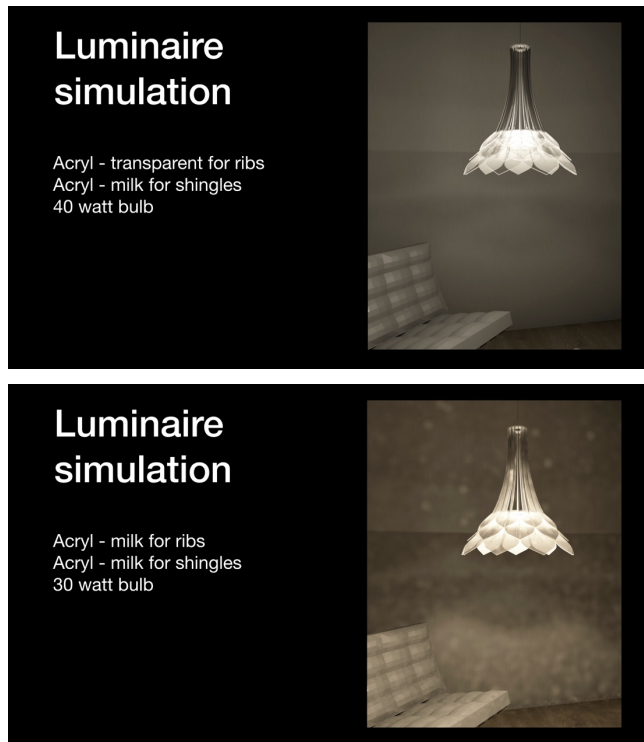

Figure 13. Light simulation, Luminaire "Shingle Dingle" – Ekaterina Ermishkina, Clemens Horvath, 2020

6 PRODUCTION AND ASSEMBLY

Luminaire parts are fabricated in a workshop which is equipped with 2 3-axis milling machines (Step Four Precise 1000U), 2 laser cutters (Trotec Speedy 500), and 4 3d printers (Renkforce RF100 v2). Allowed maximum production times are, respectively, 45 minutes for engraving of screen faces by milling, 45 minutes for cutting screen parts by laser cutting, and 4 hours for creating connector parts by 3d printers. Time slots follow a tight schedule to accommodate all teams. It is therefore mandatory for each team to have production data verified and ready at the

beginning of each production step. The final step involves manual assembly of the luminaire from its parts. Teams track steps in the assembly and keep time records for each step. Issues encountered during assembly are documented.

Figure 14. Milling, Luminaire "Flora" – Robert Zanona, Christoph Rössler, Erdinc Özcaliskan, 2019
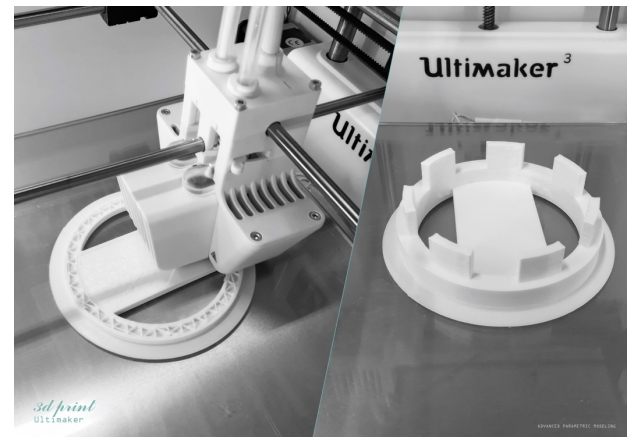
Figure 15. 3d printing, Luminaire "Flora" – Robert Zanona, Christoph Rössler, Erdinc Özcaliskan, 2019
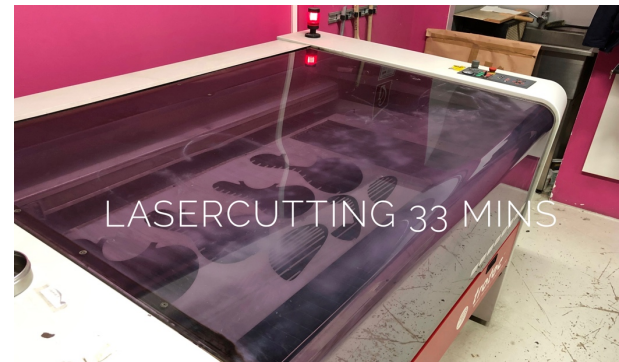
Figure 16. Laser cutting, Luminaire "Bubble" – Iurii Suchak, Lorenzo Valeriano, Pol Wagner, 2019
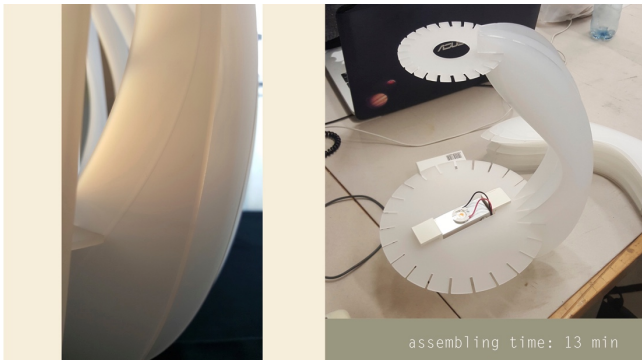
Figure 17. Assembling, Luminaire "Onion" – Igor Milekic, Fabijan Farkas, Annabella Lintl, 2019

7 RESULTS

Examples in Figure 18 give an impression of the diversity of luminaires that are produced with respect to the given design requirements. Featured luminaires are table top and ceiling hung. All examples use an off-the-shelf 10 W LED lamp module which is mounted on an aluminum sheet and provides white light with a high color rendition index (>95).

Figure 18. Luminaire examples: "Chill Bro Lamp" – Asmir Mehic, Kaspar Ehrhart, Amila Imamovic. "Blade" – Blaz Grudnik Tominc, Maximilian Schmid, Cheng Shi. "girih" – Denitsa Dimitrieva, Ada Gulyamdzhis, Behnaz Jalilifar. "MISShape" – Anes Delic, Ajla Kahric, Iuliia Radynska.

In a final presentation session all student projects were presented and the observations in the design and production process were discussed. Some student groups report that they would have changed some design aspects after their experience in the workshop. As an example, the stability of some designs was inadequate as stability simulation was not considered in the design process. Many designs used plug-in connectors to join luminaire screen panels. These had to be strengthened by with glue in several cases, thereby increasing the time to assemble the luminaire. One group created a working model out of cardboard, which helped significantly in improving the design and construction. We will will consider the need for working models in future workflows.

8 CONCLUSION

In this paper, we demonstrated how production simulation can be integrated into the design process. We consider the results as a starting point for future research. We believe that a key issue in an improved design and production workflow lies in the simulation process of the production and the feedback of the simulation results into the design process. An important factor is the generation of design variants that can be evaluated automatically. Unfortunately, at the moment many tools for simulating the production process are separated from each other as well as from the parametric design process. We will target a much tighter integration of simulation tools into the design workflow in future research. This requires fast simulation results to provide almost real-time feedback or at least short interaction cycles. The specifications of production constraints and important feedback parameters (e.g. cost, time, maximum part sizes) need to be standardized on a technical and on a semantic level. File based data exchange should be replaced by stream-based data exchange for shorter feedback loops. Simulation should not be limited to production methods, but include the many dimensions of architectural design for improving the overall design quality. We think that our teaching environment is a very well-suited test environment for an improved design workflow that will integrate simulation tools into architectural design.

REFERENCES

1. Aish, R. & Hanna, S., 2017. Comparative evaluation of parametric design systems for teaching design computation. *Design Studies,* Volume 52, pp. 144-172.

2. Austern, G., Capeluto, I. G. & Grobman, Y. J., 2018. Rationalization methods in computer aided fabrication: A critical review. *Automation in Construction,* Volume 90, pp. 281-293.

3. Celani, G. & Vaz, C. E. V., 2012. CAD Scripting and Visual Programming Languages for Implementing Computational Design Concepts: A Comparison from a Pedagogical Point of View. *International Journal of Architectural Computing,* Volume 10, pp. 121-137.

4. Imbert, F. et al., 2013. *Concurrent Geometric, Structural and Environmental Design: Louvre Abu Dhabi.* Vienna, Springer Vienna, pp. 77-90.

5. Oxman, R., 2017. Thinking difference: Theories and models of parametric design thinking. *Design Studies,* 6.Volume 52.

6. Woodbury, R., 2010. *Elements of Parametric Design.* s.l.:Routledge.

7. Wortmann, T. & Tunçer, B., 2017. Differentiating parametric design: Digital workflows in contemporary architecture and construction. *Design Studies,* Volume 52, pp. 173-197.

8. Grasshopper, McNeel Associates, https://www.rhino3d.com/. As of 11 January 2020.

9. Dynamo, https://dynamobim.org/. As of 11 January 2020.

10. Ultimaker Cura, 3D printing software, http://ultimaker.com/software/ultimaker-cura/. As of 9 January 2020.

11. VisiCut, A userfriendly tool to prepare, save and send Jobs to Lasercutters, http://visicut.org/. As of 9 January 2020.

12. CAMotics, Open-Source Simulation & Computer Aided Machining, http://camotics.org/. As of 9 January 2020.

Life Cycle & Retrofitting

Climate Change Impact on Multi-Objective Optimization: A Case Study on Educational Buildings

Şahin Akın[1,a], Orçun Koral İşeri[1,b], Çağla Meral Akgül[2], Bilge Erdoğan[3], İpek Gürsel Dino[4]

[1] Ph.D. Students, First Co-Authors	[2] Assist. Prof. Dr.,	[3] Assoc. Prof. Dr.,	[4] Assoc. Prof. Dr.,
METU, Ankara, Turkey	METU, Ankara, Turkey	HWU, Edinburgh, UK	METU, Ankara, Turkey
akin.sahin[a], koral.iseri[b] @metu.edu.tr	cmeral@metu.edu.tr	b.erdogan@hw.ac.uk	ipekg@metu.edu.tr

ABSTRACT

The changing weather conditions due to global climate change is expected to have a direct impact on buildings' energy demand and occupant comfort. These conditions are estimated to become more challenging for educational facilities due to their high occupant density and the students' sensitivity to heat. This study aims to present an approach for a comparative analysis for multi-objective optimization results that are projected under different climate change conditions. Two separate optimization processes were performed using NSGA-II for an existing educational building, with the goal of minimizing occupant discomfort and energy use. The differences between the resulting Pareto-sets were analyzed based on the hypervolume difference and statistical evaluations, including the T-test and the distribution of properties. The results of the two optimization processes showed that future weather conditions should be considered on the retrofit process as two Pareto-set have resulted differently in terms of decision variable values and the hypervolume calculation. The discomfort hours ($Y_{1\text{-Current}}$) for optimization with current data resulted in lower values compared to the optimization with projected data ($Y_{1\text{-Projected}}$), on the contrary, the total energy demand ($Y_{2\text{-Current}}$) objective results have resulted in higher values than the projected data results ($Y_{1\text{-Projected}}$).

Author Keywords
Climate Change; Multi-Objective Optimization; Occupant Comfort; Building Retrofits

ACM Classification Keywords
I.6.4 [Simulation and Modeling]: Model Validation and Analysis; I.6.5 [Simulation and Modeling]: Model Development – Modeling methodologies

1 INTRODUCTION
The greenhouse gas emission level has reached its highest level since the industrial revolution, and it continues to rise [10]. The increasing emissions have a negative influence on Earth's climate system, and this situation threatens the livability in the built environment. Concerning this situation, many buildings are required to be adapted [13,14], and the building sector needs performance-based planning and renewal to cope with the changing climate conditions. Particularly, buildings' energy performance parameters (e.g.,

the efficiency of the lighting fixtures, infiltration, and surface heat transfer rates) should be reconsidered for better-performing buildings in the future. Building retrofits plays a critical role in upgrading existing buildings' performances. It should be conducted in a well-planned way to maximize performance and minimize environmental impact. In the literature, it is reported that there is a trade-off between performance parameters, and reaching optimal solutions based on contradicted performance objectives is challenging [7]. However, optimization algorithms could provide an advantage to facilitate achieving optimal solutions. Therefore, optimization is an important technique for the identification of the performance parameter values that can satisfy performance objectives [10].

Energy performance and occupant comfort of buildings may change over the years as a consequence of global climate change. Therefore, the planned retrofit activities should also be responsive to climate change. Notably, it is vital to observe these changes in educational buildings since these buildings host many different users. Even though studies on energy efficiency and reduction of environmental impact in the built environment are increasing, only a few studies are conducted for educational facilities to evaluate and observe energy demand and occupant comfort levels [18].

Comparative studies play a significant role in retrofit planning of the buildings [11,17]. Due to global climate change, the outside temperature conditions are changing, and it has critical effects on the indoor comfort, energy demand, and environmental impact of buildings [8]. Therefore, changing weather conditions should also be considered during retrofit processes by comparing the retrofit scenarios' current and future performances. In this work, the authors realize this comparison with a multi-objective optimization process for an educational facility.

The study aims to realize a multi-objective optimization methodology with the objectives of the calculation of discomfort hours (Y_1) in the ASHRAE 55-2004's summer or winter clothes region [3] and the total energy demand (Y_2) for two climate conditions, e.g., current and projected future weather. The two Pareto-set of the optimizations were analyzed based on Pareto-set hypervolume calculation and the statistical data distribution differences. The presented

approach may help decision-makers making sense of optimization results under different climatic conditions.

2 LITERATURE REVIEW

The building retrofit studies focus on thermal comfort, visual comfort, acoustic performance, and indoor air quality, depending on passive and active interventions to architecture [1,6]. Educational facilities are complex buildings that contain various space types such as classrooms, studios, offices, common areas, requiring different demands. Retrofit processes' performance objectives require alternative interventions for responding to the varying demands of these spaces. Retrofit actions' impact on environmental performance, the relationships between performance objectives, and performance parameters can be realized through optimization studies effectively [10,12].

2.1 Climate Impact on Buildings

Weather information is essential in building energy simulations, as it has a significant influence on energy usage and occupant comfort. As global climate change is expected to have a drastic impact on the environmental conditions in the long term, the energy performance of the buildings is expected to change accordingly. Recent studies have aimed to analyze the effects of climate change for buildings under future climate conditions, to propose improvements at building systems [16,20]. Besides, the potential changes in weather conditions may have some economic consequences that stakeholders of buildings required to deal with [19].

2.2 Optimization and Building Energy Efficiency

Buildings consist of multiple systems and components; therefore, many different retrofit scenarios can be proposed to improve user comfort, reduce energy demand, and environmental impact [10]. Optimization algorithms are widely used in the literature, as they can change multiple variables in a single model at the same time and compute the solution set based on multiple objectives [9,15]. Consequently, designers may select the optimal results according to their priorities from the solution clusters, namely Pareto-set, during the design process [10]. However, decision-making based on different Pareto-sets is a hitherto unaddressed problem. Besides, changing weather conditions may also adversely affect the selection of optimum parameters. Therefore, this study aims to present an approach to interpret the differences between current and future climate conditions in the optimization processes.

3 METHODOLOGY

This section explains the methodological framework for the comparison of two optimization processes for building energy simulations with different climate conditions based on the hyper-volume technique and statistical analysis of the Pareto-set distributions. In this regard, a case study building was selected, through which the approach was presented.

3.1 Description of the Case Study

The case study was selected as the Faculty of Architecture building at the Middle East Technical University, Ankara. Anakara is in Ashrae climate zone 4B (CDD10ºC ≤ 2500 AND HDD18ºC ≤ 3000) [2] and has a dry continental climate with hot summers and cold, snowy winters. Exposed concrete curtain walls, large glass surfaces, brise soleil façades, flat roof, and open plan are the architectural features of this building. The building contains several departments, including architecture, industrial design, and city planning (Figure 1). These spaces are heated from late fall to late spring using a central heating system. The heating system is central to the whole campus, and it runs on natural gas. Fan coil units are used for larger spaces, and radiators were used for smaller spaces to heat the building. Besides, the building does not have any active cooling or mechanical ventilation systems, and the simulations were performed, excluding mechanical ventilation and air conditioning's energy demands.

3.2 Energy model setup

For the building energy simulations and optimization, DesignBuilder [21] was used. For optimizations, the building was only partially modeled in building energy software. The total simulated area was 3733 m^2, including four different zone types such as classrooms, design studios, meeting rooms, and circulation areas, which are occupied dependent on the course hours. However, some places, especially the studios, are extensively occupied even after the course hours.

The building contains numerous volumes and thermal zones, which can result in elongated simulation periods that can challenge the multi-objective optimization process. Therefore it was necessary to first simplify the energy model by combining zones with similar characteristics into a single zone. For instance, classes next to each other were lumped to a single zone. Another simplification act was performed on the shading elements on the west-facing zones. The perforated geometry of these elements was found to increase the simulation time significantly. Therefore, the shading elements were replaced with planar surfaces with reduced opacity. These planar elements ensured that the simulation complexity is reduced while emulating the shading behavior of the actual shading elements. Several energy simulations were conducted to determine the opacity level required for building geometry. The opacity parameter of the surface was decreased gradually until achieving a realistic alternative (Table 1).

Shading	Solar gain (kWh/m2)
Baseline Detailed	197.02
Combined Class, % 60.4 opaque	197.16

Table 1. Analysis for simplified shading geometry for the optimization model

The simulation analysis period was set to six months. As seen in Figure 3, February has the coldest temperature values, and there is symmetrical temperature distribution before and after August. Therefore, the simulation period was set to February 1 – August 1.
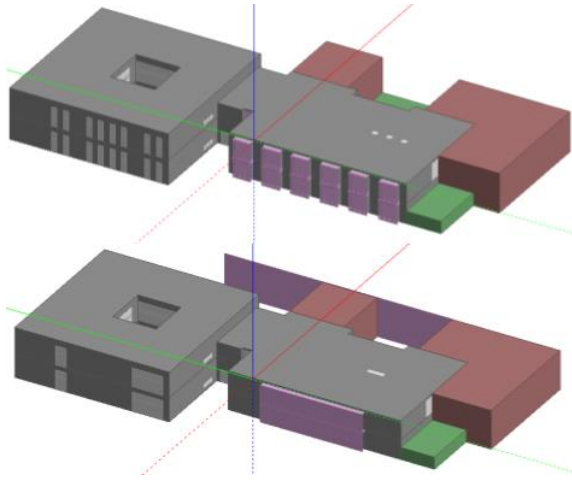
Figure 1. Detailed energy model of METU's architecture faculty (top); Simplified model of the faculty (bottom)

3.3 Optimization and The Selection of Parameters

In terms of energy efficiency and occupant comfort, the variables were determined separately as a single scenario; then, all scenarios are combined under an optimization model. The combination of scenarios in the optimizations was found useful to increase the retrofit actions' impact and to understand how the different scenarios work together. Table 2 shows the selected decision variables and optimization objectives.

As seen in Table 2, the decision variables can take continuous or discrete values, and as-is building's conditions are changed in the simulations by different decision variables. The seven decision variables' values were assigned as discrete values: X_2 (0.00001, 0.0006, 0.000140, 0.000160 kg/s.m@1Pa), X_4 (with combined parameter values, including U-values ranging from 0.6 to 2.6; SHGC values ranging from 0.5 to 0.8), X_5 (with different types, including low to high transmittance; low to high reflectance values for the shading material), X_6 (with different types, including overhangs ranging in length 0.5 to 1 m or sidefins ranging from 0.5 to 1 m; louvres ranging from 0.25 to 0.5 m), X_7 (with combined parameter values, including insulation thicknesses ranging from 4 cm to 16 cm; outmost coating with reflective or non-reflective paint), X_8 (with combined parameter values, including insulation thicknesses ranging from 0 to 16 cm; outmost coating with reflective or non-reflective paint), X_9 (with combined parameter values, including aluminum, wooden, PVC materials; frame thicknesses ranging from 5 to 60 mm).

The solutions are evaluated by two objective functions, namely the minimization of the total energy demand (kWh) and minimization of the discomfort hours. It is important to state that there is also a trade-off between these two objectives. The discomfort hours (Y_1) is calculated by summing up the exceeded degree differences from 28°C, which is determined as the top limit for the comfortable indoor temperature for all seasons [3]. The objective of total energy demand (Y_2) contains heating and lighting demands

only since the building does not have active mechanical ventilation and cooling system. Also, the defined retrofit scenarios do not focus on changing electricity consumption; therefore, the Y_2's results can only be evaluated over changes in heating energy consumption.

	Selected Retrofit Scenarios	As-is building	Range
X_1	Opening Ratio of Interior Windows to increase cross Ventilation	0%	0% - 90 %
X_2	Air Tightness (infiltration)	0.001 kg/s.m @1Pa	4, Discrete
X_3	Opening Ratio of Exterior Windows to increase natural ventilation	15%	10% - 90 %
X_4	Glazing Types (with different SGHC, U-value)	2.6 U-value,	36, Discrete
X_5	Shading Types, Interior (with different transmittance and reflectance values)	Non-reflect. Drapes	6, Discrete
X_6	Shading Types, Exterior (with different depth and shading type)	Brise-soleils	6, Discrete
X_7	Roof Insulation (with reflective coating layer)	4cm non reflective insulation	8, Discrete
X_8	Wall Insulation (with reflective coating layer)	3.316 U-value	8, Discrete
X_9	Window Frame Types (with different thickness and material type)	40 mm Aluminum	9, Discrete
Selected Objective Functions			
Y_1	Discomfort Hours (hours) according to ASHRAE [3]		
Y_2	Total energy demand (kWh)		

Table 2. Decision variables and objectives

3.4 Optimization Algorithm and Energy Simulations

In this study, DesignBuilder's optimization module that implements the Non-Dominated Sorting Genetic Algorithm-II (NSGA-II) was used. NSGA-II is a multi-objective genetic algorithm that can maintain a spread of solutions and converge in the obtained nondominated front [7]. In this study, the algorithm aims to minimize the total energy demand (Y_1) and minimize the discomfort hours (Y_2) simultaneously.

Figure 2 presents the flowchart of the NSGA-II genetic algorithm's working process. In the first step, the population (R_t) is generated with P_t and Q_t. The population is sorted based on the non-domination rule with elitism. The F_1 set is formed by best-nondominated solutions. The members of the new population P_{t+1} are selected from F_1, and if needed, from F_2 and F_3. Then, the new population P_{t+1} with the size N is used for selection, crossover, and mutation to form a new

population Q_{t+1} of size N as the next generation. The selection criteria of members works according to the crowded-comparison rule in which calculating the rank and crowded-distance of each solution as generating P_{t+1}.
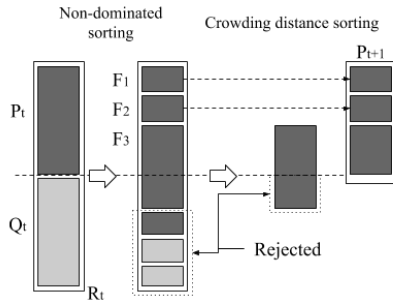

Figure 2. NSGA-II optimization loop [7]

3.5 Weather Data

The account for the impact of climate change on building performance, the future weather data was generated using Weathershift [5]. Weathershift transforms existing weather datasets (.epw files) based on different greenhouse gas emission projections in a procedure named 'morphing'. The process alters current weather variables to produce future weather time series. Morphing preserves the actual weather sequences of the current weather data. Offset data is a climate projection generated by Coupled Model Intercomparison Project Phase 5 (CMIP5). These models are defined as Representative Concentration Pathways (RCP). For instance, RCP 8.5 is the business-as-usual scenario as a worst-case scenario. In this study, the RCP 8.5 scenario was used to observe the climate impact on the retrofit process.

The current and projected weather data were used to compare climate change impact on educational facilities in terms of retrofit scenario evaluation using evolutionary algorithms. The two weather files represent two different periods. The first one represents the current weather conditions. The second represents the future weather conditions that are projected for the year 2060. The current and projected outside dry bulb temperature values can be seen in Figure 3.
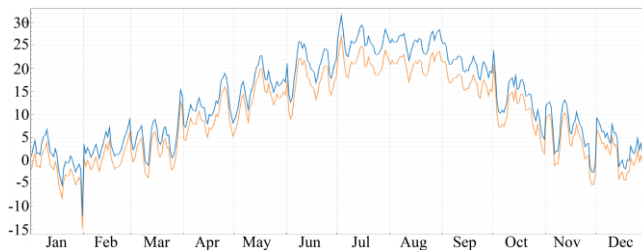
Figure 3. Dry-bulb temperature (C°); Current weather data (Orange), Projected weather data (Blue)

3.6 The Hypervolume Difference Indicator

The hypervolume indicator is used in this paper to compare the two Pareto-front sets of the two optimization results. The method calculates the volume of the dominated region of the distribution of the objective functions [4]. The comparison of two optimization results is the difference between the hypervolume area based on the total energy demand and the total number of discomfort hours (1). This sort of comparison method was used to compare different optimization sets between each other only since it is unitless.

$$\Delta I = |I_i - I_f| \qquad (1) \, [4]$$

where ΔI is the region difference between two multi-objective optimization areas based on the previous formula. I_i is the area of optimization with current weather data, and I_f is the optimization with projected weather data.

3.7 The statistical Evaluation of the Pareto-sets

As an alternative comparison method, the distribution of the two Pareto-set is analyzed based on statistical tests, i.e., kurtosis, skewness value comparison. The results also were presented on the box plot. The dominated solutions were not included in the observation. In addition to the skewness and kurtosis analysis, the Paired Sample T-test were used. T-Test presents univariate comparative descriptive statistics (mean, std. deviation) for each variable of the test [14]. The test is a statistical method for understanding whether the mean difference between the two situations is significant.

4 RESULTS AND DISCUSSION

This section presents the results of two optimization processes and their differences in terms of hypervolume and statistical analysis. The execution time was approximately four days for each optimization process. The population and initial random population sizes were determined as 100 individuals. After the 70th generation, no improvement was observed. It is important to note that a brute-force calculation would take approximately 9538 days for the same parameter set. As a result, a significant reduction in the computation time was achieved by applying the optimization algorithm. For comparison of the optimization results, the highest values on the Pareto-set in terms of objective functions were presented. In Figure 4, the highest value according to Y_1 was circled and shown as 1. Besides, the highest value according to Y_2 and shown as 2. Subsequently, statistical analyses were applied to compare the distribution of Pareto-set results. Finally, the area differentiation between two Pareto-sets was calculated by the hypervolume difference.

4.1 Optimization with the Current Weather Conditions

The optimization process of the current weather condition has produced 49 optimum results. As a result of the optimization with the current dataset, while the discomfort hours ($Y_{1\text{-Current}}$) were found between 1005 and 2321 hours, the total energy demand ($Y_{2\text{-Current}}$) demand values were between 37.06 kWh/m², and 47.16 kWh/m² (Figure 4). The discomfort hours resulted in more than 1000 hours. These results could be explained by the fact that the selected building was built in the *1960s*, and does not have any wall or floor insulation. Moreover, the lack of air conditioning and inefficient natural ventilation also contributed to high overheating.
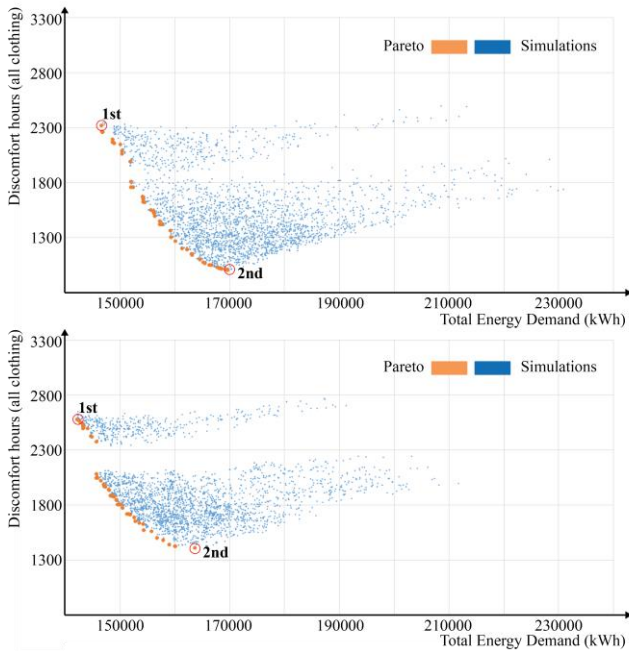
Figure 4. Current Weather Data optimization results (top); Projected weather data optimization results (bottom)

According to the simulation results, $Y_{1\text{-Current}}$ was calculated as 2261 hours, and $Y_{2\text{-Current}}$ was calculated as 37.08 kWh/m^2 for the highest value, according to Y_1. These results from the Pareto-set tend to increase envelope thermal resistance in terms of selecting the highest-performance of infiltration (X_2) and decreasing external window opening ratios (X_3) to 0%. Also, the algorithm chose decision variable alternatives for benefiting from solar gains, such as increasing the SGHC value for the glazing types (X_4), and using shorter exterior overhang shading alternatives (X_6). On the other hand, $Y_{1\text{-Current}}$ emerged as 1010 hours, and $Y_{2\text{-Current}}$ emerged as 42.78 kWh/m^2, as it is the highest value, according to Y_2. This optimum result consists of the envelope with less thermal resistance, medium level natural ventilation, and infiltration decision variable selection, e.g., low-performance infiltration selection (X_2) and increasing external window opening ratios (X_3) to 90%. On the contrary, the level of solar gain benefit could be reduced by choosing the longer overhang shading alternative (X_4) and by decreasing the SGHC value for the glazing type (X_4).

Finally, similar decision variable values were observed for the two Pareto-set comparisons. For instance, the internal window opening ratio (X_1) resulted in a 90% openness ratio, and the window frame thickness (X_4) was 60 mm for all Pareto results. Besides, non-reflective outmost materials for the flat roof construction (X_7) resulted in both optimizations.

4.2 Optimization with Future Weather Conditions

During the optimization process with the future weather file, 49 different Pareto solutions were generated. The 2nd optimization Pareto-set resulted between 1407 hours and 2576 hours for $Y_{1\text{-Projected}}$; $Y_{2\text{-Projected}}$ results varied between 35.92 to 45.50 kWh/m^2 (Figure 4). By observing the Pareto-front results, the decision variables, including the infiltration (X_2), opening ratio exterior window (X_3), shading types-exterior (X_6), window frame types (X_9) had multiple values instead of converging only on a single value. On the other hand, the decision variables, including the opening ratio of interior windows (X_1), windows' blind types (X_5), roof insulation (X_7), wall insulation (X_8) were positioned in a narrow scale by having similar values, e.g., "90% opening ratio of the interior windows with high reflectance – low transmittance window blind types".

To explain the distribution of the results in more detail, the presentation of the two extreme points on the Pareto-set helps. The first extreme resulted in one of the highest points of the Pareto-set, where $Y_{1\text{-Projected}}$ was 2573 hours, and $Y_{2\text{-Projected}}$ was 39.57 kWh/m^2. This solution has the lowest infiltration rates among the other solutions (X_2), and the thickest window frame type option (X_9) in the decision variables. Similarly to the optimization with current weather data, there was a tendency to increase thermal resistance on the envelope. In parallel with this situation, alternatives were chosen to provide more benefit from solar gain in terms of decision variables, i.e., reducing the SHGC value (X_4) and using shorter overhang shading depth (X_6). Lastly, the non-reflective wall insulation options were found effective to receive heat gains from vertical surfaces.

The second extreme had one of the lowest values in the Pareto-front trend line and had $Y_{1\text{-Projected}}$ of 1559 hours, and $Y_{2\text{-Projected}}$ of 43.30 kWh/m^2. On the contrary, this result consists of the highest infiltration rates among the other solutions (X_2), and has a thinner window frame thicknesses as 50 mm compared to the first extreme's X_9 decision variable. Some of the results were close to the Y_1 by generating less heat resistance on the envelope. On the other hand, some of the results concluded as more focused on Y_2 by using thicker insulations.

4.3 Comparison of Pareto-set solutions

According to Figure 4, it can be seen that the Pareto-sets were horizontally separated into two parts because of the differences in decision variables' effects on the simulations' performances. For the optimization with current weather data, the upper part of the Pareto-set resulted in between 1956 to 2321 hours, and the lower part of the Pareto-set resulted in between 1005 to 1807 hours in terms of the discomfort hours objective. For the optimization with projected data, the upper part resulted in between 2371 to 2575 hours, and the lower part resulted between 1407 to 2083 discomfort hours.

The projected weather data contains higher dry-bulb temperatures and irradiation values as compared to current weather data. Therefore, the total energy demand values were resulted in lower compared to the projected weather data optimization since the building demands lower heating energy consumption. On the other hand, the discomfort hours

were found higher in number because of the increased outside temperature values for the projected weather data.

According to Figure 5, the calculated region of all solutions for current weather data was found as 83247140, and it was found as 62607314 for the optimization with projected data (1). Therefore, the hypervolume difference between the two (ΔI) was found as 20639826 (2). The current weather data's Pareto-front set positioned closer to the origin (0,0) compared to the optimization with projected-data. For two optimization processes, the Pareto-set distributions were performed similarly in terms of the ranges of objective functions, instead of the distribution values. Nevertheless, the distribution of the Pareto-set was different for two optimizations as the current data's results were lower for Y_1. On the contrary, for the projected weather data, Y_2 resulted in lower values. The optimization with the current weather data's solution cluster dominantly covered more areas than the projected weather data's solution cluster. This can be explained with the high outside temperatures in projected weather data, which affect the distribution of the solution cluster.
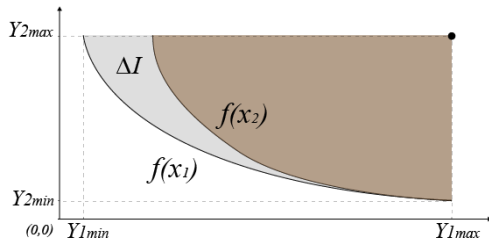
Figure 5. Hypervolume Indicator (Hv) difference between two multi-objective optimization Pareto-front set

4.4 An analysis of the decision variables on the Pareto-fronts
In the Pareto-sets, the frequency rate of several variables was found higher than the rest. This finding indicates that the recurring variables in the Pareto-sets have the potential in terms of high performance for both optimizations. Accordingly, the opening ratio of the interior windows (X_1), window types (X_4), roof insulation (X_7), wall insulation (X_8) decision variables' values were observed similarly in both of the optimizations. Particularly for X_7 and X_6, compared to the non-insulated existing conditions of the building, the highly insulated alternatives in terms of high resistance to thermal transfer perform better for the current and projected weather of Ankara. Similarly, the window alternatives with low U-values were observed more in the Pareto-sets. Also, the opening ratio of the interior windows (X_1), which helps reduce interior discomfort hours, was selected with a %90 openness rate for both of the optimizations. On the other hand, for some decision variables, there was no definite dominant selection for both optimizations, i.e., window frame types (X_2) and interior shading types (X_5) and exterior shading types' (X_6).

Apart from this, when the X_3 decision variable was examined in both Pareto-sets, the number of discomfort hours increased as the openness rate approached 0%, and the

number of discomfort hours decreased as it approached 90%. The Pareto-sets' results showed that some decision variables were more impactful in terms of changing the total energy and discomfort hours. Strikingly, the infiltration performance parameter was observed as the most dominant decision variable among the other scenarios, which drastically effects the overall building performance. Also, exterior shading types (X_6) more dominantly had a more impact on the objective values for both optimizations compared to the interior shading types (X_5).

On the other hand, window frame types' (X_2) material parameters were generally observed with PVC since the material has better heat transfer values compared to the other defined alternatives. However, the frame thickness parameter varied between 0,005 to 0,060 meters. This variation was observed more, especially in the Pareto-set with the current weather data.

4.5 Statistical Comparison of the Two Generated Data Sets
Multi-objective optimization aims to search for non-dominated results based on decision variables. Therefore, objective functions should be analyzed based on the difference between the two Pareto-sets. Figure 6 shows the Pareto-set results of the optimizations on the box plot with groups of data based on quartiles.

The optimization with current data's Pareto-set's $Y_{1\text{-Current}}$ took values between 1300 and 1650 hours, and the optimization with projected data's Pareto-set's $Y_{1\text{-Projected}}$ took values between 1800 and 2050 hours for the interquartile range. Both objectives of the optimizations, the outlier values were positioned closer to the upper-extreme values. The Pareto-set of current data values started from up to 2000 discomfort hours, and for the Pareto-set of the projected data values, they started from up to 2400 discomfort hours.
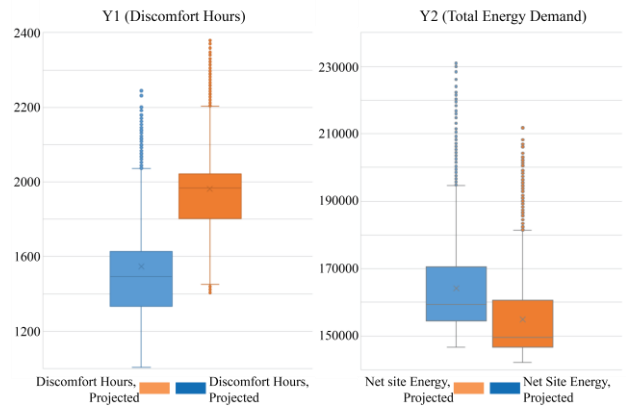

Figure 6. Objective functions distributions on box-plot; Y_1 based on ASHRAE 55 calculation (left), Y_2 (right) [20]

Further, $Y_{2\text{-Current}}$ results were distributed in a broader range between 43.10 kWh/m^2 to 47.27 kWh/m^2. However, $Y_{2\text{-Projected}}$ values took values between 40.87 kWh/m2 and 44.50 kWh/m2. The difference for two optimizations caused by a similar reason that was the higher values of weather conditions in terms of outside dry-bulb temperature (C°). For

the outlier values, the 1st optimization objective values started from 52.84 kWh/m², and the 2nd optimization objective values started from 44.50 kWh/m², so its results ended up in a narrower area.

	N	Mean	SD	Skewness	Kurtosis
$Y_{2\text{-Current}}$	49	39.78	1.70	0.225	-1.011
$Y_{2\text{-Projected}}$	49	35.93	1.29	0.837	0.300

Table 3. Distribution Properties of Objective Pareto Set (Y_2)

As seen in Table 3, the distribution properties of $Y_{2\text{-Current}}$ are positively less skewed; on the contrary, the results with $Y_{2\text{-Projected}}$ were found as highly positively skewed. Both optimization distributions were generally populated to the lower total energy demand values. For the kurtosis evaluation, $Y_{2\text{-Current}}$ had a negative value, which can be explained that the distribution had a high peak than the usual. However, $Y_{2\text{-Projected}}$ had a positive kurtosis value that was flatter than the usual, which means that they were equally distributed between the boundaries of the total energy demand results.

	N	Mean	SD	Skewness	Kurtosis
$Y_{1\text{-Current}}$	49	1537	416	0.411	-1.097
$Y_{1\text{-Projected}}$	49	1929	342	0.534	-0.581

Table 4. Distribution properties of objective Pareto-set (Y_1), ASHRAE 55 [20]

In Table 4, the normal distribution properties of the discomfort hours objective function (Y_1) were positively skewed for both of the results with $Y_{1\text{-Current}}$ and $Y_{1\text{-Projected}}$. Consequently, the optimization algorithm could reach lower values for Y_1. For kurtosis evaluation, each objective function in two optimizations valued as negative, which means the distribution has a high peak than the normal. It can be said that the objective values were evaluated close to average.

$Y_{1\text{-Current}}$ & $Y_{1\text{-Projected}}$	
Mean	-538
SD (σ)	94
%95 C.I. Difference	-566 (Lower), -511 (Upper)
t_{49}	-39.781
Df	48
Sig. (2-tailed)	.000
Correlation	0.964

Table 5. Paired sample t-test of discomfort hours, ASHRAE 55 function Pareto-set

Table 5 presents the results of Y_2 for the two Pareto-sets (r=0.964, p<0.001). The distribution of two Pareto-set was strongly correlated similarly to the Y_1 comparison. There was a significant difference in the average values between two Pareto-front distributions (t_{49} = -39.781, p < 0.001). On

average, the objective function of the current weather data optimization was 538 discomfort hours lower than the optimization with the projected weather data (95% CI [-566, -511]). The lower temperature values for the projected weather data directly affected the Y_1's Pareto-set values.

In Table 6, the two optimization processes' total energy demand values were positively and strongly correlated to each other (r=0.970, p<0.001). Also, there was an important statistical difference in average values between two Pareto-front distributions (t_{49} = 36.313, p < 0.001).

$Y_{2\text{-Current}}$ & $Y_{2\text{-Projected}}$	
Mean	1.47
SD (σ)	1.88
%95 C.I. Difference	0.88 (Lower), 1.96 (Upper)
t_{49}	31.799
Df	48
Sig. (2-tailed)	.000
Correlation	0.970

Table 6. Paired sample t-test of total energy demand objective function Pareto-set

This study aimed to interpret the differences between two different optimizations' Pareto-sets, which were simulated with different weather conditions, e.g., current and future weather data. According to the optimizations, climate change has a dramatic impact on the selected educational building's energy consumptions and occupant comfort. Therefore the retrofitting processes should not be conducted by just considering the current weather conditions since this approach provides limited insight. The projected weather data contains higher outside dry bulb temperature and global irradiation values, which affects both energy demand and occupant comfort in buildings. The two optimizations' results showed that the optimization with $Y_{2\text{-Projected}}$ has lower values for the performance of convergence, but $Y_{1\text{-Projected}}$ was found higher in terms of higher dry-bulb temperature values. There were significant changes in both objective functions and the performance parameters of the optimizations. Therefore, the optimization or retrofit processes should consist of long periods of planning and should also include worst-case scenarios for better adaptation to future climate conditions.

5 CONCLUSION
In this paper, an approach for the comparative evaluation of the Pareto-sets conducted with current and future climate conditions for an educational facility is presented. For optimization, the main objectives were selected as total energy demand (Y_2) and discomfort hours (Y_1). The optimization results were found a high number of discomfort hours, according to Ashrae 55 [3], since the selected building was historic and had not any insulations on the envelope.

In the current weather data optimizations, $Y_{2\text{-Current}}$ had high values compared to $Y_{2\text{-Projected}}$, since the comparatively low outside dry-bulb temperatures caused high heating demands in the selected building. For the same reason, $Y_{1\text{-Current}}$ had lower values compared to $Y_{1\text{-Projected}}$. Oppositely, the optimization results for the projected weather data had higher $Y_{1\text{-Projected}}$ values and had lower $Y_{2\text{-Projected}}$ values due to the increased outside dry-bulb temperature and global irradiation (W/m^2). The Pareto-set of the two optimization processes were compared with the hypervolume indicator and statistical methods. By comparing the two Pareto-front results, the alteration occurring between the two optimizations was possible to be detected. In this way, the planned retrofit scenarios may be oriented to global climate change and may function more effectively for the educational facilities in the future.

REFERENCES
1. Ainger, C., Bernier, P., and A. Fenner, R. Assessing the sustainability merits of retrofitting existing homes. *Proceedings of the ICE - Engineering Sustainability*, *163*, (2010), 197–207.

2. ASHRAE. ASHRAE climatic design conditions 2009/2013/2017. http://ashrae-meteo.info/index.php?lat=39.950&lng=32.883&place=%27%27&wmo=171300.

3. ASHRAE. *ASHRAE Standard 55-2004 -- Thermal Comfort*. 2004.

4. Auger, A., Bader, J., Brockhoff, D., and Zitzler, E. Hypervolume-based multiobjective optimization: Theoretical foundations and practical implications. *Theoretical Computer Science*, *425*, (2012), 75–103.

5. Belcher, S.E., Hacker, J., and Powell, D.S. Constructing design weather data for future climates. *Building Services Engineering Research and Technology*, *26*, (2005).

6. Berardi, U., Manca, M., Casaldaliga, P., and Pich-Aguilera, F. From high-energy demands to nZEB: The retrofit of a school in Catalonia, Spain. *Energy Procedia*, *140*, (2017), 141–150.

7. Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, *6*, 2 (April 2002), 182–197.

8. Dino, I.G. and Akgül, C.M. Impact of climate change on the existing residential building stock in Turkey: An analysis on energy use, greenhouse gas emissions and occupant comfort. *Renewable Energy*, *141*, (2019), 828–846.

9. Giouri, E.D., Tenpierik, M., and Turrin, M. Zero energy potential of a high-rise office building in a Mediterranean climate: Using multi-objective optimization to understand the impact of design decisions towards zero-energy high-rise buildings. *Energy and Buildings*, *209*, (2020), 109666.

10. Hashempour, N., Taherkhani, R., and Mahdikhani, M. Energy performance optimization of existing buildings: A literature review. *Sustainable Cities and Society*, *54*, (2020), 101967.

11. Hou, J., Liu, Y., Wu, Y., Zhou, N., and Feng, W. Comparative study of commercial building energy-efficiency retrofit policies in four pilot cities in China. *Energy Policy*, *88*, (2016), 204–215.

12. Jafari, A. and Valentin, V. An optimization framework for building energy retrofits decision-making. *Building and Environment*, *115*, (2017), 118–129.

13. Konstantinou, T. *Facade Refurbishment Toolbox*. 2014.

14. Sims, B.L. and Bakens, W.J.P. Next Page CIB Performance Based Building (PeBBu) Thematic Network. (2002), 1–8.

15. Sun, Y., Ma, R., Chen, J., and Xu, T. Heuristic optimization for grid-interactive net-zero energy building design through the glowworm swarm algorithm. *Energy and Buildings*, *208*, (2020), 109644.

16. Troup, L., Eckelman, M.J., and Fannon, D. Simulating future energy consumption in office buildings using an ensemble of morphed climate data. *Applied Energy*, *255*, (2019), 113821.

17. Witt, E., Lill, I., and Nuuter, T. Comparative Analysis of Current Guidance for the Evaluation of Building Retrofit Investments. *Procedia Economics and Finance*, *21*, (2015), 321–328.

18. Yigitcanlar, T. *Technology and the City: Systems, Applications and Implications*. Routledge, New York, 2016.

19. Zhai, Z.J. and Helman, J.M. Implications of climate changes to building energy and design. *Sustainable Cities and Society*, *44*, (2019), 511–519.

20. Zheng, Y. and Weng, Q. Modeling the effect of climate change on building energy demand in Los Angeles county by using a GIS-based high spatial- and temporal-resolution approach. *Energy*, *176*, (2019), 641–655.

21. DesignBuilder Software Ltd - Home. https://designbuilder.co.uk//.

Lifetime Energy Performance of Residential Buildings:
A Sensitivity Analysis of Energy Modeling Parameters

Diba Malekpour Koupaei[1], Manon Geraudin[2] and Ulrike Passe[1]

[1]Iowa State University
Ames, USA
{malek, upasse}@iastate.edu

[2]INSA Lyon
Lyon, France
manon.geraudin@insa-lyon.fr

ABSTRACT
Traditional building energy simulation tools often assess performance as a function of the unique climate, physical characteristics, and operational parameters that define specific buildings and communities, planned or existing. This paper presents the results of a sensitivity analysis on the input parameters (relating to both the building and climate) that affect the annual energy consumption loads of an existing residential neighborhood in the U.S. Midwest over the anticipated service life of its buildings using the Urban Modeling Interface (*umi*). Accordingly, first, the effect of multiple building construction characteristic packages and inclusion of outdoor vegetation, are investigated under typical meteorological climate conditions. Afterwards, since typical climate conditions may not adequately describe the potential extreme conditions that will be encountered over the entire service life of these buildings, alternative weather datasets were also utilized in the sensitivity analysis. The study's findings suggest that cooling loads are expected to increase dramatically over the next five decades, both due to changes in the climate and the more wide-spread use of air-conditioning units. Since the results showed that trees can effectively reduce cooling loads by up to 7%, it is recommended that urban vegetation should be considered as an effective adaptation measure for facing the growing cooling demands.

Author Keywords
Future typical meteorological year data; building construction characteristics; sensitivity analysis; residential building energy consumption.

ACM Classification Keywords
I.6.1 SIMULATION AND MODELING

1 INTRODUCTION
In the US, residential buildings accounted for more than 21% of total energy consumption, 36% of total electricity use, and 19% of total greenhouse gas (GHG) emissions in 2018 [1]. Moreover, energy consumption from residential buildings is projected to increase by a national average of 0.1% per year for the period of 2018–2050 under a business as usual scenario [2]. Thus, the building energy sector, in general—and the residential building stock in particular—represents a significant opportunity for accelerating the energy transition

and ensuring a low-carbon future [3]. Moreover, buildings as part of the infrastructure will need to withstand changing climatic conditions for long timespans (50–100 years) [4]. This requires current and future building stocks to perform satisfactorily under changing climatic conditions [4]. Consequently, the prediction of buildings' energy use via simulation tools, both current and future, is highly important. These simulation tools commonly utilize a combination of a building model and a weather file to account for the impact of climate on the aforementioned building [5]. The building model itself consists of building design and construction characteristics as well as energy-related user behavior and operational inputs [5].

This study conducted a sensitivity analysis on the input parameters that affect the annual energy consumption loads of an existing residential neighborhood in the U.S. Midwest over the anticipated service life of its buildings. These parameters are related to both the building model and the climate-related input. The goal of this sensitivity analysis was to help the researchers and community stakeholders understand the relative influence of each set of input parameters on the annual energy consumption loads of the selected case study. This analysis is not only based on the current state but is also considering the possible changes in consumption that can be anticipated over the buildings' entire service lives.

It is important to note that the majority of past simulation efforts to predict building's energy use, only use typical climate conditions based on climate data of the recent past as their input. The problem with the use of such weather files as input for the simulation is that typical climate conditions for the 20th century do not adequately describe the potential extreme conditions that will be encountered over the lifetime of buildings constructed today or those existing [6]. Thus, in recent years, a growing number of other studies have also tried to address this gap in the academic literature by attempting to understand the impact of future climate on energy performance predictions for risk management [6]. For instance, Crawley (2008) found that climate change would substantially influence buildings' energy performance in different climate zones. The author concluded that unless more comprehensive and accelerated changes for building design and operation are initiated, "building owners will

experience substantial operating cost increases and possible disruptions in an already strained energy supply system" [7]. Similarly, Kalvelage et al. (2014) predicted reductions in heating demand in contrast to increases in cooling demand and concluded that the resulting mixture of overall increase and decrease in energy demand for a future climate depends on the location considered and the energy source available [6]. Therefore, the identification of the building and system characteristics that have the most impact on energy demand, can help building owners in different locations make more informed investment decisions for future retrofits. Moreover, the value derived from reducing energy and operating costs is a decision that affects not only the building's performance, but the occupants' health, safety and welfare [6]. This perceived value is even greater for low-income households in urban areas, such as the population of the selected case study, who are already facing a high energy burden. These households need to allocate a disproportionate share of their income to energy expenditures due to energy inefficiencies in their homes [8], [9]. As Jagani et al. (2017) state "the existing residential building stock in inner urban neighborhoods is often not well equipped for the climatic challenges" and their energy inefficiencies are typically related to "little insulation, older windows, and leaky envelopes" [10]. Therefore, identification of building characteristics that have the most impact on energy demand can help low-income households become more resilient in the face of upcoming climatic challenges.

2 METHODOLOGY

In the previous section, it was stated that this study conducted a sensitivity analysis on the input parameters that affect the annual energy consumption loads of an existing low-income residential neighborhood in the U.S. Midwest over the anticipated service life of its buildings. In the upcoming sections of this manuscript, first, the general characteristics of the selected case study and the development procedure for the building model is discussed in detail. Then, the parameter variables used in the sensitivity analysis, including those related to the climate as well as the ones associated with the building model, are defined. Finally, the results of the sensitivity analysis are provided and conclusions are made based on these results.

2.1 Case Study

This study focuses on the energy use simulation of an inner-city neighborhood in the Midwest of the U.S. that is identified to be predominantly residential and low-income [11]. This neighborhood was selected primarily due to its social and economic characteristics which were in line with the previously mentioned goals of this study.

2.2 Building Model Related Parameters

The building model used for this study was developed in a Rhinoceros-based urban modeling design tool called Urban Modeling Interface (*umi*) [12]. *Umi* is able to efficiently model multiple buildings, approximate microclimatic effects and consider multiple sustainable performance metrics and is therefore suitable for this type of study [13]. Figure 1

below shows an overview of the developed neighborhood model in the *umi* environment.
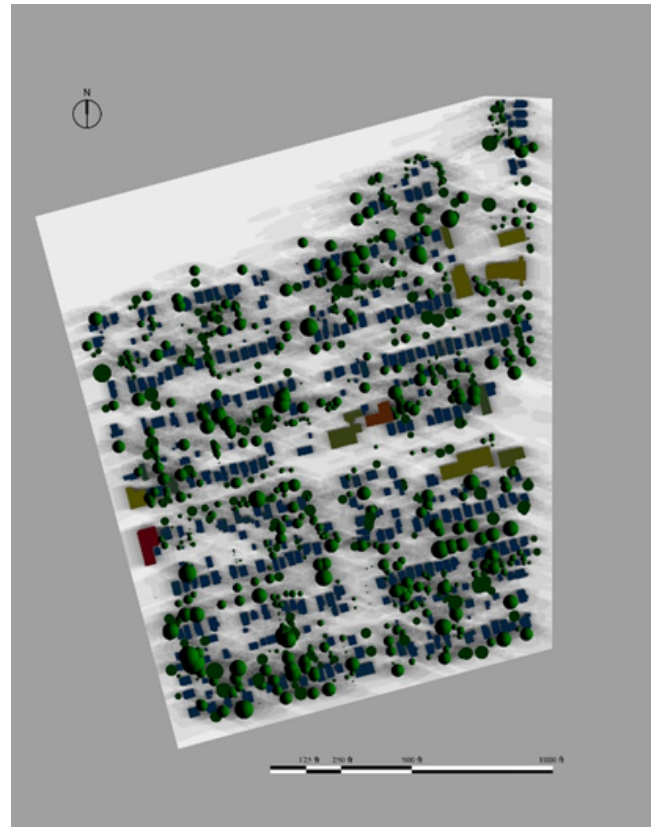

Figure 1. Top view of the Capitol East neighborhood as modeled in the *umi* environment [14].

This neighborhood model consisted of 340 buildings (323 residential and 17 nonresidential) and 1,142 trees of eight canopy shape categories. The buildings were modeled with the help of obtained GIS shapefiles for Capitol East Neighborhood's buildings which included information on buildings' footprints, elevations, and parcels. The Polk County assessor data provided another layer of information for the model which provided each building's address, parcel number, number of building stories, date of construction, number of separate residences contained within, information about the type of construction materials used, and type of occupancy [10], [15]. Accordingly, a total number of 14 building templates, which each included a set of construction material definitions and schedules, were defined in the *umi* template library and assigned to the modeled buildings based on the information provided by the Polk County assessors. Since previous work by the authors has already investigated the effect of more representative and sophisticated occupancy schedules on the model (and can be found in Malekpour Kouapei et al. (2019a) and Malekpour Koupaei et al. (2019b) [14], [16]). In addition the prediction of future occupancy profiles throughout the entire service life of buildings is quite difficult, if not impossible. Thus, for the specific purposes of this study, the schedules in all the defined templates are based on the American Society of

Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) 90.1 standard for residential buildings [17]. It is important to note that the only difference taken into account is if a specific template represents buildings that are air-conditioned (AC) during the cooling period of the year or are naturally-ventilated (NV) instead. Therefore, different building templates in this study basically represent different building construction properties and hereafter are referred to as "building construction templates" which include information on the construction properties of exterior wall, roof, ground floor, internal floor, external floor, basement wall, glazing and window to wall ratio, partition, thermal mass type and ratio, and the availability of air-conditioning systems [18].

As for the outdoor vegetation modeled in the *umi* environment, tree data was collected for 1,142 neighborhood trees during the summer of 2017 and this information was catalogued using a Trimble Geo 7X Handheld GNSS receiver [19]. The data collected in this step, which included tree species, trunk diameter, tree height, canopy shape/height, canopy width in two dimensions, and latitude/longitude coordinates, was categorized into eight canopy shapes as follows: (1) spheres, (2) ellipsoids, (3) cylinders, (4) cones, (5) horizontal rectangular cuboids, (6) vertical rectangular cuboids, (7) umbrella shapes, and (8) paraboloids [19]. Figure 2 illustrates these different tree shape types and more information on the development procedure of these categories can be found in Hashemi et al. (2018) [19].
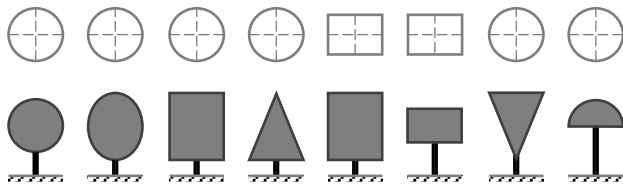
Figure 2. The eight representative canopy shapes that are used to represent trees in the model [19].

The parameters related to the building model that are studied are as following:

- The sensitivity of the model to different building construction templates is studied in detail. Of the 14 building construction templates defined in the *umi* template library, 6 can be considered AC while 8 are only NV. The analysis presented in the results section of this manuscript takes these differences into account and could help authorities identify the low-income households that are currently facing the greatest energy burden based on their general housing characteristics.

- The sensitivity of the model to the availability of trees in the neighborhood is also studied. This analysis can be considered as a level of detailing analysis that shows how much the inclusion of outdoor vegetation in the model can impact the annual simulation results.

2.3 Climate Data Related Parameters

In the previous sections, it was stated that energy simulation tools combine the building model with a weather file to study the dynamic interaction between building systems and external climate [4]. According to Bhandari et al. (2012), "There are three main classes of weather data with traditional use cases for each: "typical" weather data (representative of some location over an arbitrary period of time) often used for design and performance conditions over the life of a building, "actual" weather data (at a specific location for a specific period of time) used for simulation calibration to energy bills, and "future" weather data used for adaptive control of a building" [20]. For each class, there are a multitude of representative weather datasets that can be used depending on the purpose, location, and simulation engine that is being used [20]. To represent and compare all three of the proposed dataset types, in this study the following five weather datasets are used in the reported sensitivity analysis:

(1) A typical weather data file in the Typical Meteorological Year (TMY3) format for the Des Moines International Airport that consists of 12 typical meteorological months (January through December), with individual months selected from different years of the period of record (1991-2005) [21]. This dataset is obtained from the official EnergyPlus website [22].

(2) An actual weather file for the year 2017 in the selected location (41.53° N, 93.65° W) that is obtained from the National Solar Radiation Database (NSRDB) and formatted according to the TMY3 manual [23], [24]. From hereafter, this dataset is referred to as "Actual Meteorological Year" or "ACM".

(3-5) Three future weather files are used for the simulation of the future energy consumption of the residential building stock. These Future Typical Meteorological (FTMY) datasets were prepared by Patton (2013) who combined the projected changes in climate with existing TMY3 data to create FTMY datasets to represent high, medium and low emission scenarios of the FTMY for the 2041–2070 period [25]. In this manuscript, these three datasets are referred to as "FTMY-High", "FTMY-Medium", and "FTMY-Low" respectively.

The findings provided by this sensitivity analysis could help authorities identify the low-income households that are likely to suffer the greatest from changes in the regional climate over the projection period and take appropriate steps accordingly. The inclusion of the ACM dataset gives the researchers a chance to evaluate the current state of energy consumption in the neighborhood realistically. The results, thus, are a response to the concerns that have previously been expressed by professionals in the field about the reliability of the TMY data [26], [27].

3 RESULTS AND ANALYSIS

In total, 140 simulation runs of the model with different possible combinations of the variable parameters were run for this study. Table 1 shows an overview of the parameters studied and the total number of all the possible inputs for each.

	Building Model		Climate Data
Variable Parameter	Trees	Building Construction Templates	Weather Datasets
Number of Possible Values	2	14	5

Table 1. Defined parameters and their total number of all possible inputs.

Earlier, it was mentioned that a total number of 14 building construction templates were defined in the *umi* template library based on the assessors' data. In the baseline scenario, trees are included as shading geometry in the model and TMY3 data for Des Moines International Airport is used as the weather database. The impact of changes in the building construction templates in the neighborhood was assessed by comparing the heating and cooling loads in terms of their annual Energy Use Intensities (EUI)s. EUI, which is simply the annual energy consumption divided by the area of the building, is one of the widely used energy benchmarking and comparison methods as it is simple, and easy to compute and interpret [28], [29]. It should be noted that since the heating and cooling energy sources in this region are not the same (generally houses are cooled by electricity, while their heating energy is most commonly provided by natural gas instead [30]), an analysis of changes in the total annual energy use is less meaningful and thus avoided. Figure 3 shows the results of the 14 simulation runs for this baseline scenario where in each simulation, one of the 14 defined templates is assigned to all the residential buildings in the model. Out of these 14 building construction templates, 6 corresponded to air-conditioned buildings while the other 8 were representative of naturally-ventilated ones.

It can be seen that overall, changes in building construction characteristics can cause the annual heating loads to fluctuate by more than 10 kWh per m^2. Considering that the mean normalized annual heating load between all of these 14 simulations is 81.6 kWh per m^2, this 10.3 kWh per m^2 difference can be translated into a 13% increase in heating expenses for the households that live in less insulated and more leaky houses (for instance those represented by the AC-5 template) when compared to the more energy-efficient houses in the same neighborhood (for instance those represented by the NV-7 and NV-8 templates). Those designing and implementing weatherization assistance programs can benefit from this analysis and target the most vulnerable housing groups for maximum benefits.
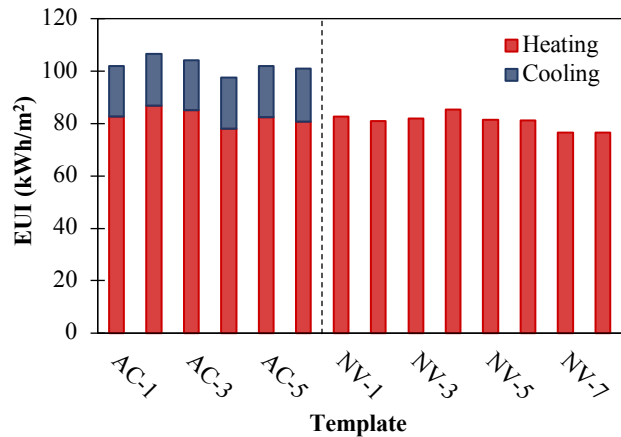
Figure 3. Normalized annual heating and cooling loads in the baseline scenario.

The changes in cooling loads, however, were more subtle and the comparative analysis showed a 1.2 kWh per m^2 range between the most and least energy-efficient templates defined. This means that, considering the flat-rate based utility billing scenario, the most efficient houses in the neighborhood (for instance those represented by the AC-1 template) only use 6% less energy for cooling expenses when compared to the less efficient ones (for instance those represented by the AC-6 template). Regardless, future work should investigate the correlation between the level of insulation and infiltration rates with both annual heating and cooling loads to determine the most optimum characteristics for houses to be built or even retrofitted in this neighborhood or those in similar settings.

In the next set of simulation efforts, the impact of the inclusion of trees in the model is assessed by removing the trees from the model developed in the baseline scenario. As can be seen in Figure 4, this change resulted in an average of nearly 7% increase (between 0.9 to 1.6 kWh per m^2) in annual cooling loads in all 6 of the models that represent air-conditioned scenarios (templates AC-1 to AC-6 assigned to all the residential buildings in the model). These results are in line with the findings of previous studies that had linked urban greening with a reduction in building cooling loads due to shading and evapotranspiration effects [19], [31]. The reduction witnessed in the results points to the importance of this level of detailing in building energy performance practices in the urban scale (nearly as important as assigning the right set of building construction characteristics) and suggests that the inclusion of urban vegetation in the model can profoundly increase the accuracy of the predictions. Moreover, these results are in line with the previous findings that suggest, to combat climate change and face the predicted hotter and longer cooling seasons in the future, more urban vegetation is recommended in the residential neighborhoods in similar climates [32], [33]. Another important implication of these results is that as more houses become equipped with air-conditioning units, this measure can help minimize the increases in electricity demand over the cooling months of the year in the future [34].
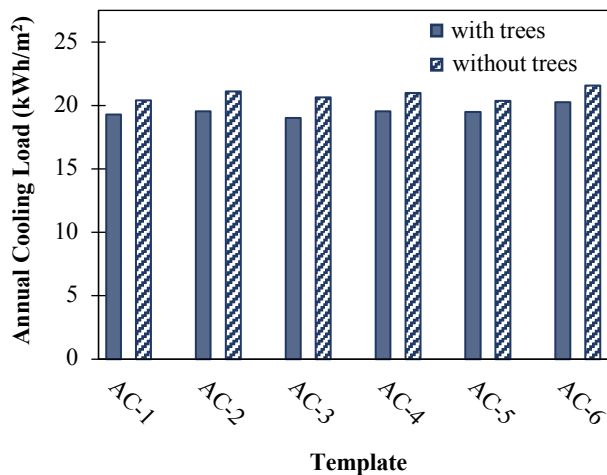
Figure 4. Decrease in annual cooling loads due to the inclusion of trees in the model.

The last set of simulation efforts in this study, focus on the sensitivity analysis of the baseline model with regards to different weather datasets. These datasets correspond to multiple phases within a typical service life of a residential building (the period for which a building is actually in use [35]). In these simulations, the appropriate building construction templates, based on the assessors' data, are assigned to all the residential buildings modeled. As can be seen in Figure 5 (a-b), heating loads for all residential buildings (AC and NV alike) are predicted to reduce by 10.9-16.1 kWh per m^2 in the years 2041-2070. The intensity of this reduction, however, is directly linked to the intensity of changes in the climate. Accordingly, the highest climate change impact scenario causes the highest reductions in heating loads. Moreover, as previous studies had also suggested, it is evident that the impacts of climate change on heating loads can already be seen in the data that represents the actual energy consumption in the year 2017 [30]. As Figure 5 (a-b) shows, heating loads in 2017 were about 7.3 kWh per m^2 less than the baseline scenario that represents the typical meteorological year in this location.

On the other hand, the cooling loads are expected to increase rather sharply. An increase of more than 7.7 kWh per m^2 compared to the baseline scenario (19.5 kWh per m^2) is predicted for all three future weather scenarios. This means that the current energy demand for cooling in this neighborhood will increase by nearly 40% in the next five decades to come. Specifically, this increase can be highly problematic as more residential houses are being equipped with air-conditioning units [34]. These findings suggest that energy-efficiency and resiliency measures for reducing current and future cooling demands are and will be of great importance in this climatic region and the largest energy cost for maintaining desired levels of health and comfort in the future at these locations will be attributed to managing higher ambient humidity levels [6], [36].
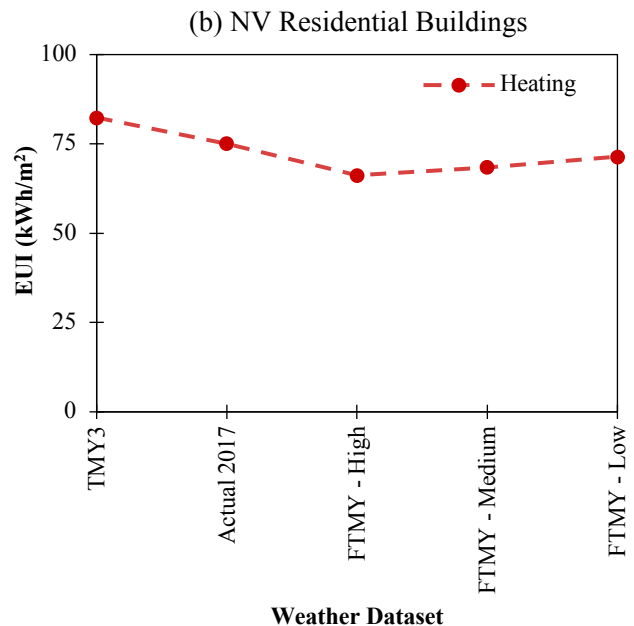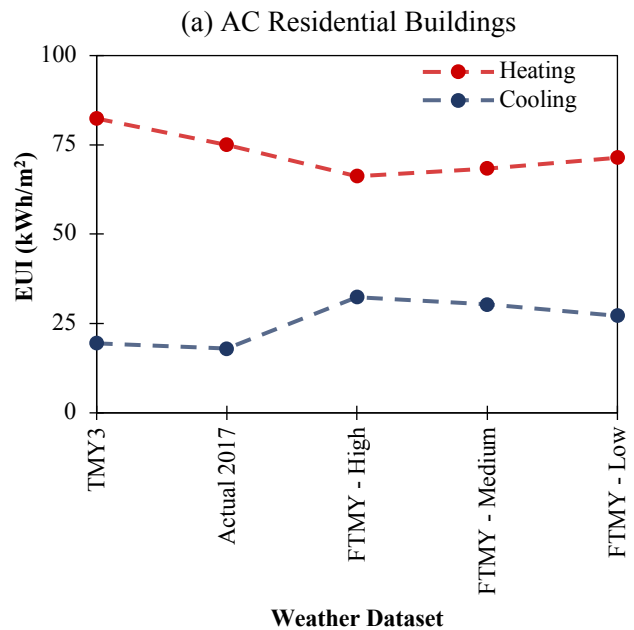
Figure 5. Lifetime energy load predictions for residential buildings.

4   CONCLUSION

In this study, the sensitivity of energy simulation results of a residential neighborhood to a varied set of simulation parameters is investigated. The parameters studied include building construction templates, outdoor vegetation, and alternative weather databases. The findings suggest that the use of different building construction packages in modelling can cause changes in heating loads as high as 13%. In contrast, while changes in climatic conditions are expected to have a profound impact (an increase of nearly 40%) on the cooling demand over the summer months, cooling load

589

calculations seem to be relatively less reliant on building construction templates. This is consistent with the findings of previous studies on the impact of climate change on the future energy loads of residences in the Midwest U.S., that had suggested the largest energy cost for maintaining desired levels of health and comfort in the future at these locations will be attributed to managing higher ambient humidity levels [6], [36]. The inclusion of trees in the energy simulation model was also found to significantly influence the results of the model and resulted in a 7% decrease in annual cooling loads for air-conditioned houses.

Current limitations of the presented work are that the findings have yet to be validated with actual metered energy consumption data. Future work will use aggregated energy use data (by zip code) provided by the utility companies involved in the region to address this shortcoming. Moreover, since the results presented do not account for the impact of longer and/or shorter heating or cooling periods possible in the future, another set of limitations are related to the identification and use of the updated cooling and heating periods based on the FTMY data. Future work should also account for the leaf shedding seasons and the effect of leaf loss on the shading properties of the modeled trees that are identified as deciduous throughout the year.

ACKNOWLEDGMENTS

REFERENCES

1. U.S. Energy Information Administration (EIA), "November 2019 Monthly Energy Review," 2019.

2. U.S. Energy Information Administration (EIA), "Annual Energy Outlook 2019 (with projections to 2050)," 2019.

3. Y. Zhang, X. Bai, F. P. Mills, and J. C. V Pezzey, "Rethinking the role of occupant behavior in building energy performance: A review," *Energy Build.*, vol. 172, pp. 279–294, 2018.

4. L. Guan, "Preparation of future weather data to study the impact of climate change on buildings," *Build. Environ.*, vol. 44, no. 4, pp. 793–800, 2009.

5. H. Yoshino, T. Hong, and N. Nord, "IEA EBC annex 53: Total energy use in buildings—Analysis and evaluation methods," *Energy Build.*, vol. 152, pp. 124–136, 2017.

6. K. Kalvelage, U. Passe, S. Rabideau, and E. S. Takle, "Changing climate: The effects on energy demand and human comfort," *Energy Build.*, vol. 76, pp. 373–380, 2014.

7. D. B. Crawley, "Estimating the impacts of climate change and urbanization on building performance," *J. Build. Perform. Simul.*, vol. 1, no. 2, pp. 91–115, 2008.

8. D. Hernández, "Energy insecurity: a framework for understanding energy, the built environment, and health among vulnerable populations in the context of climate change." American Public Health Association, 2013.

9. D. Hernández and S. Bird, "Energy burden and the need for integrated low-income housing and energy policy," *Poverty & public policy*, vol. 2, no. 4, pp. 5–25, 2010.

10. C. Jagani and U. Passe, "Simulation-based sensitivity analysis of future climate scenario impact on residential weatherization initiatives in the US Midwest," in *Proceedings of the Symposium on Simulation for Architecture and Urban Design*, 2017, p. 38.

11. Iowa State University Planning Team, "Capitol East Neighborhood Plan 2014," 2014.

12. MIT Sustainable Design Lab, "Urban Modeling Interface (umi)." .

13. C. Reinhart, T. Dogan, J. A. Jakubiec, T. Rakha, and A. Sang, "Umi-an urban simulation environment for building energy use, daylighting and walkability," in *13th Conference of International Building Performance Simulation Association, Chambery, France*, 2013, vol. 1.

14. D. Malekpour Koupaei, F. Hashemi, V. Tabard-Fortecoëf, and U. Passe, "Development Of A Modeling Framework For Refined Residential Occupancy Schedules In An Urban Energy Model," in *Building Simulation 2019: 16th Conference of IBPSA*, 2019.

15. Polk County Administration, "Polk County Assessor." .

16. D. Malekpour Koupaei, F. Hashemi, V. Tabard-Fortecoëf, and U. Passe, "A Technique for Developing High-Resolution Residential Occupancy Schedules for Urban Energy Models," in *The Symposium on Simulation for Architecture and Urban Design (SimAUD)*, 2019, pp. 95–102.

17. ASHRAE, "ASHRAE Standard 90.1.1989 Energy Efficient Design of New Buildings Except Low-Rise Residential Buildings," 1989.

18. C. Cerezo, T. Dogan, and C. Reinhart, "Towards standardized building properties template files for early design energy model generation," in *Proceedings of the ASHRAE/IBPSA-USA building simulation conference*, 2014, pp. 25–32.

19. F. Hashemi, B. Marmur, U. Passe, and J. Thompson, "Developing a workflow to integrate tree inventory data into urban energy models," in *Proceedings of the Symposium on Simulation for Architecture and Urban Design*, 2018, p. 34.

20. M. Bhandari, S. Shrestha, and J. New, "Evaluation of weather datasets for building energy simulation,"

Energy Build., vol. 49, pp. 109–118, 2012.

21. S. L. Rabideau, U. Passe, and E. S. Takle, "Exploring alternatives to the" typical meteorological year" for incorporating climate change into building design," *ASHRAE Trans.*, vol. 118, no. 1, p. 384, 2012.

22. EnergyPlus, "Weather Data Download - Des Moines Intl AP 725460 (TMY3)." .

23. S. Wilcox and W. Marion, "Users manual for TMY3 data sets," 2008.

24. U.S. Energy Information Administration (EIA), "NSRDB Data Viewer." .

25. S. Patton, "Development of a future typical meteorological year with application to building energy use," 2013.

26. D. B. Crawley and L. K. Lawrie, "Rethinking the TMY: is the 'typical'meteorological year best for building performance simulation?," in *Conference: Building Simulation*, 2015.

27. G. Pernigotto, A. Prada, and A. Gasparella, "Development of Extreme Reference Years for Building Energy Simulation Scenarios," in *Applied Mechanics and Materials*, 2019, vol. 887, pp. 129–139.

28. J. Y. Park, X. Yang, C. Miller, P. Arjunan, and Z. Nagy, "Apples or oranges? Identification of fundamental load shape profiles for benchmarking buildings using a large and diverse dataset," *Appl. Energy*, vol. 236, pp. 1280–1295, 2019.

29. C. Filippın, "Benchmarking the energy efficiency and greenhouse gases emissions of school buildings in central Argentina," *Build. Environ.*, vol. 35, no. 5, pp. 407–414, 2000.

30. R. J. Brecha, A. Mitchell, K. Hallinan, and K. Kissock, "Prioritizing investment in residential energy efficiency and renewable energy—A case study for the US Midwest," *Energy Policy*, vol. 39, no. 5, pp. 2982–2992, 2011.

31. C.-M. Hsieh, J.-J. Li, L. Zhang, and B. Schwegler, "Effects of tree shading and transpiration on building cooling energy use," *Energy Build.*, vol. 159, pp. 382–397, 2018.

32. D. J. Nowak, N. Appleton, A. Ellis, and E. Greenfield, "Residential building energy conservation and avoided power plant emissions by urban and community trees in the United States," *Urban For. urban Green.*, vol. 21, pp. 158–165, 2017.

33. U.S. Environmental Protection Agency (EPA), "Climate Impacts in the Midwest." .

34. U.S. Energy Information Administration (EIA), "Air conditioning in nearly 100 million U.S. homes," 2011. .

35. A. Rauf and R. H. Crawford, "Building service life and its effect on the life cycle embodied energy of buildings," *Energy*, vol. 79, pp. 140–148, 2015.

36. B. Mann, U. Passe, S. Rabideau, and E. S. Takle, "Future context for thermal comfort: Impact of a changing climate on energy demand and human thermal comfort," 2012.

Organizing Committee

Angelos Chronis

General Chair

Angelos is the head of the City Intelligence Lab at the Austrian Institute of Technology and teaches at the Institute of Advanced Architecture of Catalonia. His main research interests lie in integrating simulation, optimization and performance drive in the design and fabrication process, with particular expertise in computational fluid dynamics (CFD).

Gabriel Wurzer

Program Chair

Gabriel researches and teaches early-stage planning of complex buildings (especially hospitals) using agent-based simulation at TU Wien. He also the developer of ProceeDings, a web-based editing system used since 2014 e.g. by eCAADe, SiGrADI, CAADRIA and ARENA.

Wolfgang E. Lorenz

Workshops Coordinator

Wolfgang is a researcher and lecturer at TU Wien (Digital Architecture and Planning). His research spans a wide range of topics ranging from fractal analysis to visualization of different options in urban zoning.

Scientific Chairs

Christiane M. Herr

Scientific Chair

Christiane is a researcher and educator at Xi'an Jiaotong-Liverpool University focusing on structural design, digitally supported design, conceptual design, design studio teaching and traditional Chinese approaches to creative thinking.

Ulrich Pont

Scientific Chair

Ulrich is researching and lecturing in architecture and building science at TU Wien, focusing especially on building performance simulation, building ecology, building construction, building informatics, and building retrofit.

Dana Cupkova

Scientific Chair

Dana is a co-founder and a design director of EPIPHYTE Lab, an interdisciplinary architectural design and research collaborative. She is also assistant professor at CMU where she serves as Track Chair for the Master of Science in Sustainable Design (MSSD).

Gabriel Wainer

Scientific Chair

Gabriel is professor and associate chair for graduate studies at Carleton University. He is well known for his research in Discrete-Event Simulation as well as Real-Time Systems. He is member of numerous boards and steering committees, including SCS and SimAUD.

International Scientific Committee

Haşim	Altan	Final International University in Cyprus
Mohamed	Aly	Rensselaer Polytechnic Institute
Armin	Amirazar	UNC Charlotte
Spyridon	Ampanavos	Harvard GSD
Alpha Yacob	Arsano	MIT
Arianna A	Stolfi	Politecnico di Milano
Shady	Attia	University of Liège
Amber	Bartosh	Syracuse University
Johannes	Braumann	Robots in Architecture
Michael	Budig	SUTD singapore
Giorgia	Chinazzo	EPFL
Joon-Ho	Choi	University of Southern California
Joseph	Choma	Clemson University
Ruchi	Choudhary	University of Cambridge
Angelos	Chronis	Austrian Institute of Technology
Salmaan	Craig	McGill University
Drury	Crawley	Bentley
Dana	Cupkova	Carnegie Mellon University
Daniel	Davis	University of Hartford
Francesco	De Luca	TalTech
Timur	Dogan	Cornell University
Max	Dölling	University of Texas at San Antonio
Bing	Dong	University of Texas at San Antonio
Elif	Erdine	Architectural Association
Pau	Fonseca	UPC, Spain
Tomohiro	Fukuda	Osaka University
Jose Luis	Garcia Del Castillo	Harvard GSD
Jeff	Geisinger	Utile
David	Gerber	USC
Marjan	Ghobad	University of Cape Town
Betti	Giovanni	HENN Architects
Rhys	Goldstein	Autodesk
Apoorv	Goyal	HOK
Jonathan	Grinham	Harvard GSD
Kenny	Gruchalla	NREL
Chenghe	Guan	Harvard GSD
Burak	Gunay	Carleton University

Navid	Hatefnia	TU Munich
Mohammad	Heidarinejad	Illinois Institute of Technology
Ryan Luke	Johns	Greyshed/Columbia/Princeton
Nathaniel	Jones	ARUP
Alexandros	Kallegias	Architectural Association/Zaha Hadid
Alireza	Karduni	UNC Charlotte
Reinhard	Koenig	Bauhaus-University Weimar
Odysseas	Kontovourkis	University of Cyprus UCY
Anas	Lila	Cardiff University
Thorsten	Loemker	Payette Architects
Wolfgang	Lorenz	Tu Wien
Christiane	M Herr	Xi'an Jiaotong-Liverpool University
Ed	Manley	UCL Bartlett
James	McNeill	AEI
Clayton	Miller	NUS
Shreshth	Nagpal	MIT
Taro	Narahara	NJIT
Liam	O'Brien	Carleton University
Konstantinos-Alketas	Oungrinis	Technical University of Crete
Bertug	Ozarisoy	University of East London
Mine	Ozkar	Istanbul Technical University
Krista	Palen	Transsolar NYC
Dimitris	Papanikolaou	UNC Charlotte
Brady	Peters	University of Toronto
Ulrich	Pont	Vienna University of Technology
Vinu Subashini	Rajus	Carleton
Siobhan	Rockcastle	University of Oregon
Cristina	Ruiz Martin	Carleton
Azadeh	Sawyer	University of Michigan
Davide	Schaumann	Technion
Sven	Schneider	Bauhaus-University Weimar
Matthias	Schuss	TU Wien
Stefano	Shiavon	UC Berkeley
Ondrej	Sikula	TU Brno
Jelena	Srebric	University of Maryland
Rudi	Stouffs	National University of Singapore
Spyros	Stravoravdis	University of Greenwich
John	Sullivan	Spagnolo Gisness & Associates
Farhang	Tahmasebi	UCL
Martin	Tamke	Royal Danish Academy of Fine Arts

CONFERENCE PROCEEDINGS

Walid	Tizani	University of Nottingham
Irmak	Turan	MIT
Michela	Turrin	TU Delft
Thanos	Tzempelikos	Purdue
Peter	Von Buelow	University of Michigan
Milena	Vuckovic	Austrian Institute of Technology
Christoph	Waibel	ETH Zurich
Gabriel	Wainer	Carleron University
Karen	Walkerman	Second Law
Gabriel	Wurzer	Vienna University of Technology
Arta	Yazdanseta	Parsons
Tea	Zakula	University of Zagreb
Andrzej	Zarzycki	NJIT
Matias	Del Campo	University of Michigan
Robert	Cox	UNC Charlotte
Greg	Schleusner	HOK

Student Volunteers

Dominik	Fill	TU Wien / Architecture
Myrto	Karnoutsou	TU Wien / Architecture
Androniki	Pappa	UCL / Bartlett
Karim Yosef	Rezk	TU Wien / Architecture

Organizational Support

Oletha	Darensburgh	SCS
Carmen	Ramirez	SCS
Bianca	Braun	TU Wien
Peter	Kompatscher	TU Wien
Harald	Klingenstein	Stadt Wien
Shaber	Hossain	ZOOM
Tobias	Judmaier	Iss mich!
Ulli	Zeiler	TIP TOP TABLE Catering
Jela	Babic	Mensen
Andrea	Eisbach	Mensen
Joana	Haiduck	Rathauskeller
Birgit	Zehetmaier	Rathauskeller
Anastasia	Gassia	Efpalinos Publishing

Sponsors

Exclusively sponsored by

In cooperation with

SCS	The Society for Modeling & Simulation International
acm	The Association for Computing Machinery
	The International Building Performance Simulation Association
COCOS	The Center for Computational Complex Systems at TU Wien
dap	Research Unit of Digital Architecture and Planning at TU Wien
[ef] Efpalinos Publishing	Efpalinos Publishing

CONFERENCE PROCEEDINGS

Index of Authors

Symposium on Simulation for Architecture and Urban Design
2020