Spatial and Temporal Characteristics of the Impact of TOD Built Environment on Rail Transit Riders' Travels

Junyi Zhao, Changfeng Zhu, Kewei Zhao

***Abstract*—This study investigates the spatial and temporal distribution of passenger flow at rail stations, focusing on selected lines within the Xi'an Metro's main urban area. By integrating multi-source data and utilizing a multi-scale geographically weighted regression (MGWR) model based on a "5D" index system of Transit-Oriented Development (TOD) built environment, the research quantitatively examines the impact of the built environment on passenger flow during evening peak hours. The results reveal significant spatial heterogeneity in the distribution of passenger flow at Xi'an rail stations. The MGWR model demonstrates a higher goodness-of-fit compared to the traditional geographically weighted regression (GWR) model, effectively capturing the complex spatial and temporal interactions between various factors and passenger flow. Additionally, the influence of different factors on passenger flow varies considerably across space. The study suggests that TOD strategies should adopt differentiated development policies: enhancing scale and intensity at suburban stations, while focusing on optimizing the quality of development at central city stations.**

***Index Terms*—built environment; multi-source data; peak passenger flow; MGWR model; spatial heterogeneity**

I. INTRODUCTION

The sustained and rapid economic growth and acceleration of urbanization have led to an increase in the prevalence of urban transportation problems, which are now considered to be one of the key factors restricting sustainable urban development. Traffic congestion, environmental pollution, housing shortages, social pressure, and the continuous expansion of cities into surrounding areas urgently demand effective planning and management strategies. In response, Transit-Oriented Development (TOD), an innovative urban planning and traffic management concept, has increasingly attracted the attention of planners and researchers. TOD aims to promote high-density development, mixed land use, and public transportation-centered design principles. These strategies seek to reshape residents' travel behavior, reduce dependence on private vehicles, improve transportation efficiency, and foster sustainable urban development.

In their foundational study, Cervero et al. [1] introduced the "3D" framework as the primary defining characteristic of the Transit-Oriented Development (TOD) built environment, focusing on density, diversity, and design. Later, Ewing et al. [2] expanded this framework to the "5D" model by adding distance to transit and destination accessibility. The findings of the current research suggest that the 3D or 5D framework provides a valuable approach for analyzing the patterns of urban rail transit passenger flow and the factors that influence it.

The advancement of location-based service technology has sparked growing research interest in modeling methods using fine-grained Point of Interest (POI) data. Ma et al. [3] developed a Geographically Weighted Regression (GWR) model based on the traditional least squares method, utilizing POI data to represent and analyze the built environment characteristics of various transportation districts. This approach aimed to explore the influence of urban built environment factors on subway passenger flow from a micro-level perspective. Wang et al. [4] also applied a GWR model to examine the effect of the built environment on public bicycle usage, differentiating station types based on the scale of the POIs.

Yang et al. [5] developed a three-level random intercept binary logistic regression model and a geographically weighted binary logistic regression model to analyze the complex relationship between the built environment and the elderly's travel propensity. Du et al. [6] applied a spatial difference-in-differences (SDID) model alongside a GWR model to examine the spatial variation in the impact of new transit lines and surrounding land use on passenger flow at existing stations. In their study, Li et al. [7] used a Principal Component-Based Geographically Weighted Regression (PCA-GWR) model to explore the factors influencing passenger flow at railway stations, effectively addressing the issue of multicollinearity among explanatory variables.

However, the traditional GWR model assumes a uniform optimal bandwidth for all explanatory variables, disregarding the spatial scale differences in how TOD built environment factors affect the distribution of passenger trips. This can result in instability in the estimation results. To

Manuscript received June 13, 2024; revised November 22, 2024.

This work was supported in part by the National Natural Science Foundation of China (No. 72161024), "Double-First Class" Major Research Programs, Educational Department of Gansu Province (No. GSSYLXM-04), and Science and Technology Program of Gansu Province(22ZY1QA005).

Junyi Zhao is a postgraduate student at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: 1057391880@qq. com).

Changfeng Zhu is a Professor at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (Corresponding author, phone: +86 18919891566, e-mail: cfzhu003@163. com).

Kewei Zhao is a postgraduate student at School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, China. (e-mail: zkw198198@163. com).

address this issue, some scholars have begun using the enhanced GWR model to explore the spatial and temporal characteristics of the built environment's impact on passenger trips.

Peng et al. [8] applied both a global constant parameter model and a local variable parameter model, based on POI data, to investigate the interdependence between commuter flows and fine-grained land functions. Gao et al. [9] used a Multi-scale Geographically Weighted Regression (MGWR) model to analyze the spatial characteristics of the TOD built environment in relation to its impact on morning peak outbound passenger flow, integrating the 5D features. Pan et al. [10] constructed an MGWR model to reveal how built environment factors influence the demand for long- and short-distance internet taxi trips, as well as its spatial heterogeneity. Xu et al. [11] proposed a spatio-temporal geographically weighted random forest model (GTWR-RF) to analyze the spatial and temporal heterogeneity and nonlinearity of the influence of built environment features on passenger flow.

In light of the aforementioned considerations, this paper selects Xi'an as the study area and adopts a multi-source data approach, utilizing population data, road network data, POI, and other data sources, to construct a "5D" indicator system for the built environment. The impact of the built environment on rail transit passenger behavior is then analyzed using the MGWR model. The study also examines the mechanisms and outcomes of these impacts, along with their significance for achieving the goal of sustainable urban development. Furthermore, the paper discusses how these findings can provide urban planners and decision-makers with a more scientific basis for promoting the development of a greener and more efficient urban transportation system.

II. Selection of Indicators

A. Variable Selection

Considering the characteristics of rail transit travel in Xi'an and the challenges of data collection, the study selects 11 variables as explanatory factors for rail transit passenger travel behavior, classifying them according to the "5D" principle: density, diversity, design, transit distance, and destination accessibility [12]. The variables used in the study are listed in Table I.

B. Research Methods

B.1. Multicollinearity Test

First, the Pearson correlation coefficient (PCC) between the explanatory variables was calculated, and a correlation test was conducted to exclude variables with coefficients greater than 0.7. Second, the Variance Inflation Factor (VIF) was used to assess multicollinearity among the explanatory variables. The VIF is commonly used to determine the severity of multicollinearity. When the VIF values of the explanatory variables exceed a predefined threshold of 5, those variables are considered to be affected by multicollinearity.

B.2. Spatial Autocorrelation

Moran's I test is a crucial tool for evaluating the autocorrelation of spatial data, used to determine the correlation of a specific variable within a spatial distribution. The value of Moran's I statistic ranges from -1 to +1. A positive value indicates that the variable exhibits a spatially clustered distribution, a negative value indicates a spatially dispersed distribution, and a value close to zero suggests that the variable's distribution is nearly random within the spatial context. Moran's I is calculated using the following equation:

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

Where: x_i and x_j represent the values of an independent variable within the study units; n is the total number of regions, and w_{ij} is the spatial weight between study units i and j.

TABLE I
METRICS FOR THE 5D ELEMENTS OF THE BUILT ENVIRONMENT

Variable Category	Variable Symbol	Variable Meaning
Density	V_1	Number of urban residents per square kilometer of the study unit (persons/km²).
	V_2	Density of residential neighborhoods, commercial properties, and other residences within the study unit (residences/km²).
	V_3	Number of companies per unit area (companies/km²)
	V_4	Density of scientific, educational, cultural, medical, and other public services in the study unit (public services/km²)
	V_5	Density of dining, shopping, lodging, entertainment, and other commercial enterprises in the study unit (commercial enterprises/km²).
Diversity	V_6	Indicates the degree of land-use mix, calculated using the entropy index method, $H = -\sum p_{ih} \ln(p_{ih})$ where p_{ih} represents the proportion of land type h in transportation subarea i.
Design	V_7	Total road length per unit area (km/km²).
	V_8	Number of intersections per unit of area (intersections/km²)
Distance to transit	V_9	Average distance from the center of mass of the study unit to the nearest bus stop (km).
Destination accessibility	V_{10}	Number of bus stops per unit area (bus stops/km²)
	V_{11}	Distance from the center of mass of the research unit to the city center (km).

V_1=Population density, V_2=Residential density, V_3=Corporate density, V_4=Public service density, V_5=Commercial density, V_6=Land use mix, V_7=Road density, V_8=Intersection density, V_9=Distance to bus stops, V_{10}=Bus stop density, V_{11}=Distance to city center.

B.3. Regression Models

The GWR model can generate localized spatial statistical parameter estimates that reflect the spatial variation in regression contributions. The calculation formula is as follows:

$$Y_i = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \tag{2}$$

Where: Y_i represents the evening peak inbound (or outbound) passenger flow at station i , x_{ik} denotes the k-th built environment factor impacting the i-th study unit, $\beta_0(u_i, v_i)$ is a constant term, $\beta_k(u_i, v_i)$ is the k-th regression coefficient for the i-th study unit, and ε_i is a random error term.

The MGWR model overcomes the limitation of the classical GWR model, which only allows all variables to share the same bandwidth. The MGWR model allows for distinct bandwidths to be set for each variable, thereby reflecting their varying scale characteristics. The calculation formula is as follows:

$$Y_i = \beta_{bw0}(u_i, v_i) + \sum_{k=1}^{p} \beta_{bwk}(u_i, v_i) x_{ik} + \varepsilon_i \tag{3}$$

Where: $\beta_{bw0}(u_i, v_i)$ represents the regression constant term at bandwidth $bw0$; $\beta_{bwk}(u_i, v_i)$ represents the k-th regression coefficient for the i-th study unit at bandwidth bwk ; and ε_i is the random error term.

The model weights are calculated using a Gaussian distribution, and the optimal bandwidth is determined by minimizing the AICc criterion. The AICc value is calculated as follows:

$$A = \{2t - 2\ln[L(\hat{\theta}_i, x)]\} / i \tag{4}$$

Where: A is the AICc value; t is the number of independent parameters in the MGWR model; $\hat{\theta}_i$ is the maximum likelihood estimate; $L(\hat{\theta}_i, x)$ is the likelihood function for $\hat{\theta}_i$.

III. Empirical Analysis

A. Study Area

As a transportation hub in Northwest China, Xi'an is connected to several railways, highways, and air routes running in both east-west and north-south directions. Xi'an is a key node in China's high-speed rail and highway networks, commonly referred to as the "four verticals and four horizontals" and the "eight verticals and eight horizontals," respectively. By the end of 2022, the city had 11 districts and 2 counties under its jurisdiction, covering a total area of 10,108 square kilometers and housing a resident population of over 13 million. The study area focuses on six principal urban areas of Xi'an, with the stations of subway lines 1-6 serving as the main subjects of analysis. In this paper, the radius of a station's sphere of influence is set to 500 meters, and the intersection of the station's circular area with its Thiessen (Voronoi) polygon is determined. This intersection defines the station's sphere of influence. An overview of the study area is provided in Figure 1.

B. Data Sources

The study data include morning and evening peak passenger flow at stations, provided by Xi'an Railway Transportation Construction Group Co. on November 15,

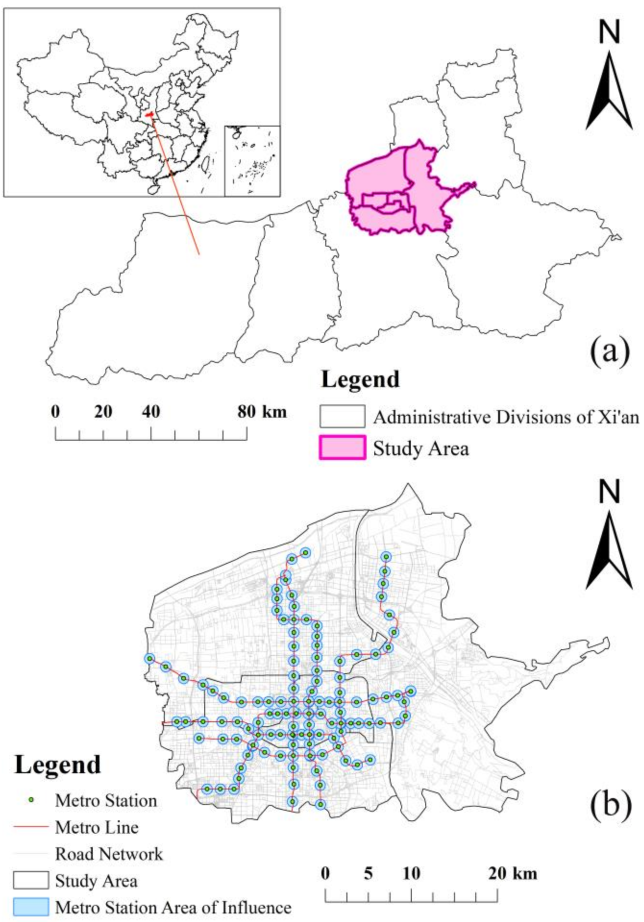

Fig. 1. Overview of the study area

2023. The specific point data for rail transit stations, bus stations, and various POI facilities around the stations were obtained from the Baidu Map application programming interface (API). Population density data were sourced from the WorldPop website (https://www.worldpop.org).

C. Multicollinearity Test and Spatial Autocorrelation Test

The Pearson correlation coefficients between the independent variables are analyzed in Table II. It shows that the correlation coefficient between "residential density" and "public service density" exceeds 0.70 with a significance level of $P < 0.05$, and the correlation coefficient between "residential density" and "distance to city center" is less than -0.70. Additionally, the correlation coefficient between "commercial density" and "public service density" is 0.782. All other correlation coefficients meet the required thresholds. To avoid multicollinearity among the independent variables, "residential density" and "commercial density" are excluded.

The multicollinearity and spatial autocorrelation tests for the screened variables are presented in Table III. The results show that all screened variables are statistically significant, with variance inflation factor (VIF) values ranging from 1.084 to 2.905, all below 5. This indicates that multicollinearity is not a concern among the selected independent variables. The Moran's I test for each variable produced p-values of less than 0.05 and positive z-values, indicating significant spatial correlation and clustering among the selected variables.

TABLE II
ANALYSIS OF PEARSON'S CORRELATION COEFFICIENT AMONG INDEPENDENT VARIABLES

	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}	V_{11}
V_1	1										
V_2	0.692**	1									
V_3	0.406**	0.389**	1								
V_4	0.532**	0.776**	0.340**	1							
V_5	0.430**	0.626**	0.391**	0.782**	1						
V_6	0.348**	0.413**	0.083	0.323**	-0.001	1					
V_7	0.150	0.270**	0.149	0.101	0.151	0.126	1				
V_8	0.148	0.432**	0.121	0.311**	0.327**	0.195*	0.466**	1			
V_9	0.061	0.065	0.022	0.040	0.049	0.109	0.280**	0.251**	1		
V_{10}	0.137	0.135	0.046	.215*	0.173	0.100	0.077	0.099	-0.001	1	
V_{11}	-0.541**	-0.749**	-0.194*	-0.645**	-0.560**	-0.471**	-0.386**	-0.489**	-0.211*	-0.077	1

Note: **, P<0.01, significant correlation; *, P<0.05, significant correlation.

TABLE III
MULTIPLE COVARIANCE TEST AND SPATIAL AUTOCORRELATION TEST

Variables	VIF	Spatial autocorrelation			
		Moran's I	E(I)	Z	P-value
V_1	1.798	0.707825	-0.007937	5.574722	0.000000
V_3	1.298	0.640999	-0.007937	5.139557	0.000000
V_4	2.188	0.894086	-0.007937	7.059460	0.000000
V_6	1.320	0.282628	-0.007937	2.318863	0.020402
V_7	1.484	0.566301	-0.007937	4.472254	0.000008
V_8	1.568	0.900454	-0.007937	7.068247	0.000000
V_9	1.124	0.251502	-0.007937	2.023912	0.042979
V_{10}	1.084	0.409252	-0.007937	3.261221	0.001109
V_{11}	2.905	1.023265	-0.007937	8.012699	0.000000

TABLE IV
RESULTS OF MODELING OPERATIONS

	Model	*RSS*	*AICc*	R^2	$Adj.R^2$
Morning peak inbound volume	GWR	85.030	333.756	0.330	0.279
	MGWR	41.948	320.481	0.670	0.544
Morning peak outbound volume	GWR	70.032	309.111	0.449	0.406
	MGWR	40.799	288.492	0.679	0.591
Evening peak inbound volume	GWR	49.590	265.275	0.610	0.579
	MGWR	27.762	245.815	0.781	0.717
Evening peak outbound volume	GWR	69.283	307.745	0.454	0.413
	MGWR	42.476	293.187	0.666	0.575

TABLE V
TABLE OF REGRESSION COEFFICIENTS FOR EACH VARIABLE

Variables	Morning peak inbound volume				Morning peak outbound volume			
	Mean	*Min*	*Median*	*Max*	*Mean*	*Min*	*Median*	*Max*
V_1	0.145	-0.527	0.183	0.596	0.073	-0.267	0.111	0.311
V_3	0.048	-0.422	0.102	0.727	0.027	-0.024	0.021	0.096
V_4	0.069	0.042	0.072	0.078	0.158	-0.128	0.210	0.252
V_6	0.108	0.085	0.095	0.164	0.154	0.122	0.162	0.174
V_7	-0.017	-0.117	-0.014	0.025	0.022	-0.102	0.032	0.125
V_8	0.021	-0.121	0.022	0.14	0.145	0.097	0.15	0.167
V_9	0.119	-0.120	0.133	0.251	-0.061	-0.080	-0.070	-0.013
V_{10}	-0.059	-0.180	-0.043	0.046	0.087	-0.122	0.085	0.274
V_{11}	-0.547	-1.064	-0.786	0.506	-0.318	-0.347	-0.326	-0.258

Variables	Evening peak inbound volume				Evening peak outbound volume			
	Mean	*Min*	*Median*	*Max*	*Mean*	*Min*	*Median*	*Max*
V_1	0.076	-0.211	0.072	0.394	0.063	-0.039	0.070	0.148
V_3	0.103	-0.011	0.088	0.246	-0.002	-0.017	-0.002	0.011
V_4	0.128	0.085	0.135	0.142	0.107	0.064	0.112	0.139
V_6	0.081	0.031	0.090	0.110	0.118	0.111	0.118	0.133
V_7	-0.017	-0.040	-0.014	-0.003	-0.087	-0.104	-0.086	-0.076
V_8	0.109	0.064	0.115	0.133	0.170	-0.107	0.113	0.731
V_9	-0.117	-0.156	-0.131	-0.042	-0.040	-0.100	-0.056	0.083
V_{10}	0.098	-0.083	0.092	0.296	0.092	-0.092	0.097	0.279
V_{11}	-0.466	-0.498	-0.480	-0.402	-0.388	-0.427	-0.393	-0.344

D.Comparison of Model Results

The GWR model and MGWR model were established based on the screened variables, and their fitting results were compared. Using MGWR2.2 software, the Spatial Kernel was set to Adaptive Bisquare, and Bandwidth Searching employed the Golden Section method. To eliminate the effects of different scales, Z-score standardization was applied to the variables. After running the MGWR model, the regression results were obtained. These were then compared with the GWR model, and the results of both models are shown in Table IV. Taking morning peak inbound passenger flow as an example, the goodness-of-fit R^2 of the MGWR model was 0.67, approximately 34% higher than that of the GWR model. Additionally, the residual sum of squares (RSS) and Akaike Information Criterion (AICc) were both reduced. This indicates that the MGWR model, which considers the spatial correlation of variables, more accurately reflects the relationship between inbound and outbound passenger flow and influencing factors across different time periods.

A comparison of the MGWR model's fitting effect (R^2) for evening peak inbound passenger flows reveals that the model has a stronger explanatory power for inbound flows than for outbound flows. The analysis of evening peak travel behavior shows that most outbound passenger flow consists of commuters returning to the station. This may be due to the exclusion of 'residential density' in the screening of independent variables, which has weakened the model’s explanatory ability for outbound flows.

E. Analysis of Model Results

The regression coefficients for each variable are shown in Table V. From the table, it is clear that different variables have varying degrees of impact on station ridership. Overall, high population density, mixed land use, and transportation accessibility (such as intersection density and bus stop density) all have a positive effect on public transit usage during peak hours. In contrast, the distance from the station to the city center and the distance to the nearest bus stop have a significant negative impact on ridership, indicating that commuting distance and the distribution of transit nodes greatly influence public transit ridership.

For the morning peak, population density and land use mix have a significant impact on station ridership, indicating that balanced development between residential and work areas is a key factor in determining morning peak traffic. The strong negative effects of the distance between the station and bus stops, as well as the distance to the city center, highlight the critical role that proximity to the city center and public transit nodes plays in commuting patterns. The density of companies and public services has a more pronounced effect on outbound ridership, suggesting that these areas are primary destinations for commuters.

For the evening peak, population density, public service density, and land use mix all have a significant positive impact on station ridership, indicating high mobility in these areas and an increased demand for travel during the evening rush hour. The negative effects of the distance to the city center and the distance to bus stops on evening peak ridership are notable, suggesting that geographic location and the accessibility of transit hubs are key factors influencing ridership. The positive impact of intersection density and bus stop density highlights that improved transportation convenience increases passengers' inclination to use public transit during the evening peak.

From a transportation perspective, the connectivity between public transit systems significantly influences the choice of rail transit. In some areas, despite a high road density, public transit still accounts for a substantial proportion of travel. In regions with higher road density, the likelihood of motor vehicle travel increases, creating a competitive dynamic with rail transit and impacting station ridership.

E.1. Density

The distribution of the population density impact coefficient is visualized in Figure 2. The influence coefficient of this variable on inbound passenger flow ranges from -0.21 to 0.39, while for outbound passenger flow, it ranges from -0.03 to 0.14. This reflects the transition of regional population density from aggregation to dispersion during the evening peak hour. Stations with higher population density for both inbound and outbound passenger flows are concentrated in the central and southwestern areas, with a weaker impact on stations at the eastern and northern edges. This is likely related to the location of the city's high-tech industrial development zone.

The distribution of corporate density impact coefficients is shown in Figure 3. This variable has a positive impact on inbound passenger flow, while the impact on outbound passenger flow is negative and has a smaller range of influence. Stations with a significant impact on inbound passenger flow are concentrated in the east, while those affecting outbound passenger flow are distributed in the west and south. The absolute value of the impact coefficient on inbound passenger flow is higher than that on outbound passenger flow, which, when analyzed alongside Figure 2, partially reflects the direction of evening peak commutes. Corporate density has a weaker influence on rail travel behavior in the southwestern region.

The distribution of the public services density impact coefficient is shown in Figure 4. This variable has a positive effect on passenger flow. Stations with a significant impact on inbound passenger flow are distributed across large areas in the central and southern regions, while those significantly impacting outbound passenger flow are primarily concentrated in the southwestern region.

E.2. Diversity

The distribution of the coefficient of influence for land use mix is shown in Figure 5. This variable has a positive effect on passenger flow, with a stronger impact on outbound passenger flow than on inbound passenger flow. The stations experiencing the greatest influence are primarily located in the central region, indicating that the diversity of land use in the city center is more attractive to rail transit passengers. Enhancing the diversity and balance of land use will encourage more residents to commute via subway.

E.3. Design

The distribution of the road density impact coefficient is shown in Figure 6. This variable primarily exerts a negative influence on passenger flow, indicating that in areas with higher road density, residents are more likely to choose alternative modes of transportation that compete with rail transit. The figure shows that road density has a stronger impact on inbound passenger flow, particularly in the northern and western regions, while its influence on outbound passenger flow is more significant in the western and southwestern areas.

The distribution of the intersection density impact coefficient is shown in Figure 7. This variable has a positive effect on inbound passenger flow and a particularly strong positive effect on outbound passenger flow in the southeastern region. In contrast, its influence in the western and northern regions is relatively minor, where it manifests as a negative impact.

E.4. Distance to Transit

The distribution of the coefficient of influence for distance to bus stops is shown in Figure 8. This variable predominantly has a negative effect on passenger flow. The distance between bus stops and subway stations significantly influences the potential to attract passengers to the subway. Conversely, the higher the density of bus stops near a subway station, the stronger the attraction of passenger flow. Stations with the greatest influence on passenger flow are located in the southern region.

The distribution of the coefficient of influence for bus stop density is depicted in Figure 9. The impact of this variable on passenger flow shows considerable spatial variability. It has a significant positive effect on subway passenger flow near the city center, while in the surrounding areas, its effect is relatively small and negative. This reflects the varying bus transfer situations in different regions.

E.5. Destination Accessibility

The distribution of the coefficient of influence for distance from the city center is shown in Figure 10. This variable negatively impacts passenger flow. Analysis of the figure reveals that it has a relatively minor effect on passenger flow at metro stations in the northern and western suburbs of Xi'an, while exerting a more pronounced influence in the southwest and central areas of the city.

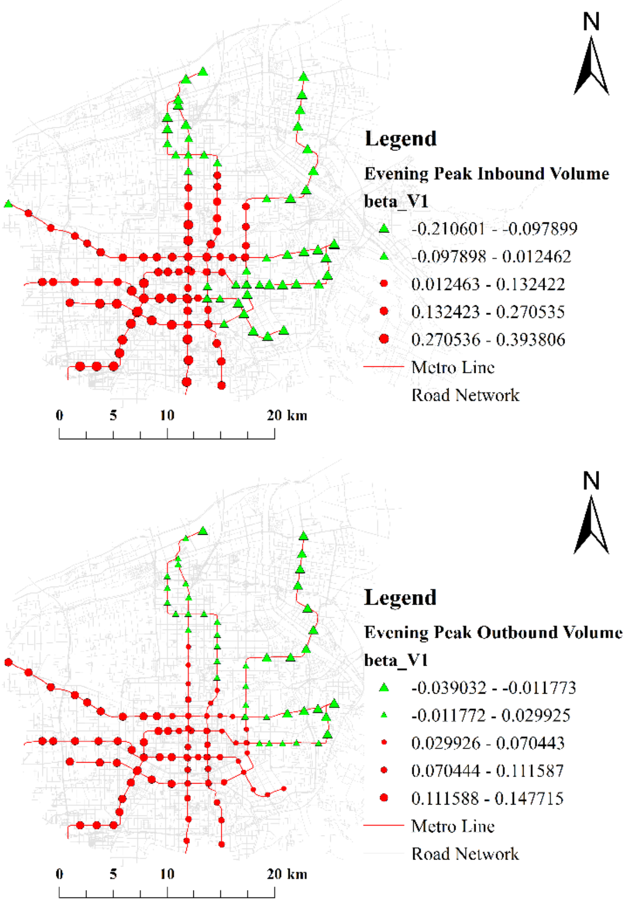

Fig. 2. Distribution of coefficients of population density impact

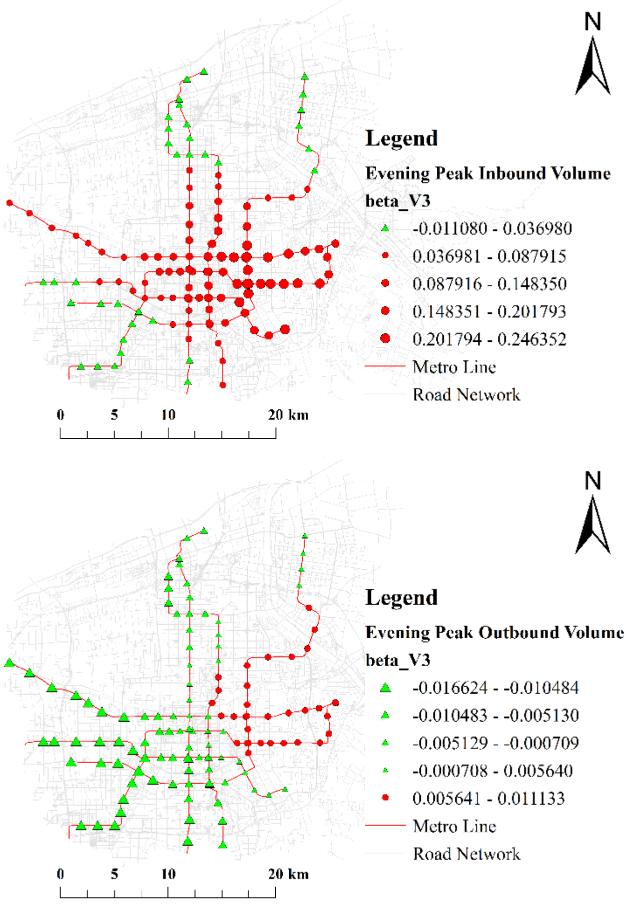

Fig. 3. Distribution of impact coefficients of corporate density

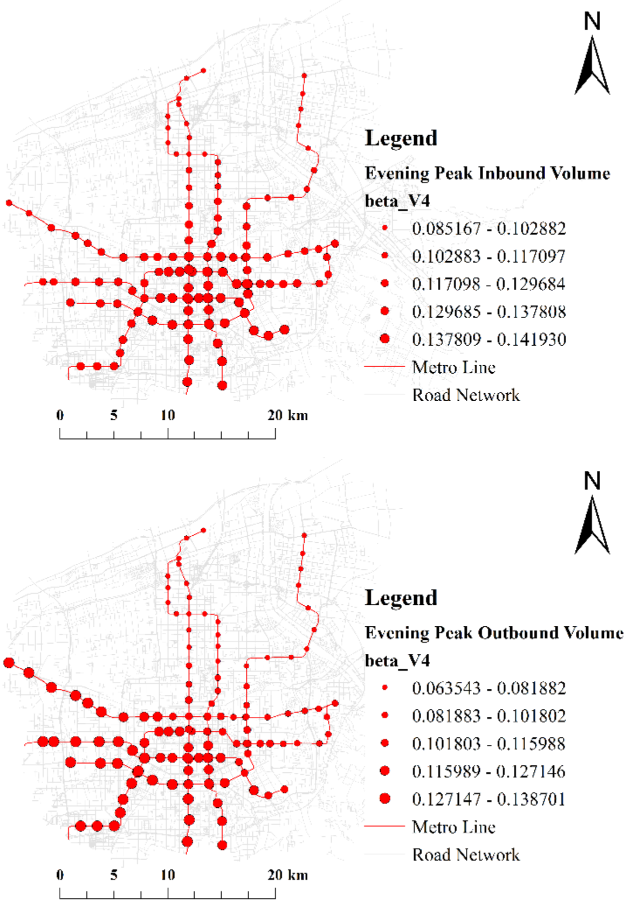

Fig. 4. Distribution of impact coefficients of public service density

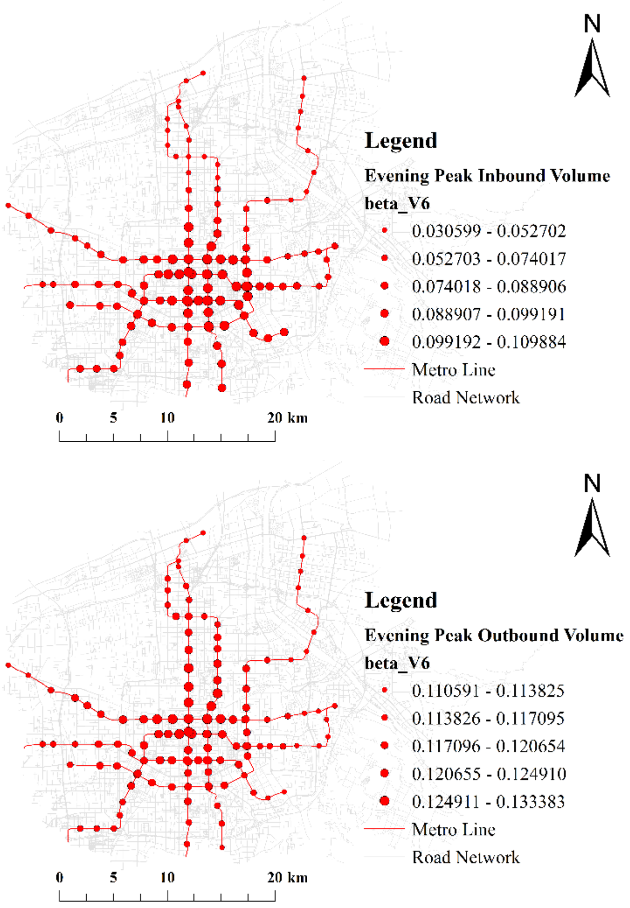

Fig. 5. Distribution of coefficients affecting land use mixing degree

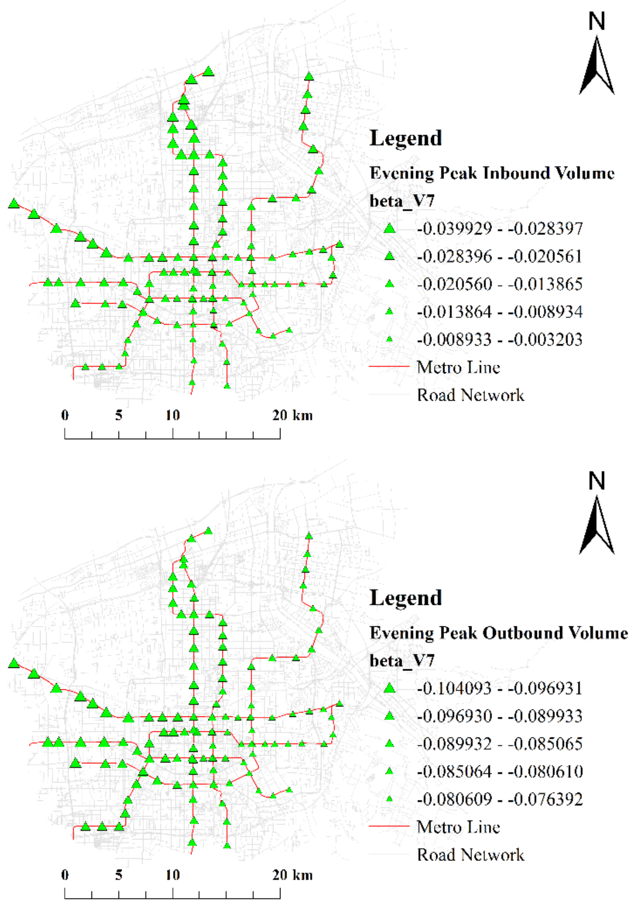

Fig. 6. Distribution of impact coefficients for road density

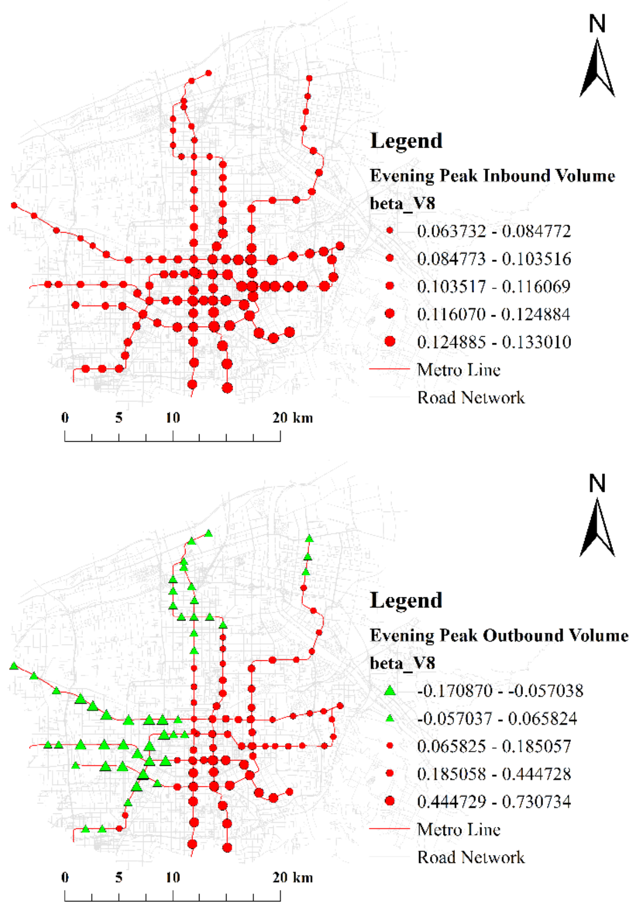

Fig. 7. Distribution of impact coefficients for intersection density

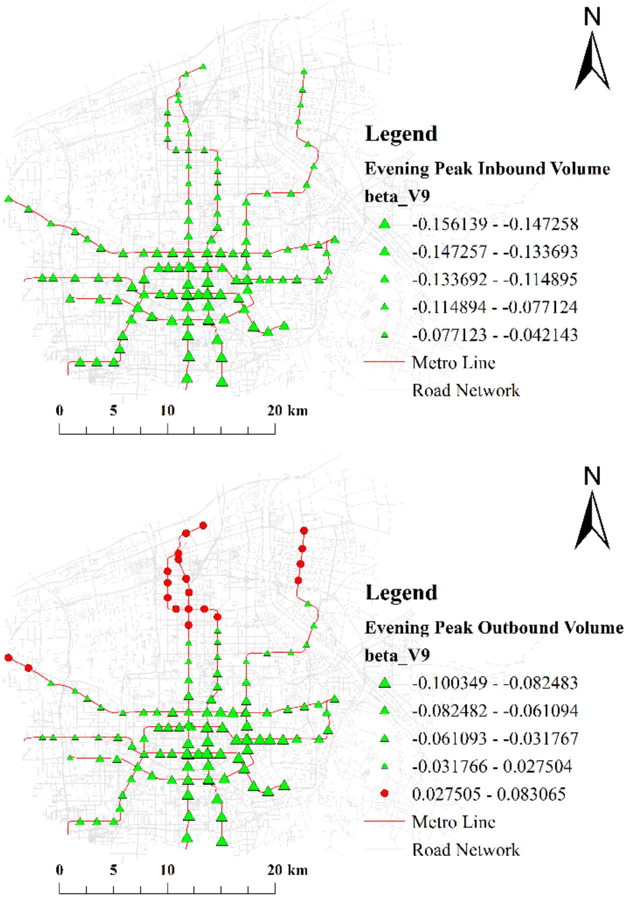

Fig. 8. Distribution of impact coefficients for distance to bus stops

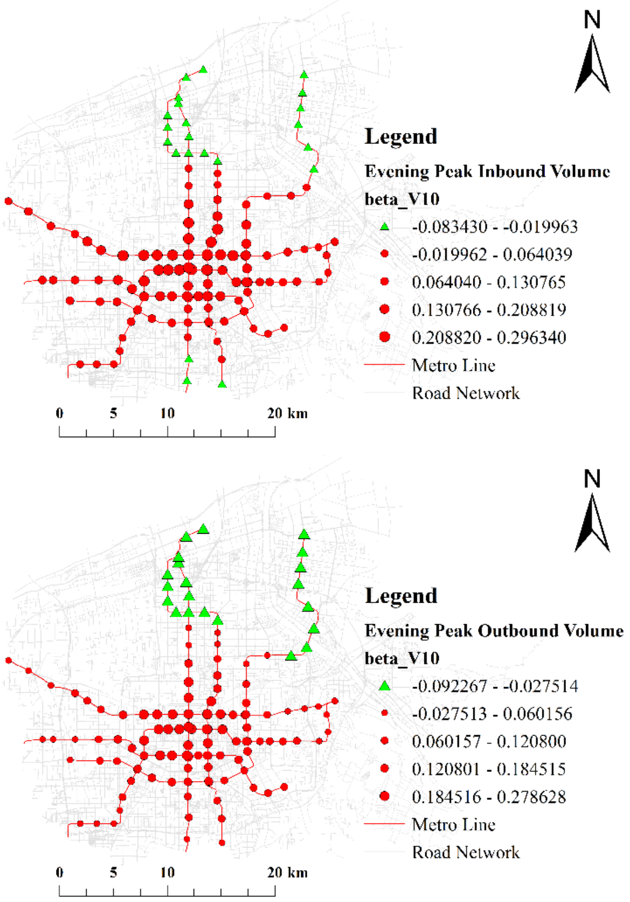

Fig. 9. Distribution of impact coefficients for bus stop density

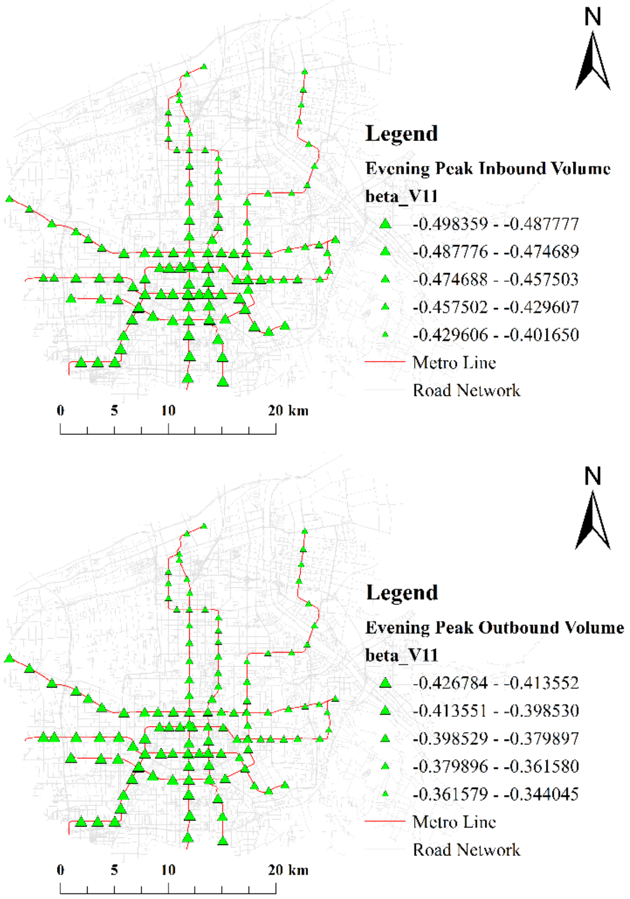

Fig. 10. Distribution of impact coefficients for distance to city center
The size of the symbols in Figure 2-10 reflects the absolute value of the influence coefficient, with red circles representing positive influence and green triangles representing negative influence.

IV. Conclusions And Discussions

This study examines the spatial distribution and scale effects of the built environment on metro ridership during peak hours in Xi'an, using a Multi-Scale Geographically Weighted Regression (MGWR) model. By analyzing the interactions between the "5D" built environment factors and metro ridership, this study provides a deeper understanding of how the built environment influences public transportation usage. The key findings are as follows:

(1) **MGWR model's superiority in capturing spatial heterogeneity**: Compared to the Geographically Weighted Regression (GWR) model, the MGWR model demonstrates higher explanatory power and precision in modeling the spatial heterogeneity of built environment influences on ridership. The MGWR model achieved an explanatory power of 78.1%, which is significantly higher than the GWR model. This improvement reflects MGWR's capacity to handle variations in spatial scales, providing a more detailed understanding of the spatial interactions between ridership and built environment factors during peak hours. The enhanced model performance underscores the importance of considering spatial heterogeneity in urban transportation research.

(2) **Built environment factors significantly shape metro ridership**: The results highlight that several key built environment factors have strong and significant impacts on metro ridership during peak hours. High population density, diverse land use, and improved transport accessibility (such as greater intersection and bus stop

density) positively influence metro usage, indicating that these factors make public transportation more attractive and accessible. In contrast, greater distances from the city center and bus stops negatively impact ridership, emphasizing the importance of proximity to transit hubs for enhancing public transport use. These findings align with transit-oriented development (TOD) principles, which emphasize compact, mixed-use development close to transit stations to encourage sustainable transportation.

(3) **Regional disparities call for tailored development strategies**: The influence of TOD-built environment factors varies significantly across different regions. Central urban areas benefit most from improvements in land-use diversity, while suburban stations rely more on enhancing transport connectivity to attract ridership. As a result, the study recommends adopting differentiated development strategies. In densely populated central areas, urban planning should focus on optimizing land use, improving the quality of development, and balancing employment with residential functions to ensure high efficiency of the transportation system. Meanwhile, in suburban areas, efforts should prioritize expanding the scale and intensity of development, improving transport networks, and enhancing connectivity between transit stations and surrounding areas to increase accessibility.

(4) **Implications for urban planning and transport policy**: This study provides valuable insights for urban planners and policymakers, particularly in designing more effective TOD strategies. The findings suggest that improving public transportation accessibility through strategic land use and infrastructure development is critical for increasing metro ridership. Planners should prioritize population density, land-use diversity, and transport connectivity when designing TOD policies, particularly around metro stations. In areas with lower ridership, improving access to transit and reducing distances between public transport nodes and residential or commercial areas will likely enhance public transportation uptake.

(5) **Evaluating the effectiveness of TOD policies**: The study's findings underscore the need for a nuanced approach to TOD policy implementation. Central areas with high density and diverse land use should focus on enhancing development quality, while suburban areas with lower density should emphasize expanding transport infrastructure and improving connectivity. These tailored strategies will help optimize public transportation usage across different regions, reduce traffic congestion, and promote sustainable urban development.

(6) **Future research directions**: While this study provides significant insights, future research could explore additional dimensions. For example, analyzing ridership patterns during off-peak hours or weekends could further enrich the understanding of how different time periods influence public transport usage. Moreover, applying the MGWR model to other cities with varying urban forms and transport systems could validate its broader applicability and identify unique regional characteristics that influence metro ridership. Future studies could also investigate the interaction between other modes of transportation, such as cycling and walking, with metro ridership, providing a more comprehensive picture of urban mobility.

REFERENCES

[1] Cervero R, Kockelman K. Travel demand and the 3Ds: Density, diversity, and design[J]. Transportation Research Part D,1997,2:199-219.

[2] Ewing R, Cervero R. Travel and the built environment: A synthesis[J]. Transportation Research Record,2001,1780,87-114.

[3] Ma Shuhong, Liao Guomei, Huang Yan et al. Heterogeneous effects of built environment on underground commuter flows in transit neighborhoods[J/OL]. Journal of Jilin University (Engineering Edition),1-10[2024-03-22]. https://doi.org/10.13229/j.cnki.jdxbgxb20221236.

[4] Wang Tao, Ji Xiaofeng. Analysis of the impact of built environment on public bicycle travelling mode based on GWR[J]. Journal of Hefei University of Technology (Natural Science Edition),2023,46(07):963-971.

[5] Yang Linchuan, Zhu Qing. Spatial heterogeneity in the influence of built environment on travel behavior of the elderly[J]. Journal of Southwest Jiaotong University,2023,58(03):696-703.

[6] Du Q, Huang Y, Zhou Y, et al. Impacts of a new urban rail transit line and its interactions with land use on the ridership of existing stations[J/OL]. Cities, 2023, 141: 104506. DOI:10.1016/j.cities.2023.104506.

[7] Li Yijun, Luo Ziyu, Zhou Tao et al. Based on PCA-GWR method to explore the impact of built environment on passenger flow of rail stations[J]. Railway Transport and Economy,2024,46(02): 159-166.DOI:10.16668/j.cnki.issn.1003-1421.2024.02.20.

[8] Peng Shiyao, Chen Shaokuan, Xu Qi et al. Spatial characteristics of land use and rail passenger flow based on POI[J]. Journal of Geography,2021,76(02):459-470.

[9] Gao D. H., Xu Q., Chen P. W. et al. Spatial characterization of urban rail passenger flow and fine-scale built environment[J]. Transportation Systems Engineering and Information,2021,21(06):25-32.DOI:10.16097/j.cnki.1009-6744.2021.06.004.

[10] Pan Yiyong, Xu Jiacong, You Yiwen et al. Multi-scale spatial heterogeneity analysis of factors influencing the demand for online car travel[J/OL]. Journal of Jilin University (Engineering Edition):1-9[2024-03-20]. https://doi.org/10.13229/j.cnki.jdxbgxb.20230789.

[11] Xu Xinyue, Kong Qingxue, Li Jianmin et al. Analysis of the influence of built environment on the spatiotemporal heterogeneity of rail passenger flow[J]. Transportation Systems Engineering and Information,2023,23(04):194-202+281.DOI:10.16097/j.cnki.1009-6744.2023.04.020.

[12] Xia Zhengwei, Zhang Ye. From "5D" to "5D+N": a study on the influencing factors of TOD effectiveness in English literature[J]. International Urban Planning,2019,34(05):109-116.

Junyi Zhao was born in Shaanxi, China, in 2000. He obtained his Bachelor degree in Traffic and Transportation from Lanzhou Jiaotong University, Lanzhou, China, in 2022.

He is currently pursuing a master's degree in Traffic and Transportation (Traffic and Transportation Planning and Management) at Lanzhou Jiaotong University. His research interests include Transit Oriented Development and Transportation Planning and Management.