https://doi.org/10.1016/j.jocs.2019.101046

Agent- and Activity- based Large-Scale Simulation of Enroute Travel, Enroute Refuelling and Parking Behaviours in Beijing, China

Chengxiang Zhuge [1,2,3], Chunfu Shao [4,*] Xiong Yang [4]

[1] Department of Geography, University of Cambridge, Downing Place, Cambridge, UK

[2] Tyndall Centre for Climate Change Research, University of East Anglia, Norwich, UK

[3] Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, China

[4] Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, 3 Shangyuancun, Xizhimenwai, Beijing, China

* Corresponding Author: cfshao@bjtu.edu.cn

Abstract

This paper develops an agent- and activity-based large-scale simulation model for Beijing, China (MATSim-Beijing) to explicitly simulate enroute travel, enroute refuelling and parking behaviours, as well as the associated vehicular energy consumption and emissions, based on MATSim (Multi-Agent Transport Simulation), which is a typical integrated activity-based model. In order to take into account heterogeneous parking and refuelling behaviours, the MATSim-Beijing model incorporates several Multinomial Logit (MNL) models to predict individual choices about the maximum acceptable times of walking from trip destination to parking lot, of diverting to a refuelling station and of queuing at a station, using the data collected in a paper-based questionnaire survey in Beijing. A Sensitivity Analysis (SA) -based calibration method was used to estimate the model parameters by searching for an optimal parameter combination with the objective of minimize the gap between simulated and observed traffic flow data, exhibiting a relatively good performance of decreasing the Mean Absolute Percentage Error (MAPE) by around 23%. Further, the calibrated model was used to investigate whether and how the population scaling and network simplification, which were two commonly used approaches to speeding up large-scale traffic simulations, might influence model accuracy and computing time. The results indicated that both approaches could to some extent influence model outputs, though they could significantly reduce computing time.

Keywords: agent-based modelling; activity-based model; model calibration, population scaling; network simplification; Beijing

1 Introduction

1.1 Activity-based Travel Demand Modelling

Travel demand analysis and modelling is the fundamental research question in transport studies, as travel demand is associated with a large variety of decision-makings, for example, in policy making and infrastructure planning. Traditionally, the four-step method (or trip-based model), was used for the analysis and modelling of travel demand (Ahmed, 2012; McNally, 2008). However, this method is limited in behavioural inadequacy (Bhat and Koppelman, 1999). In response, many efforts have been made to develop activity-based travel demand models (Rasouli and Timmermans, 2014), which look at the travel demand of each individual at the micro level. The basic concept of activity-based model is that people need to travel in order to perform their daily activities (e.g., work and shopping). In other words, travel is considered as a demand derived from the need to perform activities (Pinjari and Bhat, 2011).

Activity-based modelling is generally coupled with agent-based modelling (Macal and North, 2010), which is another approach to simulating individual behaviours in complex dynamic system, such as transport systems (Bazzan and Klügl, 2014; Zhuge, 2019): agent-based modelling can simulate heterogenous behaviours (e.g., parking and refuelling behaviours), according to agents' attributes (e.g., income); Furthermore, agents also interact with each other in their daily activities and travel, and the interactions can be either direct or indirect. For example, people from the same household may negotiation with each other in allocating maintenance activities (e.g., shopping) (Srinivasan and Bhat, 2005), and also household members may perform joint activities and travel (e.g., car sharing) (Bradley and Vovsha, 2005).

In many cases, activity-based model was also coupled with Dynamic Traffic Assignment (Peeta and Ziliaskopoulos, 2001), resulting in an integrated activity-based model able to simulate how people perform their daily activities and travel from one activity location to another through transport networks (e.g., road and public transport networks) (Zhuge et al., 2019c). The outputs of the integrated model include traffic/passenger flow and daily plans of each person which contain both activity and travel information. Among the integrated models, MATSim appears to be more favourable to transport modellers and planners, as evident from its numerous case studies across the global (Horni et al., 2016). Therefore, this paper chooses MATSim as the base to develop the Beijing activity-based travel demand model (or the MATSim-Beijing model). Although the model is particularly for Beijing, the approaches to be introduced below (including parking and refuelling models and calibration method) can also be applied into other large-scale scenarios.

1.2 Parking Behaviour Modelling

For a car trip, individual travel behaviour can be decomposed into three parts, namely enroute travel, (possible) enroute refuelling and parking behaviours. However, most of the activity-based simulations were focused on enroute travel, paying significantly less attention to parking and refuelling behaviours. Some attempts have been made to explicitly simulate parking behaviour within activity-based models (Benenson et al., 2008; Dieussaert et al., 2009; Horni et al., 2013; van der Waerden, 2012), resulting in the spatiotemporal distribution of parking demand and the usage of each parking lot. Such spatially and temporally explicit results would be particularly useful for the design of parking-related policies (e.g., parking fee) and optimizing the layout of parking lots.

MATSim has a parking extension (Bischoff and Nagel, 2017; Horni et al., 2016), which is able to explicitly simulate individual parking behaviour. A utility function is used to score and rank candidate parking lots for each driver when they approach trip destinations, and divers are assumed to always choose the parking lot with the highest utility. Agents may fail to park due to the limited parking capacity and they are assumed to keep searching for available parking spaces with the searching radius being increased gradually (Horni et al., 2016). However, this assumption is unrealistic, as agents may only try those parking lots within a specific radius around the trip destinations, considering the maximum acceptable walking time (or distance) from their trip destinations and parking lots. Therefore, the MATSim-Beijing model will incorporate a constraint on walking time/distance: those parking lots falling out of the range will not be listed as candidates. As a result, agents may fail to find a parking space and have to park illegally around their destinations, which tends to be more realistic. In order to take heterogeneous parking behaviours into account, the proposed MATSim-Beijing model will incorporate a Multinomial Logit (MNL) model to predict the maximum acceptable walking time for each driver according to their attributes (e.g., income), using the data collected in a questionnaire survey in Beijing (see Section 3.2 for more details).

1.3 Refuelling Behaviour Modelling

Simulating enroute refuelling behaviour could be useful for quantifying the demand for refuelling stations, which can be further used to locate new refuelling stations or optimize the existing layout of refuelling stations (Zhuge and Shao, 2018). Compared with parking modelling, enroute refuelling modelling has received relatively scant attention, though the charging behaviour of electric vehicle drivers at trip destinations (e.g., at home or workplaces) have been explicitly modelled (Galus et al., 2012; Gonzalez et al., 2014; Knapen et al., 2011; Knapen et al., 2012; Waraich, 2013; Waraich et al., 2014). This is likely because electric vehicles have received increasing attention over the past few years across the global (IEA, 2019; Zhuge and Shao, 2019). However, refuelling/charging behaviours of conventional and electric vehicles can be significantly different: on one hand, individual behaviour can be influenced by vehicle types used, and the approaches to simulating charging behaviour cannot be straightforwardly applied into the modelling of refuelling behaviour; on the other hand, the destination-based charging behaviour of electric vehicle could differ from the enroute refuelling behaviour. Therefore, the proposed MATSim-Beijing model will simulate how drivers refuel their conventional vehicles (e.g., petrol car) on their journeys, with a focus on the behaviour of choosing refuelling stations. In order to take heterogeneous refuelling behaviours into account, the MATSim-Beijing model will incorporate another two MNL models to predict the maximum acceptable diverting and queueing times for each driver, according to their attributes, again using the data collected in the same questionnaire survey mentioned above.

1.4 Approaches to Speeding Up Large-Scale Simulations

This study will use Beijing as a case study, which had a population of around 20 million in 2010. It would be computationally expensive to explicitly simulate travel behaviour of all individuals. Essentially, the high computing time is largely attributed to the large number of agents involved, complex and detailed individual travel behaviour, and large and complex road network. Many approaches have been used to speed up such large-scale agent- and activity-based simulations, including parallel computing, efficient traffic simulators (Charypar et al., 2006; Charypar et al., 2007; Waraich et al., 2015) and fast shortest-path searching algorithms (Balmer, 2007; Lefebvre and Balmer, 2007). Among them, population scaling and network simplification are two commonly used approaches. Specifically, population scaling is to scale down the population (using a small portion of the population) in the simulation and then to scale up the simulation results accordingly (Flötteröd et al., 2012; Horni et al., 2016). Network simplification is to simplify the road network by merging two or more adjacent links with some specific rules (or constraints), resulting in a simplified network with less nodes and links. As a result, the traffic simulator and shortest-path search algorithm in the integrated activity-based model will process less nodes and links and thus can speed up the simulations. However, the potential influences of these two approaches on the model results are not well understood. In order to quantify the influence, the proposed MATSim-Beijing model will be tested with different population scaling rates and simplified networks.

1.5 Calibration of Activity-based Models

Activity-based models were generally calibrated with the objective of minimizing the gap between the resulting and observed traffic flow data (Agarwal et al., 2017; Flötteröd et al., 2012; Gonzalez et al., 2014). Cadyts (Calibration of dynamic traffic simulations) appears to be one of the most-used calibration tools for such integrated activity-based models (Flötteröd et al., 2012; Gonzalez et al., 2014; Horni et al., 2016), and has been applied into several MATSim-based scenarios (Flötteröd et al., 2011; Flötteröd et al., 2012; Horni et al., 2016). Cadyts uses a Bayesian framework to select those optimal daily plans in each agent' memory, with the objective of minimizing the gap between the simulated and observed traffic flow data (Flötteröd, 2009; Flötteröd et al., 2011; Flötteröd et al., 2012).

However, Cadyts does not search for an optimal set of model parameters, and thus is not behaviourally sound. In response, this paper will try another Sensitivity Analysis (SA)-based calibration method, which is capable of searching for an optimal parameter combination in an efficient way, based on the results of parameter SA (Saltelli et al., 2008; Zhuge et al., 2019c).

1.6 Research Gaps and Aims

Activity-based model has become a promising approach to investigating travel demand at the individual level. In general, activity-based models were focused on enroute travel behaviour, but some of them were extended to incorporate refuelling and parking behaviours. This paper will develop an agent- and activity-based model to simulate enroute travel, refuelling and parking behaviours, as well as the associated vehicular energy consumptions and emissions in Beijing, based on MATSim, resulting in an MATSim-Beijing model. The model can output rich fine-grained results, which would be useful for policy analysis (e.g., parking fee) and infrastructure planning (e.g., layouts of refuelling stations and parking lots).

The contributions of this paper are twofold: First, the MATSim-Beijing model will integrate several Multinomial Logit (MNL) models to simulate the decision-making of drivers on parking and refuelling, in order to be more behaviourally realistic and thus more accurate. Second, this paper will also evaluate the performance of some methods (in terms of model accuracy and efficiency), which aim to make large-scale simulations computationally feasible, including the Sensitivity Analysis (SA)-based calibration method, population scaling method and network simplification. In brief, both contributions would make integrated activity-based large-scale simulations more behaviourally sound, computationally feasible and accurate.

2 Methodology

2.1 Framework of MATSim-Beijing

Originally, MATSim is composed of three core modules (see Figure 1), namely Execution, Scoring and Replanning, which are used to optimize daily plans through iterations, considering the influence of dynamic traffic flow on activity scheduling (Horni et al., 2016). Each module is introduced in detail in the official MATSim book by Horni et al. (2016). The original MATSim was updated to the MATSim-Beijing model by incorporating parking and refuelling extensions, as well as vehicular energy consumption and emission modules. To be behaviourally realistic, three MNL models were developed and integrated into the MATSim-Beijing model. The MNL models were used to predict parking- and refuelling- related choices of drivers, including the maximum acceptable time of walking from parking lot to trip destination, of diverting to a refuelling station and of queueing at a station (Zhuge et al., 2019a; Zhuge et al., 2019b).

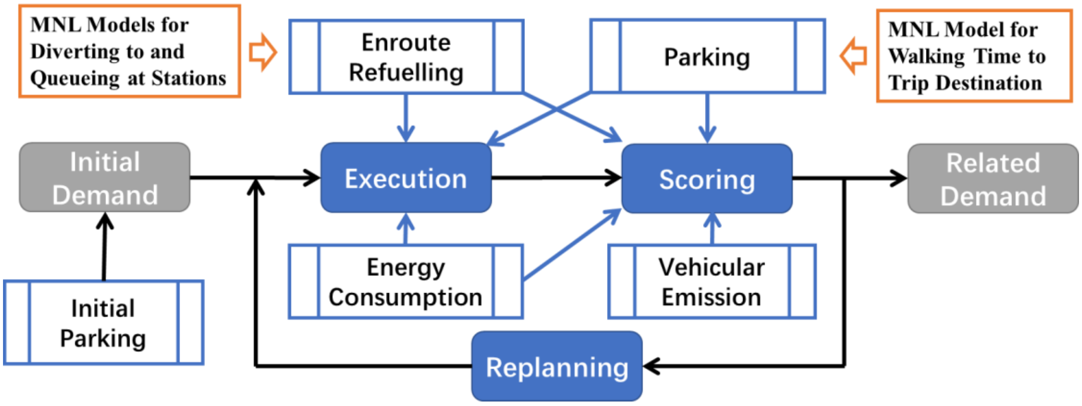


Figure 1 Framework of MATSim-Beijing (Source: Adapted from (Waraich, 2013) and (Zhuge and Shao, 2018))

In particular, the Scoring module, which contains a utility function (see Equation (1)), is used to evaluate the performance of each daily plan in terms of travelling and performing activities (Horni et al., 2016).

$$U_{DailyPlan} = \sum_{j=0}^{J} U_{Activity,j} + \sum_{j=0}^{J} U_{Travel,j} \tag{1}$$

Where, $U_{DailyPlan}$ denotes the utility of a daily plan; $U_{Activity,j}$ denotes the utility of performing j th activity; $U_{Travel,j}$ denotes the utility of j th trip between two activity locations (Zhuge et al., 2019c).

In order to simulate parking and refuelling behaviours, the original travel utility (equivalent to the utility of enroute travel here, $U_{EnrouteTravel,j}$) needs to be extended by incorporating the utilities of parking and enroute refuelling ($U_{Parking,j}$ and $U_{EnrouteRefuelling,j}$), resulting in a new travel utility function, as presented by Equation (2). $U_{Parking,j}$ and $U_{EnrouteRefuelling,j}$ will be introduced in Sections 2.2 and 2.3, respectively. More detailed introductions to $U_{Activity,j}$ and $U_{EnrouteTravel,j}$ can be found in the MATSim book (Horni et al., 2016).

$$U_{Travel,j} = U_{EnrouteTravel,j} + U_{Parking,j} + U_{EnrouteRefuelling,j} \tag{2}$$

2.2 Agent-based Modelling of Parking Behaviour in Beijing

The new parking model in MATSim-Beijing was developed based on the parking extension of MATSim (Horni et al., 2016; Waraich, 2013; Waraich et al., 2012), as well the empirical findings and conceptual model by (Zhuge et al., 2019b). Essentially, the parking model is used to simulate how agents choose parking lots at their trip destinations. The new parking model mainly differs from the parking extension of MATSim in considering 1) the initial target parking lots, 2) failing to park and 3) the constraint on the maximum acceptable time of walking from parking lot to trip destination (Zhuge and Shao, 2018).

Specifically, an initial target parking lot will be selected for each driver agent when the model is initialised, given a set of candiate parking lots that fall into the maximum acceptable walking time of the agent ($T_{Walking}$) (Zhuge and Shao, 2018). As illustrated in Figure 2, Parking Lots A, B and D fall into the range and thus will become candidate parking lots to this agent. The maximum acceptable walking time ($T_{Walking}$) may vary from one agent to another. In order to take heterogeneity into account, a Multinomial Logit (MNL) Model will be developed to predict the walking time for each driver agent, according to their attributes (e.g., sex and income), based on the empirical findings in (Zhuge et al., 2019b). Among the candiate parking lots, the one with the highest utility, which can be calcualted by Equation (3), will be selected as the initial target parking lot. In this example, Parking Lot A will be selected as the initial target parking lot.

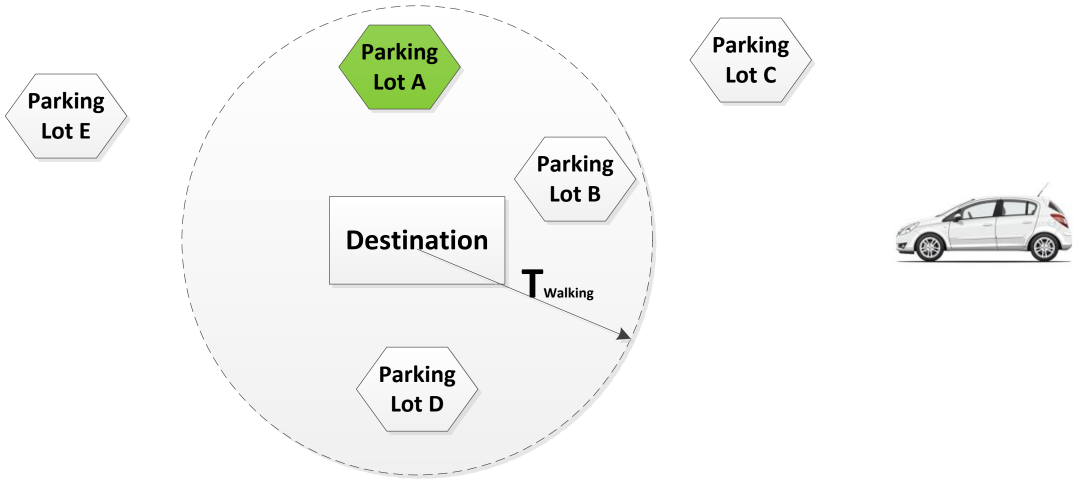


Figure 2 Illustration of Searching for a Parking Lot at Trip Destination

In the simulation, agents will first drive directly to their initial target parking lots, and they may fail to park when all parking spaces are occupied. Then they will try other candiate parking lots (Parking Lots B and D in the example above) to see if there are any parking spaces avaialble. If so, then the agents will choose the new parking lot instead; Otherwise, they may have to park illegally at their trip destinations, which will introduce a big negative utility into the utility of parking ($U_{parking}$).$U_{parking}$ is composed of the utilities for walking ($U_{Walking}$), diverting ($U_{Diverting,Parking}$) and parking fee ($U_{Parking\ fee}$), as presented by Equation (3) (Horni et al., 2016; Waraich, 2013; Zhuge et al., 2019b).

$$\begin{aligned} U_{parking} &= U_{Walking} + U_{Diverting,Parking} + U_{Parking\ fee} \\ &= \beta_{Walking} \cdot t_{Walking} \cdot V_{Time} + U_{EnrouteTravel} + \beta_{Parking\ fee} \cdot m_{Parking\ fee} \end{aligned} \tag{3}$$

Where, $U_{Walking}$ is the utility of walking from parking lot to trip destination, which is a function of walking time ($t_{Walking}$) and individual time value (V_{Time}); $U_{Diverting,Parking}$ denotes the negative utility for diverting when an agent fails to park at its target parking lot. $U_{Diverting,Parking}$ is assumed to be equivalent to the utility of enroute travel ($U_{EnrouteTravel}$); $U_{Parking\ fee}$ is the utility of parking fee ($m_{Parking\ fee}$). $\beta_{Walking}$, $\beta_{Parking\ fee}$ and $\beta_{Chargingposts}$ are the parameters to be estimated.

2.3 Agent-based Modelling of Refuelling Behaviour in Beijing

The refuelling model in the MATSim-Beijing is used to simulate the refuelling behaviour of a driver agent when it becomes aware of the low states of fuel, based on the empirical findings by Zhuge et al. (2019a), involving in the behaviours of choosing a refuelling station and queueing at a station. The assumption here is that only those refuelling stations, which fall into the maximum acceptable time of diverting to a station ($T_{Diverting}$) and the maximum acceptable time of queueing at a station ($T_{Waiting}$), can become its candidate stations. In other words, those stations which either are far away or have many vehicles queueing, will not be considered. As illustrated in Figure 3, only Refuelling Stations A, B and C fall into the driver's maximum acceptable diverting time ($T_{Diverting}$). However, Station B does not fall into its maximum acceptable waiting time ($T_{Waiting}$) due to too many vehicles queueing. In addition, both $T_{Diverting}$ and $T_{Waiting}$ may vary from one agent to another. In order to take heterogeneity into account, two MNL models will be developed to predict $T_{Diverting}$ and $T_{Waiting}$, for each driver agent according to their attributes, using the empirical findings in (Zhuge et al., 2019a). Furthermore, among the candidate stations, the driver agent is assumed to always choose the one with the highest utility. The utility function (see Equation (4)) comprises a utility of diverting to a refuelling station ($U_{Diverting,Refuelling}$, equivalent to $U_{EnrouteTravel,\,j}$), utility of queueing at a refuelling station ($U_{Waiting}$) and utility for the duration of refuelling a vehicle ($U_{Refuelling}$). Here, individual time value (V_{Time}) will also be considered.

$$\begin{aligned} U_{EnrouteReplenishing} &= U_{Diverting,Refuelling} + U_{Waiting} + U_{Refuelling} \\ &= U_{EnrouteTravel} + \beta_{Waiting} \cdot t_{Waiting} \cdot V_{Time} + \beta_{Refuelling} \cdot t_{Refuelling} \cdot V_{Time} \end{aligned} \tag{4}$$

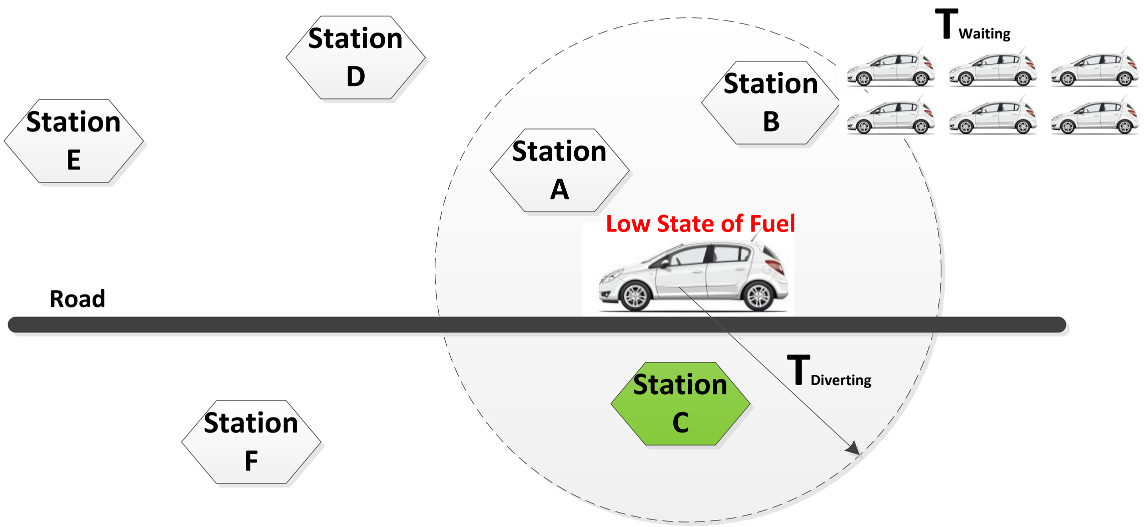


Figure 3 Illustration of Searching for a Refuelling Station on a Journey

2.4 Energy Consumption and Emission Models for Vehicles in Beijing

The MATSim-Beijing model also incorporates an energy consumption factor (FC , kg/km) and emission factor (EF , g/km), which are used to simulate the energy consumptions and vehicular emissions, respectively, as presented by Equations (5) and (6) (Zhuge et al., 2019d). We employed the factors developed by Yao and Song (2013) and Yao et al. (2013) for Beijing. The energy consumption model is linked to refuelling behaviour model, as drivers need to decide when to refuel based on the state of energy. The information on vehicular emissions can be further used for both global and local environmental impact assessments. In a simulation, the energy consumptions and vehicular emissions can be calculated when a vehicle leaves a link. Therefore, the disaggregate results by vehicle can be further aggregated into link-based statistics. Both FC and EF are functions of average speed, and the model parameters can be estimated based on real driving experiments, as detailed in Yao and Song (2013) and Yao et al. (2013).

$$FC = \frac{a}{v} + b \cdot v + c \cdot v^2 + d \tag{5}$$

$$EF = \frac{e}{v} + f \cdot v + g \cdot v^2 + h \tag{6}$$

2.5 Sensitivity Analysis (SA)-based Calibration Method

As reviewed in Section 1.5, Cadyts is one of the most-used calibration tools for MATSim (Flötteröd, 2009; Flötteröd et al., 2011; Flötteröd et al., 2012). Due to its behavioural inadequacy, this paper will use a Sensitivity Analysis (SA)-based method instead. The method tries to search for an optimal parameter combination which can minimize the gap between the simulated and observed data (Zhuge and Shao, 2018; Zhuge et al., 2019b; Zhuge et al., 2019c), so as to be behaviourally sound.

Specifically, the SA-based calibration here will only four influential model parameters to generate the candidate parameter combinations, including the probability of performing reroute when agents adjust their daily plans, the time step used in the traffic simulation (or the Execution module), the number of iterations and the scaling factor for the flow capacity, which were identified through both global and local SAs (Saltelli et al., 2008) within a smaller scale Chinese city, Baoding.

3 Case Study: Beijing, China

3.1 Scenario Description

Beijing, China was used as a case study. The MATSim-Beijing model was applied to simulate the enroute travel, enroute refuelling and parking behaviours in 2010. An agent- and GIS-based virtual creator by (Zhuge et al., 2018) was used to generate input data and initialize the model, primarily using the 2010 Household Travel Survey Data. The resulting virtual Beijing contains a synthetic population, physical environment, and linkages between agents and between agents and facilities. The synthetic population contains persons and households, as well as their characteristics (e.g., gender, age and come); Each person has a daily plan which contains detailed travel and activity information (e.g., activity location, type and duration). The physical environment is composed of a road transport network, activity facilities (e.g., residential buildings and leisure activities) and transport facilities (e.g., parking lots and refuelling stations). Note that the private parking lot generator was updated here to additionally consider those parking demand which is associated with the first car-based trips of each driver but is not based at their homes. More detailed introduction to the 2010 virtual Beijing can be found in the work of (Zhuge et al., 2018). It should be noted that a population scaling factor of 4% was used. This means that all the associated elements in the simulation, including the population and physical environment, need to be scaled down. Post-simulation the results need to be scaled up accordingly (Zhuge et al., 2019d). The potential impacts of population scaling on the model results will be assessed later in Section 3.5.

3.2 Survey-based Calibration of Refuelling and Parking Modules

Prior to calibrating the Beijing-MATSim model against the traffic flow data (see Section 3.3), the data on parking and refuelling behaviours, which was collected from a paper-based questionnaire survey in Beijing from September, 2015 to March, 2016, was firstly used to calibrate the Multinomial Logit (MNL) models (see Sections 2.2 and 2.3), as detailed in (Zhuge et al., 2019a) and (Zhuge et al., 2019b). Note the survey was on the parking and refuelling behaviour of both conventional and electric vehicles, but only the results about conventional vehicles were used here for parameter estimations. The calibrated MNL models were further used to predict individual choices about the maximum acceptable times of walking from parking lot to trip destination ($T_{Walking}$), of diverting to a refuelling station ($T_{Diverting}$) and of queueing at a station ($T_{Waiting}$).

The hitting ratio, which is the ratio of the number of accurately predicted cases to the total number, is a commonly used indicator to describe the accuracy of an MNL model. However, the MNL model generally has relatively low prediction accuracy (or small hitting ratio), especially in those cases where the number of alternatives is large. This paper therefore proposes another indicator of Hitting Degree (HD), which can be viewed as a variant of hitting ratio but is looser. HD indicates the degree to which the predicted choice is different from the observed one, which is mathematically formulated as Equation (7). If the model makes an exact prediction, then HD will be zero.

$$HD = \frac{|A_{Perdicted} - A_{Observed}|}{A_{Max} - A_{Min}} \tag{7}$$

Where, $A_{Perdicted}$, $A_{Observed}$, A_{Max} and A_{Min} denote the predicted, observed, maximum and minimum choices, respectively.

Table 1 shows the prediction accuracies of the three different MNL models for the maximum acceptable walking time (when parking), diverting time (when refuelling) and waiting time (when refuelling). The predictive ability of each MNL model was quantified through cross- validation using the two indicators, namely hitting ratio and hitting degree. Specifically, in order for cross-validation, the sample was divided into training and validation datasets, which accounted for 80% and 20%, respectively. Taking walking time for example, the hitting ratio is 49.1%, meaning that 49.1% of the predictions could exactly find the right parking choice about walking time. Furthermore, the hitting degree is 14.8%, meaning that the predicted alternative on average is less than one-alternative bias, which is 33.3% in this four-alternative case, suggesting that the predictive ability of the MNL model is relatively acceptable. The hitting ratios for the diverting and waiting times are 38.2% and 38.0%, respectively, which are relatively low, but their hitting degrees are acceptable (17.3% and 16.3%, respectively). Overall, the predictive ability of the MNL models is not very strong (in part because of the relatively high number of alternatives), but is satisfactory (in terms of hitting degrees). Therefore, the MNL models could be used to predict the parking and refuelling choices in the MATSim-Beijing simulation.

Table 1 Prediction Accuracies of MNL Models for Walking, Diverting and Waiting Times

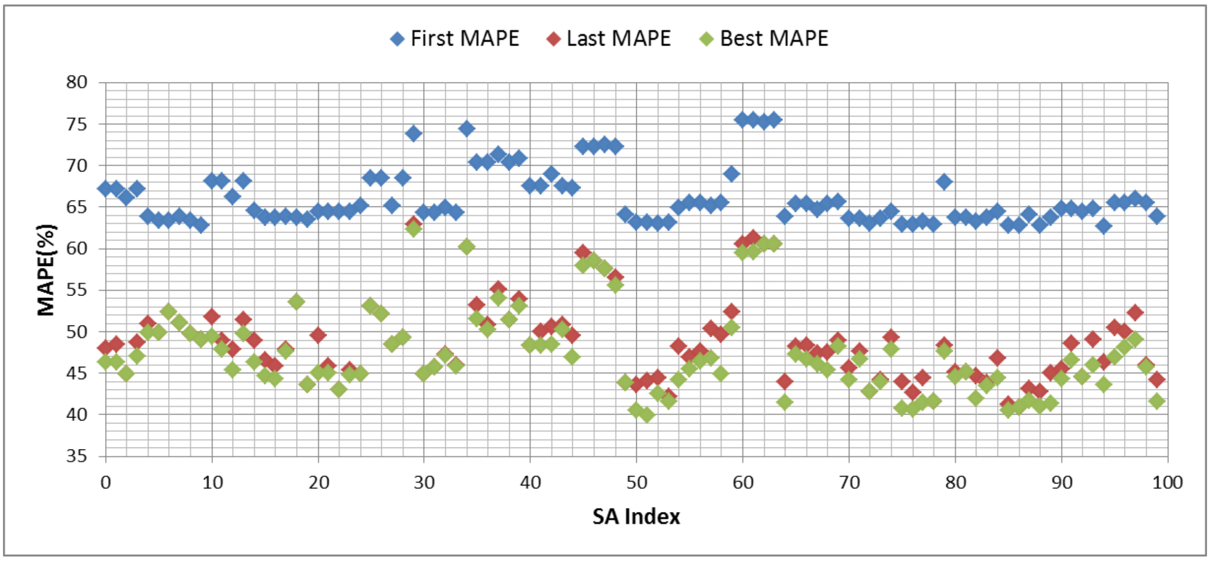
	Hitting Ratio	**Hitting Degree**	**Alternatives (or Choices) in the MNL Model**
Walking Time (Parking Module)	49.1%	14.8%	Choices 1-4: 5, 10, 15 and 20 min;
Diverting Time (Refuelling Module)	38.2%	17.3%	Choices 1-5: 3, 5, 10, 15, and 20 min
Waiting Time (Refuelling Module)	38.0%	16.3%	Choices 1-5: 0, 3, 5, 10 and 15 min

3.3 Sensitivity Analysis (SA)-based Model Calibration

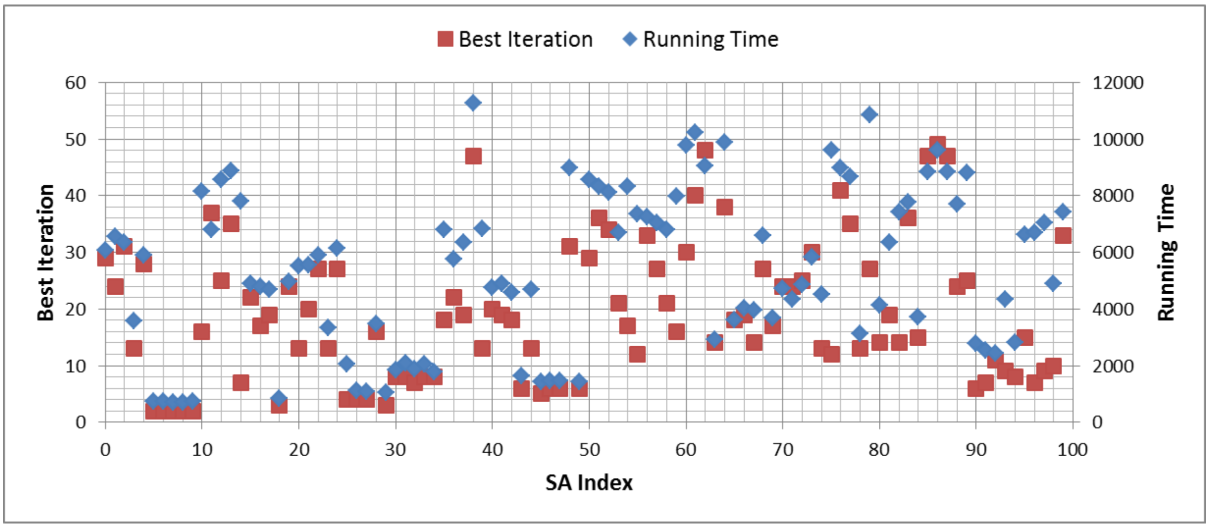
The Sensitivity Analysis (SA)-based calibration here is to search for an optimal parameter combination that can minimize the gap between simulated and observed traffic flow data, given 100 parameter combinations generated by only varying the four most influential parameters that were identified in SAs, as introduced in Section 2.5. The observed traffic flow data for the calibration was collected on 144 links at the peak hour (9AM) in 2010.

The calibration results of MATSim-Beijing are shown by Figure 4. Specifically, Figure 4-(a) presents the first, best and last Mean Absolute Percentage Errors (MAPEs) of each parameter combination. Figure 4-(b) shows the corresponding total computing time and the best iteration numbers with lowest MAPEs. Here, MAPE and computing time are used to quantify the model accuracy and efficiency, respectively. In some cases, a trade-off between model accuracy and efficiency needs to be made. It can be found from Figure 4-(a) that the best MAPEs range widely from around 40% to 62%, suggesting that the SA-based calibration method works effectively and is able to find the optimal parameter combination minimizing the gap between the simulated and observed data. In addition, there appears to be no significant relationship between the calibration performance and iteration number, as clearly evident from the parameter combinations indexed from 50 to 59 that have relatively small MAPEs, but have different best iteration numbers ranging from around 10 to 40. The best parameter combination indexed as 51 has the smallest MAPE of 39.9%.

As aforementioned, Cadyts is a typical calibration tool for dynamic traffic simulator, especially for MATSim. In the case study of the city of Zurich (Horni et al., 2016), Cadyts was able to decrease the MAPEs from around 30% to 15% for the period from 8AM to 7PM. The SA-based calibration method here has comparatively good performance. Specifically, the difference between the smallest and first MAPEs of the best parameter combination is around 23%, compared with the decrement of 15% in the case study of Zurich. In addition, the SA-based calibration method has some other advantages (e.g., more behaviourally sound) over Cadyts, as discussed in Section 2.5.



(a) First, Best and Last MAPEs of Each Parameter Combination

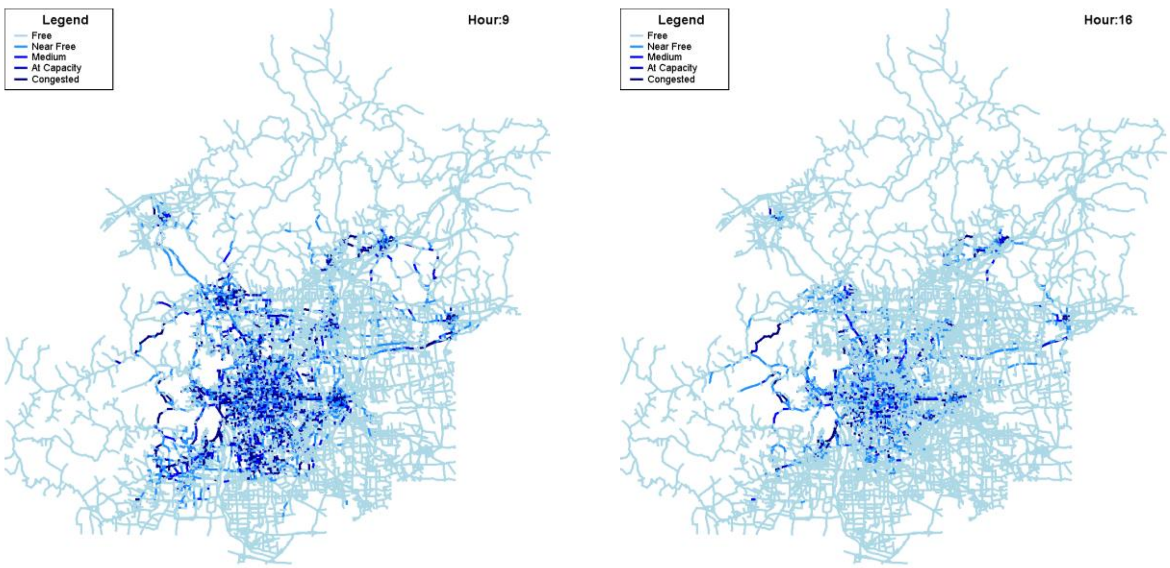


(b) Running Time and Best Iteration Numbers of Each Parameter Combination

Figure 4 SA-based Calibration Results of the MATSim-Beijing Model

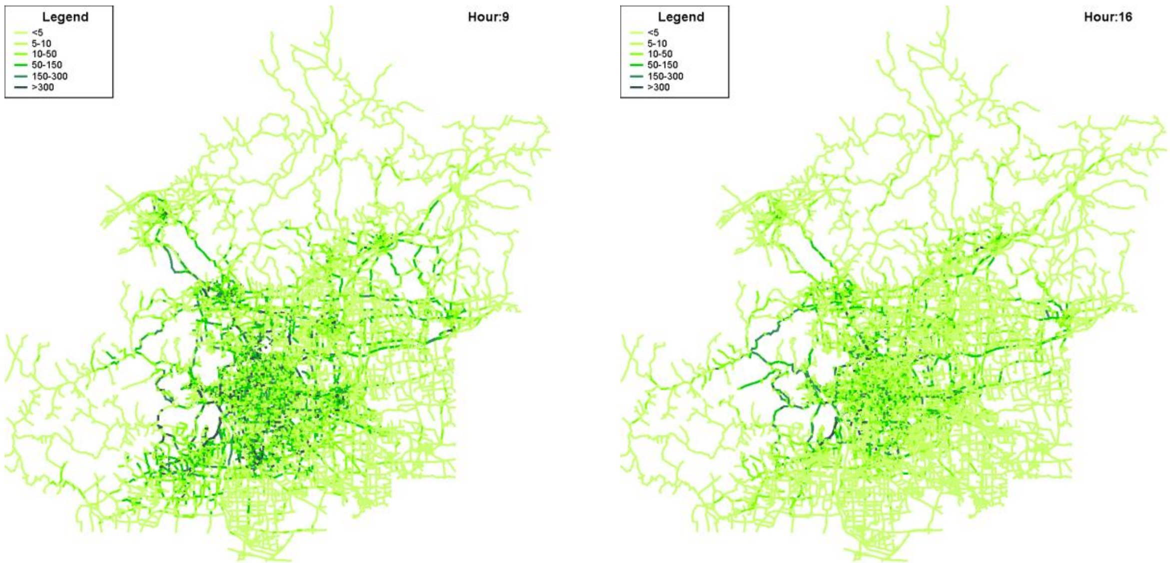
3.4 Simulation Results

As the MATSim-Beijing simulation is able to trace the moving trajectories of each agent, as well as the usage of transport facilities, such as parking lots and refuelling stations, lots of useful information can be extracted from the simulation results through aggregation and can be further used, for example, to optimize the layout of transport facilities (e.g., refuelling stations). In this case study, the hourly link states, hourly link-based vehicular emissions, and usage of parking lots and refuelling stations were aggregated and mapped. Figure 5 and Figure 6 respectively show the link states and link-based vehicular emissions at a morning peak hour (8-9AM) and an afternoon off-peak hour (3-4PM). The detailed link states and vehicular emissions throughout the whole day are shown in the Supplementary Materials. It can be clearly seen that the differences between peak and off-peak hours in traffic flow and vehicular emissions are very significant. Figure 7 spatially presents the average occupied time and average number of vehicles served at both private and public parking lots. Note that those parking lots with no vehicles served were not mapped here. By comparing Figure 7-(a) and -(c), it can be found that the majority of parking lots are occupied for a long time, but the public parking lots appear to be occupied longer than the private ones, as evident from more dark green dots in Figure 7-(a). For the average numbers of vehicles served at public and private parking lots (see Figure 7-(b) and -(d), respectively), those parking lots with more vehicles served tend to be located at the central districts, and the public parking lots appear to serve more vehicles than the private ones. Figure 8 maps the usage of refuelling stations in terms of average waiting time and average numbers of vehicles served at refuelling stations. It can be found from Figure 8-(a) that the average waiting times are relatively long for most of the refuelling stations, and those refuelling stations with less waiting time are mostly located at the outer districts. According to Figure 8-(b), the refuelling stations at the central districts, on average, appear to serve more vehicles than those located at the outer districts.



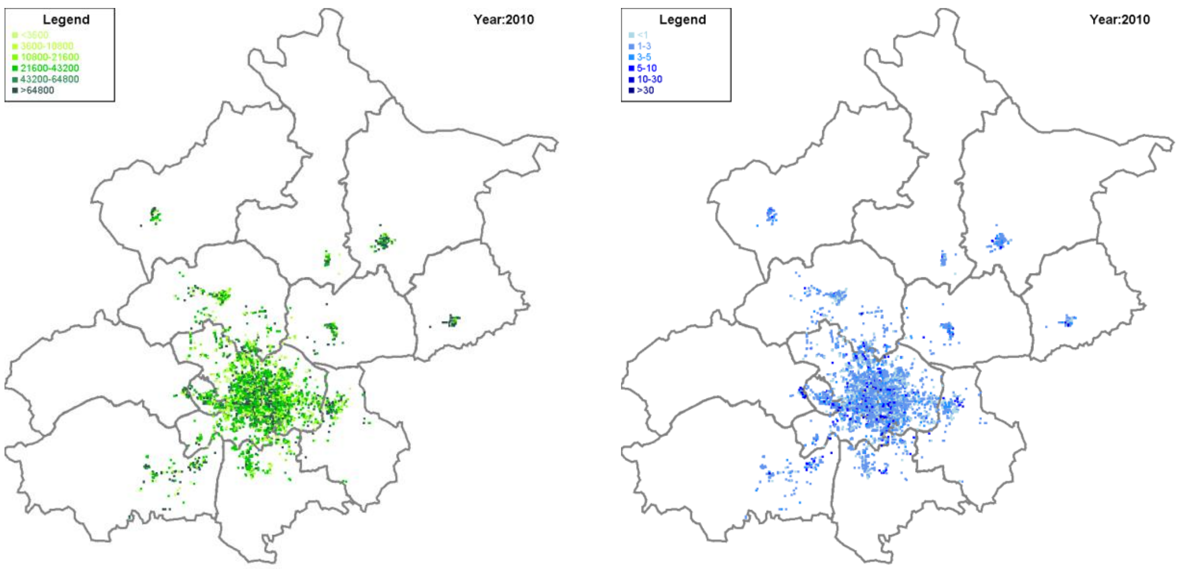
(a) 8-9AM (b) 3-4PM

Figure 5 Link States by Hour



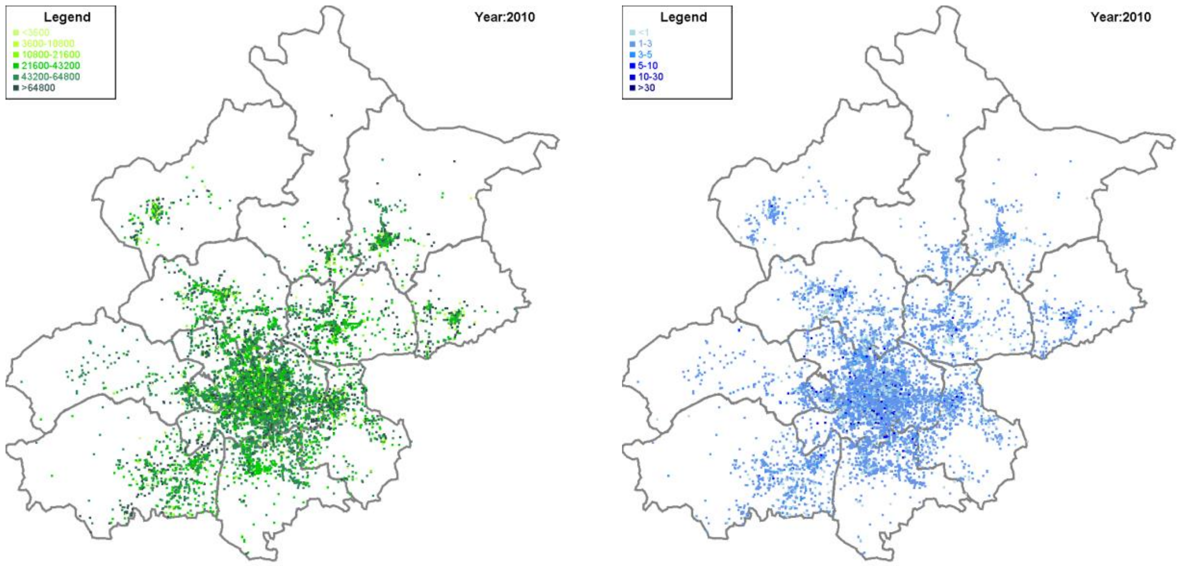
(a) 8-9AM (b) 3-4PM

Figure 6 Link-based Vehicular Emissions by Hour (Kilogram)



(a) Average Occupied Time of Public Parking Lots (seconds)

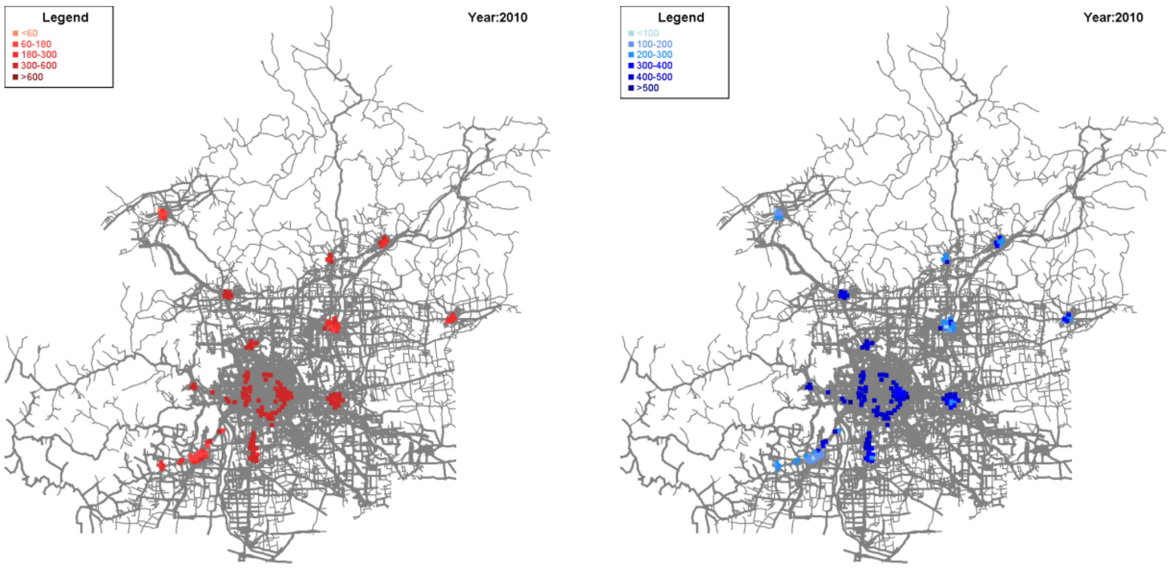
(b) Average Numbers of Vehicles Served at Public Parking Lots



(c) Average Occupied Time of Private Parking Lots (seconds)

(d) Average Numbers of Vehicles Served at Private Parking Lots

Figure 7 The Usage of Parking Lots



(a) Average Waiting Time at Refuelling Stations (seconds)

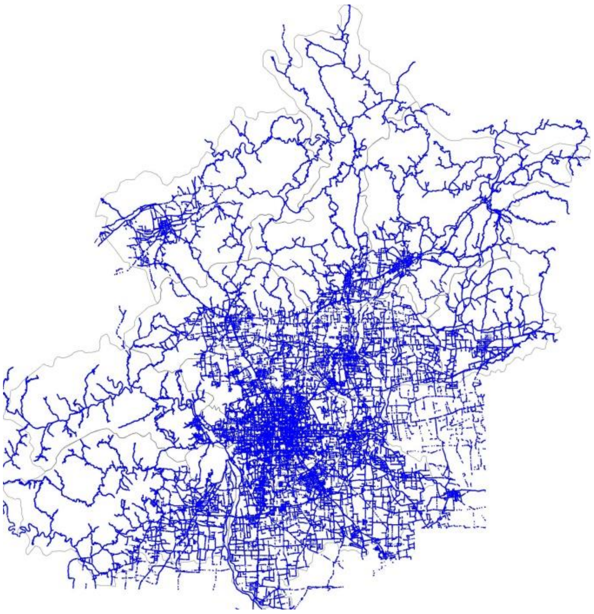
(b) Average Numbers of Vehicles Served at Refuelling Stations

Figure 8 The Usage of Refuelling Stations

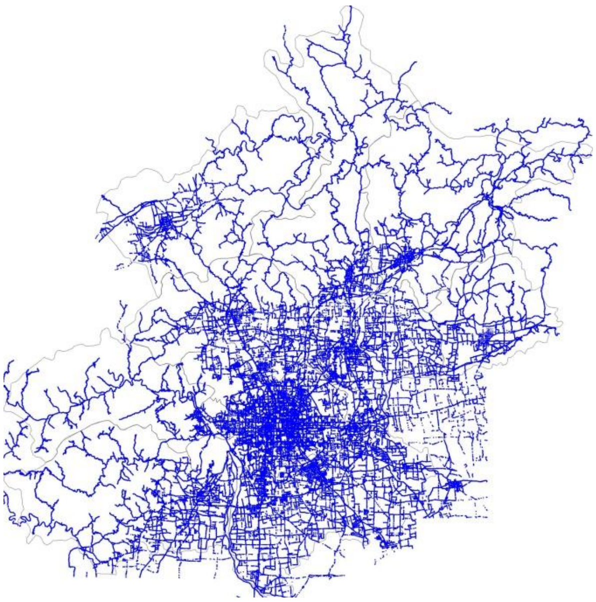
3.5 Impacts of Population Scaling and Network Simplification

As aforementioned, population scaling is generally expected to speed up a simulation and to reduce computing time. However, it has been found that the population scaling factor is an influential parameter, which can heavily influence the outputs of interest, in the SA of MATSim by Zhuge et al. (2019c). In addition to the population scaling, network simplification, which simplifies road network by merging short links, is another effective method to speed up these simulations, because running MATSim-Beijing with a simplified network can reduce the computing time for both shortest path searching and network loading (or simulating the movement of agents). These two common approaches can make simulations more efficiently and would be particularly useful for large-scale scenarios, such as the Beijing scenario here, though the extent to which they could influence the simulation outputs has not been well understood. Therefore, this paper attempts to quantify the influence by testing different population scaling factors and simplified networks within the following experiments:

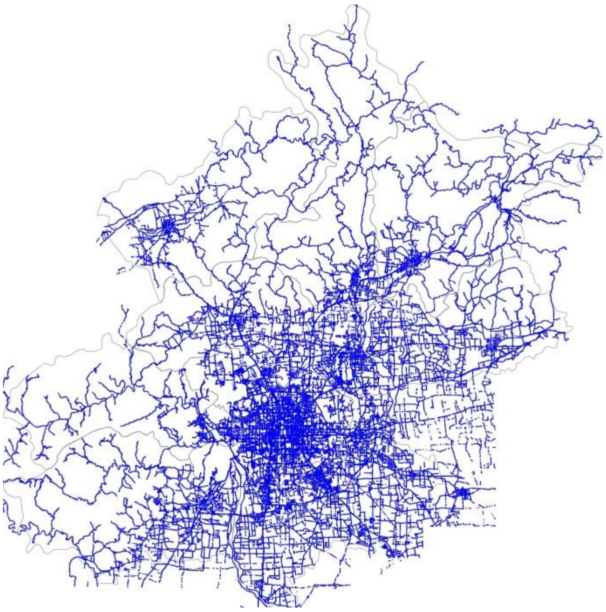
- **Experiment-Pop**: is to test the influence of population scaling within several sub-experiments using scaling factors ranging from 0.01 to 0.1 with an interval of 0.01, as well as factors of 0.15 and 0.2.
- **Experiment-Net**: is to test the influence of network simplification using six different network types. They differ from each other in the constraint on the maximum merging length used for simplification. Specifically, the simplifying method merges two or more adjacent links only if the total length of these links does not exceed the maximum merging length. Figure 9 shows the simplified networks with maximum merging lengths of 100, 300, 500, 1000, 1500 and 2000 meters (Figure 9(b)-(g), as well as the original one (Figure 9-(a)).

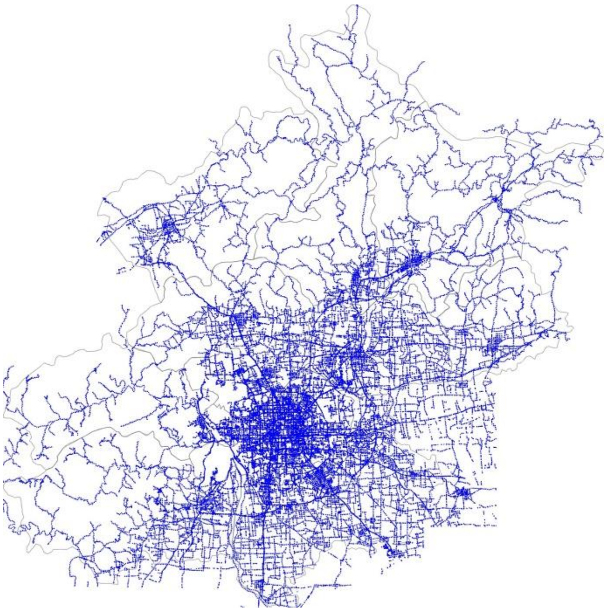


(a) Original Network with 443279 nodes and 759396 links



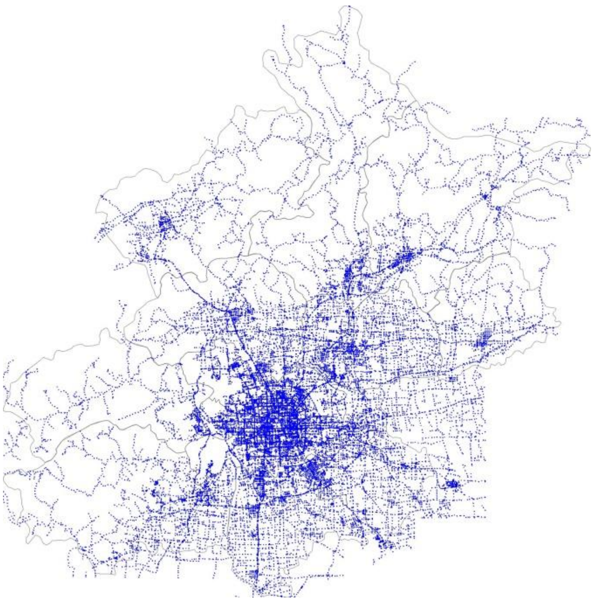
(b) Simplified Network with 201480 Nodes and 352041 (Max length= 100 meters)

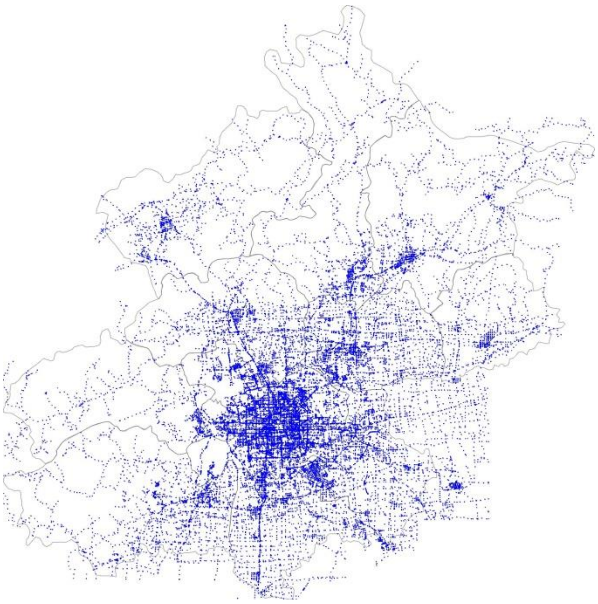




(c) Simplified Network with 90112 nodes and 167207 links (Max length= 300 meters)

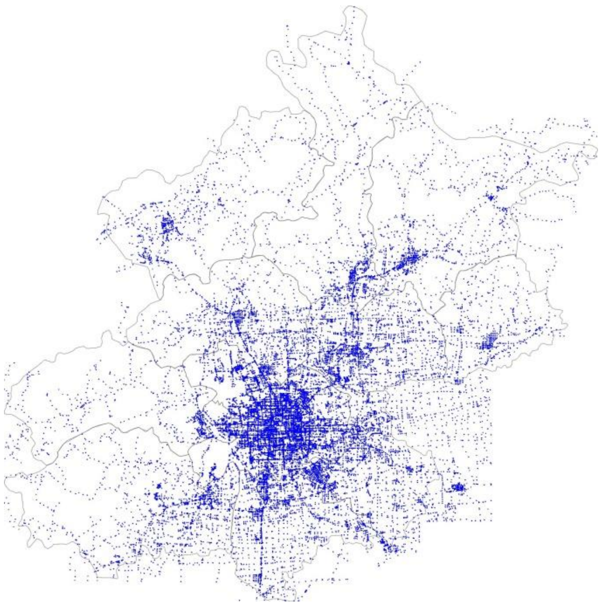
(d) Simplified Network with 61419 nodes and 120407 links (Max length= 500 meters)





(e) Simplified Network with 42560 nodes and 88616 links (Max length= 1000 meters)

(f) Simplified Network with 37748 nodes and 80164 links (Max length= 1500 meters)



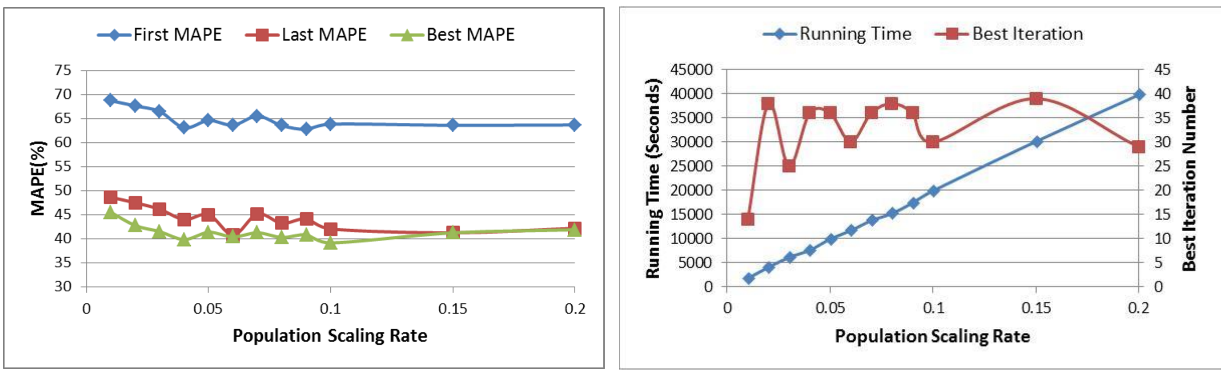
(g) Simplified Network with 35899 nodes and 76852 links (Max length= 2000 meters)

Figure 9 Maps of Original and Simplified Road Networks in 2010 (Note that only road nodes were drawn in the maps)

Figure 10 and Figure 11 show the impacts of population scaling and network simplification on the simulation results, respectively. The impact assessment uses two indicators, namely MAPE and computing time, to characterise the effectiveness and efficiency of each simulation, respectively.

(1) Impact of Population Scaling on Simulation Results

According to Figure 10-(a), it can be found that the population scaling could influence heavily the model accuracy if a too small scaling rate (below 1%) is used, but there are no significant differences in model accuracy for scaling rates from 2% to 20%. However, as shown in Figure 10-(b), the computing time is directly proportional to the scaling rate. In other words, using a small scaling rate can significantly decrease the computing time. For the number of iterations required to reach the smallest (or best) MAPE, it ranges widely from around 15 to 40, suggesting that there appears to be no significant direct relationship between the scaling rate and best iteration number, as the best iteration number fluctuates across the scaling rates from 1% to 20%.

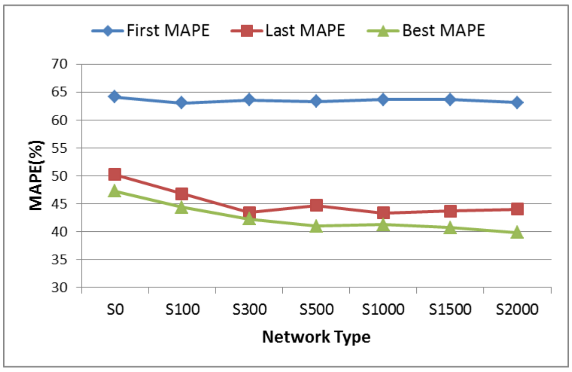


(a) Model Accuracy (b) Computing Time

Figure 10 Impact of Population Scaling on Simulation Results

(2) Impact of Network Simplification on Simulation Results

As shown by Figure 11-(a), using the original network (S0) or less simplified networks (such as S100 and S300) could increase MAPE. However, for the networks with longer maximum merging length (S500-S2000), they tend to have smaller MAPEs. There may be several possible reasons. One likely reason may be that the calibrated MATSim-Beijing may only work properly with simplified networks that are not significantly different from the one (S2000) used for the MATSim-Beijing calibration, as using those less simplified networks could heavily change the resulting traffic flow and thus cannot match the simulated and observed data. In addition, those networks with longer maximum merging lengths (S500-S2000) get significantly shorter computing time below 10,000 seconds, as shown by Figure 11-(b). There appears to be no direct relationship between the maximum merging length used for network simplification and the best iteration number.



(a) Model Accuracy (b) Computing Time

Figure 11 Impact of Network Simplification on Simulation Results

4 Conclusions

An agent- and activity-based car travel demand model was developed for Beijing, China with the explicit simulation of enroute travel, parking and enroute refuelling behaviours, primarily using the 2010 Household Travel Survey Data and the data on parking and refuelling behaviours collected in a questionnaire survey in Beijing from September, 2015 to March, 2016. In order to consider the heterogeneous parking and refuelling behaviours, Multinomial Logit (MNL) models were developed to predict the individual choices about the maximum acceptable time of walking from parking lot to trip destination, of diverting to a refuelling station and of queueing at a station. The simulation can obtain rich fine-grained outputs, including the usage of refuelling stations and parking lots and the spatial and temporal distributions of traffic flow and vehicular emissions. Therefore, the model would be useful for policy analysis and infrastructure planning. For example, the model can be used to evaluate how a parking policy (e.g., parking fee) may influence parking behaviour and further traffic flow, energy consumption and vehicular emissions at multiple resolutions (e.g., link-, zone- and district- levels). Such information can aid the relevant decision-makings.

A Sensitivity Analysis (SA)-based calibration method was applied to calibrate the model, exhibiting a satisfactory performance, compared to a typical calibration tool, Cadyts. Furthermore, the calibrated model was also applied to explore the influences of population scaling and network simplification on model outputs, which are two common approaches to speeding up large-scale simulations. The results indicated that both approaches could to some extent influence the model outputs, though the computing time could be significantly reduced. Therefore, these two approaches are suggested to be used with caution. As High Performance Computing (HPC) machines become more accessible, parallel computing, which divides a heavy computing task into several sub-tasks and allocates them to different nodes, may become a promising approach to speeding up such large-scale simulations, as an alternative to the two common approaches. Therefore, the model is planned to be run on HPC machines with a whole population and complete road network, which could help fully understand the impacts of population scaling and network simplification on model results.

The future work on the Beijing activity-based model is discussed as follows: First, the Beijing model will be further extended by considering more transport modes, such as public transit and freight transport, so as to capture the interactions between different transport modes on the transport network. This is expected to make the traffic simulation more realistic and accurate; Second, we used the fixed random seed for each simulation, so as to mitigate the impact of randomness across scenarios as far as possible. The potential impact of randomness on model results can be assessed by running the model with a specific number of random seeds and then comparing the model results. The assessment, however, could be computationally expensive, especially to large-scale scenarios. HPC machines could be particularly useful here for reducing the total computing time. Third, we calibrated the model using link-based traffic flow data. The calibration can be improved by using big data, such as GPS trajectory data on private car. For example, the model can be calibrated at the disaggregate level by matching the simulated and observed parking and refuelling behaviours. However, such big data is not always accessible due to privacy issues.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (Grant No. 51678044), the Fundamental Research Funds for the Central Universities (NO. 2017JBZ106), China, the Hong Kong Polytechnic University [1-BE2J], and the ERC Starting Grant #678799 for the SILCI project (Social Influence and disruptive Low Carbon Innovation). We would also thank Dr. Mike Bithell for discussing with us about the model development and test.

References

Agarwal, A., Ziemke, D. and Nagel, K., 2017. Calibration of a heterogeneous traffic scenario in an agent-based simulation framework, 4th Conference of the Transportation Research Group of India (CTRG), Mumbai, India.

Ahmed, B., 2012. The traditional four steps transportation modeling using a simplified transport network: a case of Dhaka city, Bangladesh. International Journal of Advanced Scientific Engineering and Technological Research, 1(1): 19-40.

Balmer, M., 2007. Travel demand modeling for multi-agent transport simulations: Algorithms and systems, Doctoral Dissertation, ETH Zurich, Swizerland.

Bazzan, A.L. and Klügl, F., 2014. A review on agent-based technology for traffic and transportation. The Knowledge Engineering Review, 29(3): 375-403.

Benenson, I., Martens, K. and Birfir, S., 2008. PARKAGENT: An agent-based model of parking in the city. Computers, Environment and Urban Systems, 32(6): 431-439.

Bhat, C.R. and Koppelman, F.S., 1999. Activity-based modeling of travel demand, Handbook of Transportation Science Springer, pp. 35-61.

Bischoff, J. and Nagel, K., 2017. Integrating explicit parking search into a transport simulation. Procedia Computer Science, 109: 881-886.

Bradley, M. and Vovsha, P., 2005. A model for joint choice of daily activity pattern types of household members. Transportation, 32(5): 545-571.

Charypar, D., Axhausen, K. and Nagel, K., 2006. Implementing activity-based models: Accelerating the replanning process of agents using an evolution strategy, 11th International Conference on Travel Behaviour Research, Kyoto.

Charypar, D., Axhausen, K.W., Nagel, K., Axhausen, K.W., Axhausen, K.W., Nagel, K. and Nagel, K., 2007. An event-driven parallel queue-based microsimulation for large scale traffic scenarios. ETH, Eidgenössische Technische Hochschule Zürich, IVT.

Dieussaert, K., Aerts, K., Steenberghen, T., Maerivoet, S. and Spitaels, K., 2009. SUSTAPARK: an agent-based model for simulating parking search, AGILE International Conference on Geographic Information Science, Hannover.

Flötteröd, G., 2009. Cadyts-A free calibration tool for dynamic traffic simulations, 9th Swiss Transport Research Conference, Ascona, Switzerland.

Flötteröd, G., Bierlaire, M. and Nagel, K., 2011. Bayesian demand calibration for dynamic traffic simulations. Transportation Science, 45(4): 541-561.

Flötteröd, G., Chen, Y. and Nagel, K., 2012. Behavioral calibration and analysis of a large-scale travel microsimulation. Networks and Spatial Economics, 12(4): 481-502.

Galus, M.D., Waraich, R.A., Noembrini, F., Steurs, K., Georges, G., Boulouchos, K., Axhausen, K.W. and Andersson, G., 2012. Integrating power systems, transport systems and vehicle technology for electric mobility impact assessment and efficient control. IEEE Transactions on Smart Grid, 3(2): 934-949.

Gonzalez, J., Alvaro, R., Gamallo, C., Fuentes, M., Fraile-Ardanuy, J., Knapen, L. and Janssens, D., 2014. Determining electric vehicle charging point locations considering drivers' daily activities. Procedia Computer Science, 32: 647-654.

Horni, A., Montini, L., Waraich, R.A. and Axhausen, K.W., 2013. An agent-based cellular automaton cruising-for-parking simulation. Transportation Letters, 5(4): 167-175.

Horni, A., Nagel, K. and Axhausen, K.W., 2016. The Multi-Agent Transport Simulation MATSim. Ubiquity, London.

IEA, 2019. Global EV Outlook 2019. Retrieved on 5th August 2019 from www.iea.org/publications/reports/globalevoutlook2019/.

Knapen, L., Kochan, B., Bellemans, T., Janssens, D. and Wets, G., 2011. Activity based models for countrywide electric vehicle power demand calculation, Smart Grid Modeling and Simulation (SGMS), 2011 IEEE First International Workshop on. IEEE, pp. 13-18.

Knapen, L., Kochan, B., Bellemans, T., Janssens, D. and Wets, G., 2012. Activity-based modeling to predict spatial and temporal power demand of electric vehicles in Flanders, Belgium. Transportation Research Record: Journal of the Transportation Research Board(2287): 146-154.

Lefebvre, N. and Balmer, M., 2007. Fast shortest path computation in time-dependent traffic networks. ETH, Eidgenössische Technische Hochschule Zürich, IVT, Institut für Verkehrsplanung und Transportsysteme.

Macal, C.M. and North, M.J., 2010. Tutorial on agent-based modelling and simulation. Journal of Simulation, 4(3): 151-162.

McNally, M.G., 2008. The four step model, Center for Activity Systems Analysis. Institue of transportation study University of California Irvine.

Peeta, S. and Ziliaskopoulos, A.K., 2001. Foundations of dynamic traffic assignment: The past, the present and the future. Networks and Spatial Economics, 1(3-4): 233-265.

Pinjari, A.R. and Bhat, C.R., 2011. Activity-based travel demand analysis. A Handbook of Transport Economics, 10: 213-248.

Rasouli, S. and Timmermans, H., 2014. Activity-based models of travel demand: promises, progress and prospects. International Journal of Urban Sciences, 18(1): 31-60.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. and Tarantola, S., 2008. Global sensitivity analysis: The primer. John Wiley & Sons.

Srinivasan, S. and Bhat, C.R., 2005. Modeling household interactions in daily in-home and out-of-home maintenance activity participation. Transportation, 32(5): 523-544.

van der Waerden, P., 2012. PAMELA: Parking analysis model for predicting effects in local areas. Doctoral Disseration, Eindhoven University of Technology.

Waraich, R.A., 2013. Agent-based simulation of electric vehicles: Design and implementation of a framework. Doctoral Dissertation, ETH Zurich, Zurich.

Waraich, R.A., Charypar, D., Balmer, M. and Axhausen, K.W., 2015. Performance improvements for large-scale traffic simulation in MATSim, Computational Approaches for Urban Environments. Springer, pp. 211-233.

Waraich, R.A., Dobler, C., Weis, C. and Axhausen, K.W., 2012. Optimizing parking prices using an agent based approach. Arbeitsberichte Verkehrs-und Raumplanung, 794.

Waraich, R.A., Georges, G., Galus, M.D. and Axhausen, K.W., 2014. Adding electric vehicle modeling capability to an agent-based transport simulation. Data Science and Simulation in Transportation Research: 282-318.

Yao, E. and Song, Y., 2013. Study on eco-route planning algorithm and environmental impact assessment. Journal of Intelligent Transportation Systems, 17(1): 42-53.

Yao, E., Yang, Z., Song, Y. and Zuo, T., 2013. Comparison of electric vehicle's energy consumption factors for different road types. Discrete Dynamics in Nature and Society, 2013.

Zhuge, C., 2019. Research resource review. Progress in Physical Geography: Earth and Environment, 43(4): 601-603.

Zhuge, C. and Shao, C., 2018. Agent-based modelling of locating public transport facilities for conventional and electric vehicles. Networks and Spatial Economics, 18(4): 875-908.

Zhuge, C. and Shao, C., 2019. Investigating the factors influencing the uptake of electric vehicles in Beijing, China: Statistical and spatial perspectives. Journal of Cleaner Production, 213: 199-216.

Zhuge, C., Shao, C. and Li, X., 2019a. A comparative study of en route refuelling behaviours of conventional and electric vehicles in Beijing, China. Sustainability, 11(14): 3869.

Zhuge, C., Shao, C. and Li, X., 2019b. Empirical analysis of parking behaviour of conventional and electric vehicles for parking modelling: A case study of Beijing, China. Energies, 12(16): 3073.

Zhuge, C., Shao, C., Wang, S. and Hu, Y., 2018. An agent- and GIS-based virtual city creator: A case study of Beijing, China. Journal of Transport and Land Use, 11(1): 1231-1256.

Zhuge, C., Shao, C., Wang, S. and Hu, Y., 2019c. Sensitivity analysis of integrated activity-based model: using MATSim as an example. Transportation Letters, 11(2): 1-11.

Zhuge, C., Wei, B., Dong, C., Shao, C. and Shan, Y., 2019d. Exploring the future electric vehicle market and its impacts with an agent-based spatial integrated framework: A case study of Beijing, China. Journal of Cleaner Production, 221: 710-737.