The Prediction of the Guangzhou Housing Market Based on the ARIMA Model

Yanfu Wang[1, *]

[1]Department of AP & AL, Guangzhou Foreign Language School, Guangzhou, China

*Corresponding author: bxzt681@ksu.edu.bi

Abstract. Since the increasing trend of the Guangzhou housing price stopped after the Covid-19 epidemic in early 2020, investors and citizens are hesitant about purchasing, keeping, or selling their real estate in Guangzhou. This paper intends to utilize the ARIMA model to forecast the Guangzhou real estate market from September 2022 to August 2023 based on the Guangzhou housing price data from the last 10 years, and compare the yearly profits of purchasing houses in Guangzhou and the profits of storing money in the Bank of China for a year. The result is counter-intuitive that purchasing houses in Guangzhou is much more profitable than storing cash in the Bank of China since the Guangzhou housing market is experiencing a trough after the Covid-19 epidemic, which created a great potential for its future recovery. To sum up, this paper suggests local citizens and investors increase the property proportion of real estate in Guangzhou in 2023 due to its upcoming profit.

Keywords: ARIMA, Guangzhou, Housing market, Time series.

1. Introduction

1.1 Research background

The housing market in China is one of the most popular topics in recent years as it developed rapidly after the Chinese reforming and opening policies proposed in 1978. In detail, on the morning of October 20, 1978, leaders in Beijing asked when inspecting dozens of newly built apartment buildings, "Can houses become a commodity?" In fact, houses did not become commodities at that time. The year 1980 was the budding stage of real estate in China when houses were defined as a commodity and formed an industry. But it was not until 1987 that real estate in China began to be truly commercialized. For the next decade, the real estate market began to boom. In August 2003, the State Council Document 18 clearly proposed that the real estate industry is a pillar industry of the national economy. In subsequent years, housing prices rose with continuous regulation, which is the most glorious golden decade of real estate development. However, since the housing market is overheated which may threaten economic stability, the government suggested that houses should be used to live instead of investing and proposed a series of policies to control the real-estate bubble in 2017 [1]. Also, the housing demand gradually eased, and the shortage was not as serious as previous. Therefore, the housing market was cooling down at that time. Most essentially, the covid-19 epidemic appeared in late 2019 in Wuhan and spread out to the whole nation in early 2020, which greatly threatened the housing market. To be specific, in order to relieve the spreading tendency of the epidemic, governments began to restrict citizens' mobility, which negatively affects their careers and incomes, so less people were able to purchase houses [2]. During the Spring Festival holiday in 2020, the area of commercial properties sold in large and medium-sized cities was 3.75 million square meters, which dropped 81% comparing with sales in 2019 (19.55 million square meters) and even much lower than the 7.61 million square meters sold in 2018 [3]. Moreover, since the sales of real estate dropped unprecedently, the housing companies were threatened by the scarcity of cash flow. Therefore, some of them failed to offer houses to customers who had already paid, so the credibility of the real estate companies declined, which negatively affected the sales. In conclusion, after the pandemic, the Chinese housing market seems to cool down evidently.

The situation mentioned above indispensably occurred in Guangzhou, one of the first-tier cities in China which has an expensive housing value. Due to the periodicity of the covid-19, the housing price

in Guangzhou became fluctuant in the last 2 years. A great number of citizens are concerned about the prospect of their house value, but recent research and prediction about such a local housing market are still absent. As a result, it is necessary to analyze the situation they are confronting and give them practical guidelines. Such guidelines aim to help them optimize their choice of investment and property structure, which can reinforce their personal worth in such an unstable market and special era.

1.2 Research contents

Under the background mentioned above, how to accurately predict the trend of housing prices and establish an appropriate model has been a hot concern of scholars. ARIMA model is a kind of random time series model that excludes complex influencing factors and only focuses on looking for self-correlation of historical data to predict the future, which is relatively suitable for doing housing price prediction [4]. Therefore, in order to give a guideline for citizens and investors in Guangzhou, this paper decides to utilize the ARIMA model to forecast the Guangzhou housing price per square meter and the corresponding rent price per square meter from September 2022 to August 2023, which is a one-year prediction. Then the predicted consequences above are used to calculate the predicted monthly housing return rate and the ratio between the rent price and the housing price in 2023. The sum of these values every month would be compared with the interest rate of the one-year fixed-time deposit started in August 2022. In the end, such a comparison is able to give an accurate answer about whether to invest in houses or store the money in banks to wait for other investments, which is a meaningful suggestion for potential investors and house owners.

2. Methodology

2.1 Data

The data used in this paper for establishing the model are separately the monthly Guangzhou housing price per square meter and the rent price per square meter from September 2012 to August 2022. The first data, which is about the Guangzhou housing price, is collected from Gotohui.com. This website is a platform that collects and summarizes data about countries and their regions—such as their GDP, incomes, government finances, housing prices, and so on [5]. The other data used by this paper, which is about the Guangzhou rent price, is collected from a website called creprice.cn. This platform is a data application platform guided by the China Real Estate Association. It publishes the actual house price, rent, supply, and demand information of each city based on Xitai data. Both of these platforms used by this paper are professional in serving their data for academic research, property evaluation, consulting services, marketing planning, and so on [6]. Therefore, the data they presented are reliable.

Fig.1 The Guangzhou housing price from 2012 to 2022

After receiving the Excel of these data, it is required to plot these data by utilizing the R studio. As shown in Figure 1 and Figure 2, the left graph above indicates the time series of the monthly Guangzhou housing price from September 2012 to August 2022. This plot contains 120 samples and has a monthly recording frequency. It shows that in the first eight years of this series, the housing price increased at a relatively constant rate, which indicates an apparent trend. However, in the last two years, the housing price became volatile and strongly fluctuated between values about 42,000 yuan. The reason is that during the last two years, the sales of the housing market are greatly affected by the epidemic policy in Guangzhou. Sometimes citizens and investors had to stay at home due to the temporary local policy, so nearly no trade can be dealt with at that time, which means that the housing price would drop significantly. By contrast, after the temporary movement limitation is removed, the demand for purchasing houses is accumulated before and released at that time. Thus, the housing price would rise dramatically in a short period. The right graph above plots the time series of the Guangzhou rent price from September 2012 to August 2022. Same as the left one, it contains 1200 samples with a monthly frequency. Compared with the housing price, the time series of the rent price looks more fluctuant. It slowly increased in the first 4 years, and then dropped in the fifth year. On the contrary, after the fifth year, it rose dramatically in a relatively short term. Then the rent price decreased at a slow rate for the next years. In the last two years, which is the epidemic era, the rent price rose to an unprecedented level in a short term and then fluctuated around that level, which is about 60 yuan per meter square.

Fig.2 The Guangzhou rent price from 2012 to 2022

2.2 ARIMA model

The autoregressive integrated moving average (ARIMA) model, one of the core models of the time-domain analysis, is established by American statistician G.E.P.Box and British statistician G.M.Jenkins. Therefore, it is also called as Box-Jenkins model. The ARIMA model can be used to fit differential smooth series, which are non-stationary in nature and tend to exhibit the properties of a stationary series after the differencing operation. The structure of the ARIMA (p,d,q) model is as follows:

$$X_t=\varphi_0+\varphi_1 X_{t-1}+\ldots+\varphi_p X_{t-p}-(\varepsilon_1+\theta_1\varepsilon_{t-1}+\ldots+\theta_q\varepsilon_{t-q}) \quad (1)$$

where $E(\varepsilon_t)= 0$, $Var(\varepsilon_t) = \sigma_\varepsilon^2$, $E(\varepsilon_t\varepsilon_s) = 0$, $s \neq t$; q denotes the qth order difference [7].

In this case, the ARIMA model would be utilized to predict the Guangzhou housing price and rent price from September 2022 to August 2023 based on the ten-year' time series. Then the real estate price time series is required to be transformed into a return rate time series which is used to predict the housing return rate from September 2022 to August 2023. Accordingly, this paper intends to build

3 ARIMA models to separately forecast the Guangzhou real estate price, the rent price, and the real estate return rate for the subsequent analysis.

Since the Guangzhou government intends to stabilize the housing price, the result of the forecasting might be the housing price would gradually return to an average price during the pandemic era which is 42,000 yuan per square meter. Therefore, the return rate of the housing price might not be large but positive. By the way, the rent price might still increase in 2023 due to the inflation during the pandemic era. Also, since the housing price would not drop dramatically in a short term, people who are unable to purchase a house still have a great demand for renting houses? Therefore, the prediction of the Guangzhou rent price might still rise. About the comparison between the house earning rate (combining its return rate and the ratio of the monthly rent price and its housing price) and the interest rate of the one-year fixed-time deposit started in August 2022 (which is 1.75% declared by the Bank of China), the one-year interest rate might be greater than the house earning rate since the return rate might be too small [8].

3. Results

3.1 General process of establishing ARIMA models

First of all, most codes in the following process are supported by the book called R in Action [9]. This paper separates the construction of an ARIMA model into 4 steps: verifying the sequence stationarity, Identifying reasonable models, evaluating the model fit, and Making forecasts. Such steps mentioned above would guide the following prediction in order to get an accurate result. In step 1, the stationarity of the time series needs to be proved as the ARIMA model can only forecast stationary series. In step 2, potential models would be compared with each other based on their AIC value, and the one with the smallest AIC value would be chosen as the most appropriate model. In step 3, after selecting a model, it is necessary to evaluate its fitness by plotting various graphs—such as histograms and Q-Q plots—and tests that can determine their residual distribution. If step 4 indicates that the model has a good fit, the last step could be done which is making forecasts with graphs and values. By the way, all the functions this paper used are supported by the R studio.

3.2 Forecasting the Guangzhou housing price

As this paper mentioned above, in predicting the Guangzhou housing price, 4 steps would be required to complete.

To begin with, the time series of the previous Guangzhou housing price shows an apparent tendency that is constantly increase. Therefore, it is absolutely unstationary which is required to be differenced. So the ndiffs function, which is used to measure how many times the time series needs to be differenced, answers that the series only need to be differenced with one time. After differencing, the ADF test (the Augmented Dickey-Fuller Test) shown above, which is a test to determine whether the series is stationary or not, indicates that the differenced time series is stationary since the p-value is less than the critical value. In conclusion, the time series is stationary when d=1.

After verifying the sequence stationarity, the next step is to select an appropriate model by determining the p (the order of the autoregressive model) and q (the order of the moving-average model) values. In this case, the auto.arima function recommends the best model is the ARIMA(1,1,0) with drift, which owns the smallest AIC (Akaike information criterion, an indicator to measure the goodness of statistical model fitting) value compared with other models.

In Figure 3, the fitness of the model is required to be evaluated. In this case, the qqnorm and qqline function is applied in order to produce the normal Q-Q plot as shown below. It shows that the residuals seem to have a thick tail distribution, which means the extreme value is more probable to occur. Therefore, the confidence interval forecasted by the auto.arima function—assuming that the residual has a normal distribution—would be inaccurate, but the point forecast should be accurate as the extreme value below and above the line is symmetric.

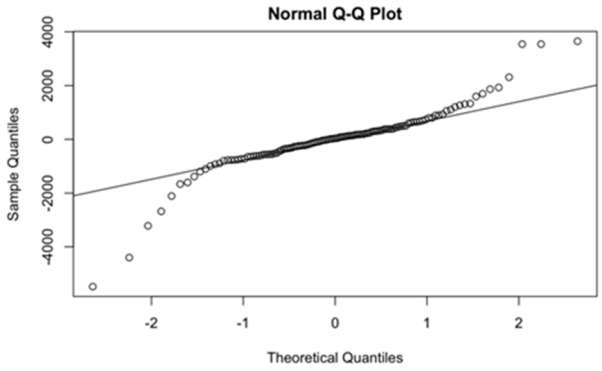

Fig.3 Normal Q-Q plot of the Guangzhou housing price

Furthermore, the histogram on the next page plotted by the hist function indicates that the peak of the residuals seems to be much higher than the normal distribution in Figure 4. Therefore, in order to confirm the statement mentioned above, the kurtosis function is applied to judge the kurtosis of the residual distribution. The result of this function is 5.580184 (that means it is a fat-tail distribution), which indicates that its pinnacle is much higher and sharper than the normal distribution. Getting back to the point, moreover, the Box-Ljung test could be applied to testify that the autocorrelations of residuals are all zero. In this case, the result of this test is not significant (p-value is greater than 0.05), suggesting that the autocorrelations don't differ from zero. Accordingly, despite those extrema, this ARIMA model appears to fit the data well in use to do the point forecast but needs to be adjusted in evaluating the confidence interval. However, since this paper only focuses on the point forecast, such an error mentioned above can be neglected.

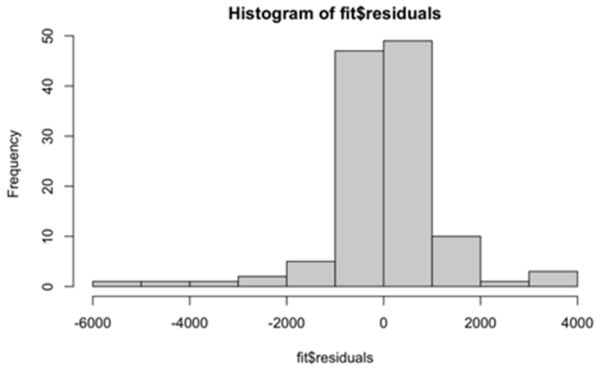

Fig.4 Histogram of fit & residuals of the Guangzhou housing price

The last step is making a forecast about the monthly Guangzhou housing price from September 2022 to September 2033 in Table 1 and Figure 5. The result is shown below by using the forecast function.

Table 1. The forecast about the monthly Guangzhou housing price

Date (M/Y)	9/22	9/23	9/24	9/25	9/26	9/27	9/28	9/29	9/30	9/31	9/32	9/33
Price	40417	40755	40907	41136	41333	41543	41748	41955	42161	42368	42574	42781

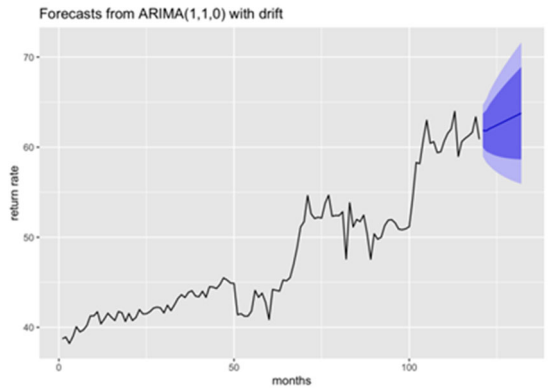

Fig.5 Forecasts from ARIMA (1,1,0) with drift of the Guangzhou housing price

3.3 Forecasting the Guangzhou rent price

Same as the process in 3.2, the time series of the Guangzhou rent price is un-stationary initially and is required to be differenced. After differencing one time, the time series passed the ADF test as shown below.

Then in step 2, the auto.arima function still recommends the best model is the ARIMA (1,1,0) with drift due to the smallest AIC value.

In Figure 6, the situation in 3.2 occurs again. Both the Q-Q plot, the histogram, and the kurtosis test show that the residuals have a thick-tail distribution, so the confidence interval predicted by the auto.arima is inaccurate. By contrast, since the extreme value below and above the line is symmetric, the point forecast is valid.

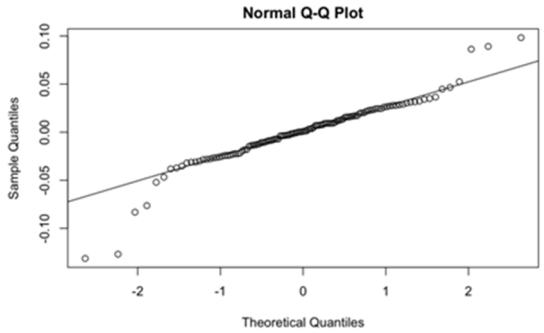

Fig.6 Normal Q-Q plot of the Guangzhou rent price

Also, the Box-Ljung Test is passed, suggesting that the autocorrelations don't differ from zero in Figure 7. In conclusion, the point forecast is still practical in predicting the rent price.

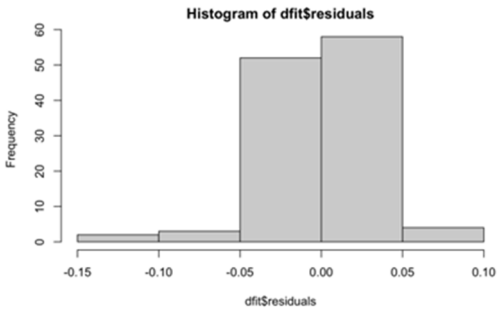

Fig.7 Histogram of fit & residuals of the Guangzhou rent price

Last but not least, the predicted Guangzhou rent price per square meter from September 2022 to September 2033 shown below could be created by using the forecast function in Table 2 and Figure 8.

Table 2. The predicted Guangzhou rent price per square meter

Date (M/Y)	9/22	9/23	9/24	9/25	9/26	9/27	9/28	9/29	9/30	9/31	9/32	9/33
Rent	61.84	61.81	62.067	62.239	62.44	62.626	62.817	63.008	63.199	63.391	63.582	63.773

Fig.8 Forecasts from ARIMA (1,1,0) with drift of the Guangzhou rent price

3.4 Forecasting the Guangzhou real estate return rate

In order to predict the Guangzhou real estate return rate, it is indispensable to transform the time series of the housing price into the return rate time series. In detail, after guaranteeing that the data is numeric, a combined function including diff and log is utilized to turn the gold price into the log return, which is similar to the rate of return.

After transforming the data, the following process would be similar to previous situations. First of all, since the time series have been differenced and turned into the log scale, it is stationary which is claimed by the ADF test. After that, the auto.arima function recommends the best model is the ARIMA (1,0,0) with non-zero mean (since the time series have already been differenced once, the d value is actually 1 for the original time series). In Figure 9, which is evaluating the model fit, the model recommended by the auto.arima function performs extremely well since most of the residuals are on the line in the normal Q-Q plot.

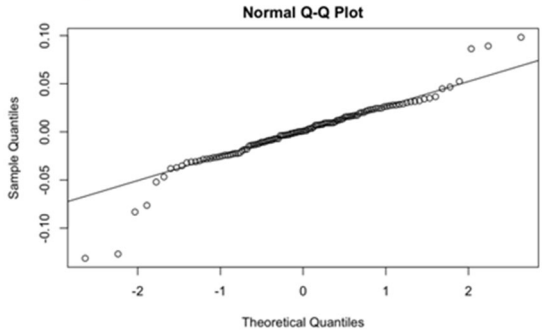

Fig.9 Normal Q-Q plot of the Guangzhou real estate return rate

Also, the histogram shows that these residuals are distributed almost normally in Figure 10. Thus, the model appears to fit well in predicting the return rate.

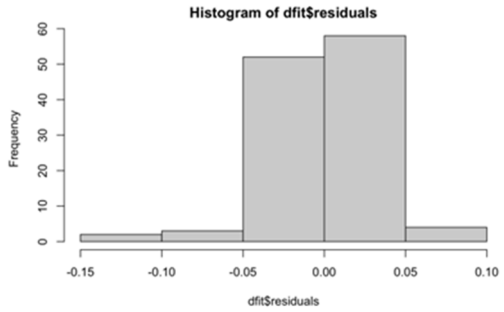

Fig.10 Histogram of fit & residuals of the Guangzhou real estate return rate

In the end, the predicted Guangzhou real estate return rate from September 2022 to September 2033 including graphs and values is shown in Table 3 and Figure 11.

Table 3. The predicted Guangzhou real estate return rate

Date (M/ Y)	9/22	9/23	9/24	9/25	9/26	9/27	9/28	9/29	9/30
Rent	0.003506 179	0.009113 851	0.007579 352	0.007999 257	0.007884 353	0.007915 795	0.007907 191	0.007909 546	0.007908 902

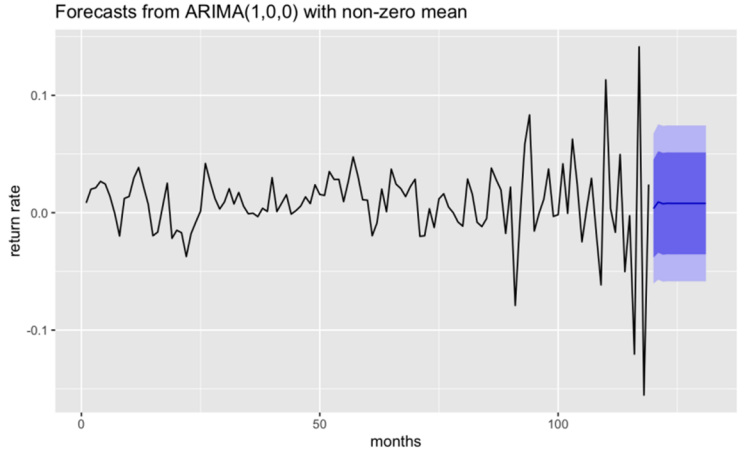

Fig.11 Forecasts from ARIMA (1,0,0) with drift of the Guangzhou real estate return rate

3.5 Analysis and comparison

In order to calculate the monthly house earning rate, it is required to calculate the ratio between the monthly predicted rent price and the house price. The result is shown in Table 4, which could be concluded that the ratio is about 0.15% every month. Then the certain monthly ratio is combined with the return rate, which is the monthly earning rate shown below. After that, all the monthly house earning rate is combined together to calculate the yearly house earning rate which is about 10.9%. Therefore, surprisingly, the profit rate created by Guangzhou houses is extremely higher than the one created by the one-year fixed deposit in banks which is 1.75%. To sum up, most pre-judgments written in the last paragraph of 2.2 are corresponding to the result, but the final conclusion is entirely different.

Table 4. The ratio between the monthly predicted rent price and the house price

month	rent price	house price	ratio	return rate	earning rate
9/2022	61.84302	40417.15	0.001530118	0.003506179	0.005036297
10/2022	61.81117	40755.32	0.001516641	0.009113851	0.010630492
11/2022	62.06679	40906.72	0.001517276	0.007579352	0.009096628
12/2022	62.23935	41136.15	0.001513009	0.007999257	0.009512266
1/2023	62.4359	41332.98	0.001510559	0.007884353	0.009394912

2/2023	62.62553	41543.43	0.001507471	0.007915795	0.009423266
3/2023	62.81715	41748.19	0.001504668	0.007907191	0.009411859
4/2023	63.0082	41955.32	0.001501793	0.007909546	0.009411339
5/2023	63.19941	42161.46	0.001498985	0.007908902	0.009407887
6/2023	63.39058	42368.02	0.001496189	0.007909078	0.009405267
7/2023	63.58176	42574.41	0.001493427	0.00790903	0.009402457
8/2023	63.77293	42780.86	0.001490688	0.007909043	0.009399731
Sum	752.79179	499680.01	0.018080824	0.091451577	0.109532401

4. Discussion

4.1 Reasons analysis

The pre-judgment on 2.2 explains most reasons for the predicted result. First of all, the housing price would increase to an average level of the price after the pandemic since the local government intends to stabilize the real estate market. Also, during the epidemic era, people's income becomes unstable, so they would be less likely to purchase houses by making loans. As a result, the strong increasing trend before the epidemic would not occur in 2023 due to the serious economic recession. In conclusion, as the housing price dropped in the last season, it is more likely to recover to an average level of housing price after the pandemic, which corresponds to the prediction shown on the left forecasted by ARIMA. By the way, as the housing price would recover, the return rate would be positive.

Also, the rent price in Guangzhou would increase since the people who are unable to afford purchasing houses would tend to rent houses. Therefore, due to the increasing demand, it is reasonable that the rent price in 2023 tends to increase.

However, the final result is counter-intuitive, which is opposite to the pre-judgment. The pre-judgment mentioned in 2.2 underestimated the annual return rate of the housing price. In fact, in the last season, the housing price dropped about 12.3%, which reserves a great potential for the upcoming recovery. Also, other research mentioned that Guangzhou Centaline Research and Development Department believes that it is unlikely that although Guangzhou will introduce relaxed policies in the second half of the year, buyers are expected to increase demand due to the economic recovery and optimization of the managemant, driving the market out of the doldrums [10]. Therefore, in the prediction, the housing price would totally recover about 9.14% in a year. Although it is still lower than the failing level in the last season, it is a great amount compared with the yearly interest rate proposed by the Bank of China which is 1.75%. Also, even the sum of the ratio (monthly rent price/monthly housing price) is about 1.8%, which is similar to the yearly interest rate. This result is also counter-intuitive since the monthly ratio is small compared with the yearly interest rate. But the monthly ratio would be added together in order to conclude a yearly ratio since the rent can be received with a monthly frequency. Therefore, the sum of them would be even greater than the interest rate which can only be received with a yearly frequency.

4.2 Suggestions

Since the forecasting result indicates that the housing price would be recovered to an average level after the pandemic, this paper recommends local citizens keep their houses in 2023. For investors, this paper even suggests they purchase the houses in Guangzhou as a short-term investment because it is currently experiencing a trough and their values would rise in 2023. Finally, for everyone who is able to buy houses in Guangzhou, such expenditures could resist the inflation caused by the epidemic. To sum up, taking all the analysis into account, this paper recommends that purchasing houses in Guangzhou would be a wise choice in 2023.

4.3 Limitation analysis

As this paper mentioned in part 3, the predicted confidence intervals in the prediction of the housing price and the rent price are inaccurate. Therefore, this paper failed to evaluate a valid confidence interval for measuring the risk of keeping or buying houses people need to undertake. Also, since the housing price tends to be much more fluctuant than previous, making a practical confidence interval would be meaningful for measuring future fluctuation.

5. Conclusion

Based on the ARIMA model, this paper successfully predicts the housing price, the rent price, and the housing return rate in Guangzhou from September 2022 to August 2023. The forecasting results indicate that the housing price would increase in a year, which means that the return rate would also be positive. Also, the rent price would also slightly increase. In the analysis part, this paper concludes that both the sum of the monthly return rate and the sum of the ratio between the predicted rent price and the house price are greater than the interest rate of the one-year fixed-time deposits. Therefore, this paper recommends investors and citizens keep or even purchase houses in Guangzhou due to the great upside potential caused by the recession in the last season. Such research fills a vacancy in using time series models to predict and analyze Guangzhou housing prices in 2023. It presents a guideline for people who are hesitant about purchasing, investing, or selling real estate in Guangzhou, which is volatile since the epidemic occurred. However, since the ARIMA model is unable to capture the non-linear relation, who frequently occurs in the housing price, the predicted result has residuals with extreme value. Therefore, an accurate confidence interval would be significant in order to include such variations. Unfortunately, due to the thick-tail distribution of the residuals, this paper failed to predict a valid confidence interval for investors and citizens to measure the upcoming fluctuation level of the Guangzhou housing price. Therefore, in future research, the neural network models that are able to capture non-linear relationships, such as the LSTM model, can be applied to improve the accuracy of the prediction.

References

[1] JiJing. China's real estate wind and cloud 40 years. China's real estate, 2019 (02) : 10 to 11.

[2] Yin X T. Study on problems and countermeasures of Chengdu's new housing deal after the COVID-19 epidemic. University of Electronic Science and Technology of China,2022.

[3] LIU Y. Current situation, problems and countermeasures of real estate market in post-epidemic period. Market in China, 2021 (8) : 13-16.

[4] Qiwei Fu, Yanchun Yi, Shijia Zhang, Minlan Tang, Jia Ding. Science and technology innovation and application,2021,11(31):47-50.

[5] Gotohui Homepage, https://fangjia.gotohui.com, last accessed 2022/9/19.

[6] National Price Quotes Homepage, https://www.creprice.cn/city/gz.html, last accessed 2022/9/19.

[7] You Meifang, Huang Min, Cheng Li. Application of ARIMA Model in Housing Price Forecasting: An Empirical Analysis of the Future Trend of Sichuan Commodity Housing Price Index. China Prices,2009(06):38-40.

[8] Bank of China, from https://www.bankofchina.com/, last accessed 2022/9/19.

[9] Robert Kabacoff. R in Action: Data Analysis and Graphics with R, S. R in Action: Data Analysis and Graphics with R. Manning Publications (2015).

[10] China Business News Homepage, https://baijiahao.baidu.com/, last accessed 2022/9/19.