buildings

MDPI

Article

Spatial Cluster Pattern and Influencing Factors of the Housing Market: An Empirical Study from the Chinese City of Shanghai

Yuhua Zhang [1,*] and Boyana Buyuklieva [2]

[1] The Bartlett Centre for Advanced Spatial Analysis, Faculty of the Built Environment, University College London (UCL), London W1T4TJ, UK
[2] Department of Information Studies, University College London (UCL), London W1T4TJ, UK; boyana.buyuklieva@ucl.ac.uk
* Correspondence: yuhua.zhang.21@alumni.ucl.ac.uk

Abstract: Infrastructure and amenities have an evident effect on differentiated urban structures and house prices. However, few studies have taken into account the spatial heterogeneity of large-scale urban areas. Regarding this issue, the present study proposes a novel spatial framework to quantify the impacts of built environment factors on the housing market. We aim to answer: how does a specific factor impact house prices across different spatially autocorrelated neighbourhood clusters? The city of Shanghai, the economic centre of China, is examined through the transaction data from the China Real-estate Information Center (CRIC) are analysed. Firstly, spatially autocorrelation clusters were explored to identify high/low housing prices in concentrated areas in Shanghai. Secondly, using the development-scale house prices as the dependent variable, we employed ordinary least squares (OLS) linear regression and geographically weighted regression (GWR) models to examine the impact of built environment facilities on the house prices across these spatial autocorrelation regions and Shanghai more generally. The results suggest the following: (1) There are significant spatially autocorrelated clusters across Shanghai, with high-value clusters concentrated in the city core and low value concentrated in the suburban fringes; (2) Across Shanghai and its spatially autocorrelated clusters, transportation accessibility and service amenities factors can affect house prices quite differently, especially when focusing on the city centre and the suburban areas. Our results highlight the importance of optimising the city's polycentric structural framework to foster a more balanced regional development. Differentiated approaches to the distribution of public service facilities should be adopted to address the diverse needs of residents across various regions.

Keywords: spatial analysis; housing market; GIS; autocorrelated clusters; geographically weighted regression; influencing factor

check for updates

Academic Editor: Yung Yau

Received: 31 January 2025
Revised: 18 February 2025
Accepted: 18 February 2025
Published: 23 February 2025

Citation: Zhang, Y.; Buyuklieva, B. Spatial Cluster Pattern and Influencing Factors of the Housing Market: An Empirical Study from the Chinese City of Shanghai. *Buildings* **2025**, *15*, 708. https://doi.org/10.3390/buildings15050708

Copyright: © 2025 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

The marketisation of China's real estate industry began with the 1998 housing reform, the effects of which are remarkable [1,2]. From 1998 to 2018, the annual transaction volumes increased by 14.1 times, which led to China overtaking Germany to become the fourth-largest real estate market in the world in 2018, only being surpassed by the US, Japan, and the UK [3]. This marketisation of housing not only gives residents more residential options, but also, as the real estate industry is characterised by a long supply chain and many related sectors, it has a strong positive indirect effect on other industries [4,5]. The estimation shows that the housing market contributes more than 15% of the GDP of the Chinese economy.

One consequence of rapid urban development and investment is the inevitable rise in housing prices. For instance, the economic centre of China, Shanghai City, experienced a 270 percent increase in housing prices in the first decade of the twenty-first century [6]. By the end of 2021, the average annual house price in the most expensive neighbourhoods had already exceeded CNY 160,000 (GBP 15,000) per square metre; housing affordability has become an issue [7,8]. Since 2016, the authority has worked hard to avoid flipping property behaviours and high-leverage situations from both the supply and demand sides, aiming to reduce the investment characteristic of the real estate, and providing affordable housing to those who really need an apartment. However, due to the multifaceted shocks caused by the COVID-19 pandemic since 2019, various factors have caused changes in people's housing preferences [9]. In May 2022, the average house prices in China's 70 major cities declined by 0.1 per cent for the first time since September 2015. Evidence in multiple forms has led people to believe that the Chinese housing market is facing unprecedented challenges. Therefore, it is beneficial to investigate the latest housing market performance of a metropolis in China through the built environment perspective, as it will provide the authority with the most intuitive consumer information to identify the living habits and the most concerning issues of residents, which can be used to optimise the urban construction together whilst distinguishing the consumption and investment behaviours when setting up policies [10–12].

The influencing factors on housing prices have been widely investigated from two perspectives, namely national and city scales. Regarding the first aspect, these factors are commonly associated with fiscal, housing policies, and monetary [13,14]. For the housing prices at the level of the city, the influencing factors mainly include urban planning and development, population, housing supply, and demand [15–17]. In the past two decades, it was found that the combination of urban expansion issues and 'skyrocketing' house prices can cause significant financial crises [18]. In addition, social inequality problems are also one of potential consequences, such as social polarisation, residential segregation, and poor housing availability [19,20]. Therefore, optimising and adjusting urban development is a crucial subject for researchers, organisations, and government authorities. However, studies in this area have primarily concentrated on government policies, developer status and housing bubbles [21,22]. And, there is less research that has been performed on the spatial dynamics of the housing market, particularly the influence of public transportation and service amenities [23]. Therefore, it is imperative to examine the specific impact of urban infrastructures such as service amenities and public transportation on housing prices to provide dynamic insights into government policy. Moreover, in recent years, although many studies have focused on the city-level influencing factors in cities in China [24–26], none of them focused on the impact of those factors in spatially autocorrelated clusters. This study creatively uses a new database, namely the China Real Estate Information Corporation (CRIC) database, together with a discussion of the impact factors from a new perspective, with the intention of contributing to a better understanding of the spatial controlling mechanism of the housing market, and propose advice on the allocation of amenities in different areas.

Our study aims to address the central question as follows: how do built environment factors—such as proximity to service amenities and public transportation—influence the house price across spatially autocorrelated neighbourhood clusters?. The economic centre of China, the city of Shanghai, is selected as the study area. We followed mature methods adopted by some similar studies (e.g., [17]), but also proposed a creative structure for considering the spatial clusters. Specifically, our objectives include (i) determining whether there are clusters of spatially autocorrelated neighbourhoods in Shanghai and identifying those clusters. (ii) Using regression analysis to measure how the proximity to service

amenities and public transportation in the built environment can have spatially different impacts on house prices across Shanghai.

2. Materials

2.1. Study Area

Shanghai City is selected as the study area. It is located in the eastern part of China and regarded as the economic centre of the country. It covers an area of approximately 6340 km^2, with a total population of around 24.9 million, equivalent to a density of 3927 people/km^2. As indicated in Figure 1a, the building area is divided into 16 administrative districts, among which seven constitute so-called 'downtown Shanghai'. In addition, the authorised Central Business District 'Lujiazui' is in the Pudong New Area, so the very north part of the Pudong New Area is also considered part of the city centre. In this study, we further divide the city into 114 neighbourhoods at the neighbourhood's scale based on its real estate market divisions. Figure 1b shows the boundaries of these 114 neighbourhoods; since the names of the neighbourhoods are too difficult to display on a single map, only the neighbourhoods located in the city centre are labelled with names.

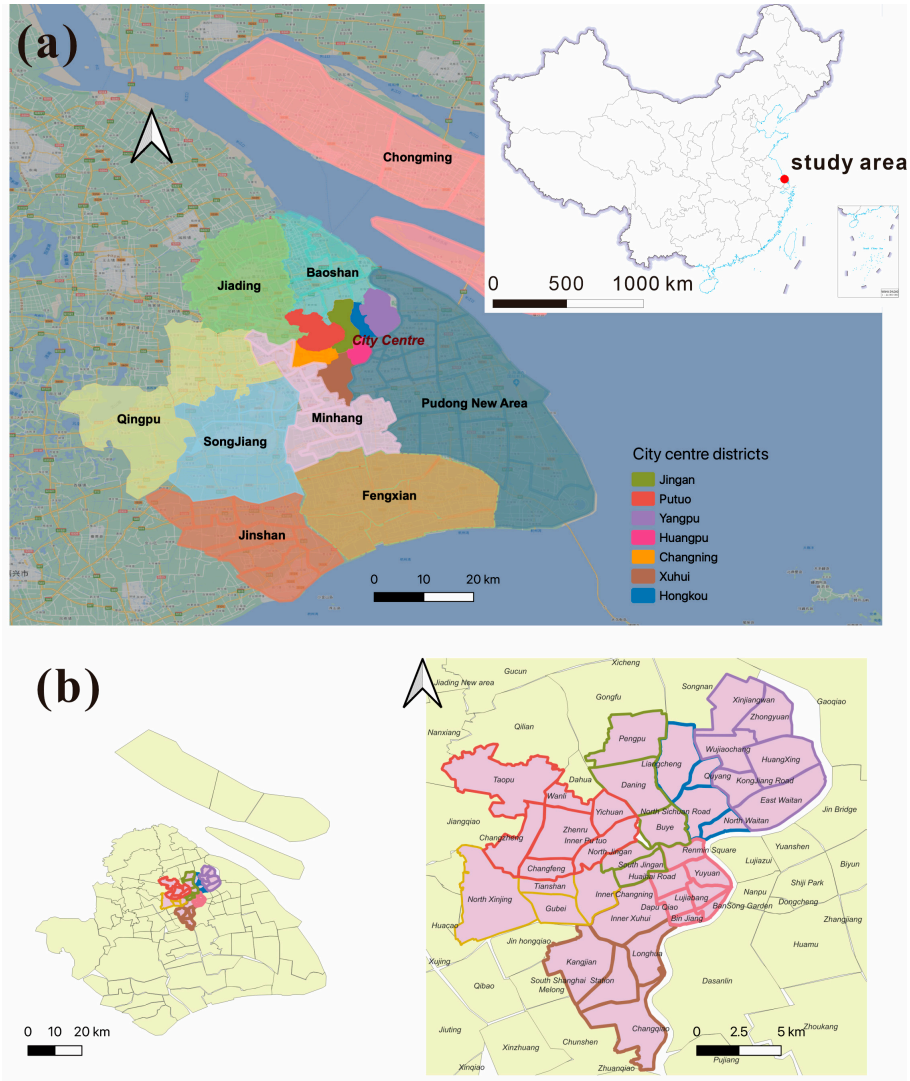


Figure 1. (**a**) Location and administrative map of the study area; and (**b**) Boundaries of the 114 neighbourhoods in Shanghai.

2.2. Data Sources

We study the Shanghai housing market with a particular emphasis on first-hand transactions by adopting two types of house price indicators from the China Real Estate Information Corporation (CRIC) database (https://vip.cric.com/gateway (accessed on 11 September 2024)). The CRIC (founded in 2006) is a premium Chinese database which provides comprehensive real estate-related information. CRIC has high credibility as the largest real estate information provider in China, serving more than 95% of the Top 100 real estate developers and some government sectors. There are several reasons to set newly built house prices as our major concern: First, a major portion of housing transactions concentrate in the newly built projects in Chinese housing markets currently, which is also an indispensable stage in other nascent housing markets' initial developments [27]. Second, one official house price index, so-called the "Average Selling Price of Newly-Built Residential Buildings" is regularly updated in China [28]. Hence, we can conclude that transaction prices of the newly built houses are rather important for Chinese housing market. Moreover, although the CRIC database contains information for both primary and second housing markets, a rather large part of data are collected from newly-built residential buildings from real estate developers. Moreover, the database only allows us to check the house prices of newly built houses for a specific project, and the data regarding second-hand house prices are considered insufficient.

Hence, in this study, we collected selling information about areas, prices, and developers of first-hand transactions. In addition, all other datasets that had been employed in the study are in Table 1.

Table 1. Data source summary used in this study.

Dataset	Description	Geographical Scale	Data Sources
CRIC Annual Average House Price Data	Annual average house price in 114 neighbourhoods	Polygons for Shanghai neighbourhoods	https://vip.cric.com/gateway (accessed on 11 September 2024)
CRIC New Development Transaction Price	The average transaction price for newly built developments including information about housing structures	Points data (SHP format)	
Gaode Map Data	Accessibility and public amenities data in Shanghai	Points data (SHP format)	https://lbs.amap.com/ (accessed on 11 September 2024)

2.2.1. CRIC Development Variables

In this study, two types of housing prices from the CRIC database are considered, namely the annual average house price (AAHP) and the new development transaction price (NDTP). The first type of dataset is mainly used for identifying the significant clusters of spatial autocorrelations in Shanghai. To better understand the patterns and developing trends of the housing market, the AAHP in recent years for all 114 neighbourhoods are taken into account.

The second dataset is used for stage II regression analysis, which records the total transactions in 787 newly built housing developments in 2021. There are two reasons for using this dataset. Firstly, it will give a larger sample size and more accurate built environmental information than those neighbourhood house prices. Secondly, the transaction price is more accurate than the listed price, given that listed prices are commonly highly overestimated. As the resale property prices in China are scattered across different real estate platforms and difficult to obtain [29–31], we choose the newly constructed (first-hand) developments information on CRIC database, which can overcome this issue. Table 2 integrates the field information that is used in this study.

Table 2. CRIC new developments transaction dataset overview.

Type of Field	Fields	Used for
General	Year	Demonstrate the time scale
Basic information	Development address', 'district', 'neighbourhoods'	Identify the geometry of those developments
	'Total dwelling number', 'total construction area', 'total construction/total land use ratio', 'parking plots ratio'	Independent variables
Sales data	'Total transactions supplied', 'average transaction price per unit' 'average transaction volume'	Independent variables; dependent variables (in bold)

2.2.2. Gaode Map

Despite the prices and structural information contained in the CRIC database, it does not provide the build environment factors. In our analysis, this type of information is collected from the Gaode Map Application Programming Interface (API). It is a Chinese mapping software which can provide locational information including name, description, and coordinates of observations. Specifically, the Gaode API v2022 provides the variables associated with accessibility to transportation and public amenities data all over Shanghai. Regarding the former type of data, the information on high-speed rail stations and underground stations is analysed. The public amenities include hospitals, primary schools, kindergartens, shopping centres, and parks. It should be noted that the education facilities include basic education (public schools) and premium education (top schools). Therefore, apart from kindergartens and primary schools, we also include 60 high-ranked middle schools in Shanghai as the indicator of public amenities. These middle schools are ranked by university acceptance rate and published by the Shanghai Education Bureau (https://edu.sh.gov.cn/ (accessed on 11 September 2024)). The locations are also retrieved from the Gaode map v2022. All the information mentioned above is provided as the point format in Gaode API v2022, the GIS platform is used to calculate the distance from them to houses and neighbourhoods.

3. Methodologies

3.1. Overview of the Analysis Strategy

The overall aim of this study is to address how service amenities and public transportation influence house prices across spatially autocorrelated neighbourhood clusters and Shanghai more generally. This question is divided and answered into two stages (Figure 2).

Stage I aims to determine whether there are spatial autocorrelation effects in the Shanghai housing market. We will use choropleth maps to examine the house price distribution of 114 neighbourhoods. Then, using the global Moran's I index to confirm whether spatial autocorrelation exists quantitatively. Since the index results indicated a statistically significant positive correlation, a local Moran's I's test is applied to find out which cluster areas have both high house prices on their own and high-value neighbours (high–high cluster), and by contrast, which areas are adjacent with lower prices (low–low).

The division of neighbourhood-scale data is limited by a relatively small sample size making it difficult to measure accessibility within the built environment. To address this problem, Stage II uses development-scale data to more accurately measure proximity to points of interest in the built environment and how these affect house prices in Shanghai. This is performed through a global regression and a geographically weighted regression

(GWR) model conducted on different variables. Both multicollinearities between variables and the global Moran's I index are tested as the prerequisites of GWR. Once the model of a house price, as measured in development transactions (discussed in the following sections), is fitted with 13 explanatory variables, the significant variables in the global regression are interpreted. The significance level set in this project is the same as in the literature [17]. In addition, it should be noted that, when interpreting the GWR coefficients, we bring in the idea of spatially clustered areas to investigate whether coefficients in those regions have significant patterns adding more nuance to the interpretation of the quantitative results. The analytical process in Stage I and Stage II is conducted in both Python 3.9 and QGIS 3.20. The Python notebook can be checked in GitHub repository (https://github.com/luciaaa77/Shanghai_analysis (accessed on 11 September 2024)).

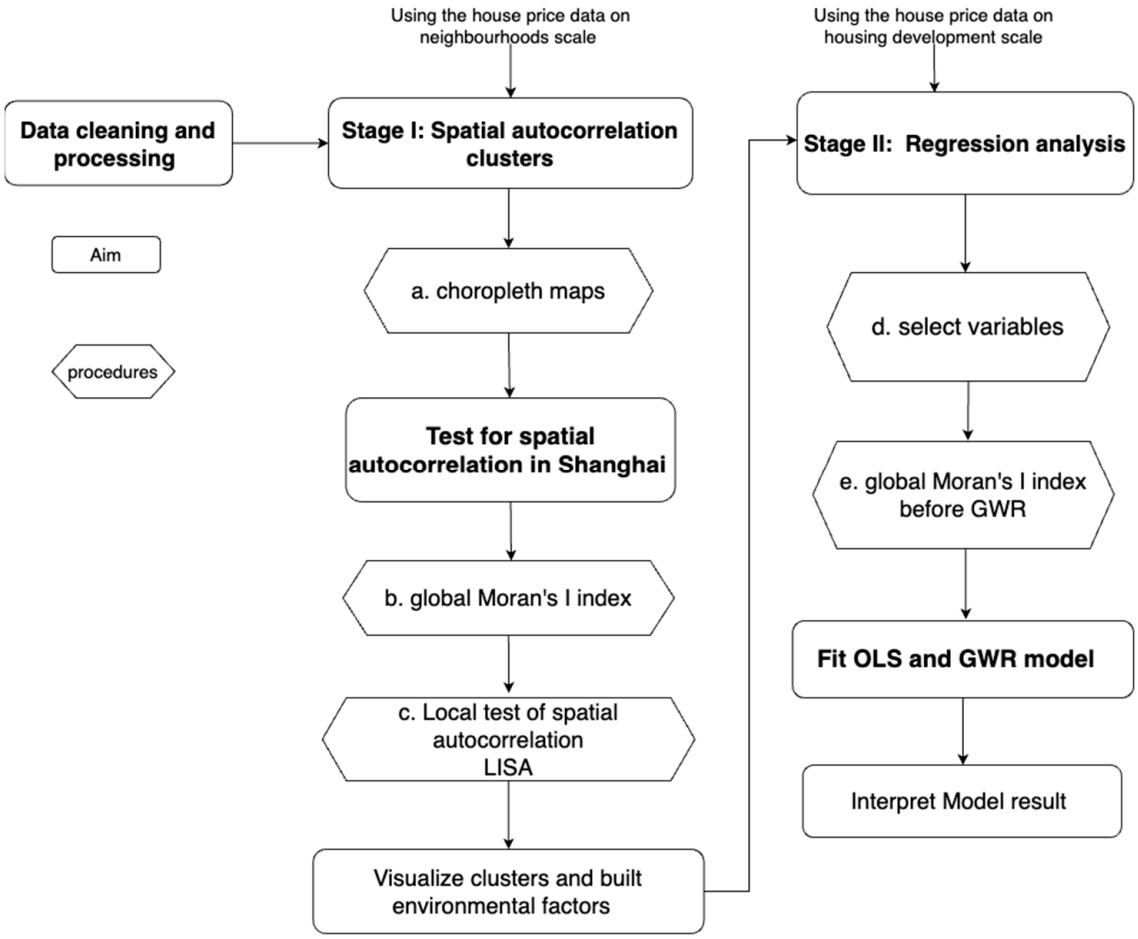


Figure 2. The methodology flow chart of this study.

3.2. Identification of the Spatial Clusters

3.2.1. Data Processing

Stage I uses the first dataset in CRIC, which contains the house price information on the neighbourhood scale. Table 3 shows the detailed information and summary statistics of the final dataset that has been employed for analysis, where (CNY) for Chinese Yuan (namely RMB), (GBP) for British Pounds, and the exchange rate is according to the latest exchange rate in 2024. The original data contains null values; with consideration, those are not dropped since they stand for areas with no transaction records throughout that year. At the end of stage I, we also employed the Gaode accessibility and services amenities data, including all transportation and public amenity locations mentioned in Session 2.2.

Table 3. Neighbourhood-scale house price.

Column Used	Average (CNY)	Average (GBP)	Standard Deviation (CNY)	Standard Deviation (GBP)
Unit house price_2021	40,164	4982	39,759	4932
Unit house price_2019	35,449	4397	27,521	3413
Unit house price_2017	31,783	3942	24,927	3091
Unit house price_2015	28,021	3476	20,467	2539

3.2.2. Moran's I Index

Choropleth maps are used to explore the general trend; however, there are cases where we think there is a spatial pattern, but in fact, what we are seeing is largely random [32]. To actually prove and measure the spatial pattern, spatial autocorrelation analysis can be introduced. Spatial autocorrelation describes the relationship between spatial variables and explains the distribution and relationship of spatial elements within a given environment [33]. As such, one of the most popular indexes of global autocorrelation analysis is the Moran's I index [34–37], which quantifies the degree of spatial autocorrelation by comparing the value of each feature with the values of neighbouring features. The formula of the global Moran's I index can be driven as the equation below:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} z_i z_j}{\sum_i z_i^2} \quad (1)$$

where I is the Moran's I index; n stands for the number of observations; z_i, z_j are the standardised value of the observation at location i, j; w_{ij} refers to the ith row and jth column of a W spatial weights matrix. A spatial weight matrix w is used to formulate the spatial information, which is necessary when testing for autocorrelation. This matrix is a $n \times n$ square matrix, and each element w_{ij} in this matrix indicates how strongly each of the n spatial observations is related spatially.

The global Moran's I value has a range between -1 and 1, and when it is greater than 0, it indicates that a given area has a positive correlation with its neighbours and shows spatial agglomeration. If the index is less than 0, it suggests that the regions tend to have geographical dispersion. In the case of $I = 0$, it shows there is no spatial connection between the observations.

The above global spatial autocorrelation analysis aims to give an overall conclusion of whether spatial autocorrelation exists in the study area; however, it is unable to identify whether the observation values have spatial clusters [38]. Therefore, to determine those spatially clustered areas, we investigate the local spatial heterogeneity between a specific area and its neighbours using local spatial autocorrelation analysis [36,39,40]. This study conducted local Moran's I index as one type of local indicator of spatial association (LISA) [41] to test the local autocorrelation, The idea of it is to identify whether the value of an observation and the average of its neighbours is positively autocorrelated (HH or LL) or negatively correlated (HL, LH), since we only interested in the areas where clusters are positively correlated, i.e., HH and LL clusters. The equation of the local Moran's I index is as follows:

$$I_i = \frac{z_i}{m_2} \sum_j w_{tj} z_j; m_2 = \frac{\sum_i z_i^2}{n} \quad (2)$$

where I_i equals the local Moran's I index; m_2 is the variance of the distribution of values in the data; Σ_j and z_j are the standardised values of the observation at location i and j; while w_{ij} refers to the ith row and jth column of a W spatial weights matrix; n means the number of observations. LISA can quickly return clusters where values are concentrated and provide evidence about potential processes [42]. However, neighbourhoods should be

statically significant to be considered, and we only take neighbourhoods that have at least one high-value neighbour and p-value < 0.05 into account.

3.3. Impacts of the Environment Factors on House Price

3.3.1. Data Processing

The data used in stage II serve as the basis for the regression analysis. Particularly, we used the following two datasets:

(1) CRIC annual new-built developments data: The development transaction data contains 787 newly built developments' transaction data in 2021 as the dependent variable. The original data were cleaned by removing all the empty values, therefore leaving the values for developments that have transaction records throughout that particular year. Despite the considerable structure-related information in the raw data, it only contains the actual address of the developments without longitude and latitude, therefore, geocoding through Gaode Map API v2022 is conducted. Moreover, for the consistency of the housing structure, housing developments that are not commercial properties, such as compensation housing and subsidised housing, were also removed.

(2) Gaode Map accessibility and amenities data: To transform the accessibility and public amenities data to the independent variables, feature engineering was conducted to convert these points to the shortest distance between development and the different public facilities. This is achieved by distance matrix analysis on QGIS 3.20.

(3) Factor Selection: There are 14 variables in total, including both independent (13) and dependent (1) variables, the new development transaction price is the only independent variable. For independent variables, there are two categories: structural attributes and built environment attributes. Table 4 provides the detailed information for each variable.

Table 4. Regression variables.

Variables	Definition (Name and Unit)	Abbrev.	Mean	Code
House price	average transaction price for newly built developments (CNY/m^2)	House price	34,610.74	
Structure attributes	Number of total dwellings	Total_D	909	1
	Construction area/land use area ratio	C_ratio	1.86	2
	Parking plots per household	Park_P	1.17	3
	Average transaction volume (CNY)	Av_TV	4,691,659	4
	Average apartment areas (m^2)	Av_AA	99.67	5
Built environment attributes	Distance to the nearest primary school (m)	D_School	1538	6
	Distance to the nearest undergrounds station (m)	D_Underground	2910	7
	Distance to the nearest hospital (m)	D_Hospital	13,788	8
	Distance to the nearest municipal park (m)	D_Park	1686	9
	Distance to the nearest train station (m)	D_Train	16,589	10
	Distance to the nearest top 60 high school (m)	D_Top	5657	11
	Distance to the nearest kindergarten (m)	D_Kindergarten	663	12
	Distance to the nearest shopping centre (m)	D_Shoppingcentre	13,037	13

We followed international literature and expert knowledge to select these variables. However, it should be noted that not only literature regarding Shanghai were referred, so the finally selected variables were a little different from similar studies. For instance, we noted that some studies excluded the distance to central business districts (CBDs) (e.g., [22,43]), or some studies found that this variable had no significant relationship with housing price (e.g., [31]), so we did not consider the distance to CBDs. Moreover, we focus on the accessibility of service amenities. In China, CBDs are rather considered as places for business activities or working, instead of a kind of service amenity. Hence, we finally did not consider this variable, although some studies have proven its important

role in shaping the house prices. Last, this study aims to mainly investigate the role of service amenities and public transportation in shaping house price, so we did not consider other kinds of variables, including structural factors (e.g., building age, number of floors), economic factors (e.g., income level, mortgage interest rates), or demographic factors (e.g., household size).

(4) As we stated in Section 3.2, the currency unit is the Chinese Yuan represented by CNY, as it has been widely adopted in the housing market in China.

3.3.2. Factor Multicollinearity Analysis

Before running the GWR model, it is necessary to carry out a spatial autocorrelation analysis and obtain a significant result. Otherwise, it will be inappropriate to use this type of model [44]. In addition, the multicollinearity between variables should also be checked before running the GWR model. Therefore, we test the global spatial autocorrelation of the CRIC development dataset under the same method in Stage I, where multicollinearity is checked using the Pearson correlation coefficient and variance inflation factor (VIF), which is helpful to drop highly correlated variables. Generally, a VIF value of less than 5 suggests independence among the factors, while a value exceeding 5~10 indicates multicollinearity, necessitating the exclusion of the correlated factor [45,46]. The formula for calculating VIF is as follows:

$$VIF = \frac{1}{1 - R_j^2} \tag{3}$$

where R_j^2 is the determination coefficient for the regression analyses of all other variables.

3.3.3. Ordinary Least Squares (OLS) Linear Regression Model

Before conducting the geographically weighted regression, it is also helpful to explore the relationship between variables and select variables using an ordinary least squares (OLS) linear regression model [44]. In this study, a hedonic pricing model has been applied, and the hedonic pricing model is a type of multiple linear regression that regresses on a given price by characterising its many characteristics using a variety of explanatory variables [17]. The hedonic price model has been extensively utilised to determine the values of houses all over the world for the last three decades [7]. The formula of the hedonic model is written as:

$$Y = C + \beta X + \varepsilon * \varepsilon \sim N(0, \sigma^2 I) \tag{4}$$

where Y equals the developments' unit price; C is the constant term; β stands for the parameter for explanatory variable X; and ε^* refers to the error term.

3.3.4. Geographically Weighted Regression

OLS offers a comprehensive overview of the Shanghai real estate market's dynamics. However, the fundamental distinction between different parts of Shanghai is still not unclear. There is substantial evidence in the literature that the housing market exhibits both spatial autocorrelation and spatial heterogeneity [47–49]. If those spatial effects exist, the hedonic price model will be unable to be compatible with the normal distribution, which brings challenges to the spatial analysis and potentially produces insignificant model estimates [30]. Thus, GWR is introduced, it is also a linear regression model but pays may more attention to the local regressions [44]. The linear regression (LR) and spatial lag model (SLM) are also widely used econometric models to explore the neighbourhood influence of the dependent variable [50] (Elhorst, 2010). However, they are both global and stationary forms where coefficients are assumed to be stationary and fixed across space. Hence, we selected the GWR model to capture the spatially heterogeneous features in data relationships, which fits better with the settings of our study area [51].

The basic idea of GWR is to explore how the relationship between a dependent variable and explanatory variables will change geographically. It assumes that regression coefficients are a function of the observation point's geographic location, so it is possible to analyse the spatial differences between the dependent variables [17,52,53]. The limitation of the GWR would be it still highly sensitive to multicollinearity [54], and the dependence between spatial errors [55]. The model formula is as follows:

$$y_i = \beta_0(\mu_i, v_i) + \sum \beta_k(\mu_i, v_i) x_{ik} + \varepsilon_i; i = 1, 2, \ldots, n, \quad (5)$$

where (μ_i, v_i) refers to the coordinate of the observation point i; $\beta_k(\mu_i, v_i)$ equals its kth regression parameter.

The way that GWR works is by sequentially moving a search window from one point in the dataset to another; when the search window sits on a sample point, all other points that are around it and within the search window are identified as regions. Regression points in each region are then fitted using a regression model, with the points closest to the centre of the model receiving the most weight [56]. Our data contain 787 observations, and GWR will fit 787 weighted regression models. Choosing the size of the 'search window' is one of the key steps when fitting the model, which can be achieved by bandwidth selection. Our study is conducted by searching for the optimal bandwidth using the Sel_BW class in the mgwr package. By selecting their default values, a fixed distance is used to calculate the bandwidths of the Gaussian kernel.

4. Results

4.1. Exploratory Data Analysis

4.1.1. Variable Distribution

Figure 3a shows that the distribution of house prices is positively skewed rather than normally distributed. To increase model interpretability, a log transformation of the development house price was conducted (Figure 3b). The distribution of each individual variable is shown in Figure 4; no further log transformation is performed. Both independent variables and dependent variables are standardised before fitting the regression model.

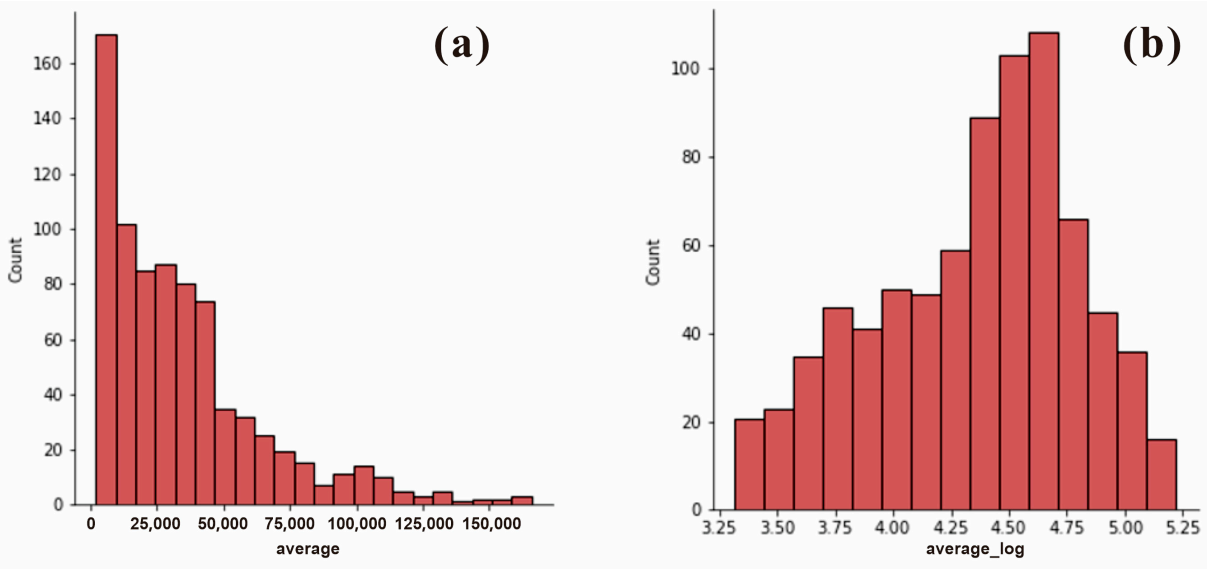


Figure 3. Histogram of house price: (**a**) Histogram of house price; (**b**) Histogram of logged price.

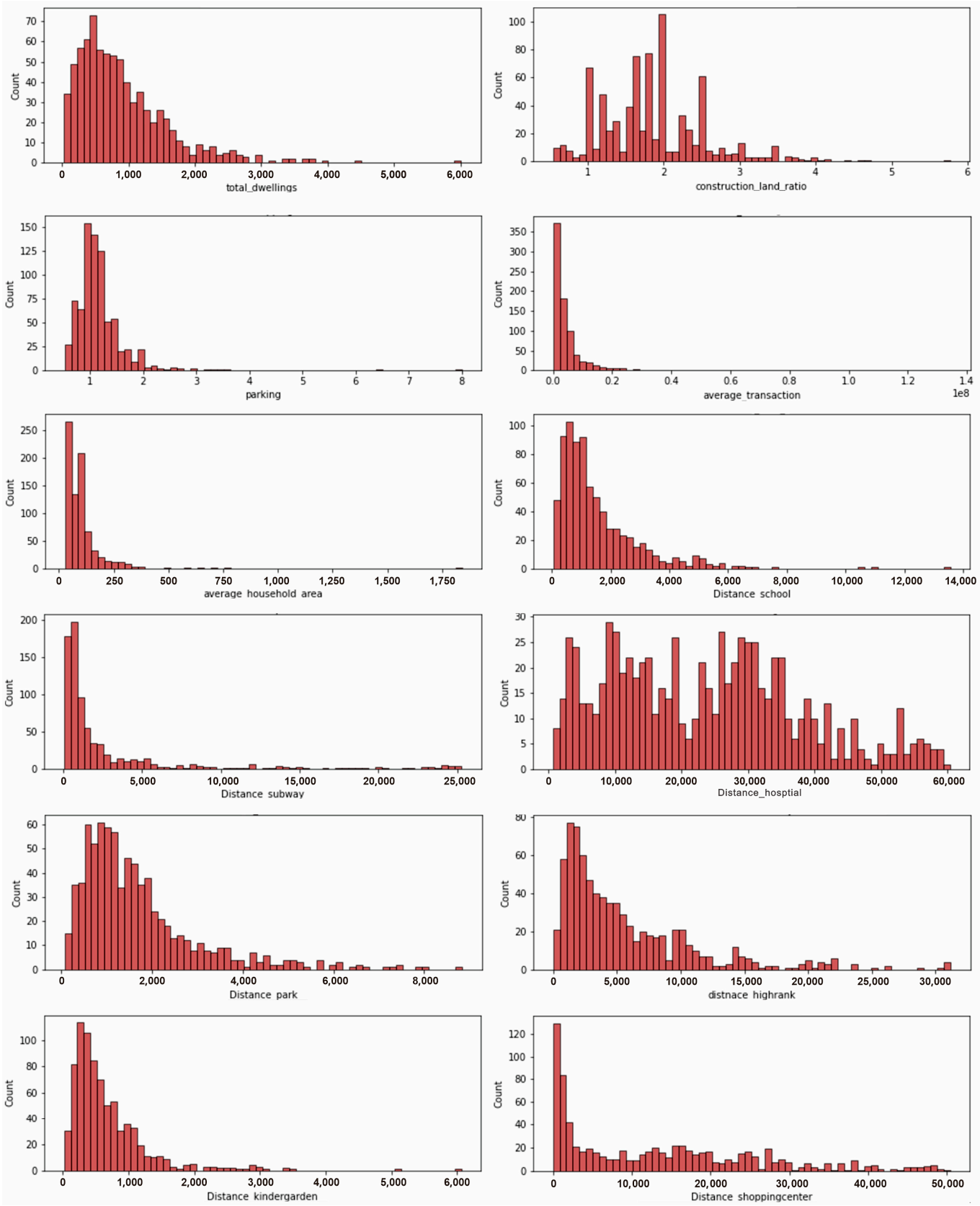


Figure 4. Distribution of independent variables multicollinearity, where the units for the variables have been addressed in Table 4.

4.1.2. Multicollinearity Analysis

The multicollinearity between variables is checked, Figure 5 shows there are highly correlated variables between some variables, e.g., 'distance to the train' and 'distance to hospital'; 'distance to kindergarten', and 'distance to hospitals'. We further use the variance inflation factor (VIF) to remove the variables that have a high value above 5, and the variable that is removed is 'distance to hospital'. The correlation between 'distance to the train station' and 'distance to kindergarten' was also relatively high, but less than 5, so they were kept in the analysis. Please note that, herein, the VIF values were the normalised values. Next, we utilised an automatic procedure recently developed in the literature [57] to test the local multicollinearity in each spatial cluster (see Section 4.2 for the details of the spatial cluster). The results showed that all of the obtained VIF values were smaller than 5, indicating that the local multicollinearity did not exist in our analysis.

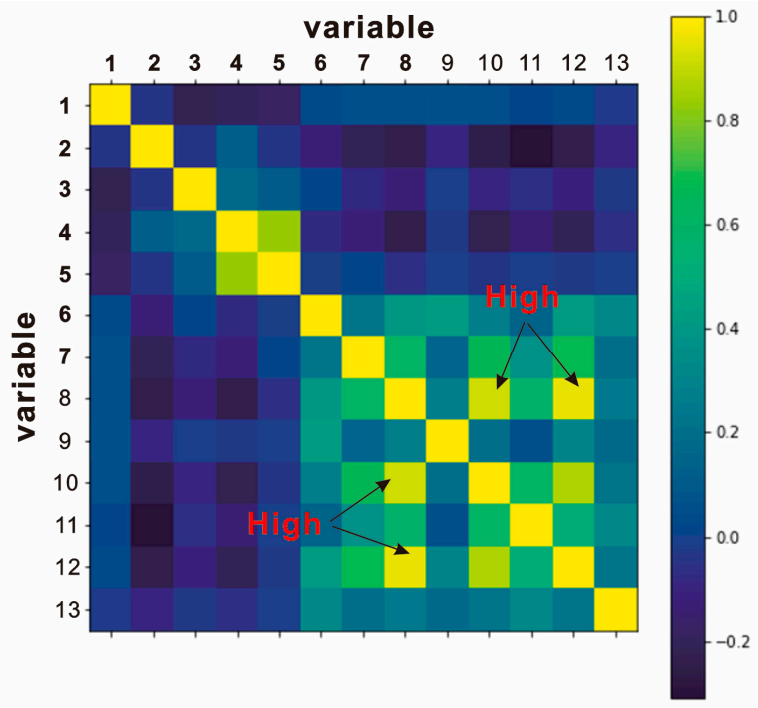


Figure 5. Results of the multicollinearity analysis showing the correlation of independent variables, where the code for variables have been clarified in Table 4.

4.2. Spatial Cluster Pattern of Housing Price

The distribution of the average annual housing price in Shanghai from 2015 to 2021 has been illustrated in Figure 6, which shows that housing prices in Shanghai vary spatially, with high house prices concentrated in the city centre and expanding from the centre to the outskirts from year to year.

In 2015, most of the neighbourhoods had an average unit price of below CNY 30,000 per m^2; neighbourhoods Huaihai Road' and 'South Jingan' are the only two areas that reached an average house price of at least CNY 100,000 (roughly GBP 12,193) per m^2. Moreover, neighbourhoods in the city centre area, particularly those two neighbourhoods, always drive a sharp increase in the average house price across these years. Neighbourhoods such as 'Huaihai Road', 'BiYun', and 'North Waitan' are the ones with the highest house prices, which reached a price range of CNY 100,000 to CNY 160,000 in 2021, while most of the neighbourhoods in the outer ring still experienced house prices below CNY 30,000. Generally, the high-price neighbourhoods are concentrated in the central city districts. However, in recent years, some neighbourhoods have higher house prices in the outskirts than those neighbourhoods relatively closer to the city centre.

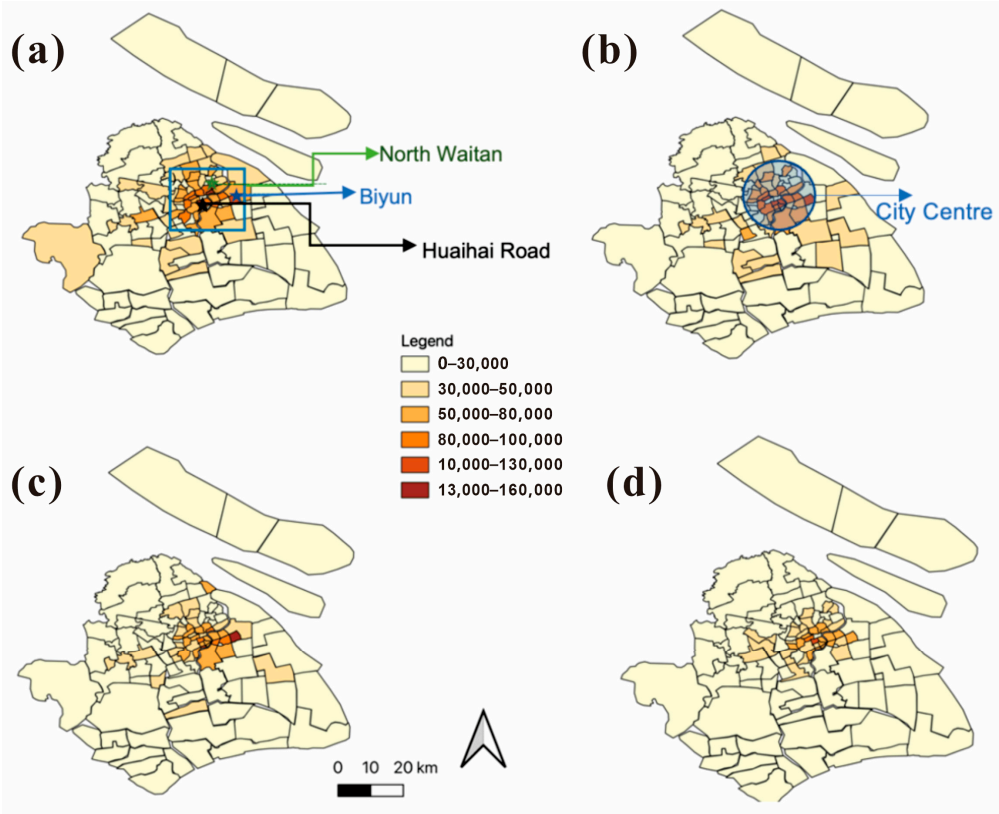


Figure 6. Distribution of house prices in 114 neighbourhoods during 2015 and 2021: (**a**) 2021 neighbourhoods with significant high house price neighbourhoods labelled; and (**b**) 2019 neighbourhoods with the outline of the city centre (**c**) 2017. (**d**) 2015.

Figure 7 and Table 5 show the results regarding the global Moran' I index analysis. Figure 7 displays a positive spatial autocorrelation of the standardised house price, as the line of best fit displays a positive relationship between both the standardised house price and its spatial lag. It means the overall trend is for high values to be close to other high values (the first quadrant), and low values surround low values. This suggests a positive spatial autocorrelation relationship that exists in the 2021 Shanghai neighbourhoods, which means neighbourhoods with high house prices are always adjacent to each other.

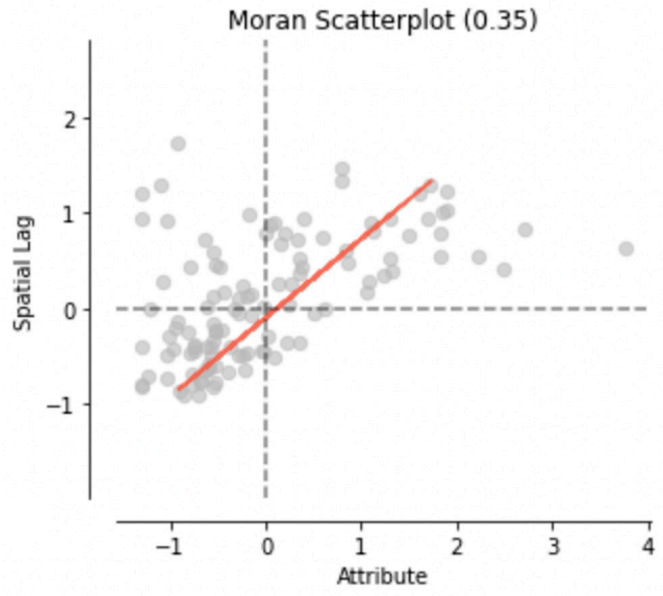


Figure 7. Moran scatterplot.

Table 5. Result of the global Moran's I index analysis for 2021 neighbourhoods house price.

Index	Value
Moran's I index	0.35165
Exp.index	−0.00887
Variance	0.00169
Z-score	8.76536
p value	0.00000

According to Table 5, Moran's I index is greater than zero (0.35) in 2021, which further illustrates at the city level, Shanghai experiences positive spatial autocorrelation characteristics and can demonstrate a significant statistical phenomenon of spatial clustering (Z-score > 1.96 and *p* value < 0.05), that is, in the significant positively correlated areas, if neighbourhoods' own house price is high, then the surrounding neighbourhoods are correspondingly high, and vice versa.

The Moran cluster map (Figure 8) shows whether the neighbourhoods where the *p*-value of the index < 0.05, is statistically significant. It finds which parts of Shanghai are spatially positive autocorrelated clusters. Regions in the red colour stand for neighbourhoods that have high house prices neighbours and are high house prices on their own. The HH clustering area is mainly concentrated in the city centre districts (e.g., Jingan district, Huangpu district, and Putuo district); however, a small part of Pudong New Area is also in HH clusters. The significant low–low clustering area is concentrated in the southern part of Shanghai, mainly located in the Jinshan, Fengxian, Qingpu, and Pudong New areas, which are all districts in the outermost parts of Shanghai. There are also LH regions in light blue in the centre, indicating negative spatial autocorrelation between themselves and their neighbours. Table 6 shows the names of the significant neighbourhoods in different clusters and in which specific districts they are located. Since only the high–high (HH) clusters and low–low (LL) clusters stand for positive autocorrelation relationships, we only take significant HH and LL clusters into further analysis.

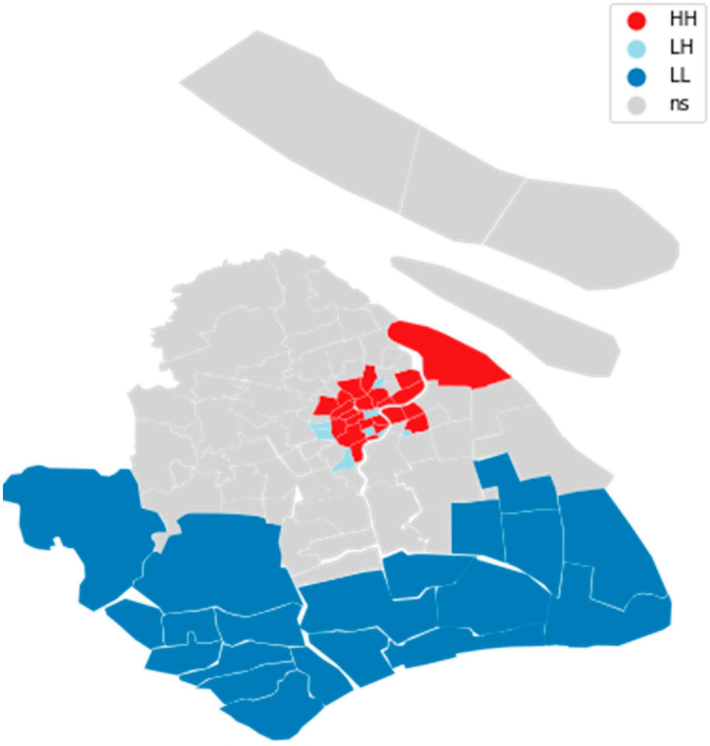


Figure 8. Moran cluster map for significant LISA values.

Table 6. Neighbourhoods in each cluster.

Labels	Neighbourhoods	Count	Districts
High–high (HH)	Yuyuan, Binjiang, Buyecheng, Inner Putuo, Daning, Shiji Park, East Waitan, Gaoqiao, Huaihai Road, Huangxing, North Jinnah, South Jingan, Liangcheng, Longhua, LujiaBang, Lujiazui, Inner Changning, Inner Xuhui, North Sichuan Road, Siping Road, North Waitan, Yuanshen, Zhenru, Caoyang, Yichuan	24	Seven Central Districts, Pudong New Area
Low–low (LL)	Ganxiang Town, Lvxiang Town, Zhangyan Town, Caojing Town, Tinglin, Fengjing, Zhujing Town, Huinan, Lingang New City, Shuyuan, Jinhui, Nanqiao, Fengcheng, Jinshan New City, Songjinag New City, Xinchang, Tuolin, Haiwan, Zhujiajiao, Zhuqiao	20	Songjiang, Fengxian, Qingpu, Pudong New Area

After determining significant clusters, we used the Gaode Map accessibility and amenities data to calculate the numbers of the built environment in each cluster. As seen in Figure 9, the radio plot shows the average number of service amenities and public accessibilities per neighbourhood in both the HH and LL clusters, which can indicate whether each cluster has a different determination on house price. Since the total numbers of different facilities could vary, for some facilities, we took the logarithm and labelled them with *. Generally, it shows that the HH cluster neighbourhoods are more convenient to access the facilities.

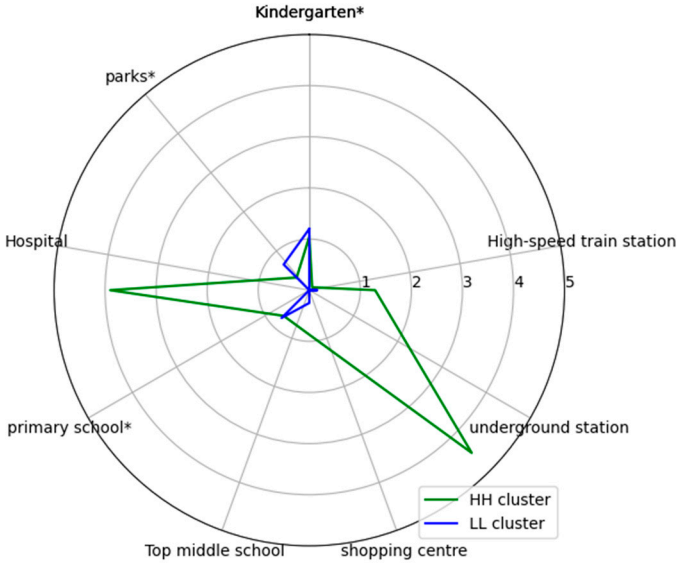


Figure 9. Radio plot for average built environment facilities in each cluster. The label * represents the logarithm of the variable values.

Generally, the results show that the HH cluster represents neighbourhoods that are more convenient when accessing transport, amenities, and leisure amenities, i.e., the average number of general hospitals, underground stations, shopping centres, top middle schools and high-speed train stations in the high–high cluster is significantly greater than those in LL cluster. On the other hand, the number of some basic amenities such as parks, kindergartens and primary schools in the LL areas is greater than those high house price clusters.

However, this graph only characterises the different number of facilities in each cluster area; it cannot be used to prove that there is a relationship between those variables and house prices. Moreover, in Figure 8 and Table 6, it is evident that each neighbourhood has different areas, which may show that counting the average number of facilities has

limited meaning. This raises the need for further regression analysis explicitly using the development scale data (where we can calculate the Euclidean Distance to facilities as X variables).

4.3. Effects of the Built Environment Factors on House Price

The analysis results of the OLS regression model are tabulated in Table 7. The explanatory power of the chosen variables on average development prices is represented by R^2 values. In 2021, the overall R^2 was 36.5%, with an adjusted R^2 of 35.6%. We noted that the fitting level of our results was moderate, lower than some similar studies [31], but higher than some others [22,51,58,59]. The *p*-value of the model is <0.01, which meets the requirement for the highest significance level, showing that the OLS model is robust. However, in the total of 12 variables, only 7 of them are the significant influencers of the house price, which are the construction area/land use area ratio, total dwellings number, average transaction volume, and distance to the nearest high-speed railway station/undergrounds stations/shopping centres/municipal parks. The average effect of coefficients is minor in the global OLS model. The intercept is not significant, meaning that, by removing all the variables, the developments will have no value.

Table 7. Result of the OLS model, where ** represents significant, and *** means very significant.

Variable	Abbrev.	*p*-Value	Significance
Total dwellings number	Total_D	0.002	***
Construction area/land use area ratio	C_ratio	0.000	***
Parking plots per household	Park_P	0.278	Not significant
Average transaction volume	Aver_TV	0.000	***
Average apartment areas	Aver_AA	0.790	Not significant
Distance to the nearest primary school	D_School	0.849	Not significant
Distance to the nearest undergrounds station	D_Underground	0.018	**
Distance to the nearest municipal park	D_Park	0.013	**
Distance to the nearest train station	D_Train	0.005	***
Distance to the nearest Top 60 high school	D_Top	0.661	Not significant
Distance to the nearest kindergarten	D_Kindergarten	0.763	Not significant
Distance to the nearest shopping centre	D_Shoppingcentre	0.000	***
R^2 = 0.356; F-Stat = 37.15; P(F-Stat)= 2.86×10^{-183}; AICc = 1901.390			

The null hypothesis for the global Moran's I index is that the spatial pattern is complete randomness. Our result of the global Moran's I index in 2021's development house price has a Z-score > 2.96 and *p* value < 0.05), which can reject the null hypothesis. Together with the positive Moran's I index, there is a positive spatial autocorrelation in Shanghai, which is thus appropriate for the GWR model. Table 8 shows the result of the GWR model. Compared to the OLS, the overall adjusted R^2 has been increased to 0.48, meaning that this model can better explain the relationship between built environment factors and house price. Figure 10 further illustrates the range of R^2 for all the features in the dataset, showing an R^2 between 0.435 and 0.566. On average, the LL cluster has a better R^2 than the HH cluster, suggesting that our chosen variables can better explain the house price in LL clusters. The north high–high cluster, south Baoshan district, and Jiading district are regions where built environment variables have the highest explanatory power. The bandwidth is 308 m, which shows the smoothness of the model. The optimal result of

bandwidth is estimated by the AICc methods; the AICc is 1812. The residual square is 365. The residual square and AICc for the GWR model are smaller than OLS, indicating that it fits the data better. Since in the GWR, calculating the *p*-values for independent variables is unsuitable. Interpret only the significant variables in the OLS model could be a suitable solution. Therefore, only seven variables are significant, four of which are accessibility and service amenities (D_shoppingcentre, D_Park, D_Underground, and D_Train).

Table 8. Result of the 2021 GWR model, where variables with * are significant variables.

Variable	Mean	Min	Median	Max
R^2	0.483	0.435	0.486	0.566
Intercept	−0.105	−0.981	−0.005	1.059
Distance to the nearest train station *	0.102	−0.026	0.106	0.271
Distance to the nearest underground stations *	−0.044	−1.642	0.042	0.807
Distance to the nearest top 60 high school	0.041	−0.207	0.067	0.441
Distance to the nearest primary school	−0.019	−0.171	−0.021	0.279
Distance to the nearest kindergarten	0.012	−0.225	0.025	0.183
Distance to the nearest shopping centre *	−0.357	−1.554	−0.276	0.591
Distance to the nearest municipal park *	0.08	−0.101	0.048	0.270
Average apartment areas	0.03	−0.508	0.062	0.802
Construction area/land use area ratio *	0.133	0.022	0.122	0.267
Parking plots per household	0.035	−0.041	0.028	0.152
Total dwellings number *	−0.027	−0.157	−0.034	0.103
Average transaction volume *	0.582	0.018	0.634	1.284
R^2 = 0.537; Adjusted R^2 = 0.483; AICc = 1812; Residual Square = 365; Bandwidth = 308				

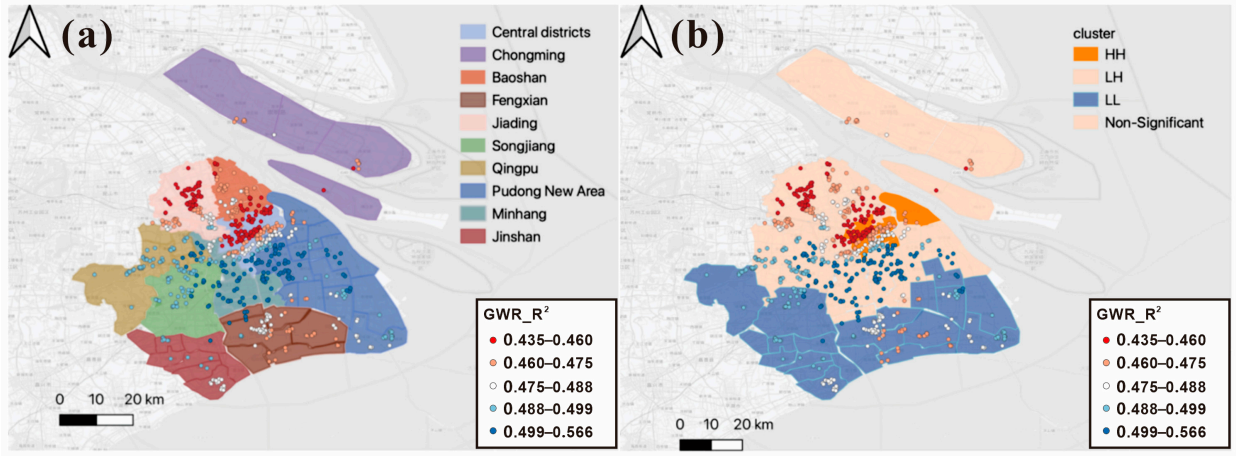


Figure 10. (**a**) R^2 map against different districts; and (**b**) R^2 map against different clusters.

Figures 11 and 12 both illustrate the spatial variation for the coefficients of the significant built environment factors. In Figure 11, variables are plotted based on the Shanghai districts, and in Figure 12, coefficients are plotted against the distribution of spatially autocorrelated clusters. Together with Table 8, in Shanghai, generally, D_shoppingcentre shows the strongest negative correlation (−0.357) to the house price, which indicates that the closer a development is to the large shopping centre, the more expensive the property is. The highest negative coefficients of this variable are concentrated in the city centre (high–high cluster) and expand to the Baoshan district. Moreover, the effect of the shortest distance between the parks varies; the highest negative values appear in the HH cluster areas and expand to the southwest, while in the rest of the suburban regions, including the low–low cluster, the correlation is positive. For public transportation, there is a generally positive effect on the distance to railway stations, which indicates that, in most of the

regions in Shanghai, an increase in the distance to railway stations will decrease the house price, but in Songjiang, Qingpu, and the middle part of Pudong New Area, the effect is in contrast. Finally, the effect of distance to the underground stations is negative overall (−0.044), but in parts of the urban core, it shows a positive relationship.

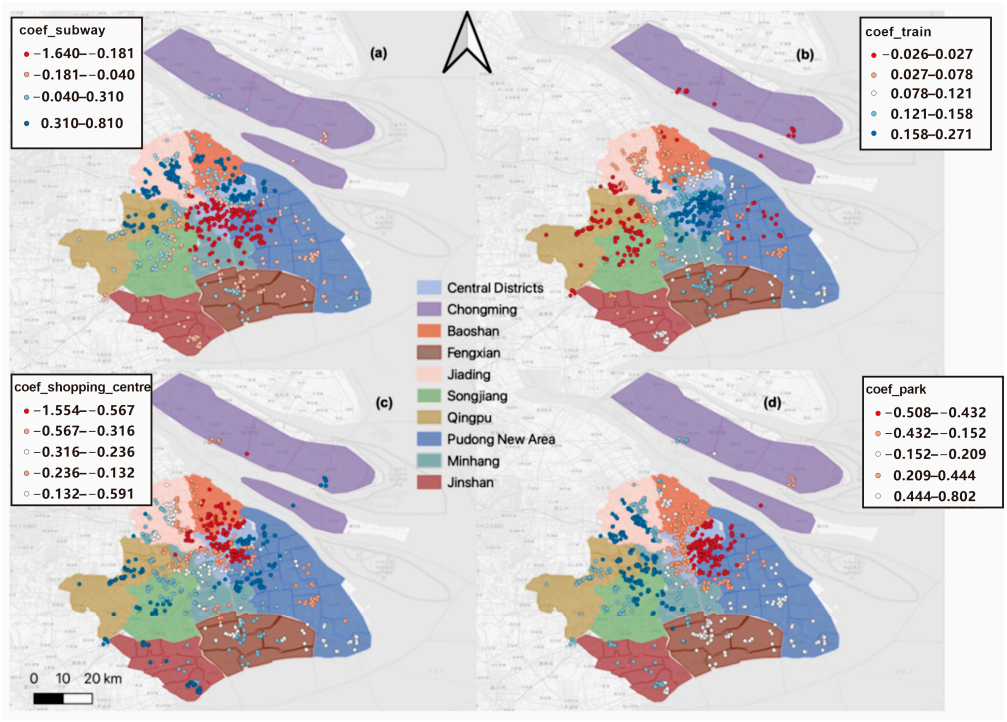


Figure 11. Correlation plot of independent built environment variables against districts: (**a**) Spatial variations of the underground station; (**b**) Train; (**c**) Shopping centre; and (**d**) Parks.

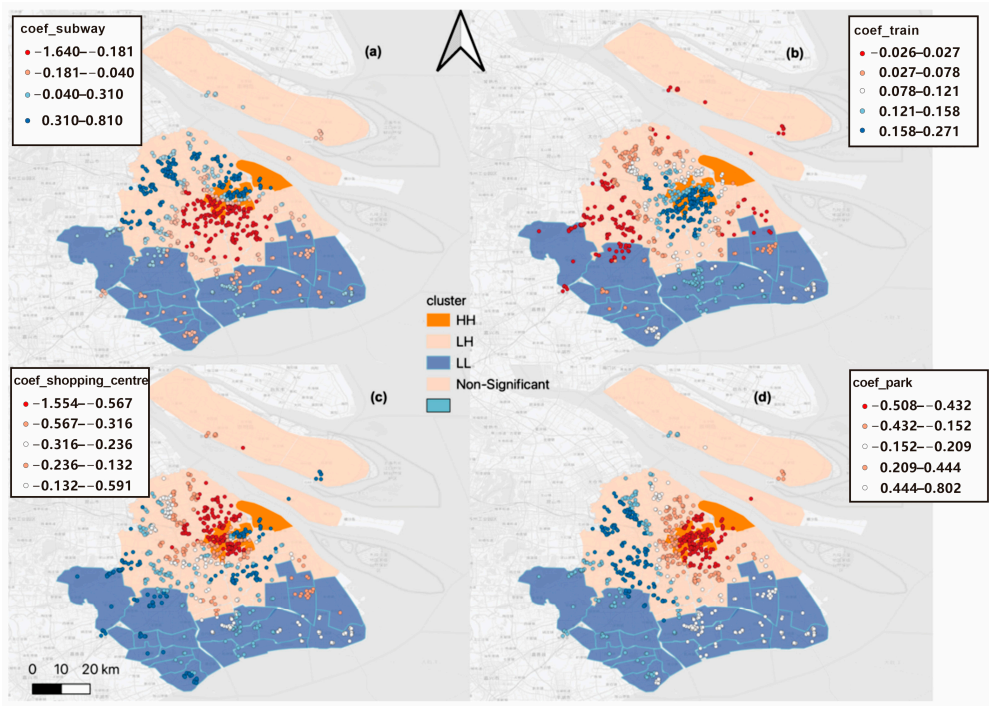


Figure 12. Spatial differentiation map of significant built environment factors: (**a**) Spatial variations of the underground station; (**b**) Train; (**c**) Shopping centre; and (**d**) Parks. The values of the dot represent the correlation coefficient between the variable and housing price. All the dots are divided into five categories, which contain the same number of dots (same principle for Figures 10 and 11).

Focussing on the cluster results (Figure 12), in the high–high cluster, the relationships between accessibility to public transportation and house prices are always negative, and the relationship between accessibility and service amenities is always positive. For example, one unit increase in the distance to the underground stations will increase the development price by at most 0.31–0.81 std (blue dots in Figure 12a); and one unit increase in the D_Train will cause house price to increase by 0.146—0.27 std (Figure 12b). On the other hand, a unit rise in the D_shoppingcentre and D_Park will decrease the development transaction price by around 1.5 and 0.5 namely (Figure 12c,d), meaning that the decrease in the distance to service amenities has a positive impact on house price.

The low–low cluster can be called 'suburban fringes' as it includes districts in the fringes such as Jinshan, Fengxian, and the southern part of Pudong New Area. The overall effect is similar to other suburban regions, but the magnitude of the coefficients is always in the middle range (Figure 12). For instance, in contrast to the values in high–high clusters, an increase in one unit of the distance to underground stations will reduce house prices for most of the developments in this area by 0.04 to 0.18, but an increase in the distance to the railway station in LL areas will have a different impact across the LL cluster, in which neighbourhoods in Songjiang and Qingpu shows an oppositive trend. In the service amenities attribute, similar to the urban core, increasing the distance to the shopping centres creates a higher house price, but the effect of increasing the distance to the urban parks has a negative impact on the house price.

5. Discussion

The above result section suggested that the distance to the nearest high-speed railway station, distance to the nearest underground station, distance to the shopping centre, and distance to parks are the four significant built environmental factors that impact the development transaction prices, and that the effects may vary across the city. This chapter expands on the result we found and attempts to give detailed reasons to answer how built environment factors—such as proximity to service amenities and public transportation—influence the house price across spatially autocorrelated neighbourhoods' clusters and Shanghai more generally? This is answered by (1) how and why the spatially autocorrelated neighbourhoods' clusters are distributed in this way; and (2) how different factors influence the house prices across those spatially autocorrelated neighbourhoods' clusters and more generally Shanghai in a spatially different way.

5.1. Interpreting the Spatial Clusters

In 2021, Shanghai experienced significant spatial segmentation, where the spatially autocorrelated high–high clustering areas are always concentrated in the city core, but the low–low cluster appeared in the southern fringe of Shanghai (Figure 8). Since there is no significant HH cluster in the suburban areas, this indicates that there is an obvious monocentric structure in Shanghai's housing market, and that this structure is only concentrated in downtown Shanghai, with no spillover effect to its neighbouring suburban areas. This is inconsistent with the city's planning, which envisions a polycentric structure where 11 new towns will be developed in each suburban district to stimulate local economic activity. This result suggested that Shanghai has failed to build these new towns or at least failed to meet the goal of letting the new towns drive up the local real estate market, as the market still having greater confidence and expectations in the central areas. In addition, there are not only the housing attributes, in terms of employment densities and land use, but as shown by other recent studies, Shanghai also still has a monocentric development rather than a leapfrog [60].

One of the main reasons for this monocentric structure (high–high clustering concentrated in the urban core) would be the city 'redevelopment' projects in Shanghai. For instance, unlike the redevelopment projects in developed countries, the redevelopment model in Shanghai is mainly profit-driven and aims to promote local economic growth by attracting investment in the housing market and property development [7,61]. Therefore, the approved redevelopment projects are mainly located in the inner-city centre with better access to public facilities, thus having better attractiveness for investors and consumers and accelerating the over-concentration of the city centre [62].

The formation of the low–low cluster in the southern parts can be interpreted as the result of government policies on transportation and infrastructure, as it resulted in lower house prices in less accessible, underdeveloped districts. Northern Shanghai is adjacent to Suzhou, the core city of the 'Yangtze River Delta Economic Zone, so the traffic networks between these regions are more intensive, and the government is also vigorously promoting the development of the northern districts. The average house price in such districts is continually increasing and always higher than in the southern part (Figure 4). In contrast, the low–low cluster for house prices in the southern part can be explained by the weaker transport accessibility due to the unfavourable policy in this region. This result is similar to previous literature [28], which found that the apartments in outer Shanghai are the particular 'cold spot' on the market by performing the Getis–Ord Gi* test on Shanghai developments.

As we already explained the existence and formation of those clusters, further investigation on how the proximity to transportation facilities and service amenities are linked to housing prices can be implemented. With a much smaller area than rural districts, the downtown areas still have significant advantages in terms of subway stations, general hospitals, shopping centres and a high level of education (Figure 9). Therefore, the rich public service facilities are one of the characteristics of high house price neighbourhoods, but whether the correlation exists needs to be examined in stage II.

5.2. Exploring the Effects of the Built Environment Factors on House Prices

The existence of the neighbourhood clusters reveals that the accessibility to the facilities may be the essential factor in forming spatial patterns and differentiating house prices. From this perspective, this study takes 787 developments in Shanghai as area units, aiming to explore the different spatial correlations between house prices and built environmental factors, and then further characterise the heterogeneous characteristics of house prices across the whole of Shanghai and its spatially autocorrelated clusters.

- **Distance to the Underground Stations**

Generally, the developments in the suburbs, including the low–low cluster, show negative relationships between the distance to underground stations and house prices. In these areas, convenient public transportation can positively affect the price (Figure 12a). This can be explained as residents in the suburban areas are more demanding on public transit as they may work in the city centre, and undergrounds can provide cheap and convenient transportation, thus favoured by many white-collar workers [63]. Another example that can explain this is that in districts such as Jiading and Baoshan, which have plenty of industrial parks within walking or cycling distance, residents intend to be less demanding on undergrounds for the daily commute, and thus, house prices show a positive correlation between the proximity to undergrounds.

However, the correlation coefficients for the D_Underground in the high–high clusters are positive, i.e., the house prices do not increase as accessibility to underground increases (Figure 12a). This can be explained by people living in the city centre that can access high-quality facilities and amenities within walking distance reducing their reliance on

transportation to reach their destinations. Moreover, the scarcity of underground stations in suburban areas makes it more favourable in the urban fringes than in the city core when considering house prices. This finding is in line with the observations in other big cities in China. For example, Li et al. [31] found that the role of underground stations in outer Shenzhen's housing market is more critical than in Inner Shenzhen (southern part of China).

- **Distance to the Railway Stations**

The effect of an increase in distance to a railway station on house prices may be based on when the region is constructed. Most districts show a positive correlation between the distance to the train stations and house prices (unfavourable), and the developments in the HH cluster show the most unfavourable attitude towards the train station (Figure 12b). As explained in stage I, the redevelopment plan first started in the city centre, which had made this area densely built and with an older average building age. In these areas, public transportation is less scarce, but the high-speed railway produces noise pollution and increases the traffic intensity level in the urban area, and thus decreasing house prices [64]. However, housing developments in Songjiang, Qingpu, and Pudong New Area show a slightly positive preference for the train stations. This might be due to these areas that are considered 'new towns' in the city planning, which developments in these areas are less dense, later built, and have more land areas and advancing soundproofing infrastructure, so that residents will be less sensitive to noise and traffic.

Moreover, compared with other suburbs without railway stations and infrastructure, the landmarking high-speed railway stations can attract more reason for the financial sectors to invest in commercial, infrastructure, and real estate. Thus, bringing ample parking areas, underground stations and restaurants to the area, providing convenient living services for these underdeveloped areas. Therefore, the closer apartments in these areas are to the station, the higher the housing price will be.

- **Distance to the Shopping Centre**

A large shopping mall makes it more convenient for the residents, thus promoting an increase in house prices across the whole city (Figure 12c). This relationship is in line with the studies regarding other big cities in both China and other countries (Canada and the United States) [65–67]. However, using the exact scope of data, our study is slightly different from the literature. For example, Liu et al. [67] found that large shopping centres could promote the increase in house prices generally in Wuhan (China), but in cases where houses are too close to the densely distributed shopping mall, travel congestion and noise pollution will decrease the house price. In contrast, in our circumstances, house prices in central city areas appear to react more positively to the high concentration of shopping malls. Considering the GDP gap between the two cities, it may reveal that the different building volumes and city sizes of Shanghai and Wuhan make the need for convenient service amenities outweigh the noise impact when evaluating house prices.

- **Distance to the Park**

Our result shows that properties in the city core will have higher market prices if they are closer to parks; in contrast, positive correlations appear in the rest of the districts. This indicates that, in suburban areas, an increase in D_park cannot increase house prices (Figure 12d). On the one hand, the unequal distribution of service facilities between urban cores and suburban areas may drive up housing prices in certain regions. In particular, 'blind spots' in the layout of parks within central urban areas make parks scarce commodities, further inflating housing prices in these areas. While since the sub-urban areas have cheaper land, the property and development sizes will be larger, so parks are more available in this area. On the other hand, as a symbol of psychological needs rather

than basic needs, parks can provide more leisure enjoyment. Therefore, in suburban areas where people care more about fulfilling basic needs such as public transportation, people will be less demanding of entertainment facilities. Meanwhile, those who live in the city core will be more demanding of the service facilities.

- **Distance to Educational Facilities**

Interestingly, all the educational amenities were deemed insignificant, including the shortest distance to kindergarten, the Top 60 high schools, and primary schools. Some potential reasons can explain this result: (i) In metropolises like Beijing and Shanghai, basic educational amenities such as kindergarten and primary schools are widely accessible. And, in the case of top high schools, admission is typically determined through examinations rather than proximity; (ii) The overall spatial accessibility of basic education resources in Shanghai is favourable. For instance, Yuan and Yang [68] (2023) found that all the residents in Shanghai could reach the nearest primary school, junior high school, and high school within a 30 min travel time. This means that purchasing properties near these schools holds little significance. (iii) Since the beginning of this century, both in urban and suburban areas, Shanghai has adhered to the principles of "fair admission" and "admission based on proximity" and has promptly constructed new schools in areas with insufficient basic education resources. The above phenomenon has led to smaller inequalities in terms of education resources in big cities like Shanghai than in in rural areas, which is also supported by previous studies (e.g., [69] Song, 2023). It is difficult to directly compare our analysis with similar studies. However, we still found that the contribution of the accessibility of education resources on housing prices highly vary with studies (e.g., [44,70]). This lets us conclude that the effect of education amenities may significantly depend on the datasets and regional settings.

In summary, local authorities' city planning has generally impacted the spatial pattern of property values at the intra-city scale [71]. The OLS and GWR models both indicate that the distance to the shopping centres is the only factor that has an overall positive effect on house prices. In different regions, the effect of shopping centres all shows the highest premium. More specifically, developments in the high–high cluster (urban core) are always the hot spots of the housing market, as they are equipped with superior urban facilities and accessibility. This cluster always has the highest demand for service amenities such as parks and shopping centres and the lowest demand for underground stations, indicating that a 'gentrification' region is formed as people are more concerned about spending and entertainment than meeting their basic needs (public transport).

Moreover, the rest of the areas, including low–low clusters, show different reactions to different factors, but generally, the factor that can drive up house prices the most is the accessibility to the underground stations and the shopping centres. Indicating that the demand for convenient life essentials (public transportation and shopping centres) is higher than the demand for leisure amenities (parks). Understanding this specific spatial distribution could better help authorities plan infrastructure in differentiated clusters and aid housing buyers in choosing the most suitable neighbourhoods based on their needs.

5.3. Limitations

There are limitations in the analytical and designing processes of this study. Firstly, for the CRIC new development transaction data, there is a risk that the value of transactions could be underestimated. In 2012, Shanghai Housing Management Bureau announced that once the land is sold for constructing commercial properties, no less than 5 percent of government-subsidised housing against the total dwellings should be constructed. Therefore, the average transaction prices where by land is obtained for development after 2012 might be underestimated because the transaction price collected by the company

inevitably recorded cheaper social housing. Next, given that the housing transactions in the newly built projects play an dominant role in current Chinese housing markets, the second-hand housing prices was excluded in our analysis, which might fail to fully capture the overall market dynamics. We employed the dataset consisting of transaction prices for 787 newly built housing projects in 2021 for regression analysis. This dataset offers a larger sample size than neighbourhood-level housing prices, thus providing more precise environmental information. We note that some similar studies have been conducted with much larger sample sizes. However, it should be mentioned that the number "787" represents the number of building projects (like communities), not for individual apartments. Here, in China, it is normal to contain more than 1000 apartments in one single building project. This means that the current analysis covers the transaction prices at a hundreds of thousands buildings. Hence, we think the current dataset approximately represents a full picture of the residential market in Shanghai. To further clarify the reliability of the present sample size, we compared the utilised dataset with those in the literature, and we found that our sample size was smaller than some studies [31,72], but larger than some others [73]. Although the overall scale of the studies vary, the order of magnitudes of datasets related to the city-level were similar.

The term "built environment" is often used within the 5Ds framework, which includes density, diversity, design, destination accessibility, and distance to transit [74]. It is difficult to directly define the attribute for each variable used in this study, because they may be associated with more than one attribute. We primarily focus on the effects of proximity to service amenities and public transport on housing prices, but we still accounted for other attributes. The variable C_ratio considers the ratio between construction area and land use area, which can represent the density of a community to a certain degree. Similar situations also exist when it comes to Park_P, namely parking plots per household. Moreover, the varying functional designations of each administrative district can indirectly influence the design concepts and approaches of neighbourhoods. As we discussed in Section 5.2, people in different clusters of the city have varying design requests for communities. Unfortunately, we failed to quantify the relationship between the factors and their attributes, especially those associated with diversity.

Regarding the choice of service amenities variables, this study, with many studies focusing on the Chinese housing market, failed to include the grocery stores where people purchase daily necessities. This is due to the lack of data classification on the Gaode Map, where there are numerous small grocery stores in China that cannot be recognised. The GWR analysis highlighted the variable of proximity to shopping centres as an important factor. However, the omission of variables related to commuting distance, specifically the proximity to central business districts (CBDs) [63,75,76], represents another significant limitation of this study. However, as we explained in the Section 3.2, we emphasise the role of service amenities, and CBDs in China are normally considered as a kind of place for business activities or working, which is different from our research objectives. And, some studies pointed out that the current literature overemphasised the significance of the distance to CBD, and residents' preferences with respect to urban amenities have been largely neglected [58]. Finally, we included a total of 12 variables in the analysis, which was similar with some other studies [11,22]). Moreover, considering the insufficient variables, we dropped the D_hospital as it has high multicollinearity between variables, so we did not interpret the effect of those two variables.

Last but not least, we used the Euclidean distance to assess the accessibility of public transportation and service amenities. However, this might be inaccurate as the range of potential distance metrics is much broader than the conventional Euclidean option in the real-world scenario [77]. This may raise concern that the accessibility for some

developments is under- or overestimated, thus affecting the overall model performance. Hence, one of the future works could do the same analysis using the non-Euclidean distance as the ED metrics may fail to reflect true spatial proximity, and instead, the non-ED metrics such as road ND, TT, water distance or landscape distance can better describe the accessibility to locations.

5.4. Policy Recommendation

Our results showed that Shanghai presented a monocentric urban structure, with evidently higher housing price in the city centre. However, this is contrary to the multi-centre development goal of local authority, thus posting potential risks to urban development. Hence, it is important to optimise the city's polycentric structural framework to foster a more balanced regional development. New land supply strategies and incentives, such as tax discounts and subsidies, should be provided to encourage developers and enterprises to invest in suburban areas and form new economic growth poles.

We found that the impact of underground and high-speed rail stations on housing prices varied across neighbourhoods, particularly in suburban areas, in which the transportation accessibility had a more noticeable effect on property values. Therefore, it is suggested that the authorities emphasise the Transit-Oriented Development (TOD) model by planning high-density mixed-use areas around the stations, integrating commercial, office, and residential spaces. Additionally, the coverage of public transportation must be increased in the southern and western areas, where transport options are relatively weak, to improve accessibility.

There are significant differences in the demand for public amenities across the study area. Residents in the urban core are more focused on leisure and entertainment facilities (e.g., parks and shopping centres), while those in suburban areas place more importance on basic living needs (e.g., public transportation). Therefore, regulatory bodies should implement differentiated strategies in urban planning and land management, adjusting the allocation of such amenities based on the region characteristics. In urban core areas, the priority should be enhancing the quality and variety of entertainment facilities (e.g., city parks and cultural venues). In suburban areas, the focus should be on improving essential services, such as adding subway stations.

6. Conclusions

The present study addressed the research question of how service amenities and public transportation factors influence house prices across spatially autocorrelated neighbourhood clusters to discuss how this might generalise in big cities. As the economic centre of China, Shanghai City is used as the test area. The analysis is designed into two stages, among which the first stage aims to determine the spatial cluster pattern of neighbourhoods and house prices in Shanghai. The results from Moran's I index confirmed significant positive spatial autocorrelation in Shanghai neighbourhoods. Secondly, using Local Moran's I test to cluster neighbourhoods that have positive autocorrelation on house price, we found there are 24 neighbourhoods in the high–high house price clusters (i.e., neighbourhoods that are both expensive and have high house price neighbourhoods) and 20 neighbourhoods in low–low cluster. The high–high cluster is in the centre of Shanghai, and the significant low–low cluster is located in the southern part of Shanghai. Moreover, by analysing the average number of public facilities, we found that there are differences in the average amount of facilities in different clusters.

The second of our work examined a finer scale to determine the correlation between built environment factors and property developments. We asked: how does this correlation change across different areas in Shanghai? We used OLS and GWR models, of which

the former is used to test the variables' significance level. The R^2 for the GWR model is between 0.435 and 0.566. There are seven significant variables in the model, four of which are built environment factors, namely the distance to underground stations, the distance to railway stations, the distance to parks, and the distance to shopping malls. The result shows that different variables have different correlation coefficients across Shanghai and their significance spatially autocorrelated clusters.

Overall, being closer to a shopping centre is the only factor that can positively impact house prices for the whole city, and it has the most significant premium (0.357) by the absolute number. The rest of the built environment factors, including public transportation accessibility and park accessibility, have different impacts spatially. In the case of the high–high cluster, because it is considered to have better accessibility to the surrounding amenities and rich public transportation facilities, the correlation between the proximity to public transportation (underground stations and high-speed railway) is positive, showing that easier access to public transport cannot increase the property prices in HH clustered areas correspondingly. However, the house prices in this area are affected by the accessibility to service amenities more positively. The low–low cluster always shows some different trends from the HH cluster. One of the most notable features is the greater impact of public transportation than the service amenities on the house price: In these areas, the closer a house is to an underground station, the higher that house's price will be. Furthermore, although decreasing the distance to the shopping mall can also increase house prices in this cluster, the premium of the shopping centre on house prices in the HH cluster will still be higher than it is in the LL cluster. In fact, the magnitude of the coefficients in the LL areas did not stand out across Shanghai but always aligned with the general trend of the rest of Shanghai besides the urban core.

Our study provides insights into the patterns of Shanghai neighbourhood clusters under the view of spatial autocorrelation and heterogeneity. It shows how different built environment variables will affect the housing market differently based on the type of places and provides insights to guide policymakers in the planning of services amenities to maximise the value of local housing market.

Author Contributions: Conceptualisation, Y.Z. and B.B.; methodology, Y.Z. and B.B.; validation, Y.Z.; formal analysis, Y.Z.; resources, Y.Z.; data curation, Y.Z.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z. and B.B.; visualisation, Y.Z.; supervision, B.B.; project administration, Y.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original data presented in the study are openly available in China Real Estate Information Corporation (CRIC) database at https://vip.cric.com/gateway, accessed on 10 September 2024.

Acknowledgments: We acknowledge the dataset of housing price and neighbourhood distribution provided by Weiqiao Hua.

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

References

1. Zhu, Y.; Lin, B. Sustainable housing and urban construction in China. *Energy Build.* **2004**, *36*, 1287–1297. [CrossRef]
2. Chen, J.; Hu, M.; Lin, Z. China's Housing Reform and Labor Market Participation. *J. Real Estate Financ. Econ.* **2023**, *67*, 218–242. [CrossRef]
3. Chen, Z.; Wang, C. Effects of intervention policies on speculation in housing market: Evidence from China. *J. Manag. Sci. Eng.* **2022**, *7*, 233–242. [CrossRef]
4. Wang, S.Y. Credit Constraints, Job Mobility, and Entrepreneurship: Evidence from a Property Reform in China. *Rev. Econ. Stat.* **2012**, *94*, 532–551. [CrossRef]

5. Zou, Y. Paradigm shifts in China's housing policy: Tug-of-war between marketization and state intervention. *Land Use Policy* **2022**, *122*, 106387. [CrossRef]
6. Qian, X.; Qiu, S.; Zhang, G. The impact of COVID-19 on housing price: Evidence from China. *Financ. Res. Lett.* **2021**, *43*, 101944. [CrossRef] [PubMed]
7. Li, D.K.; Mei, C.L.; Wang, N. Tests for spatial dependence and heterogeneity in spatially autoregressive varying coefficient models with application to Boston house price analysis. *Reg. Sci. Urban Econ.* **2019**, *79*, 103470. [CrossRef]
8. Long, Y.; Wu, Y.; Huang, L.; Aleksejeva, J.; Iossifova, D.; Dong, N.; Gasparatos, A. Assessing urban livability in Shanghai through an open source data-driven approach. *NPJ Urban Sustain.* **2024**, *4*, 7. [CrossRef]
9. He, C.; Lv, J.; Li, D.; Wu, S. How public emergencies impact Chinese households' housing consumption: Evidence from the COVID-19 pandemic. *J. Hous. Built Environ.* **2024**, 1–25. [CrossRef]
10. Tang, W.; Wang, Y. Incomplete information and real estate development strategy: Evidence from Hangzhou, China. *Habitat Int.* **2017**, *63*, 1–10. [CrossRef]
11. Tam, V.W.Y.; Fung, I.W.H.; Wang, J.; Ma, M. Effects of locations, structures and neighbourhoods to housing price: An empirical study in Shanghai, China. *Int. J. Constr. Manag.* **2022**, *22*, 1288–1307. [CrossRef]
12. Zhang, L.; Zhu, L.; Shi, D.; Hui, E.C.M. Urban residential space differentiation and the influence of accessibility in Hangzhou, China. *Habitat Int.* **2022**, *124*, 102556. [CrossRef]
13. Favara, G.; Imbs, J. Credit supply and the price of housing. *Am. Econ. Rev.* **2015**, *105*, 958–992. [CrossRef]
14. Hilber, C.A.L.; Vermeulen, W. The impact of supply constraints on house prices in England. *Econ. J.* **2014**, *126*, 358–405.
15. Basu, S.; Thibodeau, T.G. Analysis of Spatial Autocorrelation in House Prices. *J. Real Estate Financ. Econ.* **1998**, *17*, 61–85. [CrossRef]
16. Bourassa, S.C.; Cantoni, E.; Hoesli, M. Spatial dependence, housing submarkets, and house price prediction. *J. Real Estate Financ. Econ.* **2007**, *35*, 143–160. [CrossRef]
17. Chen, S.; Zhuang, D.; Zhang, H. GIS-Based Spatial Autocorrelation Analysis of Housing Prices Oriented towards a View of Spatiotemporal Homogeneity and Nonstationarity: A Case Study of Guangzhou, China. *Complexity* **2020**, *1*, 1079024. [CrossRef]
18. Sakuragawa, M.; Tobe, S.; Zhou, M. Chinese housing market and bank credit. *J. Asian Econ.* **2021**, *76*, 101361. [CrossRef]
19. Shatkin, G. Global cities of the South: Emerging perspectives on growth and inequality. *Cities* **2007**, *24*, 1–15. [CrossRef]
20. Liu, C.; Song, W. Perspectives of Socio-Spatial Differentiation from Soaring Housing Prices: A Case Study in Nanjing, China. *Sustainability* **2019**, *11*, 2627. [CrossRef]
21. Deng, L.; Shen, Q.; Wang, L. The Emerging Housing Policy Framework in China. *J. Plan. Lit.* **2011**, *26*, 168–183. [CrossRef]
22. Liang, C.; Liu, Y.; Qiu, T.; Jing, Y.; Fang, F. The effects of locational factors on the housing prices of residential communities: The case of Ningbo, China. *Habitat Int.* **2018**, *81*, 1–11. [CrossRef]
23. Haila, A. Real estate in global cities: Singapore and Hong Kong as property states. *Urban Stud.* **2000**, *37*, 2241–2256. [CrossRef]
24. Wu, F. Sociospatial Differentiation in Urban China: Evidence from Shanghai's Real Estate Markets. *Environ. Plan. A Econ. Space* **2002**, *34*, 1591–1615. [CrossRef]
25. Wen, H.; Tao, Y. Polycentric urban structure and housing price in the transitional China: Evidence from Hangzhou. *Habitat Int.* **2015**, *46*, 138–146. [CrossRef]
26. Wang, Y.; Feng, S.; Deng, Z.; Cheng, S. Transit premium and rent segmentation: A spatial quantile hedonic analysis of Shanghai Metro. *Transp. Policy* **2016**, *51*, 61–69. [CrossRef]
27. Wu, J.; Deng, Y.; Liu, H. House Price Index Construction in the Nascent Housing Market: The Case of China. *J. Real Estate Financ. Econ.* **2014**, *48*, 522–545. [CrossRef]
28. Wu, J.; Gyourko, J.; Deng, Y. Evaluating the risk of Chinese housing markets: What we know and what we need to know. *China Econ. Rev.* **2016**, *39*, 91–114. [CrossRef]
29. Wen, H.; Jin, Y.; Zhang, L. Spatial heterogeneity in implicit housing prices: Evidence from Hangzhou, China. *Int. J. Strateg. Prop. Manag.* **2017**, *21*, 15–28. [CrossRef]
30. Huang, Z.; Chen, R.; Xu, D.; Zhou, W. Spatial and hedonic analysis of housing prices in Shanghai. *Habitat Int.* **2017**, *67*, 69–78. [CrossRef]
31. Li, H.; Chen, P.; Grant, R. Built environment, special economic zone, and housing prices in Shenzhen, China. *Appl. Geogr.* **2021**, *129*, 102429. [CrossRef]
32. Traun, C.; Mayrhofer, C. Complexity reduction in choropleth map animations by autocorrelation weighted generalization of time-series data. *Cartogr. Geogr. Inf. Sci.* **2018**, *45*, 221–237. [CrossRef]
33. Dong, Y.H.; Peng, F.L.; Li, H.; Men, Y.Q. Spatial autocorrelation and spatial heterogeneity of underground parking space development in Chinese megacities based on multisource open data. *Appl. Geogr.* **2023**, *153*, 102897. [CrossRef]
34. Soleimani, M.; Bagheri, N. Spatio-temporal analysis of head injuries in northwest Iran. *Spat. Inf. Res.* **2023**, *31*, 329–344. [CrossRef]
35. Xue, M.; Zhong, J.; Gao, M.; Pan, R.; Mo, Y.; Hu, Y.; Du, J.; Huang, Z. Analysis of spatial–temporal dynamic distribution and related factors of tuberculosis in China from 2008 to 2018. *Sci. Rep.* **2023**, *13*, 4974. [CrossRef]

36. Li, Y.; Teng, M.; Liu, Y. Spatial dynamics of transport accessibility and regional efficiency in the Yangtze River Delta Urban Agglomeration. *Sci. Rep.* **2024**, *14*, 19628. [CrossRef] [PubMed]
37. Rezaee, M.; Alamdar, M.; Taheri, E.; Badiee Azandahi, M. Exploring spatial analysis of the voting patterns in the Afghanistan president elections of 2019. *GeoJournal* **2024**, *89*, 146. [CrossRef]
38. Wang, W.C.; Chang, Y.J.; Wang, H.C. An Application of the Spatial Autocorrelation Method on the Change of Real Estate Prices in Taitung City. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 249. [CrossRef]
39. Libório, M.P.; de Abreu, J.F.; Ekel, P.I.; Machado, A.M.C. Effect of sub-indicator weighting schemes on the spatial dependence of multidimensional phenomena. *J. Geogr. Syst.* **2023**, *25*, 185–211. [CrossRef]
40. Barjoee, S.S.; Azizi, M.; Khaledi, A.; Kouhkan, M.; Soltani, M.; Farokhi, H. Street dust-bound metal(loid)s in industrial areas of Iran: Moran's spatial autocorrelation distribution, eco-toxicological risk assessment, uncertainty and sensitivity analysis. *Int. J. Environ. Sci. Technol.* **2023**, *20*, 8509–8536. [CrossRef]
41. Anselin, L. Local indicators of spatial association: LISA. *Geogr. Anal.* **1995**, *27*, 93–115. [CrossRef]
42. Gozalo, G.R.; Morillas, J.M.B.; González, D.M. Perceptions and use of urban green spaces on the basis of size. *Urban For. Urban Green.* **2019**, *46*, 126470. [CrossRef]
43. Sisman, S.; Aydinoglu, A.C. A modelling approach with geographically weighted regression methods for determining geographic variation and influencing factors in housing price: A case in Istanbul. *Land Use Policy* **2022**, *119*, 106183. [CrossRef]
44. Wang, D.; Li, V.J.; Yu, H. Mass Appraisal Modeling of Real Estate in Urban Centers by Geographically and Temporally Weighted Regression: A Case Study of Beijing's Core Area. *Land* **2020**, *9*, 143. [CrossRef]
45. Guo, Z.; Tian, B.; Zhu, Y.; He, J.; Zhang, T. How do the landslide and non-landslide sampling strategies impact landslide susceptibility assessment?—A case study at catchment scale from China. *J. Rock Mech. Geotech. Eng.* **2024**, *16*, 877–894. [CrossRef]
46. Guo, Z.; Tian, B.; He, J.; Xu, C.; Zeng, T.; Zhu, Y. Hazard assessment for regional typhoon-triggered landslides by using physically-based model -A case study from southeastern China. *Georisk Assess. Manag. Risk Eng. Syst. Geohazards* **2023**, *17*, 740–754. [CrossRef]
47. Morawakage, P.S.; Earl, G.; Liu, B.; Roca, E.; Omura, A. Housing Risk and Returns in Submarkets with Spatial Dependence and Heterogeneity. *J. Real Estate Financ. Econ.* **2023**, *67*, 695–734. [CrossRef]
48. Chae, J.; Bera, A.K. Spatial Market Inefficiency in Housing Market: A Spatial Quantile Regression Approach. *J. Real Estate Financ. Econ.* **2024**, *69*, 70–99. [CrossRef]
49. Hyun, D.; Lee, H.K. Measuring spatial heterogeneity of air quality on apartment transaction prices in Seoul, South Korea. *Asia-Pac. J. Reg. Sci.* **2024**, *8*, 681–703. [CrossRef]
50. Elhorst, J.P. Applied spatial econometrics: Raising the bar. *Spat. Econ. Anal.* **2010**, *5*, 9–28. [CrossRef]
51. Lu, B.; Ge, Y.; Shi, Y.; Zheng, J.; Harris, P. Uncovering drivers of community-level house price dynamics through multiscale geographically weighted regression: A case study of Wuhan, China. *Spat. Stat.* **2023**, *53*, 100723. [CrossRef]
52. Tasyurek, M.; Celik, M. 4D-GWR: Geographically, altitudinal, and temporally weighted regression. *Neural Comput. Appl.* **2022**, *34*, 14777–14791. [CrossRef]
53. Yum, S. Analyses of human responses to Winter storm Kai using the GWR model. *Nat. Hazards* **2023**, *116*, 1805–1821. [CrossRef]
54. Wheeler, D.; Tiefelsdorf, M. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *J. Geogr. Syst.* **2005**, *7*, 161–187. [CrossRef]
55. Fotheringham, A.S. The Problem of Spatial Autocorrelation and Local Spatial Statistics. *Geogr. Anal.* **2009**, *41*, 398–403. [CrossRef]
56. Oshan, T.; Li, Z.; Kang, W.; Wolf, L.J.; Fotheringham, A. mgwr: A Python Implementation of Multiscale Geographically Weighted Regression for Investigating Process Spatial Heterogeneity and Scale. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 269. [CrossRef]
57. Guo, Z.; Wang, H.; He, J.; Huang, D.; Song, Y.; Wang, T.; Liu, Y.; Ferrer, J.V. PSLSA v2.0: An automatic Python package integrating machine learning models for regional landslide susceptibility assessment. *Environ. Model. Softw.* **2025**, *186*, 106367. [CrossRef]
58. Li, H.; Wei, Y.D.; Wu, Y.; Tian, G. Analyzing housing prices in Shanghai with open data: Amenity, accessibility and urban structure. *Cities* **2019**, *91*, 165–179. [CrossRef]
59. Zou, Y.; Zhong, N.; Chen, H.; Pu, C. Housing burden across Chinese cities: Spatio-temporal patterns and influential factors. *Appl. Geogr.* **2024**, *166*, 103259. [CrossRef]
60. Qiu, R.; Xu, W. Modes of land development in Shanghai. *Land Use Policy* **2017**, *61*, 475–486. [CrossRef]
61. He, S.; Wu, F. China's emerging neoliberal urbanism: Perspective from urban redevelopment. *Antipode* **2009**, *41*, 282–304. [CrossRef]
62. Yang, Y.R.; Chang, C. An urban regeneration regime in China: A case study of urban redevelopment in Shanghai's Taipingqiao Area. *Urban Stud.* **2007**, *44*, 1809–1826. [CrossRef]
63. Amorim, J.M.; Silva, J.A. The influence of residential location and public transit options on commuters' intention to use travel apps for different travel purposes. *Travel Behav. Soc.* **2025**, *38*, 100913. [CrossRef]
64. Tian, G.; Wei, Y.D.; Li, H. Combined effects of accessibility and environmental health risk on housing price: A case of Salt Lake County, UT. *Appl. Geogr.* **2017**, *89*, 12–21. [CrossRef]

65. Rosiers, F.D.; Laganà, A.; Thériault, M.; Beaudoin, M. Shopping centres and house values: An empirical investigation. *J. Prop. Valuat. Invest.* **1996**, *14*, 41–62. [CrossRef]
66. Hui, E.C.M.; Chau, C.K.; Pun, L.; Law, M.Y. Measuring the neighboring and environmental effects on residential property value: Using spatial weighting matrix. *Build. Environ.* **2007**, *42*, 2333–2343. [CrossRef]
67. Liu, F.; Min, M.; Zhao, K.; Hu, W. Spatial-Temporal Variation in the Impacts of Urban Infrastructure on Housing Prices in Wuhan, China. *Sustainability* **2020**, *12*, 1281. [CrossRef]
68. Yuan, H.; Yang, X. Spatial distribution of urban basic education resources in Shanghai: Accessibility and supply-demand matching evaluation. *Open Geosci.* **2023**, *15*, 579–618. [CrossRef]
69. Song, Z. Disparity in Educational Resources Between Urban and Rural Areas in China. *J. Adv. Res. Educ.* **2023**, *2*, 64–69. [CrossRef]
70. Huang, J.; Wang, Y.; Wu, K.; Yue, X. Livability-oriented urban built environment: What kind of built environment can increase the housing prices? *J. Urban Manag.* **2024**, *13*, 357–371. [CrossRef]
71. Yuan, F.; Wu, J.; Wei, Y.D.; Wang, L. Policy change, amenity, and spatiotemporal dynamics of housing prices in Nanjing, China. *Land Use Policy* **2018**, *75*, 225–236. [CrossRef]
72. Qu, W.; Huang, Y.; Deng, G. Identifying the critical factors behind the second-hand housing price concession: Empirical evidence from China. *Habitat Int.* **2021**, *117*, 102442. [CrossRef]
73. Wang, Z.; Zhang, Q. Fundamental factors in the housing markets of China. *J. Hous. Econ.* **2014**, *25*, 53–61. [CrossRef]
74. Liang, J.; Huang, S.; Bao, W.; Pu, Y.; Liu, M.; Yang, X.; Wu, Z.; Chen, J.; Dong, G.; Chen, Y. Longitudinal relationship between built environment and physical fitness in Chinese children and adolescents: Findings from the COHERENCE study. *Build. Environ.* **2025**, *267*, 112216. [CrossRef]
75. Huang, H. Creating sustainable urban built environments: An application of hedonic house price models in Wuhan, China. *J. Hous. Built Environ.* **2015**, *30*, 219–235. [CrossRef]
76. Chikhmous, A.; Rahman, M.T. Examining the effect of apartment attributes on their sale prices in Riyadh, Saudi Arabia. *Spat. Inf. Res.* **2024**, *32*, 411–424. [CrossRef]
77. Lu, B.; Charlton, M.; Harris, P.; Fotheringham, A. Geographically weighted regression with a non-Euclidean distance metric: A case study using hedonic house price data. *Int. J. Geogr. Inf. Sci.* **2014**, *28*, 660–681. [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.