

Article

Study on the Relationship between the Spatial Distribution of Shared Bicycle Travel Demand and Urban Built Environment

Lili Yang ¹ , Simeng Fei ¹, Hongfei Jia ^{1,*}, Jingdong Qi ², Luyao Wang ³ and Xinning Hu ¹

¹ College of Transportation, Jilin University, Changchun 130022, China; yanglili@jlu.edu.cn (L.Y.); feisimeng0855@jlu.cn (S.F.); huxn22@mails.jlu.edu.cn (X.H.)

² Changchun Municipal Engineering Design & Research Institute, Changchun 130022, China; qijingdong1991@163.com

³ Shenyang Urban Planning & Design Institute, Shenyang 110004, China; trafficplanning@163.com

* Correspondence: jiahf@jlu.edu.cn; Tel.: +86-0431-85095370

Abstract: As a green and sustainable trip mode, shared bicycles play an essential role in completing short-distance trips in cities. This paper proposes a method to analyze the impact of the urban built environment on the distribution of shared bicycles in a small-scale space. First, the Fishnet function of ArcGIS is utilized to divide the study area into grids of 500 m × 500 m. Then, three indicators are proposed to describe the characteristics of the urban built environment, including point of information (POI) comprehensive index, the intensity of public transportation coverage, spatial accessibility, providing them the ways to be assigned to the grids. Finally, the multivariable linear regression model and support vector regression (SVR) models are applied to reveal the impacts of built environment factors on the spatial distribution of shared bicycles. Results show that SVR models based on linear kernel function, Gaussian radial basis kernel function, and Polynomial kernel function can achieve better analysis results. The SVR model based on the Gaussian radial basis function shows higher explanatory power (adjusted $R^2 = 0.978$) than the multivariable linear regression model (adjusted $R^2 = 0.847$). This paper can aid in understanding the demand and supply of shared bicycles and help operators or governments to improve service quality.

Keywords: shared bicycle; support vector regression; built environment; space syntax



Citation: Yang, L.; Fei, S.; Jia, H.; Qi, J.; Wang, L.; Hu, X. Study on the Relationship between the Spatial Distribution of Shared Bicycle Travel Demand and Urban Built Environment. *Sustainability* **2023**, *15*, 13576. <https://doi.org/10.3390/su151813576>

Academic Editors: Luca D’Acierno and Armando Carteni

Received: 2 August 2023

Revised: 31 August 2023

Accepted: 7 September 2023

Published: 11 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Shared bicycles can not only provide convenient trip services for urban residents, but also can help them to obtain access to public transportation. They play a vital role in improving the efficiency of short-distance trips, solving the first and last mile travel problem of urban traffic, and alleviating a city’s traffic stress [1,2]. However, while sharing bicycles facilitates residents’ trips and improves trip efficiency, there is also an imbalance between supply and demand. For example, few bicycles are in areas with high demand, while too many are in areas with low demand [3]. When shared bicycles are placed, only the area’s land use is generally considered, so areas such as commercial districts, subway stations, and bus stations tend to put more bicycles [4]. However, although the land use of some areas is the same, the demand for bicycles may be quite different. This is because different urban built environments have different levels of adaptability to shared bicycles; however, land use is only one factor. A series of factors reflecting the urban built environment, such as the accessibility of road network, the distribution of POI, and public transport coverage, will affect the use of shared bicycles. In this study, we discuss the impact of urban built environment factors on the spatial distribution of shared bicycles and explore their usage. The findings of this paper guide the understanding of the demand and supply of shared bicycles and improve service quality.

The studies related to public bicycle trips mainly focus on predicting demand [5], scheduling optimization [6], cycling OD identification [7], and path selection [8]. The first

two types of studies mainly focus on docked public bicycles and pay less attention to dockless shared bicycles. The practical significance of the research on OD identification and path selection is to guide bicycle network planning, and there is little research on operation and management. The characteristics of the urban built environment are the critical factors affecting the use of public bicycles [9–12]. This study shows that bicycle trips are affected by the 3D elements of urban form, namely density, diversity of land use, and urban design. Density mainly refers to the density of public transport stations, kinds of POI density, shared bicycle station density, etc.; it mostly indicates land-use diversity. Design mostly represents urban spatial structure design. Based on this, some scholars gradually expanded the urban built environment elements to “5D” and added “destination accessibility” and “distance to transit”. Later, some studies incorporated “demand management” into influencing factors to form “6D”. Indicators related to population statistics have gradually entered the vision of relevant researchers and put forward “7D”. Weliwitiya [13] set a linear model to analyze the impact of a built environment on the utilization rate of bicycles. It was found that the density of non-motorized lanes and the passenger flow of subway have a positive impact on it. Noland et al. [14] used the public bicycle operation data of New York City to determine that the utilization rate of the public bicycle is related to factors such as population density, distance from bus stops, and catering outlets. In addition, some studies have proved that mixed land use can reduce travel distance and improve the utilization rate of shared bicycles [15,16]. For example, commercial places mixed with public buildings positively impact the use of shared bicycles [17]. The higher the density of catering outlets and shopping outlets near the station, and the closer the station to the subway station, the greater the number of public bicycles used [18,19]. With the gradual development of studies, more scholars have been analyzing the impact of the urban built environment on shared bicycles from aspects such as infrastructure [20], population density [21], land use [22,23], and traffic accessibility [24].

In general, the studies on the impact of the urban built environment on shared bicycles have not yet formed a complete theoretical method system. Most existing studies either take the whole study area as the analysis unit to analyze the relationship between the distribution of shared bicycles and the built environment, or make a horizontal comparison between different areas, which cannot focus on the smaller spatial scale, such as the street level.

Meanwhile, in terms of research methods, most of the existing studies rely on linear regression models to analyze the built environment’s impact on the distribution of shared bicycles, such as multiple linear regression [25,26] and least squares regression models [27]. However, the regression effect of these models is not very good, and the regression coefficients of most studies are between 0.5 and 0.7. These linear regression models cannot measure the complex relationship between the distribution of shared bicycles and influencing factors, nor can they analyze the impact of different built environments on the distribution of shared bicycles in a small-scale space.

This paper uses Spatial Syntax theory and ArcGIS software 10.3 to assign the urban built environment elements and shared bicycle borrowing and returning location data on the urban spatial grid network. It describes the spatial distribution of shared bicycles with the grid as a unit, obtains spatial geographic information data, analyzes the spatial distribution characteristics of shared bicycles, and discusses the influencing factors. The road network axis map is drawn and imported into Space Syntax software DepthMap X0.7.0 to obtain the accessibility index of the road network, which is used to accurately analyze the impact of the road network on the distribution of shared bicycles. By combining the factors influencing the distribution of shared bicycles, prediction models based on SVR are proposed to explore the rules between the built environment and the distribution of shared bicycles. It provides a basis for the rational planning of shared bicycle parking areas and is of great significance to improving shared bicycles’ operating status and service level.

The remainder of this paper is organized as follows: In Section 2, the situation of the study area and the collection methods employed to obtain the required research data are

introduced. In Section 3, the extraction method of urban built environment factors and its assignment method in the grid are proposed, and the SVR model is introduced to explore how an urban built environment affects the distribution of shared bicycles. In Section 4, the analysis results of the model are given, and the effectiveness of the proposed model is verified. Finally, a summary of our findings is concluded in Section 5.

2. Data

2.1. Study Area

When selecting the study area, it is necessary to consider the specific usage of shared bicycles in the area. It requires a relatively stable state regarding the number of shared bicycle users and cycling behavior. Shanghai is one of the four municipalities directly under the central government in China, and it is also the center of our economy, culture, trade, science, and technology. It has a superior geographical location and dense population, and is perfectly developed. Shared bicycles are developing well in Shanghai; the number of daily active vehicles of shared bicycles is stable at 450,000 to 550,000, the daily turnover rate of bicycles is about 5.0 times, and the user's riding behavior is stable. Therefore, this paper selects downtown Shanghai as the study area based on this condition, as shown in Figure 1. The research area includes seven significant districts, Yangpu District, Putuo District, Hongkou District, Jingan District, Changning District, Xuhui District, and Huangpu District, with about 266 km² and more than 7 million people by 2020.

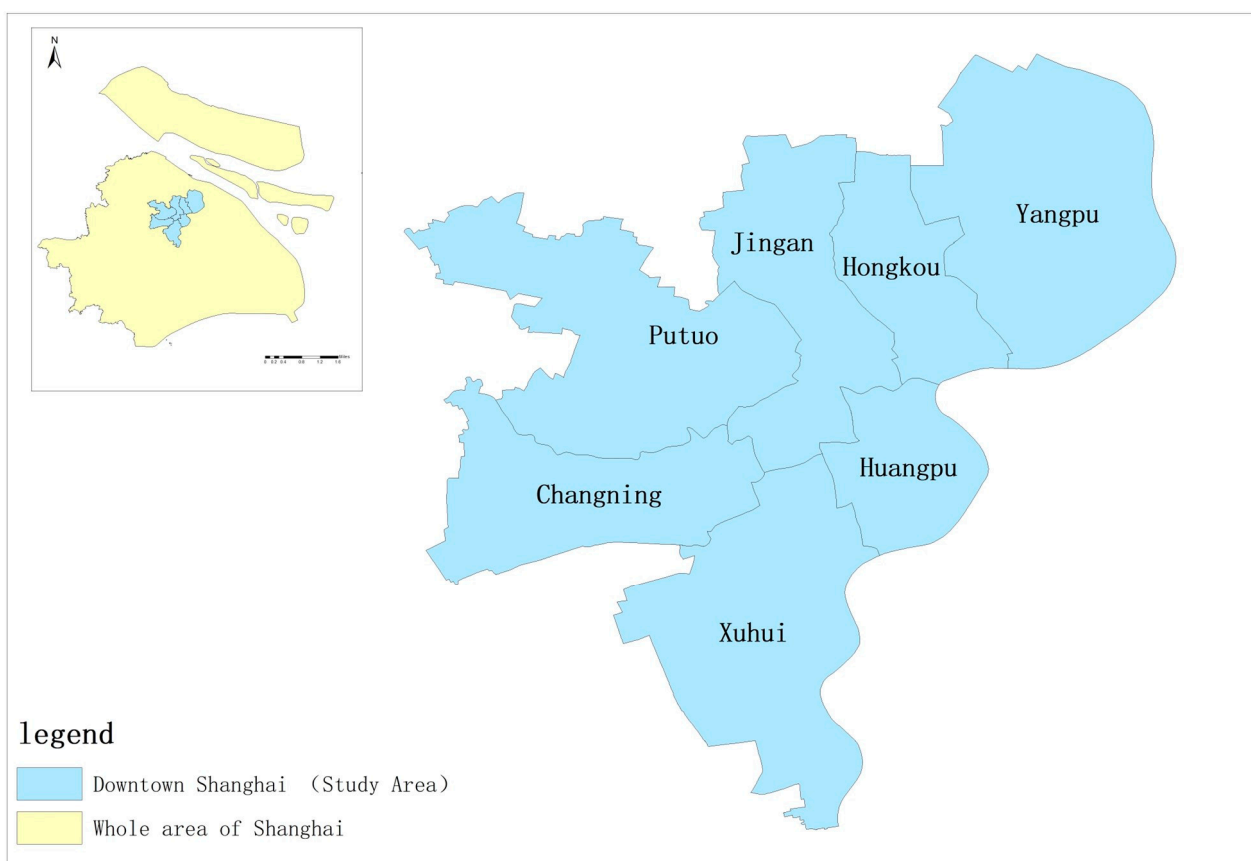


Figure 1. Study area.

2.2. Data Collection

2.2.1. Shared Bicycle Travel Data

This paper obtained 8,453,497 shared bicycle riding records within a week in downtown Shanghai, including bicycle ID, time information, and longitude and latitude information. The steps required to convert the GPS data of shared bicycles into borrowing and

returning quantity data include the following: 1. Calculate the GPS data of the shared bicycles' parking position into a data table according to the time information; 2. Compare the bicycle IDs in the two adjacent time tables of t , using $t + 1$ to eliminate data with the same ID. If data with a different ID exist in the statistical area at time t but are not in the statistical area at time $t + 1$, the bicycle was being borrowed; otherwise, the bicycle was being returned; 3. Compare the longitude and latitude information of the data with the same ID. If the longitude and latitude of the bicycle change during the period $(t, t + 1)$, it is the case that the bicycle that has been used. The data at time t are included in the bicycle borrowing statistics table, and the data at time $t + 1$ are included in the bicycle returning statistics table. The specific operation process is shown in Figure 2.

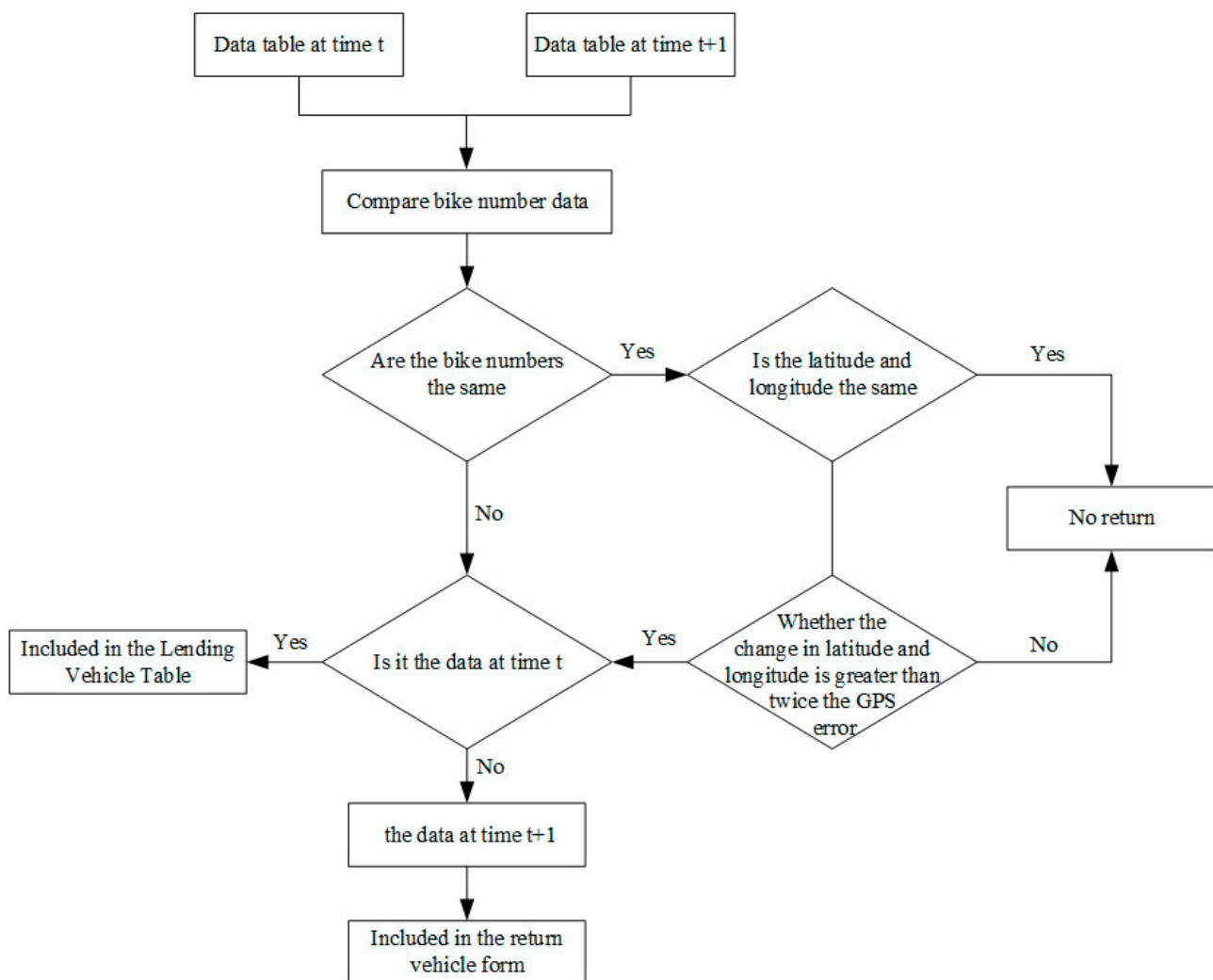


Figure 2. Calculation method of borrowing and returning position of shared bicycle.

2.2.2. Urban Built Environment Data

POI is used as an emerging data source in the study, as it can efficiently represent the situation of the urban built environment. To analyze the built environment's impact on the distribution of shared bicycles, this paper obtains the vector boundary data, bus line and station data, rail transit line and station data, and POI data of downtown Shanghai in 2020 from Baidu Maps.

3. Methodology

3.1. Extraction of Urban Built Environment Factors

Using ArcGIS software 10.3 to analyze the nuclear density, the spatial distribution of the shared bicycles in the downtown Shanghai can be obtained, as shown in Figure 3.

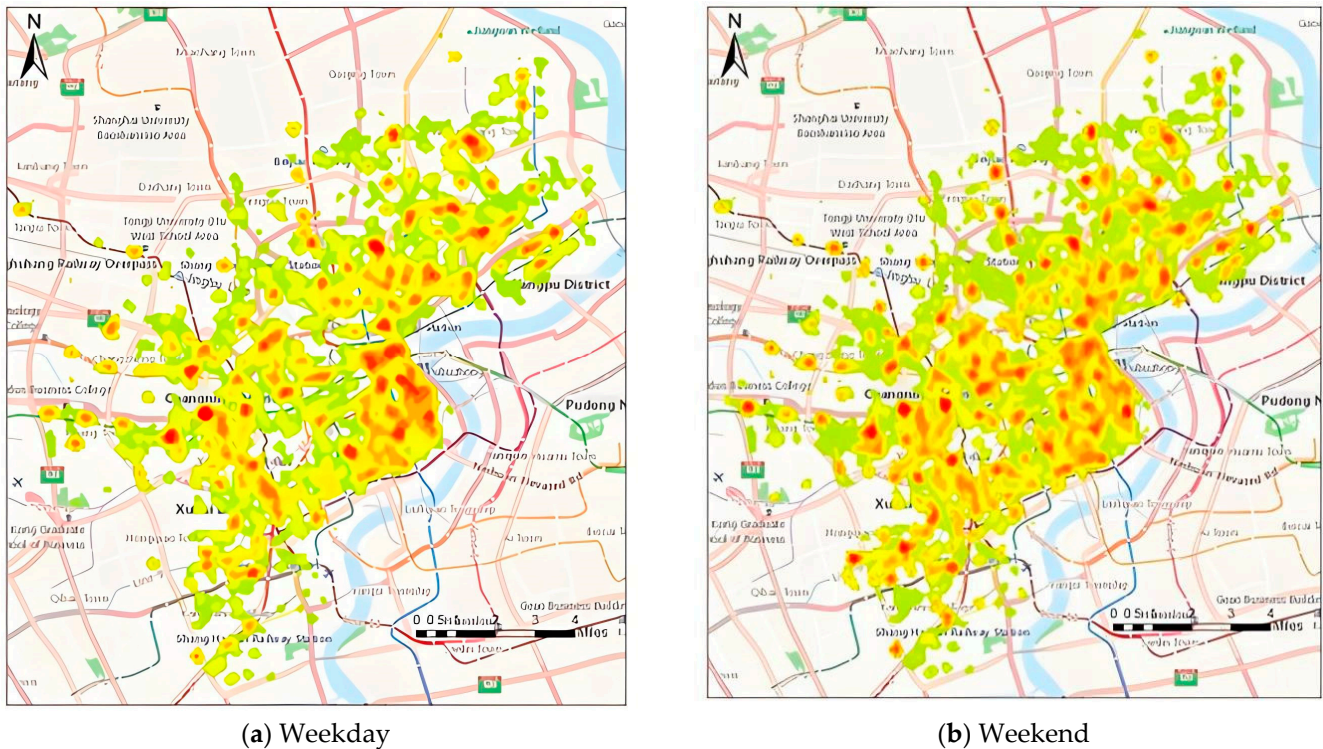


Figure 3. Spatial distribution of shared bicycles in downtown Shanghai.

As shown in Figure 3, the shared bicycles in downtown Shanghai on weekdays and weekends are relatively similar in spatial distribution and present an aggregation state. From the perspective of administrative divisions, the areas with higher shared bicycle density are Yangpu District, Huangpu District, Putuo District, and Hongkou District. At the street level, shared bicycles are concentrated in public transport stations, shopping areas, office buildings, education institutions, catering, and other service industries, as well as locations with high road accessibility.

According to the distribution characteristics of shared bicycles and the research in the existing literature, the urban built environment influence factors are divided into three aspects: building function, public transportation convenience, and road network conditions. They are characterized by the POI comprehensive index, intensity public transport coverage, and spatial accessibility.

3.2. Assign Values to the Network Grid

In order to analyze the impact of public transport and road network coverage, the paper uses the Fishnet function of ArcGIS software to divide the study area into 1297 grids of 500 m × 500 m. By assigning built environment factors to grid cells, we can analyze the relationship between the urban built environment and the spatial distribution of shared bicycles on a small-scale space.

This paper innovatively proposes three built environment indicators, including the POI comprehensive index, the intensity of public transport coverage, and spatial accessibility. Then, their characterization methods and assignment methods on the grid. Among them, the POI comprehensive index takes into account two aspects of POI density and land-use diversity. Firstly, the POI density indexes that have no significant impact on shared bicycles

are screened and excluded by SPSS, and then the comprehensive impact of various POI indexes is calculated by the entropy weight method. The intensity of public transport coverage is used to express the impact of public transport on shared bicycles. The research object of this paper is the dockless shared bicycle, and the research area is divided into a grid of 500 m × 500 m. The impact of traffic stations in the adjacent grid may be ignored by simply using indicators such as station density. Therefore, the index of public transport coverage intensity is proposed. The spatial accessibility index is used to express the impact of urban spatial structure on the distribution of shared bicycles. Previous studies mostly used road network density, non-motor vehicle density, and other indicators. This paper creatively introduces the concept of Spatial Syntax and uses the indicators of Spatial Syntax to quantify spatial accessibility.

3.2.1. POI Composite Index

Point of Interest (POI) data, representing building functions and geographic location information, can characterize the spatial distribution of urban land use to a certain extent. Eight categories of POI data closely related to residents' daily life were selected, and the Pearson correlation analysis method in IBM SPSS Statistics 24.0 software was used to explore the relationship between these categories and the distribution of shared bicycles. The results are shown in Table 1.

Table 1. Analysis of the correlation between the POIs and the distribution of shared bicycles.

POI	Correlation	Unstandardized Coefficient	Standardized Coefficient	Significance
Accommodation service	0.606	19.181	0.074	0.002
Medical institutions	0.41	6.609	0.054	0.007
Sports and leisure	0.659	0.151	0.001	0.973
Residential	0.644	17.436	0.262	0
Financial services	0.647	11.02	0.103	0
Educational institution	0.572	8.511	0.117	0
Public utilities	0.588	22.792	0.087	0
Catering services	0.736	7.17	0.32	0

As the result of data analysis shows, in addition to medical institutions, the other seven types of POI data have a significant correlation with the shared bicycle borrowing and returning quantity. The correlation value is relatively close, generally between 0.6 and 0.8. This paper introduces the POI comprehensive index to reflect the comprehensive impact of multiple POI data. It can reduce the data analysis dimension and be obtained by weighted summation of the above seven types of POI, The calculation method is:

$$A = \sum_{i=1}^n w_i x_i \quad (1)$$

where x_i is the number of POI i , and w_i is the weight of the POI. Its value is determined by the entropy method, and the specific calculation method is as follows:

Standardize POI data:

$$x'_i = \frac{x_i - x_i^{\min}}{x_i^{\max} - x_i^{\min}} \quad (2)$$

where x'_i is the standardized treatment result of POI i .

Calculate the weight:

$$\Phi_i = \frac{d_j}{\sum_{j=1} d_j} \quad (3)$$

where Φ_i is the weight coefficient of index i , and d_j is the information utility, which can be calculated based on information entropy E_j :

$$d_j = 1 - E_j \quad (4)$$

$$E_j = -\frac{\sum_{i=1}^n r_{ij} \ln(r_{ij})}{\ln n} \quad (5)$$

where E_j is the information entropy of group j , and r_{ij} is the proportion of the sample value i in index j .

The weight calculation results of POI are shown in Table 2. Since the POI data contain geographic information, the weighted POI comprehensive value can be assigned directly to the grid cell.

Table 2. Summary of weight calculation results of entropy method.

POI	Information Entropy	Information Utility	Weight Coefficient
Accommodation service	0.8848	0.1152	15.12%
Medical institutions	0.8717	0.1283	16.84%
Residential	0.9251	0.0749	9.83%
Financial services	0.8702	0.1298	17.03%
Educational institution	0.8942	0.1058	13.88%
Public utilities	0.8962	0.1038	13.62%
Catering services	0.8957	0.1043	13.69%

Through Pearson analysis in SPSS, this paper analyzes the correlation between the POI composite index and the number of shared bicycles. The result shows that the correlation between them is 0.603, and that the intensity of public transport coverage can explain the distribution of shared bicycles.

3.2.2. Intensity of Public Transport Coverage

According to relevant surveys, 69% of shared bicycle trips in Shanghai are designed to connect with public transport. Therefore, the intensity of public transport coverage has an important impact on the distribution of shared bicycles. Since the grid cell scale is small, only the grid cell at the station's location will be assigned by simply using the number of public transport stations, which is inconsistent with the actual situation. The parameter of public transportation coverage intensity is introduced in the study to characterize the spatial grid unit's convenience of public transportation connection.

$$A_{pt(i)} = \sum Q_{ij}$$

$$Q_{ij} = \begin{cases} 1, & \text{Grid } i \text{ is in the service coverage of site } j \\ 0, & \text{Grid } i \text{ is out of the service coverage of site } j \end{cases} \quad (6)$$

where $A_{pt(i)}$ is the intensity of public transport coverage of grid i . According to the coverage rate of the public transport network in downtown Shanghai, the service radius of the conventional bus station is set to 300 m, and the service radius of the rail transit station is set to 800 m. Calculate the affected area of the bus stations and rail transit stations in the unit grid, and record them as the intensity of bus coverage and the intensity of rail transit coverage, respectively. Considering the difference in service level and the attraction of passenger flow between rail transit and conventional buses, the conversion coefficient of rail transit stations is introduced. Take the ratio of the transportation capacity of rail transit and conventional bus per unit time as the conversion factor, record it as μ . Then is the following is obtained:

$$A_{pt(i)} = \mu \sum Q_{ij}R_j + \sum Q_{ij}B_j \quad (7)$$

where R_j and B_j , respectively, indicate the coverage of rail transit stations and bus stations. Figure 4 shows the assignment results of the public transportation coverage intensity of the grid network in downtown Shanghai. The data in picture (c) are the sum of pictures (a) and (b).

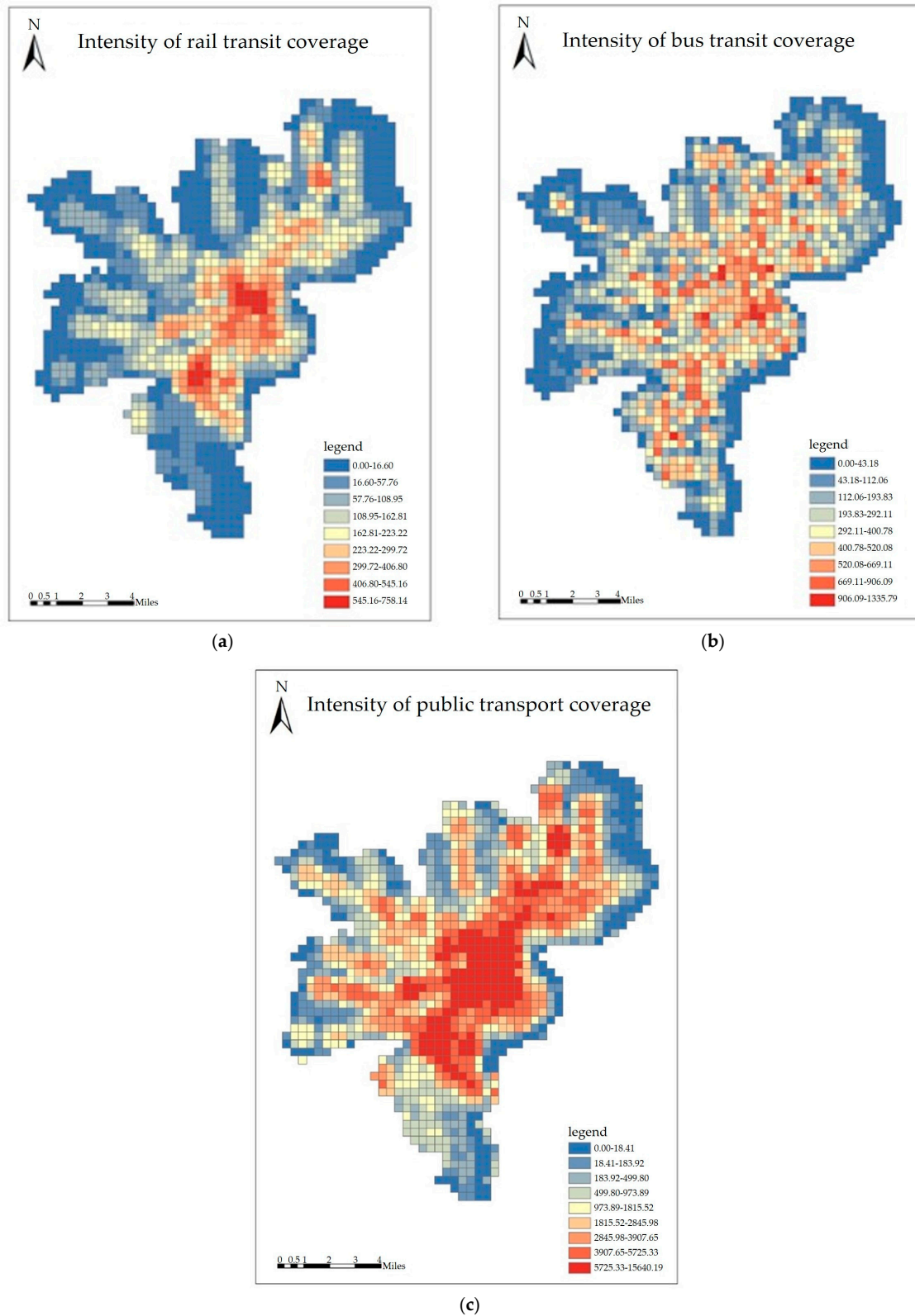


Figure 4. The intensity of grid network public transportation coverage.

Through Pearson analysis in SPSS, this paper analyzes the correlation between the intensity of public transport coverage and the number of shared bicycles. The result shows that the correlation between them is 0.786, and that the intensity of public transport coverage can explain the distribution of shared bicycles.

3.2.3. Spatial Accessibility

The influence of road network conditions on the distribution of shared bicycles essentially reflects spatial accessibility on travel choices. This paper uses the index of Space Syntax to reflect spatial accessibility. Space Syntax is a theory and method of urban spatial analysis proposed by Bill Hillier in the 1980s [28]. It essentially divides space while following the principle of graph theory, specifically abstract urban space, into intersecting line segments. Formulas have been put forward to quantify the topological relationship between line segments, explore the internal laws of spatial structure, and reveal the relationship between human activities and spatial structure [29].

The analysis variables of the Spatial Syntax model include:

1. Depth: This refers to the minimum number of steps required for a space to reach other spaces and expresses the accessibility of the space in a topological sense. The higher the depth value, the worse the accessibility, and the lower the intensity of human activity. The mean depth (MD) is the mean of the minimum steps from a node to all other nodes in the system. It is a description of the whole system. The calculation formula is:

$$MD_i = \frac{\sum_{j=1}^n d_{ij}}{n-1} \quad (8)$$

where d_{ij} is the minimum number of topology steps between the nodes.

2. Integration: This indicates the degree of aggregation or dispersion between nodes and other nodes in the entire system. If the integration value is greater than 1, the cell space has a high degree of aggregation with all other spaces in the system. If it is less than 1, it indicates that the nodes show a trend of mutual dispersion. When the value is between 0.4 and 0.6, the layout of spatial objects is relatively scattered. The calculation formula is as follows:

$$I_i = \frac{n(\log_2(\frac{n+2}{3} - 1)) + 1}{(n-1)(MD_i - 1)} \quad (9)$$

3. Choice: This indicates the frequency of the shortest topological distance between nodes of the spatial unit in the system, and measures the advantages of the spatial unit as the shortest travel path, which reflects the possibility of the space being traversed. The node with a higher choice degree is more likely to be traversed by people. The calculation formula is as follows:

$$C_i = \sum x_i \quad (10)$$

where $\sum x_i$ represents the sum of the spatial unit's times in the shortest topological distance between the two nodes.

Referring to the road network data of downtown Shanghai in 2020, the road network axis map is drawn in ArcGIS software 10.3 based on the principle of "longest and least" and imported into DepthMap X0.7.0 for data analysis to obtain the variable values of Spatial Syntax. Finally, the obtained road network index values are fed back to ArcGIS software 10.3 and assigned to the road network axis. The calculation method of spatial accessibility of unit grid cells is as follows:

$$A_{road(i)} = \frac{\sum AR_{j \times L_j}}{\sum L_j} \quad (11)$$

where $A_{road(i)}$ is the accessibility value of grid i , AR_j is the accessibility value of the road axis j in the grid, and L_j is the length of the road axis j .

In Spatial Syntax, the n -step topological distance and 3-step topological distance are representative, so they are selected for analysis. Through the Pearson correlation analysis in the SPSS, the correlation between the Spatial Syntax index and the spatial distribution of the shared bicycles is shown in Table 3. The integration (Rn) has the highest correlation with the shared bicycles, so the integration (Rn) is selected to reflect the spatial accessibility. After drawing the road network axis map of downtown Shanghai, 26,926 road axes were obtained. The result is shown in Figure 5.

Table 3. Correlation between Spatial Syntactic variables and shared bicycles.

Index	Integration (Rn)	Integration (R3)	Mean Depth	Choice (Rn)	Choice (R3)
Correlation	0.527 **	0.492 **	0.070 *	0.229 **	0.309 **

* $p < 0.05$, ** $p < 0.01$.

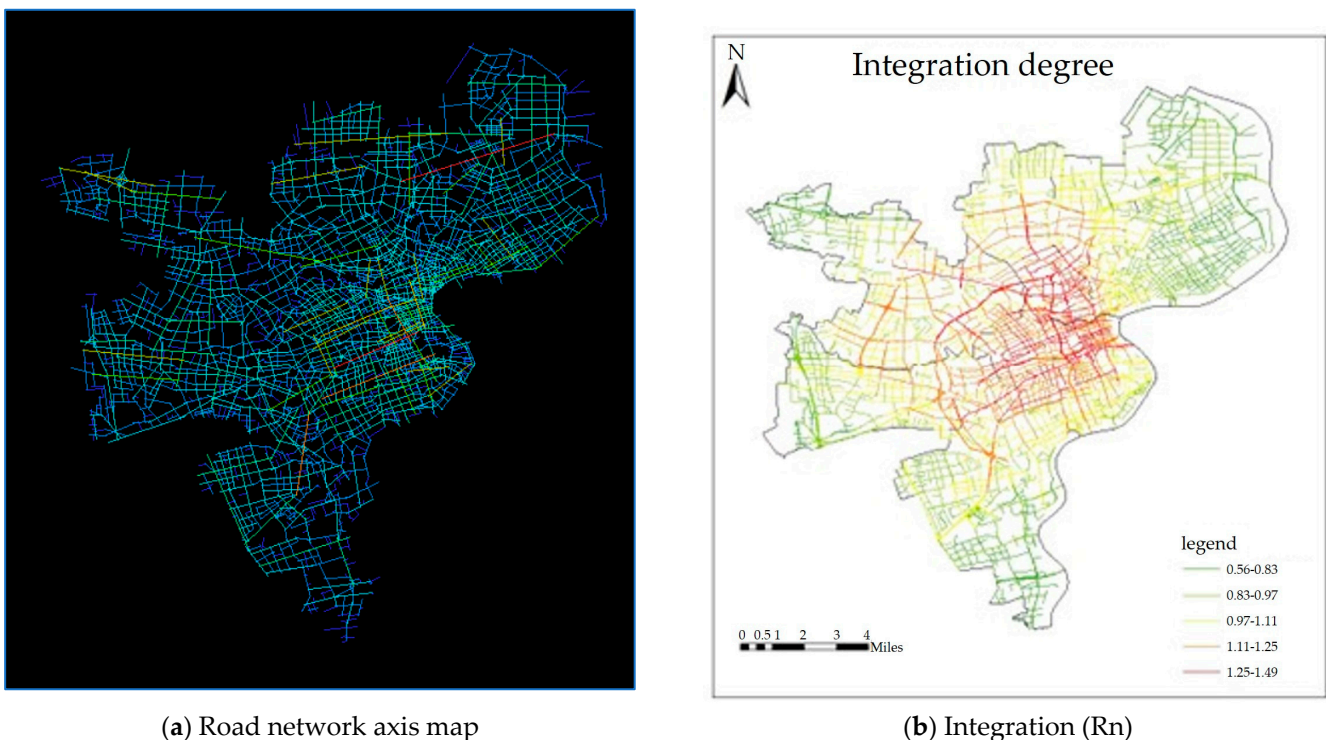


Figure 5. Analysis results of spatial accessibility in downtown Shanghai.

The integration (Rn) of the road network in downtown Shanghai presents the characteristics of the single-center structure. The core area is the junction of Jingan District, Hongkou District, and Huangpu District. The roads with high integration degrees are concentrated near the bund on the west side of the Huangpu River, forming the accessibility center of the road network in downtown Shanghai.

3.3. Support Vector Regression Model

At present, the regression prediction methods of shared bicycle demand mainly include multivariable linear regression model, geographically weighted regression, and regression prediction models based on machine learning, such as support vector machine, random forest, and SP neural network. Because there is a nonlinear relationship between the spatial distribution of shared bicycles and their influencing factors, and the distribution of shared bicycles changes periodically, this paper attempts to use the SVR model to

establish the relationship between the spatial distribution of shared bicycles and the urban built environment. As an efficient machine learning method, SVR can take into account the complexity and learning ability of the model and minimize empirical errors. It has advantages in small-sample data and nonlinear conditions [30].

The idea of SVR is to map the data to the high-dimensional feature space through nonlinear mapping and to carry out linear regression in the high-dimensional feature space to achieve the effect of linear regression in the original space. The regression estimation function is:

$$f(x) = \omega \varphi(x) + b \quad (12)$$

where ω is the weight vector, and b is a constant.

Suppose x_i is the urban built-up environmental factor affecting the distribution of shared bicycles, y_i is the number of shared bicycles, and set the training dataset $\{x_i, y_i\}, i = 1, 2, \dots, n$. The insensitive loss function is introduced into the constrained optimization problem, which can be expressed as:

$$\begin{aligned} \min_{\omega} & \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{s.t.} & \begin{cases} y_i - \omega \varphi(x) - b \leq \varepsilon + \xi_i^* \\ \omega \varphi(x) + b - y_i \leq \varepsilon + \xi_i \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned} \quad (13)$$

where $\frac{1}{2} \|\omega\|^2$ is the regularization part, $\sum_{i=1}^n (\xi_i + \xi_i^*)$ is the experience risk, C is the penalty factor, and ξ_i^* is the slack variable.

We can transform the optimization problem of the above formula into a dual problem by using the Lagrange multiplier method, then introduce the kernel function $k(x_i, x_j)$, from which we obtain:

$$\begin{aligned} \min & \left\{ \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)k(x_i, x_j) - \sum_{i=1}^n (\alpha_i^* - \alpha_i)y_i + \sum_{i=1}^n (\alpha_i^* + \alpha_i)\varepsilon \right\} \\ \text{s.t.} & \sum_{i=1}^n (\alpha_i^* - \alpha_i) = 0, \quad \alpha_i^* \leq C, 0 \leq \alpha_i \end{aligned} \quad (14)$$

where $\alpha_i, \alpha_i^*, \beta_i,$ and β_i^* are Lagrange multipliers. α_i and α_i^* can be obtained by solving the quadratic programming problem.

The final regression function expression is:

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \alpha_i)k(x_i, x) + b \quad (15)$$

In solving the problem of SVR, the selection of kernel function is the key problem. Generally, there are three types of kernel functions: (a) Linear kernel function: $k(x_i, x) = x_i \bullet x$; (b) Polynomial kernel function: $k(x_i, x) = [\langle x_i, x \rangle + 1]^q$; and (c) Gaussian radial basis kernel function: $k(x_i, x) = \exp\left(-\frac{\|x_i - x\|^2}{\sigma^2}\right)$.

This paper sets the grid as the analysis cell. Input the POI comprehensive index, the intensity of public transport coverage, spatial accessibility, and shared bicycle data during peak hours (7:30 a.m. to 8:30 a.m. and 17:30 p.m. to 18:30 p.m.) in downtown Shanghai in a week. The data within Xinhua Road, Changning District, Shanghai, were selected as the prediction sample set. The rest of the data were selected as the training sample set. In this paper, the root mean square error (RMSE) and the coefficient of determination (R^2) are used to test the prediction effect. The calculation formulas are as follows:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^* - y_i)^2} \quad (16)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i^* - y_i)^2}{\sum_{i=1}^n (\bar{y}_i - y_i)^2} \quad (17)$$

where y_i^* is the forecast value of shared bicycle demand, y_i is the measured value of shared bicycle, \bar{y}_i is the sample mean of shared bicycle demand, and n is the number of samples in the prediction sample set.

The multiple linear regression method is introduced as a comparative experiment to compare and analyze the advantages and disadvantages of SVR and traditional linear regression methods. The calculation formula is as follows:

$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m \quad (18)$$

where Y_i is the number of shared bicycles, x_1, x_2, \dots, x_m is the influencing factors of sharing bicycles, β_0 is a constant, and $\beta_1, \beta_2, \dots, \beta_m$ is the regression coefficient.

4. Results

After assigning the research data to the grid, the correlation analysis results show that the POI comprehensive index, the intensity of public transport coverage, and spatial accessibility significantly impact the spatial distribution of shared bicycles. The Linear kernel function, Gaussian radial basis function, and Polynomial kernel function are used to construct the SVR model. The forecast accuracies of the SVR models under these three kernel functions are obtained by training these models, as shown in Figures 6–8.

The root means square error and coefficient of determination of the SVR models based on the three kernel functions and the Multivariable linear regression model are shown in Table 4.

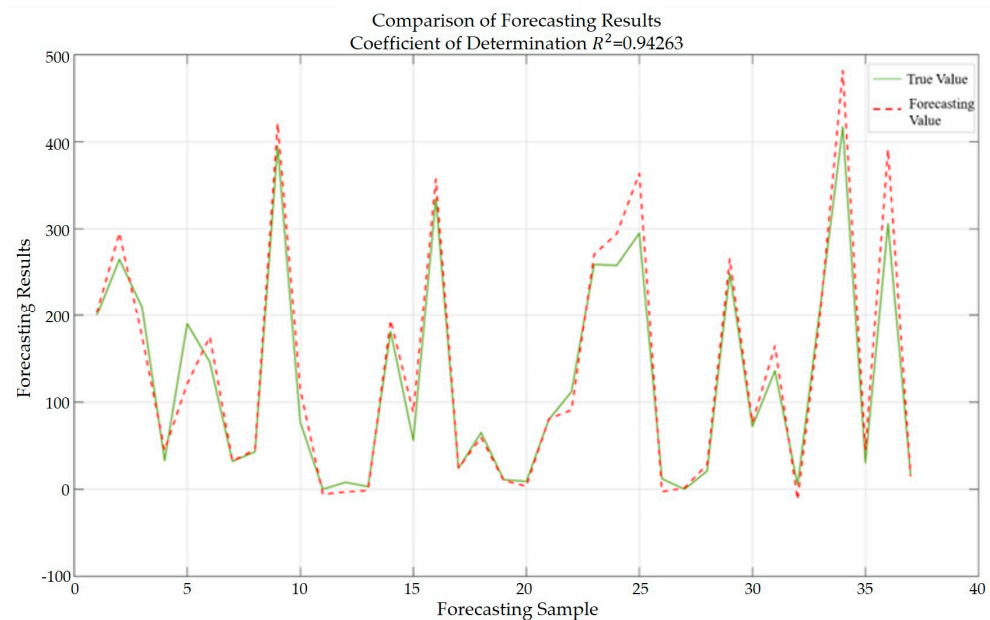


Figure 6. Forecast accuracy of Linear kernel function model.

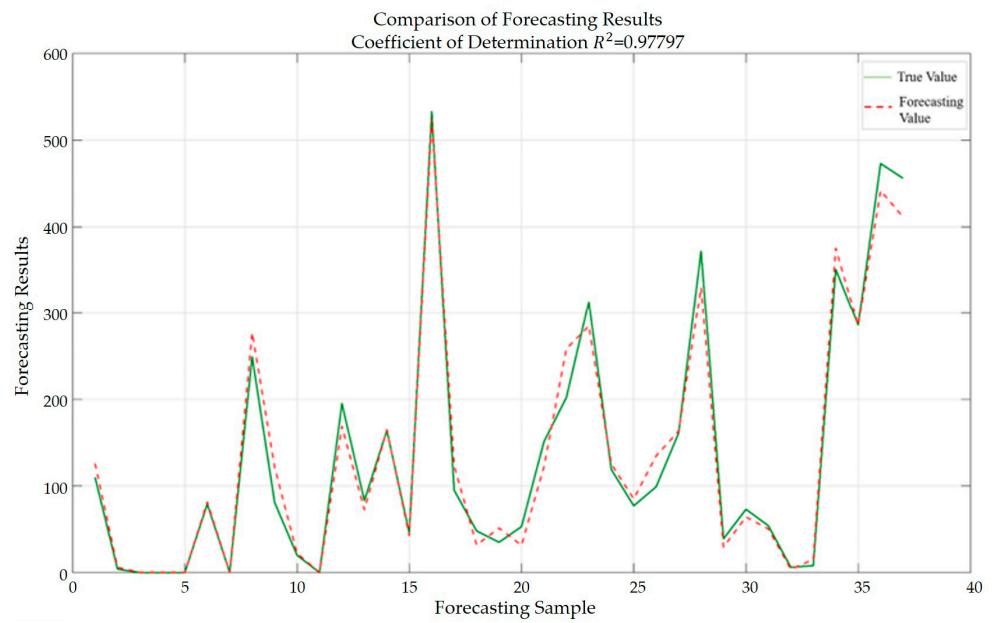


Figure 7. Forecast accuracy of Gaussian radial basis kernel function.

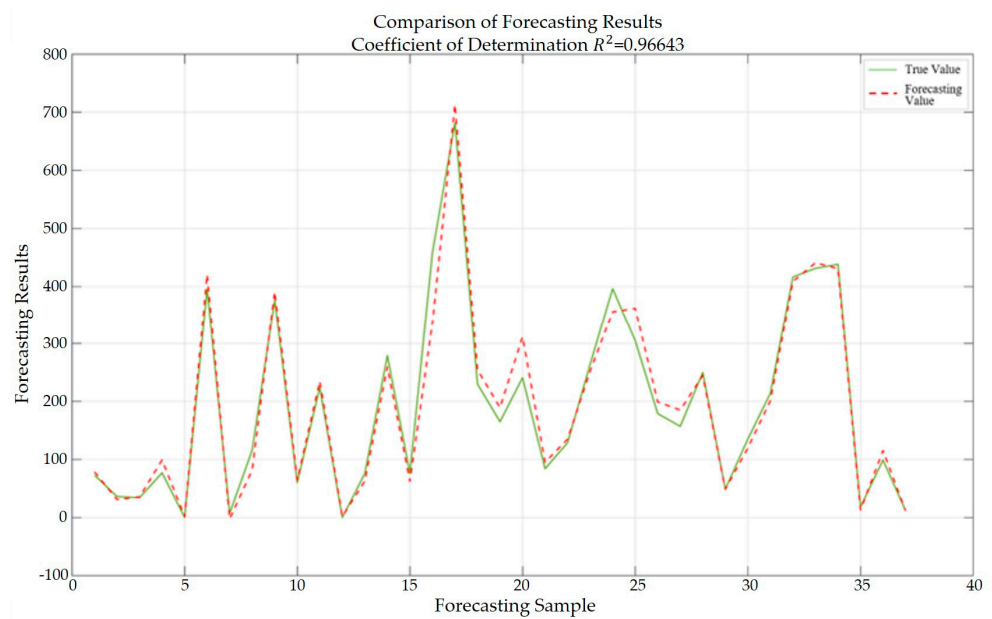


Figure 8. Forecast accuracy of Polynomial kernel function.

Table 4. Predictive effects of each model.

Model	RMSE	R ²
Linear kernel function	27.2999	0.94263
Gaussian radial basis kernel function	14.3416	0.97797
Polynomial kernel function	27.2912	0.96643
Multivariable linear regression model	67.3542	0.847

As shown in the figures, the X axis stands for forecasting sample, that is, the grid cell serial number, while the Y axis stands for forecasting results, that is, the number of shared bicycles in each grid cell. The results show that the SVR models based on these three kernel functions have achieved high forecast accuracy. The model based on the Gaussian radial basis function has the minimum errors and the best fitting effect. The

RMSE value measured by the model is 14.3416, which means that the average error between the predicted quantity of shared bicycles and the true value is about 14 vehicles in a unit fishing net of 500 m × 500 m. The value of R^2 is 0.97797, indicating that the prediction model can explain 97.797% of the distribution of shared bicycles. The RMSE obtained by the multiple linear regression model is 67.3542, and R^2 is 0.847, which is not as good as the former. The study found that the shared bicycle SVR model based on the Gaussian radial basis function can better explain the impact of urban built environment factors on the spatial distribution of shared bicycles.

5. Discussion and Conclusions

5.1. Discussion

This study provides a new perspective for investigating the impacts of the built environment on shared bicycles in a small-scale space, taking Shanghai as a case study. The main contribution of this study is threefold. First, the Fishnet function of ArcGIS was utilized to divide the study area into grids of 500 m × 500 m, and a method to calculate the borrowing and returning position of shared bicycles was proposed based on time-varying GPS data. Second, the built environmental factors are represented by building function, public transportation convenience, and road network conditions, and then the conditions are assigned to the grid. Third, SVR models are applied to explore the nonlinear relationship between the usage of shared bicycles and contributing factors.

The results show that the POI comprehensive index, the intensity of public transport coverage, and spatial accessibility significantly affect the spatial distribution of shared bicycles. The SVR model based on the Gaussian radial basis function is more effective in unveiling the complex and nonlinear relationship between shared bicycle usage and built environment elements than traditional linear models. In addition, the study finds that environmental factors, such as financial institutions and residential, have a more significant impact on shared bicycles than others.

5.2. Conclusions

Although this paper analyzes the nonlinear relationship between urban built environment factors and shared bicycle distribution in small-scale space, there are still some limitations. First, Shanghai was taken as the case study; therefore, the conclusion may not be universal. In the future, this research method can be extended to carry out case studies in different cities. Second, this paper mainly considers the impact of urban built environment on shared bicycles. Although this paper comes to the conclusion that environmental factors such as financial institutions and residential have a more significant impact on shared bicycles than other factors, considering that the population density of such POI areas is high, the impact of population density of various POIs may also play an important role in the demand and supply of shared bicycles. The influence of individual characteristics and sociodemographic attributes of residents may be ignored.

The spatial distribution law of shared bicycles has practical significance in understanding the supply and demand of shared bicycles in different built environment areas. In the next stage, the site selection of shared bicycle stations can be studied on this basis.

Author Contributions: Conceptualization, L.Y. and S.F.; methodology, L.Y. and S.F.; software, L.Y., S.F. and L.W.; validation, H.J. and J.Q.; formal analysis, S.F., L.Y. and X.H.; investigation, L.W. and X.H.; resources, L.W. and J.Q.; data curation, L.Y., H.J. and S.F.; writing—original draft preparation, S.F. and L.Y.; writing—review and editing, L.Y. and H.J.; visualization, L.Y., S.F. and X.H.; supervision, J.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Fundamental Research Funds for the Central Universities, Jilin University, 2022-JCXK-01.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: Thank you to all reviewers for their suggestions to make this paper better.

Conflicts of Interest: We declare that we have no conflict of interest.

References

1. Xiao, G.; Lu, Q.; Ni, A.; Zhang, C. Research on carbon emissions of public bikes based on the life cycle theory. *Transp. Lett.* **2023**, *15*, 278–295.
2. Koohsari, M.; Cole, R.; Oka, K.; Shibata, A.; Yasunaga, A.; Hanibuchi, T.; Owen, N.; Sugiyama, T. Associations of built environment attributes with bicycle use for transport. *Environ. Plan. B Urban Anal. City Sci.* **2020**, *47*, 1745–1757. [[CrossRef](#)]
3. Xu, C.; Ji, J.; Liu, P. The Station-Free Sharing Bike Demand Forecasting with a Deep Learning Approach and Large-Scale Datasets. *Transp. Res. Part C Emerg. Technol.* **2018**, *95*, 47–60. [[CrossRef](#)]
4. Xiao, G.; Xiao, Y.; Ni, A.; Zhang, C.; Zong, F. Exploring influence mechanism of bikesharing on the use of public transportation—A case of Shanghai. *Transp. Lett.* **2023**, *15*, 269–277. [[CrossRef](#)]
5. Sohrabi, S.; Ermagun, A. Dynamic Bike Sharing Traffic Prediction Using Spatiotemporal Pattern Detection. *Transp. Res. Part Transp. Environ.* **2021**, *90*, 102647. [[CrossRef](#)]
6. Caggiani, L.; Colovic, A.; Ottomanelli, M. An Equality-Based Model for Bike-Sharing Stations Location in Bicycle-Public Transport Multimodal Mobility. *Transp. Res. Part Policy Pract.* **2020**, *140*, 251–265. [[CrossRef](#)]
7. Song, M.; Wang, K.; Zhang, Y.; Li, M.; Qi, H.; Zhang, Y. Impact Evaluation of Bike-Sharing on Bicycling Accessibility. *Sustainability* **2020**, *12*, 6124. [[CrossRef](#)]
8. Zhao, X.; Zhang, J.; Zhang, N.; Wong, Y.D.; Zhou, Y.; Meng, M. A GIS-CA Model for Planning Bikeways upon the Footpath Network. *Sustainability* **2021**, *13*, 8926. [[CrossRef](#)]
9. Nickkar, A.; Banerjee, S.; Chavis, C.; Bhuyan, I.A.; Barnes, P. A Spatial-Temporal Gender and Land Use Analysis of Bikeshare Ridership: The Case Study of Baltimore City. *City Cult. Soc.* **2019**, *18*, 100291. [[CrossRef](#)]
10. Li, H.; Zhang, Y.; Ding, H.; Ren, G. Effects of dockless bike-sharing systems on usage of London Cycle Hire. *Transp. Res. Part Policy Pract.* **2019**, *130*, 398–411. [[CrossRef](#)]
11. Tang, Y.; Pan, H.; Fei, Y. Research on users' frequency of ride in Shanghai Minhang bike-sharing system. *Transp. Res. Procedia.* **2017**, *25*, 4983–4991. [[CrossRef](#)]
12. Wang, K.; Chen, Y. Joint analysis of the impacts of built environment on bikeshare station capacity and trip attractions. *J. Transp. Geogr.* **2020**, *82*, 102630. [[CrossRef](#)]
13. Weliwitiya, H.; Rose, G.; Johnson, M. Bicycle Train Intermodality: Effects of Demography, Station Characteristics and the Built Environment. *J. Transp. Geogr.* **2019**, *74*, 395–404. [[CrossRef](#)]
14. Noland, R.; Smart, M.; Guo, Z. Bikesharing trip patterns in New York City: Associations with land use, subways, and bicycle lanes. *Int. J. Sustain. Transp.* **2018**, *13*, 664–674. [[CrossRef](#)]
15. Maas, S.; Attard, M.; Caruana, M. Assessing spatial and social dimensions of shared bicycle use in a Southern European island context: The case of Las Palmas de Gran Canaria. *Transp. Res. Part A Policy Pract.* **2020**, *140*, 81–97. [[CrossRef](#)]
16. Guo, Y.; He, S. Built environment effects on the intergration of dockless bike-sharing and the metro. *Transp. Res. Part D* **2020**, *83*, 102335. [[CrossRef](#)]
17. El-Assi, W.; Salah Mahmoud, M.; Nurul Habib, K. Effects of Built Environment and Weather on Bike Sharing Demand: A Station Level Analysis of Commercial Bike Sharing in Toronto. *Transportation* **2017**, *44*, 589–613. [[CrossRef](#)]
18. Wang, K.; Akar, G. Gender gap generators for bike share ridership: Evidence from Citi Bike system in New York City. *J. Transp. Geogr.* **2019**, *76*, 1–9. [[CrossRef](#)]
19. Luo, S.; Zhen, F.; Yin, Q. How Built Environment Influence Public Bicycle Usage: Evidence from the Bicycle Sharing System in Qiaobei Area, Nanjing. *Sci. Geogr. Sin.* **2018**, *38*, 332–341. [[CrossRef](#)]
20. Faghih-Imani, A.; Eluru, N.; El-Geneidy, A.M.; Rabbat, M.; Haq, U. How Land-Use and Urban Form Impact Bicycle Flows: Evidence from the Bicycle-Sharing System (BIXI) in Montreal. *J. Transp. Geogr.* **2014**, *41*, 306–314. [[CrossRef](#)]
21. Scott, D.; Ciuro, C. What factors influence bike share ridership? An investigation of Hamilton, Ontario's bike share hubs. *Travel Behav. Soc.* **2019**, *16*, 50–58. [[CrossRef](#)]
22. Xing, Y.; Wang, K.; Lu, J. Exploring travel patterns and trip purposes of dockless bike-sharing by analyzing massive bike-sharing data in Shanghai, China. *J. Transp. Geogr.* **2020**, *87*, 102787. [[CrossRef](#)]
23. Chen, E.; Ye, Z. Identifying the Nonlinear Relationship between Free-Floating Bike Sharing Usage and Built Environment. *J. Clean. Prod.* **2021**, *280*, 124281. [[CrossRef](#)]
24. Ye, M.; Chen, Y.; Yang, G.; Wang, B.; Hu, Q. Mixed Logit Models for Travelers' Mode Shifting Considering Bike-Sharing. *Sustainability* **2020**, *12*, 2081. [[CrossRef](#)]
25. Mateo-Babiano, I.; Bean, R.; Corcoran, J.; Pojani, D. How Does Our Natural and Built Environment Affect the Use of Bicycle Sharing? *Transp. Res. Part Policy Pract.* **2016**, *94*, 295–307. [[CrossRef](#)]
26. Zhang, Y.; Thomas, T.; Brussel, M.; van Maarseveen, M. Exploring the Impact of Built Environment Factors on the Use of Public Bikes at Bike Stations: Case Study in Zhongshan, China. *J. Transp. Geogr.* **2017**, *58*, 59–70. [[CrossRef](#)]

27. Faghih-Imani, A.; Eluru, N. Incorporating the Impact of Spatio-Temporal Interactions on Bicycle Sharing System Demand: A Case Study of New York CitiBike System. *J. Transp. Geogr.* **2016**, *54*, 218–227. [[CrossRef](#)]
28. Hillier, B.; Hanson, J. *The Social Logic of Space*; Cambridge University Press: Cambridge, UK, 1989; pp. 30–56.
29. Kwan, M. Space-time and Integral Measures of Individual Accessibility: A Comparative Analysis Using a Point-based Framework. *Geogr. Anal.* **1998**, *30*, 191–216. [[CrossRef](#)]
30. Qin, W.; Wang, L.; Liu, Y.; Xu, C. Energy Consumption Estimation of the Electric Bus Based on Grey Wolf Optimization Algorithm and Support Vector Machine Regression. *Sustainability* **2021**, *13*, 4689. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.