WILEY | Hindawi

Research Article

Exploring Built Environment Influence on Taxi Vacant Time in Megacities: A Case Study of Chongqing, China

Guangyue Nian [ID],[1] **Jianyun Huang** [ID],[2] and **Daniel (Jian) Sun** [ID][3]

[1]*School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao Tong University, Shanghai 200240, China*
[2]*School of Design, Shanghai Jiao Tong University, Shanghai 200240, China*
[3]*Smart City and Intelligent Transportation Interdisciplinary Center, College of Future Transportation, Chang'an University, Xi'an 710021, China*

Correspondence should be addressed to Jianyun Huang; hjy2005@sjtu.edu.cn

Driven by the emission peak and carbon neutrality targets, traditional cruising taxis are also under pressure to reduce emissions, and reducing taxi vacant time is an effective way to reduce emissions. This paper aims to examine the influence mechanism of built environment characteristics on vacant time of taxi trips from the supply-side perspective. To do so, the study uses a one week of taxi trajectory data of 2019 provided by a taxi company in Chongqing, a megacity in China. We take the divided square grid cell as the study units, calculate the taxi vacant time of 168 (7×24) hourly slots for each grid, and then calculate the built environment indicators of the grids with Point of Interest (POI) and vector road network data. A generalized additive model (GAM) based on panel data is constructed to investigate the influence mechanism of built environment on taxi vacant time. The results show that the time-varying trend of taxi vacant time is opposite to the time-varying trend of taxi trips and is similar to taxi trips in spatial distribution with significant spatial dependence. Taxi vacant time is negatively correlated with taxi trips. POI mixture, the number of enterprises, governmental agencies, and shopping services are negatively correlated with vacant time, while the number of sports and leisure services, hotel services, healthcare services, financial and insurance services, and living services are positively correlated with vacant time. An interesting finding is that metro stations significantly increase the vacant taxi time in the region. Average traffic flow speed, road density, the number of bus stops, catering services, residence communities, and motor vehicle services show strong nonlinear relationships with taxi vacant time, and their influence effects are alternately positive and negative. The study provides useful insights for understanding mechanisms of the role of built environment on taxi vacant time and has important implications for driver searching strategy improvement, taxi management measures development, spatial and temporal scheduling of taxi capacity resources, and urban transportation facilities layout planning.

1. Introduction

Taxis are an essential component of the urban integrated transportation system, which play a valuable role in complementing urban transportation because of time and space flexibility during operation. Because taxi passengers are random, taxis will be unoccupied a certain percentage of the time during their service. Vacant taxis contribute to pollution, energy waste, and road traffic congestion and reduce the effectiveness of taxi service [1, 2]. The high percentage of taxi vacant time or vacant miles indicates that the demand

for taxis is not well matched with the supply, which leads to increased passenger waiting time and reduces the level of taxi service and passenger satisfaction [3, 4].

Taxi vacant time is one of the important indicators of taxi operation level, which is influenced by multiple factors [5]. When the demand and supply are relatively stable, taxi operation level largely depends on the operation capacity of drivers, which relies heavily on the empirical knowledge and real-time perception of the external environment [6]. Generally, when a taxi is in service, its external environment is divided into two types: static environment and dynamic

environment; the former mainly includes the built environment of the city, while the latter includes weather, emergencies, and changes in supply and demand. Thus, taxi vacant time is also closely related to external factors, and external factors have a more significant impact than internal factors [7].

However, the current studies related to taxi operations have mainly focused on passenger searching and delivery strategies, and the external influencing factors and mechanisms of taxi vacant time have not been fully studied. Taxi vacant time or mileage is a potential cost of taxis [5], and finding external influences on this potential cost and trying to reduce it is not only attractive to drivers but also crucial to the objective assessment of the overall performance of taxi operation [8], capacity spatial and temporal optimization and scheduling [9], passenger ride guidance, and taxi stand layout planning. Thus, to explore the correlation between the external environment and taxi vacant time, this paper conducts a study on the influence mechanism of urban built environment on taxi vacant time and takes Chongqing, a megacity in China, as a case study. In this paper, we divide the study area into square grids with a side length of 1 km and calculate the vacant time of 168 hours in one week for each grid with taxi Global Position System (GPS) trajectory data and then conduct exploratory spatial data analysis of taxi vacant time. After that, we use the Point of Interest (POI) and OSM (OpenStreetMap) road network data to calculate the built environment indicators of each grid and then construct a generalized additive model (GAM) to find the key built environment factors affecting taxi vacant time and the mechanism of built environment on vacant time.

The remainder of the paper is organized as follows. Section 2 provides a brief literature review. Section 3 introduces a description of the data and study variables. Section 4 presents the research methodologies and the model. Section 5 analyzes the results with brief explanations. Section 6 discusses the whole study. Section 7 draws the conclusions, research limits, and future directions.

2. Literature Review

The built environment is composed of various buildings and places that have been artificially constructed and modified and is a combination of land use patterns, transportation systems, and a series of elements related to urban design that can influence the behavior of residents' activities [10]. The built environment differs from the natural environment in that it is a product of human civilization, providing a spatial, temporal, and social context for human activity, and is a combination of elements related to land use, urban design, and transportation systems. A Point of Interest (POI) is a specific physical location which someone may find interesting. Restaurants, retail stores, and grocery stores are all examples of Points of Interest. POI types and densities can characterize the urban vitality of a region, and the functional areas of a city can be identified by POI [11]. Many studies have used POI to calculate built environment indicators [12, 13].

Taxi vacant time is closely related to built environment. Zhang et al. conducted a study based on New York taxis showing that built environment has significant but spatio-temporal variability influences on taxi vacant time. Their study showed that the density of road and subway stations have negative impacts on empty taxi trip duration, and the density of city bus stations has positive impacts across times outside of Manhattan. The density of bike lanes has positive and negative impacts on empty taxi trip duration within and outside Manhattan, respectively, and the ratio of green space is only significant on weekends that are negative impacts within Manhattan and positive impacts outside of Manhattan. The impacts of the density of public facilities (i.e., schools and hospitals) vary greatly across times [5]. A vacant taxi staying somewhere is an important way to search for passengers, such as staying at the roadside or taxi stand to wait for passengers, and the staying time is part of the vacant time. Ke et al. studied the staying behavior of taxis, and the results showed that drivers' staying behavior differed significantly between traffic analysis zones, and the coefficients show that the number of car repair and maintenance shops, the number of gas stations, the existence of an airport, the existence of a railway station, and POI diversity have positive associations with stay frequency. In addition, the airport has a much higher impact on stay frequency than other built environment factors [9]. Lee and Sohn selected variables related to taxi demand to analyze the factors influencing the taxi vacancy duration, and the results showed that, except for the land use types and public transport accessibility at the trip origins and destinations, the other relevant factors such as operational characteristics, weather conditions, demographics, and socioeconomic characteristics also have significant associations with vacancy duration that follows a log-normal distribution [14].

The taxi vacant time is influenced by the driver operation strategy, and the built environment is an important basis for the driver operation decision. The taxi operation process is mainly divided into two stages, that is, how to search for passengers quickly when the vehicle is vacant and how to deliver passengers to their destinations quickly when they are picked up [15]. There are two main strategies for drivers to search passengers, that is, actively search for passengers and wait for passengers at a certain location, and studies have shown that drivers who actively search for passengers can get passengers faster and earn higher income [16]. Route choice is very important for drivers in both passenger searching and passenger delivering. The route choice of drivers when vacant is usually oriented towards minimizing the time to search a passenger [17], and drivers also consider the spatial and temporal characteristics of the search route [18]. Guo et al. used path unreliability as one of the indicators to describe the route choice behavior of vacant taxi drivers searching for passengers and found that path unreliability had a significant effect on the route choice behavior of vacant taxi drivers [19]. With the accumulation of experience and the combination of traffic conditions and passenger information, the driver will continuously optimize the route to minimize the passenger search time [20]. The driver's location and route choice may be more uncertain

when searching for passengers, but the destination is certain when delivering passengers, and the driver usually chooses the route with the goal of minimizing the travel time [21, 22]. The search distance and delivery speed dominate the income level of taxi drivers [23].

Drivers have regional preferences and the built environment potentially influences their choice of regional preferences, which in turn leads to differences in taxi vacant time. Drivers tend to have regional or location preferences when searching for passengers, and these regions or locations have significant land use or built environment characteristics. For example, most drivers prefer to go to hotspots such as commercial areas, CBDs (Central Business Districts), and transportation hubs to search for passengers, where they can get passengers faster, thus reducing search time and increasing income [24]. Meanwhile, it was found that passenger demand at the pick-up area, hotspot locations, service location preference, and major transport hubs positively influence a taxi driver's next choice of passenger pick-up location [25]. Different regional or location preferences lead to differences in the probability or duration of taxi presence in the different built environment and thus affect the taxi vacant time under different built environments. Zong et al. used a zero-inflated negative binomial model to study the effects of internal and external factors on drivers' passenger-seeking behavior, and the results showed that external factors such as land use type, traffic conditions, and road hierarchy have a more significant effect on drivers' passenger-searching behavior, while internal factors such as prior passenger-searching experiences have relatively little effect, and drivers follow different passenger-searching strategies at different times of the day [7].

Urban built environment is closely related to travel behavior. For example, the built environment around metro stations has a significant effect on metro station passenger flow [26], while metro systems in return can significantly reduce the frequency of private car use [27, 28]. The distance to the city center is positively associated with auto ownership and it has a more important effect than the distance to other centers [29], and neighborhoods with well-developed bicycle riding facilities may attract more cyclists [30]. The degree of land use mixture, residential density, metro station density, and road density influence travel distance and are negatively related to road traffic emissions [31, 32]. High four-way intersection proportion, road density, and population density in residential areas can reduce the probability of driving and encourage walking, biking, and transit modes of travel in the area [33].

The built environment has a significant effect on both the production and attraction of taxi trips [34]. Taxi demand is largely affected by the internal functions of the city, and there is an obvious spatiotemporal pattern in the impact of the mixed degree of urban functions on taxi demand [35]. Taxi trips are relatively more frequent in areas with high residential concentration and higher housing prices, and road density, parking density, bus stop density, and the percentage of regional commercial and public services are positively associated with taxi demand, while the percentage of regional residential areas and land use mixture are negatively

associated with taxi demand [36]. The effect of built environment on taxi trips is significantly spatially heterogeneous, with different degrees of influence and positive and negative impact effects in different areas of the city and at different times of the day [37]. Ridesourcing demand is also associated with built environment [38]. A study by Sabouri et al. based on urban census block group showed that land use mix and bus stop density are positively associated with ridesourcing demand, and intersection density and destination accessibility are negatively associated with ridesourcing [39]. The effects of built environment on the ridesourcing demand demonstrated significant spatial variations trends from urban to suburban neighborhoods [40].

In summary, although the study of taxi vacant time is necessary and important, the literature review suggests a number of limitations in existing studies of taxi vacant time.

First, travel research has presented evidence that built environment plays an important role in travel behavior and taxi demand. For instance, the built environment has significant effects on travel mode choice, car ownership and use, and taxi demand. However, most of these studies have been conducted from the traveler's perspective (demand-side), and there is a lack of research on the impact of built environment on taxi drivers' operating behavior from the supply-side perspective. The study of taxi supply and taxi demand is equally important because taxi supply is a matter of driver income, industry efficiency, and urban environment [41].

Second, the taxi vacant time has not been measured accurately, which is a research limitation mentioned several times in the related literature. Some studies treat the time interval between the previous passenger drop-off time and the next passenger pick-up time as the searching time, which is logically fine. But, in fact, the time calculated by the above method includes the time when the driver takes a short break, eats, refills fuel, and so forth, which differs from the actual searching time [14].

Third, the influence mechanism of single factors has not been further analyzed. The results of related studies showed positive, negative, or insignificant correlations of the influencing factors. But, in fact, the effects of single factors may differ significantly at different time periods and different magnitudes, and the mechanisms between the influencing factors and the explained phenomena have not been studied deeply and adequately.

Fourth, the influence of time-varying factors is not fully considered. The built environment is a static feature, and its parameters or indicators do not change over time, which is a time-invariant factor. Related studies have often considered time-invariant factors and have not fully considered the dynamic features that change over time. Transportation system is a system with significant dynamic characteristics, and its fluctuations in time and space are more obvious, so the corresponding dynamic characteristics should be fully considered.

To fill the above-mentioned gaps, this paper conducts a study on the influence mechanism of the built environment on taxi vacant time. The contributions of the paper are as follows.

From the taxi supply-side perspective, the key built environment factors affecting taxi vacant time and their influencing mechanisms are explored, forming a closed loop of research on the correlation between taxi supply and demand and built environment with the related research on the influence of built environment on taxi demand.

Through the accurate measurement and splitting in time and space of taxi vacant time, the research basis is made more reliable, and then more practical value and guiding research conclusions are drawn.

An in-depth study of the mechanism of the built environment's influence on taxi vacant time based on a generalized additive model is conducted, the nonlinear relationship between taxi vacant time and the built environment is analyzed, and we step out of the limitations of the linear relationship assumption.

Panel data with both spatial research units and temporal dimensions are constructed under the premise of dividing the study area, and the aim is to capture the effect of time-varying factors and time-invariant factors.

3. Data and Variables

3.1. Study Area. The study area of this paper is the central city of Chongqing. As one of the megacities in China, Chongqing is located in the southwest China mainland, with a central urban area of about 5,467 square kilometers and the number of permanent residents about 10.34 million in 2020. The central city of Chongqing straddles the Yangtze and Jialing rivers and four major mountain ranges, making it a typical cluster-type mountain city.

3.2. Data Processing. The study data includes taxi trajectory data, POI data, and OSM (OpenStreetMap) vector road network data. The fields of the trajectory data are vehicle number, time, longitude, latitude, instantaneous velocity, direction angle, and trip status codes (0: vacant, 1: occupied), and the time interval of the trajectory data positioning point is 15 seconds. A full week of taxi trajectory data is selected for the study, so it can avoid possible errors in single-day studies and analyze the variability of each day of the week. The selected time period is from May 20 to May 26 (Monday–Sunday), 2019, and the weather conditions varied very little during the week, with light rain and cloudy skies predominating. It was checked that the data fluctuated relatively little during the week, thus ensuring the reliability of the data quality. The hourly GPS points of the selected dataset for the week are shown in Figure 1. The GPS points of 168 hours on 7 days fluctuate between about 570,000 and 590,000, with small fluctuations, and the data can be considered generally smooth and without abnormalities.

The POI data is collected from Amap (also known as Gaode Map), and POI data contains information such as administrative area, name, longitude, latitude, address, telephone, and classification. The data are cleaned up to remove duplicate and incomplete data records and abnormal values. After the final POI cleaning, 238,090 POI records are got with complete and accurate information in 15 types.
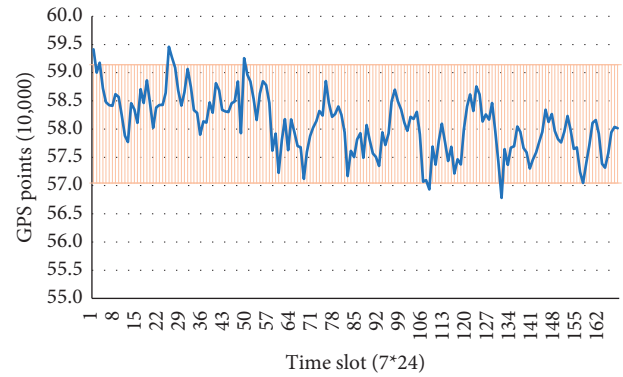

FIGURE 1: Fluctuation trend of the number of GPS positioning points per hour for 7 days. Note: the vertical line in the figure is the error line with 2 times the standard deviation.

In this paper, the taxi vacant time (TVT) is defined as the elapsed time between the drop-off time of one occupied trip and the pick-up time of the next occupied trip for one same taxi, in other words, the passenger searching time [5]. The taxi vacant time is calculated by the taxi trajectories. First, the missing values, duplicate values, and outlier of GPS positioning points are eliminated, and the positioning points with an instantaneous velocity greater than 120 km/h are eliminated [42]. Second, the taxi trajectories are constructed from the positioning points and the vacant taxi trajectories are extracted. Third, the abnormal vacant taxi trajectories are removed; that is, the vacant taxi lasts for more than 90 minutes, and the location points are removed with vehicle speed being 0 and displacement being 0 for more than 5 minutes from the vacant taxi trajectories. Fourth, the vacant taxi trajectories are matched to the map, and the vacant taxi time of each grid is calculated. Since the taxi vacant time is calculated by the actual vacant taxi trajectory and the taxi vacant time is filtered in the calculation, the taxi vacant time calculated in this paper is very close to the actual passenger searching time.

As regards the process of panel data construction, first, the study area is divided into $1 \, km \times 1 \, km$ square grids, and the vacant taxi travel time is calculated for each grid. Due to the significant time-varying characteristics of taxi demand, this paper slices the vacant travel time of each grid by hourly slots, and the vacant travel time of each grid is 168 (7×24) dimensions. It should be noted that some grids have zero vacant time in some hourly slots, and all grids with zero vacant time in all hourly slots will be excluded from the study. Next, the POIs are matched to the map, and the built environment indicators are calculated for each grid using the POI and OSM road network data. After the above process, the panel data is formed with the grid as the study unit.

3.3. Explained Variable. The explained variable is the total taxi vacant time (TVT) per hour in each grid, with each grid having a dimension in time of 168 hours (24 hours per day for 7 days), and the taxi vacant time is defined as described above. The total number of grids divided in the study area was 7,545. The total taxi vacant time (in minutes) in each grid is calculated as

$$\mathrm{TVT}_{i,j,k} = \sum_{m=1}^{n} \mathrm{TVT}_{i,j,k,m}, \qquad (1)$$

where $\mathrm{TVT}_{i,j,k}$ is the total taxi vacant time of grid i in the kth (0~23) hour of day j (1~7), $\mathrm{TVT}_{i,j,k,m}$ is the vacant time of mth taxi in grid i in the kth hour of day j, and n is the number of taxis passing through grid i in the kth hour of day j.

3.3.1. Time-Varying Characteristics.

The time series of total vacant time and total pick-ups for 168 hours in the study period is shown in Figure 2. Both vacant time and pick-ups fluctuate on a day-to-day basis. The taxi vacant time starts to increase rapidly from early in the morning and reaches a peak around 5:00 a.m., then decreases rapidly and reaches a low around 8:00 a.m., then fluctuates several times and reaches a low again around 7:00 p.m., and starts to rise slowly after 10:00 p.m. The fluctuation of taxi pick-ups, on the other hand, is basically opposite to the trend of the taxi vacant time. Poisson's correlation coefficient between vacant time and pick-ups was calculated to be −0.90, which also indicates that vacant time is highly negatively correlated with pick-ups. This is relatively easy to understand, because drivers find it relatively easier to search for passengers when taxi demand is high, and the time they spend in searching for passengers is bound to be shorter. However, we calculated the correlation coefficient between the taxi vacant time and the number of pick-ups in the hourly slots at the grid level, and the value was only −0.27. The correlation between the taxi vacant time and the number of pick-ups in the hourly slots at the grid level is much weaker, indicating a serious spatial mismatch between taxi demand and supply.

3.3.2. Spatial Characteristics.

The spatial distribution of total taxi vacant time for grids is shown in Figure 3. The grids with high vacant time are mainly concentrated in commercial centers, transportation hub areas, and residential concentrated areas.

3.4. Explanatory Variables.

The explanatory variables are the built environment characteristic indicators of the grid. The 14 types of POIs are kept unchanged as 14 built environment indicators; the transportation facility service POIs are converted into 3 types of indicators, that is, presence of metro stations (logical type, 0 means no metro stations in the grid, 1 means metro stations in the grid), number of bus stops, and road density (km/km^2). The other explanatory variables are longitude and latitude of the grid center, average traffic flow speed of the grid (km/h), and grid POI mixture. The relevant explanatory variables are defined and calculated, respectively, as follows.

Longitude and latitude of the grid center: the longitude and latitude of the grid center (World Geodetic System—1984 Coordinate System).

Road density (km/km^2): the density of roads per unit area in the grid, which is calculated based on OSM vector road network data and is calculated as follows:

$$\mathrm{DSR}_i = \frac{\sum_{r=1}^{n} \mathrm{LGR}_r}{\mathrm{AEG}_i}, \qquad (2)$$

where DSR_i is the road density of grid i, LGR_r is the length of the road numbered r in grid i (km), AEG_i is the area of grid i (km^2), and n is the number of roads in grid i.

Average traffic flow speed within the grid (km/h): the average speed of all passing taxis within a specific time slot of a given grid, which is calculated as follows:

$$\mathrm{TFS}_{i,j,k} = \frac{1}{n} \sum_{m=1}^{n} \frac{d_{i,j,k,m}}{t_{i,j,k,m}}, \qquad (3)$$

where $\mathrm{TFS}_{i,j,k}$ is the average flow speed of grid i in the kth (0~23) hour of day j (1~7), $d_{i,j,k,m}$ is the travel distance of the mth taxi passing through grid i in the kth hour of day j, $t_{i,j,k,m}$ is the travel time of the mth taxi passing through grid i in the kth hour of day j, and n is the number of taxis passing through grid i in the kth hour of day j.

Grid POI mixture: the mixture degree of POI types in the grid, where POI mixture characterizes the degree of diversity of POI, and it also characterizes the degree of land use mix to a certain extent [43], which is calculated as follows:

$$\mathrm{PMT}_i = \begin{cases} \left(\dfrac{-1}{\ln N_i} \right) \displaystyle\sum_{m=1}^{N} p_{i,m} \ln p_{i,m}, & N > 1, \\[2mm] 0, & N = 0/1. \end{cases} \qquad (4)$$

In the above equation, PMT_i is the POI mixture of the ith grid, N_i is the number of types of POIs in the ith grid, and $p_{i,m}$ is the percentage of the mth type of POIs in the ith grid to the number of all POIs in the grid. POI mixture is a dimensionless value, and its value ranges from 0 (homogeneous) to 1 (most mixed), and a larger value indicates a higher mixing degree [44]. In particular, when there is no POI in the grid or only one type of POI, the mix degree is 0 [45]. In addition, if neither POI distribution in the grid nor road density is 0, the grid is excluded.

All variable values were calculated, discriminated, and filtered, and the summary of the variables is shown in Table 1. The difference in vacant time between grids is relatively large, and about half of the grids have two or fewer bus stops, and the number of grids containing metro stations is few due to the limited number of metro stations. The average traffic flow speed is about 30 km/h, which is higher than the average peak hourly speed of the city's road network (22.90 km/h) [46]. The mean POI mixture of grids is 0.73, indicating a relatively diverse range of POIs in most grids. Shopping service POIs are the most widely distributed in most grids, while living service and catering service POIs are also widely distributed, and scenic spot and multiple-living building POIs are the fewest.

4. Methodology

4.1. Spatial Autocorrelation.

In this study, we investigated the spatial autocorrelation using the global Moran's I test [47]. The global Moran's I is an overall judgment of the correlations of all spatial units with the following expressions:
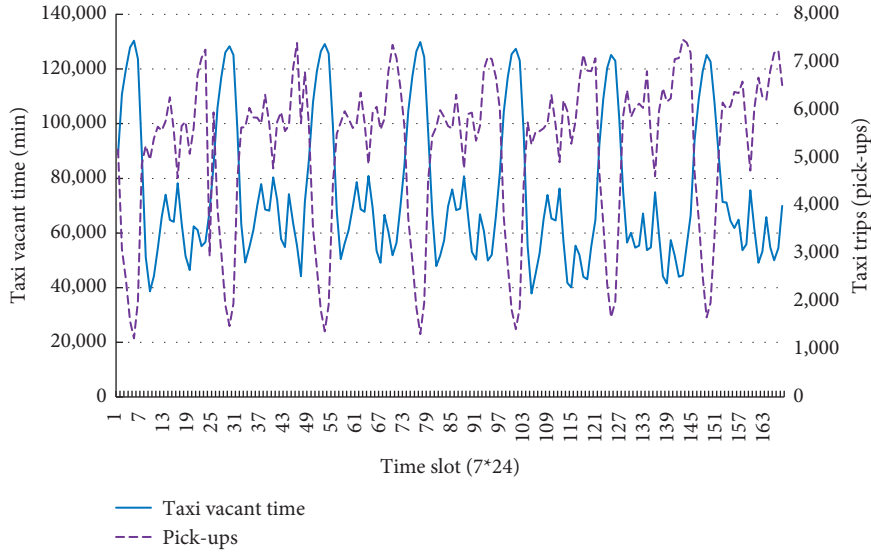
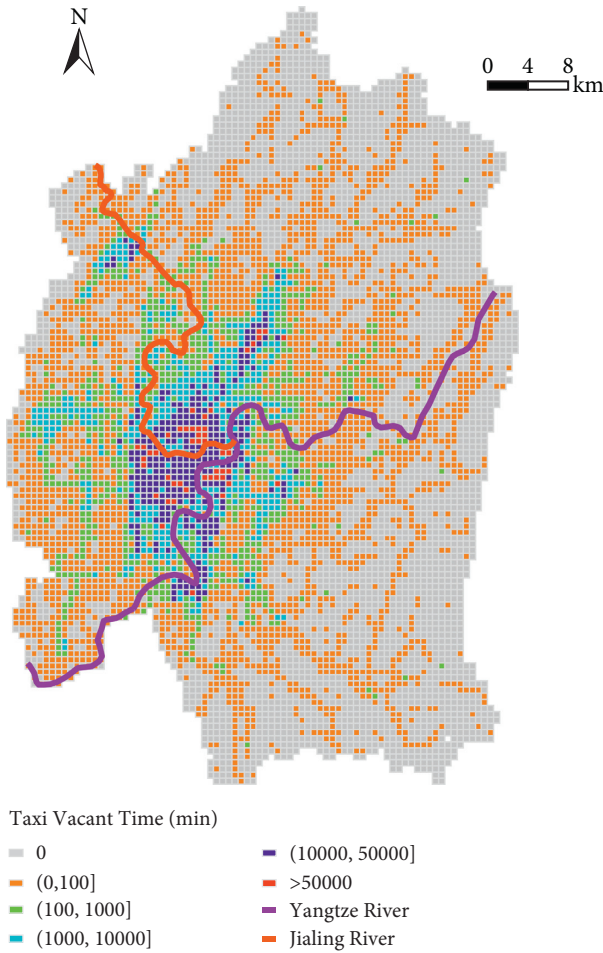FIGURE 2: Total taxi vacant time and total pick-ups per hour for 7 days.

FIGURE 3: Kernel Density Map of total taxi vacant time in grids.

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij} (x_i - \overline{x})(x_j - \overline{x})}{\left(\sum_{i=1}^{n} \sum_{j=1}^{n} \omega_{ij}\right) \sum_{i=1}^{n} (x_i - \overline{x})^2},$$

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i,$$

(5)

where n is the number of samples, ω_{ij} is the (i, j) element of the spatial weight matrix W, x_i and x_j are the observed values of spatial units i and j, respectively, and $\overline{x}$ is the mean of the observed values. Moran's I generally takes a value between [−1, 1]; a value greater than 0 indicates positive spatial autocorrelation (or indicates spatial agglomeration phenomenon), and a value less than 0 indicates negative spatial autocorrelation (or spatial dispersion). 1 indicates complete spatial clustering, −1 indicates complete spatial dispersion, and 0 indicates random distribution in space or no spatial correlation. In this paper, we construct the Queen adjacent (boundary or vertex adjacent) spatial weight matrix based on the grid for calculating Moran's I.

4.2. Generalized Additive Model. A generalized additive model (GAM) was introduced to investigate the mechanism of built environment on taxi vacant time. Initially proposed by Hastie and Tibshirani [48, 49], GAM is an extension of the linear regression model, combining parametric and nonparametric approaches to regression models, and is suitable for dealing with complex nonlinear relationships between explanatory and explained variables [50]. GAM assumes that the function is additive, which not only fits as a parametric function but also allows each explanatory variable to be transformed into an unconstrained smoothing

TABLE 1: Descriptive statistics of the variables.

Variables	Description	Min	Median	Mean	Max	Std. Dev.
TVT	Taxi vacant time of grids (min)	0.25	7.75	53.53	3397.50	120.98
DSR	Road density of grids (km/km^2)	0	6.52	7.09	22.01	4.04
BSN	Number of bus stops in grids	0	2	2.35	16	2.44
MET	Presence of metro stations within grids	0	0	0.12	1	0.33
TFS	Average traffic flow speed (km/h)	0	25.42	31.36	120.00	22.58
LON	Longitude of the center of grids	106.30	106.50	106.50	107.00	0.11
LAT	Latitude of the center of grids	29.13	29.58	29.58	30.10	0.12
PMT	POI mixture	0	0.85	0.73	1.00	0.32
CAT	Number of catering service POIs	0	2	17.78	498	43.38
SNS	Number of scenic spot POIs	0	0	1.20	57	3.58
ENT	Number of enterprise POIs	0	3	7.71	223	16.65
SHP	Number of shopping service POIs	0	5	37.16	1368	102.19
FAI	Number of finance and insurance service POIs	0	0	3.14	111	7.94
SAE	Number of scientific, culture, and education POIs	0	2	7.97	149	16.18
MVS	Number of motor vehicle service POIs	0	1	3.78	319	12.21
LFS	Number of living service POIs	0	3	18.88	527	42.10
SLE	Number of sports and leisure service POIs	0	1	6.92	191	15.64
HCS	Number of healthcare service POIs	0	1	8.41	158	15.89
GOV	Number of governmental agency POIs	0	2	7.32	181	14.71
HTS	Number of hotel service POIs	0	0	3.89	230	12.68
MLB	Number of multiple-living building POIs	0	0	1.03	43	3.41
RES	Number of residence community POIs	0	2	6.33	88	9.31

function, after which the corresponding explanatory variable is modeled in the form of a smoothing function. The form of GAM is as follows:

$$g(\mu) = \alpha + \sum_{i=1}^{p} f_i(X_i). \tag{6}$$

The hypothesis of the relationship between the explanatory and explained variables in the GAM is

$$g(\mu) = s_0 + s_1(x_1) + s_2(x_2) + \cdots + s_m(x_m), \tag{7}$$

where μ is the expected value of Y; that is, $\mu = E(Y/X_1, \ldots, X_p)$, $g(\cdot)$ is the connection function, and α is the intercept. $f_i(\cdot)$ is an arbitrary univariate function for each explanatory variable X_i. GAM is highly flexible because it does not require the specification of the form of the explanatory variables. GAM is applicable to many distribution types, and the form of the connection function $g(\cdot)$ varies for different distribution types.

5. Results and Analyses

5.1. Moran's I of Taxi Vacant Time. Moran's I of the grid vacant time was calculated as 0.328, and the significance test showed that the p value $<2.2e-16$ was significant at 0.001 significance level, indicating the existence of spatial autocorrelation of the grid vacant time. The scatter plot of grid Moran's I is shown in Figure 4, from which it can be seen that most of the grids are located in the third quadrant representing "low-low" correlation, followed by those located in the first quadrant representing "high-high" correlation. There are relatively few grids in the second quadrant representing "low-high" correlation and the fourth quadrant representing "high-low" correlation. The high impact values (square symbols in the figure) are mainly distributed in the
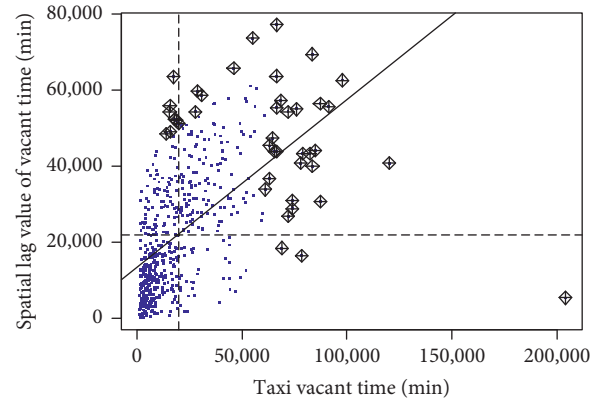

FIGURE 4: Moran's I scatter plot of grids. Note: the square symbols represent high impact values.

first quadrant, indicating a more pronounced aggregation of high vacant time grids.

5.2. Transformation of the Form of Explained Variable. From the results of the statistical description, the explained variable is identified as a nonnegative and nonzero variable. Thus, to examine the form of their distribution, the Kernel Density Estimation plot and the Empirical Cumulative Distribution Function plot are provided as shown in Figure 5.

It can be figured out that the distribution of taxi vacant time of grids is close to a power-law distribution and there are some outliers with large bias. If the variables containing outliers are applied directly in the model and the model is sensitive to the outliers, it will bias the model estimation. Thus, to solve this problem, the variables should be transformed. Combining the range of values and distribution of
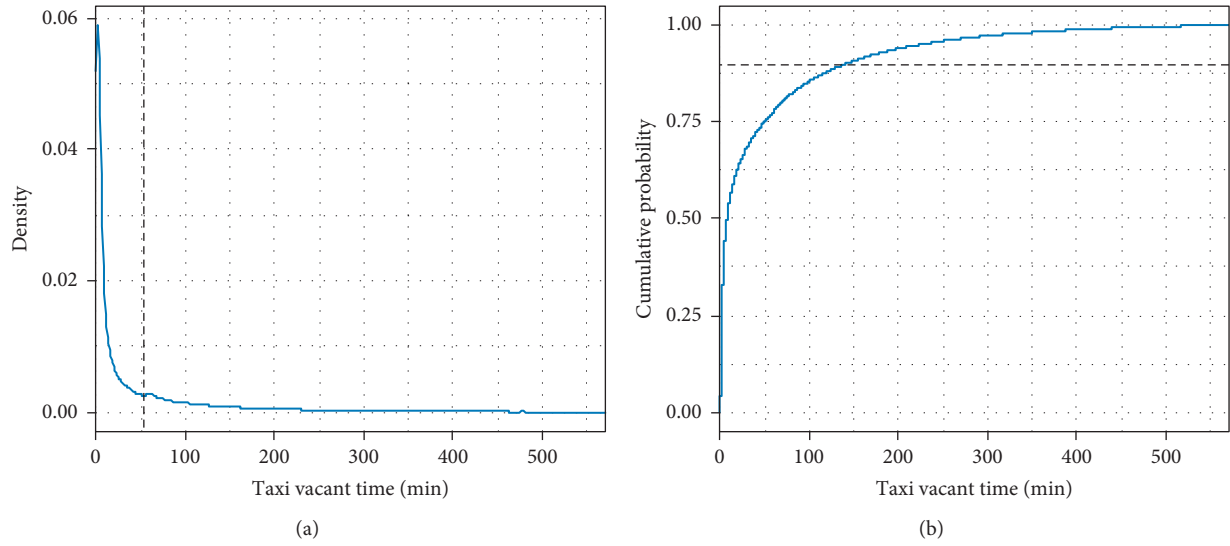
(a)

(b)

Figure 5: Distribution of taxi vacant time of grids.(a) Kernel Density Estimation. (b) Empirical Cumulative Distribution Function. Note: some outliers (top 1%) were removed to increase the readability of the plot; the vertical dashed line in Figure 5(a) is the average of the grid taxi vacant time, and the horizontal dashed line in Figure 5(b) is the dividing line of the cumulative frequency at 90%.

taxi vacant time of grids, a natural logarithmic transformation of grid vacant time is more suitable. First, the logarithmic transformation will make the range of values of vacant time much narrower, which makes the estimated value less sensitive to the presence of outliers; second, the conditional distribution of variables with vacant time for values greater than 0, which are usually strictly positive, with heteroskedasticity and skewness, but taking the logarithm will effectively improve both of them; furthermore, when the explained variable is greater than 0, models using logarithmic values of the original values are generally closer to the CML (Classical Linear Model) assumptions than models using the original values directly. Therefore, in this paper, the explained variable is converted into the form of natural logarithms; that is, the explained variable is in the form ln (TVT).

5.3. GAM Results.

Table 2 presents the estimation results of GAM model. Since both traffic demand and supply have significant time-varying characteristics, two variables reflecting time-varying characteristics, namely, day of the week (DOW, 1~7) and hour of the day (TSLOT, 0~23), are included in the model. All explanatory variables except for multiple-living building (MLB), scientific, culture, and education (SAE), and scenic spot (SNS) are significant at the 0.001 significance level. In a GAM model, the degree of nonlinearity could be measured by the effective degree of freedom (EDF). The fitted explanatory variables of the smooth spline function and smooth tensor product have significant coefficients and their corresponding estimated EDFs are high. All EDFs are greater than 8, suggesting that

Table 2: Estimation results of GAM model.

Explanatory variables	Coefficients			
	Estimate	Std. Error	t value	Pr $(>|t|)$
Intercept	2.0860	0.0083	251.450	$<2e-16$***
ENT	−0.0027	0.0002	−10.943	$<2e-16$***
FAI	0.0027	0.0007	3.751	0.0002***
GOV	−0.0013	0.0003	−4.339	$1.43e-05$***
HTS	0.0043	0.0004	10.734	$<2e-16$***
LFS	0.0018	0.0003	6.787	$1.14e-11$***
HCS	0.0034	0.0004	8.486	$<2e-16$***
MLB	−0.0012	0.0012	−0.955	0.3395
SAE	0.0001	0.0003	0.400	0.6893
SHP	−0.0002	0.0001	−3.761	0.0002***
SLE	0.0099	0.0005	20.643	$<2e-16$***
SNS	−0.0007	0.0007	−1.019	0.3081
PMT	−0.0365	0.0081	−4.487	$7.23e-06$***
MET	0.3239	0.0075	42.915	$<2e-16$***
Explanatory variables	Approximate significance of smoothing terms			
	EDF	Ref. Df	F	p value
s (TFS)	8.929	8.999	5393.5	$<2e-16$***
s (DSR)	8.927	8.998	1711.3	$<2e-16$***
s (BSN)	8.774	8.978	391.6	$<2e-16$***
s (CAT)	8.943	8.999	421.6	$<2e-16$***
s (MVS)	8.949	8.999	687.1	$<2e-16$***
s (RES)	8.591	8.931	405.5	$<2e-16$***
s (longitude, latitude)	28.945	29.000	1835.5	$<2e-16$***
t2 (DOW, TSLOT)	21.966	23.236	117.9	$<2e-16$***
Model fit	R-sq.(adj)	0.759		

Significance codes: "***" 0.001. s ( ): smooth spline function. $t2$ ( ): smooth tensor product.

significant and strong nonlinearities exist among the selected variables. The effect of metro stations (MET) on the taxi vacant time in the grid is significant and substantial, with the presence of a metro station in the grid increasing the vacant time by about 32.39%.

6. Discussion

Considering variability in the effect of built environment on the effect of taxi vacant time, the high correlation association and positive effect of metro stations (MET) and taxi vacant time may be due to the fact that most of the metro stations are located in high development intensity areas and densely populated areas of the city, where taxi demand is high [51] and thus taxi activity intensity is high. Also, the demand for transfer between the metro and taxis cannot be ignored, and these two modes of transportation have complementary roles [5]. Completing trips with taxis and metro can take advantage of the respective strengths of both modes of transportation [52]. POI mixture can reduce taxi vacant time in a region, which has two possibilities, either high taxi demand with the high mixture and thus short searching time or low taxi demand with the high mixture, which is not attractive to drivers in a searching status. There are results of land use mix in related literature studies that promote taxi trips [53] and results that reduce taxi trips [36].

Figure 6 shows the estimated degrees of freedom with smooth spline functions on taxi vacant time, in which the fits of each explanatory variable are obvious curve shapes with significant nonlinear relationship characteristics. The effect of catering service (CAT, Figure 6(a)) on vacant time is little and unstable when the number of POIs is small, and the effect on vacant time has alternating positive and negative effects, and its effect on vacant time increases significantly after the number of catering service POIs in the grid reaches about 300. Places with a high concentration of catering services are generally important sources of taxi passengers and have a high demand for taxi rides, especially during peak dining hours and weekends [43], and thus the taxi searching time can be significantly reduced.

The effect of motor vehicle service (MVS, Figure 6(b)) on vacant time also fluctuates repeatedly, but it only has a significant effect on the vacant time of the grid when the number of POIs reaches about 100; that is, the vacant time increases significantly at this time. Motor vehicle services include car sale, maintenance, and beauty, vehicle inspection, and car renting. Except for car sales, other vehicle services are generally scattered, and residents go mainly by self-drive, which has less correlation with taxi demand. Meanwhile, taxi drivers perform vehicle maintenance during their nonoperating hours and do not generate vacant time records.

The effect of residence community (RES, Figure 6(c)) is negative when the number is small, then remains positive, and finally becomes negative again, indicating that residence communities usually increase vacant time, but highly concentrated residential areas decrease it again. Residential communities are the departure points for residents to go to work and the destinations for them to leave work, and they have a large demand for taxis, which tend to go to the areas where residential communities are concentrated to search for passengers. On the other hand, however, the time-varying demand for taxis in residential communities varies greatly, with departures dominating in the morning peak and arrivals in the evening peak. Taxis often have difficulty searching for the next passenger in the vicinity after the evening peak passengers drop off in residential communities [43]. All of the above causes residential communities to positively affect taxi vacant time.

The effect of the number of bus stops (BSN, Figure 6(d)) is characterized by two stages, and the effect is nearly linear when the number of bus stops in the grid is less than 8, and the longer the vacant time is with more bus stops, after which the fitted curve fluctuates and rises, the vacant time increases significantly after the number of bus stops reaches about 13. This phenomenon is easy to understand; with a specific competitive relationship between public transport and taxis, good public transport access can significantly reduce the demand for taxis [54], as well as the increased difficulty in searching for passengers, thus leading to an increase in taxi vacant time.

The impact effect of road density (DSR, Figure 6(e)) is characterized by three stages, with two segments showing a near-linear state with a large slope and one segment showing a near-horizontal line, which is close to linear at about $10 \, \text{km/km}^2$ or less, and near-horizontal linear at about 10~17, after which the impact effect increases significantly with increasing road density. Related studies also show the uncertainty of the impact of road density on taxi trips, with some studies showing that road density promotes taxi trips [36] and other studies showing that it reduces taxi trips [54]. Since the urban road network is largely well developed, roads may no longer be a constraint on taxi demand, but high-density road network areas are usually densely populated and their taxi demand is high [36], which in turn significantly boosts regional taxi vacant time.

The average traffic flow speed (TFS, Figure 6(f)) increases taxi vacant time at low speeds and decreases it at high speeds, with a turning point around 30 km/h. The time spent by taxis, whether vacant or occupied, in areas where they can move quickly is usually reduced. Related studies also show that severe traffic congestion increases taxi vacant time [5].

The plots of the fitted taxi vacant time with temporal (day of the week, hourly slots) and spatial (longitude, latitude) interactions are shown in Figure 7. From the daily fluctuation of vacant time (Figure 7(a)), the hourly vacant time of a day has a significant change, with more vacant time in the early morning, more vacant time in the midday period compared to other hours of the day, and a continuous decrease in vacant time from the evening peak hour until late at night. As mentioned above, the total taxi vacant time is negatively correlated with the demand for taxis, with an undersupply of taxis during peak travel periods, a much shorter passenger search time, and low levels of vacant time. The opposite is true during travel valleys. The weekly fluctuations in vacant time show relatively small fluctuations from Monday to Sunday, with two relatively significant dips in the early morning hours on Monday and late afternoon on
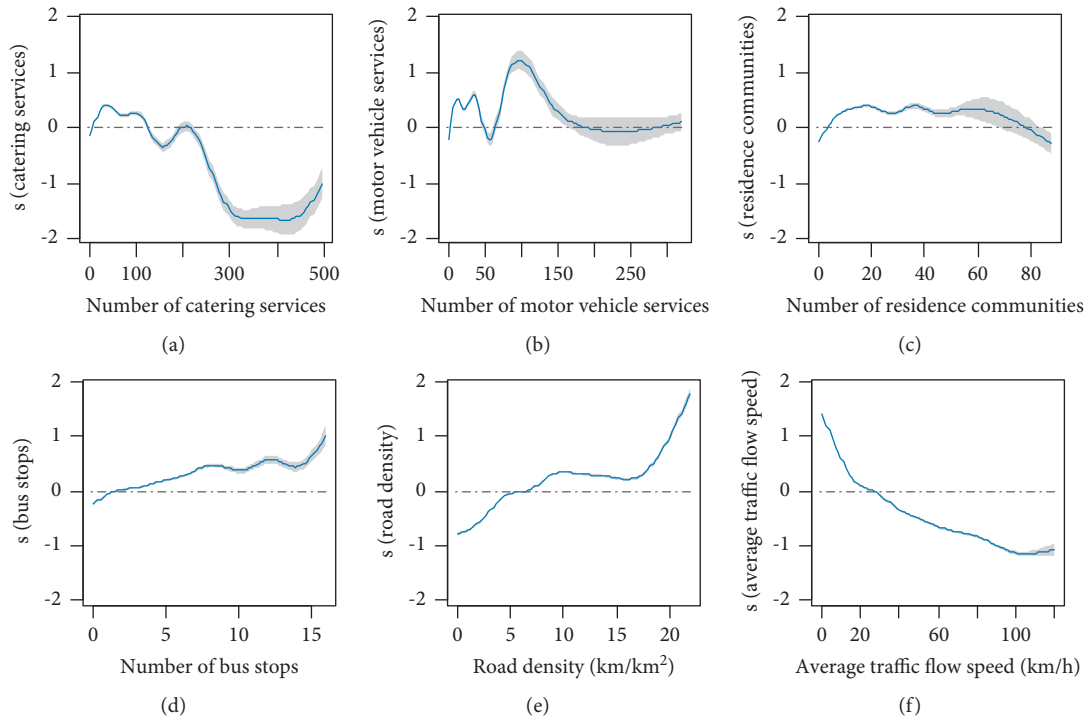
FIGURE 6: Estimated degrees of freedom with smooth spline functions on taxi vacant time for explanatory variables. (a) Number of catering services. (b) Number of motor vehicle services. (c) Number of residence communities. (d) Number of bus stops. (e)Road density (km/km^2). (f) Average traffic flow speed (km/h). Note: the horizontal coordinate indicates the actual value of each explanatory variable, the vertical axis indicates the value of the smooth function, the shading indicates the upper and lower limits of the confidence interval, and the solid line indicates the smooth fitted curve of each explanatory variable.
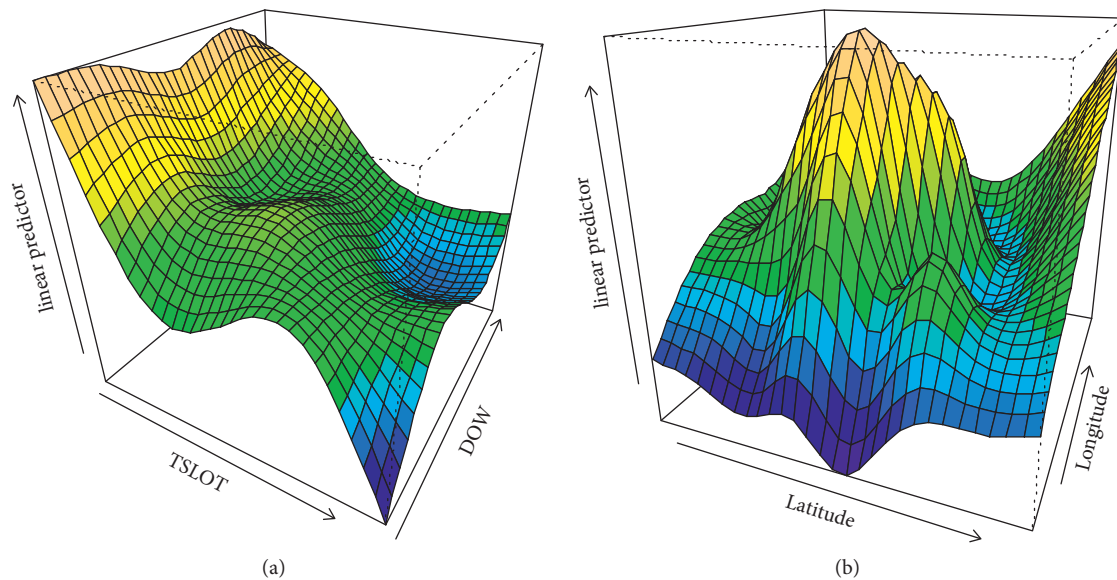

FIGURE 7: Visualization of fitted taxi vacant time with temporal and spatial interactions. (a) Day of the week and hourly slots (DOW, TSLOT). (b) Longitude and latitude (longitude, latitude).

Friday, indicating a brief decrease in vacant time during these two periods. Monday morning and Friday evenfall are the beginning and end of the workweek, respectively, and most employees are likely to commute on these two days, while the other workdays are not so concentrated.

The longitude-latitude interaction plot (Figure 7(b)) is now clearly highlighted in the middle area, which is the downtown of the city, and it is the main operating area for taxis; thus its vacant time is significantly higher than other areas. The vacant time in high taxi demand areas is not

significantly reduced, which is mainly due to the imbalance in the spatial and temporal distribution of taxi demand. Also the traffic congestion in downtown increases the taxi travel time [5].

Studies have shown that the number of taxis can be significantly reduced through operational efficiency improvement, even without changing the current operating mode and ridesharing [55]. The findings of this paper can guide urban transportation facility layout planning, taxi management policy development, and taxi operation optimization to improve taxi operation efficiency. In terms of transportation facility layout planning, a differentiated taxi stand layout should be implemented to allow for quick docking between drivers and passengers and to reduce ineffective cruising of vacant taxis [56]. In downtown areas, high-density all-weather or time-sharing taxi stands are located in the vicinity of commercial complexes, high-rise office buildings, catering service areas, and large hospitals, where taxi demand is high, while taxis are prohibited from picking up or dropping off passengers on the side of the road during peak hours on busy roads. In suburban areas, taxi stands are set up only at locations with high taxi demand. Because of the high correlation between the taxi vacant time and metro stations, the layout of taxi stands should be considered to connect with metro stations to facilitate transfers.

In terms of taxi management policy formulation and taxi operation optimization, taxi operation guiding measures and technical means should be developed to optimize operation. As mentioned above, there is a sharp increase in the vacant time of taxis during the low demand period, while there is a spatial mismatch in the supply and demand of taxis during the peak hours. Therefore, the spatial and temporal regulation of taxi capacity resources should be carried out through policy measures such as subsidies or rent reductions, and enterprises or drivers should be guided to set reasonable shift change times [57]. Benefiting from the development of information technology, increasing the ratio of ridesharing and applying intelligent and accurate driver-rider matching technology can significantly reduce the taxi vacant time as long as the policy allows [58]. From the perspective of the comprehensive transportation system, improving the level of public transportation services and enhancing the attractiveness of public transportation can not only reduce the use of taxis but also have positive significance in reducing congestion, energy saving, and emission reduction. Meanwhile, the case study city in this paper is similar to other large cities with diverse urban transportation modes and large populations, and thus the findings of this paper may provide guidance for some cities.

7. Conclusions

Unlike previous studies related to taxi demand, this paper examines the taxi travel issue from taxi supply-side perspective. The study is based on the time that taxis are available to provide service that can be reliably measured. Based on the analysis of the spatial and temporal distribution characteristics of taxi vacant time, this paper constructs a GAM model to study the influence mechanism of built environment on taxi vacant time. The distribution of taxi vacant time shows a strong spatial and temporal imbalance. Taxi vacant time is highly negatively correlated with demand (pick-ups), and sufficient passenger sources enable drivers to shorten the average passenger searching time. Taxi demand and vacant time are much higher in the downtown areas than in the suburban areas, where high land development intensity generates high taxi demand and thus attracts a large number of taxis.

The effect of built environment on taxi vacant time varies, and the effect of built environment on taxi vacant time also shows significant nonlinear characteristics. The positive effect of metro stations on vacant taxi time is the most prominent, and, in addition to considering that metro stations are located in areas that usually have high traffic demand and supply, the transfer between metro trips and taxi trips is also an important factor [52]. The degree of diversity of POI functions in the region can reduce the taxi vacant time to a certain extent. The effects of sports and leisure services, hotel services, healthcare services, finance and insurance services, and living services on vacant taxi time are relatively high and positive. The effects of enterprises and governmental agencies on vacant taxi time are relatively low and negative. The effect of catering services on vacant time is only significant when the number is high, and the positive low effect of residence communities is dominant. Road density has a negative effect on vacant time at low densities, but its positive effect increases rapidly at high road density intervals. In general, an increase in the number of bus stops in the area results in a corresponding increase in taxi vacant time. Consistent with expectations, congestion increases taxi vacant time and smooth flow significantly decreases vacant time [5].

The preference of taxi drivers for the built environment is subjective and regular, which will lead to spatial imbalance of taxi supply and demand in the case of insufficient information. The spatial and temporal distribution of taxi capacity resources can be optimized through technology and management tools, thus improving the overall taxi operating efficiency [59]. Short-term taxi demand forecasting based on vacant time can be used to guide the rational scheduling of taxi capacity [9]. At the micro level, improving driver-passenger matching through technical means can improve driver operating efficiency and thus reduce vacant time [60]. The nonlinear effect of built environment factors on taxi vacant time is significant and the mechanism of influence varies in different built environments, which have positive implications for urban functional layout planning. The close correlation between taxi vacant time and transportation facilities such as metro stations, bus stops, and road density can provide guidance for comprehensive urban transportation planning and optimization.

Although some research has been done in this paper on the mechanism of built environment influence on taxi vacant time, there are still shortcomings in this paper due to the limitation of data and methods. First, there may be differences in the mechanism of the influence of built environment on vacant time between downtown and suburban

areas. Spatial heterogeneity was not considered in this study, and spatial econometric methods could be introduced in the future to study the spatial heterogeneity of factors influencing vacant time. Second, although this study constructs a rich set of built environment indicators, it lacks socioeconomically relevant indicators, and taxi demand is associated with resident income, education, car ownership, and so forth [61]. Third, the grid division of the study area may lead to the division of areas that are closely connected spatially into different grids, which may lead to potential measurement errors. A more reasonable method of dividing study units will be explored to maintain the continuity of spatial study units to reflect the spatial characteristics of regions and reduce the influence of measurement errors on study results.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] W. Y. Szeto, R. C. P. Wong, and W. H. Yang, "Guiding vacant taxi drivers to demand locations by taxi-calling signals: a sequential binary logistic regression modeling approach and policy implications," *Transport Policy*, vol. 76, pp. 100–110, 2019.

[2] F. Chen, Z. Yin, Y. Ye, and D. Sun, "Taxi hailing choice behavior and economic benefit analysis of emission reduction based on multi-mode travel big data," *Transport Policy*, vol. 97, pp. 73–84, 2020.

[3] H. Yang and S. C. Wong, "A network model of urban taxi services," *Transportation Research Part B: Methodological*, vol. 32, no. 4, pp. 235–246, 1998.

[4] D. Q. Nguyen-Phuoc, P. T. K. Tran, D. N. Su, O. Oviedo-Trespalacios, and L. W. Johnson, "The formation of passenger loyalty: differences between ride-hailing and traditional taxi services," *Travel Behaviour and Society*, vol. 24, pp. 218–230, 2021.

[5] W. Zhang, S. V. Ukkusuri, and J. J. Lu, "Impacts of urban built environment on empty taxi trips using limited geolocation data," *Transportation*, vol. 44, no. 6, pp. 1445–1473, 2017.

[6] X. Zhan, X. Qian, and S. V. Ukkusuri, "A graph-based approach to measuring the efficiency of an urban taxi service system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 9, pp. 2479–2489, 2016.

[7] F. Zong, T. Wu, and H. Jia, "Taxi drivers' cruising patterns-insights from taxi GPS traces," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 2, pp. 571–582, 2019.

[8] L. Tang, F. Sun, Z. Kan, C. Ren, and L. Cheng, "Uncovering distribution patterns of high performance taxis from big trace data," *ISPRS International Journal of Geo-Information*, vol. 6, no. 5, Article ID 134, 2017.

[9] J. Ke, H. Zheng, H. Yang, and X. Chen, "Short-term forecasting of passenger demand under on-demand ride services: a spatio-temporal deep learning approach," *Transportation Research Part C: Emerging Technologies*, vol. 85, pp. 591–608, 2017.

[10] S. L. Handy, M. G. Boarnet, R. Ewing, and R. E. Killingsworth, "How the built environment affects physical activity: views from urban planning," *American Journal of Preventive Medicine*, vol. 23, no. 2, pp. 64–73, 2002.

[11] X. Mou, F. Cai, X. Zhang, J. Chen, and R. Zhu, "Urban function identification based on POI and taxi trajectory data," in *Proceedings of the ICBDR 2019: Proceedings of the 2019 Third International Conference on Big Data Research*, pp. 152–156, Association for Computing Machinery, Cergy-Pontoise, France, November 2019.

[12] S. Wang, D. Yu, X. Ma, and X. Xing, "Analyzing urban traffic demand distribution and the correlation between traffic flow and the built environment based on detector data and POIs," *European Transport Research Review*, vol. 10, Article ID 50, 2018.

[13] P. Zhao, Y. Xu, X. Liu, and M. P. Kwan, "Space-time dynamics of cab drivers' stay behaviors and their relationships with built environment characteristics," *Cities*, vol. 101, Article ID 102689, 2020.

[14] W. K. Lee and S. Y. Sohn, "Taxi vacancy duration: a regression analysis," *Transportation Planning and Technology*, vol. 40, no. 7, pp. 771–795, 2017.

[15] Y. Zhang and J. Sun, "Taxi income monitoring and service strategy analyzing using GPS trajectory data," in *Proceedings of the Nineteenth COTA International Conference of Transportation Professionals*, pp. 6057–6069, ASCE, Nanjing, China, July 2019.

[16] B. Li, D. Zhang, L. Sun et al., "Hunting or waiting? Discovering passenger-finding strategies from a large-scale real-world taxi dataset," in *Proceedings of the 2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pp. 63–68, IEEE, Seattle, WA, USA, March 2011.

[17] X. Hu, S. Gao, Y.-C. Chiu, and D.-Y. Lin, "Modeling routing behavior for vacant taxicabs in urban traffic networks," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2284, no. 1, pp. 81–88, 2012.

[18] J. Tang, Y. Wang, W. Hao, F. Liu, H. Huang, and Y. Wang, "A mixed path size logit-based taxi customer-search model considering spatio-temporal factors in route choice," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 4, pp. 1347–1358, 2020.

[19] Z. Guo, M. Hao, and B. Yu, "Modeling taxi customer searching behavior using high-resolution GPS data," in *Proceedings of the 2020 IEEE Twenty Third International Conference On Intelligent Transportation Systems (ITSC)*, pp. 1–6, IEEE, Rhodes, Greece, September 2020.

[20] Q. Tang, X. Hu, and H. Qi, "Modeling routing behavior learning process for vacant taxis in a congested urban traffic network," *Journal of Transportation Engineering, Part A: Systems*, vol. 146, no. 6, Article ID 4020043, 2020.

[21] Z. Wang, W.-H. Lin, and W. Xu, "A data driven approach to assessing the reliability of using taxicab as probes for real-time route selections," *Journal of Intelligent Transportation Systems*, vol. 25, no. 4, pp. 331–342, 2021.

[22] Y. Huang, D. Sun, and L.-H. Zhang, "Effects of congestion on drivers' speed choice: assessing the mediating role of state

aggressiveness based on taxi floating car data," *Accident Analysis & Prevention*, vol. 117, pp. 318–327, 2018.

[23] G. Qin, T. Li, B. Yu, Y. Wang, Z. Huang, and J. Sun, "Mining factors affecting taxi drivers' incomes using GPS trajectories," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 103–118, 2017.

[24] D. Zhang, L. Sun, B. Li et al., "Understanding taxi service strategies from taxi GPS traces," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 123–135, 2015.

[25] M. G. Demissie, L. Kattan, S. Phithakkitnukoon, G. Homem de Almeida Correia, M. Veloso, and C. Bento, "Modeling location choice of taxi drivers for passenger pickup using GPS data," *IEEE Intelligent Transportation Systems Magazine*, vol. 13, no. 1, pp. 70–90, 2021.

[26] Z. Gan, M. Yang, T. Feng, and H. J. P. Timmermans, "Examining the relationship between built environment and metro ridership at station-to-station level," *Transportation Research Part D: Transport and Environment*, vol. 82, Article ID 102332, 2020.

[27] X. Huang, X. Cao, J. Yin, and X. Cao, "Can metro transit reduce driving? Evidence from Xi'an, China," *Transport Policy*, vol. 81, pp. 350–359, 2019.

[28] S. Li and P. Zhao, "Exploring car ownership and car use in neighborhoods near metro stations in Beijing: does the neighborhood built environment matter?" *Transportation Research Part D: Transport and Environment*, vol. 56, pp. 1–17, 2017.

[29] X. Cao, P. Næss, and F. Wolday, "Examining the effects of the built environment on auto ownership in two Norwegian urban regions," *Transportation Research Part D: Transport and Environment*, vol. 67, pp. 464–474, 2019.

[30] J. E. Schoner, J. Cao, and D. M. Levinson, "Catalysts and magnets: built environment and bicycle commuting," *Journal of Transport Geography*, vol. 47, pp. 100–108, 2015.

[31] X. Cao and W. Yang, "Examining the effects of the built environment and residential self-selection on commuting trips and the related CO 2 emissions: an empirical study in Guangzhou, China," *Transportation Research Part D: Transport and Environment*, vol. 52, pp. 480–494, 2017.

[32] C. Ding, X. Cao, P. Næss, and P. Næss, "Applying gradient boosting decision trees to examine non-linear effects of the built environment on driving distance in Oslo," *Transportation Research Part A: Policy and Practice*, vol. 110, pp. 107–117, 2018.

[33] B. Sun, A. Ermagun, and B. Dan, "Built environmental impacts on commuting mode choice and distance: evidence from Shanghai," *Transportation Research Part D: Transport and Environment*, vol. 52, pp. 441–453, 2017.

[34] Y. Jiang and X. Yan, "Discovering the relationship between travel behavior and land use: a case study of Beijing, China," in *Proceedings of the 2019 Fourth International Conference on Electromechanical Control Technology and Transportation (ICECTT)*, pp. 219–224, IEEE, Guilin, China, April 2019.

[35] C. Yuan, Y. Duan, X. Mao, N. Ma, and J. Zhao, "Impact of the mixed degree of urban functions on the taxi travel demand," *PLoS One*, vol. 16, no. 3, Article ID e0247431, 2021.

[36] X. Liu, L. Sun, Q. Sun, and G. Gao, "Spatial variation of taxi demand using GPS trajectories and POI data," *Journal of Advanced Transportation*, vol. 2020, Article ID 7621576, 20 pages, 2020.

[37] C. Chen, T. Feng, C. Ding, B. Yu, and B. Yao, "Examining the spatial-temporal relationship between urban built environment and taxi ridership: results of a semi-parametric GWPR

model," *Journal of Transport Geography*, vol. 96, Article ID 103172, 2021.

[38] D. Sun and X. Ding, "Spatiotemporal evolution of ridesourcing markets under the new restriction policy: a case study in Shanghai," *Transportation Research Part A: Policy and Practice*, vol. 130, pp. 227–239, 2019.

[39] S. Sabouri, K. Park, A. Smith, G. Tian, and R. Ewing, "Exploring the influence of built environment on Uber demand," *Transportation Research Part D: Transport and Environment*, vol. 81, Article ID 102296, 2020.

[40] H. Yu and Z.-R. Peng, "Exploring the spatial variation of ridesourcing demand and its relationship to built environment and socioeconomic factors with the geographically weighted Poisson regression," *Journal of Transport Geography*, vol. 75, pp. 147–163, 2019.

[41] X. Qian and S. V. Ukkusuri, "Taxi market equilibrium with third-party hailing service," *Transportation Research Part B: Methodological*, vol. 100, pp. 43–63, 2017.

[42] D. Sun, K. Zhang, and S. Shen, "Analyzing spatiotemporal traffic line source emissions based on massive didi online carhailing service data," *Transportation Research Part D: Transport and Environment*, vol. 62, pp. 699–714, 2018.

[43] B. Li, Z. Cai, L. Jiang, S. Su, and X. Huang, "Exploring urban taxi ridership and local associated factors using GPS data and geographically weighted regression," *Cities*, vol. 87, pp. 68–86, 2019.

[44] R. Cervero, "Mixed land-uses and commuting: evidence from the American housing survey," *Transportation Research Part A: Policy and Practice*, vol. 30, no. 5, pp. 361–377, 1996.

[45] S. Hu, P. Chen, H. Lin, C. Xie, and X. Chen, "Promoting carsharing attractiveness and efficiency: an exploratory analysis," *Transportation Research Part D: Transport and Environment*, vol. 65, pp. 229–243, 2018.

[46] Chongqing Transport Planning Institute, *2019 Chongqing Urban Traffic Annual Report*, Chongqing Transport Planning Institute, Chongqing, China, 2020.

[47] P. A. Moran, "Notes on continuous stochastic phenomena," *Biometrika*, vol. 37, no. 1-2, pp. 17–23, 1950.

[48] H. Schwenk and Y. Bengio, "Boosting neural networks," *Neural Computation*, vol. 12, no. 8, pp. 1869–1887, 2000.

[49] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: a statistical view of boosting," *Annals of Statistics*, vol. 28, no. 2, pp. 337–374, 2000.

[50] S. Li, L. Zhai, B. Zou, H. Sang, and X. Fang, "A generalized additive model combining principal component analysis for $PM_{2.5}$ concentration estimation," *ISPRS International Journal of Geo-Information*, vol. 6, no. 8, Article ID 248, 2017.

[51] Q. Liu, C. Ding, and P. Chen, "A panel analysis of the effect of the urban environment on the spatiotemporal pattern of taxi demand," *Travel Behaviour and Society*, vol. 18, pp. 29–36, 2020.

[52] Y. Ni and J. Chen, "Exploring the effects of the built environment on two transfer modes for metros: dockless bike sharing and taxis," *Sustainability*, vol. 12, no. 5, Article ID 2034, 2020.

[53] M. Jina and Z. Lei, "The relationship between for-hire service pickups and built environment characteristics: evidence from New York City," in *Proceedings of the International Conference on Transportation and Development 2018: Connected and Autonomous Vehicles and Transportation Safety*, pp. 1–12, ASCE, Pittsburgh, PA, USA, July 2018.

[54] X. Zhang, B. Huang, and S. Zhu, "Spatiotemporal influence of urban environment on taxi ridership using geographically and temporally weighted regression," *ISPRS International Journal of Geo-Information*, vol. 8, no. 1, Article ID 23, 2019.

[55] M. M. Vazifeh, P. Santi, G. Resta, S. H. Strogatz, and C. Ratti, "Addressing the minimum fleet problem in on-demand urban mobility," *Nature*, vol. 557, no. 7706, pp. 534–538, 2018.

[56] R. C. P. Wong, W. Y. Szeto, and S. C. Wong, "Bi-level decisions of vacant taxi drivers traveling towards taxi stands in customer-search: modeling methodology and policy implications," *Transport Policy*, vol. 33, pp. 73–81, 2014.

[57] M. Anil Yazici, C. Kamga, and A. Singhal, "Modeling taxi drivers' decisions for improving airport ground access: john F. Kennedy airport case," *Transportation Research Part A: Policy and Practice*, vol. 91, pp. 48–60, 2016.

[58] J. Alonso-Mora, S. Samaranayake, A. Wallar, E. Frazzoli, and D. Rus, "On-demand high-capacity ride-sharing via dynamic trip-vehicle assignment," *Proceedings of the National Academy of Sciences*, vol. 114, no. 3, pp. 462–467, 2017.

[59] M. Lokhandwala and H. Cai, "Dynamic ride sharing using traditional taxis and shared autonomous taxis: a case study of NYC," *Transportation Research Part C: Emerging Technologies*, vol. 97, pp. 45–60, 2018.

[60] X. Qian, D. Kumar, W. Zhang, and S. V. Ukkusuri, "Understanding the operational dynamics of mobility service providers," *ACM Transactions on Spatial Algorithms and Systems*, vol. 6, no. 2, pp. 1–20, 2020.

[61] X. Zhang, B. Huang, and S. Zhu, "Spatiotemporal varying effects of built environment on taxi and ride-hailing ridership in New York City," *ISPRS International Journal of Geo-Information*, vol. 9, no. 8, Article ID 475, 2020.