

## Research Article

# Spatially Varying Relation between Built Environment and Station-Level Subway Passenger-Distance

Hongtai Yang <sup>1</sup>, Zhihao Zhao <sup>1</sup>, Chaozhe Jiang <sup>1</sup>, Yi Wen <sup>2</sup> and Malik Muneeb Abid<sup>3</sup>

<sup>1</sup>School of Transportation and Logistics,  
National Engineering Laboratory of Integrated Transportation Big Data Application Technology,  
National United Engineering Laboratory of Integrated and Intelligent Transportation,  
Institute of System Science and Engineering, Southwest Jiaotong University, 611756 Chengdu, China

<sup>2</sup>Department of Civil and Environmental Engineering, University of Tennessee, Knoxville 325 John D Tickle Building,  
851 Neyland Drive, Knoxville, TN 37996, USA

<sup>3</sup>Department of Civil Engineering, College of Engineering and Technology, University of Sargodha, Sargodha, Pakistan

Correspondence should be addressed to Zhihao Zhao; [zhaozhihao@my.swjtu.edu.cn](mailto:zhaozhihao@my.swjtu.edu.cn) and Chaozhe Jiang; [jiangchaozhe@home.swjtu.edu.cn](mailto:jiangchaozhe@home.swjtu.edu.cn)

Received 27 June 2021; Revised 30 December 2021; Accepted 6 January 2022; Published 20 January 2022

Academic Editor: Luis Miranda-Moreno

Copyright © 2022 Hongtai Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As a sustainable mode of transportation, subways bring great convenience to the society. Although there have been many studies examining the relationship between the built environment and the station-level ridership, those studies focused mainly on the ridership, which is defined as the number of trips for each station. While ridership is an important indicator for evaluating subway demand, passenger-distance is another critical indicator that incorporates distance into demand evaluation, which has not yet been fully explored. To fill this gap, this paper investigates the relationship between the built environment around stations and the station-level passenger-distance (SLPD). As noted in previous studies, the relationship between the built environment and travel demand can vary by space. Therefore, a geographically weighted regression (GWR) model and a mixed geographically weighted regression (MGWR) model have been used to explore this spatially varying relationship using Chengdu, China, as an example case. The results were compared with that of an ordinary least squares (OLS) model. The comparison shows that the MGWR model that considers both global and local variables has the best goodness of fit. Results also show that 11 of the 25 potential variables are significantly related to SLPD. The accessibility of the station, station type, such as transfer or terminal, number of bus stops, number of restaurants, density of building area, density of the national road network, and density of the provincial road network, all have a positive correlation with SLPD. Meanwhile, the variables, whether it is a newly opened subway station, density of living points of interest (POIs), and density of railroad network, are all negatively correlated with SLPD. Ten of the eleven significant variables (except accessibility) have spatially varying relationships with SLPD. These findings can serve a useful reference for transportation planners for the demand evaluation.

## 1. Introduction

With urbanization, a growing number of cars are occupying the roadways, which brings along a series of problems, such as vehicular traffic congestion [1], air pollution [2, 3], and fuel consumption [4]. Subway has been considered as a sustainable public transportation mode to alleviate traffic congestion [5]. As early as 1863, the world's first subway system opened in London, England, with a trunk line of

about 6.5 km, using a combination of underground and ground lines (Railway Technology). Construction of China's subway began in 1965, and in May 2020, 47 cities in China had urban rail transit (China Urban Rail Transit Association). At the end of 2016, China's urban rail transit construction investment reached 384.7 billion yuan (6.6423 billion US dollars in 2016), and the total length of urban rail transit lines under construction reached 5636.5 kilometers [6]. In recent years, transit-oriented development

(TOD) has become an effective way to alleviate traffic congestion and promote sustainable transportation modes [7, 8], and urban rail transit has been an important part of sustainable development both in developed and developing countries in cities such as New York, Hong Kong, and London [9].

The built environment has been found to have a great impact on travel behavior as well as subway ridership. The concept of built environment variables has been expanded from the original “3D” (density, diversity, and design) (Cervero and Kockelman [10]), to the “6D” (density, diversity, design, destinations accessibility, distance to transit, and demand management) [11, 12]. Therefore, it is important for transportation planners to understand the relationship between the built environment and the station-level ridership. In the literature, previous studies mainly use ridership as the response variable [8, 13–18]. However, this variable does not take into consideration the travel distance. Therefore, this may lead to the partial conclusion that when two stations have the same ridership, the station with longer passenger-distance can be regarded as the station with higher demand. To address this problem, SLPD has been taken as the dependent variable against the built environment. This study addresses this issue by exploring the relationship between the dependent variable SLPD and the built environment.

Regarding modeling, since previous studies have found that the built environment typically has a spatially varying effect on travel behaviors [17], geographically weighted regression (GWR) models are adopted in this study, which allows the coefficients of variables to vary across space to capture the spatially varying relationship. In addition, the GWR model relies on a strong assumption that the coefficient of all independent variables varies across the space. As the MGWR model allows variables to be either global or local, this model has been used in this paper, to investigate the relationship between the built environment and SLPD. The data from Chengdu, China, is used to perform the analysis.

The remainder of the paper is organized as follows: Section 2 reviews the previous studies on the relationship between the built environment and subway demand; Section 3 provides a description of the data; Section 4 explains the ordinary least squares (OLS) models, the GWR models, and the MGWR models; Section 5 presents model results with interpretations; and Section 6 summarizes the main findings of this paper.

## 2. Literature Review

There have been some studies exploring the impact of the TOD on the behavior of residents around the stations. Some studies have explored the principles of TOD by studying the relationship between the ridership and the built environment around subway stations [13, 19, 20], such as population and employment density. In addition, the concept of TOD has also been applied in the field of logistics [21–23].

The built environment is the man-made environment for human activities [24], covering land-use patterns, urban

design, and transportation infrastructure. Built environment can play an important role in influencing travel behaviors [25–27]. For example, Ding et al. [28] explored the influence of the “4D” built environment variables, which are density, design, diversity, and distance to the Central Business District (CBD), on subway ridership in Washington, D.C., and concluded that the built environment variables played a significant role with a total contribution of 34% to subway ridership estimation. A wealth of literature has also explored the relationship between the built environment and subway ridership in China. Zhao et al. [16] suggested the significant effects of eleven built environment variables on subway ridership such as business/office floor area and road length by studying 55 metro stations in Nanjing, China. Liu et al. [29] found that the type of land use around stations and the accessibility of rail stations had a significant impact on passenger flow by studying the trip data of Guangzhou Subway from 2011 to 2016. Yang et al. [30] explored the synergy between built environment variables such as land use and the city attributes based on Shenzhen subway data, and showed that the Baidu heat index (an indicator for destination popularity) for restaurants and entertainment around subway stations was higher at night than that during the day, which triggered interests for exploring the urban “night market economy.” Du and Zheng [31] studied the characteristics of different types of commercial lease-type built environment variables around the Beijing subway network, linking concentrated businesses and dispersed labor through the analysis of metrics such as subway accessibility. Li et al. [32] investigated the impact of the built environment factors on urban rail transit in Guangzhou, China, in a study that integrated multiple sources of spatial big data such as points of interest (POI), high spatial resolution remote sensing images, social media, and building footprint data. However, the above studies rarely comprehensively investigated various types of built environment variables, such as POI class, road network class, density class, etc.

When studying the effects of the built environment on the subway demand, most studies use ridership as a response variable. For example, Chen et al. [33] used daily ridership as the dependent variable in their study of daytime patterns of transit riders of the New York City subway system. However, this response variable could not reflect passenger travel distance. Based on previous reviews, there is only one study using passenger trip miles as the response variable. Iseki et al. [34] studied the determinants of passenger miles traveled (PMT) for each origin-destination (OD) station pair in Washington D.C., but the relationship between determinants and the PMT was assumed to be constant across the study area, which may be context-specific.

At present, studies on urban rail transit passengers generally use two models: spatial models and nonspatial models. Ordinary least squares (OLS) regression models are nonspatial models and the most representative linear regression models were used to initially explain the complex relationship between the built environment and subway capacity [15, 35, 36], which assumes that the relationship between independent variables and the dependent variable is

global and not spatially heterogeneous. In terms of spatial models, among the most common are spatial error models [37, 38], spatial lag models [39], geographically weighted regression models [32, 40, 41], and derivative models related to geographically weighted regression models, such as geographic time-weighted regression (GTWR) models [42], geographically weighted negative binomial regression (GWNBR) models [43], geographically weighted Poisson regression (GWPR) models [44], and mixed geographically weighted regression (MGWR) models [13, 18, 45]. For example, Li et al. [32] used a GWR model to refine the study of Guangzhou Subway. Although the GWR models yield a satisfactory fit for the study of spatial heterogeneity of subway stations, they treat all variables as local variables, assuming that all built environment variables vary across space, which can lead to a biased result. To address this, MGWR has been developed. For example, Jun et al. [13] used a stepwise regression model, and the MGWR model in their study of Seoul Subway showed that the MGWR model fits the data better, testifying for the superiority of the MGWR models in handling spatial-temporal data. In sum, few previous papers have analyzed the spatially varying relationships between the built environment and the SLPD. Therefore, it would be worthwhile contribution to literature to explore the spatially varying impact of the built environment on the subway station demand.

### 3. Data Description

**3.1. Study Area.** Chengdu is located in Sichuan Province, China, and has been known as the “Land of Heaven,” with a total area of 14,335 square kilometers, a built-up area of 949.6 square kilometers, and a resident population of 16,330,000 (Chengdu Municipal People’s Government). Chengdu had five subway lines under operation, which were Line 1, Line 2, Line 3, Line 7, and Line 10 as in March 2018, with a total of 136 stations, as shown in Figure 1. In March 2018 alone, the number of subway trips was more than 60 million. This study uses the automatic fare collection data of the Chengdu subway from March 18 to 31, 2018. The average number of trips per day during this period was 2,061,853, the average passenger travel distance was 14.5 km, and the average passenger travel time was 30 minutes.

When constructing the buffer zone, two types of distances are considered: Euclidean distance and network distance. Many studies have used the Euclidean distance [32, 46–49] while some other studies adopted the network distance [50, 51]. Guo and Bhat [52] and Schirmer et al. [53] showed that the modeling results based on the buffer zone constructed by network distance and Euclidean distance are similar. As a result, the Euclidean distance is used in this paper. The choice of the catchment area is an important consideration in spatial modeling. Relevant studies [32, 46, 47] found that rail stations generally had a catchment area with a radius of 800 m. Therefore, a circular buffer zone with a radius of 800 m is used as a catchment area in this paper. Considering the fact that some buffer zones may have overlapping areas, which could lead to double (or multiple)

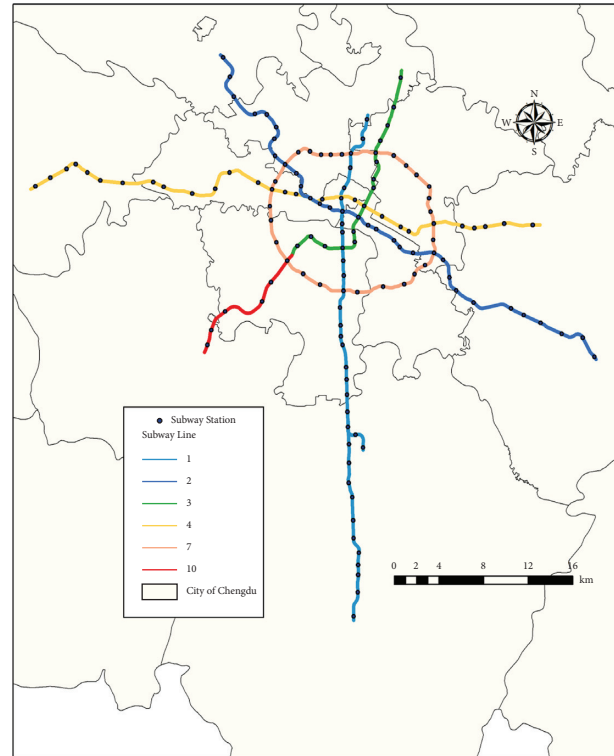


FIGURE 1: Chengdu subway network.

counting of some variables, the Tyson polygons have been used, as shown in Figure 2.

**3.2. Data.** Regarding explanatory variables, 25 built environment variables are selected. One of the major types of variables is the point of interest (POI) variables, which are commonly used in transportation related studies to reflect the land-use characteristics [36, 54, 55]. The POI data of Chengdu with longitude and latitude information have been used in this paper, and ArcGIS is used for spatial processing to retrieve POI building environment variables. The data of POI variables come from Amap (China’s Google Map equivalent) (<https://www.amap.com>). In addition, OpenStreetMap (<https://www.openstreetmap.org/>) is used to obtain the variables related to the road network.

Since the automatic fare collection (AFC) system records detailed trip information including card number, card type, trip date, boarding station name, boarding time, alighting station name, and alighting time, AFC data have been used in this study to estimate passenger travel distances at each station. The distance of each station pair along the metro line has been measured using ArcGIS. Since there are 136 stations, a 136 by 136 distance matrix has been obtained (Table 1). In the data, the origin station and the destination station of each user has been recorded. This information has been used to calculate the value of the response variable *Passenger-distance* and the value of the explanatory variable *Accessibility*.

Further, the *Passenger-distance* of each station has been obtained by the product of the OD distance and the

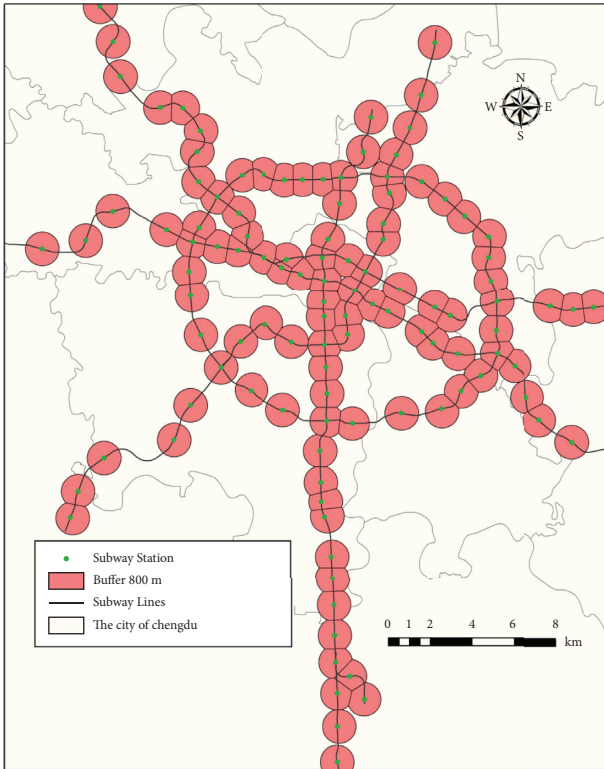


FIGURE 2: Buffer zone of subway stations.

passenger flow of the station. Since the response variable, passenger-distance, shows a highly left-skewed distribution, the variable is log-transformed to meet the normality assumption for the OLS models. The variables have been described in Table 2.

The following equation has been used for the calculation of accessibility of a station [56].

$$A_i = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^n d_{ij}, \quad (1)$$

where  $A_i$  denotes the accessibility of the station  $i$ ,  $d_{ij}$  denotes the distance from station  $i$  to station  $j$ , and  $n$  denotes the number of stations in the subway system.

**3.3. Clustering Analysis.** The subway stations are clustered using the K-means clustering algorithm based on the 25 explanatory variables. The K-means clustering algorithm is a classic clustering algorithm that was proposed by MacQueen [57] in 1967. When using the K-means clustering, the number of clusters, represented by  $K$ , needs to be determined. The commonly used methods are the Elbow diagram and the contour coefficient method [58]. The result of the Elbow diagram and the contour coefficient method are shown in Figure 3. From the Elbow diagram, the elbow points occur when the number of clusters is 4 or 5. From the contour coefficient diagram, the contour factor is the highest when the number of clusters is 5. As a result, the value of  $K$  is set as 5.

The results of the clustered stations are shown in Figure 4. Most stations in clusters 1 and 2 are located in areas far away from the city center of Chengdu. Compared to the city center, these areas have a lower density of population and employment, which may result in lower passenger flows. Cluster 3 shows the characteristics of stations in the city center, where the density of population and employment is high. Cluster 4 only has one subway station, Chunxi Road subway station, which is located in a busy commercial corridor. At the same time, this is a typical case of TOD development. Through close integration with commerce, it also drives the traffic of the site. Thus, this station may stand out as a unique station type. Cluster 5 is mainly located around the loop subway line of Line 7. These stations have similar characteristics with high percentages of residential areas.

**3.4. Passenger Flow Characteristics.** When describing the passenger flow characteristics, generally more attention is paid to the daily passenger flow and the distribution of passengers' travel time. Figure 5 (unit: person/day) shows that the weekday passenger flow is higher than that on the weekends. After calculation, the average number of trips on weekdays is 2,184,722 per day, and the average number of trips on weekends is 1,756,813 per day. Furthermore, the passenger flow of Line 1 is significantly influenced by the day of the week, indicating that Line 1 passengers are likely mainly commuters. Line 10, on the other hand, is not affected by the day of the week, as it is largely airport traffic. According to the distribution of travel time of subway passengers (Figure 6 (Note: Unit is number of people)), the travel time is mostly within the range of 0–100 minutes (99.65%). The average travel time of passengers is 29.7 minutes. Among them, the highest passenger flow was 758,264 at about 19 minutes.

According to the heat map of hourly passenger flow (Figure 7), the vertical axis represents date, which ranges from the 18<sup>th</sup> to the 31<sup>st</sup>. The horizontal axis represents time of the day, which ranges from 6 to 23. Darker color indicates higher passenger flow. The differences in passenger flow exist between weekdays and weekends. On weekdays, there are two periods of peak ridership, one in the morning (7:00–8:00) and the other one in the evening (17:00–18:00). The average passenger flow during the morning peak hours (7:00–8:00) on weekdays is 247,808.7 per hour, and the average passenger flow during the evening peak hours (17:00–18:00) is 217,769.3 per hour. The ridership in the morning peak hours is higher than that in the evening peak hours, which is consistent with the findings of Ma et al. [36]. However, on weekends, the subway ridership does not have a clear peak period, which is consistent with the findings of [40, 59]. When analyzing the hourly subway ridership, it has been found that the trips are mainly concentrated between 6:00 and 23:00, consistent with the operating hours of the Chengdu Subway.

The spatial distribution of Chengdu subway passenger flow shows that trips are concentrated around the city

TABLE 1: Example of the distance matrix (unit: meter).

	Balizhuang station	Baiguolin station	Baicaolu station	2nd beizhan road west station	Caiqiao station	Caotang road north station	Chadianzi bus terminal station
Balizhuang station	0	11714.936	18121.486	5301.106	19886.045	11549.778	12711.242
Baiguolin station	11714.936	0	10159.250	7164.942	10995.102	1553.493	4749.006
Baicaolu station	18121.486	10159.250	0	12820.380	17579.247	11712.744	5410.244
2nd beizhan road west station	5301.106	7164.942	12820.380	0	14584.939	7699.487	7410.136
Caiqiao station	19886.045	10995.102	17579.247	14584.939	0	9441.608	12169.003
Caotang road north station	11549.778	1553.493	11712.744	7699.487	9441.608	0	6302.500
Chadianzi bus terminal station	12711.242	4749.006	5410.244	7410.136	12169.003	6302.500	0

TABLE 2: Descriptive statistics of the variables.

Variable name	Description	Descriptive statistics		Data source
		Mean	S.D.	
Response variable				
Passenger-distance	The logarithmic transformation of actual line distance of cumulative passenger flow at starting station (km)	14.457	0.862	AFC
Explanatory variable				
New station	Is it a new site? 1 means yes, 0 means no	0.096	0.295	ArcGIS
Terminal	Is it a terminal station? 1 means yes, 0 means no	0.066	0.250	ArcGIS
Transfer	Is it a transfer station? 1 means yes, 0 means no	0.103	0.305	ArcGIS
Accessibility	Accessibility of the subway station (km)	16.851	5.408	ArcGIS
City center	Straight line distance from the city center (km)	9.389	6.717	ArcGIS
Bus	The density of bus stops (/km <sup>2</sup> )	6.308	3.518	Amap
Hotel	The density of hotels (/km <sup>2</sup> )	2.343	4.111	Amap
Restaurant	The density of restaurants (/km <sup>2</sup> )	1.370	2.754	Amap
Shopping	The density of shopping POIs (/km <sup>2</sup> )	1.052	2.481	Amap
Parking	The density of parking structures(/km <sup>2</sup> )	0.132	0.651	Amap
Finance	The density of financial service POIs (/km <sup>2</sup> )	2.475	8.217	Amap
Scenery	The density of scenic spots (/km <sup>2</sup> )	1.280	3.030	Amap
Car	The density of automotive service POIs (/km <sup>2</sup> )	0.458	1.645	Amap
Life	The density of life service POIs, such as barber shops, pharmacies, skin care shops, laundry shops, etc. (/km <sup>2</sup> )	0.582	1.738	Amap
Medical	The density of medical service POIs (/km <sup>2</sup> )	1.408	1.959	Amap
Recreation	The density of recreation POIs (/km <sup>2</sup> )	1.012	1.854	Amap
Buildings	Density of building areas (m <sup>2</sup> /km <sup>2</sup> )	0.628	0.517	OpenStreetMap
Highway	Density of main highways (m/km <sup>2</sup> )	73.989	334.845	OpenStreetMap
National highway	Density of national highways (m/km <sup>2</sup> )	176.961	528.822	OpenStreetMap
Sidewalk road	Density of sidewalks (m/km <sup>2</sup> )	442.783	793.432	OpenStreetMap
Township road	Density of township roads (m/km <sup>2</sup> )	5546.912	2366.141	OpenStreetMap
County road	Density of county roads (m/km <sup>2</sup> )	1339.697	1222.087	OpenStreetMap
Railway	Density of railroads (m/km <sup>2</sup> )	201.761	495.193	OpenStreetMap
Dart	Density of provincial roads (m/km <sup>2</sup> )	931.768	1301.598	OpenStreetMap
Other road	Density of other roads (m/km <sup>2</sup> )	4655.385	2559.938	OpenStreetMap

center, the terminal stations, and the southern high-tech zone, which is a residential and employment center in Chengdu. The spatial distribution of the logarithm of the subway ridership (Figure 8) and the passenger-distance (Figure 9) in Chengdu show that the passenger-distance and the ridership at subway stations have the same spatial relationship overall, i.e., the more trips are made, the more passenger miles are traveled. However, there are some stations with inconsistent performance.

## 4. Models and Methods

*4.1. OLS Model.* The ordinary least squares (OLS) model is one of the most used models to determine the linear relationship between the explanatory variables and the response variable [60]. Its function is given as

$$y_i = \sum_{j=0}^k \beta_j x_{ij} + \varepsilon_i, \quad (2)$$

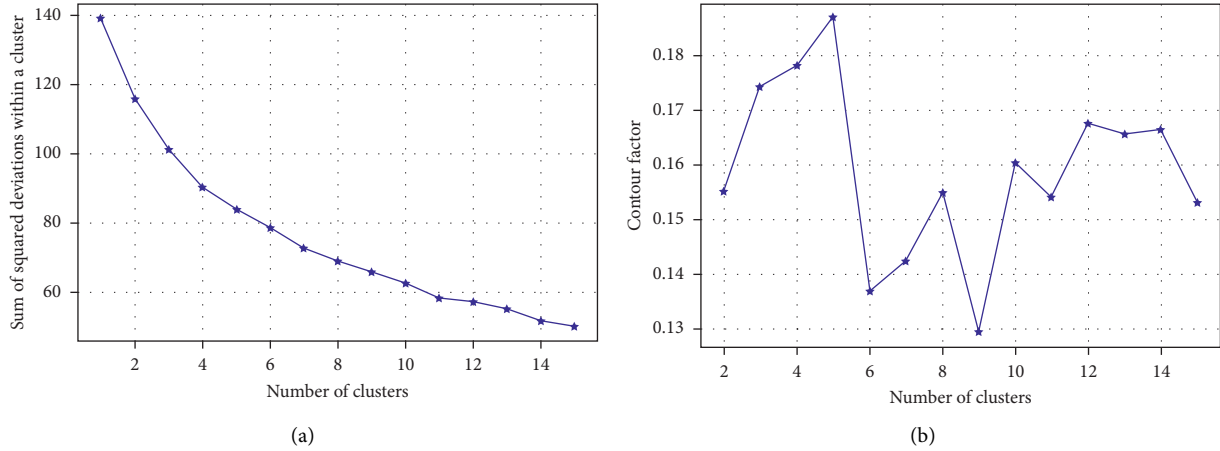


FIGURE 3: The results of the elbow diagram method (a) and the contour coefficient method (b).

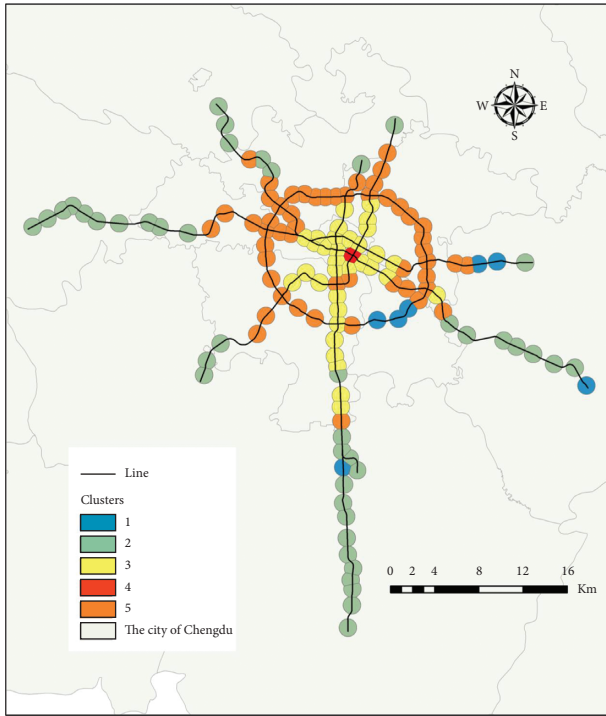


FIGURE 4: Clusters of subway stations.

where  $y_i$  denotes the  $i$ th observation of the response variable;  $\beta_j$  denotes the coefficient of the  $j$ th predicting variable;  $x_{ij}$  denotes the  $j$ th predicting variable of the  $i$ th observation; and  $\varepsilon_i$  denotes the random error term.

**4.2. GWR Model.** The OLS model assumes the relationship between the explanatory variables and the response variable to be consistent across the space. Yet, their relationship could vary across space. To overcome this issue, Brunson et al. [61] proposed the geographically weighted regression (GWR) model, which is commonly used to study the spatially varying relationships between the response and explanatory variables [62]. The function of the model is shown as

$$y_i = \beta_{i0}(u_i, v_i) + \sum_{j=1}^k \beta_j(u_i, v_i)x_{ij} + \varepsilon_i, \quad (3)$$

where  $y_i$  denotes the  $i$ th observation of the response variable;  $\beta_j$  denotes the coefficient of the  $j$ th predicting variable;  $x_{ij}$  denotes the  $j$ th predicting variable of the  $i$ th observation; and  $(u_i, v_i)$  denotes the coordinates of the  $i$ th observation;  $\varepsilon_i$  denotes the random error term.

**4.3. MGWR Model.** The MGWR model integrates the OLS model and the GWR model by allowing some variables to have a constant coefficient and other variables to have spatially varying coefficients [13]. The function of the MGWR model is shown as

$$y_i = \beta_{0(i)} + \sum_{a=1}^{p_a} \beta_a x_{ia} + \sum_{b=1}^{p_b} \beta_{ib} x_{ib} + \varepsilon_i, \quad (4)$$

where  $\beta_{0(i)}$  denotes the intercept, which can be set as a global intercept  $\beta_0$  or a local intercept  $\beta_{0i}$ ;  $x_{ia}$  is the  $a$ th global variable;  $x_{ib}$  is the  $b$ th local variable;  $\beta_a$  denotes the regression coefficient of the  $a$ th global variable; and  $\beta_{ib}$  is the regression coefficient of the  $b$ th local variable.

There are two main types of weighting functions to determine the coefficients of the MGWR model, which are (i) the Gaussian function and (ii) the double square function. For searching the bandwidth, there are four main methods, which are (i) Akaike Information Criterion (AICc) for small samples, (ii) Akaike Information Criterion (AIC), (iii) Bayesian Information Criterion (BIC), and (iv) cross-validation (CV). By exploring different weighting functions and search methods, the best model was obtained based on the CV search method and the Gaussian weighting function.

## 5. Results and Discussion

### 5.1. Model Results

**5.1.1. OLS Model Results.** Before constructing the OLS model, potential multicollinearity between variables needs to be eliminated. The variance inflation factor (VIF)

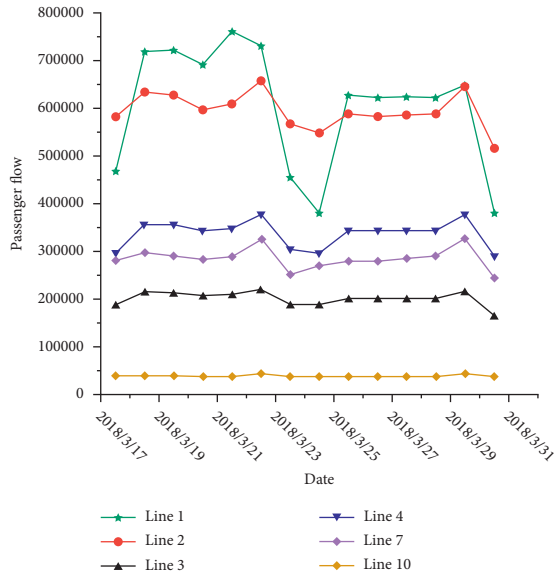


FIGURE 5: Daily passenger flow.

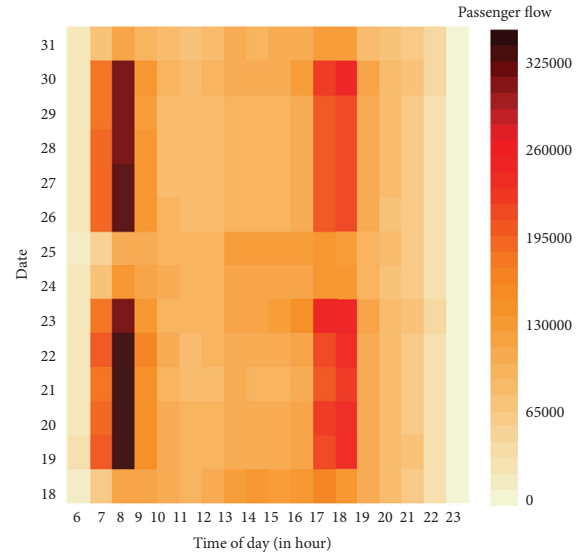


FIGURE 7: Heat map of hourly subway ridership.

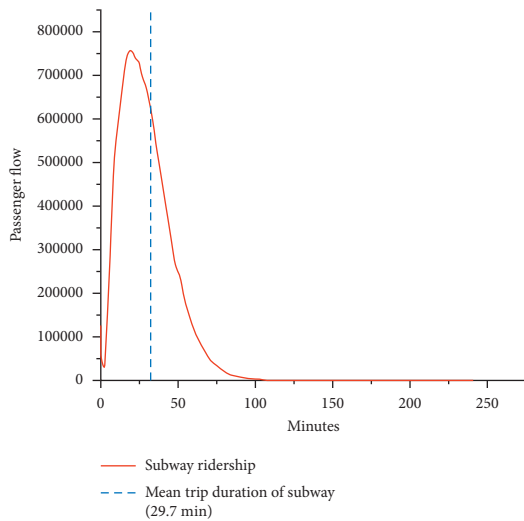


FIGURE 6: Distribution of subway travel time.

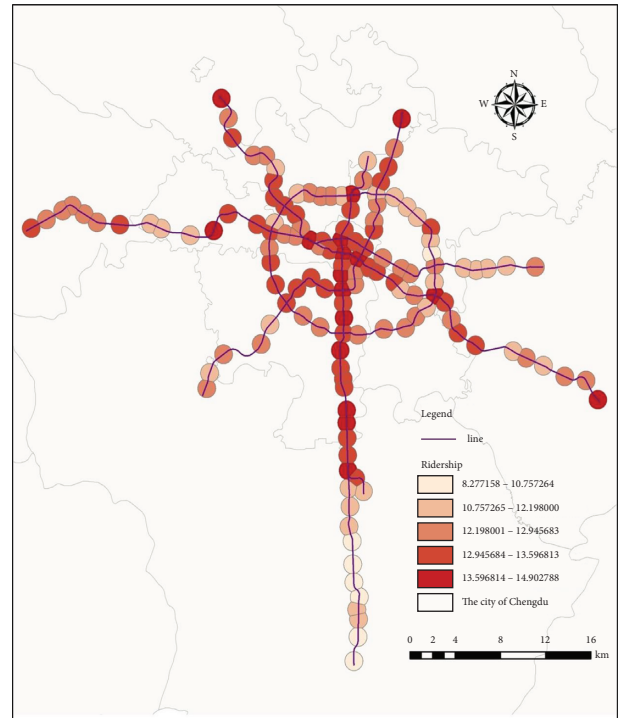


FIGURE 8: The spatial distribution of the logarithm of ridership in Chengdu subway.

is used in the paper to measure the multicollinearity among variables. According to Reference [63], when VIF is less than 10, the variables do not have a multicollinearity issue. After calculating the VIF values, the VIF values of the two variables *City center* and *Township road* were higher than 10. Thus, the *Township road* variable was removed, as its VIF value was the highest. To explore the correlation between the response variable and the explanatory variables, the correlation matrix is obtained based on the Pearson's correlation coefficient, as shown in Figure 10.

The results of the OLS model are shown in Table 3. It can be seen that 11 of the 25 explanatory variables are significant. In terms of *Passenger-distance*, significant variables at the 0.01 level include *New station*, *Transfer*, *Accessibility*, *Bus*, *Building*, and *National highway*. Among these variables, *Terminal*, *Transfer*, *Accessibility*, *Bus*, *Restaurants*, *Building*,

*National highway*, and *Dart* are positively related to the SLPD, while *New station*, *Life*, and *Railway* are negatively related. This result was compared with the result of the ridership model. There are nine variables that significantly affect the ridership at the 0.01 level, including *New station*, *Transfer*, *Restaurants*, *Bus*, *Building*, *National highway*, *Railroads*, and *Dart*.

From Table 3, it can be observed that the positive and negative effects of the response variable *Passenger-distance* and *Ridership* on the explanatory variables in the OLS model

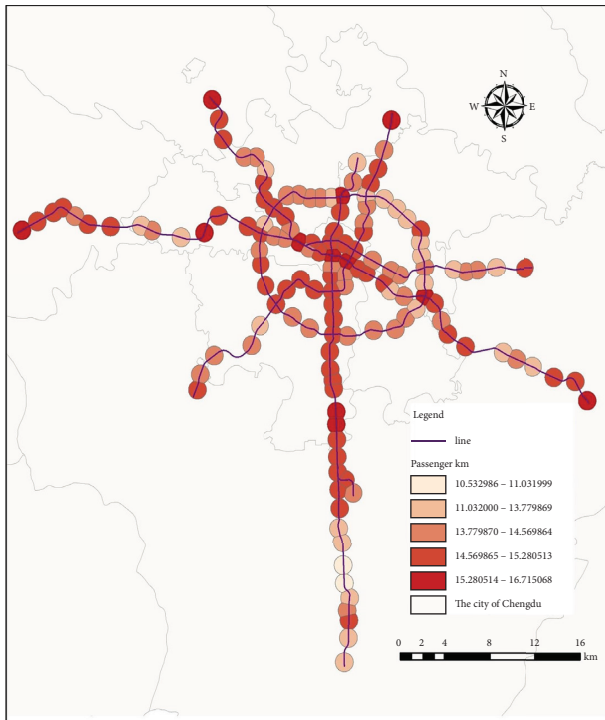


FIGURE 9: The spatial distribution of the logarithm of passenger-distance in Chengdu subway.

are comparable, except for the difference in the significance level of the variables *Lift* and *Accessibility*. In addition, the higher number of subway transit stations implies more passenger-distance, which is consistent with the findings of previous studies [35, 48].

Furthermore, it is necessary to consider the fact that passengers choose to travel by subway for different purposes. The variable *Accessibility* has a positive correlation with *Passenger-distance*. The reason could be that higher accessibility of subway lines is associated with a stronger willingness of passengers to use the station, which increases the number of trips. Consequently, it can have a positive effect on *Passenger-distance*. Unlike most studies, the *Life* has a negative effect on *Passenger-distance*, which may be because the high number of life POIs around the city reduces the need of passengers to travel by subway to reach these POIs. In contrast, the restaurant POIs have a positive effect on *Passenger-distance*, which may be explained as people in Chengdu are willing to travel for long distances to access different types of restaurants, resulting in more and longer trips to restaurant POIs.

Transportation-related variables have different effects on the travel distance of subway passengers. Generally, the more roads around a subway station, the more convenient it is for people to travel by subway [13]. For example, in the OLS model, the variables *National highway* and *Dart* both have a positive effect on *Passenger-distance* and *Ridership*. However, the railroad network can play a competing role against the subway, with nearby subway stations around the railroad station more likely to serve as an intermediate destination. Therefore, the more railway

stations there are, the fewer passenger-distance can be expected.

In addition, the subway station attribute variables *New station*, *Terminal*, and *Transfer* have different effects on *Passenger-distance*. For the variable *New station*, it has a negative relationship. The reason can be that the newly opened station is not as established as the older sites, which may result in less passenger flow in the short run. Variables *Terminal* and *Transfer* both show a positive relationship. In general, the passenger flow at the start and transfer stations of the subway is usually larger than that of other stations. These stations also have a larger passenger flow in the analysis. Therefore, more passenger distance is expected.

**5.1.2. GWR and MGWR Model Results.** When using the conventional linear regression models, estimating the relationship between variables through inefficient regression coefficients without considering the spatial autocorrelation of the data may cause a mismatch between the observation and the reality [46]. Therefore, the spatial autocorrelation of the explanatory variables needs to be measured using the spatial Moran's I index and Z score [64]. The estimation of Moran's I index by the software Geoda suggests that the variable *New station* does not have a strong spatial autocorrelation and therefore it shall be considered as a global variable. With the existence of spatial correlation of all other variables involved in the calculation, it can be concluded that the use of GWR and MGWR models is more reliable.

The GWR 4.0 software has been used for the calculation and fitting of the data using the MGWR model and the GWR model. Five important indicators Mean, STD, Min, Median, and Max for the data with the results have been presented in Table 4. For the GWR and the MGWR models, the signs of the coefficients for the global variables are similar to the signs of the variables in the OLS model. Among them, *Accessibility* is a global variable in the MGWR model and will not vary spatially.

**5.1.3. Result Comparison.** To further study the spatially varying relationship between the explanatory and the response variables, and address the problem of spatial autocorrelation, the GWR model and the MGWR model have been used to fit the data. In measuring the model performance, the results of the OLS model, GWR model, and MGWR model have been compared and analyzed using AIC, BIC, AICc, CV,  $R^2$ , and the Adjusted  $R^2$ . According to Huang et al. [65], the smaller the value of AIC, the more accurate the model is, and for different models, the difference of more than 3 in the value of AIC indicates a significant difference among the models [62]. In GWR 4.0 software, it has been observed that the model calculated by Gaussian fixed and CV is better. As shown in Table 5, the MGWR model has smaller AIC and AICc values than the OLS and GWR models, while the  $R^2$  increases to 0.671. By comparison, it has been observed that the MGWR model, which considers both global and local variables, has a better fit compared to the OLS model and the GWR model.



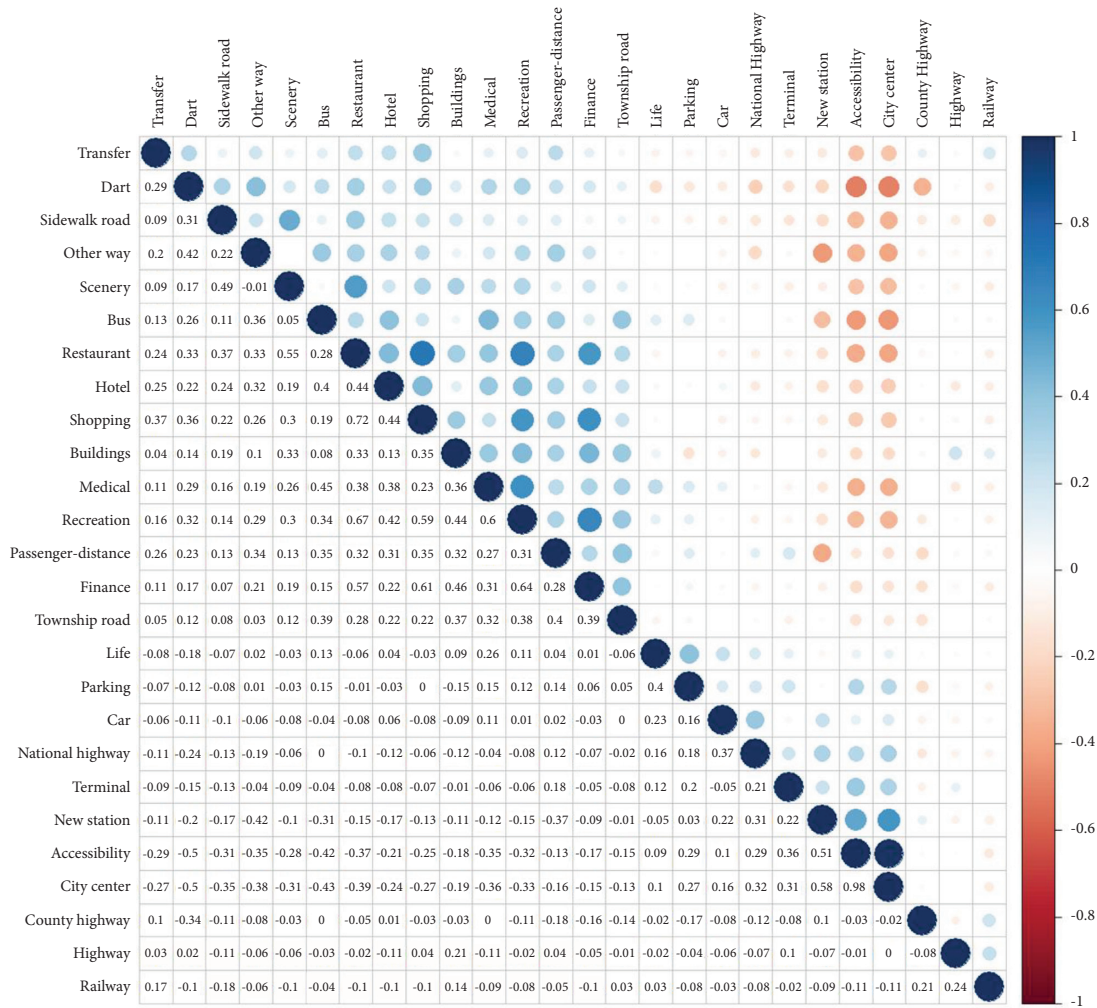


FIGURE 10: Pearson's correlation coefficient matrix.

TABLE 3: OLS model results.

	Passenger-distance			Ridership		
	Coef.	Std err	t-value	Coef	Std err	t-value
Constant	12.48	0.323	38.577***	11.8929	0.131	90.51***
New station	-1.191	0.229	-5.190***	-1.6689	0.207	-8.074***
Terminal	0.542	0.240	2.261**	0.4986	0.222	2.247**
Transfer	0.527	0.195	2.700***	0.5788	0.19	3.048***
Accessibility	0.065	0.015	4.368***	—	—	—
Bus	0.074	0.017	4.233***	0.0628	0.016	3.863***
Restaurant	0.036	0.022	1.620*	0.0362	0.022	1.681*
Life	-0.061	0.036	-1.713*	—	—	—
Buildings	0.51	0.118	4.342***	0.4815	0.114	4.223***
National highway	0.0004	0	3.318***	0.0003	0	3.174***
Railway	-0.0003	0	-2.559**	-0.0003	0	-3.1***
Dart	0.0001	0.0000519	2.186**	0.000099	0.0000462	2.145**

Note. \*indicates significance of 0.1; \*\*indicates significance of 0.05; \*\*\*indicates significance of 0.01.

5.2. Spatial Patterns. The spatial patterns of subway stations and their built environment, including distance, design, diversity, density, and accessibility, can have an impact on passenger travel behaviors [11, 40]. The research results show that the spatial distribution of the residuals of the

MGWR model has a more random distribution than the GWR and OLS models, with the GWR and OLS models showing patterns of spatial aggregation. Therefore, the GWR model and the OLS model fit the spatial data poorly. According to Figure 11, the spatial distributions of residuals

TABLE 4: Results of GWR and MGWR model.

Variable	GWR					MGWR				
	Mean	STD	Min	Median	Max	Mean	STD	Min	Median	Max
Intercept	14.461	0.004	14.451	14.461	14.474	14.438	0.131	13.722	14.469	14.532
New station	-0.416	0.004	-0.428	-0.415	-0.410	-0.396	0.092	-0.639	-0.392	-0.008
Terminal	0.174	0.023	0.097	0.178	0.212	0.197	0.080	0.034	0.189	0.397
Transfer	0.206	0.006	0.187	0.206	0.224	0.201	0.057	-0.015	0.201	0.389
Accessibility	0.257	0.007	0.251	0.255	0.287	—	—	—	—	—
Bus	0.175	0.008	0.163	0.173	0.203	0.187	0.078	0.121	0.154	0.593
Restaurant	0.112	0.002	0.105	0.112	0.117	0.130	0.035	0.019	0.138	0.274
Life	-0.049	0.007	-0.061	-0.050	-0.030	-0.018	0.045	-0.204	-0.023	0.123
Buildings	0.225	0.018	0.202	0.220	0.296	0.239	0.210	0.137	0.167	1.359
National highway	0.228	0.008	0.209	0.227	0.245	0.187	0.082	0.039	0.177	0.343
Railway	-0.076	0.010	-0.093	-0.078	-0.042	-0.050	0.137	-0.155	-0.090	0.720
Dart	0.131	0.003	0.126	0.131	0.140	0.129	0.060	-0.345	0.131	0.217

TABLE 5: Comparison of the results of the three models.

	OLS	GWR	MGWR
AIC	273.091	271.319	249.021
AICc	276.075	275.538	263.989
BIC	310.956	316.149	330.067
CV	0.461	0.459	0.436
$R^2$	0.511	0.534	0.671
Adj. $R^2$	0.464	0.471	0.567

of the MGWR model, GWR model, and OLS model are similar, but the range of residuals of the MGWR model is smaller than that of the GWR and the OLS models.

From the spatial visualization of the local  $R^2$  of the MGWR model and the GWR model (Figure 12), the spatial distributions of the local  $R^2$  of the two models show differences. However, the difference between the local  $R^2$  spatial visualization of the MGWR model and the GWR model is more obvious in the East side of Chengdu. The higher local  $R^2$  of the MGWR model is mainly distributed near subway stations in the eastern and southern regions.

From the perspective of multimodal transportation, subway trips are spatially significantly correlated with trips at bus stop densities, which is consistent with the findings of Zhao et al. [15]. Based on the results, it is observed that rail and bus have different cooperative or competitive impacts on *Passenger-distance* under different circumstances, which is similar to the findings of Chen et al. [5]. The Government of Seoul, South Korea, provides free transfers between subway and bus stations, which greatly encourage public transportation trips [35]. Therefore, it is recommended that Chengdu could develop a free or discount transfer policy between subway and bus to encourage the use of public transportation.

Moving onto station attributes, it is found that the transfer stations are generally recorded with larger passenger flows. Therefore, passenger-distance of transfer stations is higher. Figure 13 shows the spatial distribution of the three variables of station attributes *New station*, *Terminal*, and *Transfer*. The spatial distribution of the *New station* coefficient indicates a high value in the west and south but a low value in the north. This is likely due to opening of the third

phase of Chengdu Subway Line 1 (mainly located in the south) in mid-March 2018 (Chengdu Metro). As a result, the passenger-distance in the west and the surrounding areas was impacted, while the impact in the north was yet to show in the data.

As for roadway accessibility, by comparing the four submaps (a)–(c) in Figure 14, it is found that there is generally a greater demand in the southern zone of Chengdu. Since there are more job opportunities in the high-tech zone of Chengdu, this is likely generating the traffic of commuters. For the variable *National Highway*, it shows that in the eastern and southern regions, it has a strong positive correlation with passenger mileage, while for some stations in the north, it has a negative correlation. The reason for this situation can be that Chengdu's national highways are denser in the northwest; therefore, passengers in these areas are more likely to choose car-based modes such as private or ride-hailing cars. The variable *Railway* has a strong positive correlation with the southern subway station, while the northwest station has a negative correlation with the passenger-distance. The reason for this can be that the railway line is spread out in the northwest part, while in the south and east the stations are more sparsely distributed. The relationship between passenger-distance and the variable *Dart* is different from that of the previous two variables. It has a higher positive correlation in the western and southern regions. The reason can be that the provincial roadway network in the southern and the western districts is denser, while that in the northern and the eastern districts is sparser.

On the discussion of POIs, most POI class built environment variables seem to replace the equivalent types of

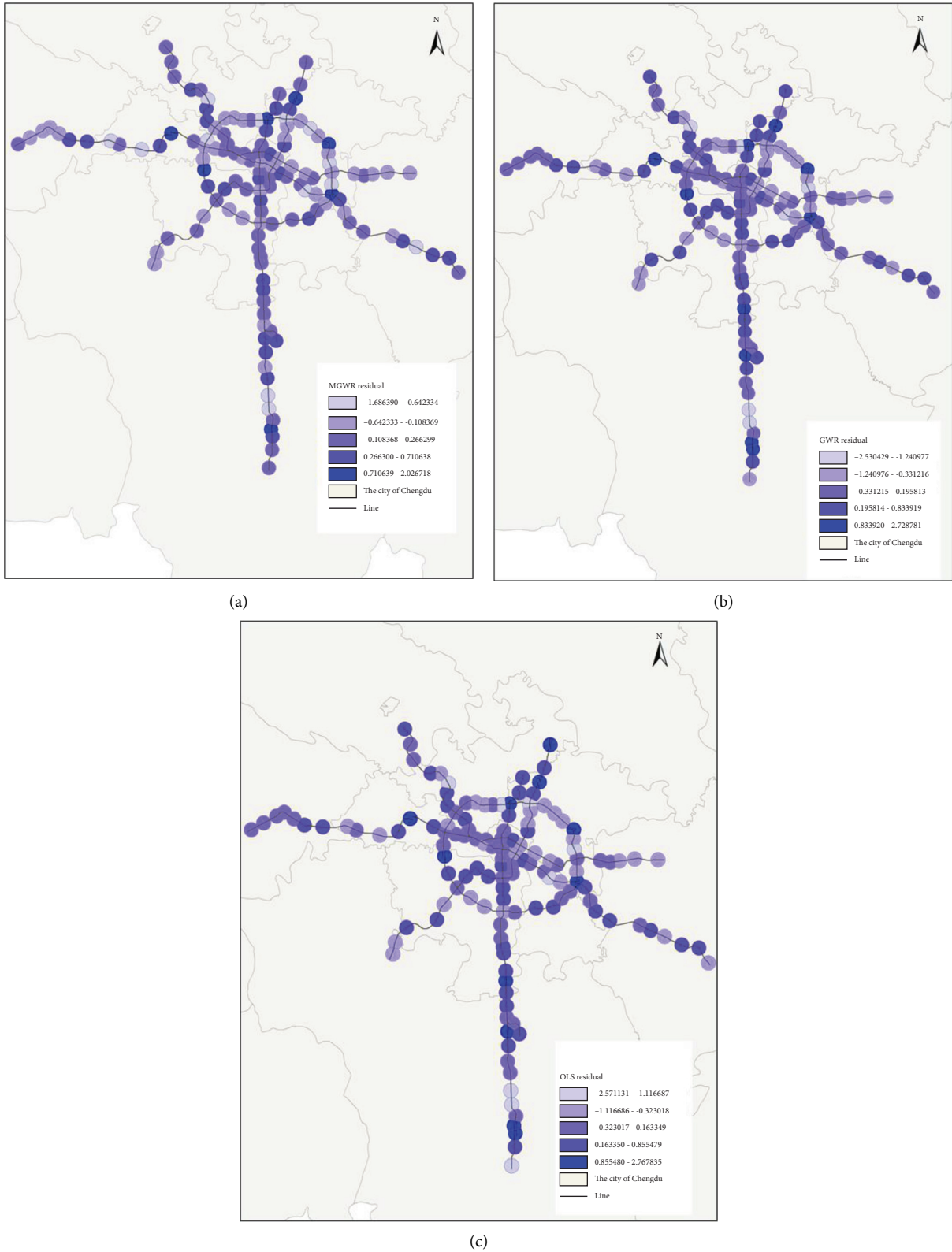


FIGURE 11: (a–c) Spatial distribution of the residuals of the MGWR, GWR, and OLS models.

land use, such as commercial POIs can replace commercial land use, and living class POIs can replace population density [5]. According to Figures 15(c) and 15(d), the

variable *Bus* expresses a positive correlation with the *Passenger-distance*. The variable *Life* has a negative correlation effect on most subway stations, but it has a positive

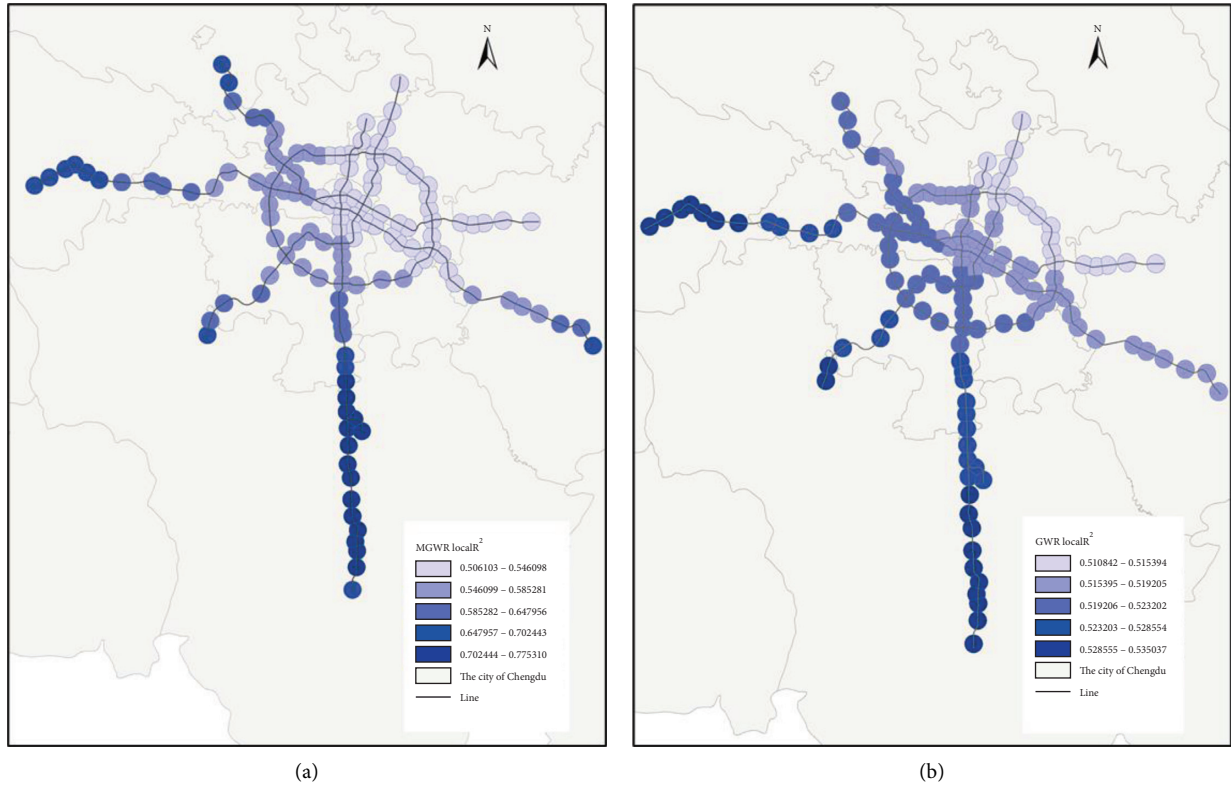


FIGURE 12: (a, b) Distribution of MGWR model local  $R^2$  and GWR model local  $R^2$ .

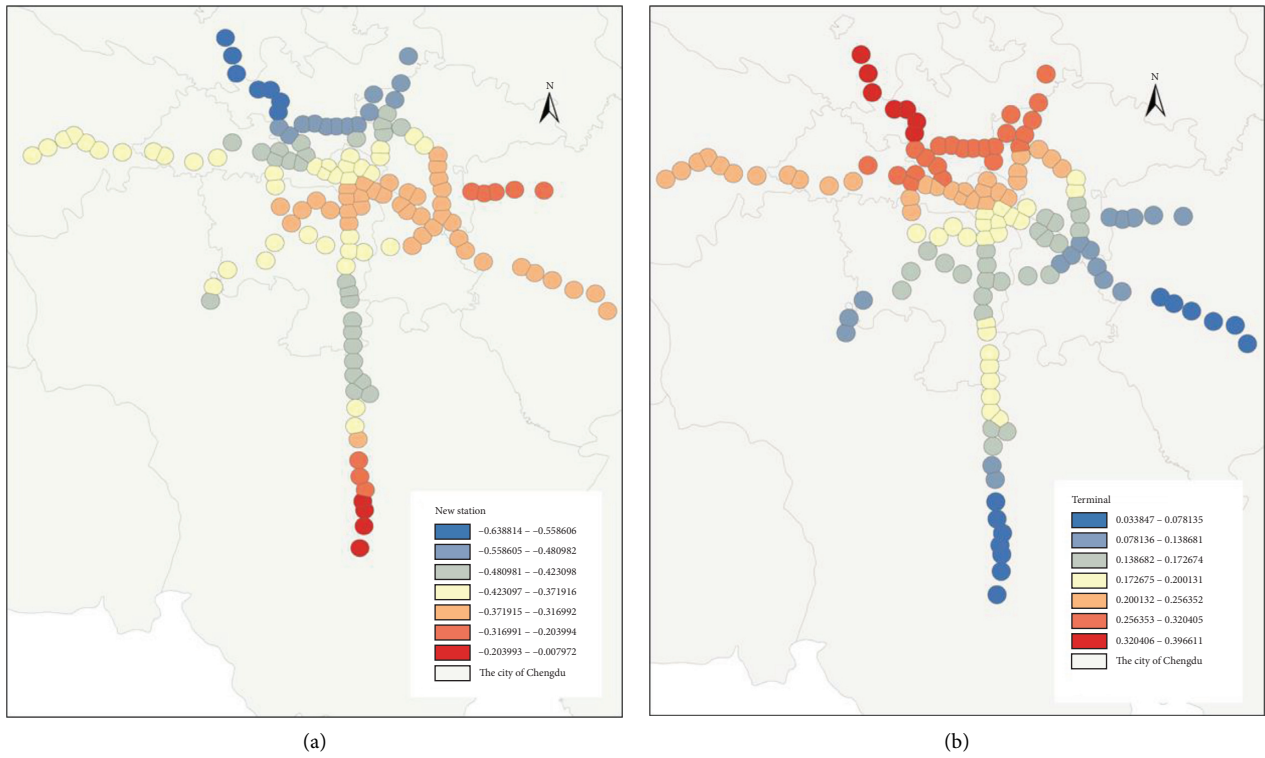
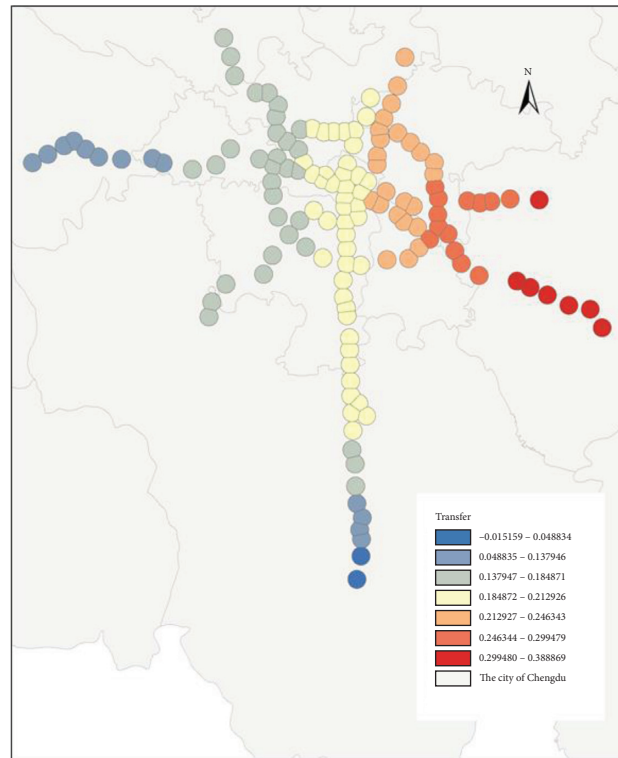
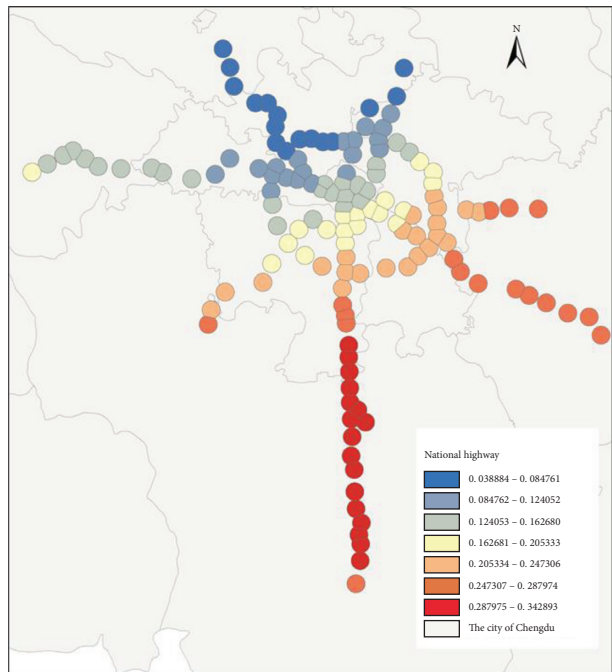


FIGURE 13: Continued.

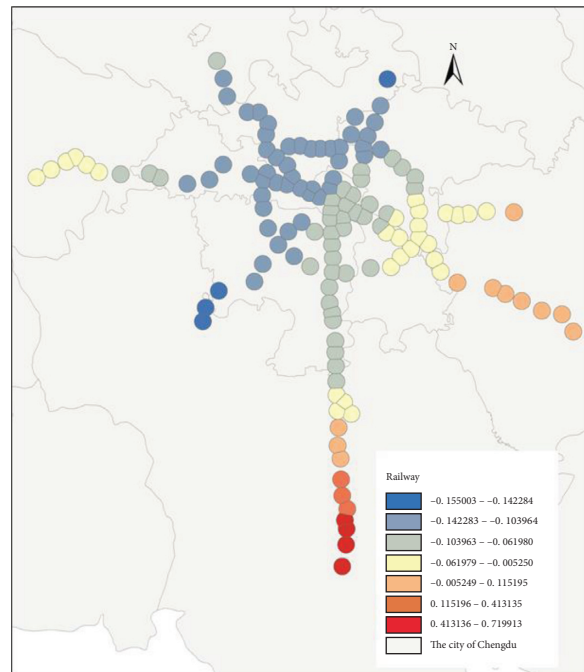


(c)

FIGURE 13: (a–c) Spatial distribution of three variables, *new station*, *terminal*, *transfer*, etc. in the MGWR model.

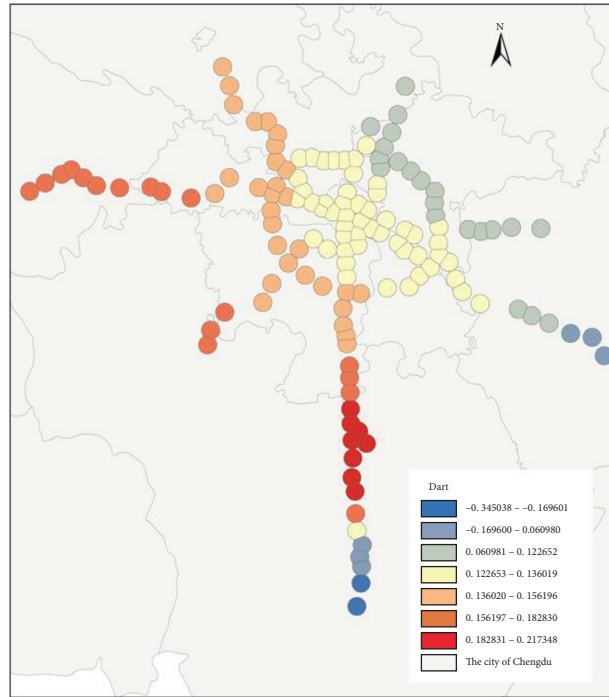


(a)



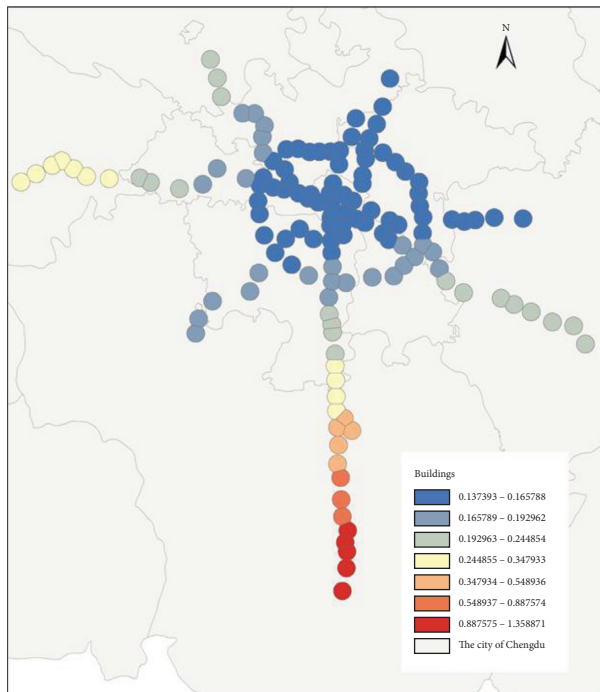
(b)

FIGURE 14: Continued.

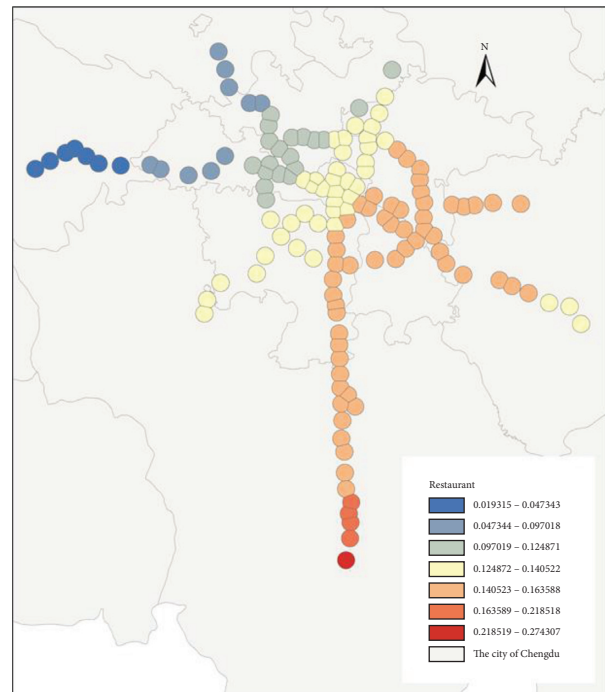


(c)

FIGURE 14: (a-c) Spatial distribution of the coefficients of the three variables: national highway, railway, and dart in the MGWR model, respectively.



(a)



(b)

FIGURE 15: Continued.

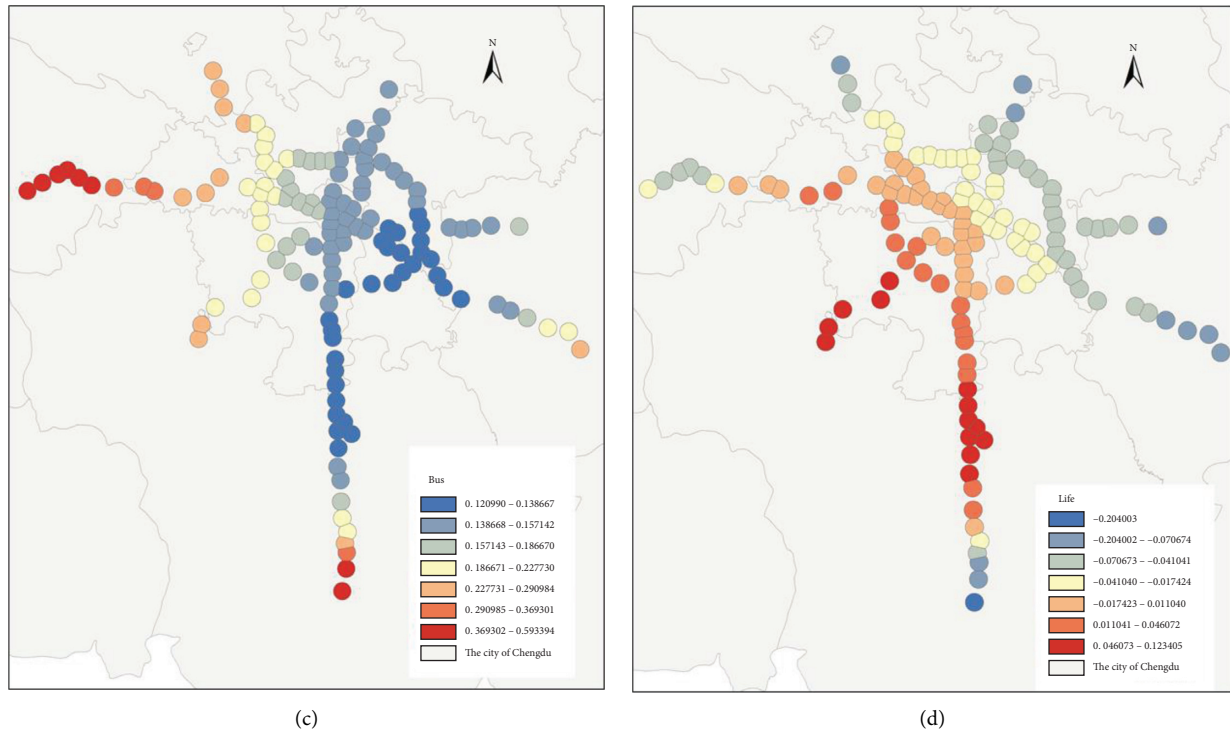


FIGURE 15: (a–d) Spatial distribution of the coefficients of the four variables of *buildings*, *restaurant*, *bus*, and *life* in the MGWR model, respectively.

correlation relationship with most subway stations in the High-tech Zone (southern region of Chengdu) and the Shuangliu District (southwest of Chengdu). The variable *Restaurant* shows a different spatial representation from the first two variables. The coefficients are higher in the eastern region, indicating that the passenger mileage and the restaurant in these regions have a strong positive correlation. In addition, the variables *Buildings* and *Passenger-distance* are both positively correlated. The difference in space is that they have a greater impact on the western and southern regions, compared to the central and eastern regions.

## 6. Conclusion

**6.1. Summary.** This paper focuses on the spatial heterogeneity between the built environment and subway passenger flow. Taking subway passenger-distance as an example, the spatial relationship between *Passenger-distance* and other key built environment factors has been investigated. To study the spatial heterogeneity of ridership at the subway station level, three models were developed for analysis and comparison, namely, the OLS model, the GWR model, and the MGWR model. Our results show that MGWR model has smaller AIC, AICc, and CV values than OLS and GWR models, while the goodness-of-fit (i.e.,  $R^2$ ) is greater. Therefore, the MGWR model, which considers both global and local variables, has a better fit to explore spatial heterogeneity through our data. The MGWR model also reveals that the relationship between the built environment variables and SLPD exhibits spatial heterogeneity.

According to the OLS model, among the eleven significant variables at the 0.1 significance level, *Terminal*, *Transfer*, *Accessibility*, *Bus*, *Restaurant*, *Building*, *National Highway*, and *Dart* are positively associated with subway SLPD, indicating that as the values of these built environment variables increase, the *Passenger-distance* of subway stations increases. In the GWR model, these 11 significant variables are all treated as local variables. Through analysis, the MGWR model considers *Terminal*, *Transfer*, *Bus*, *Restaurant*, *Building*, *Life*, *National Highway*, *Dart* and *Railway* as local variables, and *Accessibility* as a global variable. These findings show that the MGWR model fits the data better.

In addition, some policy suggestions can be proposed based on the results of the MGWR model. First, for the subway stations located on the south side of Chengdu, government agencies could increase the density of the buildings around the subway station, build denser national highways, and more bus stations to increase the accessibility and the attractiveness of the subway stations. Second, for those stations located in the southwest of Chengdu, as the number of life-related POIs is positively related to the passenger distance, the government agencies are advised to invest more life-related POIs near the stations, in consistency with the TOD strategy. Third, for the stations on the east side of Chengdu, since the passenger distance is positively related to the number of restaurants, more efforts should be spent on enriching the number and diversity of the restaurants to encourage longer travel distance on subway by the passengers. Finally, since the density of provincial highways has a significantly positive impact on the passenger

distance of subway stations in western Chengdu, government agencies could consider expanding these provincial links to increase the demand for public transportation. Although this study mainly focuses on Chengdu, China, the research framework and methodology of this paper are applicable to other cities as well. Specifically, other studies should consider the passenger-distance as the dependent variable and the use of MGWR to explore the spatially varying relationship between the station level passenger-distance and determinants.

**6.2. Limitations and Future Research.** We acknowledge several limitations of this study. First, due to the lack of data on population density and regional economy, these data were omitted from the model, which may cause some biases in the results. Second, regarding the buffer selection of subway station catchment area, a circular buffer zone with an 800 m radius and the Tyson polygons are used. This selection may result in some biases especially in the city center where many stations are densely distributed with smaller coverage areas. Despite these limitations, this study systematically explores the spatial variation patterns of subway SLPD and investigates the influence of the built environment on their spatial heterogeneity. Meanwhile, we provide some suggestions on ridership growth and management. However, although the MGWR model could be used to explore the spatially varying relationship between explanatory variables and the response variable, it cannot take into consideration the temporal dimension of the data. This is a potential direction that future work can lean towards. Models that consider both spatial and temporal scales of the relationship between the built environment and the station-level passenger distance may yield more accurate results. Another limitation of this study is that the circular buffer is used as an approximation of the actual access and egress distance. In reality, the territorial discontinuity and barriers could influence the access and egress distance, which will further influence travelers' behaviors [66]. In the future, researchers should consider the territorial discontinuity and barriers around stations to obtain more accurate estimation of the actual access and egress distance.

### Data Availability

The data used to support the findings of this study are available from OpenStreetMap and Amap: <https://www.openstreetmap.org/#map=12/51.0611/-114.1304> and <https://www.amap.no/data>.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgments

This study was funded by the National Natural Science Foundation of China (grant nos. 71704145, 51774241, and 71831006), Humanity and Social Science Foundation of Ministry of Education of China (grant no. 18YJCZH138),

and China Postdoctoral Science Foundation, Sichuan Youth Science and Technology Innovation Research Team Project (grant nos. 2019JDTD0002 and 2020JDTD0027).

### References

- [1] J. Yang, S. Chen, P. Qin, F. Lu, and A. A. Liu, "The effect of subway expansions on vehicle congestion: evidence from Beijing," *Journal of Environmental Economics and Management*, vol. 88, pp. 114–133, 2018.
- [2] P. Li, Y. Lu, and J. Wang, "The effects of fuel standards on air pollution: evidence from China," *Journal of Development Economics*, vol. 146, Article ID 102488, 2020.
- [3] S. Jiang, W. Guan, Z. He, and L. Yang, "Exploring the intermodal relationship between taxi and subway in Beijing, China," *Journal of Advanced Transportation*, vol. 2018, Article ID 3981845, 2018.
- [4] H. Wei, "Impacts of China's national vehicle fuel standards and subway development on air pollution," *Journal of Cleaner Production*, vol. 241, Article ID 118399, 2019.
- [5] E. Chen, Z. Ye, C. Wang, and W. Zhang, "Discovering the spatio-temporal impacts of built environment on metro ridership using smart card data," *Cities*, vol. 95, Article ID 102359, 2019.
- [6] C. Liu and L. Li, "How do subways affect urban passenger transport modes?-Evidence from China," *Economics of Transportation*, vol. 23, Article ID 100181, 2020.
- [7] A. Nasri and L. Zhang, "The analysis of transit-oriented development (TOD) in Washington, D.C. and Baltimore metropolitan areas," *Transport Policy*, vol. 32, pp. 172–179, 2014.
- [8] H. Pan, J. Li, Q. Shen, and C. Shi, "What determines rail transit passenger volume? Implications for transit oriented development planning," *Transportation Research Part D: Transport and Environment*, vol. 57, pp. 52–63, 2017.
- [9] R. Sharma and P. Newman, "Urban rail and sustainable development key lessons from Hong Kong, New York, London and India for emerging cities," *Transportation Research Procedia*, vol. 26, pp. 92–105, 2017.
- [10] R. Cervero and K. Kockelman, "Travel demand and the 3Ds: density, diversity, and design," *Transportation Research Part D: Transport and Environment*, vol. 2, no. 3, pp. 199–219, 1997.
- [11] R. Ewing and R. Cervero, "Travel and the built environment," *Journal of the American Planning Association*, vol. 76, no. 3, pp. 265–294, 2010.
- [12] R. Ewing and R. Cervero, "Travel and the built environment: a synthesis," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1780, no. 1, pp. 87–114, 2001.
- [13] M.-J. Jun, K. Choi, J.-E. Jeong, K.-H. Kwon, and H.-J. Kim, "Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul," *Journal of Transport Geography*, vol. 48, pp. 30–40, 2015.
- [14] H. Sung, K. Choi, S. Lee, and S. Cheon, "Exploring the impacts of land use by service coverage and station-level accessibility on rail transit ridership," *Journal of Transport Geography*, vol. 36, pp. 134–140, 2014.
- [15] J. Zhao, W. Deng, Y. Song, and Y. Zhu, "What influences Metro station ridership in China? Insights from Nanjing," *Cities*, vol. 35, pp. 114–124, 2013.
- [16] J. Zhao, W. Deng, Y. Song, and Y. Zhu, "Analysis of Metro ridership at station level and station-to-station level in



- Nanjing: an approach based on direct demand models,” *Transportation*, vol. 41, no. 1, pp. 133–155, 2014.
- [17] H. Yang, Y. Zhang, L. Zhong, X. Zhang, and Z. Ling, “Exploring spatial variation of bike sharing trip production and attraction: a study based on Chicago’s Divvy system,” *Applied Geography*, vol. 115, Article ID 102130, 2020.
- [18] H. Yang, T. Xu, D. Chen, H. Yang, and L. Pu, “Direct modeling of subway ridership at the station level: a study based on mixed geographically weighted regression,” *Canadian Journal of Civil Engineering*, vol. 47, no. 5, pp. 534–545, 2019.
- [19] J. J. Lin and C. C. Gau, “A TOD planning model to review the regulation of allowable development densities around subway stations,” *Land Use Policy*, vol. 23, no. 3, pp. 353–360, 2006.
- [20] S. Lee, C. Yi, and S.-P. Hong, “Urban structural hierarchy and the relationship between the ridership of the seoul metropolitan subway and the land-use pattern of the station areas,” *Cities*, vol. 35, pp. 69–77, 2013.
- [21] F. Russo, P. Panuccio, and C. Rindone, “Structural factors for a third-generation port: between hinterland regeneration and smart town in Gioia Tauro, Italy,” *Urban Maritime Transport XXVII*, vol. 204, pp. 79–90, 2021.
- [22] F. Russo and G. Chilà, “Structural factors for a third-generation port: actions and measures for Gioia Tauro, Italy, in the regional transport plan,” *Urban Maritime Transport XXVII*, vol. 204, pp. 17–30, 2021.
- [23] G. Musolino, C. Rindone, and A. Vitetta, “Passengers and freight mobility with electric vehicles: a methodology to plan green transport and logistic services near port areas,” *Transportation Research Procedia*, vol. 37, pp. 393–400, 2019.
- [24] S. L. Handy, M. G. Boarnet, R. Ewing, and R. E. Killingsworth, “How the built environment affects physical activity,” *American Journal of Preventive Medicine*, vol. 23, no. 2, pp. 64–73, 2002.
- [25] J. Huo, H. Yang, C. Li, R. Zheng, L. Yang, and Y. Wen, “Influence of the built environment on E-scooter sharing ridership: a tale of five cities,” *Journal of Transport Geography*, vol. 93, Article ID 103084, 2021.
- [26] H. Yang, Y. Liang, and L. Yang, “Equitable? Exploring ridesourcing waiting time and its determinants,” *Transportation Research Part D: Transport and Environment*, vol. 93, Article ID 102774, 2021.
- [27] H. Yang, G. Zhai, L. Yang, and K. Xie, “How does the suspension of ride-sourcing affect the transportation system and environment?” *Transportation Research Part D: Transport and Environment*, vol. 102, Article ID 103131, 2022.
- [28] C. Ding, X. Cao, and C. Liu, “How does the station-area built environment influence metrorail ridership? Using gradient boosting decision trees to identify non-linear thresholds,” *Journal of Transport Geography*, vol. 77, pp. 70–78, 2019.
- [29] S. Liu, E. Yao, and B. Li, “Exploring urban rail transit station-level ridership growth with network expansion,” *Transportation Research Part D: Transport and Environment*, vol. 73, pp. 391–402, 2019.
- [30] J. Yang, J. Cao, and Y. Zhou, “Elaborating non-linear associations and synergies of subway access and land uses with urban vitality in Shenzhen,” *Transportation Research Part A: Policy and Practice*, vol. 144, pp. 74–88, 2021.
- [31] R. Du and S. Zheng, “Agglomeration, housing affordability, and new firm formation: the role of subway network,” *Journal of Housing Economics*, vol. 48, Article ID 101668, 2020.
- [32] S. Li, D. Lyu, G. Huang et al., “Spatially varying impacts of built environment factors on rail transit ridership at station level: a case study in Guangzhou, China,” *Journal of Transport Geography*, vol. 82, Article ID 102631, 2020.
- [33] C. Chen, J. Chen, and J. Barry, “Diurnal pattern of transit ridership: a case study of the New York City subway system,” *Journal of Transport Geography*, vol. 17, no. 3, pp. 176–186, 2009.
- [34] H. Iseki, C. Liu, and G. Knaap, “The determinants of travel demand between rail stations: a direct transit demand model using multilevel analysis for the Washington D.C. metrorail system,” *Transportation Research Part A: Policy and Practice*, vol. 116, pp. 635–649, 2018.
- [35] K. Sohn and H. Shim, “Factors generating boardings at Metro stations in the Seoul metropolitan area,” *Cities*, vol. 27, no. 5, pp. 358–368, 2010.
- [36] X. Ma, J. Zhang, C. Ding, and Y. Wang, “A geographically and temporally weighted regression model to explore the spatiotemporal influence of built environment on transit ridership,” *Computers, Environment and Urban Systems*, vol. 70, pp. 113–124, 2018.
- [37] S. Zheng, X. Zhang, W. Sun, and J. Wang, “The effect of a new subway line on local air quality: a case study in Changsha,” *Transportation Research Part D: Transport and Environment*, vol. 68, pp. 26–38, 2019.
- [38] R. Trojanek and M. Gluszak, “Spatial and time effect of subway on property prices,” *Journal of Housing and the Built Environment*, vol. 33, no. 2, pp. 359–384, 2018.
- [39] S. Hamidi and I. Hamidi, “Subway ridership, crowding, or population density: determinants of COVID-19 infection rates in New York city,” *American Journal of Preventive Medicine*, vol. 60, no. 5, pp. 614–620, 2021.
- [40] S. Wang and R. B. Noland, “Variation in ride-hailing trips in Chengdu, China,” *Transportation Research Part D: Transport and Environment*, vol. 90, Article ID 102596, 2021.
- [41] S. Kim, S. Park, and K. Jang, “Spatially-varying effects of built environment determinants on walking,” *Transportation Research Part A: Policy and Practice*, vol. 123, pp. 188–199, 2019.
- [42] Z. Shi, N. Zhang, Y. Liu, and W. Xu, “Exploring spatio-temporal variation in hourly metro ridership at station level: the influence of built environment and topological structure,” *Sustainability*, vol. 10, no. 12, 4564 pages, 2018.
- [43] M. J. T. L. Gomes, F. Cunto, and A. R. da Silva, “Geographically weighted negative binomial regression applied to zonal level safety performance models,” *Accident Analysis & Prevention*, vol. 106, pp. 254–261, 2017.
- [44] 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.
- [45] S. Harini, “Estimation of error variance-covariance parameters using multivariate geographically weighted regression model,” *Abstract and Applied Analysis*, vol. 2020, Article ID 4657151, 2020.
- [46] O. D. Cardozo, J. C. García-Palomares, and J. Gutiérrez, “Application of geographically weighted regression to the direct forecasting of transit ridership at station-level,” *Applied Geography*, vol. 34, pp. 548–558, 2012.
- [47] H. Yang, X. Li, C. Li, J. Huo, and Y. Liu, “How do different treatments of catchment area affect the station level demand modeling of urban rail transit?” *Journal of Advanced Transportation*, vol. 2021, Article ID 2763304, 2021.
- [48] S. Chan and L. Miranda-Moreno, “A station-level ridership model for the metro network in montreal, Quebec,” *Canadian Journal of Civil Engineering*, vol. 40, no. 3, pp. 254–262, 2013.

- [49] H. Yang, C. Li, X. Li et al., "Effects of coverage area treatment, spatial analysis unit, and regression model on the results of station-level demand modeling of urban rail transit," *Journal of Advanced Transportation*, vol. 2021, Article ID 7345807, 10 pages, 2021.
- [50] R. Grannis, "The importance of trivial streets: residential streets and residential segregation," *American Journal of Sociology*, vol. 103, no. 6, pp. 1530–1564, 1998.
- [51] S. Z. Stamato, A. J. Park, B. Eng, V. Spicer, H. H. Tsang, and D. K. Rossmo, "Differences in Geographic Profiles when Using Street Routing versus Manhattan Distances in Buffer Zone Radii Calculations," in *Proceedings of the 2021 IEEE International Conference on Intelligence and Security Informatics (ISI)*, pp. 1–6, San Antonio, TX, USA, November 2021.
- [52] J. Y. Guo and C. R. Bhat, "Operationalizing the concept of neighborhood: application to residential location choice analysis," *Journal of Transport Geography*, vol. 15, no. 1, pp. 31–45, 2007.
- [53] P. M. Schirmer, M. A. B. van Eggermond, and K. W. Axhausen, "The role of location in residential location choice models: a review of literature," *Journal of Transport and Land Use*, vol. 7, no. 2, pp. 3–21, 2014.
- [54] S. Jäppinen, T. Toivonen, and M. Salonen, "Modelling the potential effect of shared bicycles on public transport travel times in Greater Helsinki: an open data approach," *Applied Geography*, vol. 43, pp. 13–24, 2013.
- [55] S. Jiang, A. Alves, F. Rodrigues, J. Ferreira, and F. C. Pereira, "Mining point-of-interest data from social networks for urban land use classification and disaggregation," *Computers, Environment and Urban Systems*, vol. 53, pp. 36–46, 2015.
- [56] J. A. Pooler, "The use of spatial separation in the measurement of transportation accessibility," *Transportation Research Part A: Policy and Practice*, vol. 29, no. 6, pp. 421–427, 1995.
- [57] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, Oakland, CA, USA, January 1967.
- [58] R. P. J. Silhouettes, "A graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.
- [59] M. Munizaga, F. Devillaine, C. Navarrete, and D. Silva, "Validating travel behavior estimated from smartcard data," *Transportation Research Part C: Emerging Technologies*, vol. 44, pp. 70–79, 2014.
- [60] H. Yang, G. Zhai, X. Liu, L. Yang, Y. Liu, and Q. Yuan, "Determinants of city-level private car ownership: effect of vehicle regulation policies and the relative price," *Transport Policy*, vol. 115, pp. 40–48, 2022.
- [61] C. Brunson, 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, 1996.
- [62] A. S. Fotheringham, C. Brunson, and M. Charlton, *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, John Wiley & Sons, Hoboken, NJ, USA, 2003.
- [63] J. Neter, W. Wasserman, and M. H. Kutner, *Applied Linear Regression Models*, Technometrics, Ahmedabad, India, 1989.
- [64] P. A. P. Moran, "Notes on continuous stochastic phenomena," *Biometrika*, vol. 37, no. 1/2, pp. 17–23, 1950.
- [65] B. Huang, B. Wu, and M. Barry, "Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices," *International Journal of Geographical Information Science*, vol. 24, no. 3, pp. 383–401, 2010.
- [66] G. Birgillito, C. Rindone, and A. Vitetta, "Passenger mobility in a discontinuous space: modelling access/egress to maritime barrier in a case study," *Journal of Advanced Transportation*, vol. 2018, Article ID 6518329, 13 pages, 2018.