Trends in Social Sciences and Humanities Research
ISSN: 2959-9938
DOI: 10.61784/tsshr3008

ANALYSIS AND FORECAST OF THE AVERAGE SALES PRICE OF RESIDENTIAL COMMERCIAL HOUSING

XinYue Dang
School of Mathematics and Statistics, Guangxi Normal University, Guilin 541006, Guangxi, China.
Corresponding Email: 845387601@qq.com

Abstract: In recent years, China's residential commercial housing market has shown pronounced regional disparities, especially between large cities and smaller urban areas. This study utilizes cluster analysis to compare market conditions in 2020 and 2002. In 2020, regions were categorized into three tiers, with Beijing and Shanghai exhibiting the most significant differences. In contrast, the 2002 data divided regions into four tiers, with Guangdong Province standing alone in the first tier, indicating a more balanced market. From 2012 to 2020, Beijing's average housing prices increased substantially, whereas changes in the Guangxi Zhuang Autonomous Region were more moderate, reflecting significant regional economic disparities. Neural network models are used to make predictions on the data. This empirical analysis underscores the diversity and economic factors influencing China's residential commercial housing market. By comparing different regions and examining temporal sequences, the study provides theoretical insights into regional development imbalances, emphasizing the need for precise regional policies to promote coordinated market development.
Keywords: Average sales price of residential commercial housing; Cluster analysis; Neural networks; Forecast

1 INTRODUCTION

Commercial residential housing refers to properties planned and developed by state real estate development institutions and various enterprises, available for sale, rental, or mortgage to private homeowners. Purchasing commercial residential housing is the primary means through which individuals and families acquire housing[1].

Real estate investment, as one of the three major sectors of fixed asset investment, holds considerable significance. Real estate is not only a key component of the social consumption market but also a pillar industry driving economic development[2]. A critical variable reflecting the development of real estate is property prices. The year 2008 marked a turning point in national policy; in response to rising property prices, the government began adjusting real estate policies to regulate the market. In 2010, to curb the rapid increase in housing prices and reduce speculative purchases, the government issued the "New Ten National Articles." This policy aimed to counteract the rapid rise in prices and speculation. Subsequently, in 2011, the "New Eight National Articles" were introduced to further control the excessive increase in housing prices[3]. After 2013, following two to three years of emergency suppression policies, the government would relax policies to adapt to a slowdown in economic growth or to address inventory reduction needs. This means that when the market declines or demand decreases, the government may appropriately relax policies to promote real estate sales[4]. In 2018, despite multiple rounds of policy regulation, housing prices continued to show a persistent upward trend[5].

In recent years, influenced by significant pandemic risks and policies in certain countries, housing prices have started to decline significantly. By October 2022, the number of cities experiencing a decrease in commercial residential property prices had markedly increased among 70 major and medium-sized cities, with prices falling month-on-month across various city tiers[6]. The 2016 Central Economic Work Conference provided a more detailed description of efforts to reduce real estate inventory, emphasizing the need for all regions to adhere to the national development strategy that "houses are for living in, not for speculation." The conference also stressed the importance of researching and accelerating the establishment of a series of foundational systems and long-term mechanisms tailored to local conditions and market dynamics.

This paper, incorporating both temporal and spatial dimensions, offers valuable theoretical insights for understanding and addressing regional development imbalances. The study's conclusions underscore the necessity of adopting more precise regional policies to promote the coordinated development of the national commercial residential property market.

2 DATA SOURCES AND DESCRIPTIONS

The research data was sourced from the National Bureau of Statistics, showcasing data from certain provinces in 2020. The prices of residential commercial properties are influenced by multiple factors, which are complex and include both quantitative and qualitative data. For the convenience of this study, the impact indicators on housing prices are primarily represented by quantifiable variables.

After preliminary data processing and initial field visits, six variables have been selected as explanatory variables for subsequent research(Table 1 and Table 2).

Table 1 Variable Description Table

variable	illustrate
Y	Average sales(Yuan)
X_1	Sales area(ten thousand/m2)
X_2	Investment in real estate development(Yuan)
X_3	Gross Domestic Product (GDP)(100 million yuan)
X_4	Year-end resident population(10,000 people)
X_5	All the people Disposable income per cianapita(Yuan)
X_6	Consumer Price Index

Table 2 Raw Data Tables

region	Y	X_1	X_2	X_3	X_4	X_5	X_6
Beijing	42684	733.59	2317.08	35943.3	2189	69434	101.7
Tianjin	16391	1220.74	2084.8	14008	1387	43854	102
Hebei	8251	5572.25	3746.74	36013.8	7464	27136	102.1
Shanxi	6877	2549.49	1431.8	17835.6	3490	25214	102.9
……	……	……	……	……	……	……	……
Hainan	16751	626.19	946.41	5566.2	1012	27904	102.3
Chongqing	8917	4814.49	3189.05	25041.4	3209	30824	102.3
Sichuan	8041	10902.37	5330.14	48501.6	8371	26522	103.2
Guizhou	5600	4929.93	2572.34	17860.4	3858	21795	102.6
Yunnan	8267	4175.88	3317.55	24555.7	4722	23295	103.6
Tibet	8824	81.62	118.47	1902.7	366	21744	102.2
Shaanxi	9624	3902.4	3225.45	26014.1	3955	26226	102.5
Gansu	6467	1863.81	1010.28	8979.7	2501	20335	102
Qinghai	8164	420.64	292.32	3009.8	593	24037	102.6

Overall, due to regional and economic development factors, the average sales prices of residential commercial housing in Beijing and Guangxi Zhuang Autonomous Region exhibited a substantial disparity around 2010. This gap widened further after 2014, and by now, the difference has more than doubled compared to 2012.

Table 3 Describe the Statistical Table

Average sales price of residential commercial h ousing(Yuan)	Beijing	Guangxi
average value	28483.3	5172.484
maximum	42684	6440
minimum	16553.48	3909.83
range	26130.52	2530.17

Just like Table 3, between 2012 and 2020, the average price of residential commercial housing in Beijing surged from 16,553.48 yuan to 42,684 yuan, an increase of 26,130.52 yuan, which is more than ten times the increase of 2,530 yuan in Guangxi Zhuang Autonomous Region and almost fourteen times the increase in Guizhou Province.

3 ANALYSIS OF THE CURRENT SITUATION OF THE AVERAGE SALES PRICE OF COMMERCI AL HOUSING

In recent years, the overall trend of housing prices in China has exhibited a pattern of being higher in the south than in the north, and higher in the east than in the west. The continuous and significant increase in residential property prices is particularly evident in the developed regions of eastern China, as well as in the central, western, and southern parts of northern China. Even within the same region, the disparity in housing prices is considerable due to uneven economic development. Additionally, policies and the impact of the recent pandemic have contributed to significant fluctuations in housing prices. This paper will conduct a detailed analysis of the average sales prices over the past few years from both horizontal and vertical perspectives.

3.1 Analysis of Regional Differences (Based on Cluster Analysis)

In order to analyze the development of residential commercial housing across provinces, cities, and autonomous regions in mainland China, relevant variables were selected. The initial analysis was conducted using 2020 data(Figure 1).

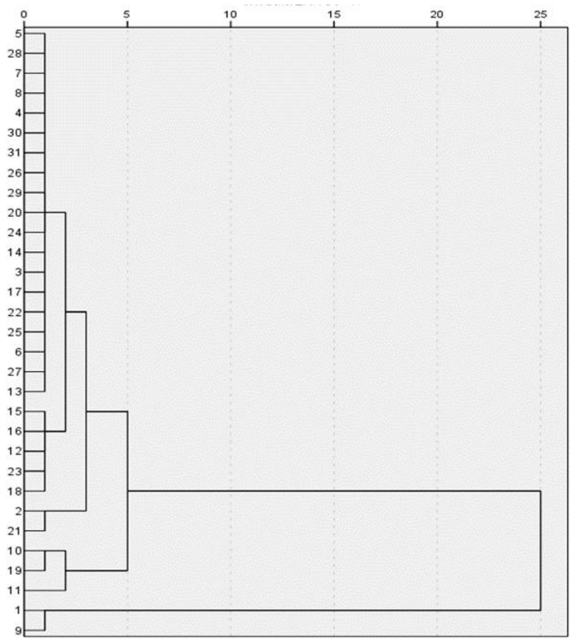

Figure 1 Phylogenetic Clustering Taxonomic Pedigree Diagram(2020)

These regions can be categorized into three groups based on housing conditions, as follows:
First Tier: Beijing, Shanghai.
Second Tier: Zhejiang, Guangdong, Jiangsu.
Third Tier: Tianjin, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Anhui, Fujian, Jiangxi, Hebei, Shandong, Henan, Hunan, Hubei, Guangxi, etc.
Among these, the two cities in the first tier exhibit the most substantial differences from other regions.
The following results of the analysis using the 2002 data(Figure 2):

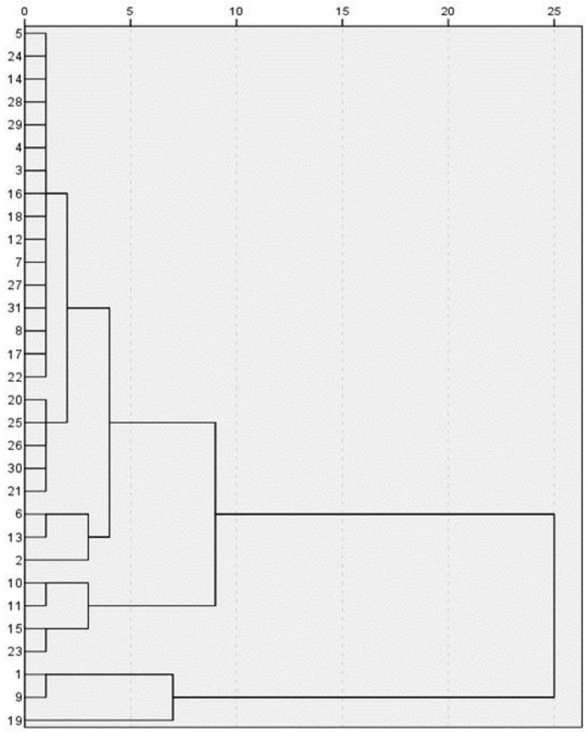

Figure 2 Phylogenetic Clustering Taxonomic Pedigree Diagram(2002)

These regions can be classified into four categories based on housing conditions, as follows:
First Tier: Guangdong Province.
Second Tier: Beijing Municipality, Shanghai Municipality.
Third Tier: Sichuan Province, Shandong Province, Zhejiang Province, Jiangsu Province.

Fourth Tier: Tianjin Municipality, Hebei Province, Shanxi Province, Inner Mongolia Autonomous Region, Liaoning Province, Jilin Province, Heilongjiang Province, Anhui Province, Fujian Province, Jiangxi Province, Henan Province, Hubei Province, Hunan Province, Guangxi Zhuang Autonomous Region, and others.

The housing market conditions were relatively more balanced in 2002. China's vast territory, large population, and uneven resource distribution generally result in high-quality resources being concentrated in major cities, while small and medium-sized cities lag behind. This development disparity has become increasingly pronounced in recent years. The rapid economic development in regions like Beijing and Shanghai has led to a corresponding swift advancement in the housing market, widening the gap from other regions.

4 NEURAL NETWORK MODEL PREDICTION

4.1 Model Building

From the previous discussion, it is evident that constructing an appropriate regression model for a practical problem is very challenging. Apart from building a regression model, one can also view the complex intermediate process of modeling as a black box and use neural network models to predict the data.

First, preliminary data processing is performed. This involves ensuring the relative completeness of the data information and checking for any missing values that may still exist. It is also hoped that the data information can be normalized.

Subsequently, the data is divided into training and testing sets, with 70% of the data allocated for training the neural network model and the remaining 30% reserved for testing it.

Generally, the number of hidden layers lies between the input and output layers, typically amounting to two-thirds of the input layer. Here, a feedforward network model with a single hidden layer containing two neurons is constructed, resulting in the model.

To evaluate the model's fitting performance, the root mean square error (RMSE) is used, which measures the closeness of the model's predictions to the actual labels by calculating the distance between predicted and actual values. The smaller the RMSE, the closer the predictions are to the true values.

First, the RMSE for the training set is calculated, yielding a result of 0.0427. Next, predictions are made on the testing dataset, resulting in the predicted values(Table 4).

Table 4 The Test Set Predicts the Outcome

region	Dependent variable
Tianjin	0.38258
Shanxi	0.07793
Inner Mongolia	0.14331
Heilongjiang	0.07033
Zhejiang	0.43241
Henan	-0.04120
Guangxi	0.00838
Hainan	0.12312
Guizhou	-0.01197
Tibet	0.04445
Xinjiang	0.02924

The correlation coefficient between the predicted and actual values, which indicates the strength of their linear relationship, is 0.832. A correlation coefficient close to 1 suggests a strong linear relationship, implying that the predictions are very close to the actual values and the predictive performance is satisfactory.

4.2 Model Modifications

Based on the original model, we will proceed to modify it by adding an additional hidden layer. We will construct a neural network model with the first and second hidden layers containing four and two hidden units, respectively. The specific model is illustrated below:

The root mean square error (RMSE) for the training set of this model is 0.0386, slightly lower than the previous model's 0.0427, but the difference is minimal. Similarly, the RMSE for the test set is 0.0782, also marginally lower than the previous model's 0.0826. Compared to the previous model, this model does not show significant improvement in predictive performance.

5 CONCLUSION

From the empirical analysis presented earlier, the following conclusions are drawn:

Regarding the development status of residential commercial housing, it is evident that the disparity in development between medium-sized and large cities is becoming increasingly apparent. In economically rapidly developing regions such as Beijing and Shanghai, the development of residential commercial housing is also accelerating, widening the gap with other areas.

There is a viewpoint that post-pandemic, the suspension of business operations and production, coupled with a decline in residents' expected income, will change the consumption patterns of some people, leading to reduced unnecessary expenditures. Consequently, residents' housing consumption and investment behavior will become more cautious and rational, causing a decline in the real estate market's sales level. However, there is also a perspective that once the pandemic stabilizes, people will engage in retaliatory consumption, and the suppression of consumption by the pandemic will be temporary.

COMPETING INTERESTS

The author have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Bai Y. Research on Influencing Factors of Average Sales Price of Residential Commercial Housing in Major Cities of China. IOP Publishing Ltd.IOP Publishing Ltd, 2020:012020.
[2] Na C, Chen-Xu T, Wei L, et al. Analysis and Forecast of the Residential Commercial Housing Price in Zhoukou. Mathematics in Practice and Theory, 2019.
[3] Wanchuan W. Prediction Model and Application of BP Neural Network. Journal of Yangzhou Polytechnic Institute, 2013.
[4] G Peter, Zhang, et al. Neural network forecasting for seasonal and trend time series. Operations Research, 2005.
[5] Marzban C, Stumpf G J. A Neural Network for Damaging Wind Prediction. Weather & Forecasting, 1998, 13(1): 151-163.
[6] Chen Y, Yang B, Dong J. Time-series prediction using a local linear wavelet neural network. Neurocomputing, 2006, 69(4/6): 449-465.