# Forecast of Housing Prices in Beijing and Analysis of Influencing Factors

Yuxi Liu*

School of Statistics, Shanxi University of Finance and Economics, Taiyuan, 030000, China

*Corresponding author: 202102030125@stu.sxufe.edu.cn

**Abstract.** Real estate is an important part of China's real economy, and housing prices are one of the basic elements in the entire real estate market. This paper analyzes the housing price data of Beijing, China, and firstly uses a multiple linear regression model to determine whether the factors that have an important impact on China's housing prices also play an effect on Beijing's housing prices. And then this paper analyzes how much each factor has an impact on Beijing's housing prices. Additionally, the ARIMA model is used to predict the trend of housing prices in Beijing in the next 10 months according to the monthly data of the past 10 years. Finally, the change point analysis method is used to study the effect of the Corona Virus Disease 2019 (COVID-19) on housing prices in Beijing. It aims to understand the operating mechanism behind housing prices and give policymakers, investors and home buyers more suggestions and instruction.

**Keywords:** Housing prices; Beijing; linear regression; ARIMA model.

## 1. Introduction

The real industry is a pillar of Chinese economy, making a significant contribution to Gross Domestic Product (GDP) and providing important support for China's economic growth. As the economic and political center of China, Beijing's real estate industry has an important impact on China's real estate market. Therefore, it is of great significance to research Beijing's real estate industry.

At present, China's real estate enterprises are in a fierce competitive environment. It is difficult for real estate companies to obtain loans from banks, and the funding cost of the capital market has increased, resulting in increased financial pressure on real estate enterprises, and some enterprises are facing the risk of capital chain breakage. Many enterprises are facing the difficult situation of falling into a liquidity crisis [1]. Liu believed that the current demand for real estate is declining rapidly, and the decline in demand exceeds the decline in supply, and boosting housing demand is an important issue [2]. Shen believed that there are real estate bubbles in some large cities in China, and the market has a great hidden crisis on a whole [3]. In order to solve many problems in the real estate market, it is important to understand the mechanism of housing price formation and the laws of the operation of the real estate market. It is essential to research the influencing factors of housing prices and what extent various factors affect housing prices.

The movement of housing prices is caused by a number of factors. Based on the Vector Autoregression (VAR) model, Xu and Ye concluded that housing prices are affected by both GDP and China Commodity Price Index (CCPI) [4]. Yang and Deng mentioned that the relationship between supply and demand is reflected in the real estate industry, and they conclude that population is one of the important factors affecting housing prices through regression analysis and post-hoc simulation [5]. Ouyang and Lyu concluded that one of the main factors is the amount of investment in real estate development [6].

In this paper, a linear regression model is used to perform regression analysis on the above factors, aiming to determine whether the effects of each factor on housing prices in Beijing is significant and the degree to which each factor affects housing prices in Beijing. To provide guidance to decision-makers such as governments, property developers, investors, and others to help them develop smarter real estate policies and investment strategies.

By analyzing historical data and influencing factors, we can build models to predict possible changes in housing prices. Wang established three prediction models to compare and research the price prediction of commercial housing in Shandong Province, and believed that the multiple linear regression model is more suitable [7]. Liu et al. predicted housing prices based on graph neural networks and long short-term memory models [8]. Jiang and Liu optimized the Back propagation (BP) neural network method based on Pearson coefficient and a firefly algorithm to predict housing prices [9].

The pandemic is a non-negligible factor that has had a huge impact on China's real estate market. In order to respond to the needs of epidemic prevention and control, phased static management has been widely adopted in some regions. Most of the projects in the real estate industry are in a state of stagnation or delay on a regular and intermittent basis. In the short term, it will lead to a decline in housing prices, and in the long term, it will lead to the relaxation of financial policy and the lowering of the threshold for buying a house to improve the vitality of the real estate market [10].

This sudden impact on property prices is subject to the change point analysis method. This paper uses the Autoregressive Integrated Moving Average (ARIMA) model, and the change point analysis method to predict the future housing price trend in Beijing. To help homebuyers make more informed home purchase decisions based on the predictions, sellers can price better, and investors can make portfolio adjustments based on the predictions.

## 2. Methods

### 2.1. Data Source

The house price data is collected from the 58.com website, and the monthly average price data on new homes in Beijing from January 2013 to October 2023 is selected, including 130 observations.

GRP data, CPI data, population data, and residential investment data are all selected from 2013 to 2022, each containing 10 observations. Among them, GPR data and population data are from the Beijing Statistical Yearbook 2022 and The People's Government of Beijing Municipality website. The CPI data comes from the Beijing Statistical Yearbook 2022, Beijing Statistical Yearbook 2018, and The People's Government of Beijing Municipality website. Residential investment data comes from the Beijing Municipal Bureau of Statistics (2013-2017) and the Huaon.com website.

### 2.2. Variable Selection

According to the researches mentioned by many scholars above, GDP, CCPI, population, and total investment in real estate development are the main factors affecting housing prices. Since the research object of this paper is a smaller range of housing prices in Beijing, other variables are also selected here as more relevant variables. The GDP is replaced by the annual GDP of Beijing, and the CCPI is replaced by the CPI of Beijing. The permanent population of Beijing and the amount of residential investment in Beijing are also selected as variables.

In the process of housing price forecasting, the average price of new houses in Beijing was selected as a variable, and the monthly data from January 2013 to October 2023 were used for forecasting.

### 2.3. Model Selection

Linear regression is a data analysis technique that uses other relevant known data values to predict the value of unknown data, and if there are two or more independent variables, it is called multiple regression. In order to explore the influence of various factors on housing prices, multiple regression method was used to analyze the influencing factors.

Since most of the data are annual data and the frequency is small, it is difficult to predict, so this paper uses the ARIMA model to predict housing prices.

# 3. Results and Discussion

## 3.1. Multiple Linear Regression

Before the regression analysis, the correlation analysis of every independent variable GRP (GDP of Beijing), POPU (permanent population of Beijing), CPI (Beijing consumer Price Index), RI (real estate investment of Beijing) and HP (Beijing housing prices) was carried out to preliminarily determine the correlation degree between each variable and the housing prices. The correlation coefficients between HP and each variable are shown in Table 1.

**Table 1**. Correlation coefficients between HP and each variable

| | GRP | POPU | CPI | RI |
|---|---|---|---|---|
| HP | 0.890 | 0.589 | 0.889 | 0.937 |

From Table 1, the correlation coefficients are all more than 0.5, so the correlation between Beijing housing prices HP and GRP, POPU, CPI and RI is significant. It can be preliminarily determined that these four factors are closely related to Beijing housing prices. Then make a scatter plot and observe the linear relationship, as Fig. 1.

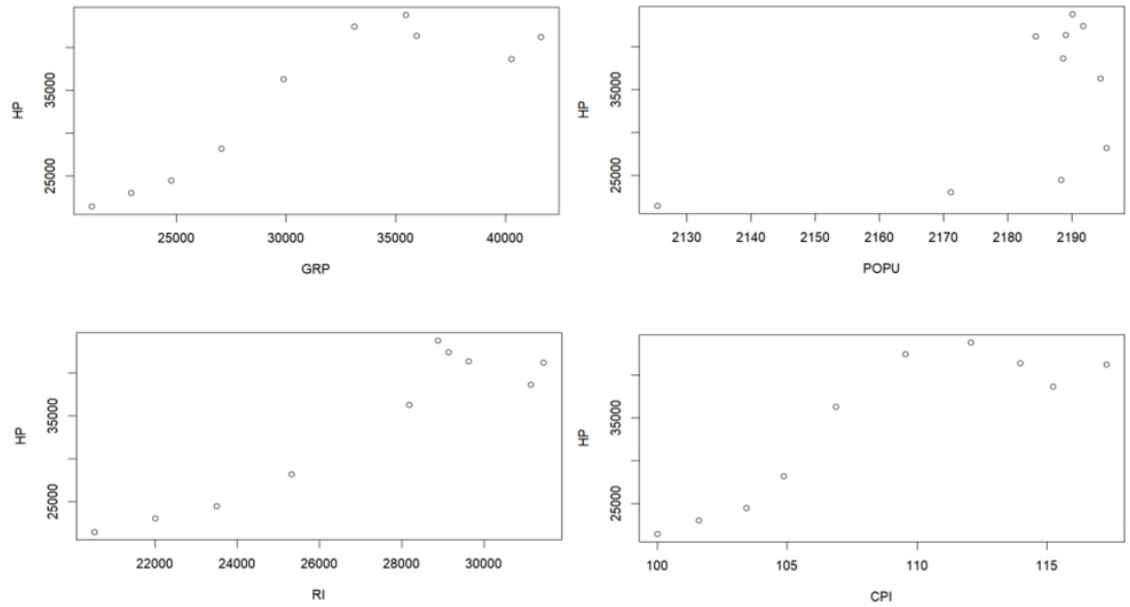


**Fig. 1** Scatter plot

It was found that there is a nonlinear relationship between POPU and HP, therefore the POPU variable was excluded, and other factors have a good linear relationship with HP. The multiple linear model was further established to study and analyze the housing prices in Beijing.

In order to further explore the functional relationship between several variables, a multiple linear regression model was established as follows and the result is shown in Table 2 below.

$$HP = \beta_0 + \beta_1 * GRP + \beta_2 * CPI + \beta_3 * RI + \varepsilon \quad (1)$$

**Table 2**. Results of regression 1

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | -1.32E+05 | 1.60E+05 | -0.825 | 0.4409 |
| GRP | -1.43E+00 | 1.83E+00 | -0.782 | 0.4641 |
| CPI | 1.22E+03 | 1.90E+03 | 0.643 | 0.5439 |
| RI | 2.89E-04 | 1.26E-04 | 2.298 | 0.0612 |
| Multiple R-squared | 0.8901 | Adjusted R-squared | | 0.8352 |
| F-statistic | 16.2 | p-value | | 0.002778 |

As can be seen from Table 2, the P values of each parameter are greater than 0.05, which means that these parameters are all of zero significance, and none of them can pass the significance test, and the results are unreliable. But the adjusted R2 = 0.8352 indicates that the model can explain 83.52% of the housing prices. The comprehensive analysis concludes that the model may have multicollinearity. Therefore, the multicollinearity test of the model was performed in Table 3.

**Table 3**. Multicollinearity test

| | GRP | POPU | CPI | RIA |
|---|---|---|---|---|
| GRP | 1.000 | 0.529 | 0.994 | 0.968 |
| POPU | 0.529 | 1.000 | 0.525 | 0.661 |
| CPI | 0.994 | 0.525 | 1.000 | 0.958 |
| RIA | 0.968 | 0.661 | 0.958 | 1.000 |

From Table 3, the correlation coefficients between the explanatory variables are high, especially the correlation coefficient between GRP and CPI is as high as 0.9943372, which confirms the existence of multicollinearity between the explanatory variables.

In this paper, a forward stepwise regression method is used to reduce the collinearity. Stepwise regression results are presented in Table 4.

**Table 4**. Result of regression 2

| | Estimate | Std. Error | t value | Pr(>\|t\|) |
|---|---|---|---|---|
| (Intercept) | -2.37E+04 | 7.69E+03 | -3.085 | 0.015 |
| RI | 2.14E-04 | 2.83E-05 | 7.586 | 6.39E-05 |
| Multiple R-squared | 0.8779 | Adjusted R-squared | | 0.8627 |
| F-statistic | 57.55 | p-value | | 6.39E-05 |

From Table 4, the adjusted R2 of HP and RI is as high as 0.8627, and the regression parameter passed to significance test, and the P value corresponding is less than 0.01, and the overall significance test of the model is passed, which has good practical economic significance.

### 3.2. ARIMA Model

The forecast of housing prices in Beijing needs to be supported by a large amount of historical data, and the data available in the market are the monthly data of the past ten years, which is presented in the following data chart, as shown in Fig. 2.

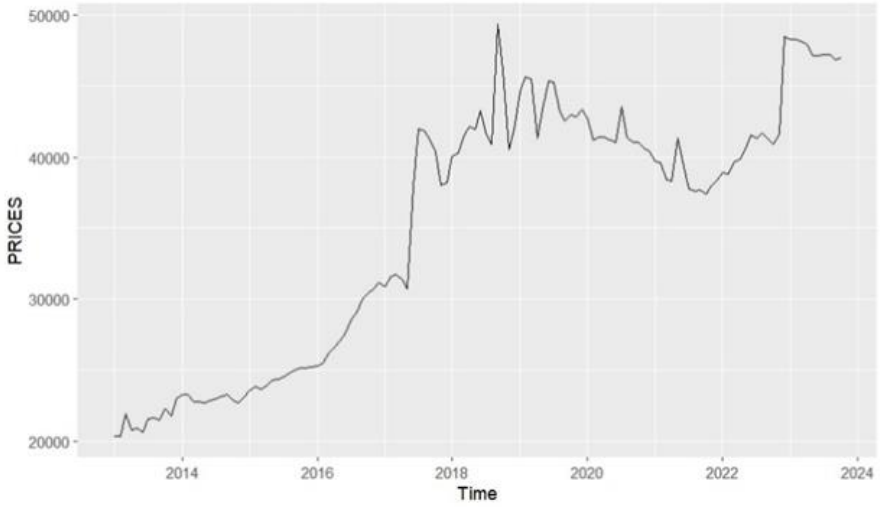


**Fig. 2** Housing Prices in Beijing

Housing prices have been on the rise in the past decade as a whole, which can be divided into a number of components with different characteristics. Housing prices began to rise slowly from 2013

until June 2017. It rose sharply from around 30,000 yuan per square meter to about 38,000 yuan per square meter over the next two months, after which prices showed slow and seasonal growth again. Housing prices continued to decline from 2020 until they rose again in 2022. This data has a clear trend and some data are seasonal, with no obvious cyclicity. Therefore, the autocorrelation detection of this data is conducted, and the ACF plot is shown in Fig. 3.

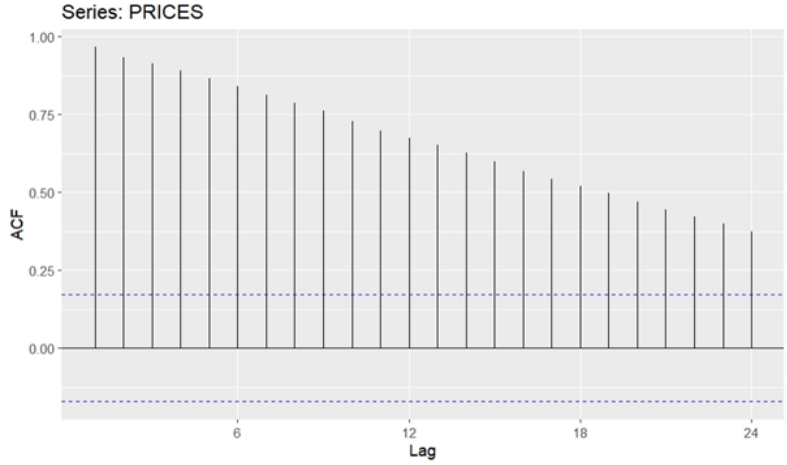


**Fig. 3** ACF plot of housing prices

There is a clear autocorrelation in this data. In this paper, the method of first differences is used to eliminate the trend and seasonality of the data, making it stationary. ACF plots and PACF plots are used to observe whether the correlation between different time series values has been removed, as shown in Fig. 4 and Fig. 5.

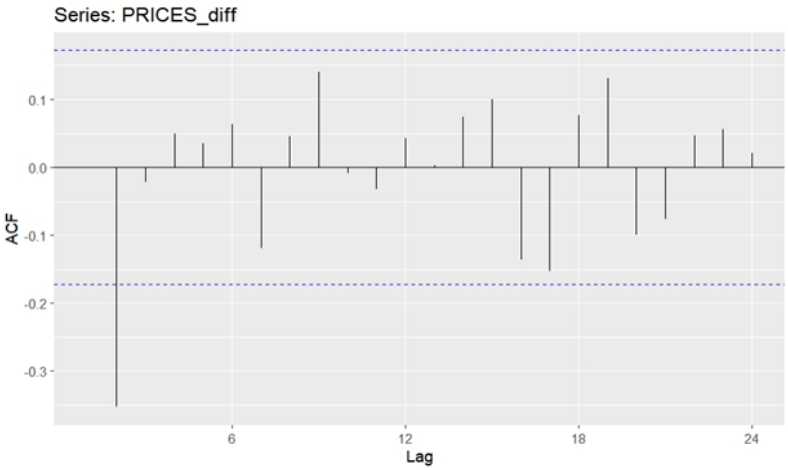


**Fig. 4** The ACF plot of the first-order differenced housing prices

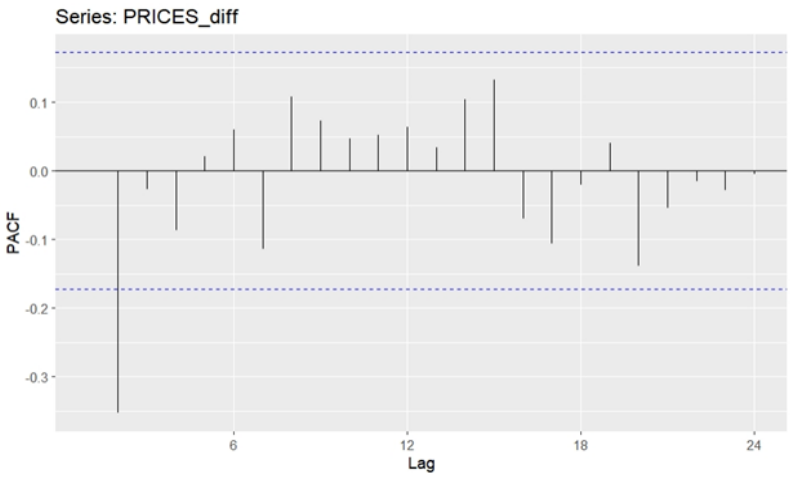


**Fig. 5** The ACF plot of the first-order differenced housing prices

Except for the first line that exceeded the confidence interval, the other periods were all within the blue dotted line, and the data autocorrelation was considered to have been removed, and the ARIMA model could be used to predict the data for the next 10 months, as shown in Fig. 6 and Table 5.

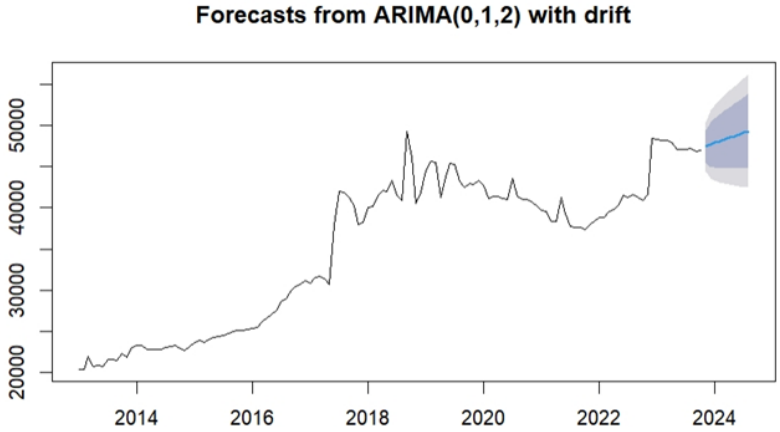


**Fig. 6** Forecasting plot with ARIMA (0,1,2)

**Table 5**. Predicting data with ARIMA (0,1,2)

| | Point Forecast | Lo 80 | Hi 80 | Lo 95 | Hi 95 |
|---|---|---|---|---|---|
| Nov-23 | 47445.39 | 45496.36 | 49394.42 | 44464.61 | 50426.17 |
| Dec-23 | 47712.32 | 44977.56 | 50447.09 | 43529.86 | 51894.78 |
| Jan-24 | 47919.65 | 44903.59 | 50935.7 | 43306.99 | 52532.31 |
| Feb-24 | 48126.97 | 44853.71 | 51400.24 | 43120.95 | 53133 |
| Mar-24 | 48334.3 | 44822.61 | 51845.98 | 42963.64 | 53704.96 |
| Apr-24 | 48541.62 | 44806.7 | 52276.54 | 42829.56 | 54253.68 |
| May-24 | 48748.95 | 44803.41 | 52694.48 | 42714.77 | 54783.13 |
| Jun-24 | 48956.27 | 44810.8 | 53101.74 | 42616.32 | 55296.22 |
| Jul-24 | 49163.59 | 44827.4 | 53499.79 | 42531.95 | 55795.24 |
| Aug-24 | 49370.92 | 44852.04 | 53889.8 | 42459.89 | 56281.95 |

According to the forecast, the trend of house prices in the next 10 months will be upward, and the rate of increase will be relatively slow. In order to test the degree of fitting of the model, the LB test is used to determine whether the residual sequence is white noise, and the results are shown in Fig. 7 below.

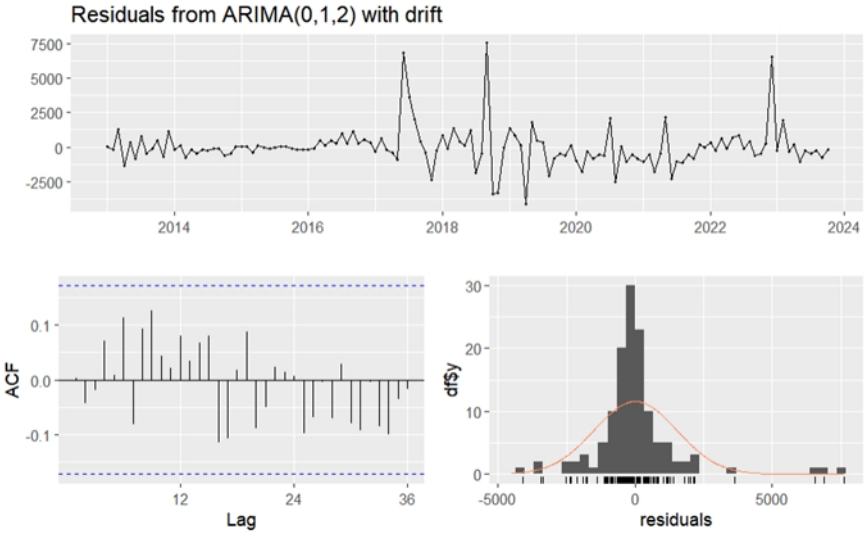


**Fig. 7** Residual testing

As shown in the ACF plot in Fig. 7, all the lines are within the blue dashed line, and the residuals are considered as white noise, and the p-value = 0.7705 which is greater than 0.05 in the Ljung-Box test, so the null hypothesis is true. It is considered that the ARIMA model is appropriate here.

At the same time, we cannot ignore the impact of the COVID-19 on housing prices. In order to explore specific impacts, this paper uses the Kolmogorov-Smirnov method to conduct a change point analysis to determine the specific change points of the data affected by the epidemic, as shown in Fig. 8 below.

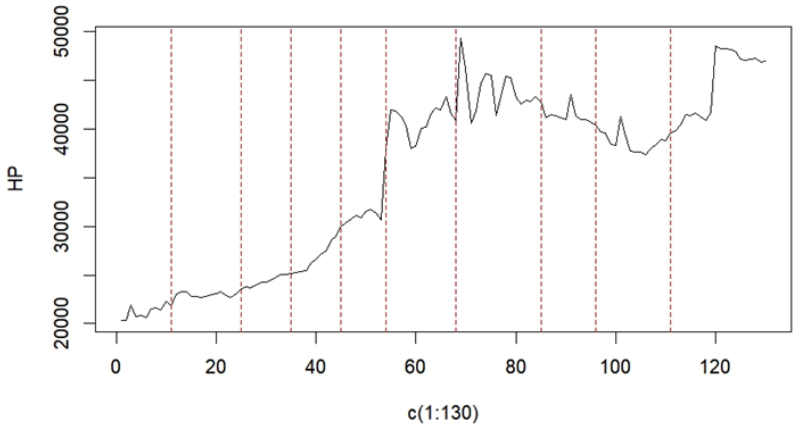


**Fig. 8** Change point plot

There are 9 change points in this data, which are the 11th, 25th, 35th, 45th, 54th, 68th, 85th, 96th, and 111th time points, of which the 85th point is January 2020, the 96th point is December 2020, and the 111th point is March 2022. These three points are the same as the time node of the epidemic.

It is speculated that the 85th time point is the point of change caused by COVID-19. In this paper, the data before this point from 2013 to 2019 are used for forecasting and comparison with the real value. The prediction results are shown in Fig. 9.

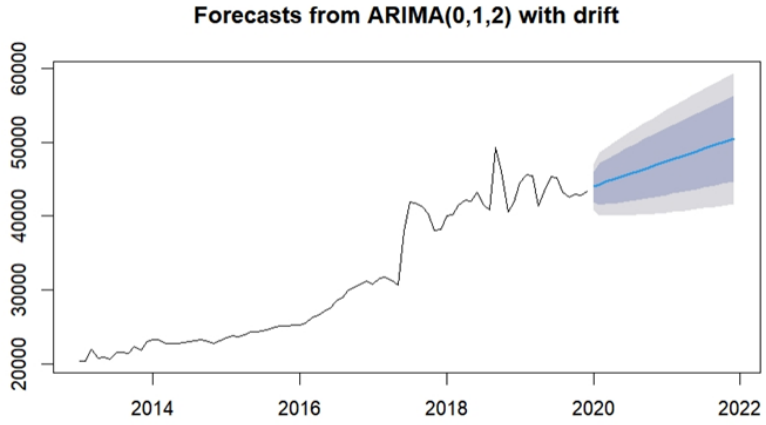


**Fig. 9** Predicting plot based on pre-pandemic data

Forecasts based on pre-pandemic data indicate that house prices will trend upwards over the next two years. But as shown in Fig. 2, the real data from 2020 to 2022 indicate that the house price data has a downward trend during this period. The forecast is completely contrary to the facts, and it can be concluded that the epidemic has suppressed the growth of housing prices and has a negative effect on the entire real estate market.

## 4. Conclusion

Through this research, the results show that among the many factors affecting housing prices, real estate development investment is an important factor affecting housing prices in Beijing, and there is a positive correlation between the two. Additionally, housing prices in Beijing will continue to rise slowly over the next 10 months, fluctuating between 4,300 and 5,700 yuan per square meter. COVID-

19 has had a huge effect on Beijing's housing prices, turning them from an upward trend to a downward trend, which has had a negative effect on the real estate market.

Based on the results of the analysis and predictions, some recommendations can be obtained. It is recommended that investors understand the government's policy adjustments on real estate development investment and adjust their investment strategies in a timely manner according to policy changes. For home buyers, it is important to carefully assess the current real estate market risks and benefits to ensure that the home purchase decision is in line with personal financial planning and risk tolerance.

## References

[1] Yang Xuying. Analysis of the problems and solutions in financing management of real estate enterprises. Business News, 2023, (20): 98-101.

[2] Liu Lin. The new situation of significant changes in the supply and demand relationship in the real estate market. China Investment (Chinese and English), 2023, (Z8): 10-13.

[3] Shen Zhuotao. Foam economy: econometric analysis of China's real estate market. Shanxi Agricultural Economics, 2021, (03): 5-11.

[4] Xu Jin, Ye Ziqing. Analysis of factors influencing the price of commercial housing based on VAR model. Statistics and Decision Making, 2017, (11): 93-97.

[5] Yang Guizhong, Deng Xuefen. Regression analysis and after- simulation of house price in Chengdu city. Value Engineering, 2007, (04): 46-50.

[6] Ouyang Xinyu, Lyv Jin. Linear regression model analysis of factors influencing real estate prices in Nanning city. Residential and Real Estate, 2023, (27): 110-112.

[7] Wang Zhaojuan. Research on the price prediction of commercial housing in Shandong province. Cooperative Economy and Technology, 2023, (17): 63-65.

[8] Liu Xin, Du Hongli, Wen Daozhou. A study on house price prediction model based on Graph Neural Network and Long Short-Term Memory model. Computer Application Research: 1-8.

[9] Jiang Yuyan, Liu Hao. A residential price prediction model based on Pearson coefficient and firefly algorithm optimized BP neural network. Journal of Nanyang Institute of Technology, 2023,15 (02): 1-6 24.

[10] Yin Lina. The impact of the epidemic on China's real estate economy and corporate response strategies. Modern Enterprise, 2023, (02): 70-72.