International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-13

https://doi.org/10.70917/ijcisim-2026-0370
Article

Innovative Research on Rural Housing Renewal and Rehabilitation Models Driven by Rural Electricity Economy

Rong Li [1,2,*], Chaonan Liu [1] and Bei Li [3]

[1] School of Information Technology, Nanchang Vocational University, Nanchang, Jiangxi, 330050, China
[2] Hanyang Graduate School of International Studies, Hanyang University, Seoul, 04763, South Korea
[3] Wuhan Shangguan Information Technology Co., Ltd., Wuhan, Hubei, 430000, China
* Correspondence author: lirong_lr2021@163.com

Abstract: Under the background of rapid development of e-commerce economy, rural housing renovation faces new opportunities and challenges. The traditional housing renovation model is difficult to meet the needs of modern rural development, with obvious spatial differentiation characteristics and complex and diverse influencing factors. Exploring the innovative model of housing renovation driven by e-commerce economy is of great significance for promoting rural revitalization, improving rural living environment, and promoting the coordinated development of urban and rural areas. This study adopts a combination of geographically weighted regression (GWR) model and random forest (RF) model to deeply analyze the spatial influencing factors and threshold effects of rural housing renovation driven by e-commerce economy, based on 800 rural housing renovation sample sites in County A. The spatial regression analysis was carried out by constructing the characteristic price model, using ArcMap 10.3 software, and producing partial dependency maps using Python 3.0 software. The results show that: the goodness-of-fit R^2 of the GWR model reaches 0.8314, which is significantly better than that of the OLS model of 0.7963, and the standard deviation decreases from 0.623 to 0.436; the prediction accuracy of the RF model is as high as 94.1%, and the average absolute error is 996.52 yuan/square meter; the housing location has the greatest influence on the remodeling price, with the contribution of 20.21%, followed by the transportation accessibility (16.32%) and park accessibility (13.63%); spatial autocorrelation analysis shows that the Moran's I value increases from 0.286 to 0.333 in 2021-2024, showing a significant positive correlation clustering characteristics. The study puts forward three countermeasure suggestions to improve the design of housing system, pay attention to the orientation of social opinion, and enhance the economic strength of farmers, which provide theoretical basis and practical guidance for the innovation of rural housing renovation mode.

Keywords: geographically weighted regression, random forest, rural housing reform, spatial influence factors, threshold effect, e-commerce economy

1. Introduction

With the rapid development of information technology, the digital economy has become a new engine to promote global economic growth. In this context, the rise of rural e-commerce is not only a new opportunity for rural economic development, but also an effective path to realize the deep integration of agriculture and the Internet [1]. Rural e-commerce breaks the geographical limitations of the traditional market through the network platform, so that agricultural products can quickly enter a broader market [2]. This mode of directly connecting producers and consumers can not only increase the sales of agricultural products, but also increase the income of farmers and directly promote the growth of rural economy [3-4]. In addition, rural e-commerce can accurately understand the market demand through big data analysis, guide farmers to scientific planting, reduce the waste of resources brought about by blind production, and improve the efficiency of agricultural production [5-6]. At the same time, the e-commerce platform provides more diversified sales channels for agricultural products, which can increase the added value of agricultural products and further promote the development of rural economy [7-8].

In the process of continuous integration and upgrading of agriculture-related industries, its industrial chain has also undergone drastic changes, and the new industrial chain closely combines urban-rural relations, and at the same time breeds diversified construction carriers [9-10]. Driven by the e-commerce economy, rural greenhouses, phytochemical plants, farms, farms, origin warehouses and other building carriers have become increasingly important in the rural industry [11-13]. At the same time, brand new production methods brought by technological innovations such as land use simulation industrial economic accounting, indoor navigation technology use, and drone logistics are being gradually piloted, promoted, and operated in the process of rural building renewal [14-16]. Therefore, it is important to explore the path of housing and other infrastructure construction and renovation in the context of rural e-commerce development, with a view to solving the pain points in the process of optimization of the industrial structure, and providing support for the high-quality development of the rural e-commerce economy.

The vigorous development of the e-commerce economy is profoundly changing the face of China's rural areas and injecting new vitality into rural revitalization. In the context of this era, rural housing, as an important carrier of farmers' production and life, its renewal and transformation is not only related to the improvement of farmers' living quality, but also a key link in promoting the process of rural modernization. However, the traditional housing renovation model often exists in a low degree of standardization, irrational spatial layout, supporting facilities lagging behind and other problems, it is difficult to adapt to the development of the e-commerce economy on the rural infrastructure and living environment put forward by the new requirements. Currently, housing renovation in rural areas in China is facing multiple challenges: first, the differentiation of housing renovation needs due to unbalanced regional development; second, the lack of capital investment constrains the scale and quality of renovation; third, the lack of systematic renovation standards and technical specifications; and fourth, the renovation process is not enough to protect the rural characteristics and culture. The existence of these problems not only affects the improvement of farmers' quality of life, but also restricts the development space of rural tourism, rural e-commerce and other emerging industries. Therefore, in-depth study of rural housing renovation model innovation driven by e-commerce economy is of great practical significance and theoretical value for cracking the current dilemma and realizing rural high-quality development.

Based on the above background, this study constructs a systematic analytical framework and adopts a research method combining quantitative and qualitative. First, the geographically weighted regression model is used to analyze the spatial influencing factors of rural housing renovation in depth, verify the superiority of the GWR model by comparing with the OLS model, and identify the functioning mechanism of each influencing factor in different geographical locations. Secondly, a random forest model is introduced to explore the threshold effect of housing renovation, and the contribution of each factor to the renovation price is quantified through variable importance analysis and partial dependency graph production, revealing the characteristics of the nonlinear relationship. Finally, combined with the spatial autocorrelation analysis, it comprehensively grasps the spatial and temporal evolution of housing renovation, provides a scientific basis for the formulation of differentiated renovation strategies, and puts forward targeted policy recommendations and implementation paths.

2. Research on spatial influencing factors for rural housing rehabilitation based on the GWR model

2.1. Overview of GWR modeling theory

2.1.1. GWR Fundamentals

Geographically weighted regression (GWR) model takes into account the location of the observation points, and conducts linear regression at all sample points with different geographic locations, respectively, and calculates the coefficient results of the impact of each factor on housing renovation at all sample points with different locations, which also allows us to obtain the impact of each factor on housing renovation in different localized areas, and then find the spatial differentiation of the impact of each factor on housing renovation [17].The principle formula of GWR model is shown in equation (1):

$$y_i = \beta_0 u_i, v_i + \sum \beta_k u_i, v_i x_{ik} + \varepsilon_i \tag{1}$$

In Eq:

u_i, v_i is the spatial location coordinates of the i th sample point, generally using latitude and longitude.

β_k, u_i, v_i the value of the continuous function at sample point i, i.e., the coefficients of each explanatory variable at each sample point.

x_{ik} Values of the explanatory variables at sample point i. ε_i The random error term at sample point i.

Let the total number of sample points be n and the number of independent variables k, then we will end up with $n*(k+1)$ coefficients. Then the above equation in matrix form is shown in equation (2):

$$y = X\beta I + \varepsilon \tag{2}$$

Eq:

X is the row vector consisting of the variables.

β is the coefficient matrix of the sample points.

I is the column unit vector.

2.1.2. Bandwidth parameters

Bandwidth is the threshold for calculating the spatial weights, so the size of the bandwidth will affect the calculation of the spatial weight parameter, points within the bandwidth range will be given spatial weights that will have an impact on the sample points, and points outside of the bandwidth range are not accounted for in the spatial weight, i.e., they are weighted to zero. The closer the points are to the sample points, the smaller the spatial weights are. Too large a bandwidth, the weights will decay more gently, the calculation of the results obtained more smooth, may cause some areas of the coefficients of the results of the significance of the deterioration, it is difficult to get effective information. Too small bandwidth will lead to the lack of correlation between the spatial sample points, which may make the coefficients of the model calculation results in the history of spatial changes in the space is drastic, and it is also difficult to get the results in line with the reality, and can not uncover effective information. Therefore, the calculation of bandwidth is of great significance, and only by choosing the appropriate calculation method and obtaining the appropriate bandwidth can meaningful model results be obtained. There are two main calculation methods for bandwidth, which are cross-validation (CV) and Akaike Information Content Criterion (AIC).

The calculation formula of the cross-validation method is shown in equation (3):

$$CV = \sum_{i=1}^{n} [y_i y_{\neq i} b]^2 \tag{3}$$

Where: $y \neq i$ is the model fit value of y_i.

When using the cross-validation method to calculate the bandwidth, the function is continuously analyzed using regression points with data from nearby sample points, and the calculation produces the respective bandwidth values and the corresponding CVs for trendline fitting until the bandwidth corresponding to the smallest CV is compared, which serves as the optimal bandwidth.

The Bare Pool informativeness criterion method is a parameterization method for baseline-improved great likelihood estimation, a model screening criterion whose formula is shown in equation (4):

$$AIC = 2K - 2\ln L \tag{4}$$

Eq:

K is the number of model parameters.

L is the likelihood function.

AIC can be regarded as a measure characterizing the effectiveness of the model fit, the smaller the AIC value obtained from the data sample calculation, the better the model fit.

There is also a difference between the cross-validation method and the AIC method in the form of bandwidth characterization of the two calculation results are different, the bandwidth obtained by the cross-validation method is expressed in the form of a distance threshold, for example, if the bandwidth is 1,000 meters, only points within 1,000 meters from the sample point are calculated. While the AIC method calculates the bandwidth results are characterized in terms of the number of nearest neighbor points involved in the calculation.

2.1.3. Spatial weighting parameters

According to the "first law of geography", i.e., things are often related to each other and have stronger correlation with things nearer than things farther away, so the calculation of spatial weights generally follows this law, and elements nearer to each other are given more weight. In this study, we use ArcMap 10.3 software from ESRI to perform geographically weighted regression analysis, which provides two kinds of calculations on spatial weight calculation, the first one is Gaussian distance weight function (fixed type), and the second one is bi-square distance weight function (adaptive type).

Among them, the first spatial weight calculation method - Gaussian distance weight function is of the form:

$$W_{ij} = e^{-(\frac{d_{ij}}{b})^2} \tag{5}$$

The second method of calculating the spatial weights - the two-square function is expressed as:

$$W_{ij} = \begin{cases} \left[1-(\frac{d_{ij}}{b})^2\right]^2 & d_{ij} \le b \\ 0 & d_{ij} \ge b \end{cases} \tag{6}$$

Eq:

W_{ij} is the weight between sample points i, j.

d_{ij} is the Euclidean distance between sample points i, j.

b is the bandwidth.

As the distance between the sample points increases, i.e., when d_{ij} is larger, the weight W_{ij} between sample point i and sample point j decreases, and when the distance between sample point i and sample point j is larger than the bandwidth threshold, the weight W_{ij} between the two points will be close to zero.

2.2. Analysis of spatial influences on rural housing rehabilitation

2.2.1. Model evaluation analysis

This subsection compares the GWR characteristic price model proposed above with the general linear regression (OLS) method. Based on the comparison results, the characteristic price model is selected to analyze the influencing factors of rural residential prices in County A driven by the e-commerce economy.

The model is constructed based on GWR, with the adaptive Gaussian function as the kernel function, the golden section method as the bandwidth search method, and the bandwidth is calculated by the modified Akaike information criterion method, and the relevant parameters of the OLS model and the GWR model are obtained by using the ArcMap10.3 software, and the analytical results of the two models are compared.A comparison of the results of the OLS and GWR results is shown in Table 1. is overall significant, but the test results of OLS residuals (Moran's I index validation) show spatial autocorrelation, which indicates that the regression results of OLS have a large error, and the global model based on the OLS method is not appropriate. The goodness of fit (R^2) of the GWR model is greater than that of the OLS model, and both the residual sum of squares and standard deviation of the GWR are significantly lower compared to the OLS model.

Table 1. OLS compared with GWR results

Model index	Residual sum of squares	Standard deviation	AIC	R^2	Adj R^2
OLS	351.63	0.623	2821.643	0.7963	0.756
GWR	286.66	0.436	2463.61	0.8314	0.793

In summary, the spatial regression analysis method based on the GWR model effectively reduces the residuals, significantly improves the simulation accuracy, and the optimal bandwidth value is also significantly improved, so this paper uses the GWR method to analyze the influencing factors of the price of rural housing in County A.

2.2.2. General statistical analysis

The basic structure of the data sample points and the characteristics of the data distribution can usually be analyzed based on the frequency histograms and normal QQ plots of the sample points.

This chapter focuses on analyzing the basic structure of the sample points of rural housing rehabilitation in County A in 2024. The histogram of frequency distribution is a visual expression form of the data, and the mean and median can reflect the size of the overall sample data as well as the reflection of the distribution of the main data according to the average distribution pattern of the data and the middle pattern of the overall distribution. The distribution pattern of the data can be reflected by the skewness and kurtosis, the skewness is close to 0, kurtosis is close to 3, when the overall changes in the sample point is a normal distribution style. The distribution of rural housing renovation in County A is shown in Figure 1. It can be seen in the 800 rural housing renovation sample points, the sample point mean, the lowest housing renovation highest housing renovation are 19,632, 55063, 5963 yuan, according to the housing renovation sample points, the kurtosis is steeper, the data is more and scattered distribution above the mean.

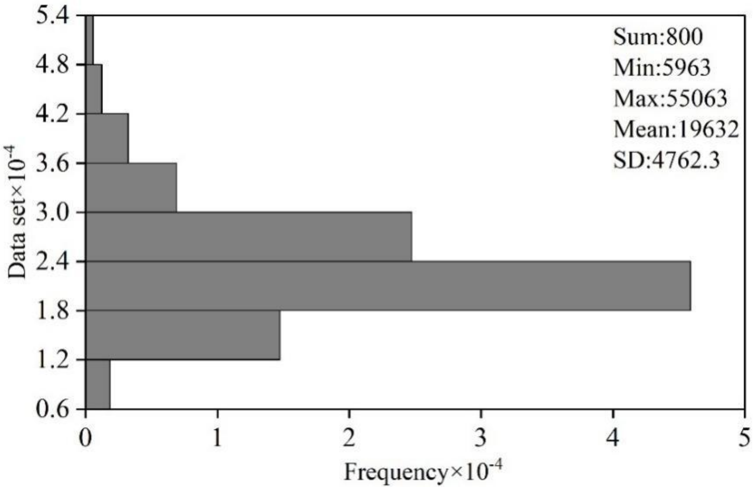

Figure 1. The distribution of housing transformation in a county and township village

After the logarithmic transformation of County A rural housing transformation normal QQ plotting is shown in Figure 2. It can be seen that the distribution of the housing transformation sample data is close to the normal distribution, more dispersed at the beginning and end of the data distribution, and the trend of the normal distribution is significantly different, but the overall has been similar to the normal distribution, housing transformation data are more centrally distributed near the middle value.

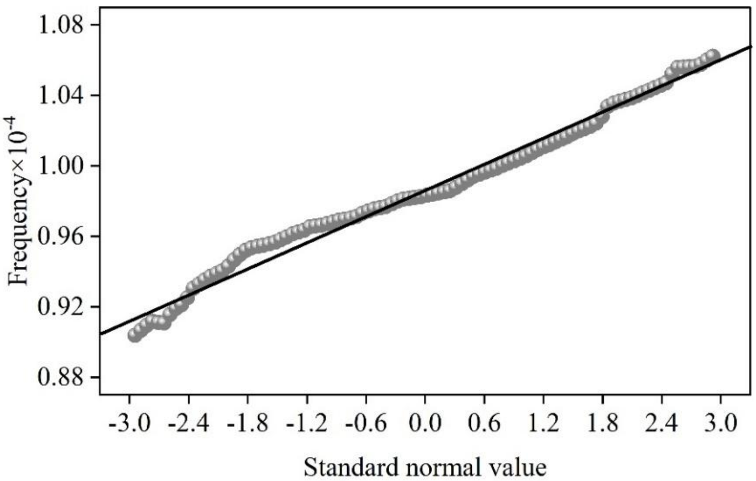

Figure 2. A county village housing changes the normal QQ map

2.2.3. Spatial autocorrelation analysis

The spatial autocorrelation characteristics of the time-series data on housing rehabilitation in rural County A reflect the time-series changes in the degree of agglomeration.Moran's I of the average price of residential housing in rural County A for the month of October, 2021-2024, was used to determine the spatial weight matrix using the Gaussian distance weight calculation method, and the spatial distance between two points was measured using the vertical axis, which objectively expresses the spatial relationship between the neighboring of rural villages of 2800 m spatial relationship of residential housing renovation within the distance threshold. As can be seen from the results table, the I_g values of the housing retrofit market in recent years are all positive and all statistically significant, indicating that the spatial positive correlation of the housing retrofit market is characterized by significant agglomeration. The housing rehabilitation market exhibits high housing rehabilitation agglomeration and low housing rehabilitation agglomeration, with significant geographic location characteristics.

Comparing the I_g values for the four years, the housing rehabilitation comparisons are shown in Table 2. 2021 to 2022 saw the largest increase and the fastest growth rate. 2022 saw rapid growth in the rural real estate sector, significant changes in localized regional economic and amenity development, and more significant differences between high and low housing rehabilitation areas. The value of I_g gradually stabilizes from October 2022 to October 2023, reflecting the steady socio-economic development in recent years, the conscious policy regulation by the government, and the tendency of full-scale real estate construction in various regions.

Table 2. Comparison of housing transformation 2021-2024

Year	I_g	Z	P	Expectation index	Variance
2021	0.286	21.949	0	-0.001	0
2022	0.307	25.326	0	-0.001	0
2023	0.324	26.442	0	-0.001	0
2024	0.333	26.752	0	-0.001	0

Rural County A has a spatially positively correlated agglomeration of housing rehabilitation markets across the range of different spatial distance thresholds. The global Moran index for housing rehabilitation at different distance thresholds is shown in Table 3. When the distance threshold is set to 0, the residential sample points in rural County A are all neighboring elements, which is complicated for the calculation of the weights and does not conform to the objective spatial proximity theory, but still meets the hypothetical test criteria and shows positively correlated agglomeration.The spatial distance threshold of 1400m has the highest value of I_g , but some of the housing renovation sample points do not have neighboring elements, which can only reflect the agglomeration of the local spatial sample points , and the statistical properties of the test may fail. As the distance threshold increases, the neighboring housing renovation sample points increase, and the I_g value decreases gradually, arguing that the more neighboring housing renovation sample points are more similar, and at 2800m passes the hypothesis test criterion of positive spatial correlation clustering, and considering the existence of sample points with a large degree of discrete degree, 2800m is adopted as a valid research range.

Table 3. Different distance threshold housing changes the global Moran index

Distance threshold	I_g	Z	P	Expectation index	Variance
0	0.086	28.057	0	-0.001	0
350	0.368	6.732	0	-0.001	0.001
700	0.319	7.83	0	-0.001	0.002
1050	0.451	11.953	0	-0.001	0.002
1400	0.415	13.833	0	-0.001	0.001
1750	0.426	15.452	0	-0.001	0
2100	0.408	18.512	0	-0.001	0
2450	0.317	26.567	0	-0.001	0
2800	0.319	26.731	0	-0.001	0

2.2.4. Analysis of factors affecting rural residential rehabilitation

In this paper, 10 variables of building age, floor area ratio, land use planning, natural landscape, transportation accessibility, housing location, historical and cultural resources, spatial layout, park accessibility, and community environment are selected to study the influencing factors of rural residential renovation.

The result obtained by GWR is a specific coefficient, i.e., regression coefficient, generated by each influencing factor for each sample, and the obtained data are counted and summarized, and the regression results of GWR model are shown in Table 4, and the numbers 1~10 represent the above variables, respectively. Comparison of the regression coefficients shows that the regression coefficients for residential remodeling are shown in the table as positive and negative, large and small, but the difference is not obvious, indicating that the various influencing factors are all present, and that the impact effects formed are both high and low.

Table 4. The GWR model returns

Factor	Mean	Min	Median	Max	SD
1	-0.147	-1.231	-0.095	0.239	0.209
2	-0.11	-0.791	-0.098	0.354	0.167
3	0.214	-0.232	0.206	0.547	0.146
4	0.098	-0.475	0.103	0.508	0.139
5	0.146	-0.212	0.114	1.03	0.197
6	-0.088	-6.602	-0.037	6.435	1.843
7	0.072	-0.406	0.041	0.814	0.18
8	0.048	-0.663	0.04	1.38	0.236
9	-1.471	-12.563	-0.789	4.673	2.552
10	0.164	-1.733	0.015	4.407	0.601

3. Research on the threshold effect of rural housing rehabilitation based on RF modeling

3.1. Overview of RF modeling theory

3.1.1. RF model fundamentals

In this paper, the Random Forest Model (RF) is used to identify the key drivers affecting rural housing rehabilitation and to accurately estimate the extent to which each driver contributes to affecting rural housing rehabilitation. The basic principle is to use Bootstrap sampling method to draw multiple samples simultaneously from the original sample, and perform decision tree modeling for each Bootstrap drawn sample. Because of the introduction of randomness in the process of generating decision trees, it is not prone to the phenomenon of overfitting, and therefore can have a better tolerance for many driving factors affecting rural housing renovation, and is a natural nonlinear model [18], and its general process is as follows:

The first step is to generate training sets. Each decision tree corresponds to a training set, to construct N trees it is necessary to extract N sub-training sets by Bootstrap, and then integrate the prediction for all sub-training sets, which have large differences between them, thus reducing the variance and ensuring more reliable prediction results.

In the second step, the decision tree is constructed. A number m is set to determine the number of splitting variables that are candidates for splitting when the node splits, and m should be smaller than the number of variables M, which is generally adopted as $M/3$. For each node, m variables are randomly selected to calculate its optimal split, and the split is repeated continuously.

In the third step, a random forest is formed. Repeat the above steps so as to build a large number of decision trees, and finally generate a random forest. As a result of resampling with put-back, each number has some unused observations, called out-of-bag observations (OOB observations), which in turn generates an out-of-bag error (OOB error rate):

$$MSE_{OOB} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_{i,OOB} - y_i)^2 \tag{7}$$

It is also possible to calculate quasi $R^2(Pseudo\ R^2)$, the degree to which the model is explainable in terms of the sample variance of the response variable, based on the OOB mean square error, with the following formula:

$$PseudoR^2 = \frac{Var(y) - MSE_{OOB}}{Var(y)} = 1 - \frac{MSE_{OOB}}{Var(y)} \tag{8}$$

where $Var(y)$ is the sample variance of the response variable y.

3.1.2. Variable Importance and Biased Dependency Graphs

Random forests contain many decision trees, but since only one variable is used in each node split, it is possible to measure the decline in the split criterion function (residual sum of squares or Gini coefficients) due to each variable in a decision tree, and then average the declines due to a variable in each decision tree to derive the importance of that variable. Each variable is plotted in order of importance to obtain a variable importance plot.

Variable importance is just a measure of how important each variable is and ranked, but sometimes we are more curious about the marginal effect of each variable on y. For the feature vector $x = (x_1 x_2 \cdots x_p)'$, assuming $y = f(x)$, the marginal effect of the first variable x_l on y is:

$$\frac{\partial y}{\partial x_1} = \frac{\partial f(x_1, x_2, \cdots, x_p)}{\partial x_1} \tag{9}$$

In equation (9), the marginal effect $\frac{\partial}{\partial x_1}$ depends on the values of the other variables $(x_1, x_2, \cdots, x_p)$, and one can consider averaging out the effect of the other variables $(x_1, x_2, \cdots, x_p)$ on y through integration:

$$\varphi(x_1) = E_{x_2, \cdots, x_p} f(x_1, x_2, \cdots, x_p) \tag{10}$$

In this case, the expectation operator $E_{x_2,\ldots,x_p}(\cdot)$ seeks the expectation for the variable $(x_2, \cdots, x_p)$ by substituting the sample mean for the overall mean $E_{x_2,\ldots,x_p}(\cdot)$ can be obtained:

$$\varphi(x_1) = \frac{1}{n} \sum_{i=1}^{n} f(x_1, x_{i2}, \cdots, x_{ip}) \tag{11}$$

In Eq. (11), given any x_1 one can compute $\varphi(x_1)$ and one can draw the image of $[k_1, \hat{\varphi}(x_1)]$, i.e., the partial dependency graph.

The advantage of Random Forest is to use Bootstrap to alleviate the problem of high variance and weaken the correlation between decision trees. The advantage of randomness in selecting features using the random subspace method enhances the generalization ability of the model. Averaging the predicted values of multiple decision trees improves the prediction accuracy and achieves accurate classification and prediction of data. Random forest method is simple and efficient, only need to adjust the number of trees in the forest and the number of features in each node to generate a reasonable model quickly and effectively. Compared with other machine learning methods, Random Forest has a strong resistance to overfitting, a high tolerance for outliers and noise, and obvious advantages in parameter optimization, variable ordering, and subsequent variable analysis and interpretation.

3.2. Analysis of threshold effects on rural housing rehabilitation

3.2.1. Importance analysis of factors affecting rural housing rehabilitation

Random forests were introduced to analyze the importance of the factors influencing rural housing renovation in County A. The ranking of the importance of the factors influencing housing renovation is shown in Table 5. The model prediction accuracy is as high as 94.1%, and the average absolute error is 996.52, which indicates that the average difference between the predicted housing renovation price and the real housing renovation price is 996.52 yuan/square meter, and the model fits well. Importance is evaluated in terms of the contribution of each feature to reducing the impurity of the decision tree. As can be seen from the table, housing location has the highest level of importance in influencing housing remodeling prices, providing a 20.21% contribution to the prediction of housing remodeling prices. This is followed by accessibility to transportation and accessibility to parks, contributing 16.32% and 13.63% to the prediction of housing retrofit prices.

Table 5. The importance of housing modification factors is ranked

Var	Importance/%	Ranking
1	3.16	9
2	10.63	5
3	4.3	8
4	7.61	7
5	16.32	2
6	20.21	1
7	1.88	10
8	9.63	6
9	13.63	3
10	12.63	4
	Mean absolute error: 996.52	
	Prediction accuracy: 94.1%	

3.2.2. Analysis of threshold effects affecting housing rehabilitation

Partial dependency analysis aims to elucidate how a selected explanatory variable affects the expected value of the model's predictions. While the previous variable importance analysis focused on explaining the magnitude of a variable's impact on the model's predictive performance, partial dependence analysis is used to explain how this variable affects the model's predictive performance. This study uses partial dependency (PD) analysis to quantitatively assess the response of rural housing rehabilitation prices in County A to changes in a single predictor variable.

Partial dependence plots of the top three influencing factors were created using Python 3.0 software, and the plotting results are shown in Figures 3 through 5. The partial dependency plot demonstrates the marginal effect of a single variable on a predictor variable after controlling for other variables in the model. In the following figure, the vertical axis is the housing rehabilitation price, the horizontal axis represents each variable affecting the housing rehabilitation price, and all fitted curves were smoothed to better show the trend of change, and the trend of change of all independent variables is consistent with the expectation.

The effect of residential location on housing remodeling price is shown in Figure 3. Residential location shows a trend of “slightly increasing, then decreasing, and then stable” on the price of housing renovation. Within 3km from the city center, the price of housing renovation rises slightly with the increase of distance, while in the range of 3km-17.5km, the price of housing renovation shows a decreasing trend with the increase of distance, which verifies that the location of housing has an important influence on the price of housing renovation. When the distance exceeds 17.5km, the slope becomes gentle again, and the housing rehabilitation price is less obviously influenced by the location factor.

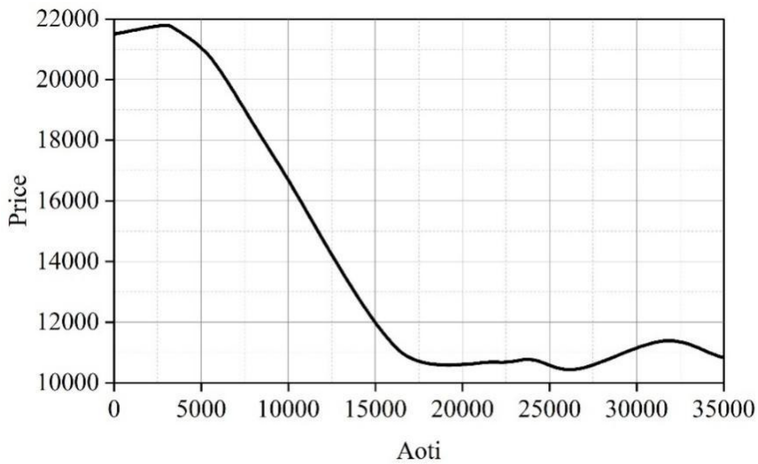

Figure 3. The influence of residential area on housing reform price

The impact of park accessibility on the price of housing renovation is shown in Figure 4, which shows that park accessibility has a “first increasing, then leveling off” trend on the price of housing renovation; when park accessibility is less than 4.1, the price of housing renovation shows an increasing trend, and once the park accessibility exceeds 4.1, the price of housing begins to remain flat.

It can be seen that the accessibility of green space to housing is not as high as it could be, and that once the number of green spaces around housing reaches a certain amount, any further increase in the number of green spaces will result in a loss of diversity, and thus will not contribute to a sustained increase in the price of housing renovation.

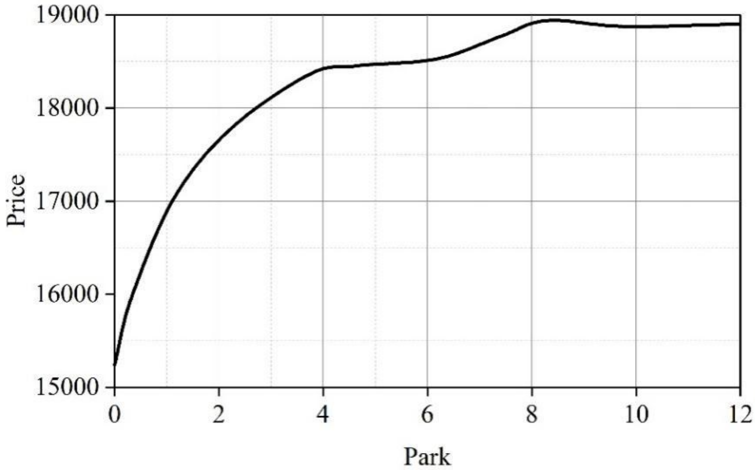

Figure 4. The impact of park accessibility on housing reform prices

The impact of public transportation accessibility on the price of housing rehabilitation is shown in Figure 5, which shows an inverted "N" trend of "decreasing, then increasing, then decreasing" in the price of housing rehabilitation, and when the accessibility of bus stops is less than 2, the price of housing rehabilitation shows a decreasing trend, which is due to the fact that most of the housing in this area is suburban housing, which is located in less-developed areas. This is because most of the housing in this area is suburban housing, which is located in less developed transit routes. When transit accessibility exceeds 2 and is less than 5, the price of housing rehabilitation shows a significant increasing trend, confirming the important impact of public transportation on residents' travel, and due to the convenience and public welfare of public transportation, homebuyers' willingness to pay for the convenience of public transportation travel, which leads to a housing premium.

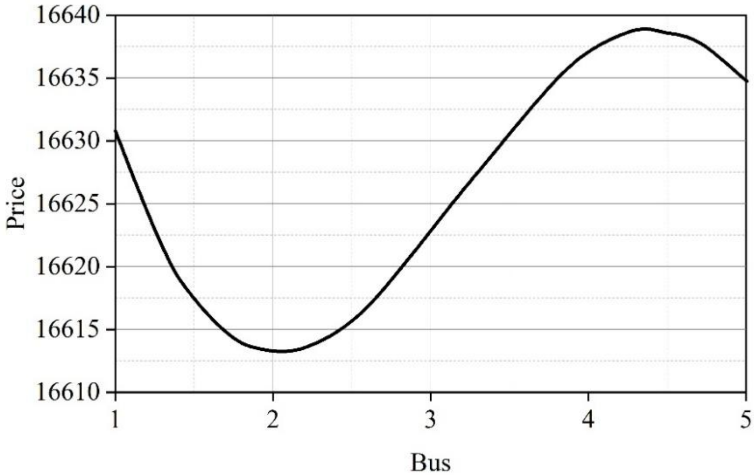

Figure 5. The impact of public transport on housing reform prices

4. Response to rural housing upgrading

4.1. Improving the design of the housing system and focusing on the housing needs of the elderly

Rural housing is mostly inhabited by the elderly. First of all, the State and the Government should further increase the degree of attention paid to the elderly in rural areas, improve the relevant systems, laws and policies, appropriately expand the coverage of social security, and better utilize the

Government's function as a guarantor to meet the housing needs of the elderly in rural areas. Secondly, with the assistance of village committees, grass-roots township governments should conduct more in-depth field visits and research at the grass-roots level, and accurately promote rural housing renovation projects based on the real family conditions of different regions and different elderly people, so as to satisfy the housing needs of elderly people in different family situations, and to ensure the fairness of the housing renovation, so as to promote the realization of the common wealth. Once again, when planning the layout of housing renovation in their villages, the two committees of each village should take into account the local conditions and needs, and plan the location and layout of different housing in the village, so as to help the construction of beautiful villages. Next, in the case of elderly people who encounter housing difficulties and are unable to solve them on their own, the relevant responsibilities and obligations of their children are stipulated, and the village committees and township governments monitor the housing situation of the elderly and urge their children to take on the relevant responsibilities and obligations. Finally, in the case of elderly people in rural areas who are in financial difficulty, it is not possible to provide them with housing subsidies or house renovations in accordance with the standards set out in previous policy documents, but rather to pay more attention to their basic needs, such as daily heating, in their newly constructed houses, so as to ensure that they are helped to the end.

4.2. Emphasize the orientation of social opinion and shape a new type of village culture

The shaping of a new village culture requires correcting unreasonable social opinions and backward concepts of old-age care, and reinforcing the main role of the elderly themselves in transforming their living environment. On the one hand, it is important to pay attention to the guidance of external social opinion, reshape the existing village culture, correct the unreasonable and excessive public opinion constraints, and replace it with a village culture that provides better conditions for the elderly to live in their old age. The government can use television, radio, newspapers and other traditional media to vigorously publicize the new village culture, and actively guide the village public opinion, so that the old people who are bound by the traditional village culture can be released, and can really improve the living environment from their own point of view to improve their personal living standards. On the other hand, changing the traditional concepts of the elderly in rural areas, and re-establishing the idea that the elderly should be nourished and have fun, so that the elderly are moderately concerned about improving their personal quality of life, and subjectively attach real importance to the improvement of their housing conditions.

4.3. Strengthening the economic power of farmers and reinforcing the family's function of providing for the elderly

Improving rural housing conditions is, for the time being, far from enough to rely solely on the Government and society; only by giving full play to the primary role of family care can the housing rehabilitation problems of the elderly in rural areas be better addressed. On the one hand, raising the economic income level of farmers as a whole is the basis for improving the housing environment and community living environment for the rural elderly. For the vast majority of rural elderly people, the improvement of their children's economic conditions and quality of life is the basic prerequisite for their willingness to improve their own living environments; therefore, raising the overall economic income level of rural families can alleviate the mental burden and economic pressure of rural elderly people's housing rehabilitation. On the other hand, to strengthen the housing suitable for the elderly transformation [19], the concept of elderly livable environment construction publicity, guide the children to pay attention to the practical housing needs of the elderly, and promote the children in the housing dimension to help the elderly to improve the quality of housing, improve the function of housing.

5. Conclusion

An in-depth analysis of rural housing rehabilitation in County A reveals that housing rehabilitation in the context of the e-commerce economy exhibits significant spatial differentiation characteristics and complex driving mechanisms. The model comparison analysis shows that the GWR model performs well in explaining the spatial variability of housing remodeling, with its residual sum of squares decreasing from 351.63 in the OLS model to 286.66, and the AIC value optimized from 2821.643 to 2463.61, which proves the importance of spatial heterogeneity in the study of housing remodeling. The spatial autocorrelation analysis reveals that the clustering characteristics of the housing renovation market have been increasing year by year, and housing renovation exhibits a significant positive spatial correlation within the 2800-meter proximity distance threshold, reflecting the trend of coordinated regional development.

The analysis of the importance of influencing factors shows that housing location, transportation accessibility, and park accessibility constitute a core system of elements influencing housing renovation. Threshold effect analysis further reveals the non-linear mechanism of each factor: the price of housing renovation rises slightly within 3 kilometers from the city center, and decreases within 3-17.5 kilometers; there exists a critical value of 4.1 for park accessibility, and the marginal utility diminishes after exceeding this value; public transportation accessibility shows an inverted N-type pattern of change, and has the most significant effect on the price enhancement in the case of accessibility from 2 to 5.

The results of the study provide a scientific basis for the formulation of rural housing renovation policies, and suggest that the government should take into full consideration the spatial variability in the process of housing renovation, implement precise renovation strategies, and strengthen the supporting infrastructure construction, so as to promote the coordination and unity of housing renovation and regional development.

Funding

This research was supported by the Jiangxi Provincial Department of Education, Jiangxi Provincial Science and Technology Research Project "Research on Precise Marketing Strategies of Big Data Empowering Rural Revitalization with Special Agricultural Products" (Project approval code: GJJ2204520).

References

1. Qin, Q., Guo, H., Shi, X., & Chen, K. (2023). Rural E-commerce and county economic development in China. China & World Economy, 31(5), 26-60.
2. Kong, S. T. (2019). E-commerce development in rural China. The Chinese economic transformation, 129.
3. Zhang, Y., Long, H., Ma, L., Tu, S., Li, Y., & Ge, D. (2022). Analysis of rural economic restructuring driven by e-commerce based on the space of flows: The case of Xiaying village in central China. Journal of Rural Studies, 93, 196-209.
4. Li, G., & Qin, J. (2022). Income effect of rural E-commerce: Empirical evidence from Taobao villages in China. Journal of Rural Studies, 96, 129-140.
5. Wei, Y. D., Lin, J., & Zhang, L. (2020). E-commerce, taobao villages and regional development in China. Geographical Review, 110(3), 380-405.
6. Zhang, H., Millan, E., Money, K., & Guo, P. (2025). E-commerce development, poverty reduction and income growth in rural China. Journal of Strategy and Management, 18(1), 148-176.
7. Liu, M., Min, S., Ma, W., & Liu, T. (2021). The adoption and impact of E-commerce in rural China: Application of an endogenous switching regression model. Journal of Rural Studies, 83, 106-116.
8. Couture, V., Faber, B., Gu, Y., & Liu, L. (2021). Connecting the countryside via e-commerce: evidence from China. American Economic Review: Insights, 3(1), 35-50.
9. Tang, W., & Zhu, J. (2020). Informality and rural industry: Rethinking the impacts of E-Commerce on rural development in China. Journal of Rural Studies, 75, 20-29.
10. Zhang, Y., & Lu, Y. (2018, September). Research on the problems and strategies of rural E-commerce in the age of internet+ agriculture. In 2018 14th International Conference on Semantics, Knowledge and Grids (SKG) (pp. 257-260). IEEE.
11. Waghulkar, S., Ganjre, K., Behare, N., & Diwan, N. (2017). A feasibility study for online marketing of agricultural greenhouse products WRT Pune District. International Journal of Management (IJM), 8(1).
12. Li, G., & Zhang, H. (2024). The Efficiency and Challenges of E-Commerce Logistics in Enhancing Market Access for Agricultural Products in Rural China. Law and Economy, 3(2), 31-43.
13. Lin, Y., & Li, C. (2023). Does rural e-commerce agglomeration help expand family farms' size? Evidence from Taobao villages in China's Yangtze River Delta. Electronic Commerce Research, 23(3), 1731-1752.

14. Min, Y., & Tan, C. C. (2022). Research and application of agricultural energy Internet intelligent system for live streaming E-commerce based on MATLAB analysis in China. Energy Reports, 8, 227-239.
15. Feng, X. (2024). Control, exploitation and exclusion: Experiences of small farmer e-tailers in agricultural e-commerce in China. Journal of Agrarian Change, 24(1), e12567.
16. Liu, C., Li, J., & Liu, J. (2015). Rural E-commerce and new model of rural development in China: a comparative study of" Taobao Village" in Jiangsu Province. Asian Agricultural Research, 7(11), 35.
17. Wen Zhao,Jie Zhong & Jiale Lv. (2024). Influencing factors and spatial differentiation of rental housing in a smart city: A GWR model analysis. Measurement: Sensors,33,101126-.
18. Rey-Blanco David,Zofío José L. & González-Arias Julio. (2024). Improving hedonic housing price models by integrating optimal accessibility indices into regression and random forest analyses. Expert Systems With Applications,235,
19. Gao Jinghao. (2021). Analysis of Architectural Design of Country Inns Transformed from Old Houses. Frontiers in Art Research,3.0(2.0),