International Journal of Applied Earth Observation and Geoinformation 122 (2023) 103368

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Agent-based modelling with geographically weighted calibration for intra-urban activities simulation using taxi GPS trajectories

Shuhui Gong [a], Xiangrui Dong [b], Kaiqi Wang [a], Bingli Lei [a], Zizhao Jia [a], Jiaxin Qin [a], Chris Roadknight [c], Yu Liu [d], Rui Cao [e,f,*]

[a] *School of Information Engineering, China University of Geosciences, Beijing, China*
[b] *School of Earth Sciences and Resources, China University of Geoscience, Beijing, China*
[c] *Department of Computer Science, University of Hertfordshire, Hatfield, UK*
[d] *Institute of Remote Sensing and Geographical Information Systems, Peking University, Beijing, China*
[e] *Department of Land Surveying and Geo-Informatics & Otto Poon Charitable Foundation Smart Cities Research Institute, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region, China*
[f] *The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen, China*

A R T I C L E   I N F O

Keywords:
Agent-based modelling (ABM)
Huff model
Geographically weighted regression (GWR)
Activity-based analysis
Taxi GPS trajectories

A B S T R A C T

Human motivations are an important factor in influencing human movement. However, most existing studies on passenger travel demand prediction focus on external characteristics of movement, but neglect the influence of activities and the motivations behind them, on the citizen's trip decisions. In this study, we proposed an agent-based model, to predict passengers' travel behaviour over a period of time, particularly when the urban structure changes. The model includes passenger characteristics, transitions in travel demands between activities over time, and their movement in space and time. In addition, we innovatively calibrated the agent-based model locally using Geographically Weighted Regression (GWR) to account for geographical variations in the parameters of destination attractiveness and travel cost in the agent-based model. We conducted a case study in Ningbo, China, using trip diaries, census data, and over 1.5 million taxi trip records. Our agent-based model showed superior performance in predicting citizens' movements and activities after a new shopping area in Ningbo was built, compared with a model without local calibration. The results also revealed the geographic sensitivity to destinations and the effects of a passenger's motivations that underpin human movement.

1. Introduction

Human mobility prediction is an increasingly important research topic (Wang et al., 2019; Shi et al., 2021). The accurate prediction of human movements could be beneficial to many applications, such as urban planning (Wang et al., 2019; Yin and Chi, 2022), traffic forecasting (Zhao et al., 2019; Wang et al., 2022), and location-based recommender system (Lim et al., 2019). However, most previous research only focuses on historical trajectories and external context information (such as traffic condition and weather) but neglects the driving force (activity) behind the trips, which could make it even more difficult to predict travel demands after changes in urban design (Gong et al., 2023; Zhao et al., 2022; Zhou et al., 2017; Cheng et al., 2011; Song et al., 2010). Some recent studies explored the possibility of integrating the purpose of travel into the prediction process (Xu et al., 2023), but how to appropriately estimate citizens' daily movements, and explore the interactions between human behaviour and city planning remains an important challenge.

Activity is defined as the internal reason for passenger travel (Koushik et al., 2020; Miller, 2021). In particular, citizens' daily visits and schedule could be estimated from activity-based analysis (Yang et al., 2014; Yin et al., 2017). Instead of a trip-based approach, activity-based analysis could describe the relationships between predecessor trips and successor trips, and ultimately, predict passengers' travel demand over a period of time (Gong et al., 2020b). With considerable assumptions (such as modelling of the relationship between activities and the human movements), activity-based analysis is able to be applied on travel behaviour analysis (Liu et al., 2012).

Agent-based modelling (ABM) can achieve good qualitative performance when used for a principled approach to activity-based analysis (Liu et al., 2012; Wu et al., 2014; Gong et al., 2020a), as well as the modelling of urban transportation (Chowell et al., 2003), daily commuting routines (Balaraman et al., 2015), travel demand (Auld et al., 2016), economy studies (Platt, 2020), spatial epidemics (Simoes,

* Corresponding author.
E-mail address: rucao@polyu.edu.hk (R. Cao).

https://doi.org/10.1016/j.jag.2023.103368
Received 23 February 2023; Received in revised form 5 May 2023; Accepted 23 May 2023
Available online 21 June 2023

2012), etc. An agent-based model can extract micro-scale patterns by simulating each agent's behaviour. In particular, it can discover the interactions among the agents, as well as the interactions between the agents and the environment (Saeedi, 2018; Crooks and Heppenstall, 2012). Although agent-based models perform well on aggregating the activity of simulated individuals, it faces strong criticisms due to the inadequacy of current validation and calibration practices (Grazzini and Richiardi, 2015). More specifically, there is insufficient research emerging on agent-based model validation with empirically-observed data, and what there is usually ignores the design of calibration methods (Grazzini et al., 2017). With effective calibration, agent-based models can achieve excellent levels of performances in various case studies (Rodrigue et al., 2012).

Traditionally, researchers utilised questionnaires or interviews for model calibration, which are labour-intensive, and time-consuming. Meanwhile, these data are largely dependent on the participants' perceptions (Gong et al., 2021). In recent years, GPS trajectory data has been widely used for model calibration. Unlike traditional data sources, large GPS data sets can be collected automatically. However, since most GPS data do not include activity information, few researchers have utilised GPS data on activity-based model calibration. As a result, how to fit and validate the agent-based models precisely is a challenging issue.

Focusing on this research gap, we proposed an agent-based model to estimate passengers' activity transitions (such as work, residence, and dining) between predecessor trips and successor trips. We simulate individual behavioural activities across a city. Furthermore, a geographically weighted calibration method is proposed to fit the parameters in an agent-based model, using the Huff model, Geographically Weighted Regression (GWR), and Monte Carlo Simulation. In particular, for each area that the agent is located in, the agent-based model would have a set of specific parameters for the agent to make trip decisions.

Our major contributions can be summarised as follows:

- We propose a Destination-aware Activity Simulation (DAS) model to simulate intra-urban activities based on agent-based modelling using taxi data, trip diaries, and census data. It is worth noting that this paper only explores the activity behaviour of taxis in cities.
- Inspired by the Huff model, we introduce the parameters of destination attractiveness and travel cost for the ABM, and propose a novel geographically weighted calibration method for ABM based on GWR, which can explicitly account for geographical variations.
- To address model validation issues, we propose two ways to validate the effectiveness of the ABM model based on isolated POI (IPOI)-based validation and a scenario test which examines the impact of introducing a new infrastructure to the simulation.
- Extensive experiments have been performed on a case study in Ningbo, China, demonstrating the effectiveness and superiority of the proposed method.

The rest of the paper is organised as follows. In Section 2, we summarise the related works on activity-based analysis, and methodologies for agent-based model calibration and validation. In Section 3, we introduce the proposed methodology in this study. In Section 4, we conducted a case study in Ningbo, China. Finally, Section 5 concludes the paper.

2. Literature review

2.1. Activity-based analysis

Although urban travel demand prediction has been well-researched, the internal motivations of most journeys are hidden (Rasouli and Timmermans, 2014). To explore passengers' travelling motivations, activity-based analysis is a good way to explain the capabilities of the forecasting (Kaelbling et al., 1996; Yang et al., 2014). Programming the internal goals of the agents, rather than the desired outcome has been shown to lead to realistic behaviour, both in a quantitative and qualitative manner (Roadknight et al., 2012).

Traditionally, the two types of methodologies that are commonly undertaken for activity-based analysis are econometric modelling and computational modelling (Koushik et al., 2020). Econometric models assume that individuals are trying to maximise their utility when making active travel decisions. However, the approach has been criticised for ignoring the imperfect selection behaviour of individuals (Koushik et al., 2020; Bowman and Ben-Akiva, 2001). Computational models define rules to explore the impact of passengers' attributes on their travel behaviours. Therefore, complex nonlinear relationships can be effectively modelled (Koushik et al., 2020). The Huff model is one of the most widely used computational models that estimates the customers' visiting probabilities to one shopping area. It considers two factors: the destination attractiveness and the travel cost (Huff, 1964). With good calibration, the Huff model can be used to predict passengers' travel behaviour. However, the Huff model is limited in several aspects: (i) the model can only analyse shopping behaviour, trips for other activities are not considered; (ii) previous work calibrated Huff model with one pair of parameters, which assumes that all passengers have similar attitude for the destination attractiveness and the travel cost. It is clearly not same as the real situation; (iii) traditional activity-based travel demand analysis cannot model the complex interactions between individuals and the environment (Yamamoto et al., 2002).

2.2. Agent-based model calibration

Agent-based modelling adopts bottom-up modelling and can be modelled at the micro-scale (Crooks and Heppenstall, 2012). Recent research mainly focuses on modelling urban flows and human behaviour to analyse behaviour–environment interactions (Walsh et al., 2013). Due to its high dimensional and hierarchical structure, calibration is always one of the most important problems in agent-based modelling (Platt, 2020). One common calibration methodology for agent-based modelling is using observation data. In particular, Sturley et al. (2018) utilised the observed number of visits to stores to calibrate the agents' shopping choice behaviour in an agent-based model; Gong et al. (2020a) filtered over 6000 trip diaries from the questionnaire data in New York to calibrate passengers' activities schedule and utilised a distance decay function to estimate the destination choices. However, calibration process in previous work often need large questionnaire data to support the real situations, which may lead to high labour cost.

2.3. GPS data-based model validation

With appropriate validation, the models can be evaluated on the prediction accuracy or their effectiveness (Gong et al., 2020b). It has also been demonstrated that GPS data can be utilised as ground truth for travel behaviour forecasting (Wang et al., 2018). For example, Gong et al. (2021) utilised taxi origin–destination (OD) data to filter all trips for shopping centres, in order to evaluate the accuracy of the Huff model on predicting the number of shopping areas visited. Liu et al. (2014) proposed a validation method using mobile phone GPS data to validate activity-based transportation models. However, since most GPS data did not include activity information, it could be challenging to validate the activity-based model using GPS data.

3. Methodology

3.1. Overview of the Destination-aware Activity Simulation (DAS) model

The overview flow chart of the proposed Destination-aware Activity Simulation (DAS) model is shown in Fig. 1. The model takes trip diaries,

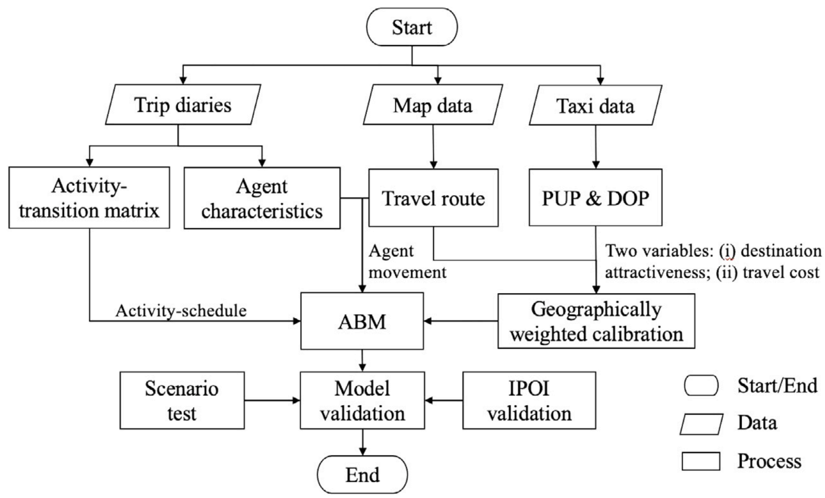


Fig. 1. Overview of the proposed Destination-aware Activity Simulation (DAS) model.

Table 1
Questions in trip diaries.

No.	Question	Example answer
1	Which district did you live in?	Beilun
2	Where is the location that you participated in this survey?	Yinzhou park
3	Where did you work at?	Shounan street
4	Was that day a weekday or a weekend?	Weekday
5	How many trips did you have planned for that day?	3
6	What was the aim of your trip?	Work
7	How long did you stay at your destination?	Over 4 h
8	If you had another trip, where did you go after your last trip?	Residence

map data (including POIs and road networks), and taxi GPS data as input.

We utilised trip diaries to estimate citizens' daily movements. More specifically, the trip diaries recorded passengers' one day trips with activities. To collect the trip diaries, all the participants are asked to write down all their trips in one day, which could be either a weekday, or a weekend day. The questions utilised in this study are shown in Table 1. For each trip of each participant, the interviews record the number of trips in that day, trip start time, end time, related activity, origin and destination location, and activity duration. The trip data is categorised into eight activity types: residence, work, shopping, dining, schooling, recreation, transport, and others. In our model, A_p represents the prior activity, $A_{s'}$ signifies a specific subsequent activity, A_s encompasses all feasible subsequent activities, and the agent begins the next activity with a defined probability after finishing a single activity. From trip records, we calculated the activity transition probabilities (see Eq. (1))

$$P(A_p, A_{s'}) = \frac{Num(A_p, A_{s'})}{\sum_s Num(A_p, A_s)} \tag{1}$$

where $P(A_p, A_{s'})$ means the transition probability from previous activity A_p to successor activity $A_{s'}$. $Num(A_p, A_{s'})$ represents the observation number of the trips moving from activity p into successor activity s'. Based on Eq. (1), we designed the activity transition matrix, to display the correlation between the previous activity and the successor activity. We input the activity transition matrix to the agent-based model to simulate the activity-schedule of the agents, which are introduced in Section 3.2 with details.

Combining trip diaries with census data, we could obtain passengers' characteristics (see Fig. 2), which includes each agent's residential location, whether they are at work or not, workplace, and working time (such as 8 am to 5 pm) in several steps: (i) after collecting trip diaries, we statistically gathered samples of citizens' residences and working areas; (ii) from census data, we accurately gathered citizens' residential distributions; (iii) finally, the residential and working places of urban residents in the model are estimated through the proportion of the distribution of residents' residences in the census data and the relationship between the residents' residences and working places in the trip diaries.

Besides citizens' characteristics, we also import POI data and road networks into the agent-based model, in order to simulate agents' movement. The POI data include the POIs' locations, related activities, and opening times. The road network data is used to calculate passengers travel route. The detail is introduced in Section 3.2.

Moreover, we utilised taxi data to calibrate the agent-based model. Specifically, we firstly extracted individual OD taxi trips from taxi GPS data, including pick-up point (PUP) and drop-off point (DOP), with both location and time. The variables destination attractiveness and travel cost (as indicated in Eq. (3)) are subsequently calibrated in a geographically weighted manner for the agent-based model. The calibration process is introduced in Section 3.3. With proper calibration, the agent-based model could simulate passengers' daily trips with activities, from which we could understand the human interaction as well as the activity patterns.

Furthermore, to examine the effectiveness of the ABM model, we propose two approaches, namely IPOI-based validation and scenario testing, as detailed in Section 3.4.

3.2. The agent-based model design

The agent-based model is the core of the proposed DAS model, which aims to simulate daily movements and activities of passengers as agents. The flow chart of the agent-based model is presented in Fig. 2. We imported three types of information into the agent-based model, including time system (from 5 am to 11 pm), map information (including POIs and road networks), and agent characteristics.

Observed from trip diaries, the activities can be further classified into two categories: (i) recurrent activity and (ii) temporary activity.

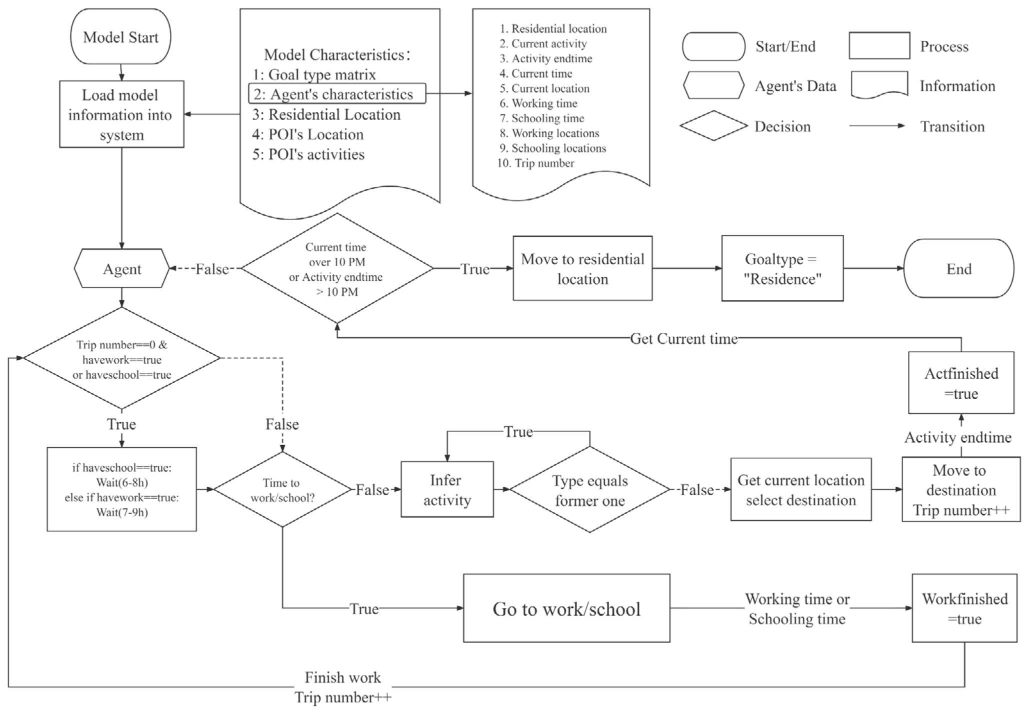


Fig. 2. Structure of the proposed agent-based model.

The recurrent activity includes residential activity, work-related activity, and school-related activity. Since in real situations, each citizen has its own permanent residential location, work location, permanent time schedule of the work, permanent school location, and permanent time schedule for school in a long time period. Therefore, we consider home, work and school activities as recurrent activity, and has been arranged as characteristics of each agent in the agent-based model. These arrangements had been statistically estimated from trip diaries and census data (introduced in Section 4). Temporary activity means agents have flexible arrangement on other activities based on their spare capacity. Passengers' activity transitions are estimated from trip diaries, while their spatial choice for destination is predicted using the Huff model, GWR, and Monte Carlo simulations.

We constrained the agent's behaviour with a series of assumptions, which are based on the agent characteristics. For each agent, we define ten attributes: (i) residential location (we initially locate all agents in their residential area and each agent has its own permanent residence), (ii) current activity (citizens' activity schedule is calculated from the activity transition matrix), (iii) activity end time (it is estimated from the trip diaries), (iv) current time, (v) current location, (vi) working time, (vii) working locations, (viii) school time, (ix) school location, and (x) trip number.

In ABM, every time step represents 15 min, from 5 am to 11 pm, we excluded the data beyond this time range since the number of night trips is relatively low and their frequency is less regular. In total, the agent-based model simulates 72 time steps, equivalent to 18 h. During every time step, the ABM estimates each agent's next activity. If the agent's previous activity has not yet completed, the agent's status will not change. Meanwhile, the ABM also checks if the current time meets the agents' schedule. If the time meets the working or schooling time, the agent will end the current activity, and head for its work place or school. Otherwise, the agent-based model will generate a successor activity, which is based on the activity transition probabilities in DAS.

3.2.1. Next activity type estimation

In the ABM, we estimate each passenger's total trip number by statistically calculating the probabilities of each trip number from trip diaries. Once trip number is determined, we utilised Monte Carlo simulation to estimate the agent's next activity, which is shown in Algorithm 1. By calculating the proportion of the current activity p to the next specific activity s in N activity categories. The activity transition probability $P(A_p, A_s)$ with the sum of 1. On this basis, we used a random decimal r in [0,1] to estimate the next activity. If there is no further trips, the ABM keep the agent at home until the end of the day.

Algorithm 1 Monte Carlo simulation

```
Input: a set of trip number
Output: one trip number
for each number do
    the probability to trip numbers N are p_1, p_2, ..., p_N    ▷ Σ_n^N p_n = 1
    set p_0 = 0
    generate random value, r ∈ [0, 1]
    for n = 0 to N − 1 do                                       ▷ decide target number
        if Σ_{j=0}^{n} p_j ≤ r < Σ_{j=0}^{n+1} p_j then
            result = number[n + 1]
        end if
    end for
end for
return result
```

The next activity is estimated using activity transition matrix (Section 3.1):

$$P(A_p^t, A_s^{t+1}) = \begin{cases} P(A_p, A_s), & (t+1) \in A_s^{opentime} \\ 0, & (t+1) \notin A_s^{opentime} \end{cases} \quad (2)$$

where t means the current time, t+1 means the next time step, $P(A_p, A_s)$ means the probability that the agent's next trip is activity s when its previous activity is p, which could be calculated from Eq. (1). The probability is directly collected from trip diaries. $A_s^{opentime}$ means the open time range of activity s.

3.2.2. Next destination estimation

For temporary activities to complete (except home and work), the agent-based model will estimate the next geographic destination for the activity. The estimation is based on the Huff model. The Huff model states that the attractiveness of various conditions of trading areas to citizens and the cost of the distance between passengers and trading areas determines the law of the size of business circles. The formula is described as follows:

$$P_{sij} = \frac{S_j^{\alpha} C_{ij}^{\beta}}{\sum_{j=1}^{m_s} S_j^{\alpha} C_{ij}^{\beta}} \tag{3}$$

where P_{sij} represents the probability that a citizen from origin location i to visit the area j for activity s; C_{ij} means the travel cost for agents from location i to area j; S_j is the attractiveness of j; α and β are, respectively, the parameters of S and C at each origin location i; and m is the total number of areas for activity s. In this study, we utilised the area's visiting volumes to estimate its attractiveness S, which is calculated from taxi data. Moreover, we consider the trip length from i to j as the travel cost C_{ij}.

3.3. Agent-based model calibration

In this study, we fit two parameters in an agent-based model, destination attractiveness (α) and travel cost (β), as introduced in Eq. (3). To estimate the Huff model parameters (in particular α and β), four attraction and cost function combinations were introduced (O'Kelly, 1999), as in Eq. (4):

$$\begin{aligned} K1 &: P_{ij} = exp(\alpha S_j - \beta C_{ij}) \\ K2 &: P_{ij} = exp(\alpha S_j - \beta LnC_{ij}) \\ K3 &: P_{ij} = exp(\alpha LnS_j - \beta C_{ij}) \\ K4 &: P_{ij} = exp(\alpha LnS_j - \beta LnC_{ij}) \end{aligned} \tag{4}$$

Moreover, Nakanishi and Cooper (1982) calibrated the model using a log-transformed-centred form of Ordinary Least Squares (OLS):

$$OLS : Ln(P_{ij}/\bar{P}_i) = \alpha_i Ln(S_j/\bar{S}) + \beta_i Ln(C_{ij}/\bar{C}_i) \tag{5}$$

where $\bar{P}_i$, $\bar{S}$ and $\bar{C}_i$ are the geometric means of P_{ij}, S_j and C_{ij} over j.

To enable comparisons with the literature, we use all the five estimation methods (Eqs. (4) and (5)) to fit one pair of DAS parameters (α and β) and make comparisons between different time periods. Considering the heterogeneity of different locations, we innovatively propose to estimate α and β using GWR model.

GWR models are extensions of general linear regression models and are widely used to analyse changes in urban travel demand (Tu et al., 2018). Different from linear regression, the parameters estimated from GWR vary spatially, instead of remaining constant across space. Therefore, GWR models are suitable for exploring spatial heterogeneity. The GWR model can be formulated as follows:

$$y_i = \beta_{i0}(u_i, v_i) + \sum_{k=1}^{m} \beta_{ik}(u_i, v_i) x_{ik} + \varepsilon_i \tag{6}$$

where y_i is the dependent variable at location i; x_{ik} is the kth independent variable at location i; m is the number of independent variables; β_{i0} is the intercept parameter at location i; β_{ik} is the local regression coefficient for the kth independent variable at location i; and ε_i is the random error at location i. Since there are two parameters in DAS to estimate, $k = 2$, and β_{ik} in Eq. (6) corresponds to α and β for each i.

3.4. Model validation

In this study, we propose two approaches to validate the performance of the proposed DAS model based on taxi trip data, i.e., IPOI-based validation and scenario testing.

3.4.1. IPOI-based validation

The key idea of the first approach is to compare the difference between the temporal arrival distributions for each activity within the agent-based model with temporal drop-off distributions for each activity captured from taxi trips. However, the challenge lies in the fact that raw taxi trips data lack labels of activity types.

To address the issue, we propose to use isolated POIs (IPOIs) to connect taxi trips with related trip purposes (activity types). We defined an IPOI as a special POI where there are no other POIs located within 500 m, measured by network distance. Previous work has demonstrated that passengers usually walk up to 500 meters for the destination (Yue et al., 2012). This suggests that, when a taxi trip DOP is located within the buffer zone of a POI, such as a large shopping area, a park, or workplace, with no other POIs located nearby, we can confidently assume that the trip destination is that POI, and the trip purpose (activity type) can be connected to the POI type. By quantitatively estimating the difference between the results of agent-based model and the taxi data, we can calculate the accuracy of the proposed DAS model.

3.4.2. Scenario testing

The second approach to validate the agent-based model is a scenario test, i.e., comparing the simulation results with the real-world data. Specifically, we first simulate the DAS model before and after a new infrastructure opened in the area. Then, we compare the output of the ABM with observed changes in activity patterns from taxi trips data (ground truth) to validate the performance of ABM. The differences between the simulated results and real-world observations can reliably evaluate the effectiveness of our proposed DAS model.

4. Case study

We took Ningbo, China as our study area. Ningbo, located in the east of China's Zhejiang province, is a rapid developing city with 7.6 million residents. It also has one of the major ports in China and is one of the leading container ports in the world. Ningbo consists of five main districts in downtown: Haishu, Yinzhou, Zhenhai, Jiangbei, and Beilun.

4.1. Study data

4.1.1. POI data

The distributions of the POIs are presented in Fig. 3. In the study, we selected 459 general POIs, which includes most of the popular areas in Ningbo. More specifically, 200 residential areas, 22 workplaces, 20 schools, 28 restaurants, 27 hospitals, 28 banks, 33 recreation areas, 29 shopping areas, 2 inter-urban transport stations, and 70 locations related to other activities. For buildings with multiple POIs, we extracted their main activity characteristics. To prove the faithfulness of the POI selection, we utilised taxi data in Ningbo from March 9 to July 25, 2017 to statistically estimate the proportion of trips that drop-off near these POIs. The results show that for each activity, the proportion of the taxi trips that aimed for the selected POIs takes up over 70% of journeys (72.5% in weekday, and 72.1% in weekend). Based on previous work, around 70% trips have strong regulations (Zhang et al., 2019; Yue et al., 2011). We therefore demonstrated the selected POIs are representative and could be simulated for most of the daily destinations.

4.1.2. Taxi data

As for taxi data, we used 1.5 million taxi OD data in Ningbo from 9th March to 6th August in 2017. The trip data includes pick-up location (longitude, latitude), pick-up time, drop-off location (longitude, latitude), and drop-off time. To verify the prediction accuracy of the proposed DAS model, we split the data into three subsets: training data in normal days (from March 9 to 18), training data in vacation (from July 21 to 27), testing data for verification (from August 3 to 6).

We exploited taxi training data to fit two parameters in the ABM: (i) destination attractiveness α and (ii) travel cost β. To estimate the attractiveness of the POI, we first define 500 m as the buffer radius of the POI. Based on previous studies, all the taxi trips that drop-off within

Fig. 3. The spatial distribution of POIs in Ningbo, with eight activity categories.

the buffer radius are possibly attracted by the POI (Yue et al., 2012). We statistically estimate the number of taxi trips within 500 m of the POI as the POI's attractiveness. The cost (measured in terms of both time and financial cost) of taxi data travel exhibits a positive correlation with the route distance. As a result, we considered the route distance as the travel cost of the trip. The route distances are directly collected from Baidu Map API.

We further segmented the training data into two subsets: weekdays and weekends. For each subset, we fitted the parameters as follows: (i) For each activity, we statistically gathered the trips, and estimate the probabilities to each POI of the activity. The probabilities are calculated from Eq. (7).

$$P_{sij} = \frac{Num_{ij}}{\sum_j Num_i} \tag{7}$$

where P_{sij} means the probabilities of the citizens from location i to POI j for activity s; Num_{ij} represents the number of the taxi trips that pick-up in i and drop-off within 500 m of POI j; $\sum_j Num_i$ means the number of the taxi trips that pick-up in location i and drop-off within 500 m of all POIs that related to activity s. (ii) After that, we calibrate ABM model based on Eqs. (4) to (6), to fit the parameters geographically.

We validated the ABM model by comparing the taxi drop-off distribution for each IPOI from the test data with the simulated activity drop-off distribution from 5 am to 12 am. Furthermore, we took a scenario test to validate the ABM model on 28th July, 2017, a new wing of In-time shopping centre opened in Yinzhou district. By comparing the ABM outcomes and the observed changes of taxi trips data in activity patterns before and after the shopping centre built, we can validate the effectiveness of the ABM.

4.1.3. Trip diaries

In the study, we also collected trip diaries in 2019 in Ningbo to calibrate the ABM model (4.59% of the selected POIs closed in 2019). In order to ensure that the trip diaries are representative, we designed the questionnaire questions according to the residents' residence, whether they work, where they work, whether they go to school, and where schools locate, so that the relevant distribution meets the data of Ningbo Municipal Bureau of Statistics (GOV.CN, 2021a). We finally obtained one day's movements for 178 participants and activities based on their memories. The day was noted as a weekday or a weekend day. The trip diaries include trip information and participants' characteristics. Trip information described the citizens' one day movements as well as their activities in every hour. There are 625 trip records in total.

From trip diaries, we capture the activity-transition probabilities from each predecessor activity to each successor activity (based on Eq. (1)). The matrix is displayed in Fig. 8(a). The activity-transition matrix is directly applied in the agent-based model to simulate passengers' daily activity schedule. Moreover, combining with census data, we estimated the citizens' residential locations and their workplace.

4.1.4. Census data

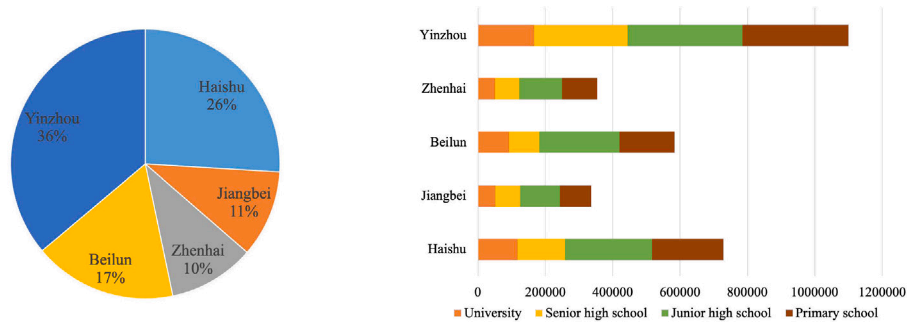
To further investigate the residents' characteristics, we collected census data from Ningbo Bureau of Statistics (GOV.CN, 2021a). More specifically, we explored the distributions of citizens' residential areas, as well as the proportion of current students in different areas. This information is used to generate the agents' characteristics in agent-based model (introduced in Section 3.2). Fig. 4 represents the statistical characteristics of the collected census data. We see that Yinzhou has the highest proportion of inhabitants (36%), while only a small percentage of citizens choose to live in Zhenhai (10%) or Beilun (11%). Since Yinzhou and Haishu are the city centre in Ningbo, the results are in line with expectations. We also noticed that most current students in Ningbo are junior high school students (35.6%) and primary school students (28.8%), while the numbers of college students are relatively small (only 15.2%). Additionally, we found that these percentages are similar to the education level of residents in Ningbo.

In addition, in order to determine critical parameters in the model, we referred to the legal provisions laid down by the Chinese government regarding work and study hours. Specifically, the working hours have been set between 7 to 9 h (GOV.CN, 2019), and the studying hours between 6 to 8 h (GOV.CN, 2021b).

4.2. Calibration results

We first perform a global calibration of the ABM model for three temporary activities (dining, shopping, and recreation) in both normal days and vacation. We defined 1 km × 1 km as a grid for model calibration. The K4 from Eq. (4) gives the best performance (with lowest Std.Error, P-value, and highest absolute T-value), we thus report the results in Table 2. It can be seen that all the estimated values of α are positive while the values of β negative; this is in line with their implications, i.e., destination attractiveness and travel cost. In addition, we can see that for shopping and recreation, the absolute values of α and β are much lower in vacation than in normal days. One interpretation is that most students do not go to school on summer vacation, they would have more time to spend in shopping and recreation areas, without considering much about travel cost. Meanwhile, Most students do not have income, they are less sensitive to the expensive goods than adults.

For every activity, the values of α are much lower at weekdays than at weekends. This implies that passengers pay more attention to the destination attractiveness at weekends. On the other hand, the absolute values of β are higher at weekdays than weekends for dining and shopping trips. This result is consistent with our experience that since people need to work and have less spare time on weekdays, they would select closer restaurants and shopping malls, in order to cut down travel times. For journeys for recreation (parks, gyms, etc.), the absolute



(a) Residential distribution (b) Current students distribution

Fig. 4. The distribution of residents and current students in five districts in Ningbo.

Table 2
Global parameter calibration on taxi data using best estimators K4 (Eq. (4)).

	Type	Time	α	β	Std.error	|T-value|	P-value
Dining	Normal day	Weekday	0.373	−0.482	0.020	22.310	2e−16
		Weekend	0.427	−0.103	0.018	14.350	−1.53e−16
	Vacation	Weekday	0.117	−0.480	0.021	13.790	2e−16
		Weekend	0.199	−0.586	0.024	13.558	2.6e−16
Shopping	Normal day	Weekday	0.353	−0.879	0.036	11.810	2e−16
		Weekend	0.396	−0.810	0.042	14.640	2e−16
	Vacation	Weekday	0.218	−0.591	0.023	15.191	2e−16
		Weekend	0.175	−0.538	0.024	10.799	2e−16
Recreation	Normal day	Weekday	0.256	−0.769	0.027	17.770	2e−16
		Weekend	0.349	−0.811	0.033	17.260	2e−16
	Vacation	Weekday	0.162	−0.560	0.019	15.725	2e−16
		Weekend	0.194	−0.499	0.022	12.070	2e−16

values of α and β are lower at weekdays than at weekends, which are consistent with trip diaries, where the number of trips for recreations in weekdays (5.1%) are much less than on weekends (10.4%). The results imply that the calibration results are reasonable.

4.2.1. Spatial distributions of α and β

Figs. 5 and 6 present the results of GWR calibration in normal days. The five downtown districts are mapped. Fig. 5 shows significant geographic variation of α for shopping, dining, and recreation activities. We see that most of the time, α takes the highest positive values (red) in Yinzhou and Haishu districts, and lower positive value, or even high negative values (green) in Zhenhai. This suggests people who live or work in Yinzhou and Haishu are more likely to prefer attractive destinations, whereas people from Zhenhai care least about the attractiveness of the destination. Moreover, we see that in Beilun, the values of α are positively high for shopping and dining activity, but negatively high for recreation-related activity. The result indicates that people in Beilun are very likely to select highly attractive shopping areas and restaurants, while they care much less about the attractiveness of the recreation areas, and even prefer an unattractive recreation place. Since we have previously defined attractive areas, which means it attracted more citizens historically. On the contrary, unattractive recreation areas tend to be located far from city centre, and the business competition is relatively low. However, these areas may locate close to suburban population centres (such as Beilun). Therefore, it is unsurprising that in Beilun, people prefer to select relatively unattractive parks or gyms in return for the convenience of shorter travel costs.

Fig. 6 shows the geographic variation of β. Generally, in most of the areas, the values of β are negative. In weekends, we see the absolute values of β are much lower (yellow) in the east of Haishu, and Beilun for dining activity. This suggest that passengers who live in these areas pay less attention on the travel costs to the restaurant. Combining with the α values in these areas in Fig. 5(e) (red), it is interesting to find that residents in the east of Haishu and Beilun are willing to travel far to attractive restaurants.

4.3. ABM simulation results

We use the agent-based model with five thousand agents to simulate one weekday and one weekend day. For the weekend day, we did not consider work and school related journeys. The agents' residential distributions are shown in Fig. 7. We can see that most citizens live in the east of Haishu and the north of Yinzhou.

4.3.1. Activity-transition analysis

Fig. 8 presents the activity-transition probabilities across Ningbo, i.e., the probabilities that citizens transfer from one activity to another in normal days (during summer vacation, the activity transitions are very similar, except that the schooling activity are not considered). To validate the performance of agent-based model, we first compare the activity-transition probabilities between the matrix from simulation (Figs. 8(b) and 8(c)) with the matrix from trip diaries as ground truth (Fig. 8(a)). Since there is only one matrix in trip diaries considering trips in both weekdays and weekends, we statistically regenerated the activity-transition matrix from simulations by considering both trips in weekdays and weekends. Mean Absolute Error (MAE) is used to measure the difference between the activity transition probabilities from simulation and the trip diaries. To be more specific, it is calculated as the MAE between the corresponding probabilities in the matrices of Fig. 8(a) and the weighted sum of which in Figs. 8(b) and 8(c) with a ratio of 5:2 since there are five weekdays and two weekend days in one week. The results show very low error (MAE = 0.015), indicating that the agent-based model accurately captures these transitions.

There are also some findings that we could capture from Fig. 8(c): (i) we see that passengers in weekends prefer to travel for recreation, shopping and dining. Most of these trips are closely related to residence (arrive from home, or leave for home); (ii) we also noticed that passengers in Ningbo have very little chance to transfer between school and work activities. Since most people can only take one job (work or study), we believe this phenomenon is reasonable and within our

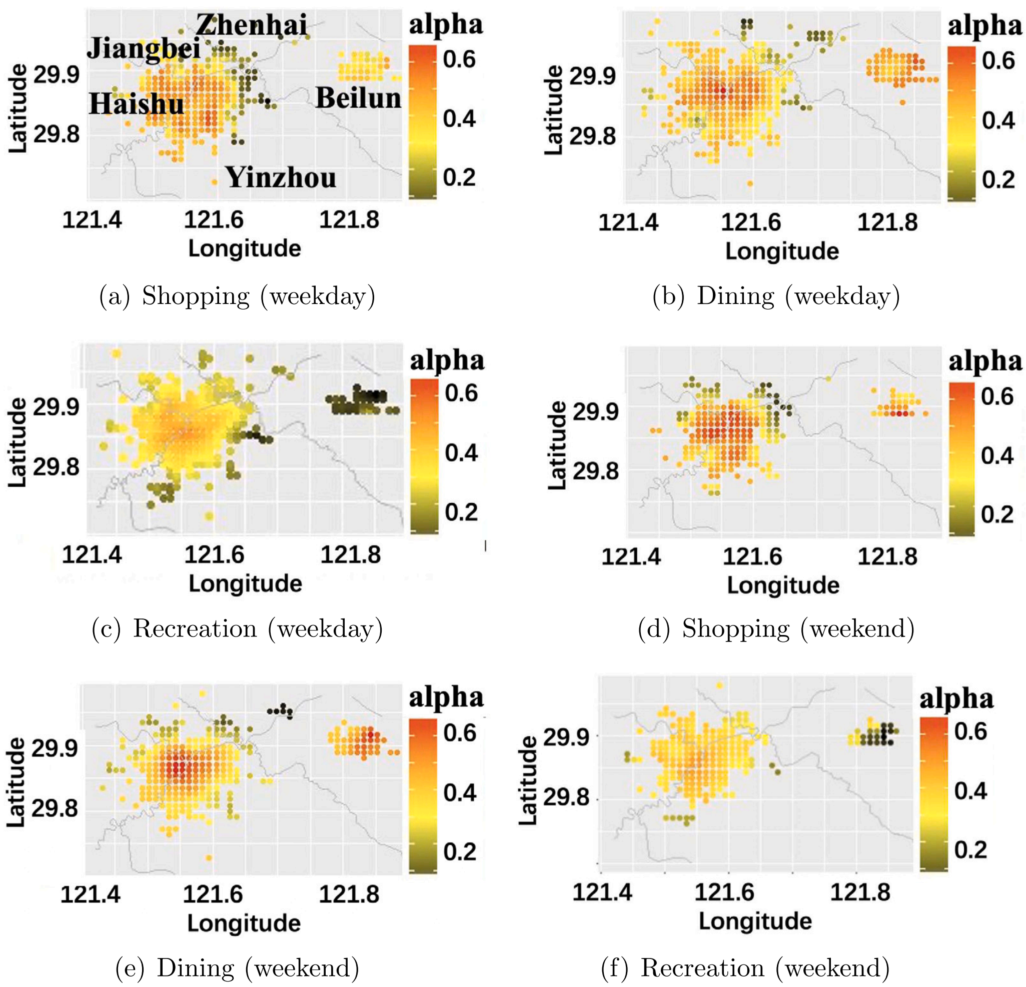


Fig. 5. The spatial distribution of α in normal days, fitted by GWR. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

expectation; (iii) after schooling activity, about 28% of the passengers would continue schooling. We interpret it as some students may take a break (such as lunch break) on the way before continuing to study at school. (iv) in Ningbo, about 27% of people will choose to work overtime after the end of the working day.

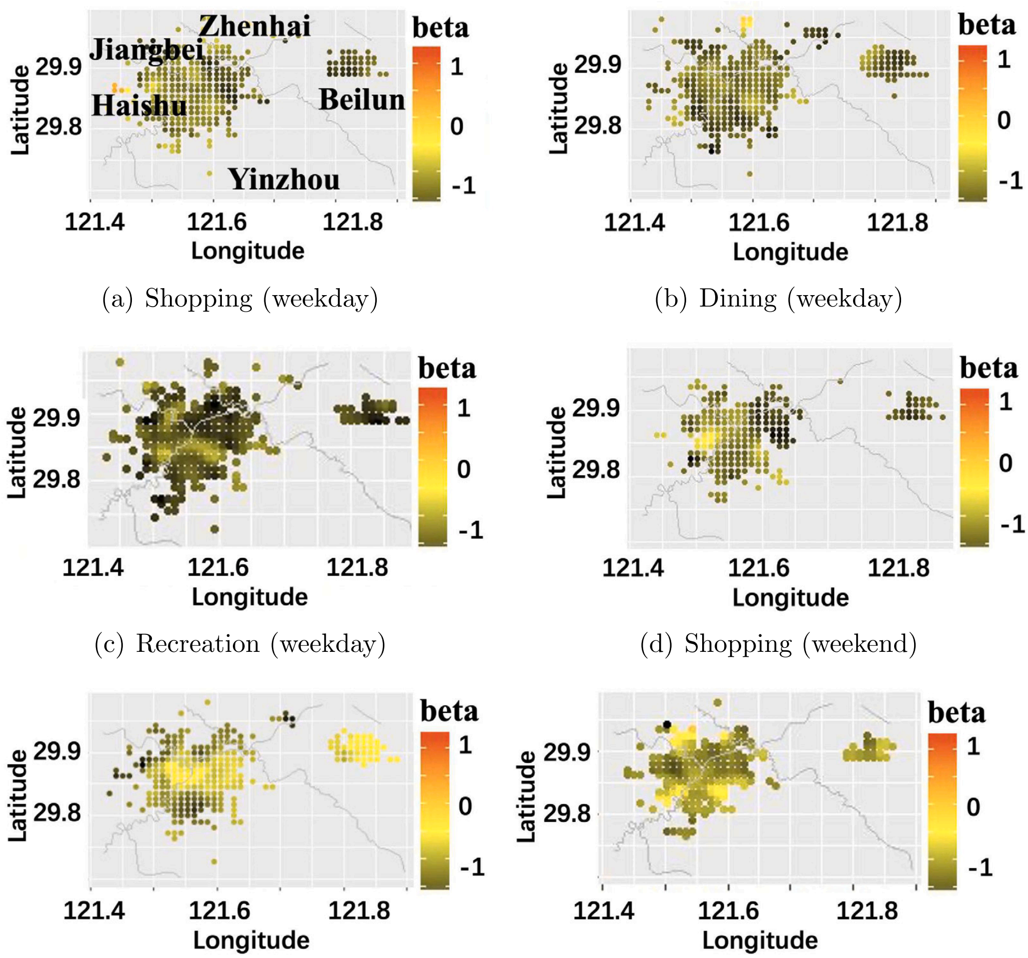
4.3.2. Activity schedule analysis

Here, we extracted some travel patterns from the DAS results. Fig. 9 presents the citizens' activities per hour from 5 am to 11 pm in normal days. It should be noted that, since most people do not go to work/school on weekends, we did not simulate work and school-related trips at weekends.

It can be seen from Fig. 9 that (i) most people prefer staying at home at noon (around 12 pm) and at night (from 8 pm to 6 am), which is in line with real situation. (ii) The shopping and dining behaviours are similar. More specifically, only a small proportion of trips from 1 pm to 2 pm are aimed for shopping areas or restaurant. On the contrary, citizens for these destinations are quite active around 11 am, from 3 pm to 4 pm, and from 5 pm to 7 pm. It provided an evidence that it is reasonable for some trading areas adding both shopping and dining infrastructures to attract customers. We therefore suggested the business strategies could be improved based on this finding. (iii) The trips aimed for inter-urban transport (e.g., railway station and airplane station) take a large proportion from 8:30 to 12:00, and 14:00 to 19:30. (iv) We also discover that most trips for temporary activities at weekends significantly outnumber that at weekdays. One interpretation is that most people need to work or go to school at weekday, and they could have more time at weekends for entertainment or other personal activities. In particular, for shopping and dining activities, more trips start before 12 o'clock, while for transport and recreation activities, more citizens prefer to start the trip in the afternoon, especially after 2 pm.

Furthermore, to explain inter-urban transport variability, we hypothesise that passengers' transport behaviours are largely influenced by the departure schedule of trains and airplanes, such that when there are more trains departure, passengers are more likely to travel by train. To test our hypothesis, we considered the number of trains' departure for each hour in Ningbo as a proxy of trains' departure arrangement, and use Pearson's correlation coefficient to identify whether there is a correlation. Pearson's correlation coefficient (PCC) is a measurement to evaluate the relationship between two distributions, which is from -1 to 1. The higher the absolute value of PCC is, the more the two distributions are related.

We collect the train departure data from government website of Haishu. The distributions of the proportions of train departure data and the transport trips volume are shown in Fig. 10. It clearly demonstrates a similar trend of inter-urban transport behaviour between the DAS simulation results and the real-world data. The Pearson Correlation between the real and simulation data is 0.757, which is significant at the 0.01 level (sigma $= 1.74 \times 10^{-4}$). It indicates a very strong relationship between passengers' start time of inter-urban trips and the number of available trips at the same time. It therefore suggests that the DAS model has good performance on travel behaviour forecasting. Also, we noticed that from 9 am to 10 am, and from 2 pm to 6 pm, there are



(a) Shopping (weekday)

(b) Dining (weekday)

(c) Recreation (weekday)

(d) Shopping (weekend)

(e) Dining (weekend)

(f) Recreation (weekend)

Fig. 6. The spatial distribution of β in normal days, fitted by GWR. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

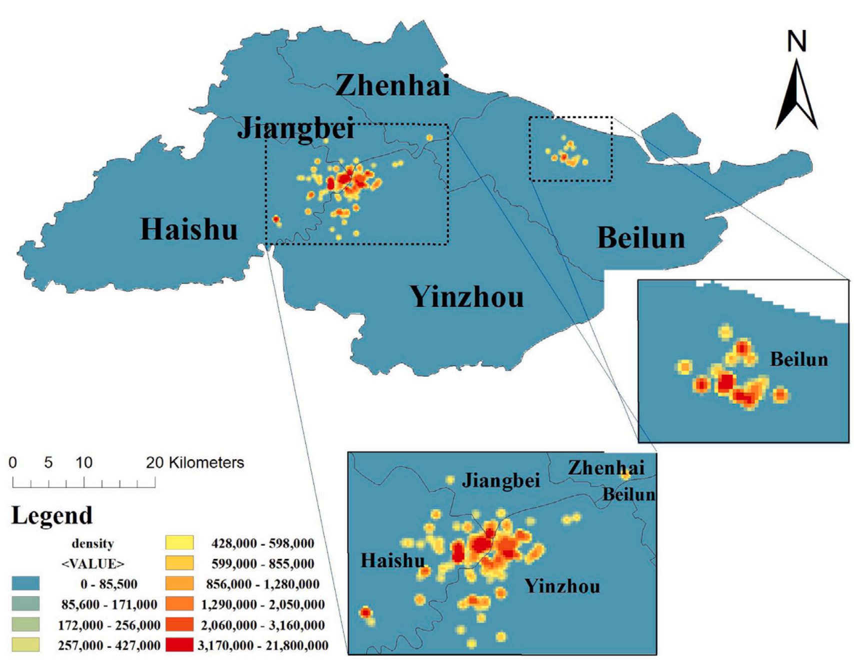
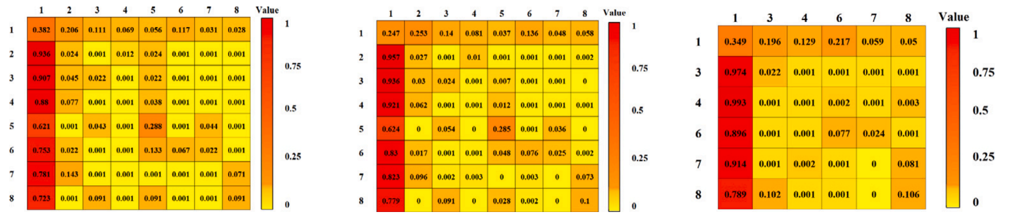


Fig. 7. Residential density in Ningbo.



(a) Ground truth (b) Simulation results (weekday) (c) Simulation results (weekend)

Fig. 8. The activity-transition matrix in normal days generated from (a) trip diaries, (b) simulation (weekday), and (c) simulation (weekend). The numbers of each row mean the predecessor activities, and of each column mean the successor activities of eight categories (1: Residence; 2: Work; 3: Shopping; 4: Dining; 5: Schooling; 6: Recreation; 7: Transport; 8: Others).

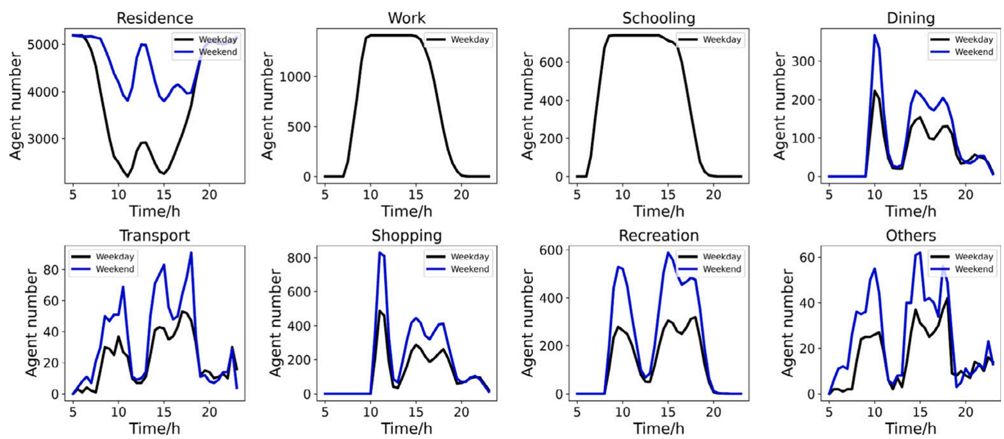


Fig. 9. Temporal variations of eight activities in normal days based on simulation. Work and school-related activities are not considered at weekend.

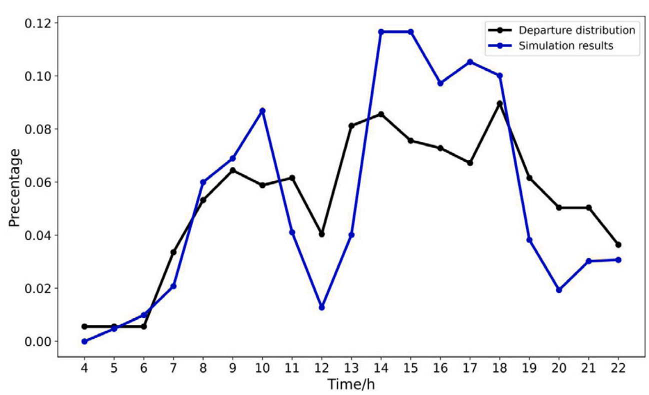


Fig. 10. The proportions of transport trips per hour from the real train departure data and the simulation results from DAS in normal days.

Table 3
Similarity comparisons between the real drop-off distributions of taxi trips to IPOIs and the DAS simulation results in normal days and summer vacation.

	Type	Time	MAE	RMSE
Dining	Normal days	Weekday	0.059	0.063
		Weekend	0.057	0.067
	Vacation	Weekday	0.025	0.030
		Weekend	0.016	0.023
Shopping	Normal days	Weekday	0.054	0.064
		Weekend	0.060	0.069
	Vacation	Weekday	0.022	0.025
		Weekend	0.031	0.042
Recreation	Normal days	Weekday	0.044	0.050
		Weekend	0.045	0.051
	Vacation	Weekday	0.028	0.032
		Weekend	0.011	0.014

around 1.5% difference between two distributions. One interpretation is that some people would travel to train station not taking train, but to pick up some people who just arrived (such as their families, friends).

4.4. IPOI-based model validation

In this study, we selected three IPOIs for validation, i.e., a restaurant (Xiangshu Restaurant), a shopping mall (Yinzhou Wanda Shopping Centre), and a recreation area (Ningbo Children's Park). Despite the limited number of IPOIs, the selected IPOIs are representative in terms of type coverage, including dining, shopping, and recreation activities, which ensures the reliability of the validation results. Specifically, for model validation, we statistically calculated the taxi drop-off numbers per hour for each IPOI from the testing dataset (taxi trips data from August 3 to August 6 in 2017), as the ground truth. We validated the DAS model by comparing the ground-truth and simulated drop-off densities for each activity in normal days (from Mar 9 to 18), and in summer vacation (July 21 to 27). Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are exploited to evaluate the similarities.

We compared the similarities of the temporal proportion of the related trips among all the trips for the same activity. The results are shown in Table 3. It can be seen that, the DAS simulation results are reliable (with very low values of MAE and RMSE).

4.5. Scenario testing

We have demonstrated that the DAS model is able to simulate passengers' activity patterns. To further examine the effectiveness of

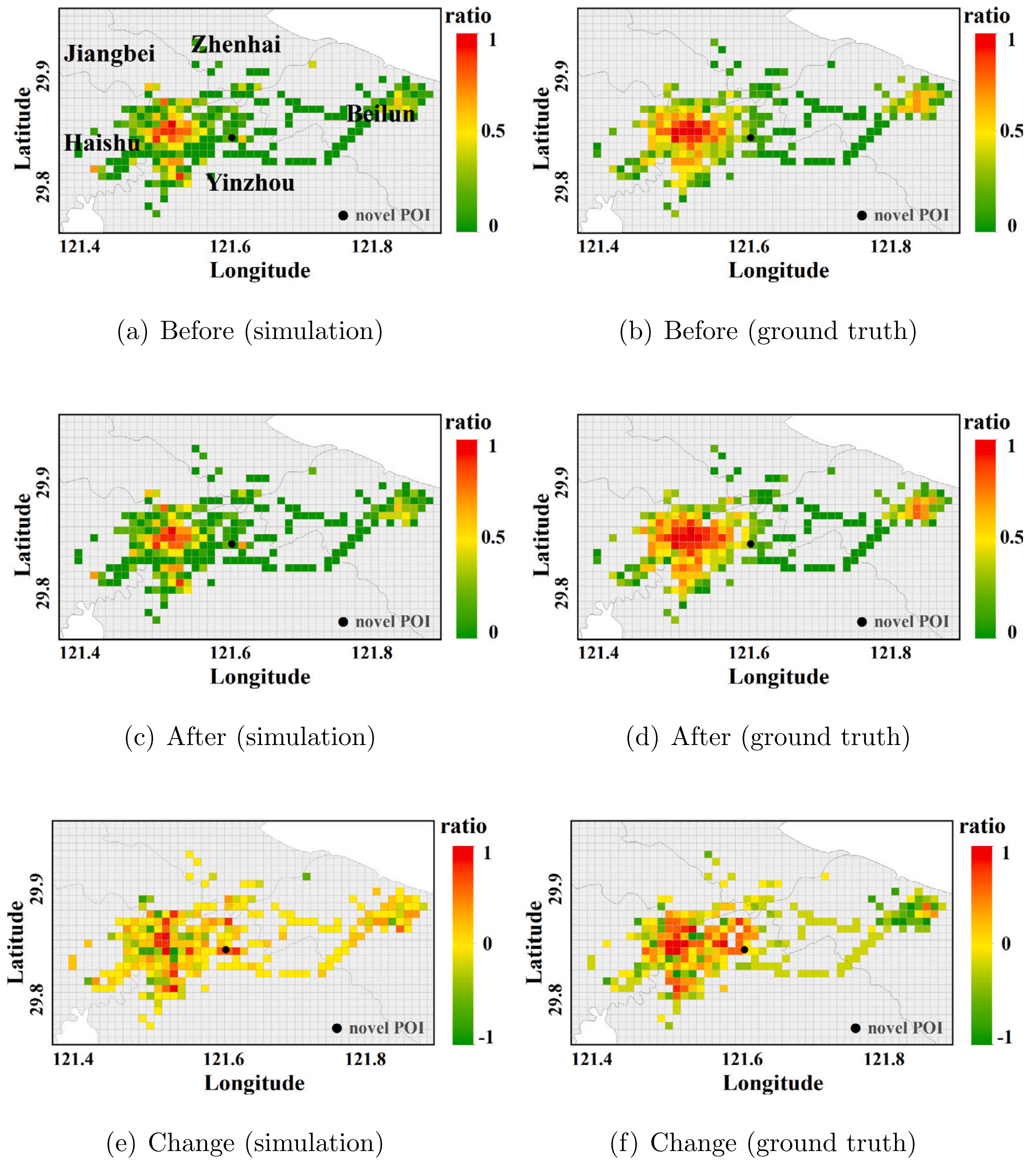


Fig. 11. Simulation and ground truth of activity volumes in weekday across Ningbo before and after a new wing of In-time shopping mall is opened in Yinzhou (marked by a black dot). DAS simulation: (a) before, (c) after, (e) change. Ground truth taxi data of drop-off volumes: (b) before, (d) after, (f) change.

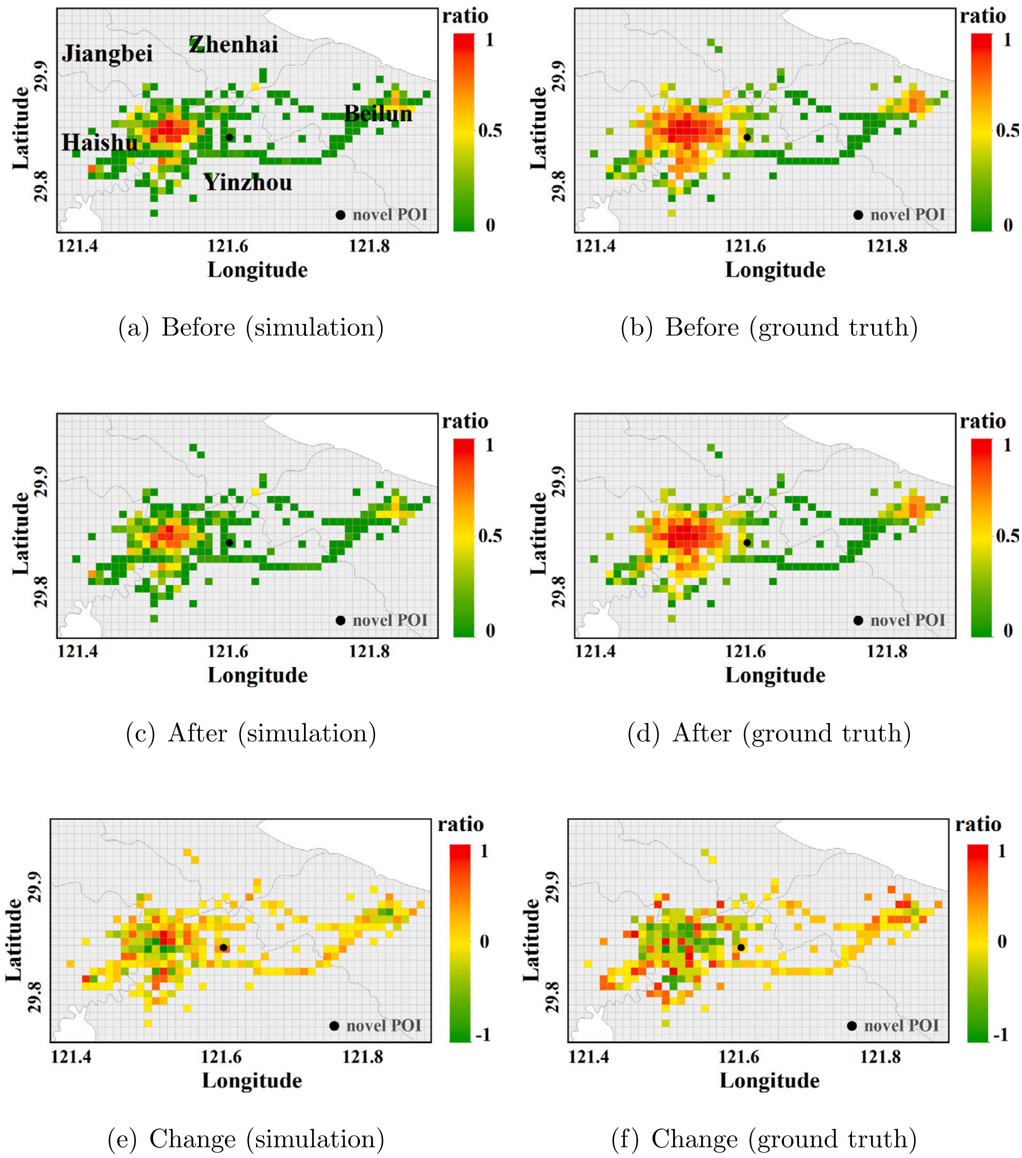
Table 4
Statistical comparisons of real vs. simulated distributions of taxi trip times as a result of a new POI.

Time	MAE	RMSE
Weekday	0.164	0.271
Weekend	0.139	0.236

DAS model, we use real-world data for scenario test. On July 28, 2017, a new wing of In-time shopping centre opened in Yinzhou, Ningbo. We collect the taxi GPS data for the period before (from July 21 to 27) and after (from August 3 to August 6) the shopping centre opened. By comparing the simulation results and the taxi trips in the change of travel behaviours before and after the opening of the new shopping centre, we can validate the effectiveness of DAS. We utilised MAE and RMSE to test the accuracy of the scenario test. The results are shown in Table 4. We see that the scenario test has good performance with low value of MAE and RMSE.

Fig. 11 shows the scenario test results of weekday. Specifically, Figs. 11(b)–11(d) shows the spatial distribution of normalisation of taxi drop-off volumes (ground truth) in Ningbo, before (Fig. 11(b)) and after (Fig. 11(d)) the new In-time shopping centre is opened. We present the difference as percentage in Fig. 11(f). Output of the agent-based model simulation of this scenario is presented in Figs. 11(a)–11(e). Generally, the result of the agent-based model is similar to reality. In particular, we see that the visiting densities are clearly increased around the new shopping mall (located in black dot). Meanwhile, we also noticed that the opening of the In-time in Yinzhou results in an increase in activity in some areas in the west of Yinzhou and in the east of Beilun, which are far from the In-time shopping mall. Although this observation is somewhat counter-intuitive, the agent-based model simulation (Fig. 11(e)) can still well predict this increase in activity, which further shows the effectiveness of our proposed method.

Fig. 12 shows the scenario test results of weekend. We also see similar distributions between DAS simulation results and the reality, which demonstrates that DAS can accurately describe citizens' movements in temporary activities (shopping, dining, etc.). Moreover, we also noticed that the opening of the in-time in Yinzhou results in an increase in activity in the area in the south west of Haishu, which is not the same as the pattern on a weekday. It is different to what we usually expect, since a new shopping centre usually attracts more customers around it



(a) Before (simulation) (b) Before (ground truth)

(c) After (simulation) (d) After (ground truth)

(e) Change (simulation) (f) Change (ground truth)

Fig. 12. Simulation and ground truth of activity volumes in a weekend across Ningbo before and after a new wing of In-time shopping mall is opened in Yinzhou (marked by a black dot). DAS simulation: (a) before, (c) after, (e) change. Ground truth taxi data of drop-off volumes: (b) before, (d) after, (f) change.

in the weekend; however, the agent-based model simulation can still capture this change. This further shows the powerful performance of ABM model.

While the real-world taxi data are not able to reveal why activities increase in the west of Yinzhou, the DAS can enable us to query this kind of detail. Specifically, we extracted visiting journeys for each activity, and discovered that the recreational activities dramatically increased during weekdays in the west of Yinzhou after the new wing of In-time shopping mall is developed. By analysing the simulation results, we surprisingly discovered that there exists business competitions between recreation areas and shopping centres. In particular, when a new shopping area has been built, some citizens who used to travel to a nearest park (30% in weekend, and 18% in weekday) would be attracted by the new shopping area, and not travel to the previous recreation area. One interpretation is that although the shopping malls can basically meet the citizens' daily necessities, most shopping centres in metropolitan city have entertainment facilities, such as cinema, KTV, etc. Since both activities (shopping, recreation) have an entertainment feature, it could be reasonable that the recreation areas could be influenced by the opening of a closely located shopping area. Searching on the unique hot spot in west of Haishu in weekend, we can attribute the pattern to the habit that people in the east of Haishu are willing to travel far to attractive restaurants according to the result of calibrated parameters. Of particular interest, we found that the increase in activities there was attributed to medical activities from categorising the change of activities in the east of Beilun. Thus it can be inferred that a newly opened shopping mall in Yinzhou can cause an increase of medical activities in Beilun, an area far away from Yinzhou. Similar situations have been documented in other literature (Gong et al., 2020a), which implies a latent correlation between medical and shopping activities that cannot be discerned solely through intuition.

The capability of examining details of human activities further demonstrates the great potential of our proposed model.

5. Conclusion

In this paper, we have introduced a Destination-aware Activity Simulation (DAS) model to map the state transitions between successive activities commonly found in cities in a probabilistic manner. An agent-based model with parameters of destination attractiveness and travel cost is core to the DAS model to simulate individual behavioural activities across city. We proposed a novel geographical calibration method for the agent-based model using GWR model to accommodate spatial variations of parameters. We calibrated the agent-based model for Ningbo, China, using 625 trip diaries and over 1.5 million taxi trips from March to August in 2017. Results demonstrated that the proposed DAS model with spatial-aware calibration achieved good performance on passengers' daily activity patterns simulation.

Furthermore, we presented a scenario test to validate the DAS model on simulating the impact of a new shopping mall developed in Yinzhou; and the simulation results are compared with the real travel behaviours change during the period before and after the new shopping mall was opened for validation observed from taxi trips. The agent-based model predicted not only intuitive outputs, such as the increased volume of shopping activities in weekday around the new infrastructure, but also counter-intuitive outputs, such as increased volume of recreation activities in weekday on the west of Yinzhou. Although some outcomes are unexpected, it is compelling to see that the DAS model could capture these changes and demonstrate accurate prediction capabilities. The results also demonstrated that there is a strong correlation between recreation activity and shopping activity. More specifically, there exists business competitions between recreation areas and the shopping malls, probably due to the fact that both shopping and recreation areas have an entertainment function.

This study has a wide range of applications. Firstly, our DAS model could provide good descriptions of the interactions between citizens and urban design. Therefore, it could provide suggestions for governments on locating new infrastructure. Secondly, with the support of road networks, DAS could predict the traffic conditions, as well as the traffic volume changes when the road design has been changed. Moreover, by using the DAS model, we have quantified the competitive relationship between shopping centres and recreation areas, this could serve as guidance for the governments in future urban design.

In the future, we consider to extend the DAS model to a framework, which would help governments locate new infrastructures, in order to reduce traffic congestion and shorten passengers' trip distances. Moreover, with the support of medical data, the DAS model could simulate and predict large scale transmission of epidemics. It is also worth developing an automatic-calibration method for the DAS model that could be widely used for different cities.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

This work is supported by the National Natural Science Foundation of China [grant numbers 42101472, 62172373, 61872325], Fundamental Research Funds for the Central Universities [Grant No. 2-9-2020-028], GHFUND B under Grant ghfund202107021958, and the Hong Kong Polytechnic University Start-Up under Grant BD41.

References

Auld, J., Hope, M., Ley, H., Sokolov, V., Xu, B., Zhang, K., 2016. POLARIS: Agent-based modeling framework development and implementation for integrated travel demand and network and operations simulations. Transp. Res. C 64, 101–116.

Balaraman, V., Athle, D., Singh, M., 2015. An agent based exploration of a relationship between daily routines and convenience store footfalls. In: Proceedings of the Conference on Summer Computer Simulation. pp. 1–9.

Bowman, J.L., Ben-Akiva, M.E., 2001. Activity-based disaggregate travel demand model system with activity schedules. Transp. Res. A 35 (1), 1–28.

Cheng, Z., Caverlee, J., Lee, K., Sui, D., 2011. Exploring millions of footprints in location sharing services. In: Proceedings of the International AAAI Conference on Web and Social Media. pp. 81–88.

Chowell, G., Hyman, J.M., Eubank, S., Castillo-Chavez, C., 2003. Scaling laws for the movement of people between locations in a large city. Phys. Rev. E 68 (6), 066102.

Crooks, A.T., Heppenstall, A.J., 2012. Introduction to agent-based modelling. In: Agent-Based Models of Geographical Systems. Springer, pp. 85–105.

Gong, S., Cartlidge, J., Bai, R., Yue, Y., Li, Q., Qiu, G., 2020a. Data-driven agent-based model of intra-urban activities. In: 2020 5th IEEE International Conference on Big Data Analytics. ICBDA, IEEE, pp. 160–166.

Gong, S., Cartlidge, J., Bai, R., Yue, Y., Li, Q., Qiu, G., 2020b. Extracting activity patterns from taxi trajectory data: a two-layer framework using spatio-temporal clustering, Bayesian probability and Monte Carlo simulation. Int. J. Geogr. Inf. Sci. 34 (6), 1210–1234.

Gong, S., Cartlidge, J., Bai, R., Yue, Y., Li, Q., Qiu, G., 2021. Geographical and temporal huff model calibration using taxi trajectory data. GeoInformatica 25 (3), 485–512.

Gong, S., Qin, J., Xu, H., Cao, R., Liu, Y., Jing, C., Hao, Y., Yang, Y., 2023. Spatio-temporal parking occupancy forecasting integrating parking sensing records and street-level images. Int. J. Appl. Earth Obs. Geoinf. 118, 103290. http://dx.doi.org/10.1016/j.jag.2023.103290.

GOV.CN, a., 2019. Labour law of PRC. http://www.npc.gov.cn/npc/c30834/201901/ffad2d4ae4da4585a041abf66e74753c.shtml.

GOV.CN, b., 2021a. Ningbo bureau of statistics. http://vod.ningbo.gov.cn:88/nbtjj/tjnj/2018nbnj/indexch.htm.

GOV.CN, c., 2021b. The regulations of school hygiene. http://www.gov.cn/zhengce/2020-12/25/content_5574078.htm.

Grazzini, J., Richiardi, M., 2015. Estimation of ergodic agent-based models by simulated minimum distance. J. Econom. Dynam. Control 51, 148–165.

Grazzini, J., Richiardi, M.G., Tsionas, M., 2017. Bayesian estimation of agent-based models. J. Econom. Dynam. Control 77, 26–47.

Huff, D.L., 1964. Defining and estimating a trading area. J. Mark. 28 (3), 34–38.

Kaelbling, L.P., Littman, M.L., Moore, A.W., 1996. Reinforcement learning: A survey. J. Artificial Intelligence Res. 4, 237–285.

Koushik, A.N., Manoj, M., Nezamuddin, N., 2020. Machine learning applications in activity-travel behaviour research: a review. Transp. Rev. 40 (3), 288–311.

Lim, K.H., Chan, J., Karunasekera, S., Leckie, C., 2019. Tour recommendation and trip planning using location-based social media: A survey. Knowl. Inf. Syst. 60 (3), 1247–1275.

Liu, F., Janssens, D., Cui, J., Wang, Y., Wets, G., Cools, M., 2014. Building a validation measure for activity-based transportation models based on mobile phone data. Expert Syst. Appl. 41 (14), 6174–6189.

Liu, Y., Kang, C., Gao, S., Xiao, Y., Tian, Y., 2012. Understanding intra-urban trip patterns from taxi trajectory data. J. Geogr. Syst. 14 (4), 463–483.

Miller, H.J., 2021. Activity-based analysis. In: Handbook of Regional Science. pp. 187–207.

Nakanishi, M., Cooper, L.G., 1982. Simplified estimation procedures for MCI models. Mark. Sci. 1 (3), 314–322.

O'Kelly, M.E., 1999. Trade-area models and choice-based samples: methods. Environ. Plan. A 31 (4), 613–627.

Platt, D., 2020. A comparison of economic agent-based model calibration methods. J. Econom. Dynam. Control 113, 103859.

Rasouli, S., Timmermans, H., 2014. Activity-based models of travel demand: promises, progress and prospects. Int. J. Urban Sci. 18 (1), 31–60.

Roadknight, C., Aickelin, U., Sherman, G., 2012. Validation of a microsimulation of the port of dover. J. Comput. Sci. 3 (1–2), 56–66.

Rodrigue, J.-P., Comtois, C., Slack, B., 2012. The Geography of Transport Systems. Langara College.

Saeedi, S., 2018. Integrating macro and micro scale approaches in the agent-based modeling of residential dynamics. Int. J. Appl. Earth Obs. Geoinf. 68, 214–229.

Shi, W., Goodchild, M.F., Batty, M., Kwan, M.-P., Zhang, A., et al., 2021. Urban Informatics. Springer.

Simoes, J.A., 2012. An agent-based/network approach to spatial epidemics. In: Agent-Based Models of Geographical Systems. Springer, pp. 591–610.

Song, C., Koren, T., Wang, P., Barabási, A.-L., 2010. Modelling the scaling properties of human mobility. Nat. Phys. 6 (10), 818–823.

Sturley, C., Newing, A., Heppenstall, A., 2018. Evaluating the potential of agent-based modelling to capture consumer grocery retail store choice behaviours. Int. Rev. Retail, Distrib. Consum. Res. 28 (1), 27–46.

Tu, W., Cao, R., Yue, Y., Zhou, B., Li, Q., Li, Q., 2018. Spatial variations in urban public ridership derived from GPS trajectories and smart card data. J. Transp. Geogr. 69, 45–57.

Walsh, S.J., Malanson, G.P., Entwisle, B., Rindfuss, R.R., Mucha, P.J., Heumann, B.W., McDaniel, P.M., Frizzelle, B.G., Verdery, A.M., Williams, N.E., et al., 2013. Design of an agent-based model to examine population–environment interactions in nang rong district, thailand. Appl. Geogr. 39, 183–198.

Wang, Z., He, S.Y., Leung, Y., 2018. Applying mobile phone data to travel behaviour research: A literature review. Travel Behav. Soc. 11, 141–155.

Wang, J., Kong, X., Xia, F., Sun, L., 2019. Urban human mobility: Data-driven modeling and prediction. Acm Sigkdd Explor. Newsl. 21 (1), 1–19.

Wang, S., Shao, C., Zhang, J., Zheng, Y., Meng, M., 2022. Traffic flow prediction using bi-directional gated recurrent unit method. Urban Inform. 1 (1), 16.

Wu, L., Zhi, Y., Sui, Z., Liu, Y., 2014. Intra-urban human mobility and activity transition: Evidence from social media check-in data. PLoS One 9 (5), e97010.

Xu, C., Li, F., Xia, J., 2023. Fusing high-resolution multispectral image with trajectory for user next travel location prediction. Int. J. Appl. Earth Obs. Geoinf. 116, 103135.

Yamamoto, T., Kitamura, R., Fujii, J., 2002. Drivers' route choice behavior: analysis by data mining algorithms. Transp. Res. Rec. 1807 (1), 59–66.

Yang, M., Yang, Y., Wang, W., Ding, H., Chen, J., 2014. Multiagent-based simulation of temporal-spatial characteristics of activity-travel patterns using interactive reinforcement learning. Math. Probl. Eng. 2014.

Yin, J., Chi, G., 2022. A tale of three cities: uncovering human-urban interactions with geographic-context aware social media data. Urban Inform. 1 (1), 20.

Yin, M., Sheehan, M., Feygin, S., Paiement, J.-F., Pozdnoukhov, A., 2017. A generative model of urban activities from cellular data. IEEE Trans. Intell. Transp. Syst. 19 (6), 1682–1696.

Yue, Y., Wang, H.d., Hu, B., Li, Q.q., 2011. Identifying shopping center attractiveness using taxi trajectory data. In: Proceedings of the 2011 International Workshop on Trajectory Data Mining and Analysis. pp. 31–36.

Yue, Y., Wang, H.-d., Hu, B., Li, Q.-q., Li, Y.-g., Yeh, A.G., 2012. Exploratory calibration of a spatial interaction model using taxi gps trajectories. Comput. Environ. Urban Syst. 36 (2), 140–153.

Zhang, H., Shi, B., Zhuge, C., Wang, W., 2019. Detecting taxi travel patterns using GPS trajectory data: A case study of Beijing. KSCE J. Civ. Eng. 23, 1797–1805.

Zhao, T., Huang, Z., Tu, W., He, B., Cao, R., Cao, J., Li, M., 2022. Coupling graph deep learning and spatial-temporal influence of built environment for short-term bus travel demand prediction. Comput. Environ. Urban Syst. 94, 101776. http://dx.doi.org/10.1016/j.compenvurbsys.2022.101776.

Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., Li, H., 2019. T-gcn: A temporal graph convolutional network for traffic prediction. IEEE Trans. Intell. Transp. Syst. 21 (9), 3848–3858.

Zhou, M., Wang, D., Li, Q., Yue, Y., Tu, W., Cao, R., 2017. Impacts of weather on public transport ridership: Results from mining data from different sources. Transp. Res. C 75, 17–29. http://dx.doi.org/10.1016/j.trc.2016.12.001.