Exploring the Effect of Built Environment Factors on Metro Station Ridership during the Holiday Season – A Case Study of the Beijing Metro System during the Chinese National Day Holidays

Zhenbao WANG[1], Yanfang HE[2], Xueqiao ZHAO[3], Yuqi LIANG[4], Shihao LI[5]

Original Scientific Paper
Submitted: 21 May 2024
Accepted: 18 Sep 2024

[1] wangzhenbao@hebeu.edu.cn, Hebei University of Engineering, School of Architecture and Art
[2] 15580590955@163.com, Hebei University of Engineering, School of Architecture and Art
[3] zhaoxueqiao@hebeu.edu.cn, Hebei University of Engineering, School of Architecture and Art
[4] yq18233082796@163.com, Hebei University of Engineering, School of Architecture and Art
[5] li15294865139@163.com, Hebei University of Engineering, School of Architecture and Art

Publisher:
Faculty of Transport and Traffic Sciences, University of Zagreb

ABSTRACT
Previous studies have primarily focused on the effect of the built environment on ridership during weekdays and weekends. This paper aims to investigate the spatial heterogeneity of the effect of built environment factors on ridership at metro stations during National Day holidays. Beijing is divided into three zones from inner to outer areas. Taking metro station boarding and alighting ridership during National Day holidays as the dependent variable, 13 built environment factors were selected as independent variables according to the "7D" dimension of the built environment. The recommended pedestrian catchments (PCA) combinations for the three zones in Beijing are 400 m_500 m_400 m by using the Multi-Scale Geographically Weighted Regression (MGWR) model. We investigated the effect of built environment factors on metro ridership and spatial heterogeneity. The influencing factors that have significant effects on both boarding and alighting ridership are building density, number of commercial facilities, bus lines density, number of entrance and exit, number of office facilities, mixed utilization of land and road density. The MGWR model results are helpful to propose targeted strategies for revitalising the built environment around metro stations.

KEYWORDS
metro station ridership, pedestrian catchment areas (PCA); multi-scale geographically weighted regression (MGWR); built environment; National Day holidays.

1. INTRODUCTION

The holiday travel has become a popular way for people to relax and unwind in recent years, and the travel demand is increasing during holidays. Beijing, being a famous tourist city in China, has a large number of tourists, and this has significantly affected the city's travel patterns. The city's mode of travel is also gradually shifting from car travel to metro transit travel [1, 2]. In accordance with the consensus on transit-oriented development (TOD), the urban transportation and planning departments actively promote the rapid development of the metro transit system [3]. TOD greatly promotes the city's green travel and ensures sustainable land use development. In the research on travel behaviour, it is found that behaviours such as urban travel behaviour and leisure consumption are highly related to holidays [4, 5]. Undoubtedly, people's travel during holidays plays a significant role in the development of the urban economy and metro transportation [6]. In China, the National Day holidays is listed as one of the longest holidays of the year. The Beijing economy's robust vigour was showcased with a flowing population, booming consumption and busy construction during National Day holidays. Compared with weekdays or weekends, metro ridership, bus and high-speed railway shows an increasing trend during the National Day holidays. During the National Day holidays, the average daily ridership in Beijing metro reaches over 5 million [7], far higher than the average daily ridership of 1 million on weekdays [8]. This indicates that Beijing's public transport system will face greater pressure during the National Day holidays. In daily operation, when the Beijing metro system faces the large ridership, some stations will take measures such as temporary traffic restriction and fare increase according to the situation. Therefore, when Beijing metro meets holidays or activities, this paper takes into account a series of flexible measures to adjust the Beijing metro ridership.

However, the current analysis of metro transit ridership primarily focuses on data from weekdays and weekends. Due to the differences in travel patterns between non-holidays (weekdays) and National Day holidays [9, 10], tourists were more time-sensitive during holidays. Fast and schedule-accurate [11] metro transit helped travellers plan their trips or activities. The metro transit experiences increased the burden during holidays, yet few studies have delved into the relationship between built environment factors and metro ridership during these periods [12]. Therefore, it is necessary to explore affecting metro ridership in the built environment factors around metro stations during holidays in Beijing, which helps predict passenger demand and develop targeted urban transportation management strategies.

Previous scholars have mainly used global regression models such as Ordinary Least Squares (OLS) [3, 13], Two Stage Least Square (2SLS) [14], Spatial error model (SEM) [15, 16], and Spatial lag model (SLM) [16], to explore the relationship between metro station-level ridership and the built environment factors. The global model was found to calculate coefficients that were not significant differences in space [17]. The Geographically Weighted Regression (GWR) model [18–20], as a local regression model based on geographical location, takes into account the spatial variability of the factors influencing metro ridership. The model was proved to have better fitting results compared to the OLS model and can reveal the factors of spatial heterogeneity in these studies [17, 21, 22]. However, the GWR model assumes that all processes are limited to operate on the same spatial scale. This limitation of the GWR model was removed by the proposal of Multi-Scale Geographically Weighted Regression (MGWR) [23, 24], which allows each influencing factor to have its optimal bandwidth, permitting different processes to be operated at different spatial scales. The MGWR model is widely used in analysing the spatial variability of residential land prices [25], the evolution of the spatial pattern of environmental and scientific factors [26, 27], and other domains. Some scholars [28, 29] applied the MGWR model to the relationship between the built environment factors and the ridership during the weekdays and weekends, but it lacks studying the ridership during holidays.

Influencing factors mainly included environmental factors, socio-economic factors and transportation environment factors. Most scholars [3, 15, 30] selected the influencing factors based on previous research results or according to their insights. Some scholars have tried to systematically select influencing factors for evaluating the built environment. Cervero and Kockelman [31] have proposed "density", "diversity" and "design" as the principles for evaluating the built environment, namely the "3D" dimension of the built environment. Later, Cervero et al. [32] added two additional dimensions: "transportation distance" and "destination accessibility," thus proposing the "5D" dimension of the built environment. Since then, many studies have used variables related to the "5D" dimension of the built environment to analyse the effect of the built environment on urban metro ridership [13, 33, 34]. Therefore, the "5D" dimension of the built environment has developed relatively maturely for a while. In recent years, Chris De Gruyter et al. [35] added the "demand management" and "demographic" dimensions to the "5D" dimension of the built environment and proposed the "7D" dimension of the built environment. Therefore, selecting the built environment influencing factors based on the built environment "7D" dimension [28] may be more helpful to comprehensively analyse the built environment's effect on metro ridership.

The station's catchment area was typically determined by the walking distance for most pedestrians accessed to metro stations, so this area was also called Pedestrian Catchment Areas (PCA) [3, 36, 37]. Metro station PCA was mostly determined based on pedestrian accessibility [38], drawing on previous studies [39] and comparing the regression goodness-of-fit [28]. Metro station PCA shapes were mostly circular buffer zones [40], Thiessen polygons or Thiessen polygons overlapping with circular buffer zones [28]. In addition, Transit Oriented Development (TOD) was analysed for stations with a circular buffer centred on the public station. In the area of dense metro station distribution, metro station PCA will overlap, but the overlapping areas may all have an effect on the ridership. Therefore, many circular buffers of different sizes were used in the literature, with radius ranging from 250 metres to 1000 metres [3, 15, 16, 40], with a wide range of sizes. There may still be significant differences in urban transit usage in different areas of the same city, and the distribution of metro stations is also different. Some scholars divided the city into three zones and used a uniform PCA for each zone to analyse the metro stations ridership [18], but did not determine different recommended metro station PCAs according to different zones. In the mega-city of Beijing, there is an imbalance in the distribution of ridership [28], and the density of inner and outer peripheral metro stations will be different. Therefore, in this study, the circular buffer zone is chosen as the metro station PCA shape, and it is necessary to divide the city into different zones and use different metro station PCA to analyse the effect of the built environment on metro stations, which may improve the explanatory power of the model.

This study contributes in three ways. Empirically, no study has explored the effect of the built environment on metro ridership during the National Day holidays. We use the MGWR model to analyse the effect degree and spatial heterogeneity of the built environment on metro ridership. Methodologically, although a few studies have considered different zones to use different PCA to collect the built environment influencing factors, some regression models have been used to explore the global and local relationships between the metro ridership and influencing factors. However, they did not determine the best PCA combination for different areas to make the regression model fit the best. Finally, although some studies point out that certain built environment influencing factors are correlated with the metro ridership, there is a lack of targeted built environment renewal strategies for specific station. To address these problems, using the Beijing metro transit as a case study, we identified the influencing factors based on the "7D" dimension of the built environment. We utilised the MGWR model to investigate the effect of the built environment on metro station ridership during National Day holidays. The main objectives are: (1) to determine the PCA combinations of metro stations with the best goodness of fit of the regression mode; (2) to study the differences in the effect of the built environment factors on metro ridership during the National Day holidays; (3) to propose built environment renewal strategies for low-vitality metro stations. With the goal of improving the accuracy of ridership prediction at metro stations during the National Day holidays, the results have a significant reference value for zoning the metro stations PCA, proposing targeted built environment renewal strategies and scientifically formulating ridership service and management plans.

2. DATA SOURCES

2.1 National Day holidays ridership data

All 291 metro stations that were been built and in service were taken into consideration during data collection in 2020. *Figure 1* shows all the metro stations in Beijing, and the colours on the map show the range of three zones: the area within the third ring road (red), the area between the third ring road and the fifth ring road (yellow), and the area outside the fifth ring road (white). We obtained hourly boarding and alighting ridership data for the Beijing Metro during the week from 1 October 2020 to 7 October 2020, as well as weekday ridership from 12 October 2020 to 16 October 2020. There are two distinct peak periods for boarding and alighting ridership on non-holidays (weekdays) [28]. However, no obvious peak hours were found in the trend analysis of hourly boarding and alighting ridership during the holidays (*Figure 2*). The total metro ridership on holidays is much higher than the weekdays ridership. It was finally determined that the study periods for both boarding and alighting ridership during the holidays were from 7:00–20:00. To minimise the effect of daily traffic fluctuations, we took the average of the seven-day boarding and alighting ridership as the dependent variable, respectively. *Figure 3* show the spatial distribution of boarding and alighting ridership at 291 metro stations. Metro stations in blue indicate stations with low ridership and low vitality.

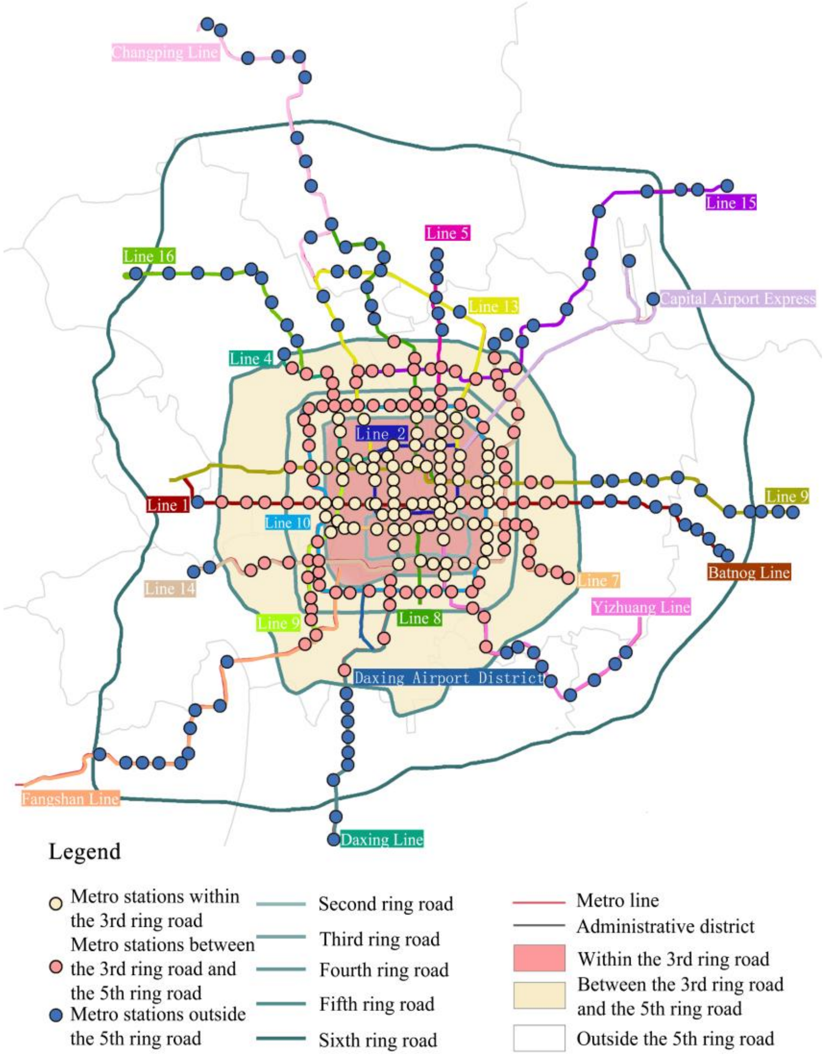


Figure 1 – Beijing metro stations distribution map

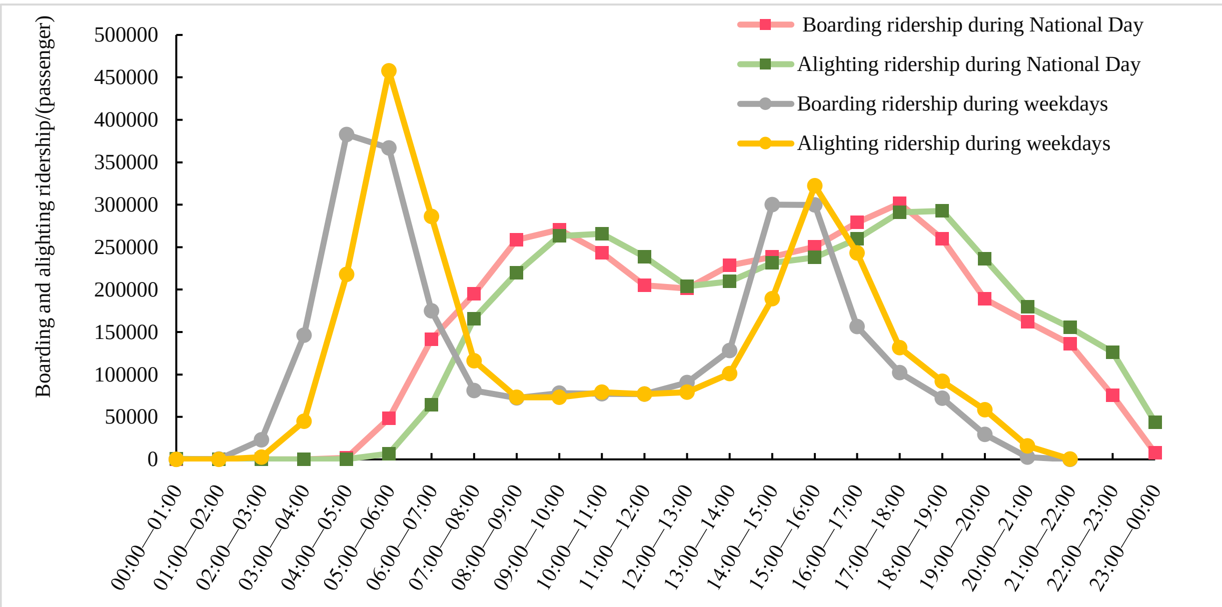


Figure 2 – Boarding and alighting ridership of metro stations during National Day holidays and weekdays in Beijing

Figure 3 – Spatial distribution of metro stations during National Day holidays: a) The boarding ridership of metro Stations during National Day holidays in Beijing; b) The alighting ridership of metro stations during National Day holidays in Beijing

2.2 Influencing factors

According to the 7D dimension of the built environment, we constructed the index system of influencing factors (*Table 1*). The road network data, building outline data and building height data were obtained from the Open Street Map (https://www.openstreetmap.org/ (accessed on 10 October 2021)). Points of interest (POIs) were obtained from the Golder API (https://lbs.amap.com/ (accessed on 10 October 2021)). Population data was obtained from WorldPop (https://www.worldpop.org/) [41, 42]. POIs can better reflect the influencing factors of metro ridership and passenger's travel purposes [43, 13]. It is divided into three main land use types: the numberof commercial facilities (the number of shopping, restaurants, scenic, and hotel facilities), the number of office facilities (the number of financial and office facilities) and the number of public service facilities (the number of science, education, medical, government and transportation facilities).

Table 1 – Built environment influencing factors

Built environment dimension	Influencing factors	Unit
Density	Building density (Den. B)	m^2/km^2
Diversity	Mixed utilisation of land (MUL)	
Design	Road density (Den. R)	km/km^2
	Floor area ratio (FAR)	
Destination accessibility	Number of commercial facilities (Num.CF)	quantity
	Number of office facilities (Num. OF)	
	Number of public service facilities (Num. PSF)	
Distance to transit	Density of bus line (Den. BL)	km/km^2
	Number of entrance and exit (Num.EE)	
Demand management	Number of parking lots (Num.PL)	quantity
	Number of bus stops (Num.BS)	
Demographics	Population density (Den. P)	persons / km^2
	Resident population (RP)	persons

3. RESEARCH METHODOLOGY

3.1 Research framework

The research framework is shown in *Figure 4*. We employ the following five steps to explore the effect of the built environment on the metro ridership during the National Day holidays by using the MGWR model. These steps include: (1) data collection, (2) metro station zoning, (3) factors construction, (4) factors selection and (5) results. Firstly, the data is prepared based on multi-source big data. Secondly, Beijing is divided into three zones, and the PCA of each zone is determined according to the development of the metro station. Subsequently, 36 PCA combinations were identified. Thirdly, according to the "7D" dimension of the built environment, the influencing factors data set is constructed. Then, the multicollinearity test and spatial autocorrelation test were used to select the influential factors with high correlation. Finally, the PCA combination suitable for this study was determined by comparing the MGWR model results. The results are visualised to analyse the effects of the built environment on metro ridership and spatial heterogeneity, and to propose urban renewal strategies for low-vitality metro stations.

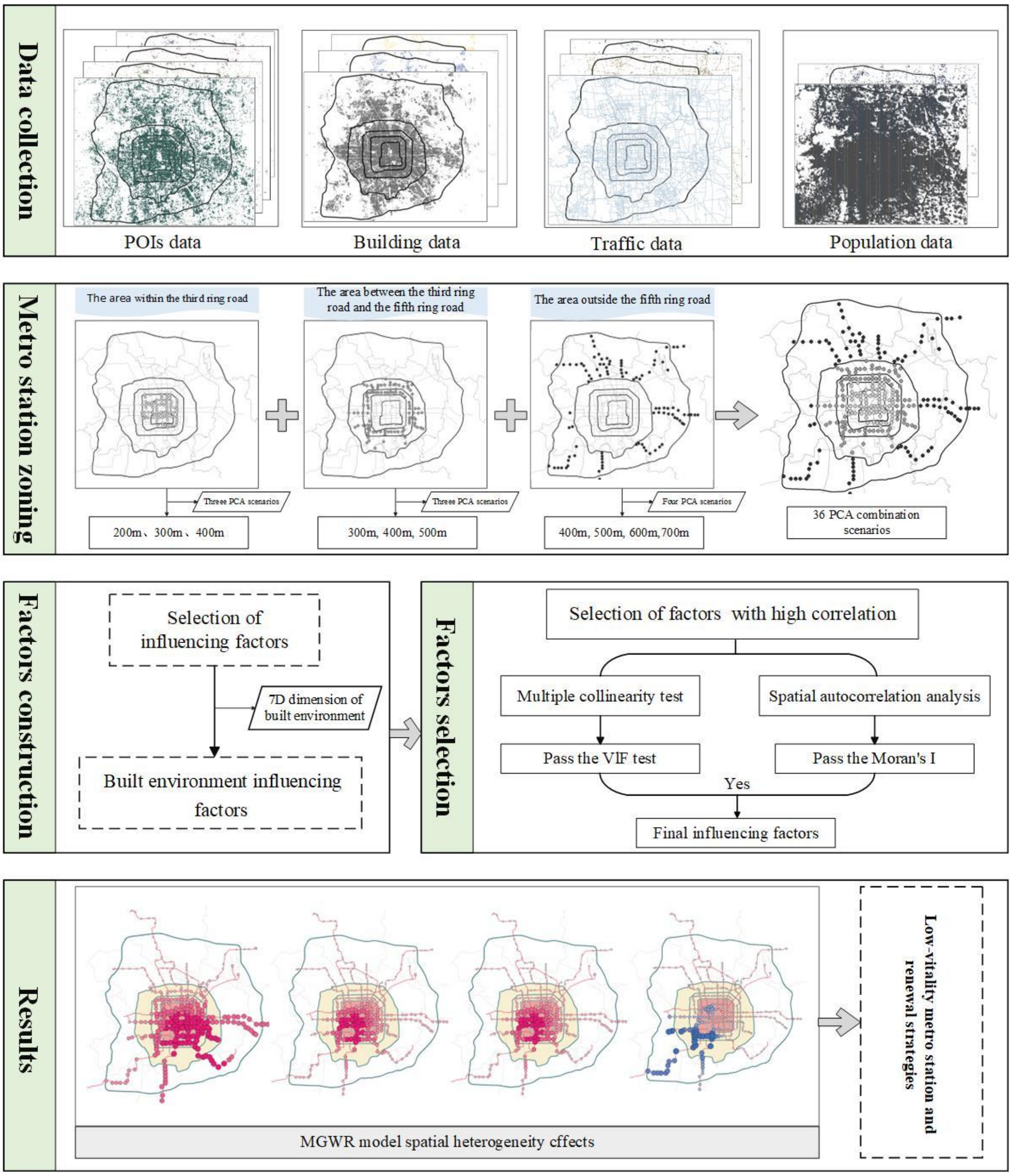


Figure 4 – Research framework

3.2 Delineation of metro station PCA during National Day holidays

A crucial step for using regression models to assess the factors influencing metro ridership is to define the catchment area of a station. Researchers usually focus on analysing metro station ridership using a uniform PCA in a city. In order to capture the variation in metro transit ridership by adopting different PCA sizes for different areas of a city, we divided Beijing into three zones based on the distribution of metro stations. The three zones are: the area within the third ring road, the area between the third ring road and the fifth ring road, and the area outside the fifth ring road. The distribution of metro stations within the third ring road is more intensive, and those between the third ring road and the fifth ring road are relatively uniform, while those outside the fifth ring road show a radial pattern, with a sparser distribution. Since on National Day holidays tourists prefer to stay in hotels near metro stations or attractions for the convenience of traveling, the tourists starting point and destination may overlap more around the metro station, resulting in a better fitting of the MGWR model for the smaller size PCA range of metro stations. Circular buffer zones with radii of 200 metres, 300 metres and 400 metres were selected for the metro stations within the third ring road, circular buffer zones with radii of 300 metres, 400 metres and 500 metres for the metro stations between the third ring road and the fifth ring road, and circular buffer zones with radii of 400 metres, 500 metres, 600 metres and 700 metres for the metro stations outside the fifth ring road. A total of 36 PCA combination scales were constructed.

3.3 Multiple collinearity test and spatial autocorrelation test of the National Day holidays influencing factors

Multiple collinearity test

Before fitting the regression model, there may be a high intercorrelation between the untested influencing factors. The 36 PCA combination scale models were tested for multicollinearity and the Variance Inflation Factor (VIF) was used as a measure of the multicollinearity severity of each influencing factor. A VIF value greater than 10 indicates that there is serious multicollinearity between the influence factor and others, which should be removed [44]. The calculation formula is as follows:

$$VIF = \frac{1}{1-R_i^2} \tag{1}$$

where R_i^2 is the coefficient of determination of the independent variable i.

Spatial autocorrelation analysis

Before building the spatial regression model, it is essential to determine whether all the influence factors have spatial autocorrelation. Moran's *I* is the correlation coefficient proposed by Patrick Alfred Pierce Moran (1950) for measuring spatial autocorrelation [45]. The calculation formula is as follows:

$$Moran's\ I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i-\bar{x})(x_j-\bar{x})}{S^2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \tag{2}$$

where S^2 is the variance; n is the total number of metro stations; x_i and x_j are the ridership of metro station i and metro station j, respectively; $\bar{x}$ is the mean value of all metro stations' ridership; and w_{ij} is the spatial weight between metro station *i* and *j*.

Multi-scale geographically weighted regression model

The multi-scale geographically weighted regression (MGWR) model adopted separate optimal bandwidths for each factor and produced more accurate local parameter values [24]. The calculation formula is as follows:

$$y_i^* = \beta_{bw0}(u_i,v_i) + \sum_{k=1}^{n}\beta_{bwk}(u_i,v_i)\cdot x_{ik}^* + \varepsilon_i \tag{3}$$

$$y_i^* = \frac{y_i-\bar{y}}{s_y} \tag{4}$$

$$x_{ik}^* = \frac{x_{ik}-\overline{x_k}}{s_{xk}} \tag{5}$$

where y_i^* is the standardised value of National Day holidays ridership at metro site i; (u_i, v_i) is the latitude and longitude coordinates of metro site i; β_{bw0} is the constant term for metro site i with coordinates (u_i, v_i); β_{bwk} is the local regression coefficient of the k^{th} influencing factor for metro station i; *bwk* is the optimal bandwidth of the k^{th} influencing factor; x_{ik}^* is the standardised value of the k^{th} influencing factor for metro station i; ε_i is the random error for metro station i; n is the number of influencing factor; $\bar{y}$ is the mean value of metro stations' ridership; s_y is the standard deviation of metro stations' ridership; $\bar{x_k}$ is the mean of the standardised value of the k^{th} influencing factor; s_{xk} is the standard deviation of the standardised value of the k^{th} influencing factor.

4. RESULTS

4.1 Results of the multicollinearity test and spatial autocorrelation test

Before constructing the MGWR model, it is necessary to determine whether there is a significant linear relationship between the explanatory variables. Therefore, the data can be processed through the multicollinearity test in order to build a better predictive model. When conducting the collinearity test, the population density VIF value was greater than 10 in the metro station PCA combination of 300m_400m_400m, indicating that the factor has a serious multicollinearity problem with the other factors (*Table 2*). Therefore, it is excluded from the regression analysis.

Table 2 – VIF results of 13 influencing factors

Buffer radius combination /m	VIF results of 13 influencing factors												
	Den. B	MUL	Den. R	FAR	Num.CF	Num. OF	Num. PSF	Den. BL	Num. EE	Num. PL	Num. BS	Den. P	RP
200_300_400	2.1	1.3	2.5	2.6	1.7	2.0	2.3	2.0	1.2	2.7	1.5	2.5	1.2
200_300_500	2.1	1.3	2.7	2.5	2.0	2.1	2.4	2.0	1.2	3.0	1.7	2.7	1.4
200_300_600	2.2	1.5	2.9	2.6	4.7	2.4	3.6	2.2	1.3	3.5	2.3	2.9	1.8
200_300_700	2.2	1.6	3.0	2.5	5.6	2.6	4.0	2.3	1.3	4.2	2.7	3.2	2.0
200_400_400	2.1	2.5	2.7	2.5	2.0	2.8	3.5	2.1	1.2	4.0	1.5	2.5	1.3
200_400_500	2.1	1.3	2.9	2.6	2.1	2.6	3.1	2.3	1.2	3.8	1.7	2.7	1.5
200_400_600	2.2	1.5	3.0	2.6	4.6	2.7	4.0	2.4	1.3	3.9	2.2	2.8	1.8
200_400_700	2.1	1.6	3.1	2.6	5.3	2.8	4.1	2.6	1.3	4.2	2.6	3.1	2.0
200_500_400	2.1	1.3	2.8	2.4	2.5	3.2	5.0	2.2	1.2	5.7	1.6	2.6	1.7
200_500_500	2.1	1.3	3.0	2.4	2.5	3.0	1.3	2.4	1.2	5.1	2.4	2.8	1.7
200_500_600	2.2	1.5	3.2	2.6	4.7	3.0	4.9	2.6	1.2	4.7	2.0	2.9	1.9
200_500_700	2.2	1.5	3.3	2.6	5.2	3.0	4.7	2.8	1.2	4.6	2.4	3.3	2.0
300_300_400	2.4	1.2	2.1	3.2	1.6	2.4	2.5	1.7	1.2	2.9	1.4	2.8	1.3
300_300_500	2.4	1.3	2.3	3.1	1.8	2.3	2.4	1.7	1.2	2.9	1.6	3.1	1.4
300_300_600	2.5	1.5	2.4	3.0	3.3	2.4	3.3	1.8	1.3	3.1	2.2	3.4	1.7
300_300_700	2.5	1.5	2.5	2.8	4.9	2.6	3.6	1.9	1.3	3.6	2.7	3.8	1.9
300_400_400	2.3	1.3	2.0	3.1	1.8	2.9	3.3	1.7	1.2	3.6	2.5	11.2	5.3
300_400_500	2.4	1.3	2.2	2.2	1.9	2.8	2.8	1.8	1.2	3.3	1.6	5.1	1.5
300_400_600	2.4	1.3	2.4	3.0	2.2	2.8	3.0	1.9	1.2	3.3	2.0	4.4	1.7
300_400_700	2.4	1.4	2.5	2.9	2.4	2.8	3.1	2.0	1.2	3.5	2.4	5.8	1.7
300_500_400	2.3	1.2	2.1	2.7	2.2	3.2	4.5	1.7	1.2	4.9	1.5	3.9	1.6
300_500_500	2.4	1.3	2.3	2.8	2.2	3.0	3.8	1.8	1.2	4.3	1.6	3.2	1.6

Buffer radius combination /m	VIF results of 13 influencing factors												
	Den. B	**MUL**	**Den. R**	**FAR**	**Num.CF**	**Num. OF**	**Num. PSF**	**Den. BL**	**Num. EE**	**Num. PL**	**Num. BS**	**Den. P**	**RP**
300_500_600	2.4	1.5	2..4	2.9	4.1	3.0	4.3	2.0	1.3	4.0	2.0	3.5	1.7
300_500_700	2.5	1.6	2.6	2.9	4.5	3.0	4.1	2.1	1.3	3.9	2.3	3.9	1.8
400_300_400	2.6	1.2	1.8	3.3	1.9	2.7	4.1	1.5	1.2	4.9	1.5	2.8	1.5
400_300_500	2.7	1.3	2.0	2.3	1.9	2.6	3.5	1.6	1.2	4.3	1.6	3.1	1.6
400_300_600	2.7	1.3	2.1	3.2	2.1	2.6	3.5	1.6	1.2	4.0	2.0	3.4	1.8
400_300_700	2.7	1.3	2.2	3.1	2.2	2.6	3.4	1.7	1.2	4.0	2.4	3.8	1.9
400_400_500	2.6	1.3	1.9	3.9	1.9	2.9	3.5	1.6	1.2	4.2	1.5	3.1	1.6
400_400_600	2.7	1.3	2.0	3.7	2.1	2.9	3.4	1.7	1.2	3.8	1.9	3.4	1.7
400_400_700	2.7	1.4	2.1	3.5	2.1	2.9	3.3	1.8	1.2	3.8	2.2	3.9	1.7
400_500_400	1.7	1.2	1.7	3.3	2.1	3.2	4.8	1.5	1.2	5.3	1.5	2.9	1.6
400_500_500	2.6	1.3	1.9	3.4	2.1	3.1	4.0	1.6	1.2	4.6	1.5	3.2	1.6
400_500_600	2.7	1.3	2.0	3.5	2.2	3.1	3.7	1.7	1.2	4.0	1.8	3.5	1.6
400_500_700	2.7	1.4	2.1	3.4	2.3	3.0	3.5	1.8	1.2	3.9	2.1	4.0	1.6

At the same time, whether the influencing factors are spatially clustered will also have a certain impact on the result. Spatial autocorrelation test can detect the significant correlation between the influencing factors and provide a basis for the feasibility of spatial models. The significance of the influencing factors was assessed by calculating Moran's I value, z-Score and p-Value. The results of spatial autocorrelation tests under one of 36 PCA combinations in *Table 3*. The Moran's I values are all greater than 0, indicating that all of the influencing factors have a positive spatial autocorrelation. In addition, z-Score and p-Value indicate that the null hypothesis can be rejected [46], which indicates that all influencing factors exhibit significant clustering.

Table 3 – Spatial autocorrelation test of all factors

Built environment dimension	Influencing factors	Moran's I	z-Score	p-Value
Density	Den. B	0.50	12.87	0.00
Diversity	MUL	0.06	1.71	0.05
Design	Den. R	0.25	6.48	0.00
	FAR	0.51	13.14	0.00
Destination accessibility	Num.CF	0.21	5.10	0.00
	Num. OF	0.43	11.04	0.00
	Num. PSF	0.47	12.16	0.00
Distance to transit	Den. BL	0.27	6.91	0.00
	Num.EE	0.21	5.37	0.00
Demand management	Num.PL	0.51	13.07	0.00
	Num.BS	0.14	3.59	0.00
Demographics	Den. P	0.81	20.95	0.00
	RP	0.30	7.86	0.00

4.2 Model fitting results and recommended PCA combinations of metro stations during NationalDay holidays

The coefficient of determination (R^2) and the adjusted coefficient of determination (Adj. R^2) is higher, indicating a better fit. Moreover, lower values of the Akaike information criterion (AICc) and the residual square sum (RSS) also indicate a better fit. The results of the goodness-of-fit of the MGWR model for the 36 different PCA combination scales of the metro stations are shown in *Table 4*. The maximum coefficients of determination (R^2) for the boarding and alighting ridership MGWR models are 0.72. The recommended metro station PCAs are circular buffer zones with a radius of 400 metres (within thethird ring road), 500 metres (from the third ring road to the fifth ring road), and 400 metres (outside the fifth ring road). Under this PCA combination, the RSS is the lowest for boarding ridership and the second lowest for alighting ridership.

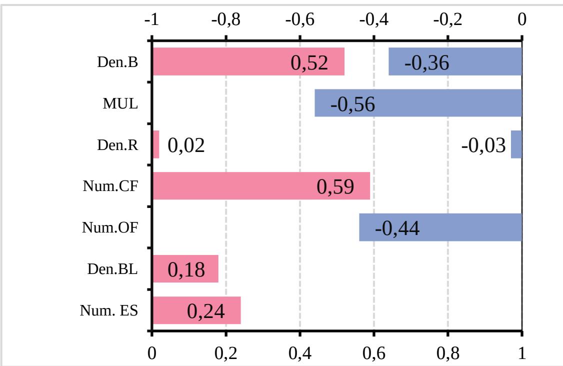
Table 4 – Goodness of fit results of the MGWR regression models under different PCA combination

Buffer radius combination /m	Boarding ridership				Alighting ridership			
	R^2	**Adj.R^2**	**AICc**	**RSS**	**R^2**	**Adj.R^2**	**AICc**	**RSS**
200_300_400	0.68	0.61	625.39	93.66	0.65	0.58	646.04	101.85
200_300_500	0.66	0.60	625.36	99.27	0.63	0.56	652.32	108.57
200_300_600	0.66	0.59	625.13	100.52	0.62	0.55	653.70	110.20
200_300_700	0.64	0.58	627.89	103.45	0.61	0.54	656.26	112.95
200_400_400	0.70	0.63	616.17	88.09	0.65	0.58	640.42	102.84
200_400_500	0.68	0.61	618.39	94.59	0.63	0.56	644.28	108.03
200_400_600	0.67	0.61	619.38	96.00	0.62	0.56	646.59	110.31
200_400_700	0.66	0.59	621.88	100.42	0.62	0.55	649.75	111.89
200_500_400	0.70	0.63	610.22	87.33	0.64	0.57	647.00	104.66
200_500_500	0.70	0.63	616.39	88.72	0.64	0.56	655.34	105.14
200_500_600	0.69	0.62	617.75	91.69	0.61	0.55	648.28	112.58
200_500_700	0.64	0.58	621.13	105.31	0.61	0.55	649.99	113.38
300_300_400	0.71	0.63	634.31	84.09	0.63	0.55	661.30	109.11
300_300_500	0.70	0.62	635.73	88.58	0.63	0.55	663.38	108.68
300_300_600	0.70	0.61	636.57	89.55	0.67	0.59	658.06	94.89
300_300_700	0.68	0.60	644.53	94.66	0.66	0.57	664.60	99.14
300_400_400	0.71	0.63	629.76	83.38	0.70	0.61	650.75	88.57
300_400_500	0.72	0.63	634.03	82.57	0.67	0.59	655.50	95.93
300_400_600	0.70	0.62	631.90	88.40	0.67	0.58	653.51	96.91
300_400_700	0.68	0.60	641.20	94.07	0.66	0.57	662.15	99.75
300_500_400	0.72	0.63	634.14	82.40	0.72	0.63	649.74	80.84

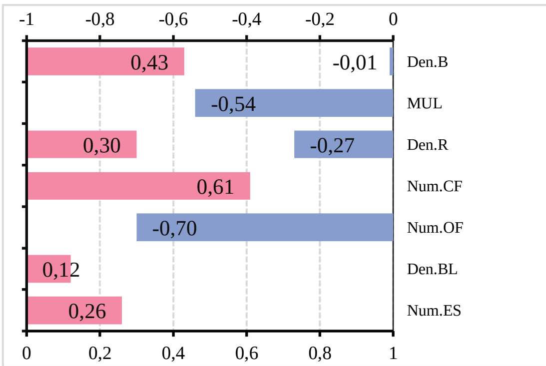
Buffer radius combination /m	Boarding ridership				Alighting ridership			
	R^2	Adj.R^2	AICc	RSS	R^2	Adj.R^2	AICc	RSS
300_500_500	0.66	0.59	639.08	97.90	0.68	0.60	654.62	92.66
300_500_600	0.66	0.59	636.38	99.10	0.65	0.57	655.51	102.36
300_500_700	0.65	0.58	641.30	102.35	0.63	0.55	660.28	108.55
400_300_400	0.70	0.60	667.22	89.72	0.64	0.54	687.39	104.70
400_300_500	0.67	0.58	674.00	96.23	0.65	0.55	692.69	101.62
400_300_600	0.69	0.59	672.98	91.70	0.65	0.55	692.82	101.23
400_300_700	0.68	0.58	677.05	93.44	0.65	0.55	695.29	101.73
400_400_500	0.70	0.60	669.47	87.96	0.63	0.54	678.15	108.07
400_400_600	0.70	0.60	668.72	87.76	0.65	0.55	679.27	102.77
400_400_700	0.66	0.57	660.43	100.20	0.65	0.56	678.32	102.35
400_500_400	0.72	0.62	665.11	80.49	0.72	0.61	683.45	81.84
400_500_500	0.72	0.62	669.56	82.15	0.71	0.61	684.22	83.69
400_500_600	0.72	0.62	665.92	82.96	0.70	0.59	687.52	88.15
400_500_700	0.70	0.61	666.55	84.82	0.70	0.60	666.48	87.11

4.3 Analysis of significant factors for boarding and alighting ridership during National Day holidays

The average values of the MGWR model coefficients of influencing factors are calculated according to positive and negative values respectively, as shown in *Figure 5*. The absolute coefficient value determines the degree of influence of the built environment on metro station ridership. The figure shows that for the boarding ridership model, the order of the influencing factors are as follows: Number of commercial facilities > Mixed utilisation of land > Building density > Number of office facilities > Number of entrance and exit > Density of bus line > Road density. For the alighting ridership model, the built environment influencing factors are as follows: Number of office facilities > Number of consumer facilities > Mixed utilisation of land > Building density > Road density > Number of entrances and exits > Density of bus line.



(a)



(b)

Figure 5 – Horizontal bar chart of positive and negative coefficients of the MGWR boarding and alighting ridership model: a) Positive and negative results of the boarding ridership MGWR model; b) Positive and negative results of the alighting ridership MGWR model.

4.4 Spatial heterogeneity effect of the built environment on metro stations ridership

The built environment's significant effect onboarding and alighting ridership are spatially visualised. *Figure 6* illustrates the correlation between different built environment influencing factors and metro ridership. Positive regression coefficients indicate that the factor is positively correlated with metro ridership (shown in red), while negative regression coefficients indicate that the factor is negatively correlated with metro ridership (shown in blue). The larger the circle and the darker the colour in the figure indicate that the influencing factor coefficient value is larger, the greater the influence degree, and the more significant effect on metro ridership. The following can be seen from the figure:

1) There are negative and positive correlations between building density and boarding and alighting ridership, as shown in *Figure 6a* and *Figure 6b*. The coefficients of building density in the southwest of the fifth ring road range from 0 to 0.82, indicating that building density strongly influences metro ridership in these regions while other factors remain unchanged. The positive correlation is mainly concentrated in the southern part of the third ring road and the southern part of the third ring road to the fifth ring road, which may be due to the higher concentration of business districts and higher building density in these areas in *Figure 6c*, which are more attractive to out-of-town tourists during the holiday season. In addition, there are large and important transportation hubs in the south-western part of the third ring road where building density is high, which will increase metro ridership. There is also a positive correlation between building density and ridership. The negatively correlated metro stations are located in the northern part of the fifth ring road, probably due to the proximity of these metro stations to mega parks with a large area and low building density, which provide a large ridership. Therefore, for the metro stations in the peripheral areas of the city, building density shows a significant negative effect.

2) The number of commercial facilities has a significant positive correlation with the boarding and alighting ridership during holidays, as shown in *Figure 6d* and *Figure 6e*. *Figure 6f* shows the distribution of commercial facilities numbers in the PCA range of metro stations. The coefficient of the number of commercial facilities has a significant effect in the northwest outside the fifth ring road, the west and south inside the fifth ring road, and a non-significant effect in the north and east inside the fifth ring road. The areas with large coefficient values of the number of commercial facilities are mainly concentrated in the western part of the fifth ring road, which is consistent with people's lifestyles. Holiday travellers mostly choose areas with more entertainment services, restaurants and shopping malls. These areas are not only for experiencing leisure and sightseeing for out-of-town tourists, but also the areas where most office workers go to relieve work fatigue. The commercial facilities POIs data includes attractions, parks, museums, historical and other facilities that attract both tourists and locals. An increase in the number of commercial facilities will attract a large number of tourists and may increase the alighting metro ridership. Therefore, the influence coefficient of the number of commercial facilities on the metro stations' boarding and alighting ridership is relatively high.

3) The distribution of the influence coefficients of bus line density on the boarding and alighting ridership of metro stations is shown in *Figure 6g* and *Figure 6h*. *Figure 6i* shows the distribution of bus line density in the PCA range of metro stations. The influence of bus line density on boarding ridership is mainly concentrated in the third ring road and the western part of the third ring road to the fifth ring road, and the north-western part of the fifth ring road. The difference is that alighting ridership shows a spatial pattern of less influence in the eastern and southern parts of the third ring road and more significant influence in the north-western part of the fifth ring road, and the regression coefficient is smaller than for the boarding ridership. The possible reason for this is that passengers are influenced by nearby activities and other travel purposes after leaving the station. They will choose to walk, take an online car-hailing or ride bike-sharing to reach their destinations by transferring to other modes. Outside the fifth ring road, people often take buses to transfer to the metro to go to the city centre, and the increase in the density of bus routes will increase the metro ridership. However, in urban areas with many shopping malls, art districts, and dining and entertainment venues, buses have the characteristics of small station spacing, flexible excursions and high frequency, so buses will take up most of the ground-level traffic. Under this influence, the metro's boarding and alighting ridership will be less in demand compared to the peripheral areas.

4) As one of the metro connection modes, the number of entrances and exits has a significant positive effect on the boarding and alighting ridership, as shown in *Figure 6j* and *Figure 6k*, respectively. Observing the coefficient change of the number of entrances and exits on boarding and alighting ridership, it is found that the number of entrances and exits in the peripheral area of the city significantly affects the boarding ridership, and the centre of the city has a greater effect on the alighting ridership. The coefficient change means that increasing the number of entrances and exits in the periphery of the city can attract more people to take the

subway. *Figure 6l* shows that the small number of entrances and exits usually means that the connectivity between the metro station and the road is poor, which causes inconvenience to passengers and reduces the metro ridership to a certain extent. Increasing the number of entrances and exits can enhance the accessibility of metro stations and subsequently increase metro ridership. This confirms the hypothesis that there is a positive correlation between ridership and the presence of entrances and exits [16].

5) The effect of the number of office facilities on the boarding and alighting metro ridership shows a significant negative correlation during National Day holidays. *Figure 6m* and *Figure 6n* show that most of the stations outside the third ring road in southwest Beijing have a significant negative relationship with boarding and alighting ridership. The influence coefficients of boarding and alighting ridership are highest outside the fifth ring road (the absolute value of local regression coefficient is more than 0.90), and the influence coefficient gradually becomes lower from the outside to the inside. Most of these stations outside the fourth ring road are low-vitality stations with small metro ridership, as shown in *Figure 3*. The built environment influencing factors are low around these metro stations, such as the number of commercial facilities (*Figure 6f*) and mixed utilisation of land (*Figure 6r*), while the building density (*Figure 6c*), bus line density (*Figure 6i*) and road density (*Figure 6u*) are relatively high. Even though there is a significant negative correlation between the number of office facilities and metro ridership, the boarding and alighting ridership are still at a very low level (*Figure 6m and Figure 6n*). The possible reason for this is that the buffer area occupied by the office facilities around the metro station is relatively large, while the residential facilities are relatively few and low. These stations are dominated by local residents, and fewer foreign tourists enter and egress these metro stations. Therefore, office facilities are negatively correlated with boarding and alighting ridership. The larger the number of office facilities, the lower the boarding and alighting ridership. The boarding and alighting ridership of the metro stations within the third ring road and the fourth ring road are relatively high (*Figure 3*). The building density, the number of commercial facilities and the mixed utilisation of land around these metro stations are relatively large, which attracts a large number of metro ridership. More commercial facilities around these metro stations may lead to less office facilities, which may be the reason why the office facilities of these metro stations are negatively correlated with boarding and alighting ridership.

6) Mixed utilisation of land can provide more choices for different passengers and promote diversified travel demand. The local regression coefficients of the effect of mixed utilisation of land on boarding and alighting metro ridership are shown in *Figure 6p* and *Figure 6q*, respectively. *Figure 6r* shows the distribution of mixed utilisation of land in the PCA range of metro stations. A negative correlation exists between mixed utilisation of land and boarding and alighting metro ridership. In areas with a high land mix, ridership decreases instead. In the southern part of the city, there is a negative correlation, which may be due to the fact that in central urban areas with high land use mix and small buildings, people prefer walking or other lightweight transportation such as bicycles. These areas are also considered pedestrian-friendly neighbourhoods. The degree of negative correlation impact is low and negative outside of Beijing's fifth ring road, which indicates a single type of land use in the outer urban areas, mainly residential areas with a large footprint and a low degree of land use mix. The reliance of the residents on the metro is relatively low, and travel is mostly by car, resulting in lower boarding and alighting metro ridership.

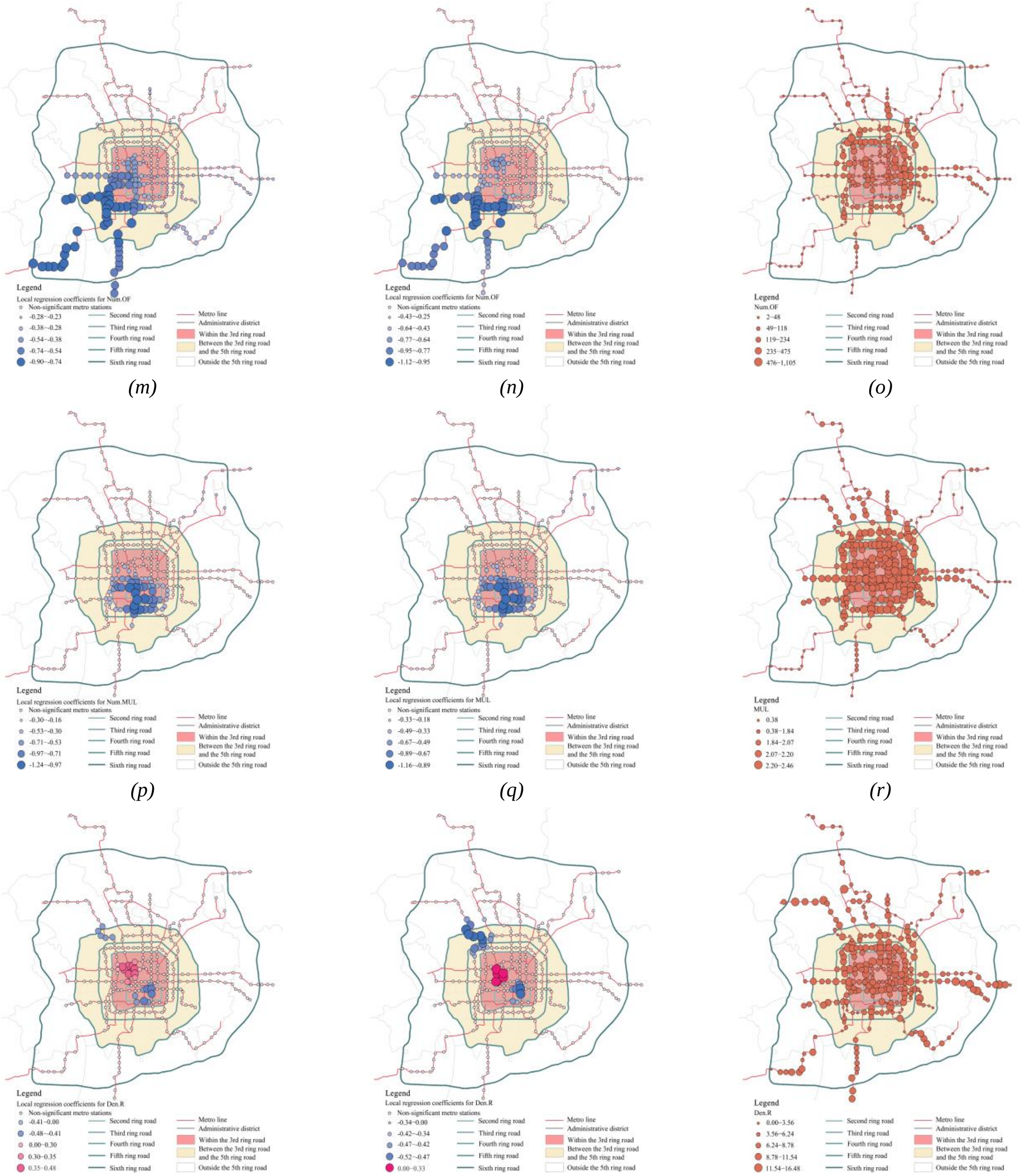
7) The distribution of the impact of road density on boarding and alighting metro ridership is shown in *Figure 6s* and *Figure 6t*, respectively. *Figure 6u* shows the distribution of road density in the PCA range of metro stations. The effect of road density on metro ridership is positive in the north-western part of the third ring road and negative in the north-western part of the fifth ring road and the south-eastern part of the third ring road. This result indicates that the emerging problems of urban traffic congestion, high energy consumption and frequent traffic accidents are caused by the rapid growth of private car ownership. People prefer to take the subway due to concerns about travel safety and longer travel times. A trend from positive to negative impacts can be observed from the urban core to the peripheral areas. The possible explanation is that on the one hand, the reduced road density, smooth roads and convenient transportation in peripheral urban areas have attracted a large number of people to choose other modes of transportation. On the other hand, due to the limitations of the metro's operating hours and routes, a part of the population prefers to choose other modes of transportation.

(a) *(b)* *(c)*

(d) *(e)* *(f)*

(g) *(h)* *(i)*

(j) *(k)* *(l)*



(m) *(n)* *(o)*

(p) *(q)* *(r)*

(s) *(t)* *(u)*

Figure 6 – Local regression coefficient distribution of the built environment influencing factors for boarding and alighting metro ridership: (a) Den. B for boarding ridership; (b) Den. B for alighting ridership; (c) Den. B; d) Num.CF for boarding ridership; (e) Num.CF for alighting ridership; (f) Num.CF; g) Den. BL for boarding ridership; (h) Den. BL for alighting ridership; (i) Den. BL; (j) Num. EE for boarding ridership; (k) Num. ES for alighting ridership; (l) Num. EE; (m) Num. OF for boarding ridership; (n) Num. OF for alighting ridership; (o) Num. OF; (p) MUL for boarding ridership; (q) MUL for alighting ridership; (r) MUL; (s) Den. R for boarding ridership; (t) Den. R for alighting ridership; (u) Den. R

4.5 Low-vitality metro station and renewal strategies

According to the spatial distribution of boarding and alighting metro ridership, the low-vitality metro stations are screened out. Targeted updating strategies are proposed according to the influencing factors that significantly affect the low-vitality metro stations. The results are shown in *Table 5*. In the table, "+" indicates a positive effect of influencing factors on the low-vitality metro station. Similarly, "-" means that the factor has a negative effect on the low-activity metro station. For updating low-vitality metro stations, we can prioritise to adjust the influencing factors that have a greater effect on the metro ridership. For example, Yongdingmenwai (a low-vitality metro station) can focus on increasing the number of entrance and exit and increasing building density to improve the metro service capacity and metro ridership. Low-vitality metro stations such as Fengtai East Street, Guozhuangzi, and Majiapu can reduce the number of office facilities or integrate residential, office, commercial and cultural uses. Changing a single land use pattern can increase mixed utilisation of land and purposefully adjust the low vitality metro stations. Low-vitality metro stations such as West Diaoyutai, Yizhuangqiao and Fuxingmen can also be combined with commercial facilities (scenic, shopping, hotels and commercial enterprises) to improve their vitality.

Table 5 – Priority update low-viability metro stations of sensitive built environment

Low-vitality station name	Den. B	Num. CF	Den. BL	Num. EE	Num. OF	MUL	Den. R
Anheqiao			+	+			
Coking plant		+	+	+	−	−	
Yizhuangqiao		+	+	+	−		
Guangyangcheng		+	+	+	−		
Yongdingmen Wai	+	+	+	+	−	−	−
Chegongzhuang West		+	+	+	−	−	+
Fengtai East Street	+	+	+	+	−	−	
Xidiaoyutai	+	+	+	+	−	−	
Beitucheng			+	+			
Malianwa		+	+				−
GuoZhuangzi	+	+	+	+	−		
Jingtai	+	+	+	+	−	−	−
Majiapu	+	+	+	+	−	−	
Fuxingmen	+	+	+	+	−	−	+

Note: "+" indicates a positive effect of influencing factors on the low-vitality metro station, "-" means that the factor has a negative effect on the low-activity metro station.

5. DISCUSSION

5.1 Advantages of studying National Day holidays ridership

Although there have been numerous studies on non-holidays ridership [3, 13, 47, 48], as well as on the characteristics of metro transit ridership during holidays [49–52], few scholars pay attention to the effect of the built environment around the metro station on the holiday's ridership. Additionally, there are significant differences in the distribution rules of ridership between holidays and non-holidays [50]. For example, we compared the ridership of Zhou and Tang [53] and Wang and Li [52], and some metro stations ridership during holidays was much higher than non-holidays ridership. This shows that the ridership usually increases during holidays, while the ridership during non-holidays is relatively stable. The crowded ridership during National Day holidays presents a significant challenge for urban metro transit. The frequent occurrence of crowd stampedes [54], the great difference in the service levels among different stations [55] and the importance of enhancing the attraction of metro stations [56] suggest that we need to pay attention to the effect of ridership during holidays. The advantages of the ridership prediction model constructed in this study are as follows: (1) It can improve the operation service level at urban metro transit stations under large ridership. In order to optimise the urban built environment, targeted renewal strategies are proposed for metro stations with insufficient capacity. (2) It is helpful for the urban metro transit operation and management department in formulating timely to deal with emergencies and ensure that the travel demands of tourists and local residents are met during holidays. (3) It can reveal the influence of commercial or scenic facilities on the metro stations ridership during holidays. For example, Wangfujing, Zoo, Fuxingmen and Yongdingmenwai stations are affected by the scenic facilities, and the boarding and alighting metro ridership increases significantly during holidays.

5.2 Characteristics of the PCA combination of metro stations during the National Day holidays

In numerous studies, the PCA size of metro stations is generally large and remains consistent during non-holidays [3, 17, 20, 33, 52, 57]. However, this study focuses on Beijing and divides the city into three zones to determine the PCA of metro stations for each zone. Some scholars have also considered dividing the city into different zones [18, 48, 58]. But their dependent variable is weekday ridership. Nevertheless, they have not taken into account how holiday ridership may lead to different results in determining the PCA size of metro stations. The findings of this study reveal that the PCA size of metro stations is relatively small during National Day holidays. Metro as one of the convenient transportation ways during holidays, can quickly connect the major scenic facilities in the city. As a result, there is a significant increase in boarding and alighting metro ridership at stations adjacent to these attractions, such as Xizhimen, Dongzhimen and Tiantongyuan North. This reflects the fact that tourists like to gather around metro stations. Not only because of the convenient transportation in these areas, but also because of the characteristics of rich cultural tourism resources and complete service functions. Choosing a hotel near the metro station can meet the travel needs during the holidays. For the above reasons, the MGWR model results indicate that the small PCA size of metro stations lead to the increase of model accuracy.

According to the research results, it is recommended that the PCA radius of the metro stations within the third ring road of Beijing be 400 metres, 500 metres from the third ring road to the fifth ring road, and 400 metres outside the fifth ring road. In this study, the recommended PCA of metro stations verifies that the PCA size of the National holiday is lower than in other studies [20, 33, 52, 57]. More than half of the office facilities in Beijing from the third ring road to the fifth ring road has driven the development of the surrounding metro stations, promoting the development of commercial, cultural and educational functions. The improved service quality of facilities within the third ring road to the fifth ring road will result in travellers not being limited to metro station vicinity. Therefore, the PCA size of the metro stations will be larger than that one within the third ring road. For metro stations outside of the fifth ring road, a recommended PCA of 400 metres is proposed. The smaller recommended PCA size of metro stations outside the fifth ring road also indicates that outside the fifth ring road, the density of metro lines is low and the spacing of metro stations is large. Therefore, tourists are likely to prefer to be closer to the metro for convenient travel during the National Day holidays. The PCA combination of metro stations improves the accuracy of the holiday ridership prediction models and provides scientific data support for enhancing Beijing Metro’s attractiveness with low ridership during holidays.

5.3 Spatial heterogeneity of the influence of the built environment on National Day holidays ridership

During the non-holidays, the most significant factor affecting the built environment is the density of residential and office facilities in the “density” dimension [15, 22, 33, 57]. During the National Day holidays, the top two factors affecting ridership are the number of commercial facilities in the “destination accessibility” dimension, as well as the mixed utilisation of land in “diversity” dimension. The results indicate that both the quantity and variety of commercial facilities play crucial roles in attracting tourists during holidays. Commercial areas establishments offering cultural and entertainment activities can provide tourists with colourful cultural experiences and tourism activities. These places can also meet the travel needs of different tourists, which will increase the ridership of some metro stations. Among the number of commercial facilities, it is worth paying attention to tourist destinations such as scenic facilities. During holidays, metro stations ridership around scenic facilities increases significantly, which may affect the surrounding road traffic and urban operation efficiency. Increase transportation connections between attractions and attractions, such as promoting the “metro + bus” mode of transportation. Enhancing transportation connectivity between attractions could also reduce pressure on metro stations around scenic facilities. Similarly, areas with a higher mixed utilisation of land can provide more facilities and services, and these areas often have more compact urban designs. Notably, during holidays and non-holidays, the similarity between some studies [3, 14] and this study is that there is a positive effect between the mixed utilisation of land and the metro stations ridership on non-holidays. However, the difference is that areas with a high mixed utilisation of land tend to have more services and recreational facilities, which assume a similar function to scenic facilities. This further suggests that areas with a high mixed utilisation of land are better able to attract tourists and local residents, especially during holidays.

In addition, existing studies has shown that the greater the number of office facilities, the greater the effect on the metro stations ridership [34, 52, 58–60]. But we found that during the holidays, the effect was reversed. Despite having a large number of office facilities, these areas are not always effective in attracting tourists and local residents. Perhaps it is the specific social and economic background of the holiday and the change of passenger travel habits. To some extent, it reflects that in areas with many office facilities, the types of land use are too single or too few. At the urban level, we propose enhancing the visual appeal of office spaces through creative space design and installation art, as well as promoting the office space opening and creating an artistic atmosphere. By combining the characteristics of holidays, holding exhibitions or commercial activities should attract more tourists.

5.4 Advantages and limitations of using the MGWR model

Recently, some scholars [61] have proposed the difference between linear influence and nonlinear influence. Among them, machine learning such as Gradient Boosting Decision Trees (GBDT) model [33, 61-63], random forest [64], deep learning [65, 66], eXtreme Gradient Boosting (XGBoost) model [48, 61, 67] and Optimal Parameters-based Geographic Detector (OPGD) [52] model are used to explore the effect of the built environment on metro stations. These models have their own characteristics and are suitable for different prediction situation and data characteristics.

The OPGD model [68] lays more emphasis on the driving force of the influencing factors, and is weak in explaining the spatial heterogeneity. Machine learning models [69] emphasise more on the degree of fitting of feature value to the results of test sets, which may be unreasonable in the process of causal relationship. The MGWR model considers the collinearity of multiple influencing factors to find the regression model with the best goodness of fit considering spatial heterogeneity. Therefore, this study uses this model to explore the effect of built environment factors on the ridership during the National Day holidays.

The limitations of this study are as follows: (1) We hope to combine nonlinear model and interpretable machine learning in future research to explore the causal inference in metro ridership. (2) The study only relied on POI data around metro stations, which emphasises information about specific points. Area of Interest (AOI) data is not used, which can express the surface of geographical entities in the map data. Using AOI data to calculate built environment variables may improve the accuracy of subway passenger flow prediction model. (3) We assume that the PCA is a circular buffer. The ways to access and egress to the metro stations in Beijing include walking, bus, taxi, online car-hailing, bike-sharing and bicycle. It is still worth discussing how to define PCA by considering multiple connection access and egress modes. (4) In the future, more studies are needed to compare research results and conclusions for holidays, weekdays and weekends.

6. IMPLICATIONS

This research carried several valuable implications for the development of urban metro transit, particularly for cities sharing similar morphology with Beijing (i.e. high density, transit-oriented and single-centre) provides significant insights. First, it is very important to predict the metro ridership during the National Day holidays to ensure travel safety, deal with emergencies and improve urban public services. The MGWR regression model was used to forecast the ridership, the present study proposed seven significant influencing factors to explain the metro station ridership in Beijing. However, the degree of these influencing factors was quite different, with some of them having a high effect in other studies, with other not. In particular, the number of commercial facilities, mixing degree of land use and building density are significantly associated with metro ridership in Beijing. In addition, the number of office facilities shows a significant negative correlation during holidays, which may be due to China's special national conditions. During holidays, some office facilities may choose to close. However, most office facilities will switch some workers their rest time in order to work normally on holidays. During this period, they will offer higher subsidies to their workers. These influencing factors provide important information for urban metro transit travel demand modellers in travel demand analysis. More consideration should be given to factors in order to improve the unbalanced metro ridership in the future.

Second, for urban metro transit planners, the recommended PCA combination can serve as a basis for calculating built environment factors around metro stations. The reasonable choice of the PCA is the key when constructing regression models for metro station ridership. According to the distribution of metro stations in the city, it is divided into three zones. Building density, the number of commercial facilities, mixed utilisation of land and the road density within the third ring road have a higher significance on the metro ridership, while the bus line density and the number of entrance and exit extend from the outside to the inside of the fifthring road, showing a clear downward trend in spatial influence. The number of office facilities exhibited an inward decline trend and demonstrates a negative correlation with metro ridership during holidays. Therefore, it is essential to implement different strategies for various areas to adjust metro ridership effectively.

Third, the analysis results can propose specific built environment strategies for low-vitality metro stations. For urban planners, the priority is to adjust the built environment influencing factors that significantly affect metro ridership and impose reasonable constraints based on control indicators in urban design. The built environment around the metro station encourages density and diversity, which has a positive impact on ridership. It is necessary to refer to the MGWR results to determine which influencing factors should be improved to increase ridership at low-vitality metro stations.

7. CONCLUSIONS

According to the density of metro station distribution, Beijing is divided into three different zones based on the built environment “7D” dimension to select the built environment influencing factors, the National Day holidays boarding ridership and alighting ridership as the dependent variable. The recommended PCA for each zone metro station was determined by using the MGWR model to select the highest goodness-of-fit results. The metro station PCA with the highest goodness-of-fit result was used to find the spatial heterogeneity of the built environment influencing factors. The main conclusions include:

1) The metro station PCA with different combination sizes selected by zoning can improve the accuracy of the MGWR model and have better explanatory power for the influencing factor’s spatial heterogeneity. The findings are valuable for defining the scope of the PCA of metro stations and accurately determining the extent of renewal needed in the built environment surrounding metro stations.

2) There is no obvious peak time for boarding and alighting ridership during the holidays, and the ridership lasts for a long time. Therefore, the variability of the MGWR model results is reflected in the fact that the size of the metro station PCA during the National Day holidays is different from many studies that take the metro station PCA as the walking area within 800 metres of the metro station. The start and end points of the tourists will overlap near the metro, which results in a smaller size of the PCA. The recommended PCA combinations are circular buffers with a radius of 400 metres, 500 metres and 400 metres for three zones, respectively.

3) The results of the spatial heterogeneity of the influencing factors can be used for developing targeted built environment strategies to improve the vitality of metro stations. The stations with low-vitality are screened, the influencing factors affecting their ridership are summarised, and strategies for updating the built environment around the metro stations are proposed in order to improve their vitality.

REFERENCES

[1] Dong X, et al. Can negative travel habits hinder positive travel behavioural change under Beijing vehicle restrictions? *Promet-Traffic&Transportation*. 2020;32(5):691-709. DOI: 10.7307/PTT.V32I5.3453.

[2] Crane R, Scweitzer LA. Transport and sustainability: The role of the built environment. *Built Environment*. 2003;29:238-252. DOI: 10.2148/benv.29.3.238.54286.

[3] Zhao J, et al. What influences metro station ridership in China? Insights from Nanjing. *Cities*. 2013;35:114-124. DOI: 10.1016/j.cities.2013.07.002.

[4] Holz-Rau C, Scheiner J, Sicks K. Travel distances in daily travel and long-distance travel: What role is played by urban form? *Environment and Planning A*. 2014;46:488-507. DOI: 10.1068/a4640.

[5] Reichert A, Holz-Rau C. Mode use in long-distance travel. *Journal of Transport and Land Use*. 2015;8:87-105. DOI: 10.5198/jtlu.2015.576.

[6] Strömblad E, et al. Characteristics of everyday leisure trips by car in Sweden–Implications for sustainability measures. *Promet-Traffic&Transportation*. 2022;34:657-672. DOI: 10.7307/ptt.v34i4.4039.

[7] Meng Z, Gao X. Beijing Metro will carry nearly 5.8 million passengers on its lines during the National Day holiday. People's Daily Online. 2021. http://bj.people.com.cn/n2/2021/0930/c82840-34939043.html [Accessed 19th Sep. 2024].

[8] Xiang B. More metro lines in smaller Chinese cities. Xinhuanet. 2017. http://www.xinhuanet.com/english/2017-09/08/c_136594680.html [Accessed 19th Sep. 2024].

[9] Yang L, Shen Q, Li Z. Comparing travel mode and trip chain choices between holidays and weekdays. *Transportation Research Part A: Policy and Practice*. 2016;91:273-285. DOI: 10.1016/j.tra.2016.07.001.

[10] Hu X, et al. Traffic volume forecasting model of freeway toll stations during holidays–An SVM model. *Promet-Traffic&Transportation*. 2022;34:499-510. DOI: 10.7307/ptt.v34i3.3885.

[11] Kim C, et al. What makes urban transportation efficient? evidence from subway transfer stations in Korea. *Sustainability*. 2017;9(11):2054. DOI: 10.3390/su9112054.

[12] Liu J, Shi W, Chen P. Exploring travel patterns during the holiday season—A case study of Shenzhen metro system during the Chinese spring festival. *ISPRS International Journal of Geo-Information*. 2020;9(11):651. DOI: 10.3390/ijgi9110651.

[13] An D, et al. Understanding the impact of built environment on metro ridership using open source in Shanghai. *Cities*. 2019;93:177-187. DOI: 10.1016/j.cities.2019.05.013.

[14] Estupiñán N, Rodríguez DA. The relationship between urban form and station boardings for Bogotá's BRT. *Transportation Research Part A: Policy and Practice*. 2008;42:296-306. DOI: 10.1016/j.tra.2007.10.006.

[15] Sohn K, Shim H. Factors generating boardings at Metro stations in the Seoul metropolitan area. *Cities*. 2010;27:358-368. DOI: 10.1016/j.cities.2010.05.001.

[16] Sung H, et al. Exploring the impacts of land use by service coverage and station-level accessibility on rail transit ridership. *Journal of Transport Geography*. 2014;36:134-140. DOI: 10.1016/j.jtrangeo.2014.03.013.

[17] Cardozo OD, García-Palomares JC, Gutiérrez J. Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. *Applied Geography*. 2012;34:548-558. DOI: 10.1016/j.apgeog.2012.01.005.

[18] Andersson DE, Shyr OF, Yang J. Neighbourhood effects on station-level transit use: Evidence from the Taipei metro. *Journal of Transport Geography*. 2021;94:103127. DOI: 10.1016/j.jtrangeo.2021.103127.

[19] Chen E, et al. Discovering the spatio-temporal impacts of built environment on metro ridership using smart card data. *Cities*. 2019;95:102359. DOI: 10.1016/j.cities.2019.05.028.

[20] Li S, 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*. 2020;82:102631. DOI: 10.1016/j.jtrangeo.2019.102631.

[21] Blainey S. Trip end models of local rail demand in England and Wales. *Journal of Transport Geography*. 2010;18(1):153-165. DOI: 10.1016/j.jtrangeo.2008.11.002.

[22] Jun MJ, et al. Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. *Journal of Transport Geography*. 2015;48:30-40. DOI: 10.1016/j.jtrangeo.2015.08.002.

[23] Yu H, et al. Inference in multiscale geographically weighted regression. *Geographical Analysis*. 2019;52:87-106. DOI: 10.1111/gean.12189.

[24] Fotheringham AS, Yang W, Kang W. Multiscale geographically weighted regression (MGWR). *Annals of the American Association of Geographers*. 2017;107(6):1247-1265. DOI: 10.1080/24694452.2017.1352480.

[25] Shen T, et al. On hedonic price of second-hand houses in Beijing based on multi-scale geographically weighted regression: Scale law of spatial heterogeneity. *Econ. Geogr*. 2020;0:75-83. DOI: 10.15957/j.cnki.jjdl.2020.03.009.

[26] Chen Z, Deng S. Research on the relationship between block morphology and heat island intensity in Guangzhou's main urban area based on multi-scale geographically weighted regression. *Intell. Build. Smart City*. 2021;10:13-17. DOI: 10.13655/j.cnki.ibci.2021.10.003.

[27] Zhou Z, Cheng X. Spatial heterogeneity of influencing factors of PM2. 5 in Chinese cities based on MGWR model. *China Environ. Sci*. 2021;41:2552-2561. DOI: 10.19674/j.cnki.issn1000-6923.2021.0267.

[28] Wang Z, et al. Spatial heterogeneity analysis for influencing factors of outbound ridership of subway stations considering the optimal scale range of "7D" built environments. *Sustainability*. 2022;14(23):16314. DOI: 10.3390/su142316314.

[29] Wang Z, et al. Multi-scale geographically weighted elasticity regression model to explore the elastic effects of the built environment on ride-hailing ridership. *Sustainability*. 2023;15(6):4966. DOI: 10.3390/su15064966.

[30] Kuby M, Barranda A, Upchurch C. Factors influencing light-rail station boardings in the United States. *Transportation Research Part A : Policy and Practice*. 2004;38(3):223-247. DOI: 10.1016/j.tra.2003.10.006.

[31] Cervero R, Kockelman K. Travel demand and the 3Ds: Density, diversity, and design. *Transportation research part D: Transport and environment*. 1997;2(3):199-219. DOI: 10.1016/S1361-9209(97)00009-6.

[32] Cervero R, et al. Influences of built environments on walking and cycling: Lessons from Bogotá. *International Journal of Sustainable Transportation*. 2009;3:203-226. DOI: 10.1080/15568310802178314.

[33] Gan Z, et al. Examining the relationship between built environment and metro ridership at station-to-station level. *Transportation Research Part D: Transport and Environment*. 2020;82:102332. DOI: 10.1016/j.trd.2020.102332.

[34] Liu C, et al. How to increase rail ridership in Maryland: Direct ridership models for policy guidance. *Journal of Urban Planning and Development*. 2016;142(4):04016017. DOI: 10.1061/(asce)up.1943-5444.0000340.

[35] Chris DG, et al. How does the built environment affect transit use by train, tram and bus? *Journal of Transport and Land Use*. 2020;13(1):625-650. DOI: 10.5198/jtlu.2020.1739.

[36] Zhao J, Deng W. Relationship of walk access distance to rapid rail transit stations with personal characteristics and station context. *Journal of Urban Planning and Development*. 2013;139:311-321. DOI: 10.1061/(asce)up.1943-5444.0000155.

[37] Chalermpong S, Wibowo SS Transit station access trips and factors affecting propensity to walk to transit stations in Bangkok, Thailand. *In Proceedings of the Proceedings of the Eastern Asia Society for Transportation Studies Vol. 6 (The 7th International Conference of Eastern Asia Society for Transportation Studies, 2007)*. 2007; pp. 232-232.DOI: 10.11175/eastpro.2007.0.232.0.

[38] Wang Z, Chen F, Xu T. Interchange between metro and other modes: Access distance and catchment area. *Journal of Urban Planning and Development*. 2016;142(4):04016012. DOI: 10.1061/(asce)up.1943-5444.0000330.

[39] Kim D, Elek P, Mirjana R. *Mapping urbanities: Morphologies, flows, possibilities*; Routledge; 2017.

[40] Yu L, Cong Y, Chen K. Determination of the peak hour ridership of metro stations in Xi’an, China using geographically-weighted regression. *Sustainability*. 2020;12(6):2255. DOI: 10.3390/su12062255.

[41] Wang Z, Wang S, Lian H. A route-planning method for long-distance commuter express bus service based on OD estimation from mobile phone location data: The case of the Changping Corridor in Beijing. *Public Transport*. 2020;13:101-125. DOI: 10.1007/s12469-020-00254-w.

[42] Ren H, Wang Z, Chen Y. Optimal express bus routes design with limited-stop services for long-distance commuters. *Sustainability*. 2020;12(4):1669. DOI: 10.3390/su12041669.

[43] Li L, et al. Spatial distribution of travel activities and its relationship with points of interest. *Promet-Traffic&Transportation*. 2021;33:1-16. DOI: 10.7307/ptt.v33i1.3462.

[44] Mansour S, et al. Sociodemographic determinants of COVID-19 incidence rates in Oman: Geospatial modelling using multiscale geographically weighted regression (MGWR). *Sustain Cities Soc*. 2021;65:102627. DOI: 10.1016/j.scs.2020.102627.

[45] Moran PA. Notes on continuous stochastic phenomena. *Biometrika*. 1950;37:17-23. DOI:10.2307/2332142.

[46] Castro MCD, Singer BHJGA. Controlling the false discovery rate: A New Application to Account for Multiple and dependent tests in local statistics of spatial association. *Geographical Analysis.* 2006;38:180-208. DOI: 10.1111/j.0016-7363.2006.00682.x.

[47] Sung H, Oh J-T. Transit-oriented development in a high-density city: Identifying its association with transit ridership in Seoul, Korea. *Cities*. 2011;28:70-82. DOI: 10.1016/j.cities.2010.09.004.

[48] Wang Z, et al. Built environment renewal strategies aimed at improving metro station vitality via the interpretable machine learning method: A case study of Beijing. *Sustainability*. 2024;16(3):1178. DOI: 10.3390/su16031178.

[49] Yao E, et al. Forecasting passenger flow distribution on holidays for urban rail transit based on destination choice behavior analysis. *Journal of Advanced Transportation*. 2021;2021:922660. DOI: 10.1155/2021/9922660.

[50] Xie B, et al. Travel characteristics analysis and passenger flow prediction of intercity shuttles in the Pearl River delta on holidays. *Sustainability*. 2020;12(18):7249. DOI: 10.3390/su12187249.

[51] Cai C, et al. Holiday destination choice behavior analysis based on AFC data of urban rail transit. *Discrete Dynamics in Nature and Society*. 2015;1:136010. DOI: 10.1155/2015/136010.

[52] Wang Z et al. Contribution of built environment factors and their interactions with subway station ridership. *Public Transport.* 2024;16:929–965.DOI: 10.1007/s12469-024-00353-y.

[53] Zhou G, Tang J. Forecast of urban rail transit passenger flow in holidays based on support vector machine model. *In Proceedings of the 5th International Conference on Electromechanical Control Technology and Transportation (ICECTT), IEEE.* 2020; pp. 585-589.

[54] Li J, et al. Risk-based crowd massing early warning approach for public places: A case study in China. *Safety Science*. 2016;89:114-128. DOI: 10.1016/j.ssci.2016.06.007.

[55] Batarce M, Muñoz JC, Torres I. Characterizing the public transport service level experienced by users: An application to six Latin American transit systems. *Journal of Public Transportation*. 2022;24:100006. DOI: 10.1016/j.jpubtr.2022.100006.

[56] Zacharias J, Zhang T, Nakajima NJUDI. Tokyo Station City: The railway station as urban place. *Urban Design International.*2011;16:242-251. DOI: 10.1057/udi.2011.15.

[57] Zhao J, et al. Analysis of Metro ridership at station level and station-to-station level in Nanjing: an approach based on direct demand models. *Transportation*. 2013;41:133-155, DOI: 10.1007/s11116-013-9492-3.

[58] Li S, et al. The varying patterns of rail transit ridership and their relationships with fine-scale built environment factors: Big data analytics from Guangzhou. *Cities*. 2020;99:102580. DOI: 10.1016/j.cities.2019.102580.

[59] Yang H, et al. Direct modeling of subway ridership at the station level: A study based on mixed geographically weighted regression. *Canadian Journal of Civil Engineering*. 2020;47:534-545. DOI: 10.1139/cjce-2018-0727.

[60] Rosenbloom S, Clifton KJ. The puzzle of income, race, and density: Preliminary evidence on transit use from the 1991 American housing survey. *Journal of public transportation*. 1996;1:87-102. DOI: 10.5038/2375-0901.1.1.5.

[61] Liu M, et al. Nonlinear effects of built environment features on metro ridership: An integrated exploration with machine learning considering spatial heterogeneity. *Sustainable Cities and Society*.2023;95:104613. DOI: 10.1016/j.scs.2023.104613.

[62] Du Q, et al. Spatiotemporal exploration of the non-linear impacts of accessibility on metro ridership. *Journal of Transport Geography*. 2022;102:103380. DOI: 10.1016/j.jtrangeo.2022.103380.

[63] Ding C, Cao X, Liu C. How does the station-area built environment influence Metrorail ridership? Using gradient boosting decision trees to identify non-linear thresholds. *Journal of Transport Geography*. 2019;77:70-78. DOI: 10.1016/j.jtrangeo.2019.04.011.

[64] Cheng L, et al. Applying a random forest method approach to model travel mode choice behavior. *Travel behaviour and society*. 2019;14:1-10. DOI: 10.1016/j.tbs.2018.09.002.

[65] Zhao X, et al. Prediction and behavioral analysis of travel mode choice: A comparison of machine learning and logit models. *Travel Behaviour and Society*. 2020;20:22-35. DOI: 10.1016/j.tbs.2020.02.003.

[66] Hagenauer J, Helbich M. A comparative study of machine learning classifiers for modeling travel mode choice. *Expert Systems with Applications*. 2017;78:273-282,.DOI: 10.1016/j.eswa.2017.01.057.

[67] Zhou S, et al. Impacts of building configurations on urban stormwater management at a block scale using XGBoost. *Sustainable Cities and Society*. 2022;87:104235. DOI: 10.1016/j.scs.2022.104235.

[68] Song Y, et al. An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: cases with different types of spatial data. *GIScience & Remote Sensing*. 2020;57:593-610. DOI: 10.1080/15481603.2020.1760434.

[69] Ke G, et al. Lightgbm: A highly efficient gradient boosting decision tree. *Proceedings of 31st Conference on Neural Information Processing Systems (NIPS 2017), Dec. 2017, Long Beach, CA, USA*. 2017. p. 3149-3157.

王振报，和艳芳，赵雪乔，梁榆淇，李世豪

探讨建筑环境因素对节假日期间地铁站乘客量的影响——以中国国庆节期间北京地铁系统为例

摘要

以往的研究更关注建成环境对工作日和周末乘客量的影响。为弥补没有关注国庆假日客流量的空白，本文将探究国庆节期间建成环境对轨道站点客流的影响程度及空间异质性。将北京市由内到外划分为三个片区。以国庆节期间地铁站上、下车乘客量为因变量，建筑环境 "7D "维度的 13 个建成环境因素为自变量。拟合优度最佳的多尺度地理加权回归模型表明三个区域行人集水区半径组合为 400 m_500 m_400 m，探讨了建成环境因素对地铁乘客量和空间异质性的影响。建筑密度、消费类设施数量、公交线路密度、出入口数量、办公类设施数量、土地利用混合度和道路密度对上、下车客流量均有显著影响。MGWR 模型的结果有助于提出针对性的轨道站点周边建成环境更新策略。

关键词

轨道站点客流量；轨道站点行人集水区；多尺度地理加权回归；建成环境；国庆假日