Research on House Price Forecast Based on Grey System GM (1, 1)

Yijia Liu[1], Ke Li[2]

[1]School of Applied Mathematics, Shanxi University of Finance and Economics, Shanxi, China

[2]Shanxi University of Finance & Economics, Shanxi University of Finance and Economics, Shanxi, China

Keywords: Taiyuan City, GM (1, 1) model, house price forecast

Abstract: House prices have always been a critical issue for people's livelihood. Strengthening the scientific forecasting and monitoring of housing price trends is an essential part of the government's implementation of the "effective control of housing prices and stable housing prices." Taking Taiyuan City as an example, this paper establishes a GM (1, 1) prediction model based on the data of the average price of commercial housing from January 2019 to October 2019. Then the error test is performed on the predicted data and the actual data. The experimental results show that the GM (1, 1) model can be used to make short- and medium-term forecasts of house prices. Finally, the model is used to predict the average price of commercial housing in Taiyuan City in the next year and a half.

1. Introduction

The level of housing prices and the rise and fall are closely related to people's livelihood. It is a hot issue of concern to the government and affects the quality of life of every citizen. The Xinhua News Agency recently released a report saying that “in accordance with the requirements for monthly analysis, quarterly evaluation and annual assessment determined by the implementation of the long-term mechanism for the implementation of real estate, the Ministry of Housing and Urban-Rural Development has issued early warning tips for cities with large fluctuations in house prices and land prices in the first quarter of 2019."This warning indicates that the central government has begun to implement real estate high-pressure regulation and control policies. At present, house prices have entered the adjustment channel, speculative demand has deviated, and real estate has gradually returned to residential properties. There is still some room for growth. How to strengthen the scientific forecasting and monitoring of housing price trends is an important part of the government's policy of “effectively controlling housing prices and stabilizing housing prices.”

The housing price issue is an extremely complex and uncertain grey system. The prediction of future housing prices should be based on a combination of factors. At present, the academic community has a variety of methods for housing price forecasting, such as multiple regression linear model, grey theory prediction model, Markov prediction model, genetic algorithm, and neural network. Some scholars at home and abroad have already studied related issues. Dipasquale, Wheaton et al. [1] used macroeconomic variables for the first time to dynamically forecast housing prices. J Luttik et al. [2] used the Netherlands as an example. The construction of the hedonic model shows that the surrounding environment is an essential factor affecting housing prices. Gerlach.Peng et al. [3] concluded that there is a long-term stable relationship between real estate prices and macroeconomics. in conclusion. Crawford, Fratantoni et al. [4] used the ARIMA transformation matrix to predict US housing prices. Liu Dajiang et al. [5] regard the dynamic process of housing prices as the Markov chain. According to the monthly statistical data of the sales price of commercial housing, the time series are smoothed to carry out housing price forecasting and research. Hu Liuxing, Wu Jiefei et al. [6] gray system theory as the theoretical basis, respectively constructed the GM (1, 1) model and the linear regression model integrated into the gray theory to predict housing prices [7]. He Li et al. conducted a medium- and long-term forecast of housing prices in Beijing based on the analysis of the ARIMA model.

  DOI: 10.25236/icfil.2019.036

As the provincial capital of Shanxi Province, Taiyuan City enjoys an advantageous geographical location. There is a Taihang Mountain barrier in the east and a Luliang Mountain as a barrier in the west. It is located on the valley plain between the two mountains. The second-largest tributary of the Yellow River, the Weihe River, traverses the entire territory of Taiyuan City from north to south. It is a combination of things in the location. With the large-scale construction of the external transportation network, the Taiyuan Economic Circle plays an essential role in linking the eastern and western regions of China and promoting the coordinated development of the regional economy. Based on the actual data of the average residential price of Taiyuan in Taiyuan City from September 2017 to April 2019 (Table 1), this paper establishes the GM (1, 1) model to predict and analyze the price of new residential buildings in Taiyuan.

2. Research methods and data sources

GM (1, 1) is a prediction model proposed by Professor Deng Julong [8] of Huazhong University of Science and Technology in the 1980s, which is a method for predicting systems with uncertain factors. Grey prediction can generate the data sequence with healthy regularity by identifying the degree of difference between the development factors of system factors, that is, performing correlation analysis, and generating raw data to find the law of system change. Then establish the corresponding differential equation model to predict the future development of things. GM (1, 1) constructs a gray prediction model by using a series of quantitative values of the characteristics of the reaction prediction object observed by the isochronous distance and predicts the feature quantity at a particular moment in the future, or the time to reach a certain feature quantity. At present, the gray prediction model GM (1, 1) model is widely used in economic analysis because of its low information and high modeling accuracy.

Table 1. Average price of new housing (2018 .01-2019.10)

Month	2018.01	2018.02	2018.03	2018.04	2018.05	2018.06
Sample average price	9971	10105	10378	10869	11294	11594
Month on month	2.85%	1.34%	0.05%	4.74%	3.91%	2.66%
Month	2018.07	2018.08	2018. 9	2018.10	2018. 11	2018.12
Sample average price	11867	12024	12084	12132	12046	11979
Month on month	2.36%	1.32%	0.5%	0.4%	0.7%	0.56%
Month	2019.01	2019.02	2019.03	2019.04	2019.05	2019.06
Sample average price	11990	12006	12030	12003	12025	12084
Month on month	0.09%	0.14%	0.21%	0.22%	0.18%	0.49%
Month	2019.07	2019.08	2019.09			
Sample average price	12113	12095	12051			
Month on month	0.24%	0.14%	0.36%			

Source: China Index Research Institute - China Real Estate Index System 100 city price index report

3. Establishment and empirical analysis of GM (1, 1) prediction model

The Empirical GM (1, 1) model predicts the principle of generating a set of new data sequences

with obvious trends for a certain data sequence, building a model according to the growing trend of the new data series, and then using the tired The subtraction method performs the inverse calculation, restores the original data sequence, and then obtains the prediction result.

3.1. Data examination and processing set up the original array with 20 observations to the Table 1.

$$X^{(0)} = \left(x^{(0)}(1),\ x^{(0)}(2), x^{(0)}(3), \ldots,\ x^{(0)}(20)\right) =$$

$$\begin{pmatrix} 9971, 10105, 10378, 10869, 11294, 11594, 11867, 12024, \\ 12084, 12132, 12046, 11979, 11990, 12006, 12030, 12003, \\ 12025, 12084, 12113, 12095, 12051 \end{pmatrix}$$

(1) Calculate the class ratio of sequence:

$$\lambda(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)} \quad k = 1, 2, 3, \ldots, 21$$

$$\lambda = \left(\lambda(2), \lambda(3), \ldots, \lambda(20)\right)$$

$$= (0.9867, 0.9737, 0.9548, 0.9624, 0.9741, 0.9770, 0.9869, 0.9950, 0.9960, 1.0071,$$

$$1.0055, 0.9991, 0.9987, 0.9980, 1.0022, 0.9982, 0.9951, 0.9976, 1.0015, 1.0036)$$

(2) Judge the class of ration:

All the $\lambda(k)$ in the $[0.9771, 1.0071]$, $k = 2, 3, \ldots, 21$. Therefore, the data sequence $X^{(0)}$ can be established by GM(1,1) model and be predicted.

3.2. Construction GM (1, 1) Model:

The steps for building this model are as follows:

(1) The first step is to add new series in order to weaken the volatility and randomness of the random sequence, and results are obtained as follows:

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), x^{(1)}(3), \ldots, x^{(1)}(21)\right)$$

$$= \begin{pmatrix} 9971, 20076, 30454, 41323, 52617, 64211, 76078, 88102, \\ 100186, 112318, 124364, 136343, 148333, 160339, 172369 \\ , 184372, 196397, 208481, 220594, 232689, 244740 \end{pmatrix}$$

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), \quad k = 2, 3, \ldots, 21$$

(2) The second step is to generate adjacent mean value sequence.

$$Z^{(1)} = \left(z^{(1)}(1), z^{(1)}(2), z^{(1)}(3), \ldots, z^{(1)}(21)\right)$$

$$= (15023.5\ , 25265, 35888.5, 46970, 58414, 70144.5, 82090, 94144,$$

$$106252, 118341, 130353.5, 142338, 154336, 166354,$$

$$178370.5, 190384.5, 202439, 214537.5, 226641.5, 238714.5)$$

$$z^{(1)}(k) = \frac{1}{2}\left(z^{(1)}(k-1)\right) + \frac{1}{2}\left(z^{(1)}(k)\right), \quad k = 2, 3, \ldots, 21$$

(3) The third step, using the least-squares method to fit the parameters a and u:

Establish difference equation:

$$X^{(0)}(t) + aX^{(1)}(k) = u$$

According to the Grey Theory, the first-order ordinary differential equation with one variable of t is established:

$$\frac{dx^{(1)}}{dt} + aZ^{(1)}(k) = u$$

$$\hat{a}=\begin{pmatrix} a \\ u \end{pmatrix}$$

where a is evolution parameter, and u is the grey parameter.

Then get $X^{(1)}(t)$ and the predictive value of $X^{(0)}$.

(4) The fourth step is to use $X^{(1)}(t)$ and $X^{(0)}$ to establish matrices B and matrices Y_{20}:

$$B=\begin{bmatrix} -z^{(1)}(2) & 1 \\ \vdots & \vdots \\ -z^{(1)}(20) & 1 \end{bmatrix}$$

(5) The fifth Solve the parameters of the gray GM (1, 1) model by the least square method (LSM)

$$\hat{a}=\left(B^T B\right)^{-1} B^T Y_n$$

$$\hat{a}=-0.0063$$

(6) Then plug the gray parameter $\hat{a}$ into $\frac{dx^{(1)}}{dt}+aZ^{(1)}(k)=u$ and solve it.

u=10957

$$\hat{X}^{(1)}(t+1)=\left(x^{(1)}(1)-\frac{u}{a}\right)e^{-at}+\frac{u}{a}$$

(7) In the last step, the prediction equation is:

$$\hat{X}^{(0)}(t+1)=\hat{X}^{(1)}(t+1)-\hat{X}^{(1)}(t)\ t=1,2,3,...,20$$

$$\hat{X}^{(1)}=\hat{X}^{(0)}(1)=x^{(0)}(1)=7720$$

$$\hat{X}^{(0)}(0)=\left(\hat{x}^{(0)}(1),\hat{x}^{(0)}(2),\hat{x}^{(0)}(3),...,\hat{x}^{(0)}(20)\right)$$

The predicted values are shown in the following table

Table 2. The forecast simulation of the average price of commodity houses

month	2018.01	2018.02	2018.03	2018.04	2018.05	2018.06
Sample average price	9971	10105	10378	10869	11294	11594
Predictive average value	9971	11054	11123	11193	11263	11334
month	2018.07	2018.08	2018. 9	2018.10	2018. 11	2018.12
Sample average price	11867	12024	12084	12132	12046	11979
Predictive average value	11405	11477	11549	11621	11694	11767
month	2019.01	2019.02	2019.03	2019.04	2019.05	2019.06
Sample average price	11990	12006	12030	12003	12025	12084
Predictive average value	11841	11916	11990	12066	12141	12218
month	2019.07	2019.08	2019.09			
Sample average price	12113	12095	12051			
Predictive average value	12294	12372	12449			

3.3 Accuracy test

In order to test the credibility of the prediction results, the model needs to be tested for accuracy. The test method is as follows:

(1) Residual test:

Absolute residual sequence:

$$\varepsilon(k) = x^{(1)}(k) - \hat{X}^{(0)}(k) \quad k = 2,3,...,21$$

Relative residual sequence:

$$\Delta(k) = \frac{E^{(0)}(k)}{x^{(0)}(k)} \quad k = 2,3,...,21$$

Average relative residua:

$$\overline{\Delta(k)} = \frac{1}{21}\sum_{k=1}^{21}\left|\Delta(k)\right|$$

$$\overline{\Delta(k)} = 0.031$$

(2) Correlation test

Relational analysis sequence:

$$r_k = \frac{\min\left(\left|\varepsilon_k\right|\right) + \rho\max\left(\left|\varepsilon_k\right|\right)}{\left|\varepsilon_k\right| + \rho\max\left(\left|\varepsilon_k\right|\right)} \quad k = 2,3,\cdots\cdots 21$$

Correlation:

$$R = \frac{1}{21}\sum_{k=1}^{21} r_k$$

ρ : Identification coefficient, often $\rho = 0.5$

$R = 0.7880$

(3) Post-test difference test:

GM (1, 1) is usually tested by the Post-test difference.

Mean:

$$\overline{X} = \frac{1}{n}\sum_{k=1}^{n} x^{(0)}(k)$$

Variance:

$$S_1 = \sqrt{\frac{1}{n}\sum_{k=1}^{n}\left[x^{(0)}(k) - \overline{X}\right]^2}$$

Mean of Post-test difference:

$$\overline{E} = \frac{1}{\text{n-1}}\sum_{k=2}^{n} E(k)$$

Variance of Post-test difference:

$$S_2 = \sqrt{\frac{1}{n-1}\sum_{k=2}^{n}\left[E(k) - \overline{E}\right]^2}$$

calculating the ratio:

$$C = \frac{S_2}{S_1}$$

$C = 0.2097$

(4) Small error probability

$p = p\left\{\left|E(k)\right| - \overline{E} \le 0.6745S_1\right\}$, $p = 1$

Table 3. Prediction accuracy level checklist

Budget accuracy level	Residual test. ($\overline{\Delta(k)}$)	Correlation test. (R)	Post-test difference test(C)	Small error probability(p)
Good (level 1)	0.01	0.9	0.35	0.95
Qualified (secondary)	0.05	0.8	0.5	0.80
Reluctant (third level)	0.1	0.7	0.65	0.70
Failed (Level 4)	0.2	0.6	0.80	0.60

Table 4. Prediction accuracy

Budget accuracy	Residual test. ($\overline{\Delta(k)}$)	Correlation test. (R)	Post-test difference test(C)	Small error probability(p)
	0.031	0. 7880	0.2097	1

In summary, the average relative error accuracy level of the model as a whole is two; the correlation degree is about two; the mean square error precision ratio is about one level; the small error probability accuracy level is equal to one level, so the overall operation result of the model is credible.

4. Forecast of the average price of commercial housing in Taiyuan City

From the error test of the prediction model, it is known that the GM (1, 1) prediction model based on the data of the average price of commercial housing in Taiyuan City from January 2019 to October 2019 is highly accurate and suitable for the average price of commercial housing in the future. Long-term forecast. The specific results are shown in Table 5.

Table 5. The Predictive average value of Taiyuan City (2019.10—2020 .06)

Month	2019.10	2019.11	2019.12	2020.01	2020.02	2020.03
Predictive average value	12527	12606	12685	12765	12845	12925
Month	2020.04	2020.05	2020.06	2020.07	2020.08	2020.09
Predictive average value	13007	13088	13170	13253	13336	13420
Month	2020.10	2020.11	2020.12	2021.01	2021.02	2021.03
Predictive average value	13504	13589	13674	13760	13847	13933
Month	2021.04	2021.05	2021.06			
Predictive average value	14021	14109	14197			

The forecast results show that: In the period from 2019.10 to 2020.06, the housing prices in Taiyuan City generally maintained a relatively stable situation with a slight increase. As long as the market expectation has not been fundamentally reversed, the price of Taiyuan City will fluctuate within a reasonable range.

5. Conclusion

The issue of housing prices has always been one of the hottest issues of concern to the government and the community. Accurate, timely and correct delivery of housing prices and changes in technology is the benchmark for conducting relevant analysis, research and formulation

of specific management and control policies.

Based on the historical data of housing prices in Taiyuan City, this paper constructs the GM (1, 1) model of housing price fluctuations and conducts medium- and long-term statistical forecasts of housing price changes in Taiyuan.

Under the strict control of the property market, the first-tier cities have the role of the property market vane, so the impact of the regulatory policies is the most obvious. As a second-tier city, Taiyuan is affected by factors such as population and demand. Under the control policy of local conditions, the hot market differentiation is more prominent. Besides, for the third- and fourth-tier cities, in the past two years, under the influence of speculative factors. The purchasing power and housing price rise in the third- and fourth-tier cities have been seriously overdrawn, and the adjustment of future housing prices may face more significant pressure.

What can be determined now is that the price control will not be relaxed for a while, and finally, it is necessary to wait until the establishment of a long-term real estate mechanism to achieve transitional linkage and guarantee the long-term stability of the real estate market. Before this, Xinhua News Agency's economic reference newspaper stated that real estate should bid farewell to the periodicity, and the property market regulation will not relax for five years. Therefore, under the tone of the continued tightening of the property market regulation, the control policy will further highlight the housing price pull in 2019, so the real estate industry in 2019 may not change much, and the overall adjustment of the property market is still in progress.

References

[1] Denise D P , Wheaton W C . Housing Market Dynamics and the Future of Housing Prices[J]. Journal of Urban Economics, 1994, 35(1):1-27.

[2] Luttik J . The Value of Trees, Water and Open Space as Reflected by House Prices in the Netherlands[J]. Landscape and Urban Planning, 2000, 48(3-4):161-167.

[3] Gerlach S , Peng W . Bank lending and property prices in Hong Kong[J]. Journal of Banking & Finance, 2005, 29(2):0-481.

[4] Crawford G W , Fratantoni M C . Assessing the Forecasting Performance of Regime-Switching, ARIMA and GARCH Models of House Prices[J]. 2003, 31(2):223-243.

[5] Zhang L, Fu S, Li B. Research on Stock Price Forecast Based on News Sentiment Analysis—A Case Study of Alibaba[C]// International Conference on Computational Science. 2018.

[6] Dong F, Qi B, Jie Y. Comparative static analysis of provincial agricultural science and technology level based on grey clustering[J]. Grey Systems: Theory and Application, 2018, 8(4):481-493.

[7] Peng A L, Dong K, Yang W N. Research on Prediction of the Amount of Regional Airport Aviation Based on GM (1.1) Model in the Gray System[J]. Applied Mechanics & Materials, 2014, 587-589(1):1901-1906.

[8] Zeng B, Meng W, Liu S F, et al. Research on prediction model of oscillatory sequence based on GM (1, 1) and its application in electricity demand prediction[J]. Journal of Grey System, 2013, 25(4):31-40.

[9] Zhang Yanjun, Yang Xiaodong, Liu Yi, Zheng Dayuan, Bi Shujun. Research on the Frame of Intelligent Inspection Platform Based on Spatio-temporal Data. Computer & Digital Engineering [J], 2019, 47(03): 616-619+637.

[10] Yi Liu, Jiawen Peng, and Zhihao Yu. 2018. Big Data Platform Architecture under The Background of Financial Technology: In The Insurance Industry As An Example. In Proceedings of the 2018 International Conference on Big Data Engineering and Technology (BDET 2018). ACM, New York, NY, USA, 31-35.