Simulation

Methodology

Simulation: Transactions of the Society for Modeling and Simulation International
1–21

DOI: 10.1177/00375497231209996
journals.sagepub.com/home/sim
Sage SCS

A data-driven approach of layout evaluation for electric vehicle charging infrastructure using agent-based simulation and GIS

Yue Zhang and Jie Tan

Abstract

The development and popularization of new energy vehicles have become a global consensus. The shortage and unreasonable layout of electric vehicle charging infrastructure (EVCI) have severely restricted the development of electric vehicles. In the literature, many methods can be used to optimize the layout of charging stations (CSs) for producing good layout designs. However, more realistic evaluation and validation should be used to assess and validate these layout options. This study suggested an agent-based simulation (ABS) model to evaluate the layout designs of EVCI and simulate the driving and charging behaviors of electric taxis (ETs). In the case study of Shenzhen, China, geographical positioning system (GPS) trajectory data were used to extract the temporal and spatial patterns of ETs, which were then used to calibrate and validate the actions of ETs in the simulation. The ABS model was developed in a geographic information system (GIS) context of an urban road network with traveling speeds of 24 h to account for the effects of traffic conditions. After the high-resolution simulation, evaluation results of the performance of EVCI and the behaviors of ETs can be provided in detail and in summary. Sensitivity analysis demonstrates the accuracy of simulation implementation and aids in understanding the effect of model parameters on system performance. Maximizing the time satisfaction of ET users and reducing the workload variance of EVCI were the two goals of a multiobjective layout optimization technique based on the Pareto frontier. The location plans for the new CS based on Pareto analysis can significantly enhance both metrics through simulation evaluation.

Keywords

GPS trajectory data, electric vehicles, charging infrastructure, layout optimization, agent-based simulation

1. Introduction

The popularization of electric vehicles (EVs) is important to upgrade the automotive industry, reduce the reliance on fossil fuels, and improve the eco-environmental quality.[1,2] The global EV sales doubled in 2021 from the previous year to a new record of 6.6 million, with China accounting for half of the growth.[3] However, the shortage and unreasonable layout of electric vehicle charging infrastructure (EVCI) have severely restricted the development of China's new energy car industry.[4] Shenzhen, as the largest-scale and widest application of pure electric buses and taxis in China, had approximately 20,000 electric taxis (ETs) in 2019, but the ratio to the EVCI is only 1:3. Some public charging stations (CSs) have serious queuing problems due to the improper layout of public EVCI, while some other stations have low utilization rates.[5,6] Other cities, such as Beijing,[7] New York,[8] and Valencia,[9] have similar problems. In the meantime, good planning and configuration of EVCI can improve charging service efficiency, increase EV users' satisfaction, reduce the distance anxiety of potential EV purchasers, and thus accelerate the popularity of the EV industry.[10,11]

The location problem of EVCI involves transportation, urban planning, energy, operation research, and other fields. The problem is of great complexity and has become a major research area in recent years. The affecting factors of the layout planning of EVCI can be classified into three

School of Management, Shenzhen Polytechnic University, China

Corresponding author:
Jie Tan, School of Management, Shenzhen Polytechnic University, No. 7098 Liuxian Avenue, Nanshan District, Shenzhen 518055, Guangdong, China.
Email: jietan@szpt.edu.cn

categories: (a) charging facility factors (e.g., private charging piles or public stations and fast or slow charging), (b) EV factors (e.g., EV types, battery types, driving and charging behaviors), and (c) external environmental factors (e.g., traffic conditions, land usage, power grids, and financial costs).[12] The EV types studied in the literature include electric buses,[13] taxies,[7] private cars,[14] and trucks,[15] while most research focuses on the planning of public fast CSs in urban environments. From the perspective of operation research, the EVCI location problem can be regarded as selecting the appropriate points in the candidate set to optimize certain objectives, which usually include one or more of the following goals: (a) reducing the overall cost (e.g., constructure cost, operation cost, electricity cost, and environmental cost),[2,16] (b) increasing EV user satisfaction (service range and distance, search time, and queuing time),[4,6,9] (c) improving EVCI utilization,[17,18] and (d) optimizing power grid performance (e.g., peak-load shifting and reduction in power losses).[19,20] With the rapid development of mobile Internet, Internet of Things and Intelligent Transportation Systems, EV trajectory, traffic, and geographic information system (GIS) data have been increasingly used in the application of mathematical optimization methodologies to EVCI location problems.[6,21,22]

The classic mathematical programming approaches may provide good EVCI layout plans with certain objectives from different perspectives. However, these algorithms might still stay at a theoretical level due to the complexity and uncertainty of this interdisciplinary problem.[23] The effectiveness of these approaches should be verified in environments close to reality (e.g., through simulation in large-scale vehicle movement scenarios).[9] Under such motivation, several agent-based simulation (ABS) models have been developed to evaluate and optimize the EVCI layout.[9,24–26] In these ABS models, potential EV purchasers, existing EV users, passengers, and fleet managers are modeled as different types of agents with a certain level of intelligence. Complex agents' behaviors, such as traveling, searching, queuing, and charging behaviors, are implemented in a GIS environment (e.g., road network, traffic conditions, and Points of interest (POI) that is close to reality in a resolution of 1 s. The EV's traveling demand is typically generated by a synthetic population, such as those who want to call ETs, current private EV users who have to commute or go shopping, or those who need to deliver goods. Then, the EV movements and charging behaviors are simulated in the GIS environment. Detailed information for each traveling and charging behavior is recorded, and system performance metrics such as EVCI utilization, operating cost, and EV user satisfaction can be calculated after the simulation ends.

These ABS models provide an effective way to access EVCI layout designs. However, the generation of EV movements and corresponding charging demand involves the implementation of some subprocedures with many assumptions and parameters. For instance, in the ABS model proposed by Jordan et al.,[9] the dispatchment of ETs is based on the negotiation process between passengers and ET drivers. Model parameters of the negotiation process, such as agents' different types of negotiation strategies and proportions of different subpopulations, have to be estimated and calibrated very well, or the generated traveling and charging demand might not reflect reality. In addition, each trip's origin and destination (OD) are determined by a weighted score of model factors that are related to the population, traffic, and social network activities of urban areas, with the weights being configured in a subjective manner. Again, these model parameters must be carefully chosen to reflect the actual circumstances.

On the basis of the abovementioned discussion, this study focuses on the layout design of public fast CSs for ETs in urban environments. An ABS model is developed in which geographical positioning system (GPS) trajectory data are used to simulate and validate the traveling and charging behaviors of ETs. The ABS model can be used for the evaluation and optimization of EVCI layout designs, with ET drivers' satisfaction and EVCI utilization rates as the main measures of EVCI layout designs. Compared with existing ABS studies, the main innovations and contributions of this work are as follows:

(a) Instead of using synthetic population and agents' negotiation to generate ET traveling behaviors and charging demand, the ET trajectory data are used to extract the spatial and temporal pattern of ET movements (e.g., OD transition probabilities for ETs' idle and busy states, interarrival time distribution, and passenger-pickup frequencies during 24 h). Based on the pattern extracted, the traveling trips and charging demand can be generated in a more direct and accurate manner.

(b) In the ABS model, the ETs search the CSs in a smarter way. Rather than simply choosing the nearest CS, EV drivers first check the availability of CS nearby (e.g., 10 km) with the help of mobile phone applications and then attempt to choose the nearest station with free charging piles. If none of the stations nearby have free piles, then the driver will head for the station with the shortest queue to reduce the waiting time. Therefore, the ET agents can make decisions based on the real-time environment, and their decisions can interact with others in this complex adaptive system. This implementation is more in line with the actual drivers' habits, which makes the queuing statistics generated by the simulation more reliable.

(c) The model is built in a GIS environment including an urban road network with a traveling speed of 24 h for each road segment, which affects the ETs' traveling and charging activities. The dynamic patterns of traffic conditions for 24 h and ET behaviors interact with each other and are validated and calibrated based on trajectory data. These factors are considered in the layout evaluation in a way that is closer to reality.

2. Literature review

2.1. Mathematical optimization approaches

The location optimization problem is a classic operation research problem that covers many aspects of public infrastructures, such as the siting of hospitals, fire stations, and shopping malls. In the earliest study by Weber[27] of location theory, the location of a warehouse is determined by minimizing the total distance from the warehouse to customers. Classic approaches to solve location problems include p-median,[28] p-center,[29] and maximum coverage[30] methods. With the rapid development of EVs in the 1990s, the demand for locating EVCIs has been increasingly growing. Relevant approaches can be classified into methods based on point demand,[31] methods based on flow demand,[32] and mixed models.[33] Typically, EVCI layout optimization is modeled as a mathematical programming problem with objectives and constraints. After the problem is defined in mathematical form, optimization algorithms such as mixed integer programming[34] are applied to obtain optimal layout solutions. Owing to the complexity of the EVCI location optimization problem, heuristic algorithms, such as the genetic algorithm (GA)[16] and particle swarm optimization (PSO),[2] have been used to solve this NP-hard problem in recent studies. However, as mentioned above, these mathematical optimization approaches may remain at a theoretical level.[23] The effectiveness of these algorithms should be validated in environments close to reality.[9]

2.2. Trajectory data-driven approaches

Currently, GPS-enabled trajectory data have been increasingly used to study the EVCI location problem because they can efficiently describe the traveling and charging behaviors of EV drivers. Most research focused on China's cities as research zones, such as Shenzhen,[6,35,36] Wuhan,[22,37] Chengdu,[38] Tianjin,[39] and Changsha.[40] In these studies, GPS trajectory data were mainly collected from ETs,[6,35,37,38,40] electric buses,[22] and freight vehicles[36] while some were from mobile phone users.[21,39] In ET's case, the GPS trajectory data are usually collected every 5–60 s and contain data fields of vehicle ID, timestamp, longitude, latitude, speed, direction, and passenger load indicator. The operational and charging patterns of the ET can be extracted based on the trajectory data.[41] The ET trajectory can be further divided into three subtrajectories: searching, charging, and traveling based on the proposed definition by Li et al.[35] Hu et al.[6] estimated the travel time between adjacent map grids based on the traveling subtrajectory, while the charging demand was measured by counting the start points of the searching subtrajectory. These estimated variables were then used to formulate the optimization problem, and heuristic algorithms were used to maximize the ET drivers' satisfaction, which is correlated to waiting and charging times. Liu et al.[38] used ET trajectory data to identify parking and charging events, based on which ET traveling patterns were obtained. The PSO algorithm was then applied to minimize the CO_2 emissions. In summary, the model variables estimated based on GPS trajectory data are more accurate and reliable, but the solution to the mathematical optimization problem still needs careful validation.

2.3. ABS approaches

The EVCI layout design is a complex problem with many dynamic and stochastic factors involved, such as traffic conditions and EV drivers' traveling, searching, and charging behaviors. In addition, these factors and behaviors interact with each other and with the environment, which causes difficulty for mathematical equations to represent them. Therefore, the real effect of the solutions provided by optimization algorithms needs to be evaluated. However, conducting actual tests in the real world to verify the EVCI layout performance is nearly impossible due to the high social risk and financial cost. Under such circumstances, simulation, especially ABS, can reproduce real-world scenarios in virtual settings to evaluate the effectiveness and rationality of different layout schemes while reducing the cost of analysis and decision-making. However, the ABS model can capture the behaviors of agents in an environment through the use of decision rules, which govern the interactions between agents in the simulation. The commonly used ABS platforms and frameworks are SimFleet,[9,42] MATSim,[14,43] EnerPol,[24] Repast Symphony,[44–47] and AnyLogic.[7]

Jordan et al.[9] proposed a simulation-optimization approach for layout designs of ET CS. An ABS model was built to simulate the negation between ET drivers and passengers, the dispatchment of fleet managers, and ETs' traveling, searching, and charging behaviors through SimFleet in the GIS environment. A GA was applied to find the optimal EVCI layout for reducing the queuing time of ETs in CSs. The model simulated the full process of ET drivers' daily work, but the generation of pickup orders contained too many details, and the model parameters involved in the process should be set and validated with great caution. For instance, the ET movements in the simulation were generated by a weighted score of population, traffic, and social network activities, while these weights were determined in a subjective way. More focus should be placed on capturing ETs' traveling and charging patterns, which are more relevant to the EVCI location problem. GPS trajectory data would definitely help if they (it) were available. In addition, ET drivers would simply choose the nearest CS when they had charging demand in the simulation. However, in reality, these drivers might search for the station with free charging piles or select the station with the shortest queue length with the help of mobile phone applications. The implementation of such searching logic might affect the evaluation accuracy of queuing statistics such as queue length and waiting time.

Pagani et al.[24] developed an ABS model to study the EVCI layout and private EV users' behaviors given different EV penetration levels from the perspective of profitability through the EnerPol simulation platform. In the ABS model, private EVs' traveling demand was generated based on a synthetic population's daily activities, such as commuting between dwellings and workplaces and going to schools and shops. Different charging strategies, such as price- and comfort-driven charging behaviors, were implemented and compared in the study. A stepwise optimizer was applied to locate CS sequentially with the goal of minimizing the queuing time of EV users and maximizing the utilization of EVCI. Similar to Jordan's ABS model,[9] the generation of EVs' traveling demand depends on a complex process with many assumptions and parameters, which need to be validated and calibrated by comparison with the actual situation. Other ABS studies include Márquez et al.'s model[25] in a long-distance cross-city transportation scenario, Zhao and Ma's[47] model to optimize the initial CS layout from the perspective of EV diffusion, and Bischoff's ABS model[26] for ET fleets.

While the focus of this study is on the EVCI location problem, it is worth noting that the GIS-based simulation methods have found utility in other disciplines. For instance, Davidson et al.[48] developed a discrete-event spatial model to simulate disease spread within a pandemic, demonstrating its adaptability and ability to provide deterministic predictions for multiple regions simultaneously. Other studies using simulation to investigate the spread of infectious diseases can be found in the work by Abadeer et al.,[49] as well as the review article by Ayadi et al.[50] Iskandar et al.[51] developed an agent-based model that incorporates realistic human behavior and urban conditions to simulate pedestrian evacuation during earthquakes at the city scale, revealing the impact of debris and human behavior on population safety and the limited capacity of open spaces to provide shelters in Beirut, Lebanon. Risco-Martin et al.[52] introduced a discrete-event simulation for real-time monitoring and management of harmful algal and cyanobacterial blooms (HABs), addressing the need for efficient detection and response to HABs threatening water quality in dynamic environments. The application of simulation-based methodology in emergency healthcare can be seen in the review article by Sahlaoui et al.[53] Similar to the EVCI location problem in this paper, these studies use simulation methods to explore real-world issues in urban environments from temporal and spatial perspectives. In these simulation models, some studies employ deterministic mathematical formulas to represent the relationships between key factors, while others utilize agent-based approaches to simulate complex systems through rules governing interactions between agents and the environment to make predictions. These methods provide inspiration for the research on EVCI layout optimization in this paper.

3. Simulation implementation and validation

3.1. Introduction of case study

Shenzhen is one of the largest cities in China, and it covers an area of 1997.47 km^2. It was designated by the Ministry of Transport as one of the earliest pilot areas for developing powerful transportation in China. In 2019, Shenzhen had a total of 20,000 pure ETs, and this situation fully realizes the pure electrification of ET in the city.[5] However, the lagging of EVCI construction and unreasonable layout design make shortages and idleness of EVCI coexist in most regions.[5,6] The mainstream model of ETs in Shenzhen is BYD e6, which usually needs approximately 2 h to charge from zero power to full charge in 4060 kW DC fast charging piles. The urban public power grid mainly undertakes the charging function of ETs while considering the occasional charging needs of other types of EVs.[5] Approximately half of the ETs use a single-shift mode, with the operation time mainly between 9:00 and 24:00, whereas the other half of the ETs use a double-shift mode with handover times of 3:00–5:00 and 15:00–18:00 when charging peaks are more likely to occur.[5] This case study focuses on the location problem of public fast CSs for ETs in Shenzhen. An ABS model was built to simulate the traveling and charging behaviors of ETs in a GIS environment based on GPS trajectory data. The simulation model was then applied to evaluate EVCI layout performance, such as EV users' satisfaction and EVCI's utilization. The ABS model was implemented using Repast Symphony,[44] which is an open-source, agent-based modeling and simulation platform. It is an object-oriented model with a source library for the creation and running of simulations, as well as for displaying and collecting data from the simulations. Geographic data, such as the data in shapefile format, can be imported into the Repast Symphony model, which can be used to control the behaviors and movements of the agents based on rule sets that exploit these data.

The ABS model in this study was built as an extension of Malleson's RepastCity prototype,[45] and its visualization snapshot can be found in Figure 1. As shown in Figure 1,

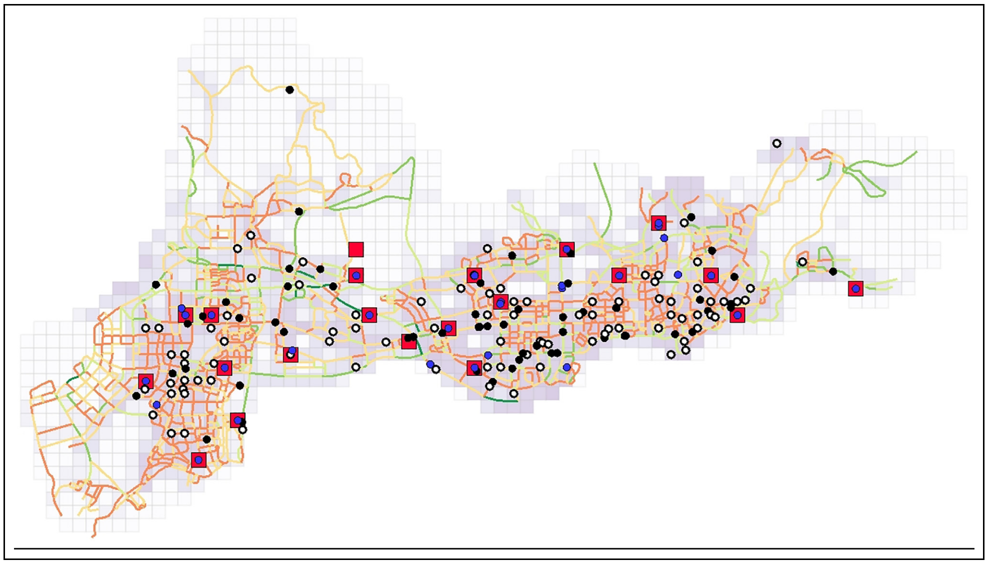

Figure 1. Visualization of agent-based simulation of electric taxis in partial districts in Shenzhen.

the study region includes the Futian, Nanshan, and Luohu districts in Shenzhen, which are divided into 1280 grids with sides of 1000 m. Roads are shown in different colors, with green color indicating normal driving and red color representing slow driving, which is probably due to traffic congestion. Circles in different colors represent ETs in different states. White and black circles indicate whether ETs are in idle or busy status, respectively, while blue circles represent ETs with charging demand. In the simulation, EVs can read the speed of the current road segment and move along it at that speed. The Dijkstra algorithm was implemented in the simulation such that ETs would select the routes that minimize the traveling time based on the OD in the road network. In this way, the traffic conditions, especially congestion's influence on ETs' moving behaviors, can be considered in the EVCI location problem. The logic of the traveling and charging behaviors of ETs can be validated through the animation of high temporal resolution in the simulation visualization.

The EVCI layout to be evaluated, which is composed of 21 public CSs, is represented by red squares in Figure 1. The exact locations of stations in Figure 1 are based on the actual situation[18] when trajectory data were collected in 2014. Based on the trajectory data of ETs, the temporal and spatial patterns of the traveling and charging behaviors are extracted and then used to generate stochastic events and movements of ETs in the simulation. The ETs consume battery power while traveling, and they search for CSs nearby based on a set of predefined rules when they demand charging. Detailed information on all the trips and charging events of ETs can be recorded in the simulation. After the simulation is completed, the layout's overall performances, such as queue statistics and EVCI utilization rates, can be obtained.

As shown in Figure 2, the input of the ABS model can be classified into GIS-related features, ET patterns extracted from trajectory data, and logic and rules based on reality, such as rules of ET searching behaviors, logic of power consumption, and charging time. Simulation output includes detailed information on each single event, such as passenger-pickup and charging events, based on which EV users' satisfaction and EVCI utilization measures can be calculated. Details of the implementation, validation, and analysis are explained in the following subsections.

3.2. Road network and traffic conditions

The road networks of the Futian, Nanshan, and Luohu districts in Shenzhen were obtained from OpenStreetMaps, which includes primary, secondary, motorway, and highway road segments. ArcGIS10.8 processing tools, such as merging, buffering, vectorization, and breaking intersection lines, were applied to the road network. In this way, multiple lanes in the road were merged, and the original complex road network was simplified, especially the entrances and exits of highways. Some errors in the shapefiles, such as unconnected road endpoints, were corrected to guarantee the connectivity of the road network. Such operations capture the major characteristics of city road networks, reduce the potential EVs' routing errors in simulation, and save computation resources for this high-resolution simulation.

The traveling speed for 24 h of each road segment was collected from the Amap (Gaode Map) API to simulate the actual traffic conditions. For each road segment, the two endpoints took turns serving as the starting and ending points of the trip, which were entered into the Amap API

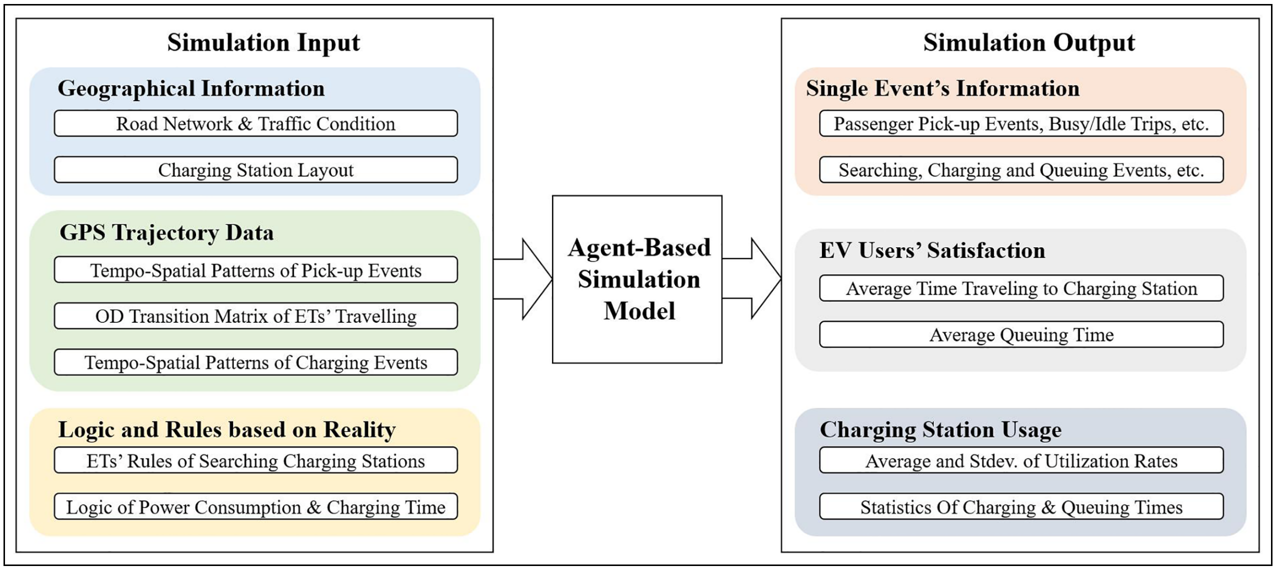

Figure 2. Input and output of the agent-based simulation model.
ET: Electric taxis

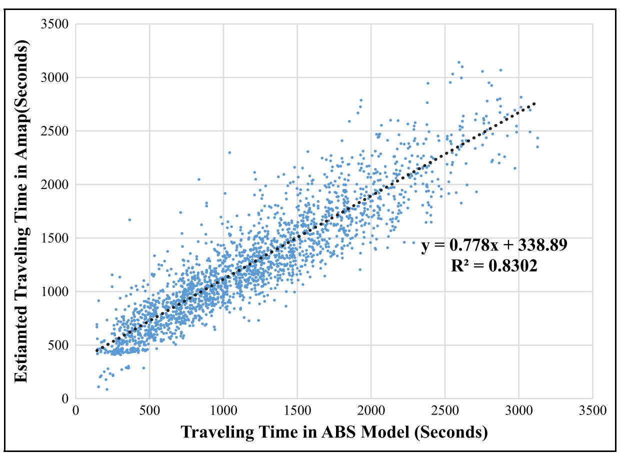

Figure 3. Comparison of traveling times between simulation and Amap estimation.
ABS: Agent-based simulation

to obtain the estimated trip time. The road speed was calculated based on the road length and the average of the two-way trip times. As shown in Figure 1, different colors of roads represent different levels of traveling speed. As the simulation processes, these speeds are updated hourly for roads based on the data collected. Vehicles read the speed of the current road and travel using that speed. A validation test was conducted based on 47 locations randomly selected from the study region to guarantee that the traveling time between any two locations in the simulation was close to the estimated traveling time in Amap. Linear regression was applied to compare 2162 trip traveling times in the simulation map and those estimated in Amap. According to the results in Figure 3, the R2 of the linear model is 0.83, which indicates a strong linear relationship between the two estimation methods. The average absolute error of traveling time is only 3.52 min. The difference might be due to the simplification of the road structure and Amap's time estimation of traffic light delays. The adjusted road structure in the ABS model gives a similar estimation of traveling times as Amap's.

3.3. GPS trajectory data

The GPS trajectory data used in this study consist of two parts: an ET dataset and a fuel taxi (FT) dataset, which are from the published dataset used in Wang et al.'s[54] and Zhang et al.'s[55] research, respectively. The ET dataset contains 1,155,654 GPS records for 664 ETs in Shenzhen in October 2014, with data fields of vehicle ID, longitude, latitude, time, and speed. After being spatially joined with map grids, 83% (964,625 records) of the original GPS records remained, which suggests that the research area of the three core districts accounted for the majority of the ET activities in Shenzhen in 2014. Based on the staying events extracted from the dataset, along with the CS locations in 2014,[41] the charging events of ETs were identified, which were used to obtain the temporal and spatial patterns of charging behaviors of ETs. Given that the data field of the passenger load indicator was unavailable in this ET trajectory dataset, the FT trajectory dataset was used as a supplement to the ET dataset. In addition to the occupancy status, all data fields in the ET dataset are also included in the FT dataset, which contains 46,927,855 GPS records of 14,728 FTs in Shenzhen in October 2013. Based on the FT dataset, the traveling patterns, such as the tempo-spatial distribution of passenger-pickup events and drop-off events, the OD transition matrixes in busy and idle status,

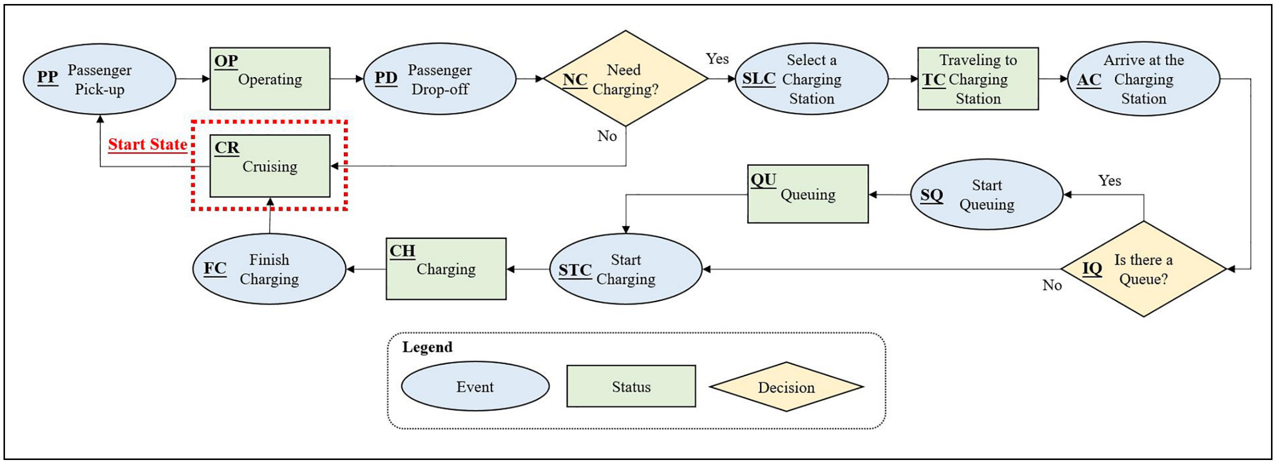

Figure 4. Flowchart of electric taxis' traveling and charging behaviors in the ABS model.

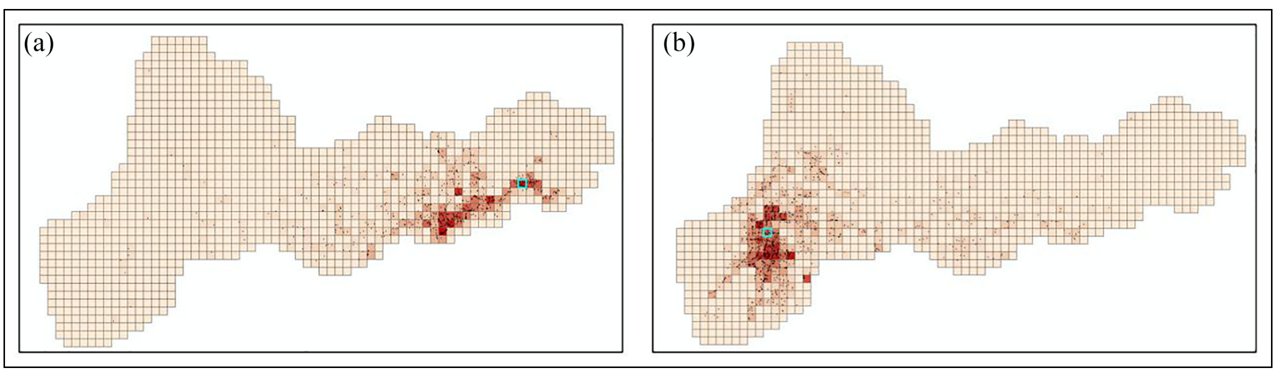

Figure 5. Heat maps of OD transition probabilities for different starting locations: (a) Transition probabilities when the starting grid index is 685. (b) Transition probabilities when the starting grid index is 303.

and the interarrival time distributions, were extracted. These extracted patterns were used to generate and validate ETs' traveling behaviors in simulation under the assumption that ETs and FTs had similar traveling patterns. The details of these behavioral patterns extracted from trajectory data are explained in the following subsections.

3.4. Behaviors of ETs during implementation

The flowchart of key events, status, and decisions of ETs in the ABS model is shown in Figure 4, with each module represented by a circled number for ease of description in the following parts.

3.4.1. Traveling behaviors of ETs. The initial positions of ETs are randomly chosen at the beginning of the simulation. Each ET is assumed to be fully charged and is in cruising status (Module CR). The destination of the cruising trip or the passenger-pickup event (Module PP) location is randomly generated based on the OD transition probability matrix OD_{DP} (Drop-off $\rightarrow$ Pickup), which is estimated from all cruising processes in the trajectory dataset. Similarly, once an ET picks up a passenger, the destination of the operating trip or the passenger-drop-off location is determined by the OD transition probability matrix OD_{PD} (Pickup $\rightarrow$ Drop-off). In Equations 1 and 2, m refers to the number of mapping grids. P_{ij} refers to the conditional probability of dropping the current passenger at grid j if the passenger-pickup location is grid i, while Q_{ij} refers to the conditional probability of picking up the next passenger at grid j if ET drops off the current passenger at grid i. For example, subplots (a) and (b) in Figure 5 show two rows of OD_{PD}, $P_{685\bullet}$ and $P_{303\bullet}$, respectively, which represent the probability of selecting each map grid as the destination of the trip, given that the starting grid index is 685 or 303. The borders of the starting grids are highlighted in green. As shown in Figure 5, the spatial distribution of possible destinations for different origins varies significantly, and the probability of selecting destination grids near the origin grid is much higher than that of the distant ones. In this study, OD_{PD} and OD_{DP} were

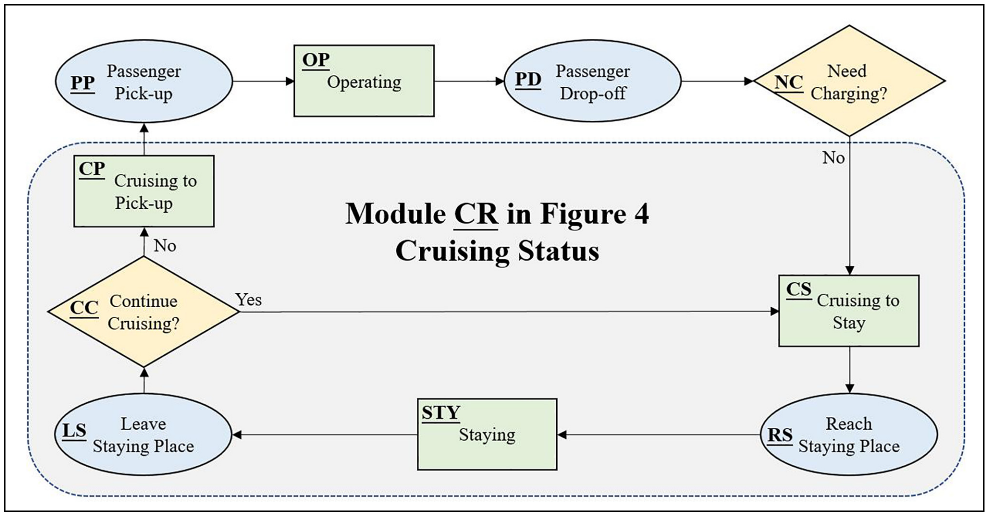

Figure 6. Flowchart of subprocesses in cruising status (Module CR in Figure 4).

estimated using data points from all 24 h at an aggregated level due to limited data volume. If sufficient data are provided, then the elements in OD_{PD} and OD_{DP} can be further extended to P_{ijt} and Q_{ijt} , respectively, where t is the hour index. In this way, hourly OD transition probability matrixes estimated from trajectory data can better characterize tempo-spatial patterns of ETs' traveling behaviors.

$$OD_{PD} = \begin{bmatrix} P_{11} & \cdots & P_{1m} \\ \cdots & P_{ij} & \cdots \\ P_{m1} & \cdots & P_{mm} \end{bmatrix} \text{with} \sum_{j=1}^{m} P_{ij} = 1,\ 0 \leqslant P_{ij} \leqslant 1 \tag{1}$$

$$OD_{DP} = \begin{bmatrix} Q_{11} & \cdots & Q_{1m} \\ \cdots & Q_{ij} & \cdots \\ Q_{m1} & \cdots & Q_{mm} \end{bmatrix} \text{with} \sum_{j=1}^{m} Q_{ij} = 1,\ 0 \leqslant Q_{ij} \leqslant 1 \tag{2}$$

In fact, an ET is intermittently in a stationary state rather than always moving. The implementation of the cruising state (Module CR) is extended to several subcomponents, which is shown in Figure 6. After dropping off the passenger, a decision module (Module NC) is available for the driver to decide whether to go to CS, which is described in detail in the following subsection. If charging is unnecessary, then the ET will reach a staying place to rest. The selection of the staying grid follows the spatial probabilities in OD_{DP}. The staying time is randomly generated based on an hourly updated normal distribution with mean (μ_{stay} in Equation 3) and standard deviation (σ_{stay} in Equation 3) estimated from staying events in the trajectory dataset. After the current staying process ends, a probability array P_{Busy} is used to randomly decide whether the ET goes to pick up the next passenger or find another place to stay at the current hour. In Module CC, if the ET decides to cruise, then the process returns to Module CS and repeats the staying process, or the ET starts a trip to pick up the next passenger (Module CP), whose destination grid is randomly determined by OD_{DP} .

$$\mu_{stay} = [\mu_0 \cdots \mu_h \cdots \ \mu_{23}] \quad \sigma_{stay} = [\sigma_0 \cdots \sigma_h \cdots \ \sigma_{23}\ ] \tag{3}$$

$$P_{Busy} = [B_0 \cdots B_h \cdots \ B_{23}]\ with\ 0 \leqslant B_h \leqslant 1 \tag{4}$$

3.4.2. Charging Behaviors of ETs. Figure 4 shows that, after the passenger is dropped off, a decision procedure (Module NC) is available for the ET driver to decide whether to go to CS. The rules in Module NC are listed as follows:

(a) If the ET's current battery level is less than SOC_{thr1} (e.g., 30%), then the process will go to Module SLC, in which the ET will select a CS.
(b) If the ET's current battery level is between SOC_{thr1} and SOC_{thr2} (e.g., between 30% and 50%), then a probability C_h that the ET will go to Module SLC to select a CS exists. C_h is an hourly updated probability based on the array P_{Char}, which is defined in Equation 5.
(c) If the ET's current battery level is greater than SOC_{thr2} (e.g., 50%), then the process will go to Module CR, and the ET will start another cruising trip.

$$P_{Char} = [C_0 \cdots C_h \cdots \ C_{23}]\ with\ 0 \leqslant C_h \leqslant 1 \tag{5}$$

Based on the rules of Module NC, if the ET driver decides to charge the vehicle, then the ET will go through a searching procedure of CS, which is Module SLC in Figure 4. The searching rules of ETs are described as follows:

(a) The set of CSs S is obtained within a prespecified search radius R_0.
(b) The subset S_{free} is obtained from S, which has free charging piles at the searching time.
(c) If S_{free} is not empty, then the ET will select the nearest one as CS.
(d) If S_{free} is empty, which implies that all stations in S have queues, then the ET will select the station with the shortest queuing length as the CS.
(e) If S is empty, then the search radius is increased by R_{step} and the same procedure is repeated.

In this way, instead of simply selecting the nearest station implemented by other ABS models, the logic and rules in this study have a certain level of intelligence with cell phone applications sharing the real-time state of CSs nearby, which is closer to the habits of ET drivers.

After the CS is selected in Module SLC, ET will take the fastest route to the CS (Module TC). Upon arrival (Module AC), the ET checks the actual line state of the station (Module IQ), which is due to that a queue may form while traveling in this direction. If a queue is formed in the station, then the ET will start queuing (Module SQ). Each station has multiple charging piles, and ET arrivals form a single queue with unlimited queuing capacity. The corresponding queuing model is M/M/c/∞ in Kendall's notation, which indicates the arrival process follows a Markovian (Poisson) distribution, the service time follows a Markovian (exponential) distribution, c servers are available, and the queue capacity is infinite. The waiting time in Module QU depends on the actual queuing status in the simulation, such as the number of vehicles in front of the queue and their charging time. If no queue is formed in the station, then the ET will start charging (Module STC), and the charging time in Module CH is a linear function of the percentage of battery to be fully charged. The empty battery takes approximately 2 h to be fully charged, which is consistent with reality. After charging is completed, the ET will be in a cruising status, which is Module CR in Figure 4.

3.5. Model calibration and validation

This simulation model includes many configuration parameters, some of which can independently determine the overall system performance, such as the temporal and spatial patterns of pickup and charging demand. For example, the parameters OD_{PD} and OD_{DP} directly determine the spatial distribution of ET destinations given a specific origin for busy or idle states, respectively. However, some system metrics are not solely determined by an individual set of parameters but are influenced by multiple groups of parameters that interact with each other. In addition, these system metrics themselves have mutual influences and need to be consistent with the real-world situation, which makes model calibration and validation challenging tasks. For example, the system performances, such as the average number of pickup tasks completed by each ET per day and the average number of charging times of each ET per day, need to be collected after a simulation run. The two system statistics represent the overall demand level for ET pickup and charging events and need to be consistent with the actual situation. They are not independent but influence each other. If fewer daily charging events occur, then more time would be available for ETs to complete additional pickup events. The goal of the calibration procedure is that not only the overall levels of the two system statistics but also the hourly proportional trends of incident counts are consistent with the real-world scenario. In this way, using the calibrated model is meaningful for further evaluation and optimization of EVCI. The level and hourly proportion trend of pickup times are mainly determined by P_{Busy} and μ_{stay}, while the main influencing factors of charging times are P_{Char} and other charging parameters such as the capacity of the ET battery and the search distance. These input and spatial parameters jointly affect the two system statistics due to the complexity of the simulation system.

In the calibration process, some parameters such as μ_{stay}, σ_{stay}, the charging capacity and the searching distances are fixed at a certain level, while P_{Busy} and P_{Char} are adjusted using an iterative trial-and-error method to achieve the calibration goal. P_{Char} and P_{Busy} are initially set based on the empirical data. After a simulation is completed, the averages of daily pickup and charging times between empirical data and simulation output are compared. The means of P_{Busy} and P_{Char} are adjusted in new simulations until the two metrics match the level in empirical data. Then, each element in the parameter array of P_{Busy} and P_{Char} is adjusted to ensure that the hourly trends of the pickup and charging times from the simulation output are close to those calculated based on empirical data. In each iteration of the calibration process, a prespecified step size is applied to the relevant hourly parameters of P_{Busy} and P_{Char} to reduce the gap between the simulation statistics and the actual level at that hour. The iterative calibration procedure can be implemented using programming language and conducted automatically.

A validation test needs to be conducted to ensure that the calibration method mentioned above does not result in model overfitting. Performing an out-of-time test is challenging due to the limited time span of the trajectory data collected. Therefore, an out-of-sample validation method is employed. The taxis in the trajectory data are randomly divided into two groups in a 2/3 and 1/3 ratio, and the corresponding trajectory data are then split into training and testing datasets. In the training phase, the parameters of

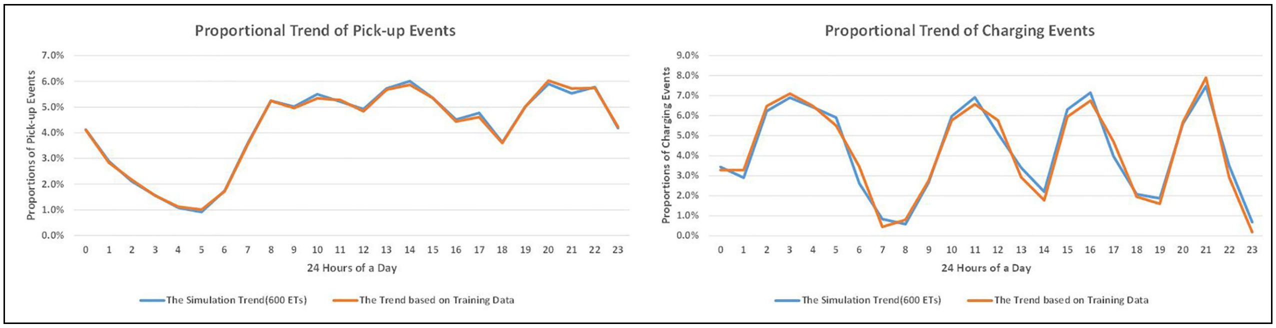

Figure 7. Comparison of proportional trends between simulation and training data.

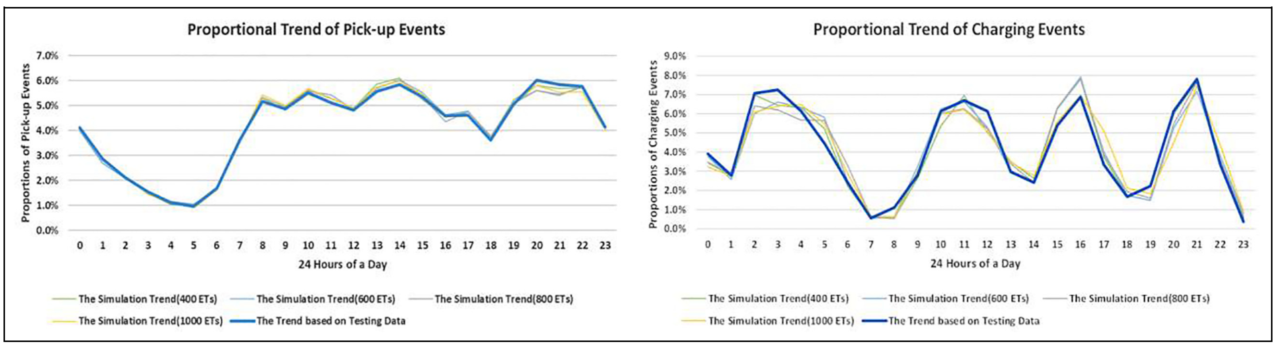

Figure 8. Comparison of proportional trends between simulation and testing data.

OD_{PD}^{Train}, OD_{DP}^{Train}, μ_{stay}^{Train}, and σ_{stay}^{Train} are first estimated and obtained. The initial values of P_{Busy} and P_{Char} are set as P_{Busy}^{Train} and P_{Char}^{Train}, respectively, which are also estimated from training data. These parameters are input into the simulation model, and other parameters are set to be at the average level. For example, the number of ETs is set to 600. Then, P_{Busy} and P_{Char} are fine-tuned iteratively until a good fitting of the hourly proportional trends is realized between the simulation outputs and those from the training data (Figure 7). Table 1 shows that the mean absolute error percentages for pickup and charging events are only 0.07% and 0.36%, respectively. Finally, the optimized parameters of P_{Busy}^{*} and P_{Char}^{*} are recorded for future tests.

In the validation phase, the parameters of OD_{PD}^{Test}, OD_{DP}^{Test}, μ_{stay}^{Test}, and σ_{stay}^{Test} are estimated based on testing data. They are input into the simulation model along with P_{Busy}^{*} and P_{Char}^{*} obtained from the training phase. The remaining parameters are set to the same values as those defined during the training phase. The ET count parameter is set at four levels, namely, 400, 600, 800, and 1000, to verify the performance of the optimized parameters under different levels of pickup and charging demand. The validation tests are conducted four times under each ET count level. The simulation results show that the average pickup times per ET per day are 27.3–27.9 for all simulations in the training and testing phases. The average charging times per ET per day are 3.2–3.3 for all simulations in both phases. This result indicates that the current simulation implementation, combined with the optimized parameters, can maintain a relatively stable level of pickup and charging demand for each ET. However, the overall demand for the system varies with the total number of ETs. Figure 8 and Table 1 show that the fittings of the hourly proportional trend of pickup and charging events are slightly worse than those in the training tests, but the absolute error still remains at a relatively low level. Overall, the simulation model can effectively reflect the demand level of pickup and charging after changes in parameters such as OD_{PD}, OD_{DP}, P_{Busy}, and μ_{stay}. The trends throughout the 24 h of the day are very close to the real scenario.

Similar hourly trends can be found in other literature, such as Wang et al.[54] and Tian et al.[41] In particular, the characteristics of four charging peaks per day in the hourly trend of charging events, as shown in Figures 7 and 8, have persisted in Shenzhen for several years from 2013 to 2017. On the basis of their study of Shenzhen ETs, the first and third charging peaks in Figure 8 are probably due to full charge between shifting and preparation for the rush hours of picking up passengers. The second and fourth peaks are around drivers' lunch and dinner times. Drivers tend to have lunch and dinner while charging the vehicles simultaneously even if the current power is not yet to the point

Table 1. The mean absolute error% of proportional trends of pickup and charging events between simulation and training/testing datasets.

Event type	Validation type	The number of ETs	The mean absolute error%
Passenger-pickup events	Sim. output vs. training data	600	0.07%
	Sim. output vs. testing data	400	0.11%
		600	0.11%
		800	0.12%
		1000	0.11%
ETs' charging events	Sim. output vs. training data	600	0.36%
	Sim. output vs. testing data	400	0.37%
		600	0.47%
		800	0.51%
		1000	0.59%

ET: Electric taxis

Table 2. Statistics of the output summary variables and simulation evaluation results.

Simulation output variables	Average	Standard deviation	Variation coefficient
Ave. charging times per ET per day	3.253	0.007	0.23%
Ave. pickup times per ET per day	27.863	0.027	0.10%
Ave. running distance per ET per day (km)	337.580	0.199	0.06%
Ave. running time per ET per day (hours)	13.185	0.007	0.05%
Ave. queuing time per ET per day (min)	19.189	0.518	2.70%
Ave. time traveling to CS per ET per day (min)	31.507	0.131	0.42%
Stdev. of charging times proportions for all charging stations	2.90%	0.02%	0.79%
Stdev. of queuing times proportions for all charging stations	6.53%	0.05%	0.76%
Ave. of utilization rates for all charging stations	35.74%	0.07%	0.19%

ET: Electric taxis; CS: Charging station

where it must be charged. Validation results show that such charging patterns are modeled well by the simulation rules and parameters.

3.6. Evaluation of one EVCI layout plan

On the basis of the calibration and validation methods proposed in Section 3.5, the use of the simulation model to evaluate an EVCI layout plan is explained in this section, along with the relevant input parameter setup, output measurements, and tests for convergence, repeatability, and stability. The EVCI layout to be evaluated by the ABS model is shown in Figure 9, which is based on actual conditions in Shenzhen when the trajectory data were collected, including one large station with 100 charging piles (the blue box), 2 median stations with 40 charging piles (green circles), and 18 small stations with 6 charging piles in each station (yellow triangles). The number around each CS in Figure 9 is the map grid index where the station lies, which can be considered the ID of the CS. The heatmap of the passenger-pickup incidents is also shown in Figure 9. Most CSs are located in areas with a relatively high pickup density. The settings of the other model parameters are listed as follows. The means of large matrix parameters are provided due to space constraints.

- Simulation length: 7,000,000 ticks (equivalent to 34 days in real time; one tick stands for 0.42 s in real time)
- Number of ETs: 600
- SOC_{thr1}: 20%
- SOC_{thr1}: 80%
- Mileage with full power M_{fp}: 300 km
- Search radius R_0: 6 km
- Average of P_{Busy}: 0.540
- Average of P_{Char}: 0.042
- Average of μ_{stay}: 10.02 min
- Average road speed: 29.23 km/h

In simulation and modeling, random events are generated using the "Random()" class in Java Repast. A predetermined seed dictates the sequence of random numbers generated during the simulation. Before evaluating the layout plan, experiments were conducted to ensure that running different simulations with the same seed produced consistent results. The initial values of some parameters in the simulation, such as the initial position of each ET, were randomly generated at the start of the simulation run. A repeatability test was conducted to demonstrate the stability and robustness of the model to ensure that the randomly generated initial parameters have little effect on the final evaluation of the EVCI layout model. The model was

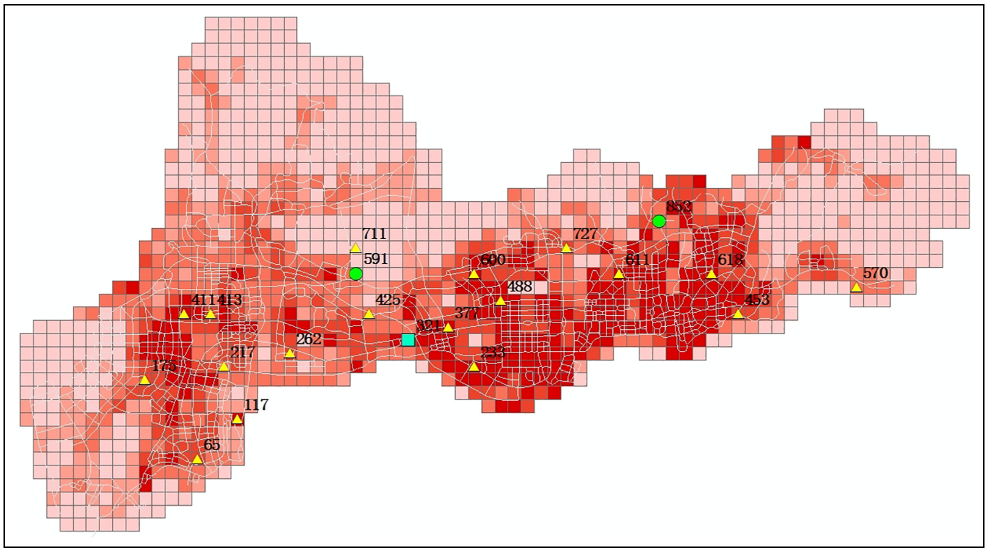

Figure 9. Heatmap of passenger-pickup events and locations of charging stations.

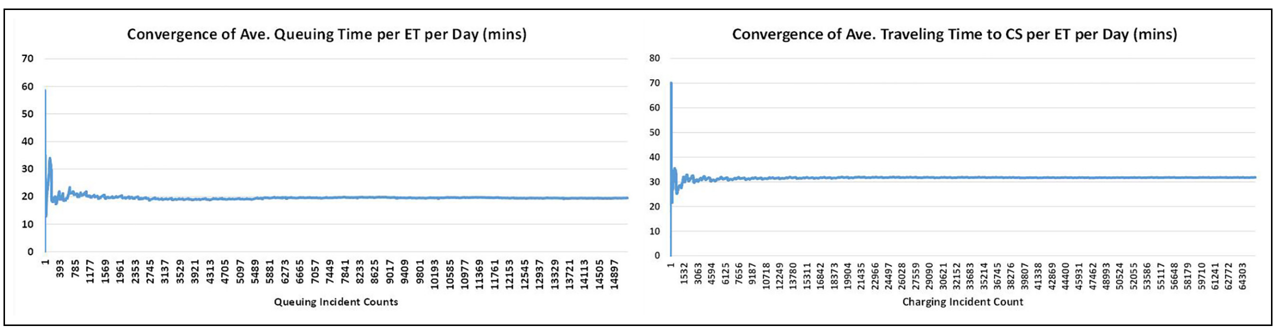

Figure 10. Convergence of simulation output variables.

run 30 times under the same parameters using different seeds, and the variation in the model's output variables was analyzed. Table 2 shows that the variation coefficient of most model outputs is less than 1%, which indicates a relatively low level of variation. The variation coefficient of the average queuing time is 2.7%, which is slightly higher than those of other output variables. However, its standard deviation is only 0.52 min, which is at a low level. Therefore, the randomness of the initial parameters has a minimal effect on the evaluation of the EVCI layout, which implies a high level of robustness and stability of the simulation model.

Given that the ABS model is a tick-by-tick high-resolution simulation model, the computational resources to execute the computer program are relatively high. Usually, obtaining a stable system performance evaluation on a normal computer takes several hours in the real world. Table 2 shows the statistics of pickup and charging events of ETs, the traveling and queuing time of ETs, and the performance metrics of CSs. Most of these variables tend to converge as the simulation progresses. Figure 10 shows the convergence of two summary variables of the ABS model. They fluctuate at the beginning of the simulation and tend to become stable when an increasing number of samples are accumulated. The stabilized values of these variables can be used to evaluate the system performances, such as the average queuing time and traveling time to CS as measures of ET users' time satisfaction and the average utilization rates for all CSs as a measure of EVCI usage efficiency. Each plan only needs to be simulated once when evaluating a large number of EVCI layout plans due to the stability of the simulation model. Moreover, the simulation length can be appropriately selected based on the convergence of the layout evaluation metrics to improve efficiency.

The performances of each CS can be evaluated by the ABS model as well. The statistics of each station's charging, queuing, and utilization situation are shown in Table 3. The utilization rates vary widely from 76% as the highest to 1.6% as the lowest. CS No. 321 and No. 591,

Table 3. The ABS model's evaluation of each charging station.

CS ID	Number of charging piles	Number of charging events	Proportion in all charging events	Number of queueing events	Proportion in all queuing events	Queuing ratio in charging events	Utilization rates
321	100 (Large)	3225	4.8%	0	0.0%	0.0%	2.6%
591	40 (Median)	841	1.3%	0	0.0%	0.0%	1.6%
852		7940	11.9%	25	0.2%	0.3%	16.2%
117	6 (Mini)	748	1.1%	0	0.0%	0.0%	10.1%
711		826	1.2%	1	0.0%	0.1%	12.1%
570		1232	1.8%	9	0.1%	0.7%	16.9%
425		1484	2.2%	6	0.0%	0.4%	19.9%
65		1727	2.6%	35	0.2%	2.0%	23.5%
217		1990	3.0%	40	0.3%	2.0%	26.1%
262		2117	3.2%	55	0.4%	2.6%	28.0%
411		2373	3.5%	137	0.9%	5.8%	32.0%
413		2603	3.9%	150	1.0%	5.8%	34.9%
175		2674	4.0%	159	1.0%	5.9%	35.8%
600		3116	4.7%	628	4.1%	20.2%	41.7%
727		3674	5.5%	1164	7.6%	31.7%	48.9%
377		3864	5.8%	1206	7.9%	31.2%	51.3%
453		5045	7.5%	1980	13.0%	39.2%	66.0%
233		5212	7.8%	2152	14.1%	41.3%	68.5%
618		5247	7.8%	2340	15.3%	44.6%	68.6%
488		5333	8.0%	2206	14.5%	41.4%	69.9%
611		5732	8.6%	2960	19.4%	51.6%	76.0%

ABS: agent-based simulation CS: Charging station; ET: Electric taxis.

Table 4. Model parameters and levels in sensitivity analysis.

Model parameters	Levels	Description
The number of ETs	[200, 400, 600, 800, 1000]	The number of electric taxis in the simulation
The capacity of ET battery	[100, 150, 200, 250, 300, 350, 400, 450, 500]	The mileage with full power
The charging pile volume	[low, mid, high, veryHigh]	The type and number of piles for each level: low (mini:3, median:20, large:50); mid (mini:6, median:40, large:100); high (mini:9, median:60, large:150); veryHigh (mini:12, median:80, large:200)
The searching distance of ETs	[1, 2, 3, 4, 5, 6]	The initial searching radius of electric taxis

ET: Electric taxis

which had more charging piles than other stations, had the lowest utilization rates of only 2.6% and 1.6%, respectively. Improper location might explain this result given that the two stations were not centrally located but on the fringes of the hot zone of pickup events. Some other stations were not far from the two stations, which further reduced the workloads. However, another median-level CS No. 852 had a higher utilization rate of 16.2%, which is probably due to that it was closer to the hotspot region and the surrounding CSs were not very close to it. CS No. 611 was located in the center of the hotspot zone. Given that it had only six charging piles, it had the highest utilization rate of all CSs. The relationship of event count proportions between charging and queuing can be found in Figure 11. A piecewise positive linear relation exists between the two. Stations with higher charging

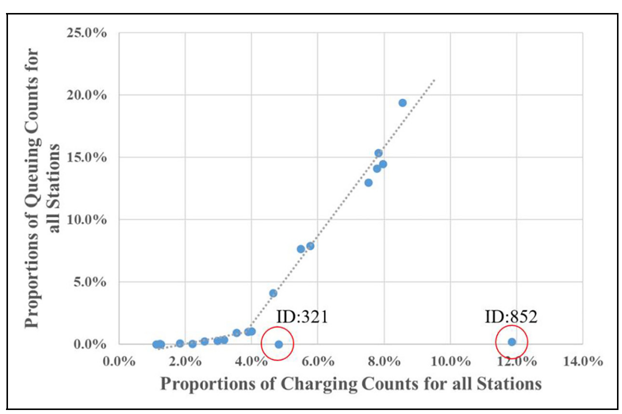

Figure 11. Piecewise linear relation between charging and queuing proportions.

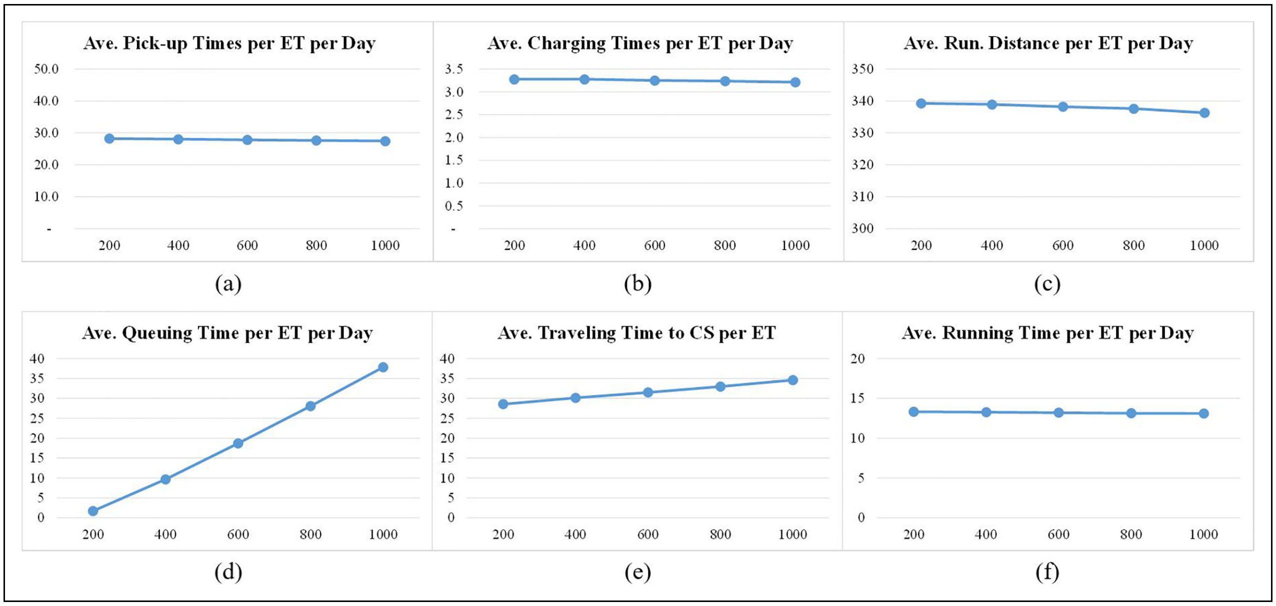

Figure 12. Sensitivity analysis for the number of ETs in the ABS model. The label of y axis for subplot (a) Ave. pick-up times per ET per day; (b) Ave. charging Times per ET per day; (c) Ave. running distance per ET per day (kilometers); (d) Ave. queuing time per ET per day (minutes); (e) Ave. traveling time to CS per ET (minutes); (f) Ave. running time per ET per day (minutes).

proportions are more likely to have more queuing proportions. However, when the charging incident proportion is less than 4%, the queuing incident proportions are kept at a low level, with a slight increasing trend such that charging activities are becoming more active. Stations No. 321 and No. 852 with red circles in Figure 11 are the outliers of this relationship because they are large and median stations with many more charging piles than others. Thus, they have difficulty forming queues.

In summary, the ABS model can provide a detailed and stable evaluation of the layout of EVCI and the traveling and charging behaviors of ETs. The layout can still be further improved in the simulation. The large and median CSs were not very well located, which led to very low utilization rates. However, some CSs with very high usage rates had serious queuing problems. This phenomenon is consistent with the description in the literature. Given that Shenzhen was in the early stage of developing new energy vehicles at that time, this EVCI layout design problem was understandable. Some stations were still under construction and were ignored in the simulation.

4. Sensitivity analysis

A sensitivity analysis was conducted for important variables in the simulation to determine the effect of model parameters on the simulation results. One set of simulation experiments was designed and conducted for different levels of one variable independently while fixing the other model parameters. The model parameters and their levels of sensitivity analysis are listed in Table 4, while the results are elaborated in the following subsections.

4.1. Number of ETs

Similar to the validation test conducted in Section 3.5, the fluctuations in the average pickup and charging times per ET per day are minimal, which are maintained at a consistent level, as illustrated in Figure 12, when the number of ETs is altered while the other parameters are fixed. Accordingly, the charging demand, traveling distance, and time for each car are maintained at the original level. However, the total pickup demand of all ETs increases proportionally with the number of ETs, which leads to a rise in the total charging demand. On the supply side, the charging capacity provided by EVCI does not change. As a result, the average queuing time and the traveling time to reach the CS increase linearly with the number of ETs.

4.2. Capacity of the ET battery

As shown in Figure 13, the charging demand decreases with the increase in ET battery capacity, which results in a decrease in the queuing time and traveling time to reach the CS. Based on the simulation implementation of ET's behavioral logics described in Section 3.4, the model does not confine the overall pickup service demand from a global perspective. Instead, it assumes that each ET decides whether the next movement after it completes a charging session or a passenger-pickup will be to pick up another passenger or to cruise based on the parameter P_{Busy}. Therefore, the decrease in charging demand allows ETs to have more time available for picking up passengers, which leads to an increase in pickup times, traveling mileage, and traveling time.

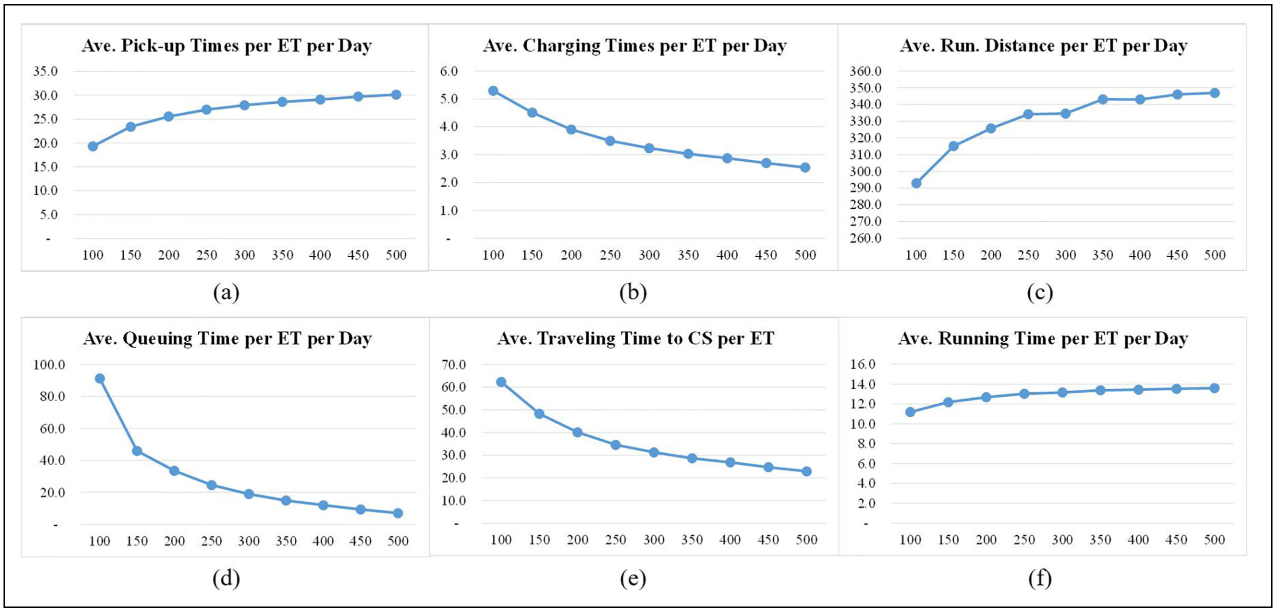

Figure 13. Sensitivity analysis for ETs' battery capacity in the ABS model. The label of y axis for subplot (a) Ave. pick-up times per ET per day; (b) Ave. charging Times per ET per day; (c) Ave. running distance per ET per day (kilometers); (d) Ave. queuing time per ET per day (minutes); (e) Ave. traveling time to CS per ET (minutes); (f) Ave. running time per ET per day (minutes).

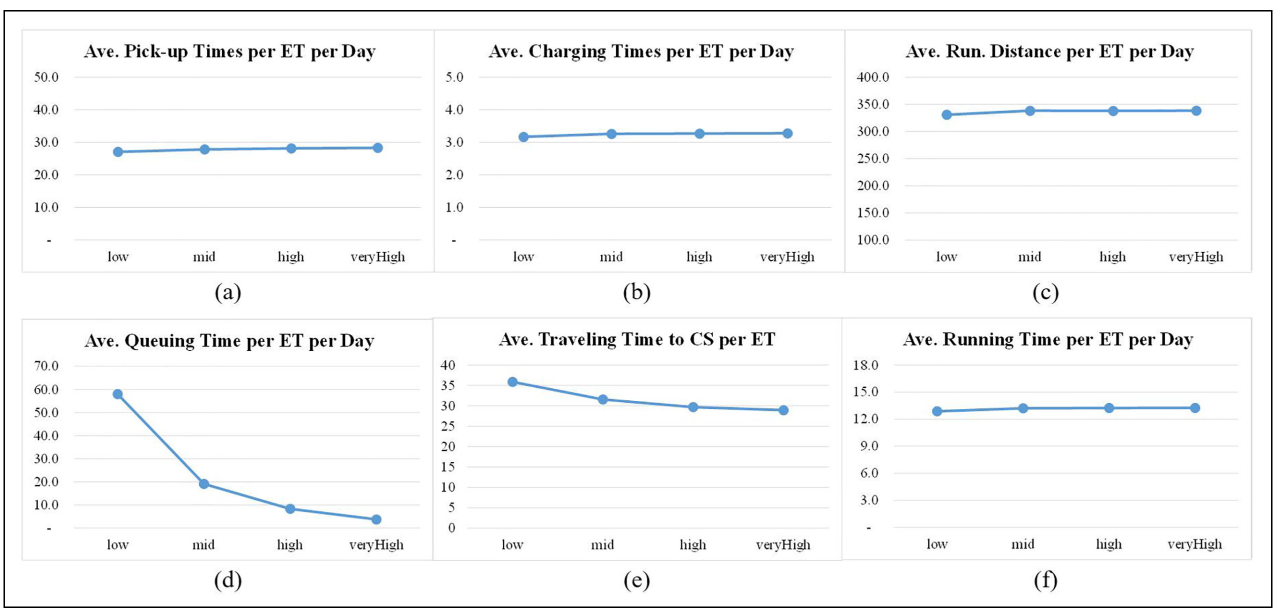

Figure 14. Sensitivity analysis for the charging pile volume in the ABS model. The label of y axis for subplot (a) Ave. pick-up times per ET per day; (b) Ave. charging Times per ET per day; (c) Ave. running distance per ET per day (kilometers); (d) Ave. queuing time per ET per day (minutes); (e) Ave. traveling time to CS per ET (minutes); (f) Ave. running time per ET per day (minutes).

4.3. Charging pile volume of CSs

Figure 14 shows that the ETs' average queuing time decreases when a CS increases its number of charging piles. As the number of charging piles increases evenly, the queuing time decreases nonlinearly, in which the magnitude of the decline gradually decreases. The traveling time to reach the CS also decreases. The reason is that ETs prefer CSs that are farther but do not require queuing. With the increase in the number of charging piles, the nearest CSs do not require queueing, which reduces the driving distance and time. The decrease in queuing time leads to more time for operation. The pickup and charging

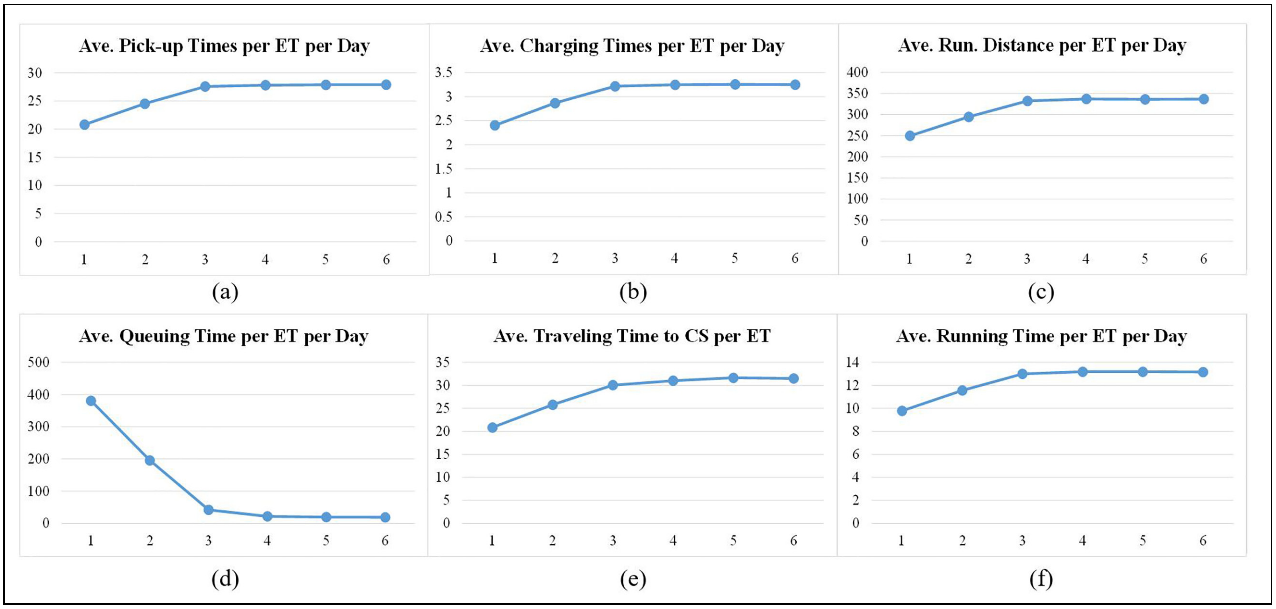

Figure 15. Sensitivity analysis for ETs' searching distance in the ABS model. The label of y axis for subplot (a) Ave. pick-up times per ET per day; (b) Ave. charging Times per ET per day; (c) Ave. running distance per ET per day (kilometers); (d) Ave. queuing time per ET per day (minutes); (e) Ave. traveling time to CS per ET (minutes); (f) Ave. running time per ET per day (minutes).

times of ETs increase slightly and are maintained at the original level.

4.4. Search distance of ETs

Finding CSs that are farther but do not require queueing is easier for ETs as the search distance increases. Therefore, the queuing time decreases, and the time to travel to the CS increases (Figure 15). Correspondingly, the reduction in time spent on charging leads to more time for picking up passengers. With the increase in the volume of pickup activities, traveling distances and time, as well as charging times, also increase. Notably, the search distance of 3 km appears to be a cut-off point. The upward or downward trend of performance variables becomes nearly nonexistent if the search distance is greater than 3 km. The reason is that the study region, especially the hotspot zone of pickup incidents, is a long and narrow area. If the search distance is more than 3 km, then the boundaries of the study area will be exceeded. The additional search radius has little practical effect.

In summary, the results of sensitivity analysis are unsurprising, which shows the correctness of the model implementation helps us understand the effect of model variables on the evaluation of system performance.

5. Optimization of EVCI layout

Considering that the ABS model gives a reasonable evaluation of the EVCI layout, it can be used for the location optimization problem. The layout designs generated by mathematical optimization methods in the literature, plans proposed by experts, and randomly generated plans can be input into the ABS model. After simulation evaluation, the best layout plan can be chosen based on one or multiple performance variables, such as average queuing time and EVCI utilization. In this study, a multiobjective optimization method based on Pareto analysis is proposed to solve the following practical problem: if the number of ETs increases to 1000 and the current EVCI layout shown in Figure 9 is given, then how can another large CS with 100 charging piles be located in the city to maximize ET users' time satisfaction and the average utilization rate of all CSs simultaneously?

EV users' time satisfaction can be measured by the sum of the traveling time to reach the station and the waiting time in the queue, which should be minimized to achieve this goal. The average utilization rate of all CSs represents the usage of the charging facilities, with a higher value indicating better usage and reduced wastage of charging infrastructure. The optimization objective is set as minimizing the opposite number of average utilization rates to facilitate subsequent Pareto analysis, and it is equivalent to maximizing the utilization rates. Thus, multiobjective optimization aims to simultaneously minimize both metrics. The possible locations of the CS to be built can be generated by selecting the top 100 map grids with the highest densities of pickup incidents based on the trajectory data of ETs. These 100 layout designs along with the current layout as the base plan are input into the ABS model for simulation evaluation.

The evaluation results regarding the time satisfaction and CS utilization percentages of these 101 layout designs are shown in Figure 16. Among the 101 layout plans, 20 are positioned on the Pareto frontier, as depicted by the blue dots in Figure 16. Under any weighting of the two measures, the 20 layout plans dominate the others. Five representative plans were selected from the Pareto frontier. They are plans with map grids No. 391 (the best layout for time satisfaction) and No. 394, No. 290, No. 697, and No. 515 (the best layout for CS utilization). These plans demonstrate varying degrees of improvement in two metrics compared with the original layout plan, which is shown in black circles in Figure 16. For example, the layout plan with grid No. 391 can increase the time satisfaction by 37.7%, and the plan with grid No. 515 can improve the average CS utilization by 6.5%. The plan with grid No. 679 can improve the time satisfaction and CS utilization by 10.5% and 4.4% simultaneously. The newly constructed CS is more likely to have a significant improvement effect on ET users' time satisfaction. The locations of the additional CSs of the five plans are shown in Figure 17. Grids No. 391, No. 394, and No. 290 are close to one another, and they are all located in densely populated areas with high demand for pickup services in Luohu District (the darker red areas in Figure 17). These locations fill the gap in surrounding areas where no CSs are available, which significantly reduces the time to reach CSs and waiting time in queues. Grid No. 617 is located on the eastern edge of Luohu District, which is not far from existing station No. 618 as one of the most utilized CSs. Therefore, Grid No. 617 can relieve the queue pressure of nearby CSs and simultaneously improve the overall average utilization of CSs. Grid No. 515 is located far from the other alternative locations on the western edge of Bao'an District, and it fills the lack of large-scale CSs nearby. Therefore, it can significantly improve the overall utilization level of CSs. However, it is far from the core areas of Futian and Luohu Districts. Thus, it does not contribute to the improvement in time satisfaction.

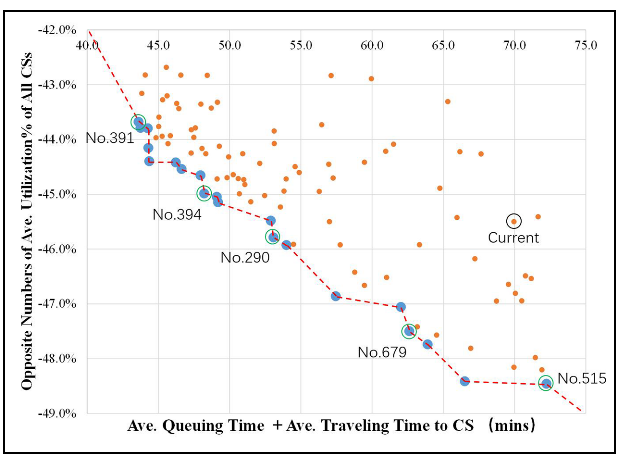

Figure 16. Pareto frontier of EVCI layout designs.

6. Discussion and limitations

The innovation of the proposed ABS model in this study is primarily the modeling and simulation of the traveling, passenger-pickup, and charging behaviors of ETs in a more realistic way using the trajectory data of ETs. The simulation system is then applied to evaluate and optimize the layout design of EVCI. The closeness between simulation and real-life situations is primarily reflected at the micro- and macrolevels. The microlevel perspective is reflected in individual ET agents' logical behaviors and decision-making processes, such as adjusting their speed based on traffic conditions, following specific spatial traveling patterns during busy or idle states, making decisions on whether to charge at different times of the day and intelligently searching and selecting CSs. From a macrolevel perspective, the model accurately captures the hourly trends of city-level passenger-pickup and ET charging incidents. This model, which is based on trajectory data, allows for a detailed simulation of the complex logic of

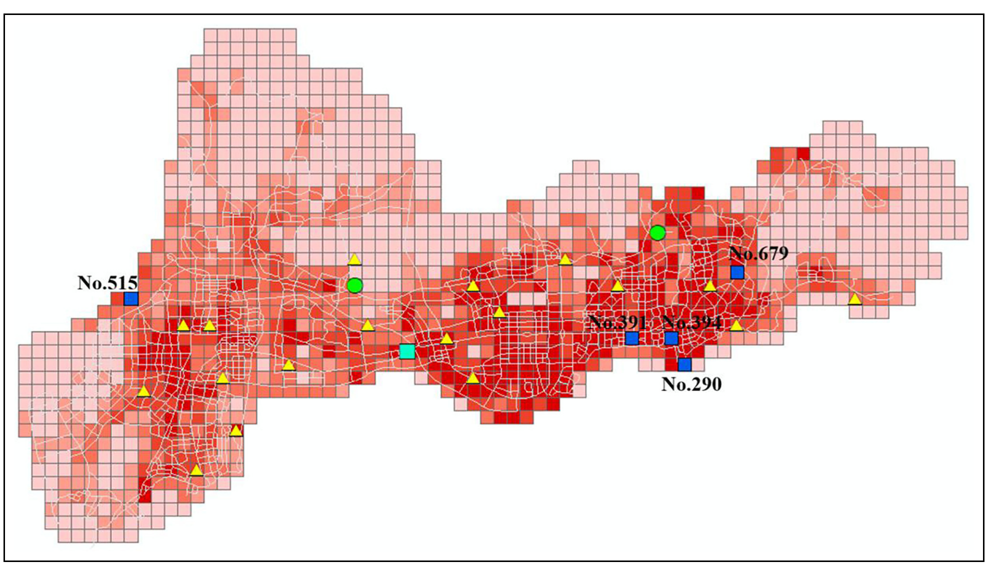

Figure 17. Locations of layout plans with map grid No. 290, No. 391, No. 394, No. 679, and No. 515.

ETs by fitting the overall levels and temporal patterns of pickup and charging demands in the entire city through a calibration process. The performance of the model in validation tests with new data verifies that it does not suffer from overfitting or excessive training, which proves that the model can capture the main temporal patterns of the city using a limited number of parameters and maintain this feature in simulating new data.

The validation tests of the model show that the current implementation can ensure that the average pickup and charging events per ET remain relatively stable when changing the ET count parameter. However, the overall demand level of the whole system will change with variations in the ET count. If the model user requires experiments to be conducted under a constant passenger-pickup service demand level, then the model's P_{Busy} and μ_{stay} parameters need to be recalibrated using a trial-and-error method to achieve the desired level. This iterative calibration method incurs a certain computational time cost. However, in the scenario of EVCI layout evaluation, the performance of CSs and their layouts under higher pickup and charging service demand intensity often needs to be tested. From this perspective, the method of adjusting the overall demand intensity by altering the ET count can also meet the needs of the model users.

The data used in this study are from 2014 due to the difficulty in collecting trajectory data. These data are(is) quite outdated and may not reflect the current situation. They are also unsuitable for conducting out-of-time dataset splitting during the validation test because of the limited time span of these data. Therefore, this study adopts an out-of-sample test approach by splitting the data based on ET IDs. If longer time span data can be collected in the future, then further validation can be conducted. The main purpose of this study is to introduce a simulation and optimization methodology based on trajectory data. From this perspective, even using data from 2014 can achieve this goal. This data format is typical trajectory data, and the data source is Shenzhen, China, which is one of the first cities in China to popularize ETs and is representative and typical in this regard. Moreover, from the findings of previous research such as Wang et al.,[54] the hourly proportional trend of charging events closely resembles the trends described in this study. The relatively stable trend of four charging peaks per day between 2013 and 2017 indicates that the hourly trend of charging incidents in a city can remain stable for a considerable period. This description further demonstrates that the calibration and validation methods of this model are effective in capturing the basic trends for future EVCI layout evaluation. However, collecting the most up-to-date data for model calibration is essential in practical applications. This study mainly focuses on Shenzhen, while the traveling and charging tempo-spatial distribution of ETs in different cities will definitely have distinct characteristics influenced by factors such as urban traffic, charging facilities, geography, culture, and travel habits. Therefore, this model can be applied to other cities, and some considerations need to be made before using the model in different cities.

7. Conclusions and future work

The proposed ABS model in this study can simulate the traveling and charging behaviors of ETs based on tempo-spatial patterns extracted from GPS trajectory data of ETs in the GIS environment of an urban road network with a traveling speed of 24 h. Complex and dynamic behaviors of ETs, such as shortest-path routing, passenger-pickups, and intelligent search of CSs, were implemented using rules and logics in the ABS model, through which interactions among agents and their interactions with the environment were modeled. The simulation was calibrated such that the dynamic patterns of important events, such as passenger-pickup and charging incidents, matched those extracted from trajectory data. The validation test confirms that the model calibration did not suffer from overtraining and overfitting problems. Instead, it demonstrates the ability of the model to capture the primary temporal patterns of the city using a limited number of parameters while maintaining this capability in simulating new data. After simulation execution, detailed and summarized data of ET behaviors and EVCI performance were obtained. Multiple simulation experiments were designed and conducted to test the sensitivity of important model parameters and settings. The results appear reasonable, which shows that the correctness of the model implementation and the influence of model variables on system performance can be further understood. A multiobjective layout optimization procedure based on the Pareto frontier was proposed to maximize ET users' time satisfaction and CSs' utilization. The location plans for the new CSs based on Pareto analysis can improve both metrics through simulation evaluation. The final decision on the location plan should be made based on practical considerations such as financial cost, power network, and environmental factors.

The performance measurements of the EVCI layout in the current ABS model are mainly ET users' time satisfaction and EVCI's utilization, which can be further extended to include other aspects such as construction cost, operating cost, and power grid performance. Currently, the layout optimization procedure is mainly based on the simulation evaluation of a set of layout plans from experts or other algorithms. Relatively good plans are selected from them, but a global optimal solution cannot be guaranteed. Integrating the current ABS model with heuristic optimization algorithms such as GAs and simulated annealing is impractical because the ABS model is a simulation of high temporal resolution, whose computational and time costs are relatively high. In the future, discrete-event simulation can be developed to fully utilize the major characteristics of the ABM model and greatly improve the execution efficiency of the simulation model. The discrete-event simulation model can filter out poorly performing solutions quickly and can be combined with a heuristic optimization algorithm to generate layout plans for further evaluation of the ABS model.

Funding

This research was funded by the University-level Scientific Research Initiation Project of Shenzhen Polytechnic, grant number 6022312065K; Foundation of Philosophy and Social Sciences of Guangdong Province, grant number GD22XYJ34; University-level Scientific Research Initiation Project of Shenzhen Polytechnic, grant number 6022312035S; University-level Youth Innovation Project of Shenzhen Polytechnic, grant number 6022310010S; Post-doctoral Later-stage Foundation Project of Shenzhen Polytechnic, grant number 6022271005S.

ORCID iD

Jie Tan https://orcid.org/0000-0001-7075-2345

Note

The code, data, and video of the Agent-based model can be downloaded from the computational model library of CoMSES Network: https://www.comses.net/codebases/68fdf788-6b8c-44a1-9c38-62cbca344b4e/releases/1.0.0/.

References

1. Kim S, Lee U, Lee I, et al. Idle vehicle relocation strategy through deep learning for shared autonomous electric vehicle system optimization. *J Clean Prod* 2022; 333: 130055.
2. Wang N, Tian H, Wu H, et al. Cost-oriented optimization of the location and capacity of charging stations for the electric Robotaxi fleet. *Energy* 2023; 263: 125895.
3. Bibra EM, Connelly E, Dhir S, et al. Global EV outlook 2022: Securing supplies for an electric future, 2022, https://trid.trb.org/view/2005689
4. Xu D, Pei W and Zhang Q. Optimal planning of electric vehicle charging stations considering user satisfaction and charging convenience. *Energies* 2022; 15: 5027.
5. Jiaqi H, Qi D, Jun Z, et al. Research on planning and layout of electric taxi charging facilities in Shenzhen. *China Transp Rev* 2021; 43: 3–8.
6. Hu D, Huang L, Liu C, et al. Data driven optimization for electric vehicle charging station locating and sizing with charging satisfaction consideration in urban areas. *IET Renew Power Gener* 2022; 16: 2630–2643.
7. Cui X, Zhang Z, Gong D, et al. Allocation of electric taxi charging: assessing the layout of charging stations based on charging frequency. *Tehnički Vjesnik* 2022; 29: 1069–1079.
8. Xiao D, An S, Cai H, et al. An optimization model for electric vehicle charging infrastructure planning considering queuing behavior with finite queue length. *J Energ Storage* 2020; 29: 101317.
9. Jordan J, Palanca J, Marti P, et al. Electric vehicle charging stations emplacement using genetic algorithms and agent-based simulation. *Expert Syst Appl* 2022; 197: 116739.
10. Wood E, Jeremy N and Burton E. *Measuring the benefits of public chargers and improving infrastructure deployments using advanced simulation tools*. Report no. 0148-7191. National Renewable Energy Lab (NREL), Golden, CO, 2015.
11. Dong J, Liu C and Lin Z. Charging infrastructure planning for promoting battery electric vehicles: An activity-based approach using multiday travel data. *Transp Res Part C: Emerg Tech* 2014; 38: 44–55.
12. Cao X, Hu P and Liu D. Progress of research on electric vehicle charging stations. *Prog Geogr* 2019; 38: 139–152.
13. Liu X, Qu X and Ma X. Optimizing electric bus charging infrastructure considering power matching and seasonality. *Transport Res Part D: Transp Environ* 2021; 100: 103057.
14. Jahn RM, Syré A, Grahle A, et al. Methodology for determining charging strategies for urban private vehicles based on traffic simulation results. *Procedia Comput Sci* 2020; 170: 751–756.
15. Bischoff J, Márquez-Fernández FJ, Domingues-Olavarría G, et al. Impacts of vehicle fleet electrification in Sweden–a simulation-based assessment of long-distance trips. In: *2019 6th international conference on models and technologies for intelligent transportation systems (MT-ITS)*, Cracow, 5–7 June 2019, pp. 1–7. New York: IEEE.
16. Zhou G, Zhu Z and Luo S. Location optimization of electric vehicle charging stations: based on cost model and genetic algorithm. *Energy* 2022; 247: 123437.
17. Ma T-Y and Xie S. Optimal fast charging station locations for electric ridesharing with vehicle-charging station assignment. *Transp Res Part D: Transp Environ* 2021; 90: 102682.
18. Tian Z, Lai T, Yi W, et al. Impact of core charging station's cease operation in the entire charging station system: a case study in Shenzhen. In: *2017 IEEE third international conference on big data computing service and applications (Bigdataservice)*, Redwood City, CA, 6–9 April 2017. New York: IEEE.
19. Tayyab M, Helm S, Hauer I, et al. Infrastructure linking for placement of charging stations using Monte Carlo simulation. In: *2020 6th IEEE congress on information science and technology (CIST)*, Agadir, 5–12 June 2021, pp. 436–441. New York: IEEE.
20. Chen J, Li F, Yang R, et al. Impacts of Increasing Private Charging Piles on Electric Vehicles' Charging Profiles: A Case Study in Hefei City, China. *Energies* 2020; 13.
21. Vazifeh MM, Zhang H, Santi P, et al. Optimizing the deployment of electric vehicle charging stations using pervasive mobility data. *Transp Res Part A: Policy Pract* 2019; 121: 75–91.
22. Tao Y, Huang M and Yang L. Data-driven optimized layout of battery electric vehicle charging infrastructure. *Energy* 2018; 150: 735–744.
23. Zhang Y, Liu X, Zhang T, et al. Review of the electric vehicle charging station location problem. In: Wang G, Bhuiyan MZA, di Vimercati DCS, et al. (eds) *Dependability in sensor, cloud, and big data systems and applications*. Singapore: Springer, 2019.

24. Pagani M, Korosec W, Chokani N, et al. User behaviour and electric vehicle charging infrastructure: an agent-based model assessment. *Applie Energ* 2019; 254: 113680.
25. Márquez-Fernández FJ, Bischoff J, Domingues-Olavarría G, et al. Using multi-agent transport simulations to assess the impact of EV charging infrastructure deployment. In: *2019 IEEE transportation electrification conference and expo (ITEC)*, Detroit, MI, 19–21 June 2019, pp. 1–6. New York: IEEE.
26. Bischoff J and Maciejewski M. Agent-based simulation of electric taxicab fleets. *Transp Res Proced* 2014; 4: 191–198.
27. Weber A. *Alfred Weber's theory of the location of industries*. Chicago, IL: University of Chicago Press, 1929.
28. Hodgson MJ and Rosing KE. A network location-allocation model trading off flow capturing and p-median objectives. *Ann Oper Res* 1992; 40: 247–260.
29. Shier DR. A min-max theorem for p-center problems on a tree. *Transp Sci* 1977; 11: 243–252.
30. Church R and ReVelle C. *The maximal covering location problem* (Papers of the regional science association). Berlin: Springer, 1974, pp. 101–118.
31. Hosseini M and MirHassani S. Selecting optimal location for electric recharging stations with queue. *KSCE J Civil Eng* 2015; 19: 2271.
32. Hodgson MJ. A flow-capturing location-allocation model. *Geograph Anal* 1990; 22: 270–279.
33. Kuby M and Lim S. The flow-refueling location problem for alternative-fuel vehicles. *Socio Econ Plan Sci* 2005; 39: 125–145.
34. Gong D, Tang M, Buchmeister B, et al. Solving location problem for electric vehicle charging stations—a sharing charging model. *IEEE Access* 2019; 7: 138391–138402.
35. Li Y, Luo J, Chow C-Y, et al. Growing the charging station network for electric vehicles with trajectory data analytics. In: *2015 IEEE 31st international conference on data engineering*, Seoul, 13–17 April 2015, pp. 1376–1387. New York: IEEE.
36. Bai X, Chin K-S and Zhou Z. A bi-objective model for location planning of electric vehicle charging stations with GPS trajectory data. *Comput Ind Eng* 2019; 128: 591–604.
37. Yang Y, Zhang Y and Meng X. A data-driven approach for optimizing the EV charging stations network. *IEEE Access* 2020; 8: 118572–118592.
38. Liu Q, Liu J, Le W, et al. Data-driven intelligent location of public charging stations for electric vehicles. *J Clean Prod* 2019; 232: 531–541.
39. Jia J, Liu C and Wan T. Planning of the charging station for electric vehicles utilizing cellular signaling data. *Sustainability* 2019; 11: 643.
40. Yang J, Dong J and Hu L. A data-driven optimization-based approach for siting and sizing of electric taxi charging stations. *Transp Res Part C: Emerg Tech* 2017; 77: 462–477.
41. Tian Z, Lai T, Chen T, et al. Understanding battery degradation phenomenon in real-life electric vehicle use based on big data. In: *International conference on big data computing & communications*, Chengdu, China, 10–11 August 2017. New York: IEEE.
42. Palanca J, Terrasa A, Carrascosa C, et al. *SimFleet: a new transport fleet simulator based on MAS.* Cham: Springer, 2019, pp. 257–264.
43. WAxhausen K, Horni A and Nagel K. *The multi-agent transport simulation MATSim*. London: Ubiquity Press, 2016.
44. North MJ, Collier NT, Ozik J, et al. Complex adaptive systems modeling with Repast Simphony. *Complex Adapt Syst Model* 2013; 1: 1–26.
45. Malleson N, Evans A and Jenkins T. An agent-based model of burglary. *Environ Plan B: Plan Design* 2009; 36: 1103–1123.
46. Zhang Y and Brown D. Simulation optimization of police patrol districting plans using response surfaces. *SIMULATION* 2014; 90: 687–705.
47. Zhao J and Ma T. Optimizing layouts of initial refueling stations for alternative-fuel vehicles and experiments with agent-based simulations. *SIMULATION* 2016; 92: 251–266.
48. Davidson G, Fahlman A, Mereu E, et al. A methodological approach for modeling the spread of disease using geographical discrete-event spatial models. *SIMULATION* 2023; 2023: 00375497231152458.
49. Abadeer M, Magharious S and Gorlatch S. Modeling and interactive simulation of measures against infection transmission. *SIMULATION* 2023; 99: 327–346.
50. Ayadi A, Frydman CS, Laddada W, et al. Combining Devs and semantic technologies for modeling the SARS-CoV-2 replication machinery. In: *2021 annual modeling and simulation conference (ANNSIM)*, Fairfax, VA, 19–22 July 2021. New York: IEEE.
51. Iskandar R, Dugdale J, Beck E, et al. Agent-based simulation of seismic crisis including human behavior: application to the city of Beirut, Lebanon. *SIMULATION* 2023; 2023: 00375497231194608.
52. Risco-Martín JL, Esteban S, Chacón J, et al. Simulation-driven engineering for the management of harmful algal and cyanobacterial blooms. *SIMULATION* 2023; 99: 1041–1055.
53. Sahlaoui F-Z, Aboueljinane L and Lebbar M. A review on simulation-based metamodeling in emergency healthcare: methodology, applications, and future challenges. *SIMULATION* 2023; 99: 989–1009.
54. Wang G, Chen X, Zhang F, et al. Experience: understanding long-term evolving patterns of shared electric vehicle networks. In: *ACM/IEEE international conference on mobile computing and networking*, Vancouver, BC, Canada, 27–30 August 2019.
55. Zhang D, Zhao J, Zhang F, et al. UrbanCPS: a cyber-physical system based on multi-source big infrastructure data for heterogeneous model integration. In: *Proceedings of the ACM/IEEE sixth international conference on cyber-physical systems*, Seattle, WA, 14–16 April 2015, pp. 238–247. New York: IEEE.

Author biographies

Yue Zhang is currently a lecturer at the School of Management, Shenzhen Polytechnic University. In 2014, he graduated with a Ph.D. in systems engineering from the University of Virginia in the United States. His research interests include agent-based simulation in a GIS environment and simulation optimization algorithms. He has published research articles in Simulation: Transactions of the Society for Modeling and Simulation International and Proceedings of the Winter Simulation Conference. After completing his Ph.D. degree, he spent seven years in financial and e-commerce companies, such as eBay and Pingan, focusing on machine learning and artificial intelligence-related work, and received several patents during this period.

Jie Tan is now a lecturer at the School of Management, Shenzhen Polytechnic University. He received his Ph.D. degree in management science from Huazhong University of Science and Technology, China. Then he worked as a postdoctoral at Sun Yat-sen University from 2019 to 2021, concentrating on the research of logistics and supply chain management, production, and operation management. He has published research articles in international journals such as Physica A: Statistical Mechanics and its Applications, Transportation Research Part E: Logistics and Transportation Review, and Journal of Cleaner Production.