Research Article

A BP Neural Network-Based GIS-Data-Driven Automated Valuation Framework for Benchmark Land Price

Lei Wu,1 Yu Zhang,2 Yongchang Wei,2 and Fangyu Chen2

1Business College, Shaoxing University, Shaoxing, China
2Institute of Operations Management and System Engineering, School of Business Administration, Zhongnan University of Economics and Law, Wuhan, China

Correspondence should be addressed to Fangyu Chen; cfyconan@126.com

Received 25 January 2022; Accepted 6 April 2022; Published 28 April 2022

Academic Editor: Yu Zhou

Copyright © 2022 Lei Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The automated valuation of benchmark land price plays an essential role in regulating land demand in Chinese real-estate market as the big data are currently accumulated rapidly. However, this problem becomes highly challenging due to the multidimension, large volume, and nonlinearity of the land price-influencing factors. In this paper, an effective data-driven automated valuation framework is proposed for valuing real estate assets by combining a GIS (geographic information system) and neural network technologies. This framework can automatically obtain the values of spatial factors affecting land price from GIS and generate training set data for training the neural network to identify the complex relationship between all kinds of factors and benchmark land prices. The effectiveness and universality of the framework is verified via the data of benchmark land prices in Wuhan. The framework can be applied for automated benchmark land price valuation in other cities.

1. Introduction

In China, the benchmark land price is one of the main means for the government to control land demand at the macro-level. The benchmark land price essentially reflects the average level of land prices in a certain area of a town, as well as its spatial distribution pattern. It can macroscopically guide land space utilization and investment behaviors and can be used as the basis for bulk land valuation. The benchmark land price is relatively stable over certain periods, generally being in line with the real estate market and urban development requirements [1, 2]. However, with the development of the social economy and the progress of urbanization, there exists a gap between the benchmark land price and the actual price of the land market. When this gap reaches a certain level, the benchmark land price needs to be reevaluated, which is a basic business in asset valuation. However, there are defects in the current commonly used benchmark land price valuation approaches. Firstly, there is a large subjectivity in the process of dealing with the benchmark land price impact factors. Secondly, when evaluating an entire city, the resulting large amount of data result in low work efficiency. Facing the huge land transaction market, the application effect of traditional benchmark land price valuation approaches has become increasingly unsatisfactory. It is therefore necessary to establish a reasonable automated benchmark land price valuation framework.

When constructing an automated benchmark land price valuation framework, we are faced with several challenges: (1) benchmark land prices often take an entire city as the valuation object, involving a large number of land parcels. Therefore, benchmark land price valuation has higher requirements for data collection and processing; (2) there are multiple dimensions of the factors affecting the benchmark land price, and many factors have no quantitative indicators. In traditional valuation, the value of these factors is often determined based on experience, which decreases the objectivity of the valuation process; (3) the mechanisms of impact factors on the benchmark land price are complicated, which cannot be reflected by simple linear relations; (4) considering the above three reasons, benchmark land price
2. Literature Review

This part will summarize the related research from three aspects: benchmark land price valuation, automated valuation technology, and artificial intelligence technology in land price valuation. Section 3 introduces the automated valuation framework; Section 4 uses the residential benchmark land price valuation in Wuhan as a data source to verify the effectiveness of the framework. The paper is concluded in Section 5.

2.1. Benchmark Land Price. Wang [5] proposed to predict the price curve of commercial and residential land through the urban population size and income level. The results showed that population size and income level had a positive impact on urban benchmark land price. Moreover, the benchmark land price for different land-use types often varies greatly; for example, the benchmark land price of general commercial land is typically higher than that of residential land. De Groot [6] developed a comprehensive valuation framework of ecological services and socioeconomic benefits. By introducing landscape ecology, this valuation framework could accurately reflect the value of the land itself. Davis and Heathcote [7] found a close relationship between the price of residential land in the United States and the current housing price. Demetriou [8], in a case study based in Cyprus, found that location characteristics, legal factors, physical attributes, and economic conditions had more impact on the value of land than other factors. Yang et al. [9] pointed out that demographic and economic factors have an influence on the residential land price in major cities of China and also found that the GDP has a great influence. Burian et al. [10] divided the impact factors of land price into three aspects: social, environmental, and economic. Compared with social and environmental factors, the impact of economic factors on land price was found to be relatively small. Glumac et al. [11] proposed a hedonic urban land price index which could quantify the spatial effect of land and the impact of macro-policies on land transaction volume. Li et al. [12] concluded that the price of urban land is not only influenced by the city itself but also by its interaction with nearby cities. They calculated the interaction intensity between cities based on a gravity model, in order to obtain the urban land price of a certain area. Yuan et al. [13] focused on the difference of land prices caused by land marketization and fiscal decentralization. Nakamura [14] proposed a land pricing model which considers six factors: entrepreneurship, nature conservation, resource recycling, social vitality, financial viability of local governments, and environmental quality. Tan et al. [15] considered that a newly opened subway station was more likely to raise land prices nearby than a central station. In addition, earthquake risk [16] and noise [17] have also been shown to affect the land price under special circumstances.

From the above analysis, it can be seen that the factors affecting land price mainly focus on the natural environment, economic society, population, infrastructure construction, and the correlations among relevant cities.
of studies in the literature. Taking Cyprus as a case study, Demetriou [23] developed an automated valuation model (AVM) using geographic information system (GIS) data, which had a higher valuation efficiency and greatly reduced the use of time and resources. The general valuation process has also begun to shift from empirical to systematic and standard. Renigier-Biliozor [24] proposed an automated valuation model with decision theory and data mining to assess real-estate characteristics and real-estate prices. In response to the multidimensional factors that affect land prices, Bencure et al. [25] adopted AHP to integrate multiple factors into one model, in order to achieve the automated valuation of land. It can be seen that the most critical process in automated valuation or batch valuation models is dealing with large-scale parameters, as well as integrating multiple parameters with different characteristics into one valuation model.

2.3. Application of Artificial Intelligence Technology in Valuation. Artificial intelligence technology has been applied in assets valuation areas such as box office revenue [26], stock value [27–29], and the auction price of paintings [30]. In real estate valuation, artificial intelligence technology has been widely used, especially in developed countries [31]. The most commonly used models are artificial neural networks (ANNs). ANNs can obtain the nonlinear relationships between real-estate characteristics and real-estate prices but do not have many parameter limits, such as in traditional statistical methods [32].

Peterson and Flanagan [33] found that, compared with the linear hedonic price index model, the artificial neural network evidently had lower pricing error and higher pricing accuracy, which can lead to better predictions in a volatile pricing environment. Follow-up studies indicated that, in different situations, ANN and multiple linear regression each have their own advantages and the advantages and disadvantages of the two are not absolute. When the homogeneity of the data set is low, a valuation method based on AI is more appropriate [34]. ANN, as a “black box” data-driven method, has the major disadvantage of complexity [35], which limits its use by valuators. Multiple regression analysis is more transparent and explanatory than ANN. Geographically weighted regression can be used to achieve an optimal balance between the accuracy of the valuation and the transparency of the methodology [36].

Furthermore, Bin et al. [37] used the long short-term memory (LSTM) model to improve an automatic valuation model based on Boosting Tree to predict house prices in the next few months or a year. Poursaeeet al. [38] believed that the internal and external appearance affect the value of a house and thus proposed to evaluate the visual characteristics of houses by use of convolutional neural networks (CNNs). In addition to artificial neural networks, other methods have also been used in the field of real estate valuation, such as Random Forest [39, 40], Fuzzy Logic System [41, 42], and Genetic Algorithms [43].

The related research illustrates that the wide application of artificial intelligence provides an opportunity to conduct automated valuation. The benchmark land price is an important tool for the government to implement macro-control on the land market, and it is a business in asset valuation. The introduction of any new method must comply with the criteria for asset valuation. As a basic approach for land valuation, the market approach selects multiple factors that affect land value, quantifies the differences between the reference objects and the evaluated land with respect to these factors after determining the weights of these factors, and finally determines the land value. In essence, it is a multi-objective decision problem, and an optimization method can be introduced to improve the decision efficiency [44, 45]. The calculation process of the BP neural network [46] conforms to the basic idea of the traditional market approach, which considers the effect of comparable impact factors. More importantly, BP neural network can make up for the shortcoming of the simple linear assumption in the market approach to the influencing mechanism of the impact factors. Moreover, there are many factors affecting benchmark land prices and, so, the first problem to be solved is data collection and integration of multidimensional factors. Therefore, based on the GIS and BP neural network, we propose a framework of automated benchmark land price valuation with the function of spatial data collection and processing.

3. Automated Valuation Framework

In this paper, we propose an automated benchmark land price valuation framework based on the BP neural network. In this framework, the land use must be determined first, as plots with different uses yield different benchmark land price types. Then, the impact factors of land price are selected, according to the land-use type, to form a value impact system. The sample data to train the neural network are collected with the assistance of the GIS. Finally, after training and testing, the neural network-based valuation model can execute the automated valuation of benchmark land prices.

The basic procedure of the framework is shown in Figure 1.

3.1. Impact Factors Selection. According to different specific uses, different functional attribute values of the same plot yield different benchmark prices. For example, the value of commercial land is mainly related to its geographical location, local economic development, lifestyles of residents, and surrounding business services; the land area, mode of land leasing, location of the district, GDP, paid-in foreign investment, unemployment, tenure of district mayor, gross industrial production, and total industrial asset affect the industrial land price [47]; meanwhile, residential land value is closely related to the surrounding cultural environment, traffic, and public facilities. After determining the specific use of the land, it is necessary to choose the corresponding impact factors. As the most common transaction object in the land trading market, the valuation of residential land has attracted the most attention [48].

For residential land benchmark pricing, impact factors exist in many aspects. Dale and McLaughlin [49] proposed
that land price depends on various factors, such as economy, society, environment, and law. By using the Delphi method, China’s local government found that the impact factors for the land price were business services, public facilities, population density, transportation facilities, and so on [1]. Residential land price is also influenced by the environment and ecology [50]. Xu and Coors [51] took socioeconomic level, regional pressure of environmental constraints, population pressure, and transportation pressure as the most important impact factors affecting urban residential land price. Yang et al. [9] used geographically weighted regression (GWR) to analyze the relationships between residential land prices and housing prices. They screened out three main impact factors—immigrant population, gross domestic product (GDP), and residential investment—and found that the impact of GDP on the residential land price was more significant than other factors in the prior three years. It has been observed that location, business services, and education resources have a great influence on the spatial layout of residential land prices. The benchmark land price is the basic reference for residential land prices. The government also considers the same factors when pricing. As the basic consensus of the industry, to meet the requirements of the evaluation criteria, in this paper, we chose business services, environment, traffic, public facilities, and GDP per capita as the five dimensions for impact factors of residential land price. According to the different connotations of these five dimensions, the specific impact factors of residential land price were summarized into the valuation system shown in Table 1.

3.2. Data Collection and Preprocessing. The premise of automated valuation is that there are a lot of standardized samples and data. As benchmark land price valuation is carried out for an entire administrative region, we imported the corresponding plot information into GIS, in order to establish a geographic information database for collecting and processing data in real time.

It should be noted that some impact factors can be quantified by GIS, such as distance to the nearest business service center and driving time to the nearest train station. The accessibility of the road network can be quantified by the number of arterial roads around the plot. The indices which are difficult to quantify, such as intensity of nearby scenic location and intensity of medical support, can be processed by an index scoring method. The standards for index scoring are shown in Table 2, which is offered by our project collaborator.

3.3. Backpropagation Neural Network Training. In our neural network, the input value of the input layer corresponds to the matrix of preprocessed data of impact factors. The output layer corresponds to the predicted value of the residential benchmark land price without renormalization. The quantified values of 13 benchmark land price impact factors were selected as the input, while the output is the residential benchmark land price. The proposed BP neural network, therefore, reflects the mapping relationship of 13 independent variables to 1 dependent variable.

Based on the above analysis, the specific operation process was as follows:

Step 1: establish the value impact system of residential benchmark land price valuation

Step 2: use GIS and index scoring method to collect impact factor data

Step 3: process the collected parameters and data by the mapminmax function

\[ X = \frac{2(x - x_{\text{min}})}{x_{\text{max}} - x_{\text{min}}} - 1, \]

where \( X \) is the value after processing, \( x \) is the value of the impact factor, \( x_{\text{min}} \) is the minimum value in this kind of impact factor, and \( x_{\text{max}} \) is the maximum value in this kind of impact factor

Step 4: establish a BP neural network model and determine the number of network layers, the number of input layer nodes, the number of hidden layer nodes, the number of output layer nodes, training functions, performance functions, and so on.
Step 5: train the model to process residential benchmark land price samples
Step 6: test the BP neural network model. When it successfully passes the test, the model can then be used to evaluate and update the residential benchmark land price in real time.

4. Case Analysis

4.1. Data Collection. We selected the most representative data samples from the 10 central districts of Wuhan, based on the latest residential benchmark land price data for 2018. The 110 groups of residential benchmark land price valuation project information were used to train and test the model. The related data sources are shown in Table 3.

All the collected information was imported to the GIS database to draw an electronic vector map of Wuhan, as shown in Figure 2. This map can be used to obtain the geographic characteristics of the impact factors, which were used to analyze the correlations between the target plot benchmark price and each impact factor.

4.2. Valuation Model Construction

4.2.1. Design of BP Neural Network Structure. The effect of the BP neural network model is affected by the number of hidden layer nodes [52]. A reasonable number of hidden
layer nodes not only leads to better performance but can also greatly save the amount of training time required. We selected the following empirical formula to determine the number of hidden layer nodes:

\[ M = \sqrt{n + m + a}, \]  

where \( m \) is the number of input layer nodes; \( n \) is the number of output layer nodes; and \( a \) is a number in \([1, 10]\).
According to this formula, the ideal number of hidden layer nodes was calculated to be between 4 and 14. After a series of training simulations and tests, the optimal BP structure was determined to be 13-10-1. The other model parameters and equations are shown in Table 6.

### 4.2.2. Training of BP Neural Network Model.

Considering the limited sample size, we used 10-fold cross-validation to train the BP network, obtaining the following 10 sets of test residential benchmark land price results (Figure 3).

The test error results of the ten groups are shown in Table 7.

From the previous results, it can be seen that there was a certain gap in the size of the MSE obtained by these 10 groups of experiments, but the overall error was within 2%, which is acceptable in practice. We randomly selected a group of data sets to evaluate the experimental results; the results are shown in Figure 4.

According to the above results, it can be seen that the model reached the best verification state at the 1st epoch, where the mean square error was at the lowest value. Moreover, the regression analysis shows that the training data set, testing data set, and overall data set all showed a linear relationship between the values, with a correlation coefficient close to 1, as shown in Figure 5.

### 4.2.3. Test of BP Neural Network Model.

In order to make the test results more intuitive and easier to understand, the training output obtained by the model was denormalized to the predicted value of residential benchmark land price, as shown in Table 8. We compared the actual value with the predicted value, and the absolute error and relative error were given as well. The relative error is calculated as follows:

$$\varepsilon = \left| \frac{y - \hat{y}}{y} \right|$$

where $y$ is the actual benchmark land price, $\hat{y}$ is the forecast benchmark land price, and $\varepsilon$ is the relative error. The result of the error analysis is shown in Table 9.

According to the obtained results, it can be seen that the error of the model itself was relatively small, further reflecting the applicability of the framework. It proves that, in addition to real estate [24], the automatic evaluation can also be used in land evaluation [8, 22, 23, 53] as well. Moreover, from the results, the validity of the five dimensions of impact factors on residential land price is consistent with the previous practice results of the valuation industry.

### 4.3. Results Analysis.

Using GIS, the error comparison data were reflected in the form of a heat map, as shown in Figure 6. Figure 6(a) illustrates that Wuhan’s benchmark land prices showed a trend of gradual decreasing from the central city to its surroundings, consistent with the distribution of land prices in most Chinese cities. Among the surrounding urban areas, it is worth noting that the prices in the Huangpi and Xinzhou districts (in the northeast) were higher than

| Parameter | Max   | Min   | Mean   | Standard deviation | Kurtosis | Skewness | Coefficient of variation |
|-----------|-------|-------|--------|--------------------|----------|----------|-------------------------|
| $D_{bs}$  | 32100 | 382   | 3211.991 | 3915.932           | 34.133   | 5.309    | 1.219                   |
| $G$       | 132141| 57810 | 8493.173 | 25478.534          | -0.655   | 0.799    | 0.3                     |
| $D_{sr}$  | 7800  | 97    | 1269.327 | 1268.117           | 9.283    | 2.734    | 0.999                   |
| $I_{sc}$  | 9     | 1     | 3.027   | 2.436              | -0.679   | 0.871    | 0.805                   |
| $D_p$     | 6200  | 32    | 1270.736 | 1053.044           | 6.696    | 2.212    | 0.829                   |
| $D_m$     | 10100 | 97    | 1435.636 | 1305.607           | 17.186   | 3.19     | 0.909                   |
| $D_h$     | 54000 | 912   | 8106.227 | 8863.934           | 9.307    | 2.647    | 1.093                   |
| $I_h$     | 10    | 2     | 4.855   | 2.559              | -0.926   | 0.546    | 0.527                   |
| $D_{bas}$ | 2100  | 31    | 482.273 | 446.313            | 3.219    | 1.769    | 0.925                   |
| $D_{sub}$ | 32600 | 29    | 4506.845 | 6943.069           | 4.835    | 2.354    | 1.541                   |
| $T_i$     | 68    | 11    | 30.845  | 190.591            | -0.366   | 0.758    | 0.448                   |
| $T_a$     | 90    | 19    | 52.1    | 15.998             | -0.688   | 0.128    | 0.307                   |
| $A_r$     | 6     | 2     | 4.155   | 0.96               | -0.648   | 0.128    | 0.231                   |
| Price     | 6059  | 408   | 2236    | 1403.913           | -0.157   | 0.866    | 0.628                   |

| Parameter | Value |
|-----------|-------|
| Network layer | 3     |
| Number of input layer nodes | 13    |
| Normalization function $(map\_min\_max)$ | $2(x - x_{min})/x_{max} - x_{min} - 1$ |
| Number of hidden layer nodes | 10    |
| Error tolerance | 0.01  |
| Hidden layer activation function | $log\_sig(x) = 1/1 + e^{-x}$ |
| Output layer activation function | $tan\_sig(x) = 2/1 + e^{-2x}$ |
| Learning rate | 0.01  |
| Mean square error (MSE) | $mse = \sum_{i=1}^{N} e_i^2/N$ |

$N$ is the vector length and $e_i$ is the error caused by the predicted value of the $i$th sample—the actual value.
Figure 3: Ten-fold cross-validation results.

Table 7: MSE (mean square error) of ten groups.

| Group | MSE     |
|-------|---------|
| Group 1|0.0186  |
| Group 2|0.013624|
| Group 3|0.015339|
| Group 4|0.0038203|
| Group 5|0.01039 |
| Group 6|0.0055963|
| Group 7|0.014904 |
| Group 8|0.011278 |
| Group 9|0.007668 |
| Group 10|0.01005 |

Best Validation Performance is 0.013612 at epoch 1

Figure 4: Model iteration curve.
Figure 5: Dynamic changes of the metrics: (a) training set, (b) test set, and (c) all data sets.

Table 8: Training value and error analysis of benchmark land price.

| ID | Actual price | Predict price | Absolute error | Relative error (%) |
|----|--------------|---------------|----------------|-------------------|
| 1  | 11270        | 10999.46462   | 270.5353781    | 2.40              |
| 2  | 10575        | 10659.02123   | 84.0212615     | 0.79              |
| 3  | 9860         | 8826.97559    | 1033.02441     | 10.48             |
| 4  | 6894         | 7187.927573   | 293.9275729    | 4.26              |
| 5  | 6522         | 7401.07505    | 879.07505      | 13.48             |
| 6  | 6487         | 7234.729173   | 747.729173     | 11.53             |
| 7  | 7037         | 6535.341643   | 501.658357     | 7.13              |
| 8  | 6659         | 7747.381349   | 1088.381349    | 16.34             |
| 9  | 6306         | 5906.933983   | 399.0660173    | 6.33              |
| 10 | 6732         | 6434.807979   | 297.1920212    | 4.41              |
| 11 | 6845         | 5251.461686   | 1593.538314    | 23.28             |
| 12 | 6725         | 5674.98795    | 1050.01205     | 15.61             |
| 13 | 1895         | 1661.540296   | 233.459703     | 13.22             |
| 14 | 1895         | 2283.207164   | 388.207163     | 20.49             |
| 15 | 1895         | 2670.394473   | 775.394473     | 40.92             |
| 16 | 1895         | 1908.185993   | 13.1859925     | 0.70              |
| 17 | 1835         | 1934.873953   | 99.873952      | 5.44              |
| 18 | 1835         | 1805.847641   | 29.1523589     | 1.59              |
| 19 | 1835         | 2855.693002   | 1020.693002    | 55.62             |
| 20 | 3919         | 4632.610073   | 713.610073     | 18.21             |
| 21 | 5187         | 3953.243253   | 1233.756747    | 23.79             |
| 22 | 5187         | 4556.753877   | 630.246129     | 12.15             |
| 23 | 4069         | 3953.086687   | 115.913312     | 2.85              |
| 24 | 4025         | 4598.903646   | 573.9036462    | 14.26             |
| 25 | 4025         | 3475.743022   | 549.2569775    | 13.65             |
| 26 | 4699         | 4949.017147   | 250.017147     | 5.32              |
| 27 | 4515         | 4217.909246   | 297.090753     | 6.58              |
| 28 | 4025         | 4361.910531   | 336.910531     | 8.37              |
| 29 | 3815         | 3931.892256   | 116.8922562    | 3.06              |
| 30 | 4463         | 3607.50109    | 855.4989099    | 19.17             |
| 31 | 3520         | 4055.871363   | 535.8713631    | 15.22             |
| 32 | 3415         | 4836.723094   | 1421.723094    | 41.63             |
| 33 | 3415         | 4104.835694   | 689.8356941    | 20.20             |
| 34 | 3415         | 3345.341087   | 69.65891264    | 2.04              |
| 35 | 3212         | 2956.042133   | 255.9578668    | 7.97              |
| 36 | 3160         | 3562.980096   | 402.9800956    | 12.75             |
| ID  | Actual price | Predict price | Absolute error | Relative error (%) |
|-----|--------------|---------------|----------------|--------------------|
| 38  | 3160         | 2629.756478   | 530.2435215    | 16.78              |
| 39  | 3031         | 2332.130699   | 698.8693008    | 23.06              |
| 40  | 3031         | 2275.117362   | 755.8826377    | 24.94              |
| 41  | 3020         | 3297.434768   | 277.4347683    | 9.19               |
| 42  | 1835         | 2039.899685   | 204.899684     | 11.17              |
| 43  | 1835         | 2265.600497   | 430.6004969    | 23.47              |
| 44  | 1835         | 2145.251377   | 310.2513772    | 16.91              |
| 45  | 1835         | 1750.940692   | 84.0593081     | 4.58               |
| 46  | 1835         | 1715.068536   | 119.068536     | 6.54               |
| 47  | 1835         | 2009.052246   | 174.052246     | 9.49               |
| 48  | 3919         | 4717.376442   | 798.3764421    | 20.37              |
| 49  | 3919         | 3936.573752   | 17.5737515     | 0.45               |
| 50  | 3878         | 4018.82728    | 140.827279     | 3.63               |
| 51  | 3878         | 4527.441176   | 649.4411755    | 16.75              |
| 52  | 3878         | 3351.890981   | 526.890981     | 15.57              |
| 53  | 3878         | 4077.41773    | 199.417731     | 5.14               |
| 54  | 3720         | 3876.882547   | 156.882547     | 4.12               |
| 55  | 3720         | 3440.412065   | 279.587935     | 7.52               |
| 56  | 3738         | 4000.18481    | 262.1848097    | 6.51               |
| 57  | 2560         | 2133.729179   | 426.7291792    | 16.65              |
| 58  | 2560         | 1420.83742    | 1139.16258     | 8.16               |
| 59  | 2432         | 2360.789      | 71.2109995     | 3.03               |
| 60  | 2432         | 2182.210975   | 249.789025     | 10.57              |
| 61  | 2314         | 2399.249307   | 85.24930743    | 3.68               |
| 62  | 2287         | 2195.412277   | 91.5877227     | 4.00               |
| 63  | 2287         | 2077.668831   | 209.331169     | 9.15               |
| 64  | 2287         | 2118.524751   | 168.475249     | 7.87               |
| 65  | 8223         | 8289.125721   | 66.1257212     | 0.80               |
| 66  | 8196         | 8811.482363   | 615.4823632    | 7.11               |
| 67  | 8346         | 8875.026878   | 529.0268777    | 6.34               |
| 68  | 8429         | 6519.027601   | 1909.027699    | 22.66              |
| 69  | 8429         | 7357.134108   | 1071.856892    | 12.72              |
| 70  | 8289         | 8249.076233   | 1355.076233    | 19.66              |
| 71  | 6725         | 7561.213903   | 836.2139027    | 12.43              |
| 72  | 6725         | 7638.856438   | 913.8564382    | 13.59              |
| 73  | 5899         | 5864.577641   | 34.4235919     | 0.58               |
| 74  | 5730         | 5774.474114   | 44.4741143     | 0.78               |
| 75  | 5000         | 5044.700666   | 44.7006623     | 0.89               |
| 76  | 5000         | 4685.049398   | 314.9506024    | 6.30               |
| 77  | 5153         | 5721.648844   | 568.6488439    | 11.04              |
| 78  | 5187         | 5610.268719   | 423.2687185    | 8.16               |
| 79  | 5187         | 4225.477747   | 961.522256     | 18.54              |
| 80  | 4289         | 3399.520371   | 889.4769294    | 20.74              |
| 81  | 4289         | 4000.231472   | 288.768528     | 6.73               |
| 82  | 3020         | 2155.479456   | 864.5205444    | 28.63              |
| 83  | 3020         | 2671.458994   | 348.541006     | 13.54              |
| 84  | 2945         | 3029.856262   | 84.8562621     | 2.88               |
| 85  | 2945         | 2858.285059   | 86.7149149     | 2.94               |
| 86  | 2880         | 2405.414772   | 474.585228     | 16.48              |
| 87  | 2880         | 2812.938556   | 67.06144369    | 2.33               |
| 88  | 2880         | 2970.800195   | 90.80019471    | 3.15               |
| 89  | 2880         | 2498.657799   | 381.342013     | 13.24              |
| 90  | 2287         | 2946.780872   | 659.7808721    | 28.85              |
| 91  | 2170         | 1862.63273    | 307.3672698    | 16.16              |
| 92  | 2170         | 2242.295835   | 72.29583527    | 3.33               |
| 93  | 2170         | 2571.335509   | 401.3355094    | 16.49              |
| 94  | 1895         | 2366.274084   | 471.2740842    | 24.87              |
| 95  | 1895         | 2828.802786   | 933.8027861    | 49.28              |
| 96  | 1895         | 1997.013051   | 102.0130515    | 5.38               |
those of the Caidian and Hannan districts (in the southwest). This is mainly because there are more natural scenic spots in the Huangpi district than in the southwest urban districts, and an important part of the nationally planned Wuhan Yangtze River Economic Belt, The Yangtze River New City, is located at the junction of Huangpi and Xinzhou districts. These factors have promoted an increase in land prices. In Figure 6(b), it can be seen that, although the predicted value

![Figure 6: Actual land price and relative error heat map: (a) actual land price and (b) relative error.](image)

Table 8: Continued.

| ID | Actual price | Predict price | Absolute error | Relative error (%) |
|----|--------------|---------------|----------------|-------------------|
| 97 | 1895         | 1751.015159   | 143.984841     | 7.60              |
| 98 | 1895         | 1975.518443   | 80.51844309    | 4.25              |
| 99 | 1895         | 1671.076755   | 223.9232445    | 11.82             |
| 100| 7037         | 5139.590024   | 1897.409976    | 26.96             |
| 101| 4654         | 3993.050653   | 660.949347     | 14.20             |
| 102| 4463         | 3198.501774   | 1264.498226    | 28.33             |
| 103| 3720         | 4700.273197   | 980.2731969    | 26.35             |
| 104| 3520         | 3110.062106   | 409.9378942    | 11.65             |
| 105| 2314         | 2125.805912   | 188.1940884    | 8.13              |
| 106| 9052         | 9357.966337   | 305.9663367    | 3.38              |
| 107| 7878         | 9116.477395   | 1238.477395    | 15.72             |
| 108| 5622         | 5651.900916   | 29.900916      | 0.53              |
| 109| 5309         | 4895.078181   | 413.9218192    | 7.80              |
| 110| 2890         | 2453.368521   | 436.6314791    | 15.11             |

Table 9: Error analysis.

| Error                          | Value  |
|-------------------------------|--------|
| Relative average absolute error (RAAEE) | 0.13   |
| Root means absolute error (RMAE)     | 22.133 |
| Coefficient of determination ($R^2$) | 0.913  |
| Standard deviation ratio (RSR)      | 416.002|
| Nash–Sutcliffe coefficient (NS)     | 0.915  |
| Root means squared error (RMSE)     | 642.659|
| Mean absolute percentage error (MAPE) | 12.665 |
has a good fit with the actual benchmark land price, there were still some deviations in some regions. The deviations indicate that the price in the northeast was overestimated and the price in the southwest was underestimated. The sensitivity of the model to environmental and economic factors caused the price of the urban area in the northeast to be overestimated. The urban area in the southwest is a national automobile industry base, and the related employed population has a high demand for housing. The impact mechanism of the industrial base on local land prices was not fully reflected in the model. However, these factors have strong localization features. Under the premise of ensuring the universality of the framework, these factors can be added separately when conducting the valuation for specific cities.

Overall, the main reasons for the deviations between the predicted value and the actual value of the benchmark land price were as follows: (1) when quantifying the impact factors of residential benchmark land price, some factors used an index scoring method, which affected the accuracy of the valuation. (2) The heterogeneity of certain factors was ignored. For example, when considering the impact of the school district, only the distance to the nearest primary and middle schools was selected as the main consideration. However, school rankings have a more obvious impact on land prices, which was ignored in the model. (3) The analysis of macrofactors was not very comprehensive; for example, the impact of some government regional support policies was not considered in depth.

5. Conclusion

As one of the main means of the Chinese government's macrocontrol of land demand, the benchmark land price needs to be adjusted, in a timely manner, along with economic development. The traditional benchmark land price valuation process has the drawback of being subjective and unable to process large amounts of data. In recent years, the development and application of the GIS technology has provided a basis for the collection and real-time update of large-scale geographic data. Therefore, in this paper, we propose an automated valuation framework based on GIS data and neural networks. This framework automatically extracts spatial data which affects land prices through the GIS database, thus reducing data distortion. Using an impact factor valuation system to avoid the incommensurability of the data, we finally designed an automated valuation model based on the BP neural network model, effectively reducing the subjectivity and uncertainty of the valuation.

In the case study, in which we took Wuhan as an example, we found that although there were still some errors in the model training and testing results, they generally met the accuracy requirements. It can be seen that the BP neural network had good performance in automated valuation. The model can also be applied to the prediction of residential benchmark land prices outside Wuhan. If there are a few factors that affect the benchmark land price significantly, the model can be further adjusted. As data continue to accumulate, the fitness of the BP neural network can be further improved. This is of great practical significance for regulating the order of the Chinese land market, formulating relevant land management laws and allowing the public to more conveniently receive land transaction information.

Due to the project schedule requirements, the number of samples data provided by the collaborator is not large, and the effectiveness of the proposed framework needs to be further verified with a larger data set in the future work. In addition, except for the spatial data obtained from the GIS, more data require manual collection. The way to introduce data mining and web crawling technology to improve the automation of the proposed framework is also a focus of future work.

Data Availability

The data are available on request to the corresponding author after the paper is published.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the major projects of National Social Science Foundation of China Study and explains the spirit of the Fifth Plenary Session of the 19th CPC Central Committee (under grant no. 21ZDA028), the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (under grant no. 21YJA630094), and Teaching and Research Project of Zhongnan University of Economics and Law in 2018 (under grant no. YB2018019).

References

[1] L. H. Li and A. Walker, “Benchmark pricing behaviour of land in China’s reforms,” *Journal of Property Research*, vol. 13, no. 3, pp. 183–196, 1996.

[2] C. r. Ding, “The benchmark land price system and urban land use efficiency in China,” *Chinese Geographical Science*, vol. 11, no. 4, pp. 306–314, 2001.

[3] S. P. Chen, D. C. Zhuang, and H. X. Zhang, “Gis-based spatial autocorrelation analysis of housing prices oriented towards a view of spatiotemporal homogeneity and nonstationarity: a case study of Guangzhou, China,” *Complexity*, vol. 2020, Article ID 1079024, 16 pages, 2020.

[4] Z. Xu and Q. Li, “Integrating the empirical models of benchmark land price and gis technology for sustainability analysis of urban residential development,” *Habitat International*, vol. 44, pp. 79–92, 2014.

[5] R. Wang, “The structure of Chinese urban land prices: estimates from benchmark land price data,” *The Journal of Real Estate Finance and Economics*, vol. 39, no. 1, pp. 24–38, 2009.

[6] R. De Groot, “Function-analysis and valuation as a tool to assess land use conflicts in planning for sustainable, multifunctional landscapes,” *Landscape and Urban Planning*, vol. 75, no. 3–4, pp. 175–186, 2006.

[7] M. A. Davis and J. Heathcote, “The price and quantity of residential land in the United States,” *SRN Electronic Journal*, vol. 54, no. 8, pp. 2595–2620, 2007.

[8] D. Demetriou, “A spatially based artificial neural network mass valuation model for land consolidation,” *Environment
Complexity

and Planning B: Urban Analytics and City Science, vol. 44, no. 5, pp. 864–883, 2017.

[9] S. Yang, S. Hu, W. Li, C. Zhang, and J. E. A. Torres, “Spatiotemporal effects of main impact factors on residential land price in major cities of China,” Sustainability, vol. 9, no. 11, p. 2050, 2017.

[10] J. Burian, K. Macku, J. Zimmermannova, and B. Kocvarova, “Spatio-temporal changes and dependencies of land prices: a case study of the city of olomouc,” Sustainability, vol. 10, no. 12, p. 4831, 2018.

[11] B. Glumac, M. Herrera-Gomez, and J. Licheron, “A hedonic urban land price index,” Land Use Policy, vol. 81, pp. 802–812, 2019.

[12] C. Li, Y. Meng, Y. Li, J. Ge, and C. Zhao, “Inter-metropolitan land-price characteristics and patterns in the beijing-tianjin-hebei urban agglomeration in China,” Sustainability, vol. 11, no. 17, p. 4726.

[13] F. Yuan, Y. D. Wei, and W. Xiao, “Land marketization, fiscal decentralization, and the dynamics of urban land prices in transitional China,” Land Use Policy, vol. 89, Article ID 104208, 2019.

[14] H. Nakamura, “Relationship among land price, entrepreneurship, the environment, economics, and social factors in the value assessment of Japanese cities,” Journal of Cleaner Production, vol. 217, no. 1, pp. 144–152, 2019.

[15] R. Tan, Q. He, K. Zhou, and P. Xie, “The effect of new metro stations on local land use and housing prices: the case of wuhan, China,” Journal of Transport Geography, vol. 79, no. 1, Article ID 102488, 2019.

[16] T. Gu, M. Nakagawa, M. Saito, and H. Yamaga, “Public perceptions of earthquake risk and the impact on land pricing: the case of the uemachi fault line in Japan,” The Japanese Economic Review, vol. 69, no. 4, pp. 374–393, 2018.

[17] J. Batog, I. Forys, R. Gaca, M. Gluszak, and J. Konovalczuk, “Investigating the impact of airport noise and land use restrictions on house prices: evidence from selected regional airports in Poland,” Sustainability, vol. 11, no. 2, p. 412, 2019.

[18] B. J. L. Berry and R. S. Bednarz, “A hedonic model of prices and assessments for single-family homes: does the assessor follow the market or the market follow the assessor?” Land Economics, vol. 51, no. 1, pp. 21–40, 1975.

[19] R. K. Pace and O. W. Gilley, “Estimation employing a priori information within mass appraisal and hedonic pricing models,” The Journal of Real Estate Finance and Economics, vol. 3, no. 1, pp. 55–72, 1990.

[20] I. A. O. A. Officers, Standard on Automated Valuation Models (AVMs). International Association of Assessing Officers, Kansas City, MO, USA, 2003.

[21] P. Aragonés-Beltrán, J. Aznar, J. Ferris-Oriate, and M. García-Melón, “Valuation of urban industrial land: an analytic network process approach,” European Journal of Operational Research, vol. 185, no. 1, pp. 322–339, 2008.

[22] S. Metzner and A. Kindt, “Determination of the parameters of automated valuation models for the hedonic property valuation of residential properties,” International Journal of Housing Markets and Analysis, vol. 11, no. 1, pp. 73–100, 2018.

[23] D. Demetriou, “Automating the land valuation process carried out in land consolidation schemes,” Land Use Policy, vol. 75, pp. 21–32, 2018.

[24] M. Renigier-Bilozor, A. Janowski, and M. d’Amato, “Automated valuation model based on fuzzy and rough set theory for real estate market with insufficient source data,” Land Use Policy, vol. 87, Article ID 104021, 2019.

[25] J. C. Bencure, N. K. Tripathi, H. Miyazaki, S. Ninsawat, and S. M. Kim, “Development of an innovative land valuation model (ilvm) for mass appraisal application in sub-urban areas using ahp: an integration of theoretical and practical approaches,” Sustainability, vol. 11, no. 13, p. 3731, 2019.

[26] D. Barman, N. Chowdhury, and R. K. Singha, “To predict possible profit/loss of a movie to be launched using mlp with back-propagation learning,” in Proceedings of the 2012 International Conference on Communications, Devices and Intelligent Systems (CODIS), 2012 International Conference on Communications, Devices and Intelligent Systems (CODIS), pp. 322–325, Kolkata, India, December 2012.

[27] Q. Cao, K. B. Leggio, and M. J. Schniederjans, “A comparison between fama and French’s model and artificial neural networks in predicting the Chinese stock market,” Computers & Operations Research, vol. 32, no. 10, pp. 2499–2512, 2005.

[28] S. S. Roy, R. Chopra, K. C. Lee, C. Spampinato, and B. M. ivatlood, “Random forest, gradient boosted machines and deep neural network for stock price forecasting: a comparative analysis on south Korean companies,” International Journal of Ad Hoc and Ubiquitous Computing, vol. 33, no. 1, pp. 62–71, 2020.

[29] A. Bose, C. H. Hsu, S. S. Roy, K. C. Lee, B. Mohammadivatloo, and S. Abimannan, “Forecasting stock price by hybrid model of cascading multivariate adaptive regression splines and deep neural network,” Computers & Electrical Engineering, vol. 95, Article ID 107405, 2021.

[30] M. Aubry, R. Kräussl, G. Manso, and C. Spoenjers, “Machine learning, human experts, and the valuation of real assets,” Tech. Rep., CFS Working Paper Series, , 2019.

[31] R. B. Abidoye and A. P. C. Chan, “Valuers’ receptiveness to the application of artificial intelligence in property valuation,” Pacific Rim Property Research Journal, vol. 23, no. 2, pp. 175–193, 2017.

[32] C. Chun Lin and S. B. Mohan, “Effectiveness comparison of the residential property mass appraisal methodologies in the USA,” International Journal of Housing Markets and Analysis, vol. 4, no. 3, pp. 224–243, 2011.

[33] S. Peterson and A. Flanagan, “Neural network hedonic pricing models in mass real estate appraisal,” Journal of Real Estate Research, vol. 31, no. 2, pp. 147–164, 2009.

[34] J. Zurada, A. Levitan, and J. Guan, “A comparison of regression and artificial intelligence methods in a mass appraisal context,” Journal of Real Estate Research, vol. 33, no. 3, pp. 349–388, 2011.

[35] G. Spellman, “An application of artificial neural networks to the prediction of surface ozone concentrations in the United Kingdom,” Applied Geography, vol. 19, no. 2, pp. 123–136, 1999.

[36] W. J. McCluskey, M. McCord, P. T. Davis, M. Haran, and D. McIlhatton, “Prediction accuracy in mass appraisal: a comparison of modern approaches,” Journal of Property Research, vol. 30, no. 4, pp. 239–265, 2013.

[37] J. Bin, S. Tang, Y. Liu et al., “Regression model for appraisal of real estate using recurrent neural network and boosting tree,” in Proceedings of the 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA), pp. 209–213, Beijing, China, September 2017.

[38] O. Poursaeed, T. A. V. S. Matera, and S. Belongie, “Vision-based real estate price estimation,” Machine Vision and Applications, vol. 29, no. 4, pp. 667–676, 2018.

[39] A. Prinzie and D. Van den Poel, “Random forests for multiclass classification: random multinomial logit,” Expert Systems with Applications, vol. 34, no. 3, pp. 1721–1732, 2008.
[40] E. A. Antipov and E. B. Pokryshevskaya, "Mass appraisal of residential apartments: an application of random forest for valuation and a cart-based approach for model diagnostics," *Expert Systems with Applications*, vol. 39, no. 2, pp. 1772–1778, 2012.

[41] H. Kuşan, O. Aytekin, and İ. Özdemir, "The use of fuzzy logic in predicting house selling price," *Expert Systems with Applications*, vol. 37, no. 3, pp. 1808–1813, 2010.

[42] S. Amri and G. A. Tularam, "Performance of multiple linear regression and nonlinear neural networks and fuzzy logic techniques in modelling house prices," *Journal of Mathematics and Statistics*, vol. 8, no. 4, pp. 419–434, 2012.

[43] J. J. Ahn, H. W. Byun, K. J. Oh, and T. Y. Kim, "Using ridge regression with genetic algorithm to enhance real estate appraisal forecasting," *Expert Systems with Applications*, vol. 39, no. 9, pp. 8369–8379, 2012.

[44] S. Teng, L. Hay Lee, and E. Peng Chew, "Multi-objective ordinal optimization for simulation optimization problems," *Automatica*, vol. 43, no. 11, pp. 1884–1895, 2007.

[45] G. Kou, H. Xiao, M. H. Cao, and L. H. Lee, "Optimal computing budget allocation for the vector evaluated genetic algorithm in multi-objective simulation optimization," *Automatica*, vol. 129, Article ID 109599, 2021.

[46] P. Samui, S. S. Roy, and V. E. Balas, *Handbook of Neural Computation*, Academic Press, Cambridge, Massachusetts, 2017.

[47] J. Cheng, "Data analysis of the factors influencing the industrial land leasing in shanghai based on mathematical models," *Mathematical Problems in Engineering*, pp. 1–11, 2020.

[48] M. Al-Omari, "The Role of Reliable Land Valuations in Land Management and Land Administration Systems efficiency," in *Proceedings of the FIG Working Week*, Stockholm, Sweden, June 2008.

[49] P. Dale and J. McLaughlin, *Land Administration*, Oxford University Press, Oxford, UK, 2003.

[50] R. B. Hammer, S. I. Stewart, R. L. Winkler, V. C. Radeloff, and P. R. Voss, "Characterizing dynamic spatial and temporal residential density patterns from 1940–1990 across the north central United States," *Landscape and Urban Planning*, vol. 69, no. 2-3, pp. 183–199, 2004.

[51] Z. Xu and V. Coors, "Combining system dynamics model, gis and 3d visualization in sustainability assessment of urban residential development," *Building and Environment*, vol. 47, no. 1, pp. 272–287, 2012.

[52] E. Bahadır, "Prediction of prospective mathematics teachers’ academic success in entering graduate education by using back-propagation neural network," *Journal of Education and Training Studies*, vol. 4, no. 5, pp. 113–122, 2016.

[53] D. Demetriou, "The assessment of land valuation in land consolidation schemes: the need for a new land valuation framework," *Land Use Policy*, vol. 54, pp. 487–498, 2016.