Relationship Between Spatial Form, Functional Distribution, and Vitality of Railway Station Areas Under Station-City Synergetic Development: A Case Study of Four Special-Grade Stations in Beijing

Yuhan Sun, Bo Wan * and Qiang Sheng *

School of Architecture and Design, Beijing Jiaotong University, Beijing 100044, China; 22121762@bjtu.edu.cn
* Correspondence: wanbo@bjtu.edu.cn (B.W.); qsheng@bjtu.edu.cn (Q.S.)

Abstract: The integration of railway stations into urban environments necessitates a detailed examination of their vitality and influencing factors. This study assesses urban vitality around four major railway stations in Beijing utilizing a variety of analytical models including Ordinary Least Squares, Geographically Weighted Regression, Multi-Scale Geographically Weighted Regression, and machine learning approaches such as XGBoost 2.0.3, Random Forest 1.4.1.post1, and LightGBM 4.3.0. These analyses are grounded in Baidu heatmaps and examine relationships with spatial form, functional distribution, and spatial configuration. The results indicate significant associations between urban vitality and variables such as commercial density, average number of floors, integration, residential density, and housing prices, particularly in predicting weekday vitality. The MGWR model demonstrates enhanced fit and robustness, explaining 84.8% of the variability in vitality, while the Random Forest model displays the highest stability among the machine learning options, accounting for 76.9% of vitality variation. The integration of SHAP values with MGWR coefficients identifies commercial density as the most critical predictor, with the average number of floors and residential density also being key. These findings offer important insights for spatial planning in areas surrounding railway stations.

Keywords: station-city synergetic; spatial form; street function; street vitality; railway stations

1. Introduction

As global urbanization accelerates, urban transportation hubs—such as train and subway stations—play a pivotal role in urban development. These hubs serve not only as central nodes in transportation networks but also as key elements in the spatial organization of cities. The redevelopment of these hubs increasingly influences urban growth [1–3]. Recently, the station-city synergetic model has become a significant strategy in urban planning worldwide. For instance, Transport for London's 2021 Sustainable Development Framework highlights the importance of effectively linking transportation hubs with urban areas to promote sustainable urban growth. Similarly, the Chinese government's "14th Five-Year Plan for the Development of a Modern Integrated Transportation System" emphasizes the synergistic effects between transportation hubs and urban development [4].

In the context of these policies and strategies, the promotion and implementation of Transit-Oriented Development (TOD) models become particularly crucial. TOD enhances the utilization of public transportation systems by encouraging high-density, multifunctional urban space configurations around transit hubs, which not only optimizes public transport usage but also boosts the vitality and quality of life in surrounding areas [5,6].

Railway stations, compared to traditional public transportation modes, possess a stronger capacity for passenger aggregation and spatial radiation, making them more applicable to TOD theory. This aggregation effect of railway stations leads to higher

land-use density and more intensive spatial development in the areas surrounding train stations [7]. With the large-scale construction of railways, the spatial morphology and functional structure of areas surrounding railway stations have fostered new development models [8]. The role of railway hubs has expanded from solely being transportation nodes to incorporating multiple functions, such as consumption and residential activities [9,10]. This transformation encourages the public to perceive train stations not merely as transit points but as vibrant, multifunctional urban spaces [11].

As a typical example of China's rapid urbanization, railway passenger stations in Beijing hold a significant position in urban planning. Although the implementation of TOD around train stations is not yet fully realized, its potential in urban planning is gradually becoming evident. As illustrated in Figure 1, passenger volumes at Beijing's train stations steadily increased from 2000 to 2019, dropped significantly in 2020 due to the pandemic, and rebounded notably in 2021. This trend underscores the need to further advance the integration of station areas with urban development to support future economic growth and urban vitality.
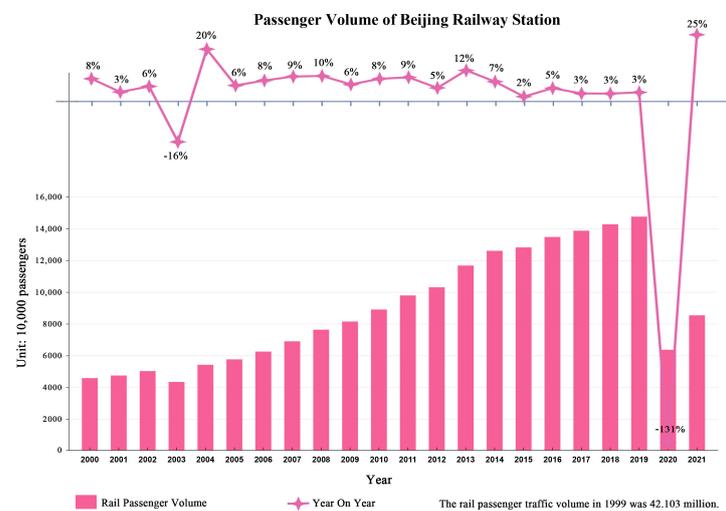
Figure 1. Passenger volume of Beijing Railway Station. Data source: China Railway Beijing Bureau Group Co., Ltd., Beijing, China.

Research on the vitality of railway stations and their surrounding areas remains limited. Traditional methods for measuring vitality struggle to capture its dynamic nature, but emerging technologies and methods now enable more comprehensive, multi-dimensional studies of spatial vitality around railway stations. While previous studies have mainly examined how the presence of stations leads to significant changes in spatial structure and land-use patterns [12], less attention has been given to exploring the relationship between spatial vitality and form within various station influence zones. Space syntax, as a tool for analyzing spatial structures, helps quantify the geometric and topological properties of street networks, revealing the potential influence of spatial form on vitality [13]. Most existing studies on vitality rely on either Ordinary Least Squares (OLS), geographically weighted regression, or machine learning methods in isolation or in pairs to assess factors influencing vitality [14]. However, the vitality of railway station areas is a multi-faceted concept, involving aspects such as spatial form, functional distribution, and transportation connectivity. To gain a comprehensive understanding of the combined influence of these factors, this study integrates OLS, Geographically Weighted Regression (GWR), Multi-Scale Geographically Weighted Regression (MGWR), and machine learning techniques (XGBoost, Random Forest, and LightGBM), offering a comparative analysis of their performance. This combined approach not only uncovers the diverse factors affecting vitality around railway stations but also identifies consistent influences across different methods.

In response to these research gaps, this study applies an integrated analytical framework to examine the relationship between vitality, spatial form, and function in areas surrounding four special-grade railway stations in Beijing. The primary contributions of this study are (1) a comprehensive analysis of the relationship between spatial form, functional distribution, and urban vitality in railway station areas; (2) the integration of spatial configuration metrics with OLS, GWR, MGWR, and machine learning approaches, offering an innovative methodology for examining factors influencing urban vitality; and (3) the identification of consistent factors that affect the vitality of special-grade railway station areas in Beijing across different analytical methods, providing valuable insights for urban planners and policymakers in the context of station area development and urban renewal.

The remainder of this paper is structured as follows: Section 2 reviews the relevant literature on station area development and urban vitality. Section 3 details the research methodology. Section 4 presents the empirical results of the analysis. Section 5 discusses the broader implications of the findings. Finally, Section 6 concludes the study and offers directions for future research.

2. Literature Review

2.1. The Vitality of Railway Station Areas

Urban vitality is a product of the interaction between human activities and spatial environments, serving as a crucial indicator of a city's developmental capabilities [15]. Therefore, urban vitality is often measured by the spatial concentration of human activities [16]. Urban vitality involves multiple scales and types of cities. Currently, the research object of urban vitality has gradually shifted from the macro spatial scale to the meso and micro spatial scales, such as urban parks, commercial centers, community centers, and waterfront spaces [17–19]. Research on the vitality of railway stations has gradually evolved into a significant subfield with the development of the TOD model. However, the aspect of vitality has received relatively little attention. Various methods have traditionally been employed to measure and analyze urban vitality, including travel surveys, observational interviews, and statistical data analysis [20–23]. However, the dynamic nature of vitality means traditional data often fall short in capturing its nuances. With advancements in information and communication technologies, new data sources such as POI, public review data, and mobile signaling data have become available, providing novel ways to measure urban vitality [24–27]. This study utilizes Baidu heatmap data to illustrate the vitality of areas surrounding Beijing's special-grade stations. Baidu heatmaps provide real-time data on population density, offering a dynamic and convenient means of capturing human activity patterns [28–30].

2.2. Factors Influencing the Vitality of Railway Station Areas

The planning and development of railway station TODs have reshaped the original urban structure, leading to a high concentration of commercial, residential, and recreational facilities. Studies have shown that spatial form [31–33] and function [34–36] are key factors influencing vitality.

Theoretically, urban vitality is closely linked to spatial form and function (i.e., the built environment) [36]. An optimized functional layout can enhance the attractiveness and convenience of railway passenger stations, promote their role as transportation hubs, attract large crowds for consumption, and effectively boost station vitality [37]. Research indicates that in the 26 railway stations of the Yangtze River Delta urban agglomeration, these stations, as significant large-scale infrastructure, facilitate the development of various industries such as commercial services, communication, entertainment, retail, real estate, and tourism through efficient transportation networks. This, in turn, drives industrial structure adjustment and population mobility [38]. When examining functional layout and vitality, the focus is often on the density and diversity of functions [39].

Spatial form provides the foundation for understanding the factors influencing urban vitality. TOD highlights how these elements enhance vitality by optimizing public transportation systems and improving the pedestrian environment. Specifically, urban morphology shapes the quality of daily life and promotes urban vitality by determining the availability of space, density, and accessibility [40,41]. However, various studies employ different calculation methods and indicators, with single indicators often failing to capture the complexity of urban morphology [42,43]. Therefore, it is essential to utilize multiple variables in analysis, including floor area ratio, building density, building height, road connectivity, street greening, and regional location [44–48]. Given China's extensive railway network and the numerous stations linking cities, assessing and analyzing the development status and spatial form of station areas is crucial. Such evaluations can optimize the spatial layout and design of passenger stations, enhance their vitality, and inform urban planning and traffic management, thereby promoting sustainable urban development [49]. Additionally, current empirical analyses tend to draw lessons from exemplary foreign stations, lacking comprehensive reviews of the status and challenges of domestic mega-city cases.

Space syntax is a set of theories and techniques used to analyze spatial configuration in buildings, urban areas, and other environments [50]. In addition, space syntax has been extensively applied in various morphology-related studies [51]. Scholars, building on road network structures, have developed a space syntax-based network analysis method to explore network configurations [52]. As transportation hubs, railway passenger stations connect different regions. Their high accessibility allows more people to conveniently reach the stations, thereby increasing foot traffic and transportation efficiency [53]. Factors such as station accessibility and transportation network connectivity significantly influence the vitality and development of railway station areas [54–56]. Moreover, by improving transportation accessibility, railway stations and their surrounding areas play a crucial role in promoting urban economic development [57–62]. Integrating space syntax analysis with node location models is essential for assessing the accessibility and attractiveness of railway stations, providing key insights into the dynamic mechanisms of vitality [63]. In general, urban spatial organization, particularly the street network, is a critical factor in shaping both urban diversity and vitality.

2.3. Methods for Studying the Factors Influencing Vitality in Railway Station Areas

Earlier research has frequently employed OLS models to examine the relationship between vitality and the factors influencing it due to the models' simplicity and ease of interpretation, making them widely favored by researchers [64–67]. Recently, scholars have turned to geographically weighted models to investigate these connections. These models have proven effective in analyzing spatiotemporal differentiation patterns and offer researchers a more nuanced spatial analysis perspective [68]. For instance, the Multiscale Geographically Weighted Regression (MGWR) model allows researchers to examine the relationship between urban vitality and various influencing factors across different geographic scales. This approach is particularly useful for datasets that display significant spatial variations or non-stationarity [69]. Beyond geographically weighted models, machine learning methods have opened new pathways for exploring the nonlinear relationships between urban vitality and its determinants [70,71]. Tree-based algorithms such as Extreme Gradient Boosting (XGBoost), Random Forest Regression (RF), and LightGBM are particularly adept at revealing the complex interactions between influencing factors and dependent variables [72,73]. Previous studies have primarily relied on single methods or combined OLS with GWR for analysis.

Overall, when researching urban vitality issues, using a multifaceted analysis framework that includes Multiple Linear Regression, Geographically Weighted Regression (GWR and MGWR), and machine learning models (such as Random Forest, XGBoost, and LightGBM) provides a comprehensive approach. Multiple Linear Regression is extensively applied in many studies on urban vitality, offering a macroscopic view and quickly identifying key factors. However, its linear assumptions may lead to oversimplifications or

incorrect calculations of the actual impact of the built environment, failing to elucidate the interactions between environment and behavior [74]. Furthermore, some variables may only show their impact after reaching a certain threshold, and their effect might saturate beyond that threshold [75]. Machine learning models excel in capturing complex non-linear relationships and spatial heterogeneity. However, models like Random Forest have limitations: they are prone to overfitting, leading to conclusions that are restricted and non-logical; they lack significant statistical indicators like those found in linear regression; and as global models, they do not recognize spatial heterogeneity [76]. Using spatially weighted regression models like GWR or MGWR can reveal spatial heterogeneity in the data. Therefore, integrating these models can achieve a complementarity of strengths, providing more comprehensive and accurate analysis results and offering a scientific basis for urban planning and policy-making.

In this context, this paper selects four top-grade railway stations in Beijing as case studies to explore the similarities and differences in their development conditions through comparative analysis. First, the spatial form of the railway station areas is quantitatively analyzed by calculating indicators such as block density and floor area ratio. Second, the functional distribution of streets is examined using POI data, and the spatial distribution of vitality is presented through Baidu heatmaps. Subsequently, an urban street network model is constructed based on spatial syntax theory, and multiple models are introduced for comparative analysis to explore their complex relationships with vitality. Based on these findings, relevant policy recommendations for urban planning are proposed.

3. Data and Methods

3.1. Study Area and Data Sources

Beijing, the birthplace of China's railway industry, holds significant historical importance and is home to one of the country's largest railway hubs—the Beijing Railway Hub [77]. This study examines four premium railway stations in Beijing: Beijing Station, Beijing West Station, Beijing South Station, and Beijing Fengtai Station. These special-grade stations are large in scale and handle significant passenger volumes. Figure 2 illustrates the passenger flow trends for the four special-grade stations in Beijing from 2011 to 2020. During this period, Beijing West Station consistently had the highest passenger volumes, peaking at 55.76 million passengers in 2019. Beijing South Station, which was renovated after China entered the high-speed rail era in 2008, saw its passenger flow steadily increase from 2011 and exceed 50 million passengers in 2019. In contrast, Beijing Station's passenger flow remained relatively stable, while Beijing Fengtai Station's passenger flow dropped to zero after it ceased passenger services in 2010. All these stations are located near the city center and exhibit a strong interaction with the surrounding urban areas [78]. Based on the concentric zone theory and considerations of walking speed [79,80], this study focuses on a 2000 m radius around each station as the core research area (Figure 3). To further clarify the study area, an 800 m service zone centered on each station is also defined. These stations vary in size, age, and location, providing a representative sample of stations in the urban core (Table 1).

Table 1. Basic information of the four major first-class stations.

Station Names	First Time of Use	Renovation Completion Time	Ranking	Distance from Station to City Center
Beijing Railway Station	1903	2004	First-class station	3.6 km
Beijing West Railway Station	1996	2005	First-class station	8.0 km
Beijing South Railway Station	1897	2008	First-class station	8.1 km
Beijing Fengtai Railway Station	1895	2022	First-class station	14.6 km
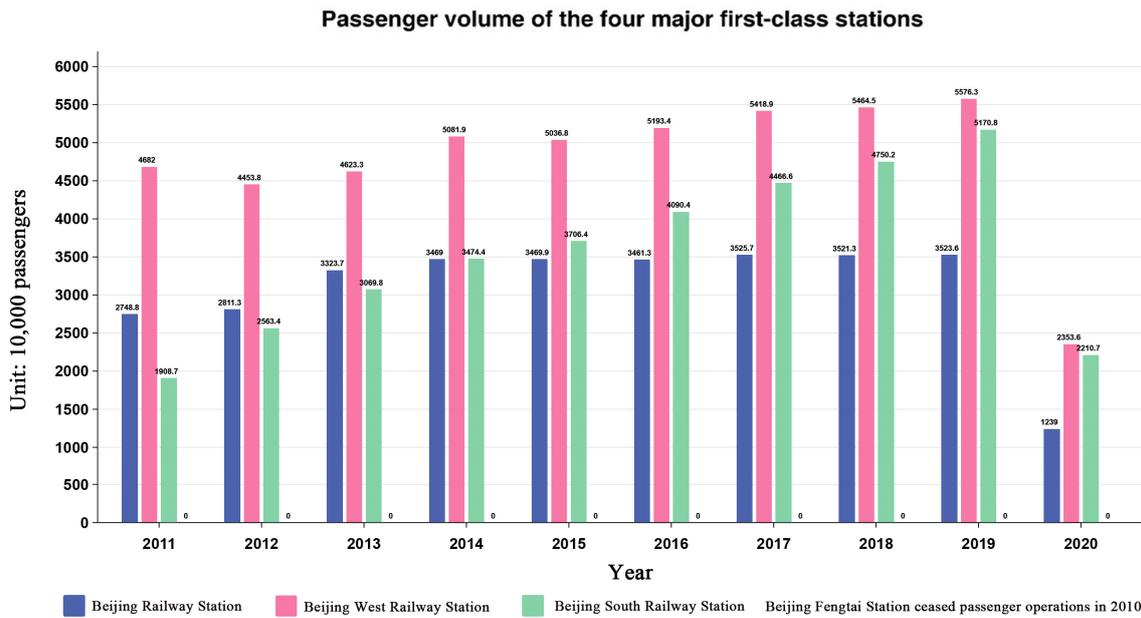
Figure 2. Passenger volume of the four major special-grade stations. Data source: China Railway Beijing Bureau Group Co., Ltd.
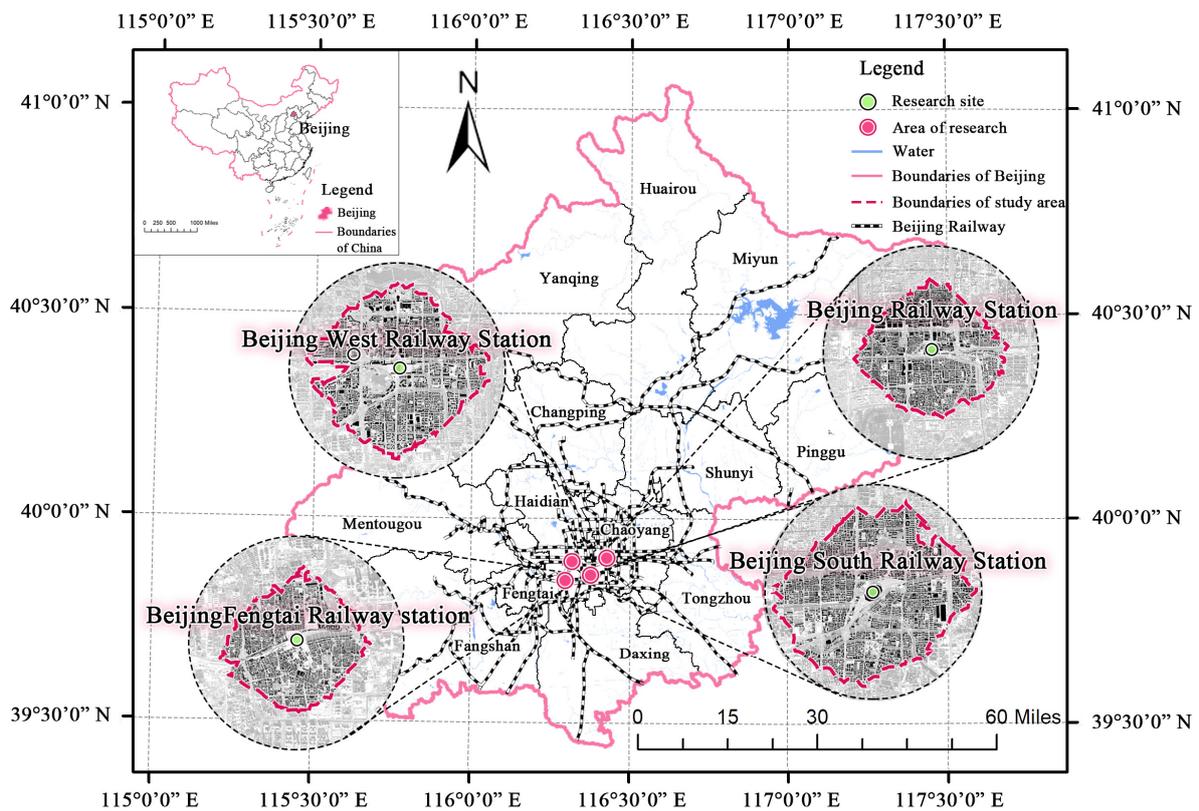
Figure 3. Study area.

The research data include population heat map data based on Baidu Huiyan, 2023 POI data for Beijing, 2023 road network data for Beijing, and 2022 building vector data for Beijing (Table 2).

Table 2. Data source statistics.

Data Type	Data Name	Data Source	Year	Link
Basic geographic data	Vector data of Chinese maps	National Geospatial Information Center	2022	http://www.ngcc.cn/
	Road network data	Official website for publicly available street maps	2023	https://www.openstreetmap.org/
Open-source data on the Internet	POI data	Gaode Map crawler	2023	https://ditu.amap.com/
	Baidu heatmap	Baidu Map crawler	2023	https://map.baidu.com/
	Building height and floorcount data	Gaode Map crawler	2023	https://ditu.amap.com/

(1)   Road Network Processed with Space Syntax

The pedestrian network data processed using space syntax were sourced from OSM maps. Initially, the Depthmap 1.0 software was employed to analyze the CAD road network through space syntax, converting the axial map into a segment model. This model then underwent preliminary generation and verification. Various configurational metrics were computed next. To ensure comparability, global and local configurational values were measured using NACH_Rn (choice), NACH_R1000m (choice), NAIN_Rn (integration), and NAIN_R1000m (integration). Building outline data were extracted from Gaode maps and converted into point features within GIS. These points were aligned with the nearest segment model and intersected to attach segment values to the point features. Finally, these points were linked with building outline data and visualized. Figure 4 illustrates the distribution of configurational measurement values within a 2 km radius around the four premium stations in Beijing. The color values indicate the degree of configuration in different areas, revealing significant spatial configuration differences across regions.
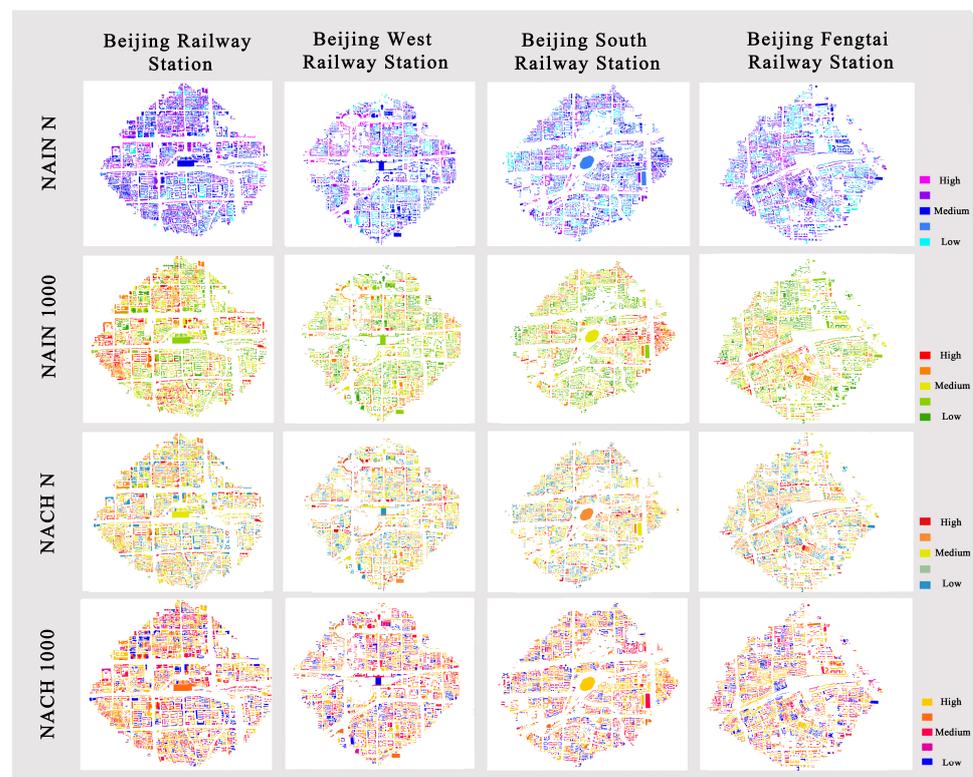
Figure 4. Spatial road network parameter values.

(2)   POI Data Reclassification

Considering the characteristics of the areas surrounding railway stations, the original POI data were reclassified. The final functions are shown in Figure 5.

POI classification	Major functional categories	Sub-functional categories	Number of POIs
Residential communities	Commercial residential	Residential area	858
Corporate office	Companies and enterprises	Companies	2467
Commercial shopping	Food and dining	Cake and dessert, coffee, Chinese cuisine, foreign cuisine, snacks and fast food, etc	8989
	Shopping and retail	Department stores, convenience stores, supermarkets, home and building materials, etc.	
	Hotels and accommodation	Five-star hotels, four-star hotels, three-star hotels, etc.	
Transportation services	Transportation facilities	Subway stations, bus stops, parking lots	2527
Tourism and leisure	Recreation and entertainment	KTV, cinema, bars, amusement parks, theaters, etc.	648
	Tourist attractions	Zoo, park, memorial hall, scenic spots, religious sites, etc.	
	Sports and fitness	Fitness center, camping site, taekwondo, multi-purpose sports arena, etc.	
Public services	Life services	Lottery sales, beauty and hairdressing, photography and printing, laundry services, etc.	3962
	Automotive-related	Charging stations, used cars, automotive repairs, automotive parts, etc.	
	Healthcare	Pharmaceutical sales, clinics, specialized hospitals, general hospitals, etc.	
	Science, education, and culture	Museums, news publishing, cultural centers, higher education, etc.	

Figure 5. POI categories.

(3)   Baidu Heatmaps

This study collected Baidu heatmap data from 27 March to 2 April 2023 at 60 min intervals yielding a total of 168 heatmaps. Additionally, to better quantify urban vitality, the study area was divided into 300 m × 300 m grid cells.

3.2. Methods

The research framework, as shown in Figure 6, comprises two main components: (1) quantifying spatial form using the floor area ratio and building density, and measuring vitality through multi-source spatial big data and (2) introducing and comparing five multivariable analysis models: Ordinary Least Squares (OLS), Geographically Weighted Regression (GWR), Multi-Scale Geographically Weighted Regression (MGWR), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Light Gradient Boosting Machine (LightGBM). This includes a comparative analysis of Multiple Linear Regression, Geographically Weighted Regression, and machine learning methods.

The geographically weighted regression model allows regression coefficients to vary spatially via a geographical weight function, capturing local effects. Machine learning models leverage their robust global predictive capabilities to assess the overall impact of independent variables on dependent variables. These models are evaluated for their accuracy, complexity, and ability to explain spatial heterogeneity under global and local effects. The most suitable model is then selected to explore the influence of spatial form and functional distribution in railway station areas on street vitality, as well as to understand variable differences under global and local effects.
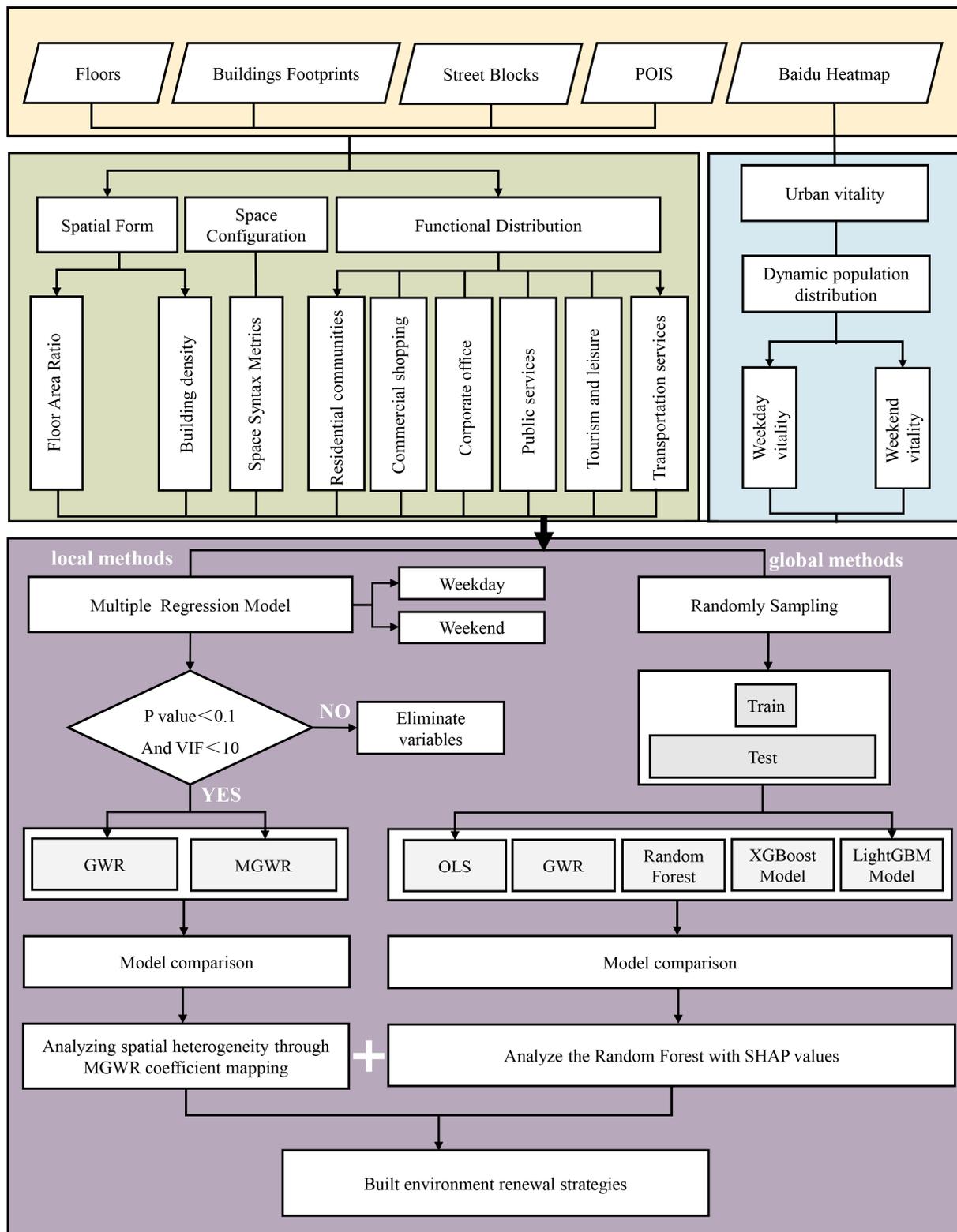
Figure 6. Research framework diagram.

This study identifies three primary categories of factors affecting vitality: spatial form, functional properties, and street network configuration. For calculating various spatial configuration indicators, different radii were used, including 250 m, 500 m, 1000 m, 1600 m, 2000 m, 5000 m, 10 km, 20 km, 50 km, and N. Table 3 presents the spatial configuration indicators that are significantly related to vitality.

Table 3. Impact factors and calculation methods table.

Category		Explanation	Calculation Method	N	Min	Max	Mean	SD
Spatial Form	Floor Area Ratio	Where Fi represents the building floor area (m^2) of the *i*-th block, BAi represents the built-up area (m^2) within the *i*-th block, and Ai represents the total area (m^2) of the *i*-th block.	$FAR_i = \frac{F_i}{A_i}$	701	0.000	6.762	1.283	0.866
	Building Density		$BCR_i = \frac{BA_i}{A_i}$	701	0.000	86.096	22.142	11.116
	Average Number of Floors	Average Number of Floors of Buildings	$F_i = \frac{\sum_{i=1}^{n} N_i}{n}$	701	0.000	22.000	4.874	2.921
Urban Function	POI Density	Density of various functional POIs on streets	$D_x = -\sum_{i=1}^{n}(p_i \ln p_i), (i1,2,3\ldots,n)$	701	0.000	1.769	1.040	0.484
	Tourism and Leisure Density	Kernel density of Tourism and Leisure POIs	$f_n(x) = \frac{1}{n\delta}\sum_{i=1}^{n}K\left(\frac{distance_i}{\delta}\right),$ where $K(x distance_i \delta)$ is the kernel density; $\delta > 0$ is the bandwidth, also known as the search radius; *n* is the number of known points; and *distance_i* represents the distance from the estimated point *X* to the sample point *Xi*.	701	0.000	56.149	12.266	10.040
	Transportation Facility Density	Kernel density of POIs related to transportation facilities		701	0.000	132.629	47.567	26.315
	Public Service Density	Kernel density of Public Service POIs		701	0.000	581.753	104.060	77.364
	Enterprise Density	Kernel density of Company/Enterprise POIs		701	0.000	284.088	45.923	42.902
	Commercial Density	Kernel density of Commercial POIs		701	0.000	2455.276	249.745	307.129
	Residential Density	Kernel density of Residential POIs		701	0.000	52.601	16.125	10.056
Spatial Configu-ration	Housing prices	The average price of a house for sale in a block per unit		701	0.000	0.000	51,654.176	48,702.214
	NAIN 5000 m	Enables cross-scale comparison between different urban systems. Length of a geodesic (shortest path) between vertices, considering the urban system's tendency to optimize travel distance.	$NAIN_\theta(x) = \frac{(n+2)^{1.2}}{(\sum_{i=1} d_\theta(x,i))}$	680	0.820	2.216	1.488	0.225
	NAIN 50 km			680	1.298	2.546	1.879	0.193
	NACH 250 m	Adjusted angular choice measure for cross-scale comparison. Calculated similarly to ACHB(x) but normalized for cross-scale comparison and considering the urban system's optimization of travel distance and cost of segregation.	$NACH_B(x) = \frac{\log(\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma(i,x,j)+1)}{\log(\sum_{i=1} d_\theta(x,i)+3)} \ (i \neq x \neq j)$	680	0.000	1.330	0.793	0.197
	MTL 5000m	Metric Total Length: This variable calculates the total length of all segments within the entire network, using actual distance measurements. This reflects the breadth of the network and the range of mobility it provides.		680	690,162.943	1,227,088.147	1,008,602.413	134,380.721
	TINT R10K	Indicates how integrated or segregated a vertex is from the urban system. Degree of integration or segregation from the urban system, both globally and locally.	$INT_i = \frac{1}{RRA_i} = \frac{D_i}{RA_i}$	680	5737.742998	24,266.305	13,451.649	3471.852
	TINT R50K			680	45,186.47918	89,824.041	65,149.993	7015.099
	MTD R5K	This variable measures the total number of nodes reachable within a spatial network based on a specific distance metric (such as meters or feet). In urban network analysis, nodes usually represent intersections or entrances to buildings.		680	368,013.0000	6,637,732.167	2,693,390.266	953,048.618
	MTN R5K	Metric Total Nodes: This variable measures the total number of nodes reachable within a spatial network based on a specific distance metric (such as meters or feet). In urban network analysis, nodes usually represent intersections or entrances to buildings.		680	12,640.3	28,517.85	21081.68	4269.267
	TTD R10K	Total Nodes: Refers to the total number of nodes that are reachable within a specific threshold. This variable helps analyze accessibility and the concentration of the network in a particular area.	$TD_i = \sum_{j=1}^{n-1} d_{ij}, \ i \neq j$	680	213,896.970	634,530.48	400,644.634	77,591.926
	TTD R5K			680	51,289.496	151,837.111	91,678.257	19,355.75
	TNC R10K	Node Count: Simply records the total number of nodes within the network. This count includes all independent nodes within the analysis boundary, reflecting the scale of the network.		680	40,586.2	96,244.272	71,912.097	13,546.586
	TNC R24K			680	234,780.2	365,161.818	309,205.451	32,381.628
	TTSL R10K	T1024 Total Segment Length: This variable represents the total length of all segments within a certain threshold. It helps understand the density and connectivity of the road network within the given range.		680	2,294,018.1	4,302,021.055	3,525,273.747	483,289.453

3.3. Formulas for Each Algorithm Model

(1)   Quantification of Vitality in Railway Station Areas

To analyze population concentration in a city over a specific time period, Baidu heatmaps can be used to visualize dynamic temporal patterns. For this, all heatmaps within the given period are overlaid, and their average is calculated to derive the population distribution for that timeframe [81]. The specific formula is as follows:

$$V_i = \frac{\sum\limits_{i=1}^{n} X_i A_i}{n \times S_i} \tag{1}$$

where V_i represents the average street vitality, A_i represents the vitality of street i during a specific time period, S_i represents the area of street i, and n = 0:00, ... 0:00 denotes the 24 time periods in a day.

(2)   Linear Regression Model

Ordinary Least Squares (OLS)
The expression for the linear regression model is as follows [82]:

$$Y = a_1 X_1 + a_2 X_2 + a_3 X_3 + b \tag{2}$$

where Y represents the vitality of the street segment, and X_1, X_2, and X_3 represent the variables of spatial form, functional attributes, and road network structure, respectively.

(3)   Geographically Weighted Regression (GWR) Model

① GWR
Geographically Weighted Regression (GWR) is a regression analysis method that considers the impact of geographical locations, capturing spatial heterogeneity and spatial dependency [69]. The expression for GWR is as follows:

$$Y_i = \beta_0(u_i, v_i) + \sum_K \beta_k(u_i, v_i) X_{ik} + \varepsilon_i \tag{3}$$

where Y_i represents the vitality intensity within the study unit (dependent variable), (u_i, v_i) denotes the coordinates of spatial unit i, $\beta_0(u_i, v_i)$ represents the intercept for unit i, $\beta_k (u_i, v_i)$ are the geographically varying coefficients for unit i, and ε_i represents the regression residual.

② Multi-Scale Geographically Weighted Regression (MGWR) Model
Brunsdon introduced spatial location information into regression models, proposing the GWR model based on the spatial attributes of the data [83]. The GWR model accounts for spatial variation in the relationship between dependent and independent variables. However, while GWR addresses spatial heterogeneity, using the same spatial bandwidth for all independent variables can result in unstable regression outcomes. To resolve this issue, the MGWR model was developed [84]. This multi-bandwidth approach constructs a more effective and realistic spatial process model.

$$y_i = \sum_{n=1}^{k} \beta_{bwn}(\mu_i, v_i) x_{in} + \varepsilon_i \tag{4}$$

where y_i is the response variable, x_{in} is the covariate, β_{bwn} is the local regression coefficient for MGWR bandwidth b_w or the n-th variable, (u_i, v_i) is the spatial location of the sample point, and ε_i is the regression residual.

(4) Machine Learning Models

① Random Forest (RF)

Random Forest constructs multiple decision trees (tree models) for prediction, then averages or sums the predictions to improve the overall model's accuracy [85]. The expression for Random Forest is as follows:

Random Forest constructs multiple decision trees (tree models) for prediction, then averages or sums the predictions to improve the overall model's accuracy [85]. The expression for Random Forest is as follows [M3]:

$$\hat{f}(x_i) = \frac{1}{N} \sum_{n=1}^{N} f_b(x_i) y_i = \hat{f}(x_i) + \varepsilon_i \tag{5}$$

where $f(x_i)$ denotes the contribution of variable X to the regression model, representing its specific effect within the overall model structure, x_i is the feature vector, N is the total number of independent trees in the forest, and $f_b(x_i)$ is the prediction of individual tree b.

② Extreme Gradient Boosting (XGBoost)

XGBoost is an open-source machine learning framework developed by Tianqi Chen et al. based on the Gradient Boosting Decision Trees (GBDT) algorithm, and it is robust to data multicollinearity [86]. The expression for XGBoost is as follows:

$$\text{Obj} = \sum_{i=1}^{n} l(\hat{y}_i, y_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{6}$$

where Obj is the target function optimized, (y_i, y_i) is the loss function, y_i is the true label of sample x_i, y_i is the predicted value, $\Omega(f_k)$ is the regularization term representing model complexity, f_k is the function expression of the k-th tree model, n is the total number of samples, and K is the total number of base models.

③ LightGBM (Light Gradient Boosting Machine)

LightGBM is a framework that implements the GBDT algorithm using a histogram-based algorithm, supporting efficient parallel training [87]. The expression for LightGBM is as follows:

$$\text{Obj}^K = \sum_{i} L(y_i, \hat{y}_i^K) + \Omega(f_K) + c^{K-1} = \sum_{i} L(y_i, \hat{y}_i^{K-1} + f_K(x_i)) + \Omega(f_K) + c^{K-1} \tag{7}$$

where the Obj^k function, like the Obj in XGBoost, is designed to be optimized during the model training process. y_i is the observed value, y_i is the predicted value after K iterations, $\Omega(f_k)$ is the regularization term for the k-th tree, c^{K-1} is the sum of regularization terms for the first $K-1$ trees, and f is the function for the K-th tree.

④ SHAP (Shapley Additive exPlanations) Framework for Interpretable Machine Learning

Models such as Random Forest, XGBoost, and LightGBM are inherently difficult to interpret [88]. Additionally, SHAP can accurately quantify the impact of these variables on vitality prediction.

$$\hat{y}_i = f_0 + \sum_{i=1}^{M} f_i \tag{8}$$

where y_i is the model prediction, f_i is the marginal contribution value for each feature, f_0 is the mean prediction of all training samples, and M is the total number of features.

⑤ Root Mean Squared Error (RMSE)

RMSE is a common error metric used to evaluate and select models in various types of model comparisons [89]. The smaller the RMSE, the better the model fit. The calculation method for RMSE is as follows:

$$RMSE = \sqrt{\frac{\sum\limits_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{9}$$

where n is the total number of predictions, y_i is the true value of the target variable i, and y_i is the predicted value of the target variable i.

⑥ Coefficient of determination R^2

The coefficient of determination (R^2) is a statistical metric that quantifies the proportion of total variance in the dependent variable that can be explained by one or more independent variables within a regression model. It serves as a critical indicator for assessing the goodness of fit of the model, particularly in the context of linear regression [90].

$$R^2 = 1 - \frac{\sum\limits_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum\limits_{i=1}^{n}(y_i - \overline{y})^2} \tag{10}$$

where y_i represents the observed values, $\hat{y}_i$ represents the predicted values, and y is the mean of the observed values.

⑦ Adjusted R^2

The adjusted R^2 is an enhancement of the R^2 statistic, which not only measures the goodness of fit but also adjusts for the number of predictors in relation to the sample size. By doing so, it penalizes the model for unnecessary complexity, reducing the likelihood of overfitting [90]. The formula for this metric is as follows:

$$Adjusted\ R^2 = 1 - \frac{n-1}{n-p}\left(1 - R^2\right) \tag{11}$$

where n is the sample size, and p is the number of predictors.

⑧ Standard Deviation

The performance of the regression models was evaluated using the standard deviation, which measures the dispersion of the data points from the mean.

$$SD = \sqrt{\frac{1}{n-1}\sum\limits_{i=1}^{n}(x_i - \overline{x})^2} \tag{12}$$

where x_i represents each data point, and x is the mean of the data points.

⑨ K-Fold Cross-Validation

K-Fold Cross-Validation is important in machine learning, as it helps reduce overfitting risks, optimize parameter selection, fully utilize data, and evaluate model stability. It is a common evaluation method in machine learning. The calculation method for K-Fold Cross-Validation is as follows:

$$CV_K = \frac{1}{K}\sum\limits_{i=1}^{K}L_i \tag{13}$$

where CV_K represents the cross-validation score using K folds, and L_i represents the performance metric calculated on fold i, where the model is trained on all folds except i.

4. Results

4.1. Spatial Form Morphological Distribution Characteristics

The land use efficiency in the areas surrounding the stations is reflected through two indicators: building density and floor area ratio (Figure 7). High floor area ratio areas are

particularly prominent near Beijing Station and Beijing West Station. Specifically, these high ratios around Beijing Station are primarily found on the north and south sides, which is influenced by land use strategies. Given the limited land resources in the city center, especially near the transportation hub of Beijing Station, various buildings have been constructed to meet residential and commercial needs. Additionally, the "Station-City Integration" development strategy proposed by the Beijing Municipal Government in 2018 has encouraged high-density land use around this station. At Beijing West Station, high floor area ratio areas are mainly located to the north, east, and south, forming a semi-enclosed pattern. To the northwest, numerous schools, hospitals, and other functional areas have been established, where building height and density are restricted to maintain open spaces. In contrast, the low floor area ratio areas around Beijing Fengtai Station are the most prominent. After its reconstruction and reopening in June 2022, the surrounding area aligns more with modern livable city planning, focusing on a balance between green space and population density. Urban planning also tends to control building density to accommodate future city development needs.
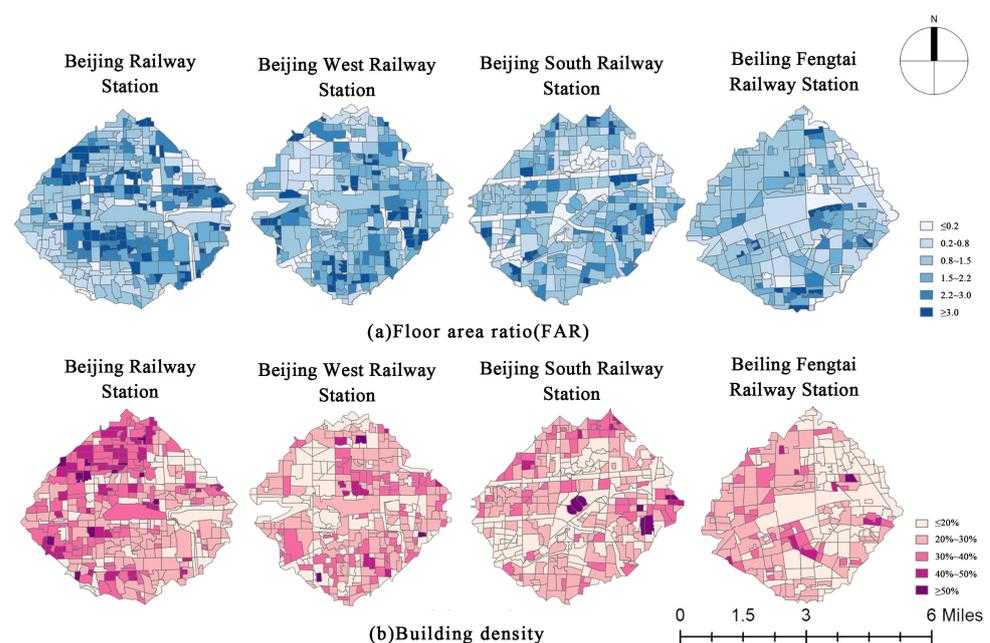
Figure 7. Development intensity.

High building density areas are notably significant around Beijing Station, especially to the northwest, where there is a marked difference in building density. As a landmark of the capital, Beijing Station, located in the city center, boasts convenient transportation, numerous business offices, and historical and cultural attractions, drawing significant foot traffic. In contrast, the buildings around the other three stations are more evenly distributed, providing residents with a good living environment and opportunities for social interaction.

4.2. Functional Distribution Characteristics

Figure 8 illustrates the distribution patterns of six types of functions within the two-level station areas across four cases. The figure reveals significant progress in mitigating railway-induced urban fragmentation at the Beijing West, Beijing South, and Beijing Fengtai Stations. Conversely, the accessible area around Beijing Station is skewed, highlighting the railway's impact on urban space division. Among the six major functions, commercial and public service functions have the highest densities, while tourism and leisure functions have the lowest. Furthermore, there is a notable increase in all functions from the core area to the outer expansion area. The total POI density at Beijing Station, Beijing West Station,

and Beijing South Station surpasses that of Beijing Fengtai Station, indicating a trend where longer renovation periods correlate with more developed functions.
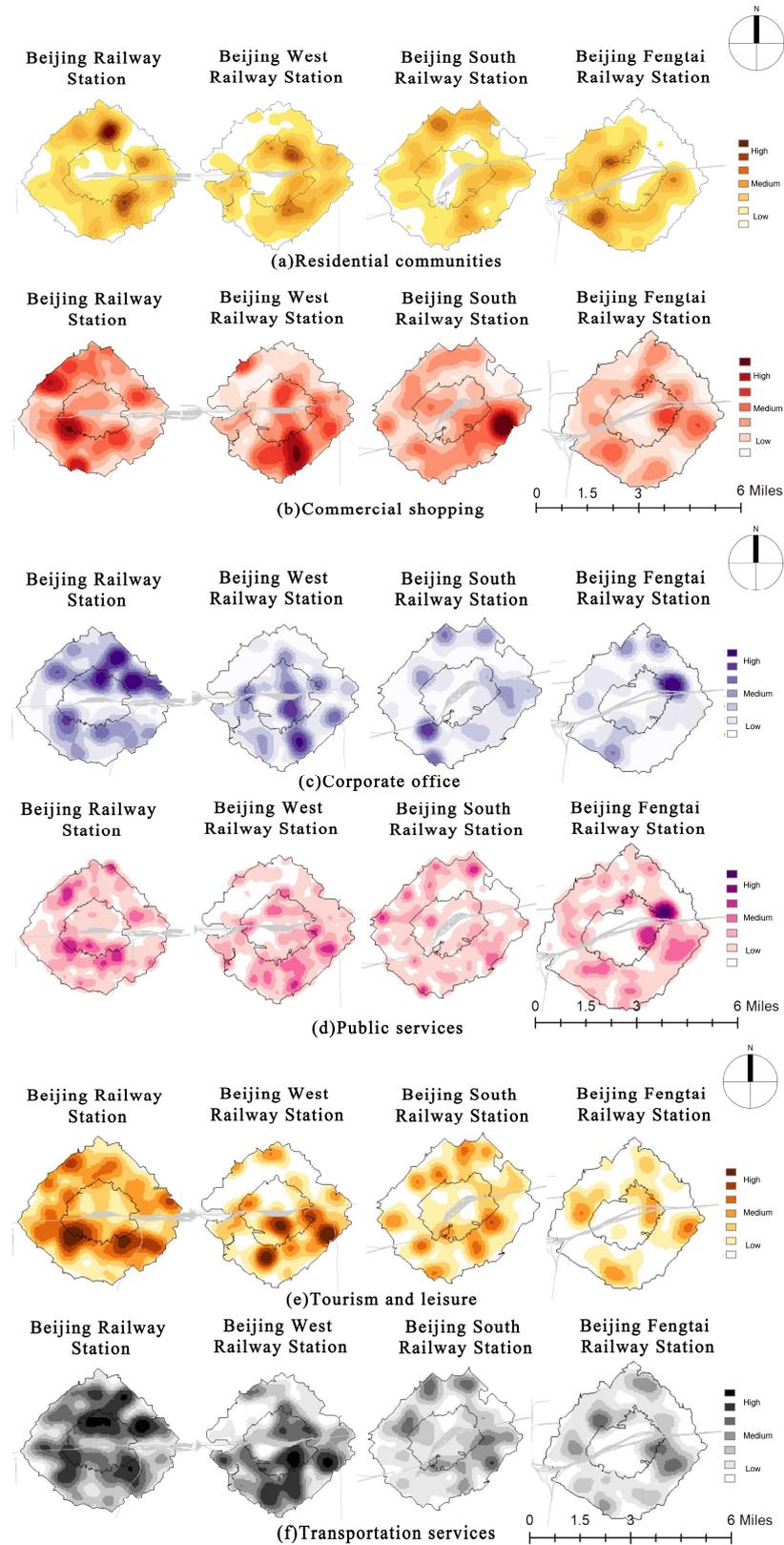
Figure 8. Kernel density values of each function.

Analyzing the spatial distribution and connectivity of the six functions reveals several insights:

1. Residential Functions: Residential functions are more prevalent around Beijing Station, forming multiple clusters, predominantly in the outer expansion layer.

2. Commercial Shopping Functions: The core and outer expansion areas of all four stations are well-covered, indicating good overall connectivity without fragmentation due to the train stations. However, the commercial coverage at the Beijing South and Beijing Fengtai Stations is less comprehensive. Beijing Station's commercial hotspots are mainly on the southern side of the outer expansion area, whereas Beijing South Station's are in the southeastern outer expansion area, extending outward. Due to Lianhuachi Park and Yuyuantan Park, the commercial activity at Beijing West Station is less concentrated than at Beijing Station. The Beijing South and Beijing Fengtai Stations differ in commercial hotspot distribution, with the former concentrated at the boundary of the core and outer expansion areas, while the latter lacks significant hotspots. This can be explained by the proximity to the city center and land rent theory for the first two stations.

3. Public Service Functions: Public service distribution correlates strongly with commercial functions, showing similar trends. Beijing Station and Beijing West Station exhibit more pronounced hotspots within station buildings due to centralized public facilities. In contrast, Beijing South and Beijing Fengtai Stations' service functions are less developed.

4. Transportation Functions: Beijing Station, Beijing West Station, and Beijing Fengtai Station have extensive, balanced hotspot coverage, while Beijing South Station shows weaker connectivity. The connection between transportation functions and Beijing South Station is mainly in the east-west directions, with the outer expansion area featuring low-density scattered points, indicating that transportation development still relies on the station itself, and the surrounding areas have not yet formed a cohesive system.

5. Office Functions: Beijing Station has a more balanced distribution and better connectivity, while the other stations display lower-density scattered distributions. Office hotspots at Beijing Station are clustered on the north and south sides, making it the largest area among the four cases, highlighting its significant business role, closely linked to its proximity to the city center.

6. Tourism and Leisure Functions: At the Beijing South and Beijing Fengtai Stations, these functions are distributed on both sides of the station road network, showing a multidirectional scattered pattern. Beijing West Station's functions are concentrated in the southeast outer expansion area. Tourism and leisure functions at Beijing Fengtai Station are relatively underdeveloped and lack a clear system. The clustering trend from Beijing Station to Beijing Fengtai Station varies over time, reflecting the ongoing development of tourism and leisure functions at Beijing Fengtai Station.

4.3. Urban Vitality Changes Around Beijing's Four Major Special-Class Stations

Figures 9 and 10 demonstrate the spatial and temporal variations in vitality around the special-grade stations on both weekdays and weekends. Overall, the four stations show distinct patterns in vitality distribution. Although the vitality distribution on weekdays and weekends is generally similar, certain areas exhibit noticeable differences. Both on weekdays and weekends, Beijing Railway Station and Beijing West Railway Station have more high-vitality areas compared to Beijing South Railway Station and Beijing Fengtai Railway Station, underscoring the greater activity levels and significance of the former two stations. Additionally, most of the high-vitality grids are located closer to the city center.
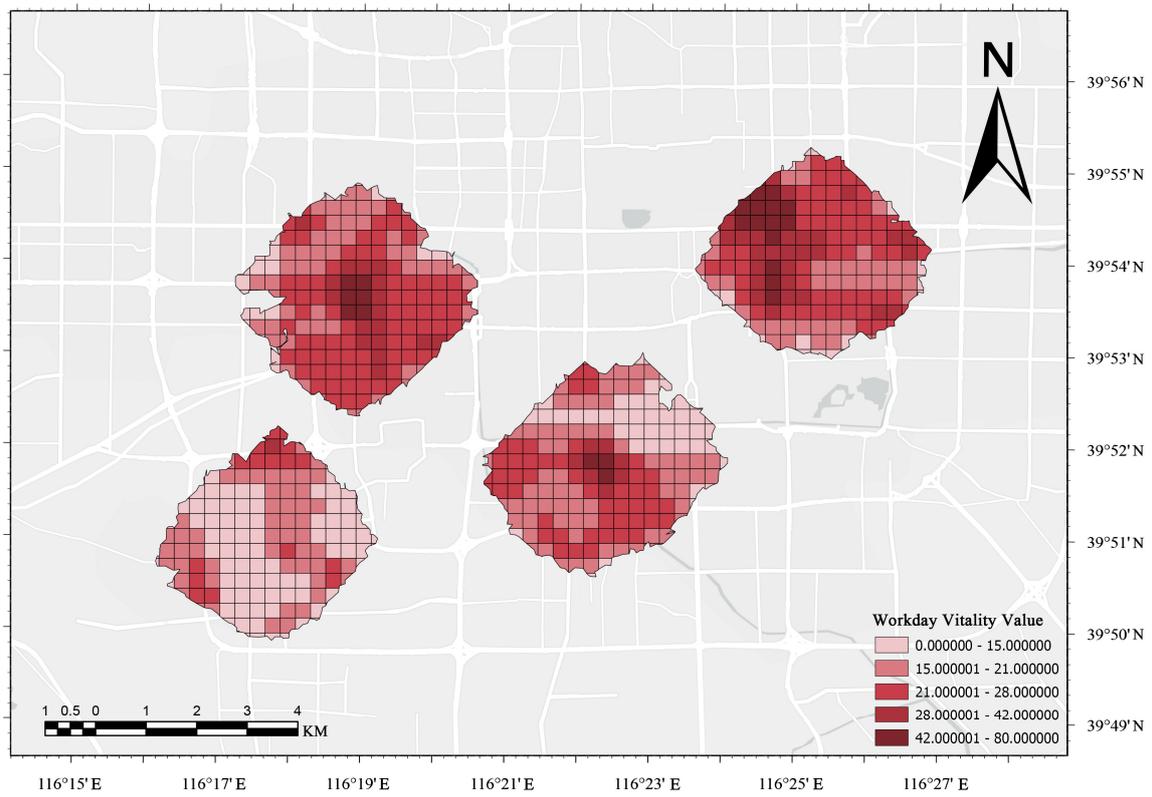
Figure 9. Spatiotemporal variation of urban vitality on workdays.
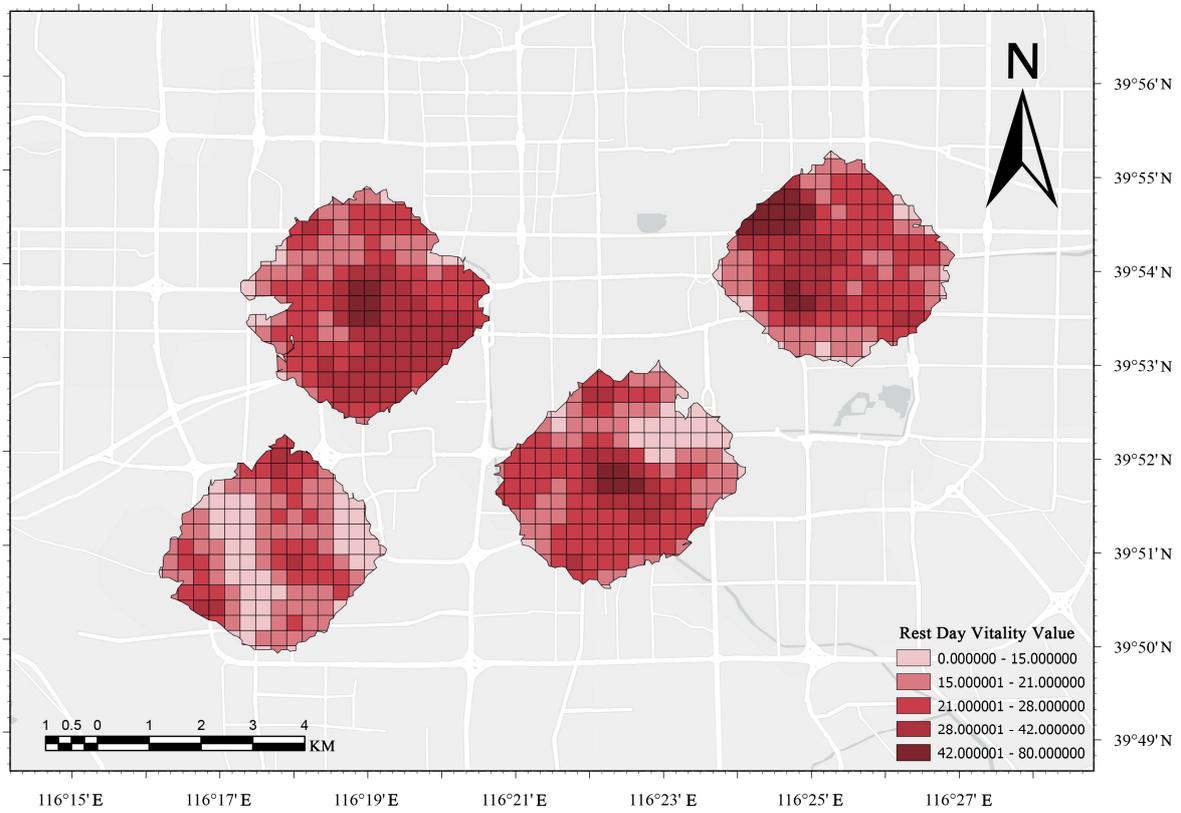
Figure 10. Spatiotemporal variation of urban vitality on weekends.

From a temporal perspective, high-vitality areas are more widespread on weekends than on weekdays, suggesting that residents engage in more frequent and broader activities on weekends. Furthermore, the concentration of vitality increases on weekends, with more regions displaying mid-to-high vitality values (denoted by darker shades), highlighting the enhancing effect of weekends on urban vitality.

Regionally, high-vitality areas around Beijing Railway Station are more concentrated during both weekdays and weekends, particularly around commercial and healthcare facilities, indicating the consistent attractiveness of these amenities throughout the week. The high-vitality areas around Beijing West Railway Station are centered in the station's core on both weekdays and weekends. However, on weekends, these areas extend southeast towards nearby communities and shopping malls, particularly around Maliandao Road, reflecting the rise in shopping and leisure activities. Similarly, Beijing South Railway Station sees an expansion of mid-to-high vitality areas on weekends, especially near Jiahe Park to the southwest, indicating increased outdoor leisure activities. At Fengtai Railway Station, mid-to-high vitality areas are more dispersed on weekdays but become more concentrated on weekends, suggesting a higher concentration of resident activities during the weekend.

4.4. Method Comparison
4.4.1. Comparison of Models Under Local Variables

After standardizing the built environment indicators, Multiple Linear Regression analysis was conducted using SPSS, with the results displayed in Table 4. Significant independent variables were identified and used to construct GWR and MGWR models for the vitality of areas around railway stations across different time periods.

Table 4. Comparison of performance of each model.

Algorithm		Linear Model	Geographically Weighted Model		Machine Learning Model		
		OLS	GWR	MGWR	Global Random Forest	XGBooost	LightGBM
Adjusted R^2	Vitality	0.515	0.807	**0.848**	/	/	/
	Vitality on weekdays	0.553	0.857	**0.887**	/	/	/
	Vitality on weekends	0.436	0.835	**0.883**	/	/	/
AICc	Vitality	1361.233	1090.66	**908.739**	/	/	/
	Vitality on weekdays	1395.504	889.478	**682.979**	/	/	/
	Vitality on weekends	1588.196	993.324	**756.137**	/	/	/
Bandwidth	Vitality	/	55	(43, 615)	/	/	/
	Vitality on weekdays	/	58	(43, 206)	/	/	/
	Vitality on weekends	/	57	(43, 110)	/	/	/
Out-of-Sample R^2	Vitality	0.446	0.557	/	**0.769**	0.674	0.670
	Vitality on weekdays	0.625	0.691	/	**0.763**	0.744	0.679
	Vitality on weekends	0.462	0.568	/	0.651	**0.657**	0.565
Out-of-Sample RMSE	Vitality	0.146	0.147	/	**0.115**	0.120	0.231
	Vitality on weekdays	0.145	0.149	/	**0.124**	0.124	0.200
	Vitality on weekends	0.145	0.147	/	0.143	**0.129**	0.323

The results indicate that the adjusted R^2 of the MGWR model improved significantly, while the AICc value decreased, suggesting that the MGWR model provides a better fit than both the GWR model and OLS, thus enhancing model stability. Specifically, the MGWR model accounts for 84.8%, 88.7%, and 88.3% of the variation in vitality values for average periods, weekdays, and weekends, respectively, around railway passenger stations, indicating a strong model fit. It is worth noting that, in the comparison across the three models, the R^2 for weekdays consistently exceeded that of the overall vitality, which in turn was higher than that of weekends. Therefore, this study adopts the MGWR model to analyze the factors affecting vitality under localized effects.

4.4.2. Comparison of Models Under Global Variables

Utilizing grid search, the optimal parameter combination was determined to optimize the model. Table 4 presents the out-of-sample results for OLS, GWR, XGBoost, Random Forest, and LightGBM, obtained through ten-fold cross-validation. Considering the out-of-sample RMSE and R^2 metrics, the global Random Forest model outperforms other models, whereas LightGBM shows relatively weaker predictive performance. Specifically, Random Forest excels in the R^2 values for average vitality and weekday vitality, surpassing XGBoost by 0.095 and 0.019 respectively. However, XGBoost slightly outperforms Random Forest in the R^2 for weekend vitality, with a marginal difference of 0.06. Compared to linear models and GWR, the global Random Forest also achieves higher a R^2 and a lower RMSE. Thus, the Random Forest (RF) model is selected for investigating the nonlinear relationships between spatial form, street function, and vitality in station areas under global effects, integrating it with SHAP analysis. For MGWR, the theories and methods regarding out-of-sample predictions are not yet mature [91], and no public implementation methods are available. Thus, we do not support out-of-sample prediction for MGWR.

Table 5 displays the optimal parameter combinations for various machine learning models. In the training process, "Val" represents the validation set used for tuning or model selection, "Train" refers to the training set for learning data features, and "Test" denotes the reserved test set used to assess model performance, which is not involved in training or tuning. The dataset is divided into training, validation, and test sets. The training set is used for model training, the validation set for tuning (optimal hyperparameters), and the test set for performance evaluation. Among the parameter combinations, "bootstrap" set to True, "max_depth" set to 20, "min_samples_leaf" to 1, "min_samples_split" to 2, and "n_estimators" to 300 performed best on the test set, achieving an R^2 score of 0.769.

Table 5. Model parameters.

Model		Params	train_R^2_Score	val_R^2_Score	test_R^2_Score
Global Random Forest	Vitality	{'bootstrap': True, 'max_depth': 20, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 300}	0.906000373	0.768849224	0.665804941
	Vitality on weekdays	{'bootstrap': True, 'max_depth': None, 'min_samples_leaf': 2, 'min_samples_split': 2, 'n_estimators': 200}	0.900778505	0.763400339	0.900992067
	Vitality on weekends	{'bootstrap': True, 'max_depth': 20, 'min_samples_leaf': 1, 'min_samples_split': 5, 'n_estimators': 300}	0.836411732	0.650810857	0.531086753
XGBooost	Vitality	{'colsample_bytree': 0.8, 'learning_rate': 0.05, 'max_depth': 5, 'min_child_weight': 10, 'n_estimators': 500, 'reg_alpha': 0.1, 'reg_lambda': 1, 'subsample': 0.8}	0.989456264	0.674649991	0.525335069
	Vitality on weekdays	{'colsample_bytree': 0.8, 'learning_rate': 0.1, 'max_depth': 3, 'min_child_weight': 5, 'n_estimators': 1000, 'reg_alpha': 0.1, 'reg_lambda': 1, 'subsample': 0.8}	0.999511538	0.744485954	0.88515299
	Vitality on weekends	{'colsample_bytree': 0.8, 'learning_rate': 0.05, 'max_depth': 3, 'min_child_weight': 10, 'n_estimators': 500, 'reg_alpha': 0.1, 'reg_lambda': 1, 'subsample': 0.8}	0.90138167	0.657759935	0.69574312
LightGBM	Vitality	{'bagging_fraction': 0.8, 'bagging_freq': 3, 'feature_fraction': 1.0, 'learning_rate': 0.01, 'n_estimators': 300, 'num_leaves': 50}	0.720894442	0.67069974	0.40130828
	Vitality on weekdays	{'bagging_fraction': 1.0, 'bagging_freq': 3, 'feature_fraction': 0.8, 'learning_rate': 0.01, 'n_estimators': 200, 'num_leaves': 30}	0.804136545	0.679186887	0.788486183
	Vitality on weekends	{'bagging_fraction': 1.0, 'bagging_freq': 3, 'feature_fraction': 0.8, 'learning_rate': 0.01, 'n_estimators': 300, 'num_leaves': 30}	0.753907118	0.565776059	0.548275682

4.5. Relationship Between Spatial Form, Street Functions, and Vitality in Station Areas

4.5.1. Ordinary Least Squares Analysis

The regression results, as shown in Table 6, indicate an R^2 value of 0.515, suggesting that the selected variables account for 51.5% of the variance in street vitality. The analysis reveals that commercial density, average number of floors, housing prices,

NAIN_R5000m, TINT_R10K, MTL_R5K, and residential density significantly influence street vitality. Among the variables positively correlated with vitality, the strength of the associations is as follows: TINT_R10K > commercial density > average number of floors. Conversely, the variables negatively correlated with vitality rank in strength as follows: NAIN_R5000m > housing prices > MTL_R5K > residential density.

Table 6. Linear regression results of variable.

Variable	Unstandardization Coefficient Beta		Standardization Coefficient Beta	t	p	Descriptive	
	B	Std. Error				Tolerance	VIF
(Intercept)	1.286	0.071		18.164	0.000 ***		
Commercial density	0.157	0.015	0.400	10.392	0.000 ***	0.795	1.258
Average number of floors	0.072	0.025	0.107	2.851	0.005 **	0.842	1.188
Housing prices	-1.000×10^{-6}	0.000	−0.182	−3.946	0.000 ***	0.553	1.808
NAIN_R5000m	−0.437	0.043	−0.647	−10.112	0.000 ***	0.288	3.470
TINT_R10K	4.700×10^{-5}	0.000	1.116	11.764	0.000 ***	0.131	7.629
MTL_R5K	0.000	0.000	−0.246	−3.702	0.000 ***	0.268	3.737
Residential density	−0.001	0.001	−0.096	−2.653	0.008 **	0.900	1.111

Model summary		
	R	0.723
	R square	0.523
	Adjust R square	0.515
	Std. Error of the Estimate	0.0977840720
	Durbin–Watson	1.998

Dependent Variable: Vitality * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$.

To more comprehensively consider the impact of the built environment on urban vitality across different times and to provide more specific, targeted recommendations for urban planning, Tables 7 and 8 present the results of Multiple Linear Regression analyses for the predictor variables of weekday vitality and weekend vitality. The results indicate that the weekday vitality model has an R^2 of 0.553, while the weekend vitality model has an R^2 of 0.436. This suggests that the selected variables can explain 55.3% and 43.6% of the variation in street vitality during weekdays and weekends, respectively, with the weekday model performing better than both the overall vitality model and the weekend model. Although an R^2 of 0.436 might appear low in other scientific fields, it is still considered a relatively good performance in this domain. Even if the R^2 values are relatively low, such values are acceptable as long as the predictors or explanatory variables in the model have statistical significance [92].

The above results indicate that weekday vitality is more predictable compared to weekends. This suggests that human activity patterns are more regular on weekdays [93]. The variables positively correlated with weekday vitality, in order of influence, are as follows: TINT_R10K, commercial density, and average number of floors. The variables negatively correlated with weekday vitality, in order of influence, are as follows: NAIN_R5000m, housing prices, MTL_R5K, and residential density. These results, except for residential density, are consistent with those in Table 5.

Table 7. Linear regression results of variables on weekdays.

Variable	Unstandardization Coefficient Beta		Standardization Coefficient Beta	t	p	Descriptive	
	B	Std. Error				Tolerance	VIF
(Intercept)	1.258	0.075		16.818	0.000 ***		
Commercial Density	0.163	0.016	0.392	10.142	0.000 ***	0.802	1.247
Average Number of Floors	0.067	0.026	0.098	2.622	0.009 **	0.853	1.173
Housing prices	-9.812×10^{-7}	0.000	−0.201	−4.341	0.000 ***	0.558	1.791
NAIN_R5000m	−0.474	0.045	−0.699	−10.546	0.000 ***	0.272	3.682
TINT_R10K	5.176×10^{-5}	0.000	1.195	12.387	0.000 ***	0.128	7.795
MTL_R5K	-2.387×10^{-7}	0.000	−0.222	−3.364	0.001 ***	0.274	3.650
Residential Density	−0.002	0.001	−0.103	−2.816	0.005 **	0.896	1.116
Model summary	R				0.749		
	R square				0.561		
	Adjust R square				0.553		
	Std. Error of the Estimate				0.0960910489		
	Durbin–Watson				1.926		

Dependent variable: Vitality on weekdays * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$.

Table 8. Linear regression results of variables on weekends.

Variable	Unstandardization Coefficient Beta		Standardization Coefficient Beta	t	p	Descriptive	
	B	Std. Error				Tolerance	VIF
(Intercept)	0.941	0.060		15.777	0.000 ***		
Commercial Density	0.168	0.017	0.423	9.687	0.000 ***	0.789	1.267
Average Number of Floors	0.103	0.028	0.157	3.747	0.000 ***	0.857	1.167
Housing prices	-7.019×10^{-7}	0.000	−0.151	−2.858	0.005 **	0.543	1.841
NAIN_R5000m	−0.088	0.032	−0.137	−2.735	0.007 **	0.605	1.654
TINT_R10K	7.416×10^{-6}	0.000	0.717	7.879	0.000 ***	0.182	5.496
MTL_R5K	-2.387×10^{-7}	0.000	−0.222	−3.364	0.001 ***	0.274	3.650
Residential Density	−0.001	0.001	−0.099	−2.417	0.016 **	0.906	1.104
Model summary	R				0.668		
	R square				0.447		
	Adjust R square				0.436		
	Std. Error of the Estimate				0.1029961884		
	Durbin–Watson				1.913		

Dependent variable: Vitality on weekends * $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$.

Typically, an appropriate residential density can enhance the vitality of station areas, as a higher number of residents brings about richer public facilities and commercial services. This, in turn, promotes increased pedestrian activity and business opportunities, driving the development and prosperity of station areas. However, if the residential density is excessively high, it can lead to insufficient public facilities, traffic congestion, and environmental degradation, which negatively impact the vitality of station areas.

The variables positively correlated with weekend vitality, in order of influence, are as follows: commercial density, TINT_R10K, and average number of floors. The variables negatively correlated with weekend vitality, in order of influence, are as follows: MTL_R5K, residential prices, NAIN_R5000m, and residential density. On weekends, commercial density has a greater impact on vitality than the road network structure. Weekends are typically reserved for rest and leisure, with many individuals choosing to engage in activities such as shopping, dining, and entertainment. Commercial facilities not only attract more visitors to the station area, thereby increasing vitality, but also function as key venues for social interaction.

4.5.2. Multi-Scale Geographically Weighted Regression (MGWR) Correlation Analysis

To determine the spatial distribution of significant variables influencing vitality, the factors affecting vitality were estimated using MGWR. However, this section does not include a time-period-specific analysis. According to the regression results presented in Table 9, the magnitude of influence is ranked as follows: TINT_R10K > NAIN_R5000m > MTL_R5K > residential density > commercial density > average number of floors > housing prices, indicating that spatial configuration exerts the most substantial impact on the vitality of areas surrounding the stations. TINT_R10K and housing prices show a significant positive correlation with vitality, whereas NAIN_R5000m exhibits a significant negative correlation. Commercial density, residential density, average number of floors, MTL_R5K, and NAIN_R5000m demonstrate contrasting effects on street vitality.

Table 9. Statistical description of MGWR coefficients.

Variable	Band-Width	Adj t-val (95%)	p	T	Mean	STD	Min	Median	Max
Intercept	43.000	3.070	0.008 **	−1.127	−0.361	0.726	−1.788	−0.404	1.229
Commercial density	43.000	3.143	0.005 **	4.338	0.343	0.243	−0.116	0.305	1.250
Average number of floors	50.000	3.140	0.011 **	3.016	0.059	0.138	−0.356	0.046	0.594
Residential density	43.000	3.142	0.009 **	−2.436	−0.105	0.171	−0.528	−0.105	0.378
Housing prices	615.000	2.192	0.019 **	2.206	0.030	0.011	0.009	0.029	0.051
MTL_R5K	43.000	3.060	0.002 **	−4.503	−0.466	0.622	−1.453	−0.325	0.799
NAIN_R5000m	183.000	2.458	0.000 ***	−5.009	−1.076	0.472	−1.864	−1.077	−0.553
TINT_R10K	43.000	3.134	0.000 ***	4.745	1.832	0.577	0.966	1.649	2.898

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.001$.

However, due to the variations in the bandwidth scales of each indicator in the MGWR model, it is necessary to apply the Adj t-val (95%) values from Table 8 (representing the confidence level of the bandwidth) as the criterion for evaluating the data, excluding insignificant values. The results, as shown in Figure 11d,e, indicate that only NAIN_R5000m and TINT_R10K are related to most areas around the four stations, demonstrating the significant impact of spatial configuration on vitality.
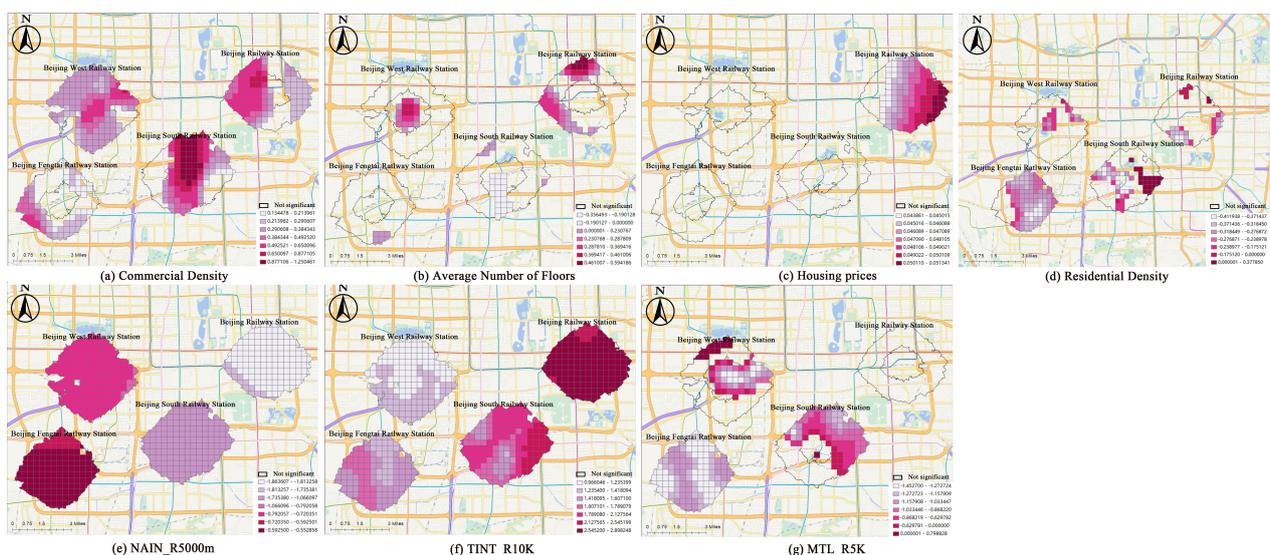
Figure 11. Spatial distribution of MGWR regression coefficients.

4.5.3. Correlation Analysis of Regression Prediction with Random Forest Model

This study utilizes the Random Forest regression model, incorporating the SHAP algorithm and partial dependence plots, to thoroughly investigate the nonlinear relationships between vitality and its influencing factors around railway stations. The results are summarized as follows:

Figure 12 shows the calculated SHAP values on the horizontal axis, with higher positive SHAP values indicating a greater positive impact on vitality. Each point in the figure represents a data point extracted from the database, stacked vertically to illustrate density and color-coded according to value, with red indicating higher values and blue-purple indicating lower values. Commercial density, TTD R10K, and TTSL R10K make the highest contributions. Enterprise density, average number of floors, and transportation facility density generally enhance vitality, while residential density tends to diminish it.
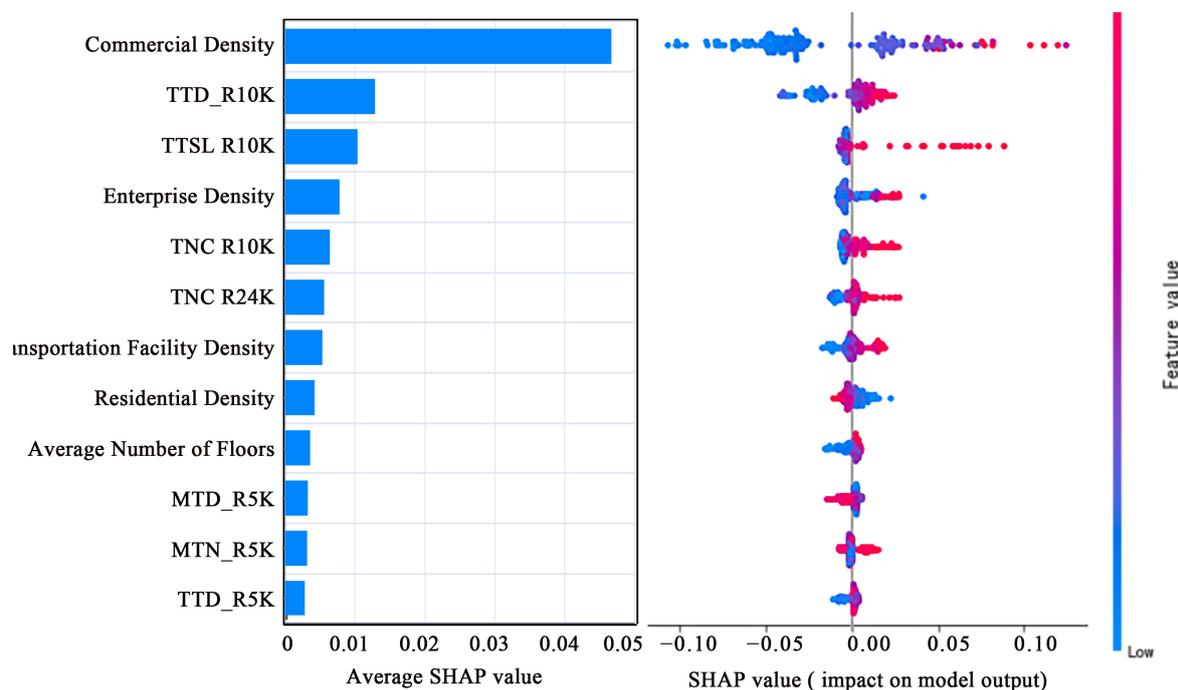
Figure 12. SHAP diagram of respective variables.

Figure 13 elucidates the complex interplay between fluctuations in factor values and their corresponding shifts in importance, based on the SHAP algorithm. Notably, factors such as the average number of floors, commercial density, transportation facility density, TNC R10K, TTD R10K, TTD R5K, TNC R24K, and TTSL R10K display a pattern where, as their values increase, their negative impact on the vitality of areas surrounding railway stations gradually diminishes, transitioning to a positive influence.

Conversely, factors like residential density and MTD R5K exhibit an inverse relationship, where their negative impact on vitality increases as their values decrease. Enterprise density shows a nuanced pattern: initially, as its value increases, the negative impact on the vitality of areas around railway stations also increases; however, upon further increases, this negative impact diminishes and eventually transforms into a positive influence on vitality. Overall, the relationship between enterprise density and MTN R5K with the vitality around railway stations demonstrates a positive U-shaped interaction.
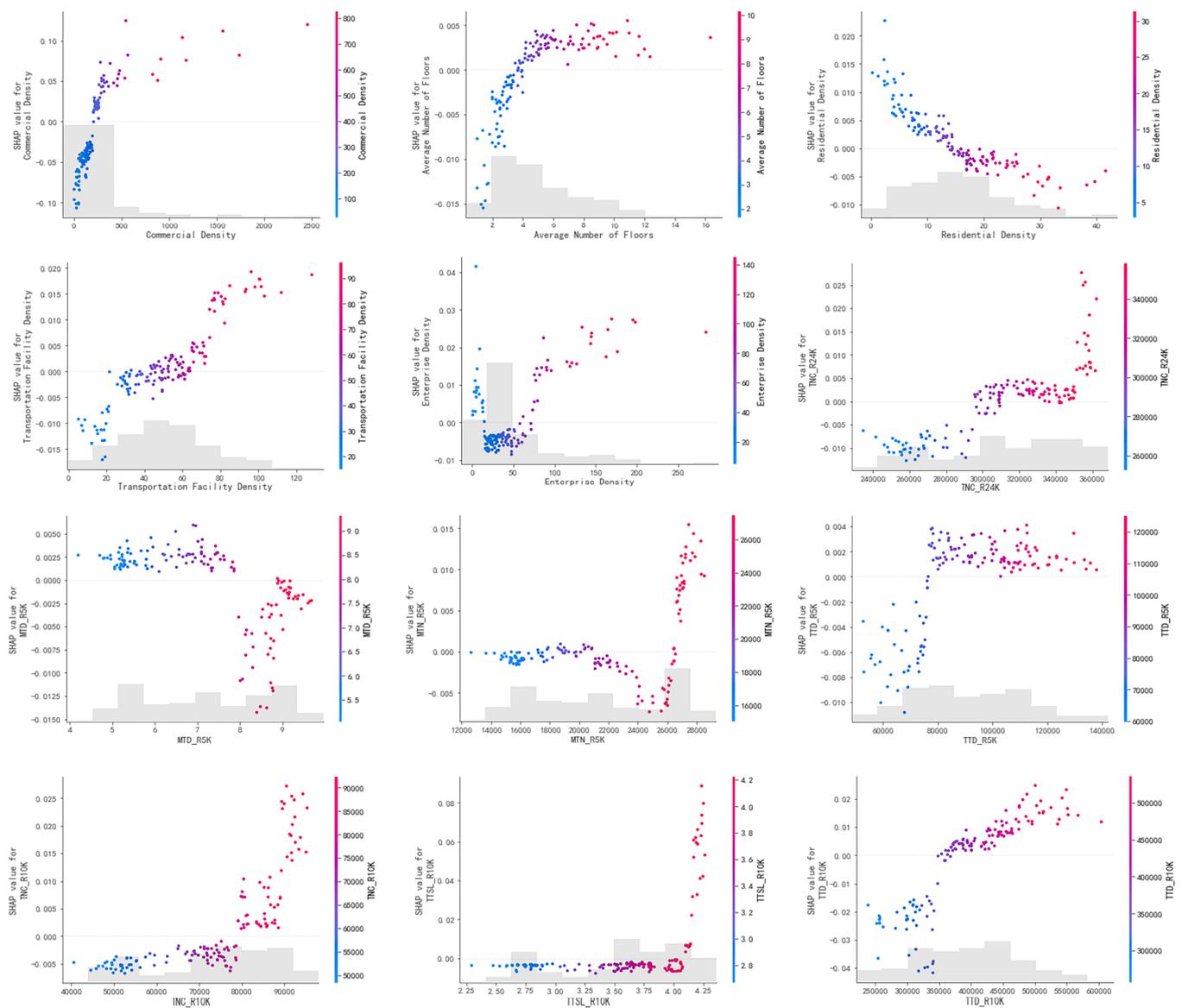
Figure 13. SHAP diagram of various factors.

4.6. Comparison of SHAP Values and MGWR Coefficients

The SHAP values derived from the global random forest model (Figures 12 and 13) were integrated and compared with the MGWR coefficients (Table 8). The findings reveal that, in both models, commercial density in the areas surrounding the stations emerges as the most influential factor, underscoring a strong relationship between the density of commercial facilities and the vitality of railway station areas. Moreover, the average number of floors and residential density serve as key predictors across both models.

5. Discussion

5.1. Factors Influencing the Areas Surrounding Railway Stations

5.1.1. Spatial Form

Based on the OLS models for urban vitality (Tables 5–7) and the random forest analysis (Figures 12 and 13), the average number of floors exerts a significant influence on street vitality [94,95] and demonstrates consistent results across both the OLS and random forest models. This consistency highlights the crucial role of the average number of floors in predicting urban vitality. In many rapidly urbanizing regions, the number and scale of high-rise buildings have historically been regarded as key indicators of socioeconomic

development [96]. Moreover, the MGWR model reveals a polarized impact of the average number of floors on the vitality around special-grade stations. In certain areas, such as those surrounding Beijing Station and Beijing Fengtai Station, this variable does not show a significant relationship with vitality. These stations, being large, purpose-built public facilities, exhibit a level of vitality that is not particularly influenced by building height. However, in the outer ring of Beijing Station, a positive correlation between the average number of floors and vitality is observed. This could be attributed to the stringent height restrictions in areas near the Forbidden City, limiting new development.

5.1.2. Urban Function

According to the multiple regression models for urban vitality (Tables 5–7) and the random forest model (Figures 12 and 13), commercial density significantly influences the vitality of areas surrounding railway stations, showing substantial effects in both the OLS and random forest models. While some studies suggest that commercial density negatively impacts vitality on weekdays in Beijing [31], others argue that it exerts a notable positive influence on street vitality [97]. In the MGWR model used in this study, commercial density is positively associated with the vitality of areas near railway stations, exhibiting a polycentric structure. The spatial distribution of its coefficients follows a "core-periphery" pattern. For instance, around Beijing South Station, the impact of commercial density on vitality decreases progressively from the station outward. As a key transportation hub, railway stations attract not only a significant number of commuters but also diverse transient populations, such as tourists and business travelers. These populations demand a wider range of commercial services, and areas with higher commercial density are better positioned to meet their needs for shopping, dining, and entertainment. Consequently, higher commercial density contributes to greater vitality in station-adjacent areas. This finding is consistent with research on the vitality of subway station areas, which shows that economically advanced cities tend to exhibit higher vitality around subway stations [98].

In the Random Forest model, increases in enterprise density and transportation facility density also exert a positive influence on the vitality of areas surrounding railway stations. This finding aligns with results from earlier research, where scholars have noted that enterprise density and a diverse range of transportation facilities can enhance urban vitality [99,100]. However, both the random forest and MGWR models reveal that residential density has a significant impact on neighborhood vitality, displaying a bifurcated trend [29].

Housing prices are positively correlated with vitality across the study areas. Previous research has also shown that residential clustering is closely linked to housing prices [101]. While the vitality of areas surrounding Beijing West Station, Beijing South Station, and Beijing Fengtai Station is not influenced by housing prices, the area around Beijing Station exhibits a ring-like spatial structure, with the impact gradually increasing from west to east. This suggests that rising housing prices in the residential communities near Beijing Station could enhance the vitality of the surrounding areas.

5.1.3. Spatial Configuration

The regression analysis reveals that accessibility has the most significant influence on street vitality, emphasizing the critical role that street networks play in sustaining vitality [102]. The MGWR model indicates that the impact of road network structure on vitality exhibits clear temporal variability [103]. NAIN_R5000m, TINT_R10K, and MTL_R5K notably affect the vitality around railway stations. NAIN_R5000m correlates with all four stations in Beijing, showing a negative effect, which diminishes progressively from east to west. This suggests that the walkability of eastern areas has a more pronounced negative influence on neighborhood vitality compared to the western areas. In Space Syntax, an increase in integration signifies greater spatial accessibility, enhancing overall convenience for residents and contributing to the improvement of street vitality [104]. TINT_R10K exerts a positive effect across all four stations, with its impact ranked as Beijing

Station > Beijing South Station > Beijing Fengtai Station > Beijing West Station. In the Random Forest model, increases in the values of TNC R10K (Node Count), TTD R10K (Total Depth), TTD R5K, MTN R5K (Metric Total Nodes), TNC R24K, and TTSL R10K (Total Segment Length) are associated with a decreasing negative impact on the vitality of areas surrounding railway stations, eventually showing a positive influence. Conversely, an increase in the value of MTD R5K (Metric Total Depth) gradually reduces its positive impact on vitality, turning it into a negative influence. These variables, which are related to road network density, indicate that greater complexity in the road network may enhance vitality; as node counts, total depth, and segment lengths increase, the vitality around railway stations is positively affected [99,100].

5.2. Implications for Railway Station Planning and Decision-Making

Based on the findings above, we propose the following planning strategies. First, amid rapid urban development, high-rise buildings in specific areas may play a pivotal role in enhancing urban vitality. However, this impact varies across regions, and planning decisions should consider the specific needs of areas surrounding the stations. For example, for stations such as Beijing Station and Beijing Fengtai Station, where large public facilities dominate, building height exerts minimal influence on vitality. In peripheral areas, however, a moderate increase in building height can positively affect vitality. Such planning should be aligned with local height restrictions.

Second, optimizing the layout of commercial facilities can significantly enhance regional vitality by addressing the needs of various transient populations, including commuters, tourists, and business professionals. Planners should aim to establish a multi-centered commercial structure both in the core and peripheral areas around stations, ensuring the diverse demands of these groups are met while preventing a decline in vitality caused by overly concentrated commercial density. Additionally, a moderate increase in housing prices can contribute to improving regional vitality.

Finally, the influence of road network structure on vitality highlights the importance of accessibility for maintaining vibrant streets. Planners should prioritize optimizing walkability and road network integration around railway stations. This involves adjusting network density and integration according to the specific needs of different areas to mitigate congestion and noise issues caused by high-density road networks.

In conclusion, the planning of areas surrounding railway stations should integrate considerations of building design, functional distribution, floor area ratio control, road network optimization, and housing prices. Such a balanced approach will ensure enhanced regional vitality and promote long-term sustainable development.

6. Conclusions

As urban railways rapidly develop and railway stations see increasing usage, these stations and their surrounding areas are becoming focal points for a variety of urban functions. Therefore, it is critical to fully understand the spatial form, functional distribution, and mechanisms by which existing railway stations impact the vitality of adjacent areas to inform future urban renewal and planning. This study conducts a quantitative analysis of urban vitality, spatial form, and functional facilities distribution using Baidu heatmaps and POI data. By evaluating the performance of different models on both local and global variables, urban planners and decision-makers can more accurately identify the factors influencing vitality around stations, leading to more effective strategy formulation. The main conclusions are as follows:

(1) The spatial distribution of vitality around railway stations reveals significant disparities and an uneven spread of vitality. Policy formulation should account for the developmental context of each station and aim to blur the conceptual boundary between station and city, fostering greater interconnection and integrated growth between railway stations and urban areas within metropolitan clusters.

(2) Multiple regression analysis reveals that commercial density, average number of floors, and road network integration are positively correlated with vitality, while housing prices and residential density show negative correlations. These findings suggest that urban planning should prioritize the enhancement of commercial density and the improvement of transportation networks to boost vitality. The factors influencing vitality vary between weekdays and weekends. Weekday vitality is more predictable, being closely linked to commercial density and transportation infrastructure, whereas weekend vitality is more influenced by commercial density alone, reflecting shifts in activity patterns and needs across time.

(3) On a global scale, the Random Forest (RF) model demonstrates superior performance in predicting vitality around railway stations compared to traditional linear regression and other machine learning models. At the local level, MGWR outperforms conventional GWR and OLS in terms of fit and robustness.

(4) Comparisons between SHAP values and MGWR coefficients reveal that commercial density is the most critical predictor, indicating that the intensity of commercial activities significantly influences the vitality of areas surrounding railway stations. The average number of floors and residential density are identified as fundamental predictors.

This study can be enhanced in several areas: First, the analysis of street vitality could be further refined by exploring its economic, cultural, and social dimensions. Future research should incorporate additional spatial data to provide a more comprehensive understanding of street vitality, thereby deepening the analysis. In the context of high-density urbanization, subjective perceptions and related indicators should be integrated into evaluations of urban vitality. Survey methods could be used to capture public perceptions of various aspects of urban life, thereby providing a unified and subjective assessment of vitality to more accurately depict urban vibrancy. Second, the range of indicators should be expanded. For instance, combining Baidu street view maps for visual assessments of the environment—such as calculating the green view index and openness—could offer a more holistic understanding of the area and support more precise analyses. Third, increasing the sample size would provide more robust insights. This study only focuses on four major special-grade stations in Beijing. To gain a deeper understanding of the relationship between spatial layout, function, and urban street vitality, future research should incorporate additional railway stations from various regions.

Author Contributions: Conceptualization, B.W. and Q.S.; Methodology, Y.S.; Software, Y.S.; Validation, Y.S.; Investigation, Y.S.; Data curation, Y.S.; Writing—original draft, Y.S.; Writing—review & editing, Y.S.; Visualization, Y.S.; Supervision, B.W. and Q.S.; Funding acquisition, B.W. and Q.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, L.; Chen, T.; Lan, T.; Chen, C.; Pan, J. The Contributions of Population Distribution, Healthcare Resourcing, and Transportation Infrastructure to Spatial Accessibility of Health Care. *INQUIRY J. Health Care Organ. Provis. Financ.* **2023**, *60*, 004695802211460. [CrossRef]
2. Jiang, L.; Lai, Y.; Guo, R.; Li, X.; Hong, W.; Tang, X. Measuring the Impact of Government Intervention on the Spatial Variation of Market-Oriented Urban Redevelopment Activities in Shenzhen, China. *Cities* **2024**, *147*, 104834. [CrossRef]

3. Sun, Y.; Cui, Y. Evaluating the Coordinated Development of Economic, Social and Environmental Benefits of Urban Public Transportation Infrastructure: Case Study of Four Chinese Autonomous Municipalities. *Transp. Policy* **2018**, *66*, 116–126. [CrossRef]

4. Ibold, S.; Xia, Y.; Wang, Y. *The Central People's Government of the People's Republic of China Notice of the State Council on Issuing the "14th Five-Year" Plan for the Development of Modern Comprehensive Transpor-Tation System*; German Society for International Cooperation (GIZ) GmbH: Beijing, China, 2022.

5. Pezeshknejad, P.; Monajem, S.; Mozafari, H. Evaluating Sustainability and Land Use Integration of BRT Stations via Extended Node Place Model, an Application on BRT Stations of Tehran. *J. Transp. Geogr.* **2020**, *82*, 102626. [CrossRef]

6. Vale, D.S.; Viana, C.M.; Pereira, M. The Extended Node-Place Model at the Local Scale: Evaluating the Integration of Land Use and Transport for Lisbon's Subway Network. *J. Transp. Geogr.* **2018**, *69*, 282–293. [CrossRef]

7. Yi, Y.; He, F.; Si, Y. Spatial Effects of Railway Network Construction on Urban Sprawl and Its Mechanisms: Evidence from Yangtze River Delta Urban Agglomeration, China. *Land* **2023**, *13*, 25. [CrossRef]

8. Lunardon, A.; Vladimirova, D.; Boucsein, B. How Railway Stations Can Transform Urban Mobility and the Public Realm: The Stakeholders' Perspective. *J. Urban Mobil.* **2023**, *3*, 100047. [CrossRef]

9. Otsuka, N.; Wittowsky, D.; Damerau, M.; Gerten, C. Walkability Assessment for Urban Areas around Railway Stations along the Rhine-Alpine Corridor. *J. Transp. Geogr.* **2021**, *93*, 103081. [CrossRef]

10. Ma, F.; Ren, F.; Yuen, K.F.; Guo, Y.; Zhao, C.; Guo, D. The Spatial Coupling Effect between Urban Public Transport and Commercial Complexes: A Network Centrality Perspective. *Sustain. Cities Soc.* **2019**, *50*, 101645. [CrossRef]

11. Du, J.; van Wesemael, P.; Druta, O. Place Quality in High-Speed Rail Station Areas. *J. Transp. Land Use* **2021**, *14*, 1165–1186. [CrossRef]

12. Küpper, M.; Seyfried, A. Analysis of Space Usage on Train Station Platforms Based on Trajectory Data. *Sustainability* **2020**, *12*, 8325. [CrossRef]

13. Fang, C.; He, S.; Wang, L. Spatial Characterization of Urban Vitality and the Association With Various Street Network Metrics From the Multi-Scalar Perspective. *Front. Public Health* **2021**, *9*, 677910. [CrossRef] [PubMed]

14. Lu, S.; Shi, C.; Yang, X. Impacts of Built Environment on Urban Vitality: Regression Analyses of Beijing and Chengdu, China. *Int. J. Environ. Res. Public Health* **2019**, *16*, 4592. [CrossRef] [PubMed]

15. Guo, X.; Chen, H.; Yang, X. An Evaluation of Street Dynamic Vitality and Its Influential Factors Based on Multi-Source Big Data. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 143. [CrossRef]

16. Wu, C.; Ye, X.; Ren, F.; Du, Q. Check-in Behaviour and Spatio-Temporal Vibrancy: An Exploratory Analysis in Shenzhen, China. *Cities* **2018**, *77*, 104–116. [CrossRef]

17. Mu, B.; Liu, C.; Mu, T.; Xu, X.; Tian, G.; Zhang, Y.; Kim, G. Spatiotemporal Fluctuations in Urban Park Spatial Vitality Determined by On-Site Observation and Behavior Mapping: A Case Study of Three Parks in Zhengzhou City, China. *Urban For. Urban Green.* **2021**, *64*, 127246. [CrossRef]

18. Guo, X.; Yang, Y.; Cheng, Z.; Wu, Q.; Li, C.; Lo, T.; Chen, F. Spatial Social Interaction: An Explanatory Framework of Urban Space Vitality and Its Preliminary Verification. *Cities* **2022**, *121*, 103487. [CrossRef]

19. Niu, Y.; Mi, X.; Wang, Z. Vitality Evaluation of the Waterfront Space in the Ancient City of Suzhou. *Front. Archit. Res.* **2021**, *10*, 729–740. [CrossRef]

20. Mushkani, R.A.; Ono, H. The Role of Land Use and Vitality in Fostering Gender Equality in Urban Public Parks: The Case of Kabul City, Afghanistan. *Habitat Int.* **2021**, *118*, 102462. [CrossRef]

21. Wu, J.; Ta, N.; Song, Y.; Lin, J.; Chai, Y. Urban Form Breeds Neighborhood Vibrancy: A Case Study Using a GPS-Based Activity Survey in Suburban Beijing. *Cities* **2018**, *74*, 100–108. [CrossRef]

22. Xu, Y.; Chen, X. Quantitative Analysis of Spatial Vitality and Spatial Characteristics of Urban Underground Space (UUS) in Metro Area. *Tunn. Undergr. Space Technol.* **2021**, *111*, 103875. [CrossRef]

23. Xu, Z.; Lu, P. Evaluation of Urban Vitality in Shandong Based on Multi-Source Data. *E3S Web Conf.* **2023**, *372*, 01008. [CrossRef]

24. García-Palomares, J.C.; Salas-Olmedo, M.H.; Moya-Gómez, B.; Condeço-Melhorado, A.; Gutiérrez, J. City Dynamics through Twitter: Relationships between Land Use and Spatiotemporal Demographics. *Cities* **2018**, *72*, 310–319. [CrossRef]

25. Zhang, E.; Xie, H.; Long, Y. Decoding the Association between Urban Streetscape Skeletons and Urban Activities: Experiments in Beijing Using Dazhong Dianping Data. *Trans. Urban Data Sci. Technol.* **2023**, *2*, 3–18. [CrossRef]

26. Ouyang, J.; Fan, H.; Wang, L.; Zhu, D.; Yang, M. Revealing Urban Vibrancy Stability Based on Human Activity Time-Series. *Sustain. Cities Soc.* **2022**, *85*, 104053. [CrossRef]

27. Yang, J.; Cao, J.; Zhou, Y. Elaborating Non-Linear Associations and Synergies of Subway Access and Land Uses with Urban Vitality in Shenzhen. *Transp. Res. Part A Policy Pract.* **2021**, *144*, 74–88. [CrossRef]

28. Dong, Q.; Cai, J.; Chen, S.; He, P.; Chen, X. Spatiotemporal Analysis of Urban Green Spatial Vitality and the Corresponding Influencing Factors: A Case Study of Chengdu, China. *Land* **2022**, *11*, 1820. [CrossRef]

29. Lv, G.; Zheng, S.; Hu, W. Exploring the Relationship between the Built Environment and Block Vitality Based on Multi-Source Big Data: An Analysis in Shenzhen, China. *Geomat. Nat. Hazards Risk* **2022**, *13*, 1593–1613. [CrossRef]

30. Zhang, S.; Zhang, W.; Wang, Y.; Zhao, X.; Song, P.; Tian, G.; Mayer, A.L. Comparing Human Activity Density and Green Space Supply Using the Baidu Heat Map in Zhengzhou, China. *Sustainability* **2020**, *12*, 7075. [CrossRef]

31. Jiang, Y.; Huang, Z.; Zhou, X.; Chen, X. Evaluating the Impact of Urban Morphology on Urban Vitality: An Exploratory Study Using Big Geo-Data. *Int. J. Digit. Earth* **2024**, *17*, 2327571. [CrossRef]

32. Wang, Z.; Wang, X.; Liu, Y.; Zhu, L. Identification of 71 Factors Influencing Urban Vitality and Examination of Their Spatial Dependence: A Comprehensive Validation Applying Multiple Machine-Learning Models. *Sustain. Cities Soc.* **2024**, *108*, 105491. [CrossRef]

33. Jiang, Y.; Chen, Z.; Sun, P. Urban Shrinkage and Urban Vitality Correlation Research in the Three Northeastern Provinces of China. *Int. J. Environ. Res. Public Health* **2022**, *19*, 10650. [CrossRef] [PubMed]

34. Liu, M.; Jiang, Y.; He, J. Quantitative Evaluation on Street Vitality: A Case Study of Zhoujiadu Community in Shanghai. *Sustainability* **2021**, *13*, 3027. [CrossRef]

35. Zeng, Z.; Li, Y.; Tang, H. Multidimensional Spatial Driving Factors of Urban Vitality Evolution at the Subdistrict Scale of Changsha City, China, Based on the Time Series of Human Activities. *Buildings* **2023**, *13*, 2448. [CrossRef]

36. Yang, L.; Ao, Y.; Ke, J.; Lu, Y.; Liang, Y. To Walk or Not to Walk? Examining Non-Linear Effects of Streetscape Greenery on Walking Propensity of Older Adults. *J. Transp. Geogr.* **2021**, *94*, 103099. [CrossRef]

37. Chen, J.; Yan, Z.; Fan, R. Elements of the Core Area Transport System Oriented to Station–City Integration. *Comput. Electr. Eng.* **2023**, *110*, 108842. [CrossRef]

38. Wang, F.; Wei, X.; Liu, J.; He, L.; Gao, M. Impact of High-Speed Rail on Population Mobility and Urbanisation: A Case Study on Yangtze River Delta Urban Agglomeration, China. *Transp. Res. Part A Policy Pract.* **2019**, *127*, 99–114. [CrossRef]

39. Ye, Y.; Li, D.; Liu, X. How Block Density and Typology Affect Urban Vitality: An Exploratory Analysis in Shenzhen, China. *Urban Geogr.* **2018**, *39*, 631–652. [CrossRef]

40. Kim, S. Urban Vitality, Urban Form, and Land Use: Their Relations within a Geographical Boundary for Walkers. *Sustainability* **2020**, *12*, 10633. [CrossRef]

41. Li, S.; Wu, C.; Lin, Y.; Li, Z.; Du, Q. Urban Morphology Promotes Urban Vibrancy from the Spatiotemporal and Synergetic Perspectives: A Case Study Using Multisource Data in Shenzhen, China. *Sustainability* **2020**, *12*, 4829. [CrossRef]

42. Najaf, P.; Thill, J.-C.; Zhang, W.; Fields, M.G. City-Level Urban Form and Traffic Safety: A Structural Equation Modeling Analysis of Direct and Indirect Effects. *J. Transp. Geogr.* **2018**, *69*, 257–270. [CrossRef]

43. Song, Z.; Yu, L. Multifractal Features of Spatial Variation in Construction Land in Beijing (1985–2015). *Palgrave Commun* **2019**, *5*, 68. [CrossRef]

44. Li, M.; Pan, J. Assessment of Influence Mechanisms of Built Environment on Street Vitality Using Multisource Spatial Data: A Case Study in Qingdao, China. *Sustainability* **2023**, *15*, 1518. [CrossRef]

45. Wu, C.; Ye, Y.; Gao, F.; Ye, X. Using Street View Images to Examine the Association between Human Perceptions of Locale and Urban Vitality in Shenzhen, China. *Sustain. Cities Soc.* **2023**, *88*, 104291. [CrossRef]

46. Xiao, L.; Lo, S.; Liu, J.; Zhou, J.; Li, Q. Nonlinear and Synergistic Effects of TOD on Urban Vibrancy: Applying Local Explanations for Gradient Boosting Decision Tree. *Sustain. Cities Soc.* **2021**, *72*, 103063. [CrossRef]

47. Li, M.; Liu, J.; Lin, Y.; Xiao, L.; Zhou, J. Revitalizing Historic Districts: Identifying Built Environment Predictors for Street Vibrancy Based on Urban Sensor Data. *Cities* **2021**, *117*, 103305. [CrossRef]

48. Wang, B.; Lei, Y.; Xue, D.; Liu, J.; Wei, C. Elaborating Spatiotemporal Associations Between the Built Environment and Urban Vibrancy: A Case of Guangzhou City, China. *Chin. Geogr. Sci.* **2022**, *32*, 480–492. [CrossRef]

49. Liang, Y.; Song, W.; Dong, X. Evaluating the Space Use of Large Railway Hub Station Areas in Beijing toward Integrated Station-City Development. *Land* **2021**, *10*, 1267. [CrossRef]

50. Yamu, C.; Van Nes, A.; Garau, C. Bill Hillier's Legacy: Space Syntax—A Synopsis of Basic Concepts, Measures, and Empirical Application. *Sustainability* **2021**, *13*, 3394. [CrossRef]

51. Ye, Y.; Nes, A.V.N. Quantitative Tools in Urban Morphology: Combining Space Syntax, Spacematrix and Mixed-Use Index in a GIS Framework. *Urban Morphol.* **2014**, *18*, 97–118. [CrossRef]

52. Hillier, B. *Space Is the Machine: A Configurational Theory of Architecture*; Cambridge University Press: Cambridge, UK, 1996; ISBN 978-0-521-56039-9.

53. Meng, L.; Ishida, T. Analysis of the Relationship between Beijing Rail Transit and Urban Planning Based on Space Syntax. *Sustainability* **2022**, *14*, 8744. [CrossRef]

54. Liu, W.; Guo, J.; Wu, W.; Cao, Y. The Evolution of Regional Spatial Structure Influenced by Passenger Rail Service: A Case Study of the Yangtze River Delta. *Growth Change* **2022**, *53*, 651–679. [CrossRef]

55. Wang, X.; Liu, J.; Zhang, W. How Does the Spatial Structure of High-Speed Rail Station Areas Evolve? A Case Study of Zhengzhou East Railway Station, China. *Appl. Sci.* **2021**, *11*, 11132. [CrossRef]

56. Jia, H.; Li, Q. Analysis of Spatial Interaction Vitality Based on High-Speed Railway Network and Highway Network. *J. Phys. Conf. Ser.* **2020**, *1682*, 012082. [CrossRef]

57. Deng, T.; Gan, C.; Perl, A.; Wang, D. What Caused Differential Impacts on High-Speed Railway Station Area Development? Evidence from Global Nighttime Light Data. *Cities* **2020**, *97*, 102568. [CrossRef]

58. Ortuño-Padilla, A.; Espinosa-Flor, A.; Cerdán-Aznar, L. Development Strategies at Station Areas in Southwestern China: The Case of Mianyang City. *Land Use Policy* **2017**, *68*, 660–670. [CrossRef]

59. He, J.; Chen, L.; Zhang, M.-Z.; Bae, K.-H. Research on the Impact of China's High-Speed Railway Construction on Regional Economy—Take Xibao High-Speed Rail as an Example. *Korean Logist. Res. Assoc.* **2019**, *29*, 75–89. [CrossRef]

60. Zheng, L.; Long, F.; Chang, Z.; Ye, J. Ghost Town or City of Hope? The Spatial Spillover Effects of High-Speed Railway Stations in China. *Transp. Policy* **2019**, *81*, 230–241. [CrossRef]

61. Jiang, M.; Kim, E. Impact of High-Speed Railroad on Regional Income Inequalities in China and Korea. *Int. J. Urban Sci.* **2016**, *20*, 393–406. [CrossRef]

62. Ke, X.; Chen, H.; Hong, Y.; Hsiao, C. Do China's High-Speed-Rail Projects Promote Local Economy?—New Evidence from a Panel Data Approach. *China Econ. Rev.* **2017**, *44*, 203–226. [CrossRef]

63. Li, X.; Li, Y.; Jia, T.; Zhou, L.; Hijazi, I.H. The Six Dimensions of Built Environment on Urban Vitality: Fusion Evidence from Multi-Source Data. *Cities* **2022**, *121*, 103482. [CrossRef]

64. Long, Y.; Huang, C. Does Block Size Matter? The Impact of Urban Design on Economic Vitality for Chinese Cities. *Environ. Plan. B Urban Anal. City Sci.* **2019**, *46*, 406–422. [CrossRef]

65. Sulis, P.; Manley, E.; Zhong, C.; Batty, M. Using Mobility Data as Proxy for Measuring Urban Vitality. *J. Spat. Inf. Sci.* **2018**, *2018*, 137–162. [CrossRef]

66. Shi, Y.; Zheng, J.; Pei, X. Measurement Method and Influencing Mechanism of Urban Subdistrict Vitality in Shanghai Based on Multisource Data. *Remote Sens.* **2023**, *15*, 932. [CrossRef]

67. Meng, Y.; Xing, H. Exploring the Relationship between Landscape Characteristics and Urban Vibrancy: A Case Study Using Morphology and Review Data. *Cities* **2019**, *95*, 102389. [CrossRef]

68. Liu, S.; Zhang, L.; Long, Y.; Long, Y.; Xu, M. A New Urban Vitality Analysis and Evaluation Framework Based on Human Activity Modeling Using Multi-Source Big Data. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 617. [CrossRef]

69. Cao, X.; Shi, Y.; Zhou, L.; Tao, T.; Yang, Q. Analysis of Factors Influencing the Urban Carrying Capacity of the Shanghai Metropolis Based on a Multiscale Geographically Weighted Regression (MGWR) Model. *Land* **2021**, *10*, 578. [CrossRef]

70. Huang, X.; Gong, P.; Wang, S.; White, M.; Zhang, B. Machine Learning Modeling of Vitality Characteristics in Historical Preservation Zones with Multi-Source Data. *Buildings* **2022**, *12*, 1978. [CrossRef]

71. Lin, J.; Zhuang, Y.; Zhao, Y.; Li, H.; He, X.; Lu, S. Measuring the Non-Linear Relationship between Three-Dimensional Built Environment and Urban Vitality Based on a Random Forest Model. *Int. J. Environ. Res. Public Health* **2022**, *20*, 734. [CrossRef]

72. Wang, R.; Chu, H.; Liu, Q.; Chen, B.; Zhang, X.; Fan, X.; Wu, J.; Xu, K.; Jiang, F.; Chen, L. Application of Machine Learning Techniques to Improve Multi-Radar Mosaic Precipitation Estimates in Shanghai. *Atmosphere* **2023**, *14*, 1364. [CrossRef]

73. Pozoukidou, G.; Ntriankos, I. Measuring and Assessing Urban Sprawl: A Proposed Indicator System for the City of Thessaloniki, Greece. *Remote Sens. Appl. Soc. Environ.* **2017**, *8*, 30–40. [CrossRef]

74. Ding, C.; Cao, X.; Dong, M.; Zhang, Y.; Yang, J. Non-Linear Relationships between Built Environment Characteristics and Electric-Bike Ownership in Zhongshan, China. *Transp. Res. Part D Transp. Environ.* **2019**, *75*, 286–296. [CrossRef]

75. Wang, J.; Wang, S.; Song, M.; Li, W.; Ma, C.; Lu, X.; Li, D. Review of Research on Impact of Urban Built Environment on Travel Behavior in Big Data Context. Available online: https://kns.cnki.net/kcms2/article/abstract?v=QHiZY5KKB7a2UjS2Bpnqe3DPf4GODhPiop1ARIEzf2GAsNeOjNA_VJ8tYPx_srMYOOFspg4zf0RUXVSlAHNlK89vhsqfUtZgxL9Ssk7WTfJPWQpN-Wu35n1T97e4QabR4JHhsF4Q39YEzDL6GG_DZqtqxY0TkHmt-4wVUeKcp4KOI-YCfBDKgPjevn_klZUD&uniplatform=NZKPT&language=CHS (accessed on 1 October 2024).

76. Wu, W.; Dang, Y.; Zhao, K.; Chen, Z.; Niu, X. Spatial Nonlinear Effects of Urban Vitality Under the Constraints of Development Intensity and Functional Diversity. *Alex. Eng. J.* **2023**, *77*, 645–656. [CrossRef]

77. Kou, X.; Zhou, J. *Beijing Chronicle·Municipal Volume·Railway Transport Chronicle*; Beijing Press: Beijing, China, 2004; ISBN 978-7-200-05098-1.

78. Yang, S.; Zhang, L.; Zhuang, Y. Relations between Function Distribution and Vitality within Railway Station Areas from Perspective of Station-City Synthetic Development: Case Studies on Four Principal Chinese Stations. *Shanghai Urban Plan. Rev.* **2021**, *161*, 106–112.

79. Priemus, H. HST-Railway Stations as Dynamic Nodes in Urban Networks. *Citeseer* **2006**, *5*, 101–121.

80. Ma, Q. Recent Studies on Transit-oriented Development in North America. *Urban Plan. Int.* **2003**, *5*, 45–50.

81. Zhang, H. Extracting Active Population Data Based on Baidu Heat Maps for Transportation Planning Appli-cations. *Urban Transp. China* **2021**, *19*, 103–111. [CrossRef]

82. Trevor, H.; Robert, T.; Jerome, F.; Hastie, T.; Friedman, J.; Tibshirani, R. *The Elements of Statistical Learning*; Springer: New York, NY, USA, 2009; p. 2.

83. Brunsdon, C.; Fotheringham, A.S.; Charlton, M. Some Notes on Parametric Significance Tests for Geographically Weighted Regression. *J. Reg. Sci.* **1999**, *39*, 497–524. [CrossRef]

84. Fotheringham, A.S.; Yang, W.; Kang, W. Multiscale Geographically Weighted Regression (MGWR). *Ann. Am. Assoc. Geogr.* **2017**, *107*, 1247–1265. [CrossRef]

85. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

86. Chen, T.; Guestrin, C. Xgboost: A Scalable Tree Boosting System. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.

87. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. Lightgbm: A Highly Efficient Gradient Boosting Decision Tree. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.

88. Lundberg, S.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. *arXiv* **2017**, arXiv:1705.07874.

89. Barnston, A.G. Correspondence among the Correlation, RMSE, and Heidke Forecast Verification Measures; Refinement of the Heidke Score. *Weather Forecast.* **1992**, *7*, 699–709. [CrossRef]

90. Fahrmeir, L.; Kneib, T.; Lang, S. *Regressionsmodelle*; Springer: Berlin/Heidelberg, Germany, 2007; ISBN 3-540-33932-9.

91. Comber, A.; Brunsdon, C.; Charlton, M.; Dong, G.; Harris, R.; Lu, B.; Lü, Y.; Murakami, D.; Nakaya, T.; Wang, Y.; et al. A Route Map for Successful Applications of Geographically Weighted Regression. *Geogr. Anal.* **2023**, *55*, 155–178. [CrossRef]

92. Ozili, P.K. *The Acceptable R-Square in Empirical Modelling for Social Science Research*; University Library of Munich: München, Germany, 2023.

93. Chen, Y.; Yu, B.; Shu, B.; Yang, L.; Wang, R. Exploring the Spatiotemporal Patterns and Correlates of Urban Vitality: Temporal and Spatial Heterogeneity. *Sustain. Cities Soc.* **2023**, *91*, 104440. [CrossRef]

94. Xuan, W.; Yao, Y.; Zhao, L.; Wang, C.; Xiao, J. The Influence Mechanism of Urban Built Environment on the SpatialDistribution of Urban Vitality from the Perspective of Multi-source Data. *Sci. Technol. Eng.* **2023**, *23*, 11349–11363.

95. Wang, N.; Wu, J.; Li, S.; Wang, H.; Peng, Z. Spatial Features of Urban Vitality and the lmpact of Built Environmenton Them Based on Multi-Source Data:A Case Study of Shenzhen. *Trop. Geogr.* **2021**, *41*, 1280–1291. [CrossRef]

96. Brueckner, J.K.; Fu, S.; Gu, Y.; Zhang, J. Measuring the Stringency of Land Use Regulation: The Case of China's Building Height Limits. *Rev. Econ. Stat.* **2017**, *99*, 663–677. [CrossRef]

97. Fan, L.; Zhang, D. Research on the Influence Mechanism and Spatial Heterogeneity Characteristics of Block Vitality in Beijing: Based on Multi-Scale Geographically Weighted Regression. *City Plan. Rev.* **2022**, *46*, 27–37.

98. Tu, W.; Zhu, T.; Zhong, C.; Zhang, X.; Xu, Y.; Li, Q. Exploring Metro Vibrancy and Its Relationship with Built Environment: A Cross-City Comparison Using Multi-Source Urban Data. *Geo-Spat. Inf. Sci.* **2022**, *25*, 182–196. [CrossRef]

99. Zhang, P.; Zhang, T.; Fukuda, H.; Ma, M. Evidence of Multi-Source Data Fusion on the Relationship between the Specific Urban Built Environment and Urban Vitality in Shenzhen. *Sustainability* **2023**, *15*, 6869. [CrossRef]

100. Yu, B.; Cui, X.; Li, H.; Luo, P.; Liu, R.; Yang, T. TOD and Vibrancy: The Spatio-Temporal Impacts of the Built Environment on Vibrancy. *Front. Environ. Sci.* **2022**, *10*, 1009094. [CrossRef]

101. Tsai, I.-C. Housing Price Convergence, Transportation Infrastructure and Dynamic Regional Population Relocation. *Habitat Int.* **2018**, *79*, 61–73. [CrossRef]

102. Li, X.; Lv, Z.; Zheng, Z.; Zhong, C.; Hijazi, I.H.; Cheng, S. Assessment of Lively Street Network Based on Geographic Information System and Space Syntax. *Multimed. Tools Appl.* **2017**, *76*, 17801–17819. [CrossRef]

103. Li, X.; Qian, Y.; Zeng, J.; Wei, X.; Guang, X. The Influence of Strip-City Street Network Structure on Spatial Vitality: Case Studies in Lanzhou, China. *Land* **2021**, *10*, 1107. [CrossRef]

104. Zhang, J.; Zhang, J.; Yu, S.; Zhou, J. The Sustainable Development of Street Texture of Historic and Cultural Districts—A Case Study in Shichahai District, Beijing. *Sustainability* **2018**, *10*, 2343. [CrossRef]