

# Prediction of Opening Price of Shanghai Stock Exchange 380 Index Based on ARIMA Model

Qi Zuo<sup>1</sup>, Yumei Wang<sup>1,\*</sup>

<sup>1</sup>School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, China

\*Corresponding author: Yumei Wang (Email: wym9808@163.com)

**Abstract:** This paper first introduces the relevant theoretical knowledge about time series analysis, then makes an empirical analysis and finally summarizes it. This paper selects 97 data of the opening price of Shanghai Stock Exchange 380 Index from June 1, 2023 to October 30, 2023, fits the ARIMA model and tests the validity of the model. Finally, the opening price of the next three trading days is predicted by using the fitted model and compared with the real value to test the prediction effect.

**Keywords:** ARIMA model, Opening price, SSE 380 Index, Time series.

## 1. Introduction

With the rapid development of our society, China's financial market is becoming more and more complete, and it has gradually become one of the most important capital markets in China, which plays an important role in China's market economic system. After more than 30 years' development, China's stock market has gone through the journey that western countries took hundreds of years to complete. In the stock market, investors will decide whether to invest in stocks according to various income indicators of stocks, which depend on investors' judgment on the future trend of stocks. The sample companies of SSE 380 Index are all companies with moderate scale, active transactions, high growth characteristics and growth potential in Shanghai stock market, reflecting the overall situation of a group of high-quality large-cap enterprises with the most market influence in Shanghai stock market, so they have very good market representation and have a guiding role in China's stock market. Therefore, the prediction of the future trend of the Shanghai Stock Exchange 380 Index plays an important role in the investment market.

With the rapid development of the stock market, more and more scholars began to use time series method to predict the stock market, and achieved good results. However, the time series analysis method only uses historical data to extract useful information to predict the future trend, and does not deeply consider the reasons for stock price changes. It belongs to intuitive analysis method and can only make short-term predictions.

The sample selected in this paper is the opening price of Shanghai Stock Exchange 380 Index from June 1, 2023 to October 25, 2023. On this basis, the ARIMA model is built to predict and analyze the time series, so as to provide investors and entrepreneurs with direction when selecting stocks and provide a better basis for the government to formulate relevant policies.

## 2. Theoretical Research

### 2.1. Time Series

Time series is a data series obtained by arranging the observed values in time order for a certain phenomenon. In real life, the observed values of many phenomena can be

regarded as time series. Establishing time series is the basis of time series model. Through time series, we can vividly reflect the development and change trend of social and economic phenomena.

### 2.2. Stationary Time Series Model

#### 2.2.1. AR Model

Because of the inertia of time series data, there is a correlation between time series data before and after. In other words, there is a correlation between current data and historical data of time series. The model with the following structure is called autoregressive model, abbreviated as AR(p):

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \dots + \varphi_p x_{t-p} + \varepsilon_t \\ \varphi_p \neq 0 \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma^2_\varepsilon, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(x_s \varepsilon_t) = 0, s < t.$$

#### 2.2.2. MA Model

Sometimes, the current data of time series has the memory of external shock value. The model with the following structure is called the moving average model and abbreviated as MA(q):

$$x_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \\ \theta_q \neq 0 \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma^2_\varepsilon, E(\varepsilon_t \varepsilon_s) = 0, s \neq t.$$

#### 2.2.3. ARMA Model

When the current data of time series not only has the memory of external shock value, but also has a certain correlation with historical data, when the model is used to describe this dynamic feature, it should include both lag term and external shock term, and this model describing dynamic feature is called autoregressive moving average model. The structure of autoregressive moving average model is usually as follows, abbreviated as ARMA(p, q):

$$x_t = \varphi_0 + \varphi_1 x_{t-1} + \varphi_2 x_{t-2} + \dots + \varphi_p x_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \\ \varphi_p \neq 0, \theta_q \neq 0 \\ E(\varepsilon_t) = 0, \text{Var}(\varepsilon_t) = \sigma^2_\varepsilon, E(\varepsilon_t \varepsilon_s) = 0, s \neq t \\ E(x_s \varepsilon_t) = 0, s < t.$$

### 2.3. ARIMA Model

A wide stationary time series usually has the mean and variance of constant values, but in the economic field, the time series studied is usually non-stationary, which requires us to establish a non-stationary time series model.

The autoregressive integrated moving average model with the following structure is abbreviated as ARIMA(p, d, q) model:

$$\begin{aligned} \Phi(B)\nabla^d x_t &= \Theta(B)\varepsilon_t \\ E(\varepsilon_t) &= 0, \text{Var}(\varepsilon_t) = \sigma_{\varepsilon}^2, E(\varepsilon_t \varepsilon_s) = 0, s \neq t. \\ E(x_s \varepsilon_t) &= 0, s < t. \end{aligned}$$

## 2.4. Steps of Establishing ARIMA Model

### 2.4.1. Testing the Stationarity of the Data

One method is to judge by time series, and to observe whether the time series has trends and periodicity by observing the trend diagram of the time series. Another method is to calculate the ADF value of statistics.

### 2.4.2. Smoothing the Data

If the series has a definite trend, we can eliminate its trend to obtain a stable time series. If there is no definite trend in the series, the trend can be removed by difference.

### 2.4.3. Determining the Order of the Model

The order of the model can be determined according to the properties of ACF and PACF.

**Table 1.** The properties of ACF and PACF

Model	ACF	PACF
AR(p)	Trailing	P-order truncation
MA(q)	Q-order truncation	Trailing
ARMA(p, q)	Trailing	Trailing

### 2.4.4. Testing the Model

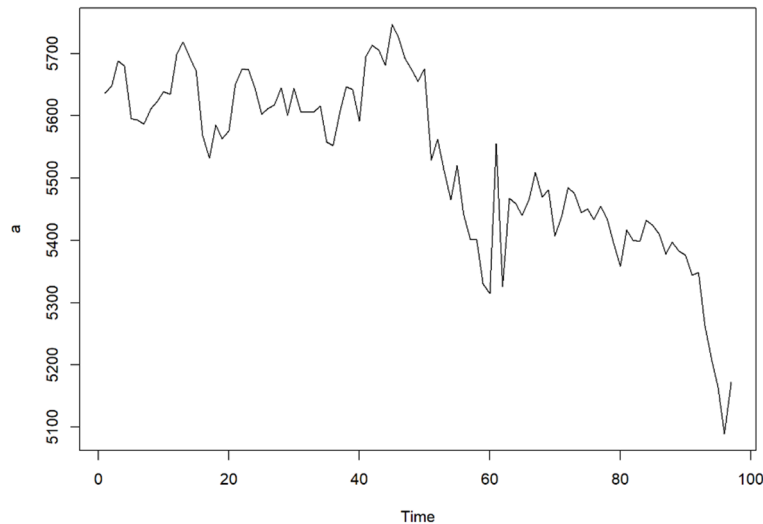
After the model identification and parameter estimation are completed, the model test is actually the residual independence test of the model to ensure that the model extracts most of the effective information in the time series.

## 3. Empirical Analysis

In this paper, the opening price of Shanghai Stock Exchange 380 Index from June 1 to October 25, 2023 is selected, with a total of 97 data, and ARIMA model is fitted to predict the opening price of the last three trading days.

### 3.1. Testing the Stationarity of the Data

Draw a time chart for 97 data of the opening price of the SSE 380 Index on all trading days from June 1 to October 25, 2023, as shown in the following figure:

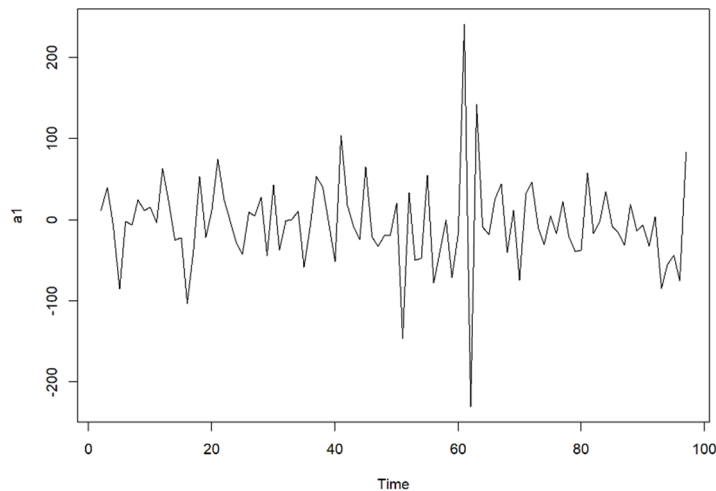


**Figure 1.** The opening price of the SSE 380 Index

The above figure shows that the series is non-stationary, so we need to do difference processing to transform the non-stationary time series into stationary time series.

### 3.2. Smoothing the Data

The time series diagram after the first-order difference is shown in the following figure:



**Figure 2.** The time series diagram after the first-order difference

The above figure shows that the sequence after the first-order difference basically fluctuates around the value of 0, and there is no obvious trend characteristic at this time, so we can think that the sequence after the first-order difference is stationary. In order to verify whether the time series after the first-order difference is really stable, we need to continue ADF test. ADF test results are as follows:

**Table 2.** ADF test results

Test statistic	Dickey-Fuller critical value		
	1%	5%	10%
<b>-14.054</b>	-4.051	-3.455	-3.153

ADF test results show that the value of statistics is less than the critical value at 1% level, so we can think that the sequence after first-order difference is stationary.

The white noise test is carried out on the stationary

sequence after the first-order difference, and the results are as follows:

**Table 3.** Results of white noise test

Lag Orders	X-squared	Df	P-value
<b>6</b>	16.89	6	0.009697
<b>12</b>	38.701	12	0.0001179

The results of white noise test show that the p values of LB statistics with lag order of 6 and 12 are all less than 0.05, so we can think that the sequence after first-order difference is not a white noise sequence.

### 3.3. Determining the Order of the Model

The ACF and PACF diagrams of the first-order differential sequence are as follows:

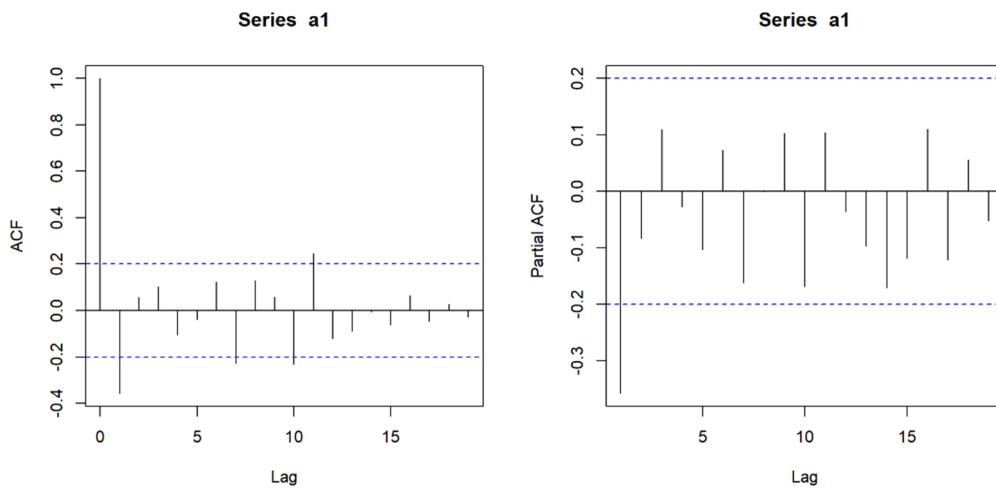


Figure 3. The ACF and PACF diagrams

Observing the graphs of ACF and PACF, we can think that ACF is trailing, PACF is truncated in the first order, and AR(1) model should be fitted. We have made the first-order difference of the series. In summary, we should choose ARIMA(1, 1, 0) model to fit the opening price series of the

Shanghai Stock Exchange 380 Index from June 1 to October 25, 2023.

### 3.4. Testing the Model

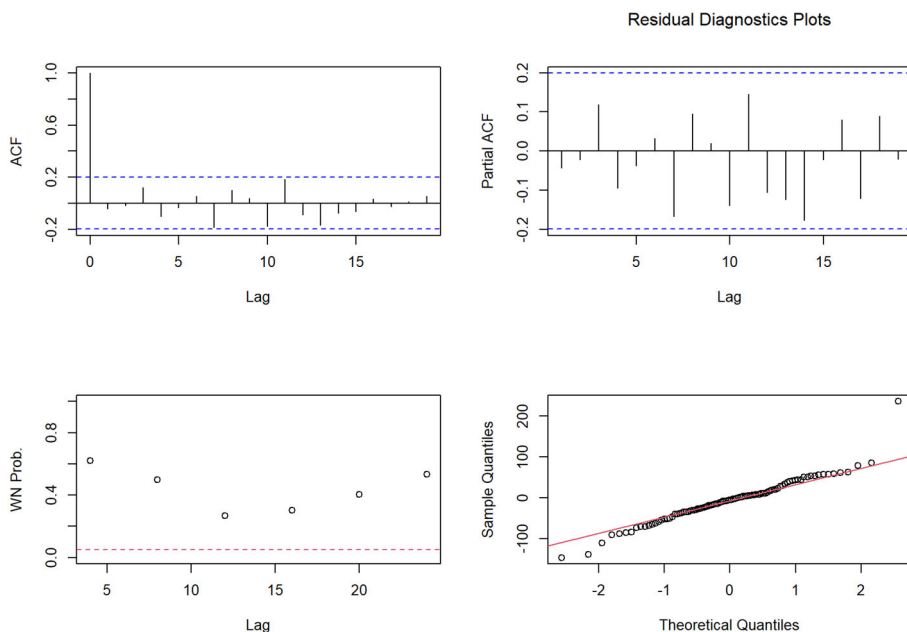


Figure 4. Results of model test

As shown in the figure on the left, the P values of all Q statistics are greater than 0.05, so we can think that ARIMA(1, 1, 0) model has passed the significance test and the model is remarkably established.

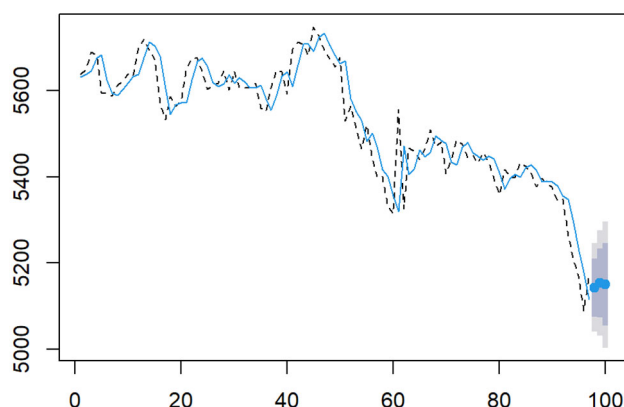
### 3.5. Making predictions

We use ARIMA(1, 1, 0) model to forecast the opening price of the next three trading days. The forecast results and the forecast effect chart are as follows:

**Table 4.** The forecast results

Transaction Date	True Value	Predicted Value	Absolute Error	Relative Error	Average Relative Error
2023-10-26	5,140.3060	5143.414	-3.108	0.0006	0.0064
2023-10-27	5,157.2980	5153.707	3.591	0.0007	
2023-10-30	5,243.6660	5150.095	93.571	0.0178	

**Forecasts from ARIMA(1,1,0)**



**Figure 5.** The forecast effect chart

The prediction results show that the average relative error is only 0.0064. In the prediction diagram, the blue solid line represents the observed value of time series, the black dotted line represents the fitting value, and the blue dot in the shadow part represents the prediction value, so it can be seen that the fitting effect is good.

## 4. Conclusion

In a word, it is feasible to use ARIMA model to predict the opening price series of Shanghai Stock Exchange 380 Index. The ARIMA model is used to analyze the opening price series. The results show that the model has a good short-term prediction effect on the opening price of the Shanghai Stock Exchange 380 Index, and the fitting model fully extracts the useful information contained in the opening price series. Of course, there are some limitations in this study, such as subjectivity and experience in the analysis of the original sequence and the differential sequence, especially the analysis of ACF and PACF of the sequence; In addition, in the process of sequence analysis, other factors such as national macroeconomic policies should be considered.

## References

- [1] Zou Zhiyuan. Analysis and prediction of Shanghai Stock Exchange Index based on ARIMA model [J]. Hebei Enterprise, 2021 (03): 36-37. DOI: 10.19885/j.cnki.hbqy.2021.03.014.
- [2] Guo Gaiwen, Wang Shihan. Stock price prediction based on grey theory and ARIMA model [J]. Journal of Henan Education Institute (Natural Science Edition), 2023,32(02):22-27.
- [3] Huang Mengting. Empirical study on stock price prediction based on ARIMA model [J]. Neijiang Science and Technology, 2023,44(03):61-62.
- [4] Zhao Ning. Short-term forecast of Vanke's stock price based on ARIMA model [J]. Financier, 2023(01):34-36.
- [5] Huang Shimin. Stock price analysis and forecast based on ARIMA model-taking China Merchants Bank as an example [J]. SME Management and Technology, 2022(11):184-187.
- [6] Zhang Yingchao, Sun Yingjuan. Empirical study on the analysis and prediction of Shanghai Stock Exchange index based on ARIMA model [J]. economic research guide, 2019 (11):131-135.
- [7] Li Xiaoxian. Forecasting the trend of Shanghai Stock Exchange Index based on ARIMA model [J]. Liaoning Economy, 2017 (07): 94-96.
- [8] Ma Yanna, Zeng Jiying. Prediction and analysis of Shanghai Stock Exchange Index-based on ARIMA model [J]. Economic and Trade Practice, 2017(03):18-19.
- [9] Chen Hao, Lan Yanhong, He Yubo. Analysis and forecast of stock opening price based on ARMA model [J]. Journal of Jishou University (Natural Science Edition), 2020,41 (05): 86-91.
- [10] Li Xiangguang. Analysis and forecast of consumption ETF based on ARIMA model [J]. Caixun, 2021,000(12):P.140-143
- [11] Wang Lu. Shanghai Stock Exchange adds the emerging blue chip index [N]. shanghai securities news, 2010-11-15(001).
- [12] Huang Zhiyi. The impact of capital market opening on the tax avoidance of enterprises [D]. Southwestern University of Finance and Economics, 2023.