International Journal of Geo-Information

MDPI

Article

Simulating the Spatial Heterogeneity of Housing Prices in Wuhan, China, by Regionally Geographically Weighted Regression

Zengzheng Wang [1], Yangyang Zhao [2,*] and Fuhao Zhang [1,2]

[1] Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu 610032, China; wzz_giser@my.swjtu.edu.cn (Z.W.); zhangfh@casm.ac.cn (F.Z.)
[2] Research Centre of Geospatial Big Data Application, Chinese Academy of Surveying and Mapping, Beijing 100830, China
* Correspondence: zhaoyy@casm.ac.cn

Abstract: Geographically weighted regression (GWR) is an effective method for detecting spatial non-stationary features based on the hypothesis of proximity correlation. In reality, especially in the social and economic fields, research objects not only have spatial non-stationary characteristics, but also spatial discrete heterogeneity characteristics. Therefore, how to improve the accuracy of GWR estimation in this case is worth studying. In this paper, a regionally geographically weighted regression (RGWR) is proposed. Using incoming dummy variables, the zoning discrimination is added to the spatial kernel function of GWR, the spatial kernel function is modified, the spatial weight is optimized, and the influence of "near heterogeneous" observation points is reduced. In this paper, the residential sale price in Wuhan City is taken as an example in the analysis of three aspects: model performance, fitting effect and influencing factors. The results show that the introduction of a zoning dummy variables can significantly improve the model accuracy of a fixed bandwidth and adaptive bandwidth. Under a fixed bandwidth, compared with the GWR model, RGWR increases R^2 and R^2adj from 0.6776 and 0.6732 to 0.777 and 0.7746, respectively, and the Akaike information criterion, corrected (AICc) standard decreases by 37.4006 compared with GWR, which proves the effectiveness of the method.

Keywords: geographically weighted regression; regionally geographically weighted regression; spatial discrete heterogeneity

check for updates

Citation: Wang, Z.; Zhao, Y.; Zhang, F. Simulating the Spatial Heterogeneity of Housing Prices in Wuhan, China, by Regionally Geographically Weighted Regression. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 129. https://doi.org/10.3390/ijgi11020129

Academic Editor: Wolfgang Kainz

Received: 12 December 2021
Accepted: 9 February 2022
Published: 11 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 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

In early studies of spatial regression, models were applied at the global level, where data relationships were constant throughout the whole study region. However, this spatial stability hypothesis is usually invalid for non-stationary spaces, which generally manifest as uncontrolled spatial variability [1]. To account for this spatial heterogeneity, spatial statistics changed from global matching to a local model, and many local regression techniques were proposed. Casetti proposed the expansion method [2], Jone proposed the multilevel models [3], McMillen and McDonald proposed nonparametric local linear regression [4,5], Elhorst proposed the panel data model [6], and Brunsdon and Fotheringham proposed geographic weighted regression (GWR) [7–9]. GWR is built on the premise of Tobler's famous first law of geography, that "everything is related to everything else, but near things are more related than distant things" [10], effectively resolving spatial non-stationarity and detecting spatial heterogeneity. Many studies have focused on extending GWR to better detect spatial heterogeneity. Lu applied the non-Euclidean distance metric to the GWR model solution [11,12], Anselin L proposed heteroskedastic GWR [13], Harris P proposed robust GWR [14], Wang N proposed local linear estimation-based GWR [15], and Zhao proposed a GWR Based on semi-supervised learning [16]. These studies improve different aspects of model-solving accuracy, and scenarios such as a flexible distance metric, adaptation to heteroskedasticity, outliers, covariance, and Semi-Supervised Learning. They

also make it easier to detect heterogeneous features of spatial relationships due to their irregular distribution in geographic space [17].

Spatial heterogeneity is one important feature in geographical phenomena, and needs to be considered when quantitatively modeling the relationship between a response variable and explanatory variables in spatial data analysis [18]. The specification of spatial heterogeneity can be classified into continuous heterogeneity and discrete heterogeneity [19]. The former specifies how the regression coefficients change over space, either following a predetermined functional form, or as determined by the data through a local estimation process [20]. The latter consists of a prespecified set of spatially distinct units, or spatial regions, such as administrative units, differences in the population densities in different areas, climates or ecological zones, and the distribution of soil types, land use, and land cover [21,22]. Model coefficients and other parameters are also allowed to vary among the regions. GWR is a major paradigm for spatial modeling to reveal spatial continuous heterogeneity. Existing GWR models can detect spatial continuous heterogeneity to some extent in practical applications by means of bandwidth optimization, but cannot solve the discrete heterogeneity [23].

In solving the problem of spatial discrete heterogeneity, the existing research on spatial analysis has conducted some explorations. Rich Harris pointed out that two points can be geographically close but socially distant because the contexts (or neighborhoods) within which they are situated are not alike [24]. This is slightly similar to spatial discrete heterogeneity. They proposed a contextualize GWR (CGWR) by incorporating contextual information in the GWR weights' matrix. However, the decay rate of spatial distance may be inconsistent with that of context variables, and spatial regions are not considered. In the early regression analysis, in order to detect spatial discrete heterogeneity, scholars used sub-area regress, which divided the research area into multiple different regions according to certain indicators, and then established a regression model in each region [25]. Some scholars used dummy variables to represent regions, and different regions are given different dummy variable values [26,27]. These effectively reflect the difference between regions, while a global regression in the same region is still unable to detect non-stationary spatial features.

The hedonic price model (HPM) is a widely used approach to studying housing prices [28,29]. It establishes a quantitative relationship between characteristics and housing prices. Such models regard houses as a composite commodity formed by structural attributes, neighborhood attributes, age of construction and other attributes. Structural attributes determine the basic functions of a house and have a large impact on its price. The housing area, age, the number of bedrooms, the residential plot ratio, residential greening ratio, and other factors are considered as structural attributes [30,31]. Neighborhood attributes reflect the accessibility of amenities and the socioeconomic status of communities. The influence of supermarkets, shopping centers, primary schools, and bus stations contribute greatly to explaining housing prices [32,33]. However, the essence of the HPM is consistent in the whole space. Since GWR and geographically and temporally weighted regression (GTWR) were used for real-estate research, the non-stationary nature of housing prices has been widely revealed [9,34]. An increasing number of studies have explored the non-stationary nature of housing prices from various perspectives [35–38]. In fact, housing prices not only have the characteristics of spatial non-stationary, but also spatial discrete heterogeneity. For example, in the context of the "nearby enrollment" policy of compulsory education in China, enrollment in primary and secondary schools is strictly in accordance with the school district where the child's registered residence is located. Many parents are, hence, willing to pay high housing prices for good schools, thereby leading to a soaring price of "school district housing" at present [39]. This phenomenon exists not only in China, but also in other countries and regions. According to the data of the UK Department of Education, the average price of performing schools' district housing is more than 18,600 GBP higher than that of non-school-district housing in the whole England. The average house price in the whole of England in 2016 was 233,000 pounds. House prices

near the 10% best-performing primary schools are 8.0% higher than that in the surrounding area. Near the 10% best-performing non-selective secondary schools, house prices are 6.8% higher [40]. Therefore, spatial discrete heterogeneity has a widespread existence in the field of house prices. Although the spatial attribute of housing prices is generally considered when constructing a hedonic price model, few studies have simultaneously investigated spatial non-stationary features and spatial discrete heterogeneity in depth.

The objective of this article is to extend the traditional GWR model to simultaneously detect the spatial non-stationary features and spatial discrete heterogeneity in housing prices. This study seeks to contribute to the literature on the topic in the following three ways. First, we propose regionally geographically weighted regression (RGWR), which uses dummy variables to reflect regional differences. Second, the algorithm flow and estimation steps of RGWR are illustrated. Third, we examine and compare the GWR and RGWR for modeling housing prices by means of a case study in the Wuhan city of China.

This paper is organized as follows: the RGWR methods are derived in Section 2; in Section 3, we carry out an experiment to assess the performance of the proposed method and make an empirical comparison with the basic GWR. Finally, the discussion and conclusions of the paper are reported.

2. Materials and Methods

2.1. Geographical Weight Regression

Geographical weight regression is an extension of ordinary linear regression. It is used to explore spatial non-stationary features, embed geographic location into regression parameters, and allow each individual point in a different location to have a different value to estimate regression parameters [9]. The model can be expressed as:

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

where (u_i, v_i) is the coordinate of the i-th sampling point and $\beta_k(u_i, v_i)$ is the k-th regression parameter at the i-th sampling point. When estimating the regression parameters of sampling point i, the importance of observations at different observation points is not the same. The closer the observation point is to point i, the higher the importance; the farther the observation point is, the lower the importance. Used the local weighted least squares approach, the regression parameter estimation $\hat{\beta}_i$ at point i is given by Formula (2):

$$\hat{\beta}_i = (\boldsymbol{X}'\boldsymbol{W}_i\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}_i y \tag{2}$$

The spatial weight matrix $\boldsymbol{W}_i$ is the core of the geographically weighted regression model; $\boldsymbol{W}_i$ is an n^*n matrix, which is calculated by the monotonously decreasing function of the geographic distance d_{ij} between the regression point i and other observation points j. The elements on the off-diagonal line of the matrix are zero, and the elements on the diagonal line represent the geographical weight between the regression point i and the observation point j, namely, $\boldsymbol{W}_i = diag(w_{i1}, w_{i2}, \ldots, w_{in})$.

The selection of the spatial kernel function is of great importance to the correct estimation of the parameters in GWR. The general spatial kernel function is of two types: the fixed kernel function and the adaptive kernel function. For the fixed kernel, distance is constant, but the number of nearest neighbors varies. For the adaptive kernel, distance varies but the number of neighbors remains constant [34]. The most common kernel is a Gaussian distance decay-based function. Its function form is as follows:

$$w_{ij} = exp(-\frac{d_{ij}^2}{b^2}) \tag{3}$$

where b is called bandwidth, which is a non-negative attenuation parameter calculated by the cross-validation (CV) approach [9]. The larger the bandwidth, the slower the weight

decays as the distance increases, and the smaller the bandwidth, the faster the weight decays as the distance increases.

2.2. Regionally and Geographically Weighted Regression

RGWR is an extension of GWR used to explore spatial non-stationary and spatially discrete heterogeneity by adding regional dummy variables to GWR that embed geographic location into the regression parameters, allowing each individual point to have different values to estimate the regression parameters, also affected by regional factors for each regression point. The model can be expressed as:

$$y_i = R\beta_0(u_i, v_i) + \sum_{k=1}^{p} R\beta_k(u_i, v_i)x_{ik} + \varepsilon_i \quad i = 1, 2, \ldots, n \tag{4}$$

where (u_i, v_i) is the coordinate of the i-th sampling point and $R\beta_k(u_i, v_i)$ is the k-th regression parameter of the i-th sampling point. In estimating the regression parameters of sampling point i, the regional factors are added on the basis of GWR's "the closer the observation point is to point i, the higher the importance, and vice versa [9]", so that when the observation point is outside the regional range, the point does not participate in the regression. Using the locally weighted least squares method, the regression parameter estimate $\widehat{R\beta_i}$ at point i is given by Equation (5):

$$\widehat{R\beta_i} = (\boldsymbol{X}'\boldsymbol{R}\boldsymbol{W}_i\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{R}\boldsymbol{W}_i y \tag{5}$$

The geographic weight between regression point i and observation point j is $\boldsymbol{RW}_i = diag(rw_{i1}, rw_{i2}, \ldots, rw_{in})$. rw_{ij} is the spatial kernel function of RGWR, and the solution of the spatial kernel function of RGWR involves the selection of RGWR bandwidth.

2.2.1. Spatial Kernel Function Calculation

GWR is essentially a partial regression, that is, it uses observation points within the bandwidth of the regression point to estimate the value of the regression point. The spatial weight effectively describes the observation point's degree of influence on the regression point as the distance changes. Sometimes, the distribution of observation points in the study area is not uniform. To ensure that a certain number of observation points participate in the fitting during the estimation of each regression point, Fotheringham proposed a fixed nuclear bandwidth and an adaptive bandwidth strategy [9]. In the fixed bandwidth strategy, the global bandwidth is a fixed value, and the number of observation points involved in the calculation is different when each regression point is estimated. Figure 1a shows a schematic diagram of the fixed bandwidth kernels. The most common adaptive bandwidth strategy is to fix the number of adjacent observation points; that is, take the number of observation points participating in the estimation of each regression point as a fixed value in the global scope, then the bandwidth will change with the change in the regression point. Figure 1b shows a schematic diagram of the adaptive bandwidth kernels. The former is suitable for a sample set with a relatively uniform spatial distribution, and the latter is suitable for a sample set with uneven spatial distribution.

The weight calculation strategy of GWR does not consider the regional factor. To better characterize the role of regional factors in the calculation of spatial weights, we assume that observation points located in the same region have the same region attributes, and observation points located in different regions have different regional attributes. Then, it is better to use homogeneous observation points for fitting. Therefore, when calculating the weight, the region attribution judgment is introduced; that is, when the observation point and the regression point are in the same region, the observation point participates in regression point estimation. When the observation point and the regression point are in different regions, the observation point does not participate in the regression point

estimation. Figure 1c shows a schematic diagram of the RGWR with fixed bandwidth kernels. Figure 1d shows a schematic diagram of the RGWR with adaptive spatial kernels.

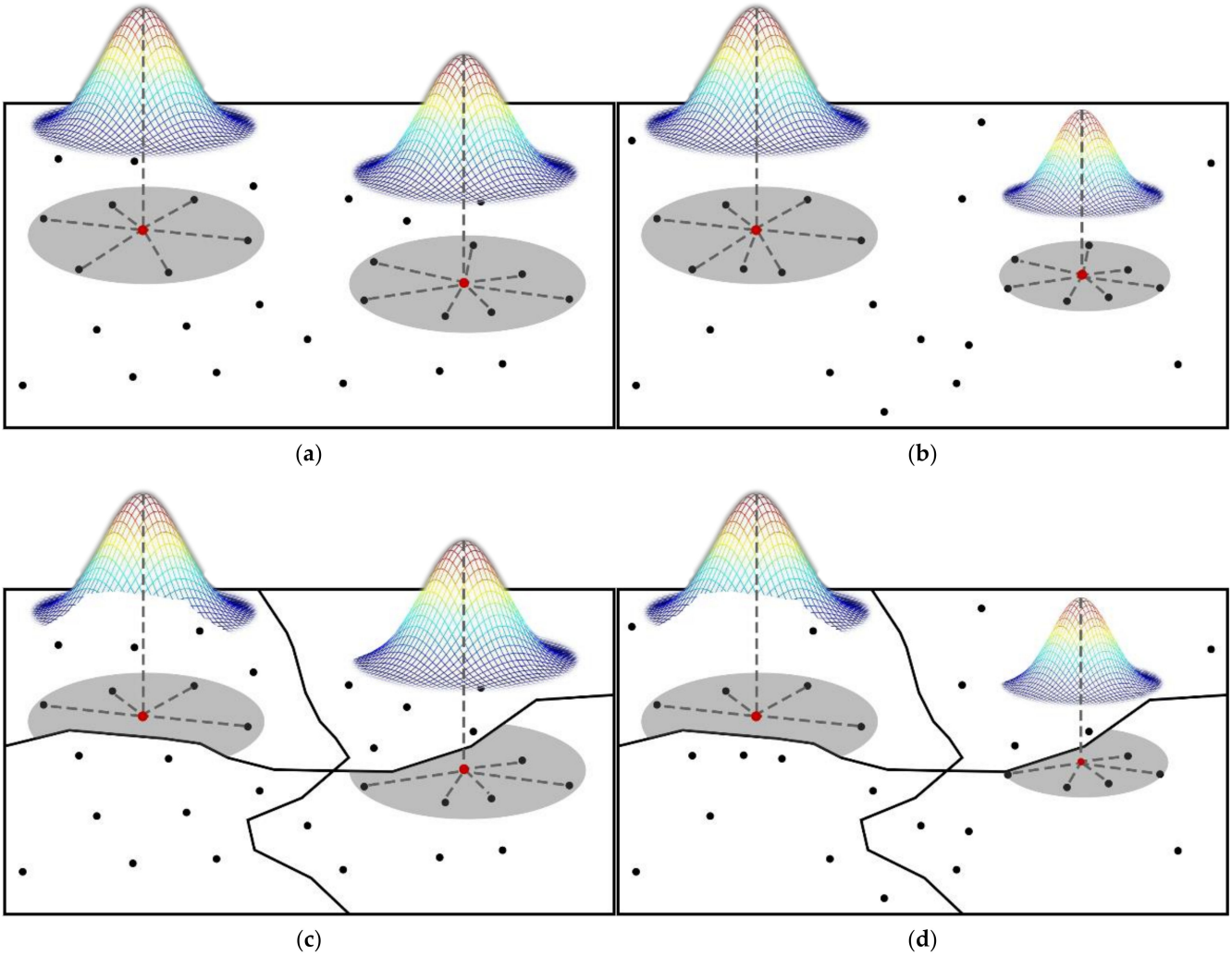


Figure 1. (**a**) GWR with fixed gaussian spatial kernels; (**b**) GWR with adaptive gaussian spatial kernels; (**c**) RGWR with fixed gaussian spatial kernels; (**d**) RGWR with adaptive gaussian spatial kernels.

Regionally geographically weighted regression adds the measurement factor of spatial regions to the traditional method of geographical weighted regression. In the weight function design, priority is given to the influence of zoning factors and then the influence of neighboring points. Corresponding to the Gauss kernel function and the bi-square kernel function, the expression of the kernel function based on the zoning is proposed.

The RGWR of the Gauss kernel function as:

$$rw_{ij} = \delta exp(-(d_{ij}/b)^2) \tag{6}$$

where b is the bandwidth, d_{ij} is the distance between point i and point j, and δ is the introduced dummy variable: the regional influence factor. When i and j are located in the same region, $\delta = 1$, which is the traditional GWR. When i and j are located in different regions, $\delta = 0$, which means that $rw_{ij} = 0$, which is the RGWR.

In addition, the bi-square kernel function is also a commonly used weight calculation method for GWR. RGWR of the bi-square kernel function is given as:

$$rw_{ij} = \begin{cases} \delta\left[1 - (d_{ij}/b)^2\right]^2 & d_{ij} \le b \\ 0 & d_{ij} > b \end{cases} \tag{7}$$

2.2.2. Bandwidth Selection

Bandwidth selection adopts Akaike information criterion, corrected (AICc) [9]. The optimal bandwidth is selected through trials: In each trial, a bandwidth is selected, RGWR is fitted using the bandwidth, and then a goodness-of-fit measure such as AICc is calculated, where AICc is defined by:

$$\mathrm{AICc} = 2n\ln(\hat{\sigma}) + n\ln(2\pi) + n\left[\frac{n + tr(S)}{n - 2 - tr(S)}\right] \tag{8}$$

where $\hat{\sigma}$ is the estimated standard deviation of the error term and $tr(S)$ is the trace of the hat matrix $\boldsymbol{S}$. The optimal bandwidth is that which minimizes AICc.

$$\hat{\sigma} = \frac{RSS}{n - tr(S)} \tag{9}$$

2.2.3. Algorithm Process

The algorithmic flow of RGWR is given in Figure 2. RGWR estimation is divided into two parts: one is the selection of the optimal bandwidth and the other is the parameter estimation, i.e., estimation of regression coefficients, fitted values, and evaluation metrics.

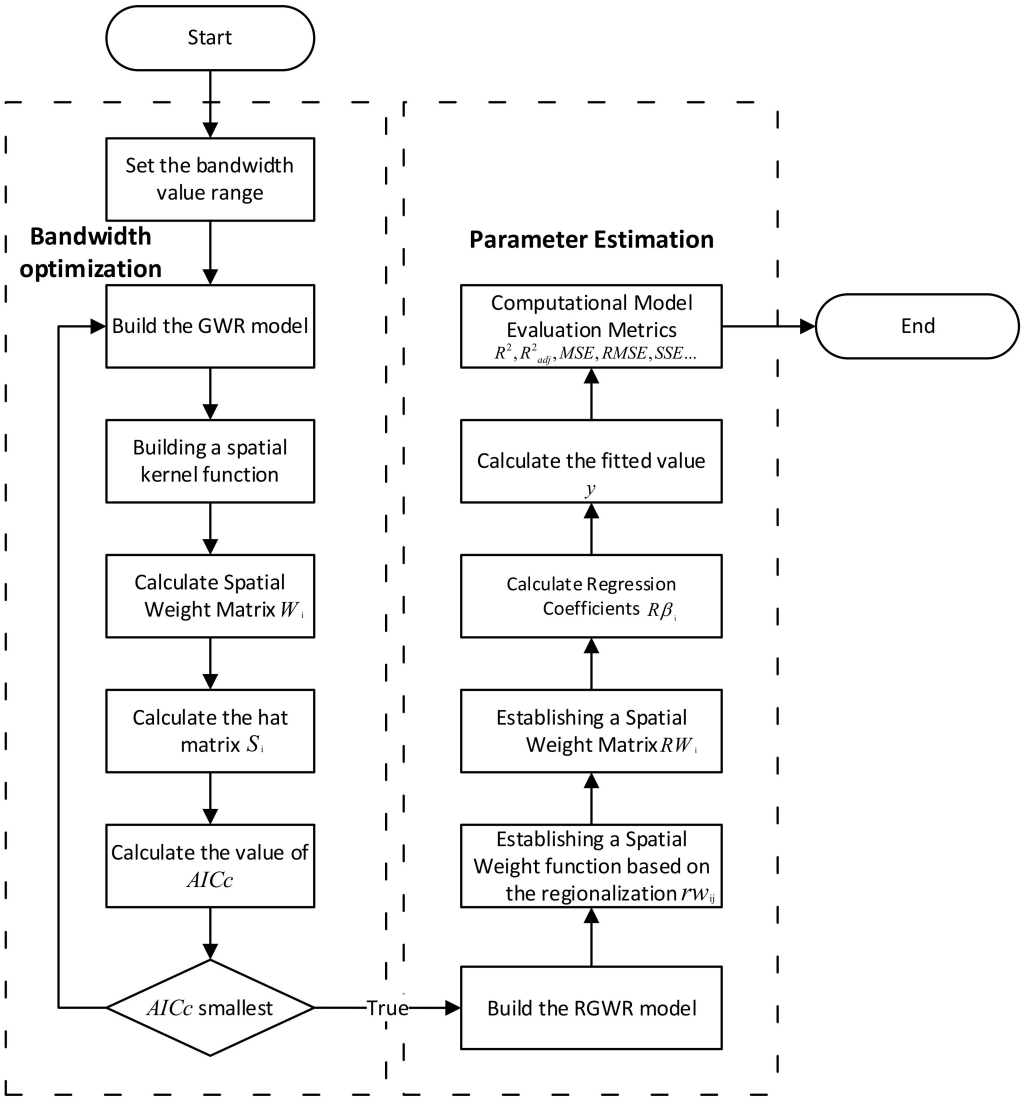


Figure 2. Algorithm process of RGWR.

The data include independent variables, dependent variables, spatial location variables, alternative bandwidths, and regional impact factors. The step-by-step process is as follows:

1. Data initialization. Set the range of bandwidth values.
2. Perform steps 3–7 for each bandwidth. At the end, perform step 8.
3. Build the GWR model using independent variables, dependent variables, spatial location variables, and bandwidths.
4. Construct a spatial kernel function for each observation using spatial location variables and bandwidth.
5. Calculate the spatial weight matrix for each observation point.
6. Calculate the Hat Matrix by using the independent variables, dependent variables, and spatial weight matrices.
7. Calculate the AICc values of the models corresponding to this set of bandwidths.
8. Select the parameters of the model corresponding to the minimum AICc value, which is the optimal bandwidth.
9. Build the RGWR model using the independent variables, dependent variables, spatial location variables, and the optimal bandwidth.
10. Construct a regional geographically weighted, spatial kernel function using spatial location variables, bandwidth, and regional factors.
11. Calculate the regional geographically weighted spatial weight matrix for each observation.
12. Calculate model regression coefficients, fitted values, and evaluation metrics by using the independent variables, dependent variables, and spatial weight matrices.

3. Experiment Analysis

In this section, the research area and related research data were selected and preprocessed; then, GWR and RGWR models were used to estimate the data under different bandwidth strategies. Then, depending on the estimation results, the model performances of GWR and RGWR under different bandwidth strategies were compared. The next step is to compare the fitting effects of GWR and RGWR under different bandwidth strategies. Finally, the main factors affecting the price of commercial housing in Wuhan were analyzed.

3.1. Study Area and Data

Wuhan is the capital of Hubei Province in central China, and serves as its political, economic, and cultural center. Wuhan consists of thirteen administrative regions: Jangan, Janghan, Qiaokou, Hanyang, Wuchang, Qingshan, Hongshan, Dongxihu, Hannan, Caidian, Jangxia, Huangpi, and Xinzhou. The housing market is one of the most active markets in China and plays a crucial role in China's economy [41–45]. The real estate prices of Wuhan are increasing at an alarming rate, associated with rapid industrialization and urbanization and consequent demands for various categories of real estate. A spatial heterogeneity analysis of real estate prices is considered crucial for revealing major issues in real estate market development, understanding effective strategies of economic macro control, and promoting the high-quality development of internal economics [46–49]. This paper uses the listed residential sale prices of Wuhan City, Hubei Province, China as characteristic price data, constructs a characteristic price model [30,50–54], and conducts an experimental analysis. The characteristic price model is used to describe the quantitative relationship between house characteristics and housing prices. Research has found that the price of commercial houses is related to factors such as house structure, surroundings, geographic location, and construction time [55]. House structure includes factors such as the indoor area and construction time, and the surroundings, such as the plot ratio, greening rate, and the distance from elementary schools and shopping malls. This paper collected 954 communities in the urban area of Wuhan as sampling points and obtained the average listing price (yuan/square meter) and construction time of each community in December 2019. Their geographic locations are shown in Figure 3. At the same time, data on points of interest were collected in Wuhan, such as data on bus stations, subway

stations, hospitals, parks, shopping malls, primary schools, middle schools, universities, fire protection, public security, and so on. Among them, characteristic house price data are from Anjuke (https://www.anjuke.com/)(last accessed on 29 June 2020), and the point of interest data are from Gaode (https://www.gaode.com/)(last accessed on 1 March 2021). In addition, the base map data of Wuhan's administrative area, including the main roads and waters, came from the China Map Publishing House, and the list of provincial model schools in Wuhan came from the Hubei Provincial Department of Education. The numerical ranking of each district is shown in Table 1.

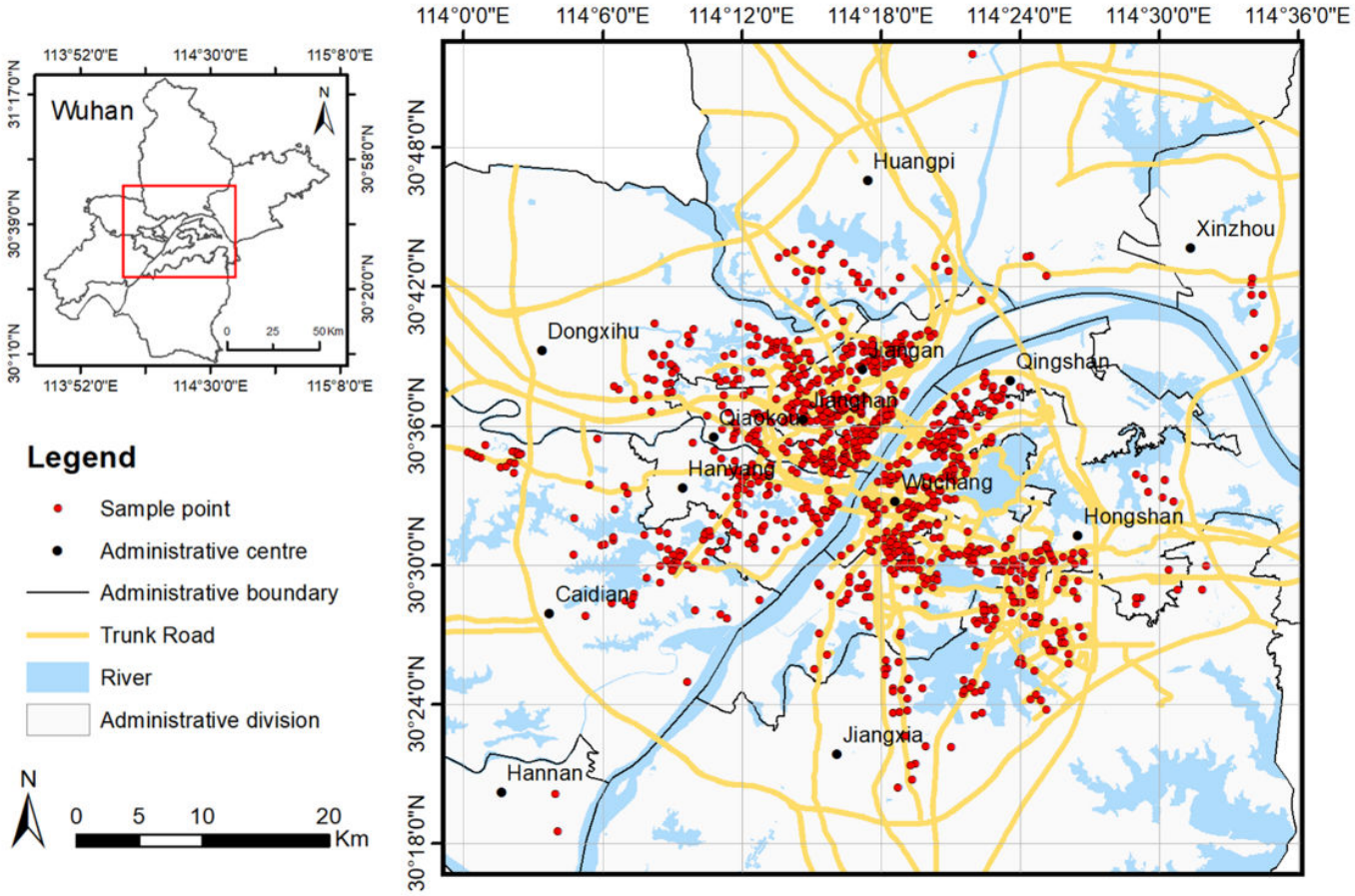


Figure 3. The distribution of sampling points in Wuhan.

Table 1. Demonstrative list of schools in various regions of Wuhan.

Region	Number of Provincial Demonstration Primary Schools	Number of Provincial Demonstration High Schools
Jangan	2	4
Janghan	1	4
Qiaokou	1	3
Hanyang	0	2
Wuchang	5	8
Qingshan	1	2
Hongshan	2	3
Dongxihu	0	2
Hannan	0	1
Caidian	0	2
Jiangxia	0	1
Xinzhou	0	1
Huangpi	0	1

3.2. Data Preprocessing

We extracted administrative regionalization, structural, neighborhood, and temporal variables to explain the variations in house prices. Before building the model, we preprocessed the data, using a logarithm operation on continuous variables. Overlay analysis was used to obtain the regional relationship between each residential plot and the administrative regions of Wuhan. Finally, we used a multicollinearity analysis and stepwise regression analysis to determine the independent variable factors.

An overview of the variables involved in housing prices is given in Table 2. The dependent variable is the sale price of the house. Unit prices are calculated in RMB. The structural characteristics of each house are described by three covariates. The natural logarithm of explanatory variables was used [34,56,57]. The plot ratio, also called the floor area ratio (FAR), is the ratio of a building's total floor area (gross floor area) to the size of the piece of land on which it is built. FAR is logarithmically transformed into LnFAR. The green ratio is the ratio of green space to the entire plot area. The green ratio is logarithmically transformed into LnGreenRatio. The management fee of the property (in RMB/m^2) is logarithmically transformed into LnPropertyFees. The temporal variable is the age of the building at the time of its building year. We record the earliest construction year as the base number 1, and the count is increased each year thereafter. We calculate the distance from each residential plot to the nearest primary school, subway station, junior high school, and university (LnPriSchool, LnSubway, LnHighSchool, LnUniversity).

Table 2. Variables used to predict housing prices in Wuhan, Hubei, China.

Abbreviation	Description	Minimum	Mean	Maximum
LnPrice	price of the house	−0.8144	0.5566	1.8498
Administrative Regionalization				
RegionalizationCode	Corresponding administrative regionalization code	-	-	-
Structural Covariates				
LnFAR	Log of the FAR	0.9516	−0.9163	3.0819
LnGreenRatio	Log of the green ratio	−2.3126	−1.077	−0.4155
LnPropertyFees	Log of the property fees	−2.3026	0.2813	1.6487
Temporal Covariates				
NormalAge	Normalized building age	1	21.6033	33
Neighborhood Covariates				
LnPriSchool	Log of the distance to the nearest primary school	6.2563	0	8.694
LnSubway	Log of the distance to the nearest subway station	3.0910	6.6235	9.9936
LnHighSchool	Log of the distance to the nearest junior high school	4.1109	6.5419	8.6477
LnUniversity	Log of the distance to the nearest university	1.6094	6.8148	8.9642

3.3. Comparison of the RGWR and GWR Models

3.3.1. Parameter Setting

We used administrative region to constrain the calculation range of geographically weighted regression. For sampling points in the same regions, geographically weighted regression calculations were normally used. Sampling points belonging to different regions were not considered in the scope of geographically weighted regression.

According to the algorithm process described in Section 2.2.3 of this paper, we used GWR and RGWR methods to establish characteristic price models. First, the AICc method

was used to determine the optimal bandwidth of GWR; the fixed type was 8516.9 and the adaptive type was 155. Then, the characteristic price models of GWR and RGWR were established using the optimal bandwidth, while hypothesis tests were performed for spatial stationarity under different bandwidth strategies [7,9,23]. The *p*-values of the hypothesis tests are all less than 1%, which is statistically significant. The spatial non-stationary characteristic test was carried out for the regression coefficient [23], as shown in Table 3, the results show that property fees, greening rate, FAR, subway station, primary school, and junior high schools all have spatial non-stationarity. Finally, the model regression coefficients, fitting values, and evaluation metrics indexes of GWR and RGWR were obtained under different bandwidth strategies. Taking GWR as the comparison method, the experiment analyzes the performance of the model, the fitting effect of the model and the influencing factors for house prices in Wuhan.

Table 3. The spatial non-stationary characteristic test of the RGWR and GWR model.

Parameter	GWR (Fixed) *p*-Value	RGWR (Fixed) *p*-Value	GWR (Adaptive) *p*-Value	RGWR (Adaptive) *p*-Value
Property fees	0.000 ***	0.000 ***	0.000 ***	0.042 **
Greening rate	0.000 ***	0.000 ***	0.000 ***	0.011 **
FAR	0.000 ***	0.000 ***	0.000 ***	0.086 *
Subway station	0.000 ***	0.016**	0.000 ***	0.055 *
Primary school	0.000 ***	0.000 ***	0.000 ***	0.052 *
Junior high school	0.000 ***	0.000 ***	0.000 ***	0.062 *
University	0.000 ***	0.000 ***	0.000 ***	0.010 **

Note: *, **, and *** denote statistically significant at the 10%, 5%, and 1% level.

3.3.2. Model Performance Comparison

Table 4 displays a comparison of the RGWR model with GWR under the fixed bandwidth and adaptive bandwidth strategies. Under the fixed bandwidth strategy, the R^2 of the RGWR model is 0.7777, which is 14.77% higher than the GWR model, R^2adj is 15.06%, MSE is 31.07%, and RMSE is 16.97%. The AICc value of RGWR model is −353.0750, which is 31.4006 smaller than GWR. Generally, a difference in AICc by of greater than three indicates that the two models are also significantly different. The smaller the AICc value, the higher the model fitting accuracy. This shows that under the fixed bandwidth strategy, the RGWR model can better simulate the sale price of residential buildings in Wuhan.

Table 4. The value of the RGWR and GWR model.

Models	Bandwidth/Nth	R^2	R^2adj	MSE	RMSE	AICc
GWR (fixed)	8516.9	0.6776	0.6732	0.0338	0.1838	−321.6744
RGWR (fixed)	8516.9	0.7777	0.7746	0.0233	0.1526	−353.0750
GWR (adaptive)	155	0.6315	0.6265	0.0386	0.1965	−451.1876
RGWR (adaptive)	155	0.6555	0.6508	0.0361	0.1900	−837.7629
RGWR (fixed)/GWR (fixed) Improvement	–	14.77%	15.06%	31.07%	16.97%	
RGWR (adaptive)/GWR (adaptive) Improvement	–	3.8%	3.88%	6.48%	3.31%	-

3.3.3. Comparison of Model Fitting Effects

By comparing the predicted value and the real value of housing prices in the RGWR and GWR models, it is possible to intuitively explore the fitting effect of the model. Figure 4 shows the fitting effect distribution of RGWR and GWR under the fixed and adaptive bandwidth strategies. The X-axis represents the predicted values of different models under different bandwidth strategies, and the Y-axis represents the real housing prices. The red dotted line in the figure indicates that the real value is the same as the predicted value. Therefore, in the figure, the closer the predicted value point distribution and the true value point position are to the red dotted line, the better the fitting effect of the model.

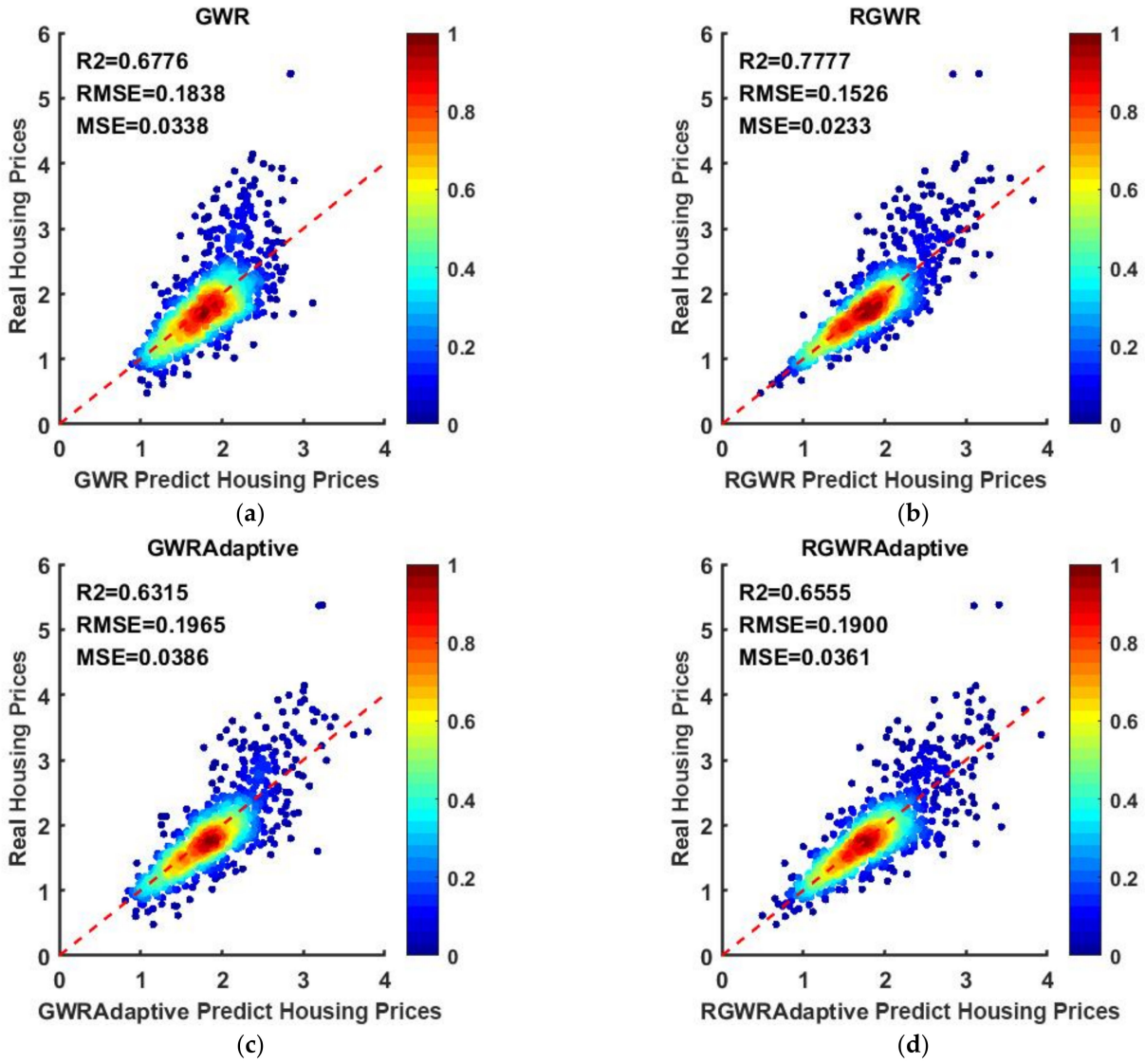


Figure 4. Fitting effect distribution of GWR and RGWR. (**a**) Predicted value and the real value of the housing prices in the GWR models under the fixed bandwidth strategies; (**b**) Predicted value and the real value of the housing prices in the RGWR models under the fixed bandwidth strategies; (**c**) Predicted value and the real value of the housing prices in the GWR models under the adaptive bandwidth strategies; (**d**) Predicted value and the real value of the housing prices in the RGWR models under the adaptive bandwidth strategies.

It can be clearly seen that under the same bandwidth strategy, the point distributions of RGWR are significantly higher than those of GWR near the red dotted line, indicating that the fitting effect of the RGWR model is significantly improved compared to that of GWR. Similarly, between the fixed bandwidth strategy and the adaptive bandwidth strategy, it can be seen that the point distribution of the RGWR fixed bandwidth is higher than that of the adaptive bandwidth near the red dotted line. At the same time, the R^2 value of the RGWR model under the fixed bandwidth strategy is 0.7777, 18.64% higher than that under an adaptive bandwidth. This indicates that, in the data environment of this article, the fitting effect of the RGWR model under the fixed bandwidth strategy is better than that of the RGWR model under the adaptive bandwidth strategy.

3.4. Analysis on the Price of Commercial Housing in Wuhan and Its Influence Factors

3.4.1. Analysis on the Price of Commercial Housing in Wuhan

Figure 5 show the price of commercial housing in Wuhan. Housing prices in the selected study area are between 4800¥ and 53,800¥, with an average price of 18,500¥. The housing estate with the lowest housing price is located in Changlejiayuan, Xinzhou District, Wuhan City, and the housing estate with the most expensive housing price is located in Tiandi Yujiang, Jiangan District, Wuhan City. It can be seen from the figure that the housing prices in the study area gradually moved from the urban fringe area to the urban core area and its surroundings, and the housing prices rose gradually, reaching a peak near the urban core area. This shows that, in the horizontal direction, housing prices show a gradual downward trend from the city center to the periphery. In areas with high housing prices (over 27,000 yuan/m^2), the distribution is concentrated in Wuchang District, Qiaokou District, Jianghan District, and Jiangan District, which shows that the high housing prices in the study area are greatly affected by the school district factors within the district.

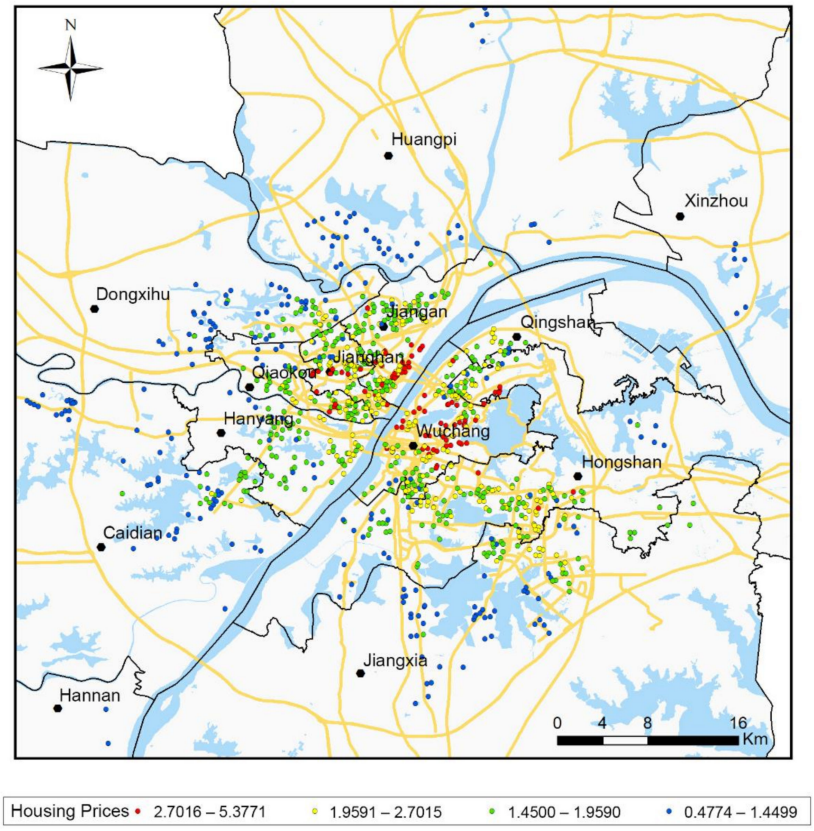


Figure 5. Price of commercial housing in Wuhan.

3.4.2. Influence Factors

Summaries of the RGWR coefficients' estimation under the fixed bandwidth strategy, including the minimum (min), lower quartile (LQ), median (med), upper quartile (UQ), and maximum (max), are presented in Table 5. When using the RGWR models, property fees are positively correlated with house prices, as shown in Table 5. In other words, as the property fees or greening rate increase, the house price increases. In contrast, the presence of a subway station is negatively correlated with house prices; as the distance to the nearest school or subway station increases, the house price decreases. However, there is no significant correlation between the variable FAR and house prices, since the coefficients of FAR have both positive and negative values. Therefore, FAR are not major factors influencing house prices at the scale of the study area. However, this result occurred in this study because FAR has different significant impacts on house prices in different

regionalization within the study area. The coefficients of school (e.g., primary school, junior school) and greening rate also have positive and negative values. Most of the schools are negatively correlated with house prices, while a few are positively correlated with house prices, while greening rate has the opposite trend, which is also related to the different influences of the same influencing factor on house prices in different regions in this study. Specifically, taking typical FAR, greening rate, junior high school, and primary school as examples, this paper analyzes the significant impact of different factors on different regions through the spatial distribution map of influencing factors.

Table 5. RGWR coefficients estimate summaries.

Parameter	Min	LQ	Med	UQ	Max
Property fees	0.00000	0.150591	0.18006	0.205149	0.235376
Greening rate	−0.08843	−0.01152	0.11429	0.223408	0.526083
FAR	−0.27336	−0.11424	−0.05788	0.036113	0.108551
Subway station	−0.133	−0.10929	−0.06097	−0.03337	0.078584
Primary school	−0.06298	−0.05845	−0.0424	−0.02978	0.280742
Junior high school	−0.18901	−0.07806	−0.02306	−0.01593	0.059032

Figure 6 is the FAR coefficient diagram of the study area. It can be seen that the coefficients of FAR range from positive to negative correlations, indicating that there is no significant correlation between FAR and housing prices in the complete study area. However, specific to different administrative divisions, it can be seen that the FAR of Jianghan District, Jiangan District, Qiaokou District, and Wuchang District are positively correlated with housing prices, while the best middle schools and primary schools in Wuhan are basically concentrated in these four districts. Districts and these four districts are the core urban areas of Wuhan, and the plot ratio is positively correlated with the housing price; that is, the higher the plot ratio, the higher the housing price. In other areas of the study, the plot ratio is negatively correlated with house prices, that is, the higher the plot ratio, the lower the house price. This is also related to the pursuit of living comfort in non-central urban areas that deviate from the central area and remove school district factors.

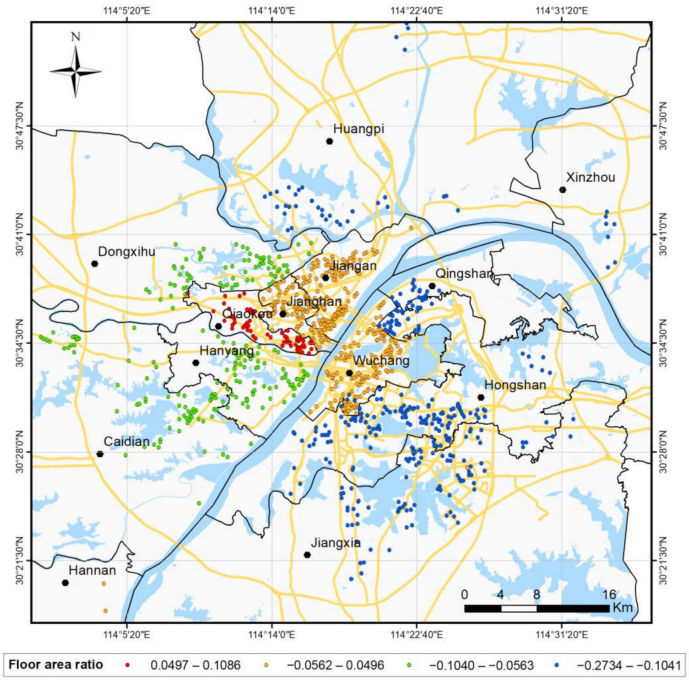


Figure 6. FAR coefficients map of the study area.

Figure 7 is the coefficient diagram of the greening rate in the study area. It can be seen that the greening rate is positively correlated as a whole. As the greening rate increases, the housing price rises. This is particularly prominent in the central area of the study area. The central area is densely built with less overall greenery, so the greening rate has a greater impact on housing prices. In the marginal area of the study area, the greening rate has an influencing factor of −0.0884–0.0003, most of which have a weak negative correlation and approaching zero. This is because the overall residential construction is relatively high in the suburbs of the city. It is sparse and has a high degree of greening, so the impact of greening rate on housing prices is almost negligible.

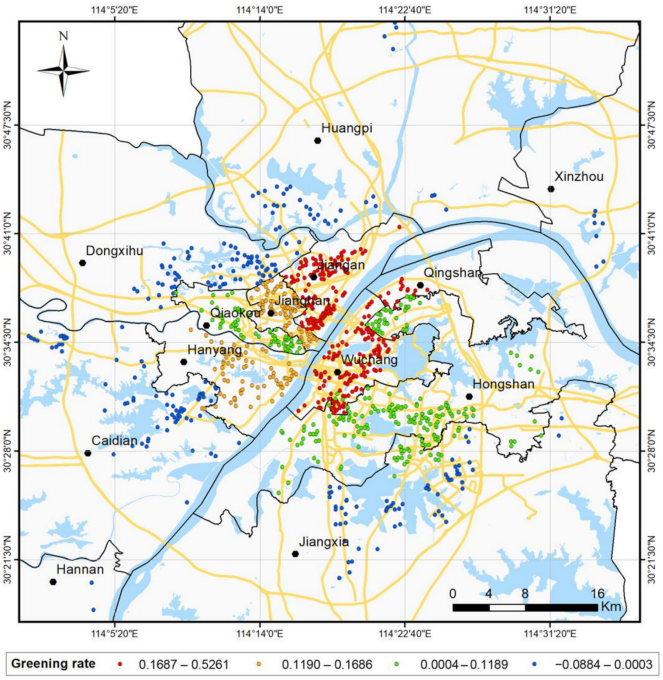


Figure 7. Greening rate coefficients map of the study area.

Figure 8 is the coefficient chart of junior and senior high schools in the study area. It can be seen that the distance of the sampling point between the middle and high schools has different impacts on housing prices in different administrative divisions in Wuhan. Among them, in Wuchang District, the overall housing price is high and there is little room for housing prices to rise. There are many key demonstration middle schools in the area and the area is small. The key demonstration middle schools cover a wide area, so the price is less affected by middle and high schools. The overall housing prices in Jianghan District, Qiaokou District, Hanyang District, and Dongxihu District belong to the middle and lower areas of Wuhan. The number of key demonstration middle schools in the region ranges from 2 to 4. Housing prices are greatly affected by middle and high schools, and the overall housing prices are negatively correlated. The overall housing prices in Jiang'an District and Qingshan District belong to the middle and high areas of Wuhan; housing prices in the area are negatively correlated with middle and high schools, and housing prices are most affected by middle and high schools. Other areas are located in fringe areas of Wuhan. There are few or no demonstration middle schools in the province, and the area is large and lacks high-quality middle schools, so the distance factor between middle and high schools has little effect on housing prices.

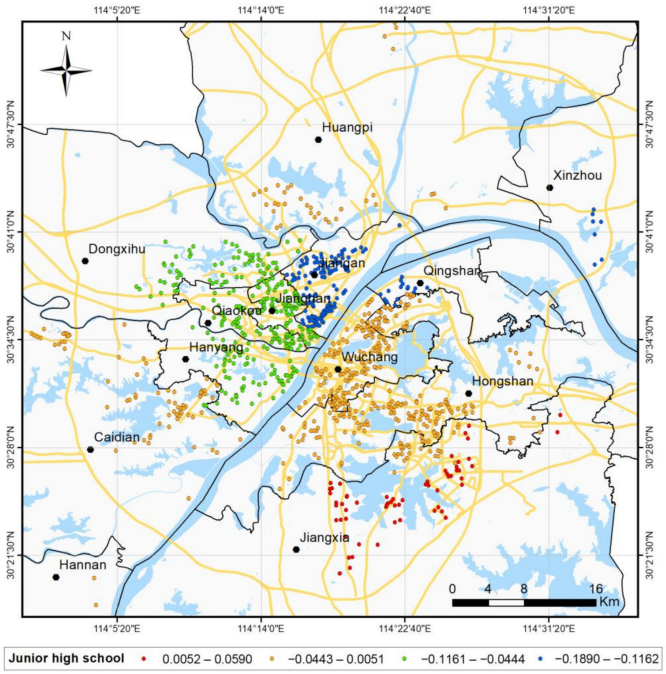


Figure 8. Junior high school coefficients map of the study area.

Figure 9 is the coefficient chart of primary schools in the study area. It can be seen that the overall distance between the sampling point and the primary school in Wuhan is negatively correlated with the housing price. That is to say, the farther away from the nearest primary school, the lower the housing price. This is particularly prominent in the center of the study area, especially in Wuchang District, Jiang'an District, and Qiaokou District. As the best primary schools in Wuhan are concentrated in these three districts, the distance between primary schools in this administrative division has become an important factor affecting housing prices. As the edge of the study area is located in the suburbs of the city, educational resources are relatively balanced, and there is a lack of high-quality primary schools, so the distance between primary schools has little effect on housing prices.

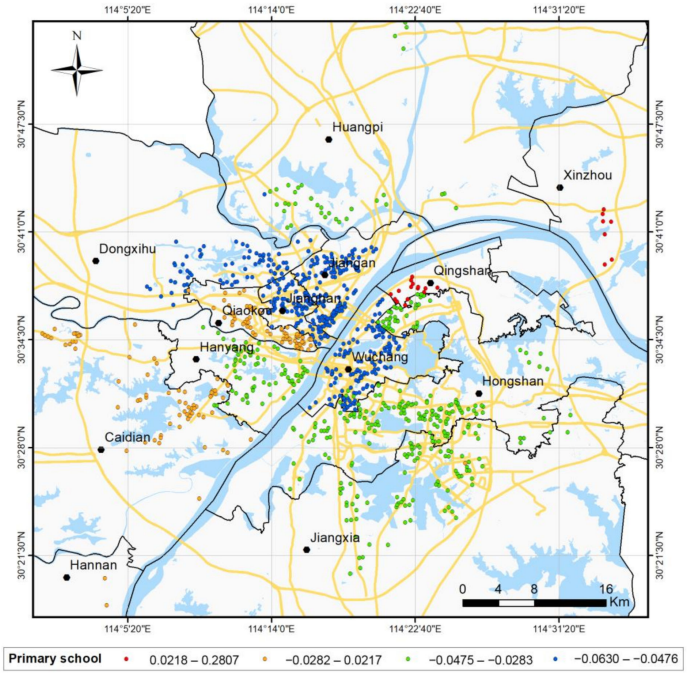
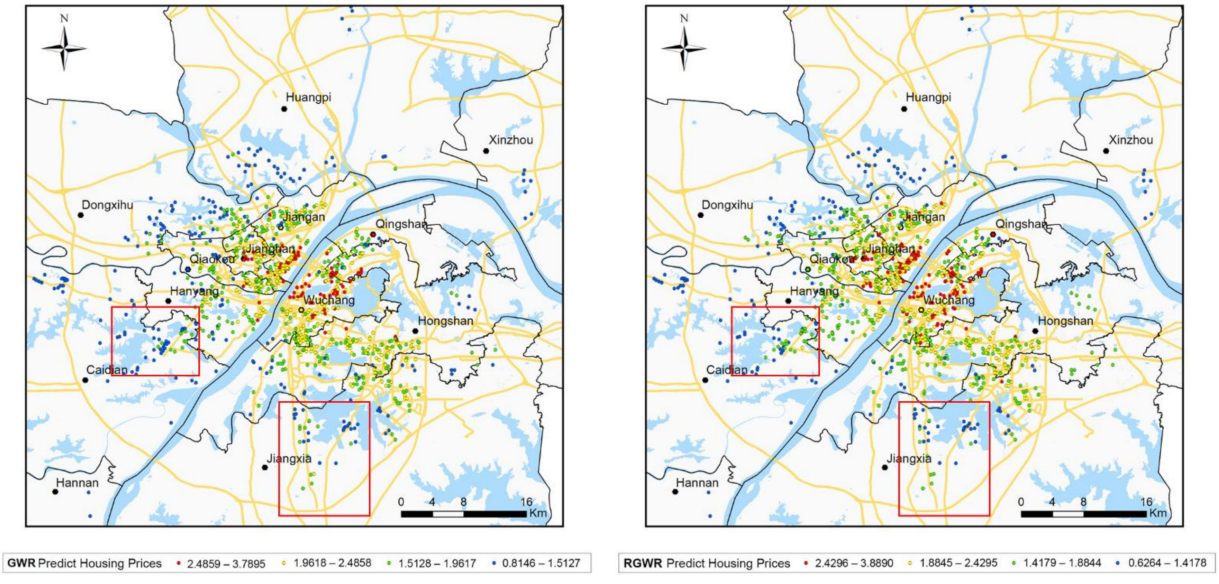


Figure 9. Primary school coefficients map of the study area.

4. Discussion

4.1. The Performance of the Model in Exploring Spatiotemporal Heterogeneity

This paper takes the housing sale price in Wuhan as an example to carry out an experimental analysis, which proves the effectiveness of the introduction of region factors. The study found that, in the Wuhan housing sale price model, under the fixed and adaptive bandwidth strategies, the accuracy of the model is improved after considering the region factors, indicating that the administrative division has a significant impact on housing sale price in Wuhan. Drawing a price distribution map based on the estimation results of the RGWR and GWR models, Figure 5 shows the real housing prices, while Figure 10a,b, respectively, show the predicted housing prices of GWR and the predicted housing prices of RGWR. First, the spatial trend of housing prices distribution in the three figures is basically the same, and prices are expanding outward from the central city area, showing a gradual downward trend. This shows that the two regression models of RGWR and GWR are close to the real situation in global trend estimation, and can objectively reveal the law of housing prices changes in Wuhan. Secondly, in areas with high housing prices (above 24,000 yuan/m^2), RGWR predicted results are distributed in Jiangan District, Jianghan District, and Wuchang district. In addition to the above-mentioned areas, the GWR prediction results also include Qiaokou District, Hanyang District, and Hongshan District. When comparing the real housing prices in Figure 10, RGWR is closer to the real housing price distribution. Third, in the surrounding areas of administrative region, the predicted values of RGWR and GWR are quite different. For example, in the border area of Jiangxia District, Wuhan, the predicted result range of RGWR is below 14,000/m^2, which is close to the true value, while the GWR values are between 15,000 and 19,000/m^2, and the predicted result is relatively high. Similarly, the same situation exists in the border areas of Caidian District and Dongxihu District, because Jiangan District, Jianghan District, Qiaokou District, and Wuchang District are the core areas of Wuhan, and housing prices are relatively high, Hanyang District is adjacent to Qiaokou District and Wuchang District; GWR did not use regions to screen sample points when estimating housing prices in Hanyang District. The selection of sample points in Wuchang District and Qiaokou District for use in the estimation led to the overall high results. RGWR uses regions to screen sample points, and the predicted results are generally close to the real housing prices. The RGWR model is meaningful.



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

Figure 10. The predicted housing prices of different models. (**a**) The predicted housing prices of the GWR model; (**b**) The predicted housing prices of the RGWR model.

4.2. Impact on Adaptive and Fixed Bandwidth

Under the fixed bandwidth strategy, the RGWR model is improved by 14.77%, R^2adj by 15.6%, MSE by 31.07%, and RMSE by 16.97%. Under the adaptive bandwidth strategy, the RGWR model is 3.8% better than the GWR model R^2. R^2adj increased by 3.88%, MSE increased by 6.48%, and RMSE increased by 3.31%. Through comparison, it can be found that the RGWR model under the fixed bandwidth strategy is much better than the RGWR model under the adaptive bandwidth strategy. This is because the selection of sampling points in the study area presents the characteristics of dense core urban areas and scattered peripheral areas. When the adaptive bandwidth strategy is adopted, because the sampling points are dense in the central area, the bandwidth is small, and the sampling points in the edge areas are sparse, and the bandwidth selection will become larger. To a certain extent, this is in line with the characteristics of small administrative divisions in the core urban areas of the study area and large administrative divisions in the fringe areas, thereby reducing the participation of cross-regional sampling points in the estimation. In the fixed bandwidth strategy, the same bandwidth is used in the research area, and a high number of cross-regional sampling points participate in the estimation. Therefore, the RGWR model improves more under the fixed bandwidth strategy than under the adaptive bandwidth strategy.

4.3. Spatial Distribution and Discrete Heterogeneity of Commercial Housing Prices in Wuhan

The overall housing sale prices in Wuhan city show a trend of gradual decrease from the city center to the periphery in the horizontal direction, with high house prices in the core urban areas and low house prices in the peripheral areas of the city, which is consistent with previous studies [58]. Property fees are positively correlated with house prices, have a global spatial scale, and are the main factor affecting housing sale prices in the whole study area: the higher the property fees, the higher the house prices, indicating a heterogeneous spatial effect between the core and newly developed areas of the city [59]. For different administrative regions in Wuhan, the high housing-price areas are concentrated in the center of Wuhan city, and they show a clear polycentric distribution in the core areas of Wuchang, Qiaokou, Jianghan, and Jiangan districts, which is basically consistent with previous studies [60], which show that the spatial distribution of housing prices in Wuhan city is a "three high-priced areas and one lower-priced areas " polycentric pattern with obvious spatial aggregation. At the same time, in the high-priced areas, the FAR and greening rate are positively correlated with housing price, which are the main factors affecting the sale price of housing in these areas. Residents in this area pursue living comfort more than those in the core urban area, so the FAR is negatively correlated with house prices, i.e., the higher the FAR, the lower the house prices. This observation of spatial heterogeneity is consistent with a recent study that suggested that residents of downtown and suburban areas may value different spatial characteristics [61]. With regard to schools as an influencing factor, the general view of previous studies is that all types of schools increase house prices in surrounding residential neighborhoods [62], while some scholars argue that only elementary and junior high schools are globally correlated with house prices [63], and there is also a recent study that used MGWR to analyze house prices in Wuhan and showed that only junior high schools and kindergartens are positively correlated, and elementary schools are negatively correlated [64]. The results of this paper differ from previous studies in that elementary schools and junior high schools are negatively correlated with house prices in high-house-price areas and are one of the main factors affecting the sale price of housing in the area: the closer to the school, the higher the house price. However, they have little effect on house prices in low-price areas, which is determined by China's school district policy. Since there are few high-quality primary and secondary schools in low-price administrative regions, house prices in the area are less affected by schools. In this regard, RGWR can better explain spatial discrete heterogeneity and better reflect the impact of China's primary and secondary school zoning policy, where high-quality schools drive a sharp increase in surrounding house prices.

5. Conclusions

In the last two decades, GWR technology has continuously developed and evolved, which has better solved the continuous heterogeneity in spatial relations caused by irregular distribution in space, but cannot solve the discrete heterogeneity. This paper proposes an RGWR model to use the influence of zoning, modify the spatial kernel function, and optimize the spatial weight. The proposed model can extend the traditional GWR model and weaken the influence of "heterogeneous" observation points on regression points to detect spatial non-stationary features and spatial discrete heterogeneity at the same time.

We use the experimental results of the Wuhan housing price case study to show that the modeling accuracy of RGWR is better than that of the GWR model. The latter only deals with spatial non-stationarity, while the former solves the problem of the discrete heterogeneity of housing prices in Wuhan and improves the accuracy of the model. Compared with the GWR model, RGWR can increase R^2 and R^2adj from 0.6776 and 0.6732 to 0.777 and 0.7746, respectively, and reduce MSE and RMSE from 0.0338 and 0.1838 to 0.0233 and 0.1526, respectively. The AICc standard is also 37.4006 lower than GWR. Statistical tests show that there is a significant difference between RGWR and GWR, so we conclude that it is meaningful to incorporate zoning factors into the GWR model.

Although RGWR originated from the study of house prices, the analysis in this paper is suitable for investigating various phenomena across spatial partitions, such as landscape dynamics, crime, and air quality [65–67]. There are also some limitations in our case study. For example, it is unclear whether RGWR will perform better when applied to data covering more observations or whether the model still has better accuracy than GWR when faced with different partition methods and partition scales. The "zoning" proposed in this paper is not just administrative zoning. Whether RGWR model still has better accuracy than GWR if we choose different zoning methods and zoning scales and with different research problems needs further study.

Therefore, future research will focus on applying RGWR to more observation data and different zoning scales. At the same time, further exploration will be carried out for zoning factors in order to improve the accuracy of detecting discrete heterogeneity. Finally, as the amount of data increases, the interpretation of partition factors between different regions will become more accurate. More research needs to be conducted on the computational efficiency bottleneck of RGWR, which is also one of the hot future research directions for GWR [17]. These must be studied in future work.

Author Contributions: Conceptualization, Zengzheng Wang and Yangyang Zhao; methodology, Zengzheng Wang, Yangyang Zhao and Fuhao Zhang; software, Zengzheng Wang; validation, Zengzheng Wang and Yangyang Zhao; resources, Yangyang Zhao; data curation, Zengzheng Wang; writing—original draft preparation, Zengzheng Wang; writing—review and editing, Zengzheng Wang, Yangyang Zhao and Fuhao Zhang; visualization, Zengzheng Wang; funding acquisition, Yangyang Zhao and Fuhao Zhang All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (NSFC, Grant No. 42001343).

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.

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

References

1. Fotheringham, A.S.; Brunsdon, C. Local forms of spatial analysis. *Geogr. Anal.* **1999**, *31*, 340–358. [CrossRef]
2. Casetti, E. Generating models by the expansion method: Applications to geographical research. *Geogr. Anal.* **1972**, *4*, 81–91. [CrossRef]
3. Jones, K. Specifying and estimating multi-level models for geographical research. *Trans. Inst. Br. Geogr.* **1991**, *16*, 148–159. [CrossRef]
4. Mcmillen, D.P. One hundred fifty years of land values in Chicago: A nonparametric approach. *J. Urban Econ.* **1996**, *40*, 100–124. [CrossRef]
5. Mcmillen, D.P.; Mcdonald, J.F. A nonparametric analysis of employment density in a polycentric city. *J. Reg. Sci.* **1997**, *37*, 591–612. [CrossRef]
6. Elhorst, J.P. Specification and estimation of spatial panel data models. *Int. Reg. Sci. Rev.* **2003**, *26*, 244–268. [CrossRef]
7. Brunsdon, C.; Fotheringham, A.S.; Charlton, M.E. Geographically weighted regression: A method for exploring spatial nonstationarity. *Geogr. Anal.* **1996**, *28*, 281–298. [CrossRef]
8. Fotheringham, S.; Charlton, M.; Brunsdon, C. The geography of parameter space: An investigation of spatial non-stationarity. *Int. J. Geogr. Inf. Syst.* **1996**, *10*, 605–627. [CrossRef]
9. Brunsdon, C.; Fotheringham, S.; Charlton, M. Geographically weighted regression. *J. R. Stat. Soc. Ser. D Stat.* **1998**, *47*, 431–443. [CrossRef]
10. Tobler, W.R. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* **1970**, *46*, 234–240. [CrossRef]
11. Lu, B.; Charlton, M.; Harris, P.; Fotheringham, A.S. 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]
12. Lu, B.; Charlton, M.; Brunsdon, C.; Harris, P. The Minkowski approach for choosing the distance metric in geographically weighted regression. *Int. J. Geogr. Inf. Sci.* **2016**, *30*, 351–368. [CrossRef]
13. Lesage, J.P. A family of geographically weighted regression models. In *Advances in Spatial Econometrics*; Springer: Berlin/Heidelberg, Germany, 2004; pp. 241–264.
14. Harris, P.; Fotheringham, A.S.; Juggins, S. Robust geographically weighted regression: A technique for quantifying spatial relationships between freshwater acidification critical loads and catchment attributes. *Ann. Assoc. Am. Geogr.* **2010**, *100*, 286–306. [CrossRef]
15. Wang, N.; Mei, C.; Yan, X. Local linear estimation of spatially varying coefficient models: An improvement on the geographically weighted regression technique. *Environ. Plan. A* **2008**, *40*, 986–1005. [CrossRef]
16. Yangyang, Z.; Jiping, L.; Shenghua, X.; Fuhao, Z.; Yi, Y. A Geographic Weighted Regression Method Based on Semi-supervised Learning. *Acta Geod. Cartogr. Sin.* **2017**, *46*, 123.
17. Lu, B.; Ge, Y.; Qin, K.; Zheng, J. A Review on Geographically Weighted Regression. *Geomat. Inf. Sci. Wuhan Univ.* **2020**, *45*, 1356–1366.
18. Li, D.; Mei, C.; 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]
19. Anselin, L. Thirty years of spatial econometrics. *Pap. Reg. Sci.* **2010**, *89*, 3–25. [CrossRef]
20. Casetti, E. The expansion method, mathematical modeling, and spatial econometrics. *Int. Reg. Sci. Rev.* **1997**, *20*, 9–33. [CrossRef]
21. Anselin, L. Spatial dependence and spatial structural instability in applied regression analysis. *J. Reg. Sci.* **1990**, *30*, 185–207. [CrossRef]
22. Wang, J.; Zhang, T.; Fu, B. A measure of spatial stratified heterogeneity. *Ecol. Indic.* **2016**, *67*, 250–256. [CrossRef]
23. Zhong, Q.W. The Basic Theoretics and Application Research on Geographically Weighted Regression. Ph.D. Dissertation, Tongji University, Shangai, China, 2007.
24. Harris, R.; Dong, G.; Zhang, W. Using Contextualized Geographically Weighted R egression to Model the Spatial Heterogeneity of Land Prices in Beijing, China. *Trans. GIS* **2013**, *17*, 901–919. [CrossRef]
25. Paez, R.M.H.A. Applied Statistical Analysis of Detailed Geographical Data with Emphasis on Spatial Effects. Ph.D. Dissertation, Tohoku University, Sendai, Japan, 2000.
26. 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]
27. Dai, J. Dummy Variable Regression and Its Application. *Stat. Decis.* **2019**, *35*, 77–80.
28. Lancaster, K.J. A new approach to consumer theory. *J. Political Econ.* **1966**, *74*, 132–157. [CrossRef]
29. Rosen, S. Hedonic prices and implicit markets: Product differentiation in pure competition. *J. Political Econ.* **1974**, *82*, 34–55. [CrossRef]
30. Páez, A.; Long, F.; Farber, S. Moving window approaches for hedonic price estimation: An empirical comparison of modelling techniques. *Urban Stud.* **2008**, *45*, 1565–1581. [CrossRef]
31. Shen, Y.; Karimi, K. The economic value of streets: Mix-scale spatio-functional interaction and housing price patterns. *Appl. Geogr.* **2017**, *79*, 187–202. [CrossRef]
32. Duan, J.; Tian, G.; Yang, L.; Zhou, T. Addressing the macroeconomic and hedonic determinants of housing prices in Beijing Metropolitan Area, China. *Habitat Int.* **2021**, *113*, 102374. [CrossRef]

33. Huang, Z.; Chen, R.; Xu, D.; Zhou, W. Spatial and hedonic analysis of housing prices in Shanghai. *Habitat Int.* **2017**, *67*, 69–78. [CrossRef]
34. Huang, B.; Wu, B.; Barry, M. Geographically and temporally weighted regression for modeling spatio-temporal variation in house prices. *Int. J. Geogr. Inf. Sci.* **2010**, *24*, 383–401. [CrossRef]
35. Liu, J.; Yang, Y.; Xu, S.; Zhao, Y.; Wang, Y.; Zhang, F. A geographically temporal weighted regression approach with travel distance for house price estimation. *Entropy* **2016**, *18*, 303. [CrossRef]
36. Liu, J.; Zhao, Y.; Yang, Y.; Xu, S.; Zhang, F.; Zhang, X.; Shi, L.; Qiu, A. A mixed geographically and temporally weighted regression: Exploring spatial-temporal variations from global and local perspectives. *Entropy* **2017**, *19*, 53. [CrossRef]
37. Wu, C.; Ren, F.; Hu, W.; Du, Q. Multiscale geographically and temporally weighted regression: Exploring the spatiotemporal determinants of housing prices. *Int. J. Geogr. Inf. Sci.* **2019**, *33*, 489–511. [CrossRef]
38. Yao, J.; Stewart Fotheringham, A. Local spatiotemporal modeling of house prices: A mixed model approach. *Prof. Geogr.* **2016**, *68*, 189–201. [CrossRef]
39. Wen, H.; Xiao, Y.; Hui, E.C.M.; Zhang, L. Education quality, accessibility, and housing price: Does spatial heterogeneity exist in education capitalization? *Habitat Int.* **2018**, *78*, 68–82. [CrossRef]
40. Department for Education. House Prices and Schools: Do Houses Close to the Best Performing Schools Cost More? Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/600623/House_prices_and_schools.pdf (accessed on 11 December 2021).
41. Solar, P.M. China's GDP: Some Corrections and the Way Forward. *J. Econ. Hist.* **2021**, *81*, 943–957. [CrossRef]
42. Harten, J.G.; Kim, A.M.; Brazier, J.C. Real and fake data in Shanghai's informal rental housing market: Groundtruthing data scraped from the internet. *Urban Stud.* **2021**, *58*, 1831–1845. [CrossRef]
43. Alm, J.; Lai, W.; Li, X. Housing market regulations and strategic divorce propensity in China. *J. Popul. Econ.* **2021**. [CrossRef]
44. Schneider, A.; Woodcock, C.E. Compact, dispersed, fragmented, extensive? A comparison of urban growth in twenty-five global cities using remotely sensed data, pattern metrics and census information. *Urban Stud.* **2008**, *45*, 659–692. [CrossRef]
45. Gong, P.; Liang, S.; Carlton, E.J.; Jiang, Q.; Wu, J.; Wang, L.; Remais, J.V. Urbanisation and health in China. *Lancet* **2012**, *379*, 843–852. [CrossRef]
46. Dunsky, R.M.; Follain, J.R.; Giertz, S.H. Pricing Credit Risk for Mortgages: Credit Risk Spreads and Heterogeneity across Housing Markets. *Real Estate Econ.* **2021**, *49*, 997–1032. [CrossRef]
47. Stockhammer, E.; Bengtsson, E. Financial effects in historic consumption and investment functions. *Int. Rev. Appl. Econ.* **2020**, *34*, 304–326. [CrossRef]
48. Qu, Y.; Jiang, G.; Li, Z.; Shang, R.; Zhou, D. Understanding the multidimensional morphological characteristics of urban idle land: Stage, subject, and spatial heterogeneity. *Cities* **2020**, *97*, 102492. [CrossRef]
49. Cabras, I.; Sohns, F.; Canduela, J.; Toms, S. Public houses and house prices in Great Britain: A panel analysis. *Eur. Plan. Stud.* **2021**, *29*, 163–180. [CrossRef]
50. Selim, H. Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. *Expert Syst. Appl.* **2009**, *36*, 2843–2852. [CrossRef]
51. Redfearn, C.L. How informative are average effects? Hedonic regression and amenity capitalization in complex urban housing markets. *Reg. Sci. Urban Econ.* **2009**, *39*, 297–306. [CrossRef]
52. Selim, S. Determinants of house prices in Turkey: A hedonic regression model. *Doğuş Üniv. Derg.* **2011**, *9*, 65–76. [CrossRef]
53. Farber, S.; Yeates, M. A comparison of localized regression models in a hedonic house price context. *Can. J. Reg. Sci.* **2006**, *29*, 405–420.
54. Peterson, S.; Flanagan, A. Neural network hedonic pricing models in mass real estate appraisal. *J. Real Estate Res.* **2009**, *31*, 147–164. [CrossRef]
55. Yang, Y.; Liu, J.; Xu, S.; Zhao, Y. An extended semi-supervised regression approach with co-training and geographical weighted regression: A case study of housing prices in Beijing. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 4. [CrossRef]
56. Mian, A.; Sufi, A. House prices, home equity-based borrowing, and the US household leverage crisis. *Am. Econ. Rev.* **2011**, *101*, 2132–2156. [CrossRef]
57. Lisi, G. Hedonic pricing models and residual house price volatility. *Lett. Spat. Resour. Sci.* **2019**, *12*, 133–142. [CrossRef]
58. D'Acci, L. Quality of urban area, distance from city centre, and housing value. Case study on real estate values in Turin. *Cities* **2019**, *91*, 71–92. [CrossRef]
59. Sheppard, S. Hedonic analysis of housing markets. *Handb. Reg. Urban Econ.* **1999**, *3*, 1595–1635.
60. Lin, Y.; Wang, L.; Liu, Y.; Gong, Y.; Xu, J. Distribution Pattern of Urban Housing Price Based on Multiple Big Data: A Case Study of Beijing, Shanghai, Guangzhou, and Wuhan. *J. Landsc. Res.* **2021**, *13*, 57–66.
61. Seo, W.; Nam, H.K. Trade-off relationship between public transportation accessibility and household economy: Analysis of subway access values by housing size. *Cities* **2019**, *87*, 247–258. [CrossRef]
62. 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]
63. Wen, H.; Xiao, Y.; Hui, E.C.M. Quantile effect of educational facilities on housing price: Do homebuyers of higher-priced housing pay more for educational resources? *Cities* **2019**, *90*, 100–112. [CrossRef]

64. Liu, L.; Yu, H.; Zhao, J.; Wu, H.; Peng, Z.; Wang, R. Multiscale Effects of Multimodal Public Facilities Accessibility on Housing Prices Based on MGWR: A Case Study of Wuhan, China. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 57. [CrossRef]
65. Paul, J.D.; Roberts, G.G.; White, N. The African landscape through space and time. *Tectonics* **2014**, *33*, 898–935. [CrossRef]
66. Brunsdon, C.; Corcoran, J.; Higgs, G. Visualising space and time in crime patterns: A comparison of methods. *Comput. Environ. Urban Syst.* **2007**, *31*, 52–75. [CrossRef]
67. He, Q.; Huang, B. Satellite-based mapping of daily high-resolution ground PM2.5 in China via space-time regression modeling. *Remote Sens. Environ.* **2018**, *206*, 72–83. [CrossRef]