WILEY | Hindawi

*Research Article*

# Understanding the Modifiable Areal Unit Problem and Identifying Appropriate Spatial Units while Studying the Influence of the Built Environment on the Traffic System State

**Tian Li [iD],[1,2,3] Mengmeng Zhang,[1,2] Haobin Jiang,[3] and Peng Jing[3]**

[1]*School of Transportation and Logistics Engineering, Shandong Jiaotong University, Jinan 250000, China*
[2]*Shandong Key Laboratory of Smart Transportation (preparation), Jinan 250000, China*
[3]*School of Automotive and Traffic Engineering, Jiangsu University, Jiangsu 212013, China*

Correspondence should be addressed to Tian Li; 450589500@qq.com

Spatially aggregated data are prone to the effects of the modifiable areal unit problem (MAUP), which applies to built environments and traffic data. Although various studies have been carried out to explore the impact of built environment factors on traffic systems, few have considered MAUPs, which may result in statistical inconsistency. The purpose of this study is to assess the effects of MAUPs on statistical variables and geographically weighted regression results when evaluating the influence of the built environment on the traffic system state. Fifty sets of spatial configurations were created using the different aggregation criteria. The variance inflation factor and spatial autocorrelation of the variables, as well as the R2 and root mean squared error of the GWR model, were used to assess the MAUP effect. The results show that the index variation is more dependent on the scale of the spatial unit than on zoning type. In the case study presented, based on the available dataset, the optimal spatial unit size for analyzing the influence of the built environment on Jinan's traffic system was $900\,m \times 900\,m$.

## 1. Introduction

Although increasing attention has been paid to the MAUP, it is still neglected in research on built environments (BEs) and traffic system states (TSSs). Some authors discuss BE study areas at the city [8, 9], community [10], or block [11] levels, but these studies generally use scale as a category rather than as a basis for MAUP discussion. Recent studies have investigated the effect of MAUP on the impact of BE on travel behaviors. Zhong et al. [12] investigated the impact of BE on travel behavior on both traffic community and buffer zone scales and found that at a neighborhood scale, the BE cannot explain all resident types' travel behavior effectively. Hong et al. [13] studied the effect of BEs on the vehicle travel distance at a transport analysis zone (TAZ) scale and at a 1 km buffer zone scale and found that the travel distance is more sensitive to the BE variables at the TAZ scale. However, the spatial scale classifications of the above study are too simple, and the SU scale and zoning type design were not carried out; therefore, the corresponding MAUP effect analyses were insufficient, while the conclusions are only applicable to the TAZ and buffer zones. In addition, the above study mainly evaluated the changes in the model parameters while ignoring the MAUP effect on the variables. In addition, the above research focuses on the relationship between BE and travel behavior, and the latter will eventually affect TSS. The influence of the MAUP on the relationship between the BE and TSS has received little attention. Thus, it is necessary to compare the influence of BE on TSS from multiple spatial scales and types to determine the effect of the MAUP. In addition, the wide availability and utilization of transportation and geographic big data provide an opportunity to study the MAUP within the context of the BE-TSS relationship from the perspective of variable variation and model parameter fluctuation.

In this study, we examined how variables and model results change when using SUs of various scales and zoning types, as well as the optimal SU choice.

## 2. The Built Environment

BE refers to various buildings and urban spaces that are different from the natural environment, especially those that can be changed through policies and human behavior. BE contains multidimensional elements. In the field of urban geography and urban planning, Cervero et al. [14] put forward the "3 Ds" element, referring to density, diversity, and design, which is widely applied. Ewing et al. [15] extended the "3 Ds" to "5 Ds," adding "destination accessibility" and "distance to transit". In this paper, we use the "5 Ds" dimensions of BE. The main indicators and calculation methods for each dimension are listed in the Table 1.

## 3. The Traffic System State

The TSS refers to the running and changing states of complex traffic systems. In a narrow sense, TSS is usually expressed through parameters such as the urban road traffic service level, vehicle speed, and road congestion index, with the focus being an accurate representation of traffic efficiency. The generalized TSS includes the state of traffic travel, safety, and comfort, and is a comprehensive representation of the multidimensional state of the traffic system (shown in Table 2). The TSS studied in this study was in a generalized state.

## 4. Methodology

*4.1. Design of SU Scale and Type.* The MAUP effect can be subdivided into scale and zoning effects [18] The scale effect refers to the fact that when the size of the SU changes owing to aggregation, the statistical results of the same dataset will be different (shown in Figure 1). The zoning effect occurs when a change in the spatial element boundaries for the same unit area results in different statistical results (shown in Figure 2).

Commonly used SUs in the study of BE and TSS interactions include the TAZ [19, 20], buffer zone [21], and grid [22]. Considering the need for multiscale design of the SU scale and types of this study, we note that once the traditional TAZ division has been completed, it is difficult to change. In addition, TAZs are too large and, therefore, unsuitable for fine-scale research. Buffers are more focused on "point/line" research and are not suitable for the whole region's "surface" research. Grids can be varied freely in scale and type and are more conducive to data clustering analysis at different scales. Considering the strong operability and wide application of square grids, a square grid was used as the basic SU in this study.

The selection of the SU scale should consider the road length and block area enclosed by the road in the urban built-up area. On February 6, 2016, the Central Committee and the State Council issued "several opinions on further strengthening the management of urban planning, construction,"

which introduced the concept of "narrow roads and dense networks," and proposed that road density should reach $8 \, km/km^2$. Based on this initiative, the resulting road spacing was between 100 m and 200 m. Therefore, in this study, the minimum SU scale selected was 100 m. For the selection of the maximum scale of SU, we referred to the literature on BE and transportation [23–27], where 400 m, 500 m, and 1000 m were used as maximum SU scales. Eventually, the 1000 m SU scale was chosen as the maximum, with 10 subscales ranging from 100 m to 1000 m at 100 m intervals.

To explore the role of SU type in the influence of BE on TSS, five zoning methods with grid aspect ratios of 1 : 1, 1 : 2, 2 : 1, 1 : 3, and 3 : 1 were adopted (shown in Figure 3). Finally, we designed 50 SU types (10 scales multiplied by 5 zoning types).

*4.2. Two Layer MAUP Effect Analysis*

*4.2.1. Analysis of MAUP Effect on Variables.* Spatial autocorrelation analysis can reflect whether the pattern reflected by spatial elements is aggregated, discrete, or random. Spatial autocorrelation analysis is often used as a partition reference in spatial data research. Marcos et al. [6] used the spatial autocorrelation of traffic accident data as the basis for dividing research units. In this study, spatial autocorrelation was also used to test the MAUP effect of variables. The metric of spatial autocorrelation is Moran's I and is calculated as follows:

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

where $i$ is the SU index, $x_i$ and $x_j$ are the attribute values of SUs $i$ and $j$, $\overline{x}$ is the mean value of all SU values, and $W_{ij}$ is the spatial weight matrix. $I > 0$ indicates that the attribute values of all SUs have a positive spatial correlation; that is, as the spatial distribution locations are aggregated, their correlation becomes more significant. $I = 0$ represents random regional distribution with no spatial correlation. $I < 0$ indicates that the attribute values of all regions have a negative correlation in space; that is, with the dispersion of spatial distribution positions, their correlation becomes increasingly significant.

In a linear regression model, multicollinearity refers to a large deviation between the linear regression result and the true value owing to a high linear correlation between variables. The scale of SU and the type of zoning also have an important influence on variable multicollinearity. The variance inflation factor (VIF) is commonly used in multivariate multicollinearity tests, with high values indicating possible collinearity between the explanatory variables. The VIF is calculated as VIF = $(1/(1 - R_j^2))$, where $R_j^2$ is the degree of linear correlation between the variable $x_j$ and the other variables.

*4.2.2. Analysis of MAUP Effect on Model Parameters.* In this section, the variation rules of model parameters under different SU scales and types are studied to explore the

TABLE 1: Main indicators of the 5Ds.

| Variables | Main indicators of previous studies | Indicator variables of this paper | Calculation method |
|---|---|---|---|
| Density | Population density, employment density, building density | Population density | Population per square kilometer |
| Diversity | Land-use mix, job-housing imbalance | Land-use mix [16] | $\text{HHI}_i = \sum_{j=1}^{n} (x_{ij}/x_i)^2$ <br> Where $\text{HHI}_i$ is the herfindahl-hirschman index of grid $i$, $x_i$ is the total POI in grid $i$ and $x_{ij}$ is the total POI of category $j$ in grid $i$ |
| Design | Intersection density, neighborhood, road density | Intersection density | Number of intersections per grid |
| Destination accessibility | Accessibility to jobs, distance to CBD | Distance to CBD | $A_i = \sum_{j=1}^{n} \frac{1}{d_{ij}}$ <br> Where $A_i$ is the destination accessibility of grid $i$, $d_{ij}$ is the distance between grid $i$ and $\text{CBD}_j$ |
| Distance to transit | Bus stop density, distance to transit stop | Bus stop density | Number of bus stops per grid |

TABLE 2: Definition and calculation method of TSS indicators.

| Variables | Definition | Calculation method |
|---|---|---|
| Traffic efficiency | Road speed | $\text{TTI}_i = v_f / 1/N \left( \sum_w^N v_w^i \right)$ <br> Where $\text{TTI}i$ is the travel time index of grid $i$ $v_f$ is the free-flow speed, $v_w^i$ is the actual speed of track point $w$ in grid $i$ |
| Traffic safety | Number of traffic accidents | Total number of incidents per year |
| Traffic comfort | Comfort level while driving | The sum of uncomfortable track points, as defined in the literature [17] |
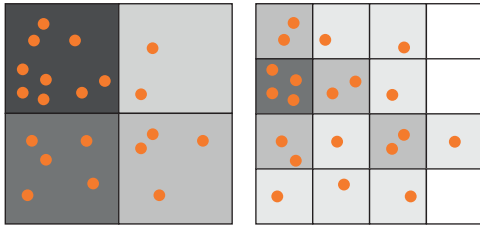


FIGURE 1: Scale effect.
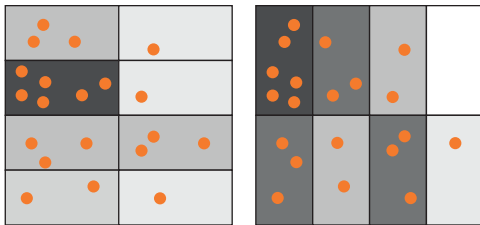


FIGURE 2: Zoning effect.

influence of the MAUP on the effects of BE on the TSS models. Since both BE and TSS data are spatial, the impact of BE on TSS may vary depending on SU, and the traditional ordinary least squares global linear regression model is not capable of explaining spatial heterogeneity. The geographically weighted regression (GWR) model fully considers the spatial characteristics of the research object, and the local regression function can be adjusted according to the spatial variation of the sample. Therefore, in this study, the GWR

model was used as a unified model. The mathematical expression corresponding to the GWR model is shown as follows [28]:

$$y_i = \beta_{0i} + \sum_{j=1}^{p} \beta_{ij} x_{ij} + \varepsilon_i, \qquad (2)$$

where $y_i$ is the $ith$ dependent variable, $x_{ij}$ is the value of the $jth$ variable at the $ith$ grid, $\beta_{ij}$ is the $jth$ regression parameter of grid $i$, and $p$ is the number of independent variables.

GWR models of BE, traffic efficiency, traffic safety, and traffic comfort were constructed and are denoted as GWR-TE, GWR-TS, and GWR-TC, respectively. To evaluate the effect of the GWR model on different SUs, the root mean square error (RMSE) of the model was used, where smaller RMSE values corresponded to a better fit of the GWR model to the observed data, as well as $R^2$, which is the determination coefficient of the GWR.

### 4.3. Multiobjective Selection Model of SU

*4.3.1. Model Construction.* In the study of the influence of BE on TSS, it should be noted that the aggregation values of variables, such as population density, land use diversity, and intersection density, are affected by the MAUP. Simultaneously, the performance of the model is also affected by MAUP, as the parameter estimates and performance indicators of different SUs are different. Therefore, multiobjective functions with both variable -and model result-related objectives should be designed. The goal of a
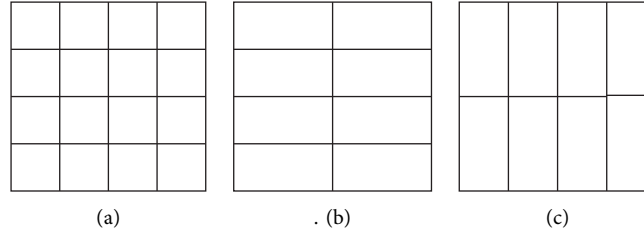
FIGURE 3: Diagram of basic SU type design: (a) $100\,m \times 100\,m$ $(1:1)$. (b) $100\,m \times 200\,m$ $(1:2)$. (c) $200\,m \times 100\,m$ $(2:1)$.

multiobjective selection model is to select an appropriate SU that minimizes the influence of the MAUP on the variables and models. We set four goals as:

Goal 1: The spatial autocorrelation of the variables is minimal to achieve homogeneity in the region and randomness in the interval. The degree of spatial autocorrelation was measured by Moran's index $I$, with smaller values corresponding to less spatial autocorrelation.

Goal 2: The collinearity degree between variables is minimal to ensure that the variables are independent. The degree of collinearity was measured using VIF, with smaller values corresponding to less collinearity.

Goal 3: The model-fitting effect is the best. The fitting effect of the GWR model was measured using $R^2$, with larger values considered more desirable.

Goal 4: The model-fitting performance, measured using the RMSE in the case of the GWR model, is optimal. Smaller RMSEs correspond to a better fit.

A multiobjective function is now established according to the above objectives.

The decision variables are the SU type ($z$).

The objective function is: $\begin{cases} \min I(z) \\ \min V(z) \\ \max R(z) \\ \min RSS(z) \end{cases}$

The constraints are as follows:

$z = 100m * 100m, 100m * 200m, 200m * 100m, \ldots,$ $3000m * 1000mI(z), V(z), R(z),$ and $RSS(z)$ are the mean values of Moran's I, VIF, $R^2$, and RMSE, respectively, for SU type $z$.

In this study, multiple objective functions were combined into a unified function via a weighted grouping method.

$$\min F(z) = w_1 I(z) + w_2 V(z) - w_3 R(z) + w_4 RSS(z), \quad (3)$$

where $w_1, w_2, w_3, w_4$ are the weights of $I(z), V(z), R(z),$ and $RSS(z)$, respectively.

*4.3.2. Model Solution.* To determine the weights while ensuring that (i) they include both subjective and objective factors, (ii) the uncertainty of subjective factors is reduced, and (iii) any discrepancy between the objective weight and the actual degree of importance is minimized, we chose a hybrid weighting method combining the analytic hierarchy process (AHP) and the entropy weight method.

The calculation process of AHP is mainly divided into three steps:

(i) The comparison matrix $B = (a_{ij})_{n \times n}$ is established, where $a_{ij}$ is a value indicating the importance of factor $i$ relative to factor $j$ and $n$ is the number of scoring objects.

(ii) The largest eigenvector of matrix U was found. The maximum eigenvalue of the comparison matrix $\lambda_{\max}$ is then obtained, followed by the corresponding eigenvector $w$ using the equation $Bw = \lambda_{\max} w'$.

(iii) A consistency test was conducted. If the random consistency ratio $C_R < 0.1$, it can be concluded that the consistency of the comparison matrix is strong, where $C_R = CI/RI$, with $CI = /\lambda_{\max} - n/n - 1$, and RI are obtained from Table 3.

The entropy weighting method determines the index weights according to their degree of difference. The main calculation process is as follows.

(i) The variables are normalized and forwarded.

(ii) The information entropy of the $j$-th index is calculated according to the equation as:

$$E_j = -\frac{1}{\ln n} \sum_{i=1}^{m} x'_{ij} \ln x'_{ij}, \quad (4)$$

where $x'_{ij}$ is the value of the $j$-th index of the $i$-th sample after normalization and forward transformation.

(iii) The difference coefficient $g_j$ of the $j$-th index is calculated as the following type.

$$g_j = 1 - E_j. \quad (5)$$

(iv) The weight index $w''_j$ is calculated as the following type.

$$w'_j = \frac{g_j}{\sum_{j=1}^{n} g_j}. \quad (6)$$

Then, the combined weight is calculated as the following type:

TABLE 3: Determination of matrix RI value.

| $n$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| RI | 0 | 0.52 | 0.9 | 1.12 | 1.24 | 1.32 | 1.41 | 1.45 | 1.49 |

$$w_j = 0.7w'_j + 0.3w'_j. \tag{7}$$

## 5. Case Study

In this study, the main urban area of Jinan is taken as the research area (shown in Figure 4), and the specific scope is the area enclosed by the Ji-Guang Expressway, the Jinan Ring Expressway, the South Viaduct of the Second Ring Road, and the Jinan Ring Expressway, representing an area of approximately 438 square kilometers (shown in Figure 5). BE data were obtained from the POI data provided by the AutoNavi API, and road network data were obtained using OpenStreetMap. TSS data were calculated using taxi tracks and accident data shared by the traffic management department. The basic BE and TSS data are shown in Figure 6 to Figure 10, respectively. According to the design scheme of the SU scale and type, the SU ranged from 100 to 1000 m, with the five zoning types being 1 : 1.1 : 2, 1 : 3, 2 : 1, and 3 : 1, resulting in a total of 50 SU types. The 500 $m \times 500$ m area is shown as an example in Figure 5.

## 6. Results and Discussion

*6.1. Effects of MAUP on Moran's I Variables.* Before analyzing the scale and type of SU effects on $R^2$ and RMSE for the application of the GWR model, an investigation of the effect of MAUP on the variables should be performed. The distributional characteristics of the variables for the different SUs are shown in Figure 11. As evident from the figure, the spatial autocorrelation of variables fluctuates with the different SU sampling scales and zoning types and shows relatively consistent regularity as follows:

(i) Larger SU scales result in weaker data differences and stronger spatial autocorrelation, whereas smaller SU scales yield weaker data similarity and spatial autocorrelation.

(ii) The effect of the SU type on spatial autocorrelation was weaker than that of the SU scale.

(iii) The Moran's I values of the variables were all positive. This means that the variables had a positive spatial correlation. The correlation became more significant as the spatial distribution locations were aggregated.

*6.2. Effects of MAUP on VIF of Variables.* The computation of the VIF led to the conclusion that multicollinearity issues were not present, as the VIF factors were all below 10 (shown in Figure 12). However, the VIF values of the independent variables fluctuated as the SU scale and type changed. In particular, the VIF values of 3 : 1 and 1 : 3 subdivisions were
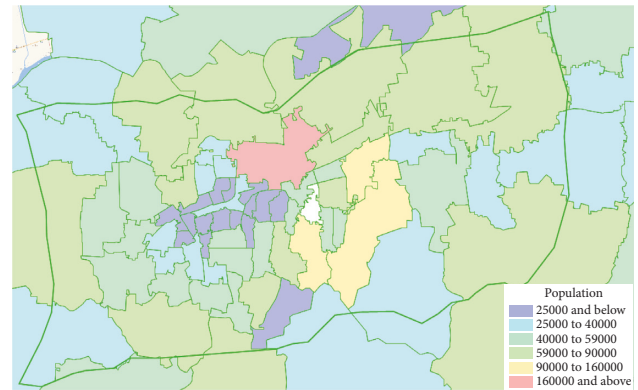


FIGURE 4: Study area.



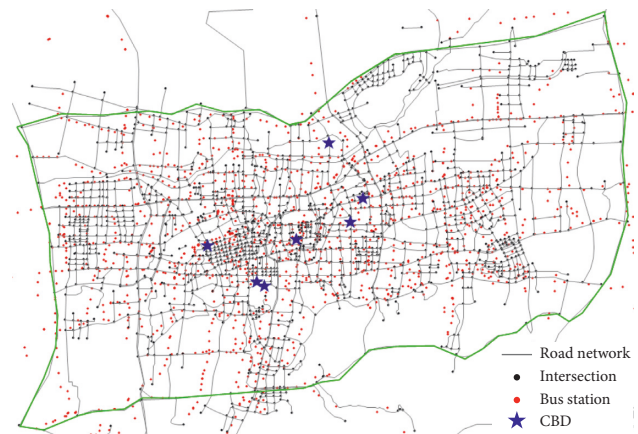FIGURE 5: Population distribution of study area.



FIGURE 6: Intersection, bus station, and CBD of the study area.

generally higher. The VIF value of intersection density was higher than 5 at the SU scale of 1000 m and the 1 : 3 subdivision, showing relative multicollinearity. Therefore, to reduce the multicollinearity of the variables, large SU types and scales should be avoided.
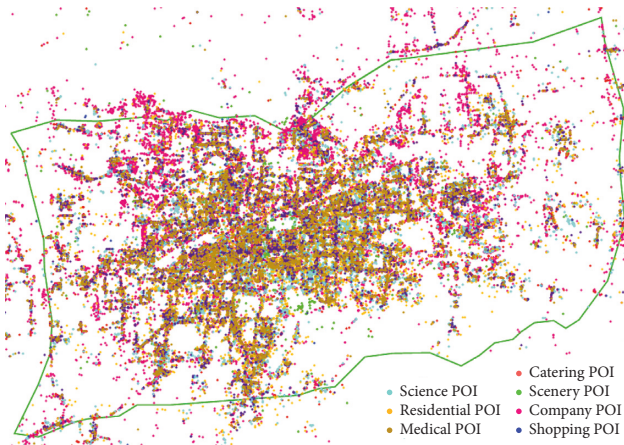
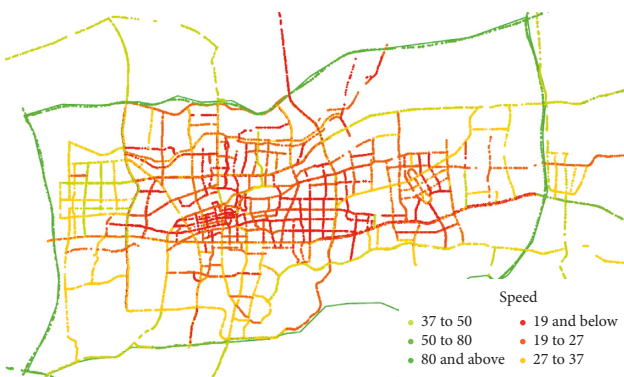FIGURE 7: POI of the study area.



FIGURE 8: POI of the study area.



FIGURE 9: Speed distribution for the study area.

6.3. *Effects of MAUP on $R^2$ and RMSE.* The distributional characteristics of the $R^2$ and RMSE for the GWR model are shown in Figures 13, and 14, respectively. The following characteristics can be observed in the figure.

(i) Under a unified SU scale, radial and zonal zoning had little influence on the correlation coefficient. For example, the $R^2$ values of the $1:3$ and $3:1$ SU types were relatively similar, as were those of the $1:2$ and $2:1$ types.

(ii) Larger SU scales tend to result in fluctuations of the correlation coefficients.

(iii) Larger SU scales caused a decrease in the RMSE of GWR-TE and an increase in the RMSE of GWR-TS and GWR-TC.

These characteristics indicate that when the SU scale and type change, the model results and goodness of fit undergo complex changes.

6.4. *Model Solution.* Nine transportation engineering experts were invited to fill in a judgment matrix questionnaire on the importance of the indicators. The weight index was calculated according to each specific judgment matrix, with that obtained from expert A (Table 4) considered as an example to illustrate the calculation process.

The index calculation results were as follows:
$\lambda_{\max} = 4.2$, $CI = 0.068$, $RI = 0.075 < 0.1$. Based on these values, the judgment matrix exhibited satisfactory consistency.

Similarly, the judgment matrices of the remaining eight experts are calculated to obtain the corresponding index weight values. Finally, the weight values of all the expert scores were averaged and normalized, and the final weights of the four indices were calculated as follows:

$$\begin{aligned} \overline{W}_1 &= 0.21\overline{W}_2 \\ &= 0.17\overline{W}_3 \\ &= 0.27\overline{W}_4 \\ &= 0.35. \end{aligned} \tag{8}$$

The index weights calculated using the entropy weight method were as follows:

$$\begin{aligned} w_1'' &= 0.19w_2'' \\ &= 0.62w_3'' \\ &= 0.13w_4'' \\ &= 0.06. \end{aligned} \tag{9}$$

The weight values of each index were calculated, and the results were as follows:

$$\begin{aligned} w_1 &= 0.21 \\ w_2 &= 0.3 \\ w_3 &= 0.23 \\ w_4 &= 0.26. \end{aligned} \tag{10}$$

The minimum value of the objective function is obtained for $z = 900\,m \times 900$ m.
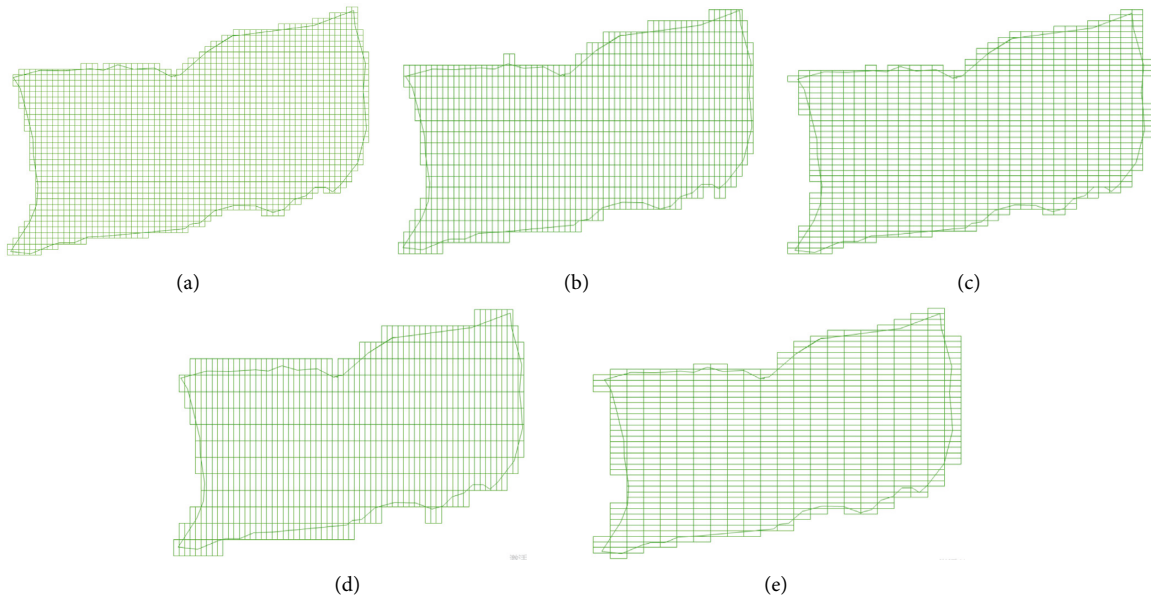
(a)

(b)

(c)

(d)

(e)

FIGURE 10: SU types corresponding for a scale of 500 m. (a) 500 m × 500 m. (b) 500 m × 1000 m. (c) 1000 m × 500 m. (d) 500 m × 1500 m. (e) 1500 m × 500 m.
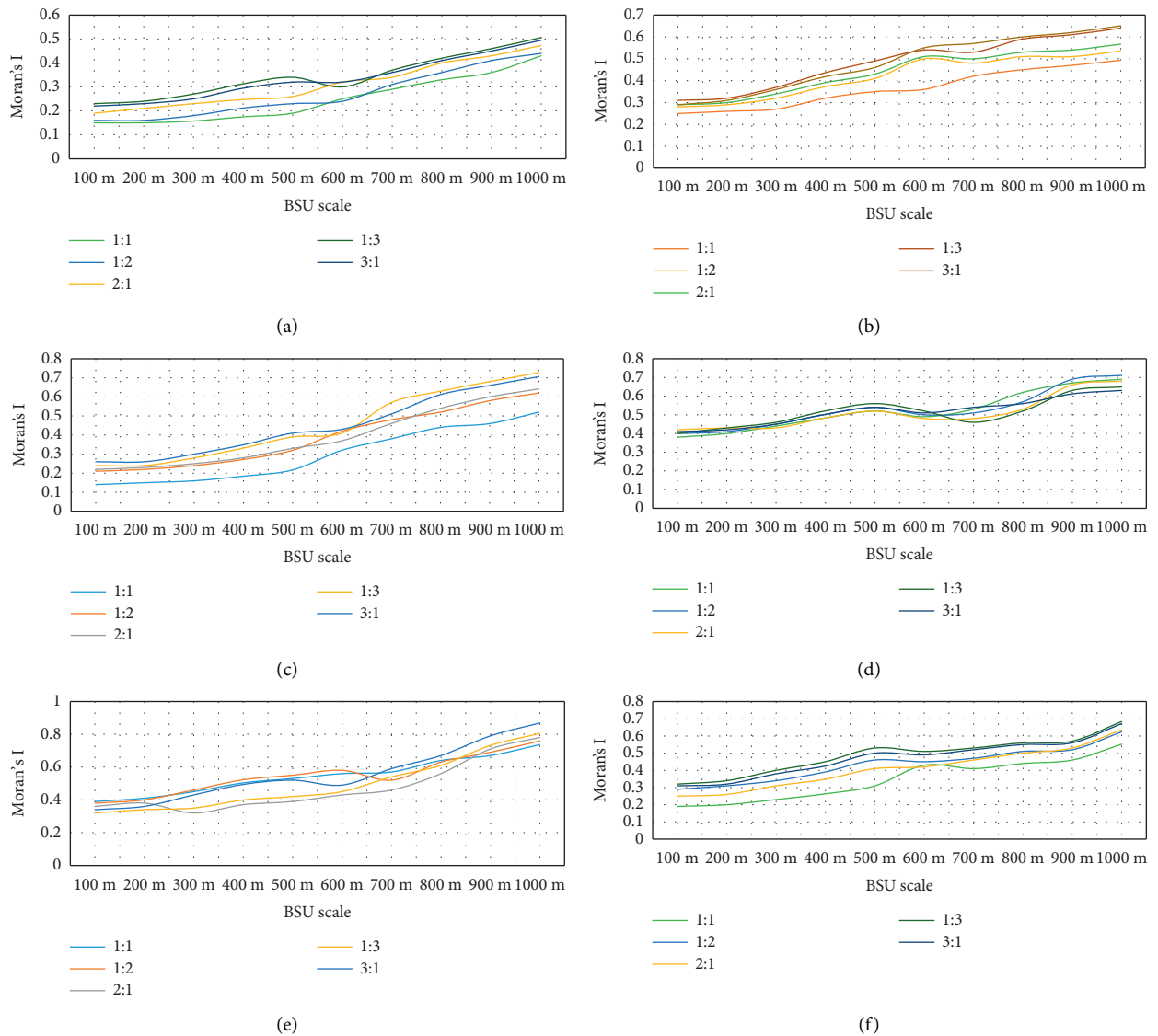


(a)

(b)

(c)

(d)

(e)

(f)

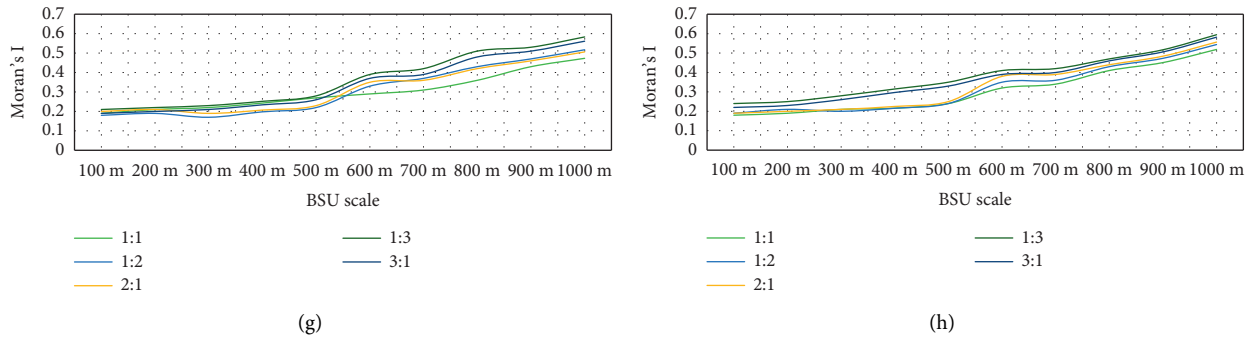FIGURE 11: Continued.

(g)



(h)

FIGURE 11: Moran's I of variables. (a) Moran's I *of HH*I. (b) Moran's I of Intersection density. (c) Moran's I of bus station density. (d) Moran's I of population density. (e) Moran's I of distance to CBD. (f) Moran's I of traffic safe. (g) Moran's I of traffic comfortable. (h) Moran's I of traffic efficiency.
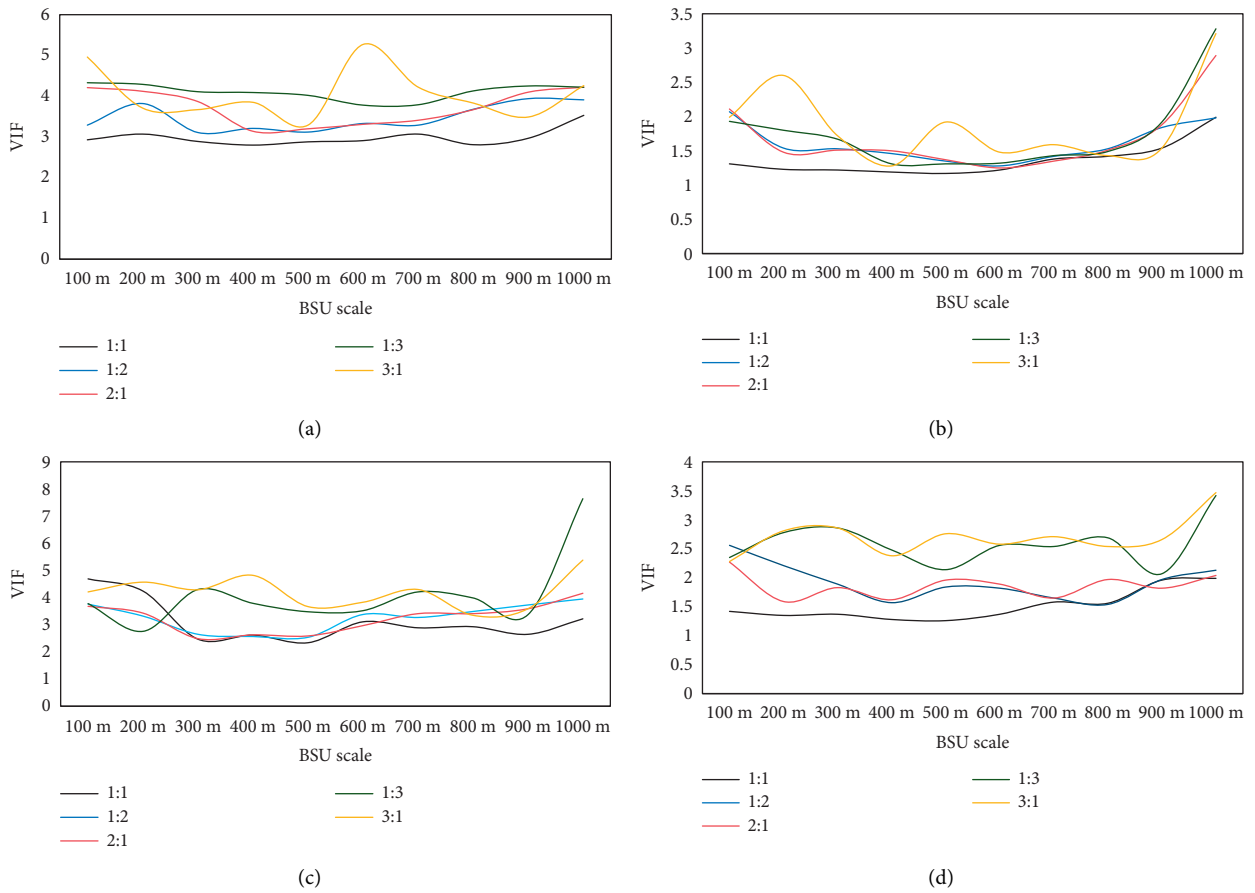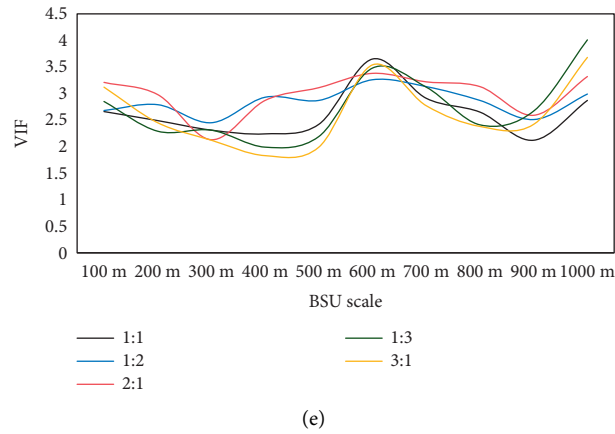


(a)



(b)



(c)



(d)
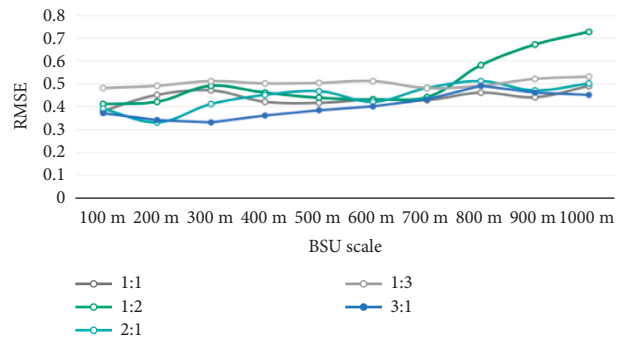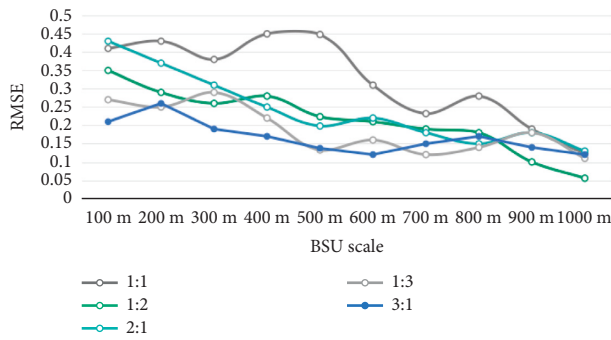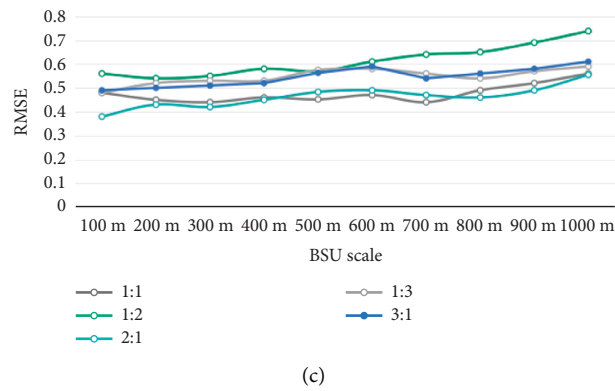
FIGURE 12: Continued.

(e)

FIGURE 12: VIF of variables. (a) VIF of Population density. (b) VIF of HHI. (c) VIF of Intersection density. (d) VIF of Distance to CBD. (e) VIF of bus station density.
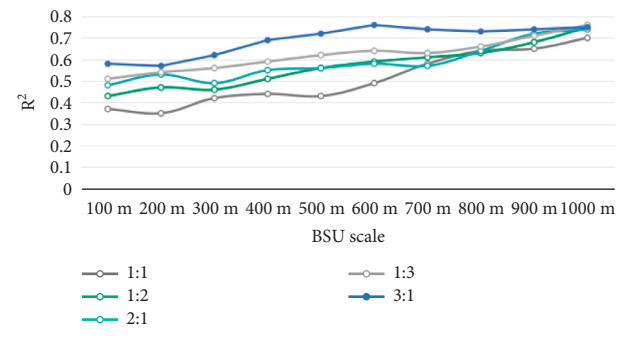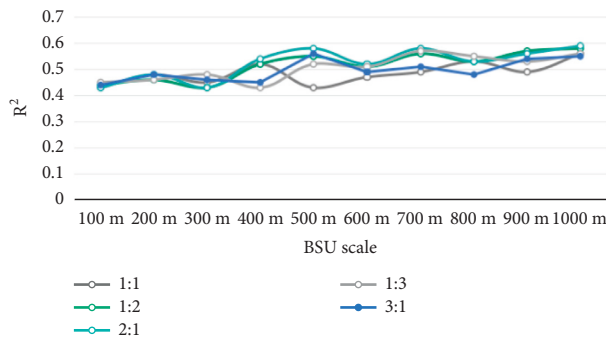


(a)



(b)



(c)

FIGURE 13: RMSE of GWR model. (a) RMSE of GWR-TE. (b) RMSE of GWR-TS. (c) RMSE of GWR-TC.
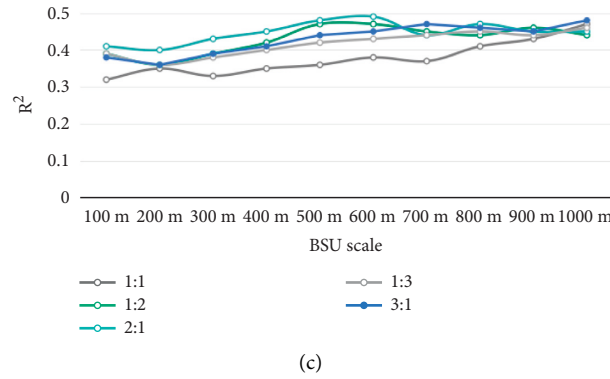


(a)



(b)

FIGURE 14: Continued.

(c)

FIGURE 14: $R^2$ of GWR model. (a) $R^2$ of GWR-TE. (b) $R^2$ of GWR-TS. (c) $R^2$ of GWR-TC.

TABLE 4: Judgment matrix filled in by expert A.

|       | $B_1$ | $B_2$ | $B_3$ | $B_4$ |
|-------|-------|-------|-------|-------|
| $B_1$ | 1     | 0.67  | 0.85  | 0.75  |
| $B_2$ | 1.5   | 1     | 0.57  | 0.5   |
| $B_3$ | 1.17  | 1.75  | 1     | 0.88  |
| $B_4$ | 1.33  | 2     | 1.14  | 1     |

*Note*. B1 represents $I(z)$, which is the Moran's I of SU $z$; B2 represents $V(z)$, which is the VIF of SU $z$; B3 represents $R(z)$, which is the $R^2$ of the GWR model of SU $z$; B4 represents $RMSE(z)$, which is the RMSE of SU $z$.

## 7. Conclusion

To the best of our knowledge, this study is the first attempt to simultaneously investigate the effect of SU scale and type on the MAUP, affecting the relationship between BE and TSS. Another capital objective was to assess how changes in SU affect the variables and model parameter data. The consequences of MAUP cannot be ignored, as the extent to which MAUP affects variables and model results is obvious, but it is difficult to explain why SU changes lead to inconsistent changes in model parameters. However, several rules can still be extracted. Longitude and latitude zoning (such as 1 : 3 or 3 : 1) has little influence on data aggregation, while SU scale variation has a great influence on both variables and model parameters. Larger SU scales resulted in weaker data differences and stronger spatial autocorrelation. These findings suggest that a change in the SU scale should prompt additional checks on the randomness of the spatial data distribution and collinearity. Comparative analysis showed that a $900 \, m \times 900$ m square grid was the optimal size for investigating the relationship between BE and TSS. However, this result is only the optimal solution for the available data, and may not be valid when the data are updated. Thus, the effect of the MAUP requires further investigation under various conditions.

Moreover, this study has some limitations. First, the SU scale and types were mainly based on square grids, whereas in the field of transportation, SUs generally include additional types, such as TAZ and Tyson polygons. In the future, the detection scope of SU types will be expanded to study the MAUP more comprehensively. Second, in this study, only the MAUP effect on the GWR model was investigated, but

the applicability of the method to different models, such as multiple linear regression or conditional autoregressive models, for determining the effects of MAUP was not considered.

## Data Availability

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

## References

[1] AS. Fotheringham and D. W. S. Wong, "The modifiable areal unit problem in multivariate statistical analysis," *Environment & Planning A*, vol. 23, no. 7, pp. 1025–1044, 1991.

[2] D. Manley, *Scale, Aggregation, and the Modifiable Areal Unit Problem*, Springer Berlin Heidelberg. Press, Berlin, BLN, GER, 2014.

[3] S. Openshaw, "An empirical study of some zone-design criteria," *Environment & Planning A*, vol. 10, no. 7, pp. 781–794, 1978.

[4] A. D. Páez and D. M. Scott, "Spatial statistics for urban analysis: a review of techniques with examples," *Geojournal*, vol. 61, no. 1, pp. 53–67, 2004.

[5] Á. Briz-Redón, F. Martínez-Ruiz, and F. Montes, "Investigation of the consequences of the modifiable areal unit problem in macroscopic traffic safety analysis: a case study accounting for scale and zoning," *Accident Analysis & Prevention*, vol. 132, Article ID 105276, 2019.

[6] M. J. T. L. Gomes and F. J. C. Cunto, "Influence of different spatial aggregations on variables implemented in macroscopic road-safety modeling," *Journal of Transportation Engineering Part A Systems*, vol. 146, no. 4, Article ID 04020014, 2020.

[7] X. Zhou, C. Sun, X. Niu, and C. Shi, "The modifiable areal unit problem in the relationship between jobs–housing balance

and commuting distance through big and traditional data," *Travel Behaviour and Society*, vol. 26, pp. 270–278, 2022.

[8] H. X. Pan, "Post-expo shanghai's LowCarbon city traffic and land use 5D model," *Shanghai City Planning*, pp. 27–32, 2011.

[9] N. Mohajeri, A. Gudmundsson, and J. R. French, "CO$^2$ emissions in relation to street-network configuration and city size," *Transportation Research Part D: Transport and Environment*, vol. 35, pp. 116–129, 2015.

[10] M. R. Majid, A. N. Nordin, and I. N. Medugu, "Influence of housing development designs on household vehicle miles traveled: a case of iskandar Malaysia," *Transportation Research Part D: Transport and Environment*, vol. 33, pp. 63–73, 2014.

[11] Y. Wei and C. Pan, "Urban land use characteristics and Commuters'Travel pattern: a case study of west hangzhou," *City Planning Review*, vol. 3, pp. 76–84, 2012.

[12] S. Zhong, Z. Wang, Q. Wang, A. Liu, and J. Cui, "Exploring the spatially heterogeneous effects of urban built environment on road travel time variability," *Journal of Transportation Engineering Part A Systems*, vol. 147, p. 1, Article ID 04020142, 2021.

[13] J. Hong, Q. Shen, and L. Zhang, "How do built-environment factors affect travel behavior? A spatial analysis at different geographic scales," *Transportation*, vol. 41, no. 3, pp. 419–440, 2013.

[14] R. Cervero and K. Kockelman, "Travel Demand and the 3Ds: Density, Diversity, and Design," *Transportation Research Part D Transport & Environment*, vol. 2, pp. 199–219, 1997.

[15] R. Ewing and R. Cervero, "Travel and the Built Environment: a Synthesis," in *Proceedings of the Redefining, reevaluating and reinventing transit: the transportation/land~use/environmental connection*, C.A ,USA, October 2001.

[16] T. Li, P. Jing, L. Li, D. Sun, and W. Yan, "Revealing the varying impact of urban built environment on online car-hailing travel in spatio-temporal dimension: an exploratory analysis in chengdu, China," *Sustainability*, vol. 11, no. 5, pp. 1–17, 2019.

[17] T. R. Jones and R. B. Potts, "The measurement of acceleration noise—a traffic parameter," *Operations Research*, vol. 10, no. 6, pp. 745–763, 1962.

[18] A. S. Fotheringham, C. Brunsdon, and M. Charlton, *Geographically Weighted Regression*, Wiley, New York, NY, USA, 2002.

[19] M. S. Rahman, M. Abdel-Aty, S. Hasan, and Q. Cai, "Applying machine learning approaches to analyze the vulnerable road-users' crashes at statewide traffic analysis zones," *Journal of Safety Research*, vol. 70, pp. 275–288, 2019.

[20] J. Lee, S. Yasmin, N. Eluru, M. Abdel-Aty, and Q. Cai, "Analysis of crash proportion by vehicle type at traffic analysis zone level: a mixed fractional split multinomial logit modeling approach with spatial effects," *Accident Analysis & Prevention*, vol. 12, p. 22, 2018.

[21] B. Sun and Z. He, "Could jobs-housing balance relieve traffic congestion?-A case study of Shanghai based on GIS buffer," *Urban planning forum*, vol. 5, pp. 98–104, 2017.

[22] T. Li, H. Jiang, P. Jing, and M. Zhang, "Analyzing the coordination relationship between urban built environment and traffic level," *Journal of Advanced Transportation*, vol. 2021, Article ID 9952306, 11 pages, 2021.

[23] C. Wang, *Study on the Built Environment Optimization Strategy of China-Singapore Eco-City under the Guidance of Low-Carbon Travel*, Ph.D. dissertation, Xi 'an University of Architecture and Technology, Chinia, 2020.

[24] Q. Niu and X. Niu, "Research on the division method of macro meso commuter traffic analysis unit based on community discovery algorithm: a case study of wuhan central city," *Traffic Information and Safety*, vol. 38, pp. 10–21, 2020.

[25] C. Ding, D. Wang, C. Liu, Y. Zhang, and J. Yang, "Exploring the influence of built environment on travel mode choice considering the mediating effects of car ownership and travel distance," *Transportation Research Part A: Policy and Practice*, vol. 100, no. 6, pp. 65–80, 2017.

[26] T. Chowdhury, D. M. Scott, and D. M. Scott, "Role of the built environment on trip-chaining behavior: an investigation of workers and non-workers in Halifax, Nova Scotia," *Transportation*, vol. 47, no. 2, pp. 737–761, 2018.

[27] A. M. Larranaga, L. I. Rizzi, J. Arellana, O. Strambi, and H. B. B. Cybis, "The influence of built environment and travel attitudes on walking: a case study of Porto Alegre, Brazil," *International Journal of Sustainable Transportation*, vol. 10, no. 4, pp. 332–342, 2016.

[28] C. Brunsdon, A. S. Fotheringham, and M. E. Charlton, "Geographically weighted regression: a method for exploring spatial nonstationarity," *Geographical Analysis*, vol. 28, no. 4, pp. 281–298, 2010.