2021 International Conference on Computer, Blockchain and Financial Development (CBFD) | 978-1-6654-1227-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/CBFD52659.2021.00046

Stock Prices Prediction Based on ARMA Model

Huanze Tang
School of Mathematics and Statistics
Wuhan University
Wuhan, China
2018302010084@whu.edu.cn

***Abstract*—The financial time series contain some information that indicates the operation law of the system. Researchers can use classic models of time series to study previous stock prices and predict a short-term trend of the volatility of the prices. In this article, we choose the adjusted closing prices of Apple Inc from 2018 to the end of 2019. Then we perform the first difference on the original data to make the sequence stationary to apply the ARMA model to predict the adjusted closing prices of Apple Inc in the next five days. The time series, which we predict, is compared to the actual value. And it turns out that the data's error rates are low, indicating that the ARMA model is suitable for the short-term prediction of the prices and further. Meanwhile, it further proves that the time series model serves as a positive catalyst in the study of finance.**

Keywords-time series; ARMA model; adjusted closing prices; Apple Inc; Python

I. INTRODUCTION

There are many time series in the financial market, such as the changing stock prices and returns, which usually contain the whole system's operation law. Hence, researchers can build up a time series model to figure out the rule of the changing data to predict the trend of the stock prices [1].

As we all know, stocks can tell if the economy is good, so stock price forecasts are of great significance to the country and investors. The ARMA model (p, q) is currently the most commonly used model for fitting stationary time series, which considers the autoregression of financial market indicators and considers the interference of random fluctuations, making it possible to predict the short-term trend with a low error rate. Simultaneously, the ARMA model has been matured in theory, which is convenient for statistical analysis and mathematical processing [1].

In this article, we will use the ARMA model to fit the historical data of stock prices and predict the adjusted closing prices of the next five days so that we can make judgments on the future trend of the prices.

II. BASIC THEORIES OF ARMA MODEL

A. The Background of ARMA Model

The autoregressive-moving-average model, known as the ARMA model, is a kind of time series model, which was put forward by George E.P. Box and Gwilym Jenkins in 1970. This model describes a stationary stochastic process in terms of two polynomials, one for the autoregression (AR) and the other for the moving average (MA).

ARMA(p, q) is simply the merger between the two models--AR(p) and MA(q): AR(p) models are trying to capture the momentum and mean reversion effects often observed in trading markets, while MA(q) models are trying to capture the shock effects observed in the white noise terms. There are usually unexpected events that can be considered as shock effects, such as a sudden increase in earnings and a terrorist attack, etc.

B. The Way to Decide Values of p and q

To fit data to an ARMA model, Akaike Information Criterion (AIC) will be used across a subset of values for p, q to find the model with minimum AIC and then apply the Ljung-Box test to determine if we have achieved a good fit, for particular values of p, q. If the p value of the test is greater than the required significance, we can conclude that the residuals are independent and white noise.

C. Modeling Steps for ARMA Model

There are four steps when it comes to modeling with the ARMA model. First, we should perform a stationary test on the original data sequence. If the sequence is not stationary, we can use a different transformation to make the sequence satisfy the stationary condition. Then, we should decide the specific form of the model by calculating the statistics, such as autocorrelation coefficient and partial autocorrelation coefficient, which describe the characteristics of the sequence. At the same time, we can determine the order of the model according to the criteria such as AIC and SC. Besides, the least-squares method should be used to estimate the parameters of the model, and the reasonableness test should be conducted. Finally, we have to examine whether the model is consistent with the observed data to figure out if the model is practical for predicting.

III. AN EMPIRICAL ANALYSIS OF STOCK PRICES

As is known to all, Apple Inc has made huge progress in the renovation of Apple productions since 2017. Meanwhile, COVID-19 has spread worldwide and had a negative effect on every country's economy. Therefore, I have acquired the daily adjusted closing prices of the company Apple Inc from 2018 to the end of 2019, 502 samples in total, to make sure that these data are as stationary as possible. Then we will use the ARMA model to predict the adjusted closing price of Apple Inc's stock.

A. Smooth Process of Raw Data

We can use Python to draw the plot of the data sequence as below:

978-1-6654-1227-8/21/$31.00 ©2021 IEEE
DOI 10.1109/CBFD52659.2021.00046

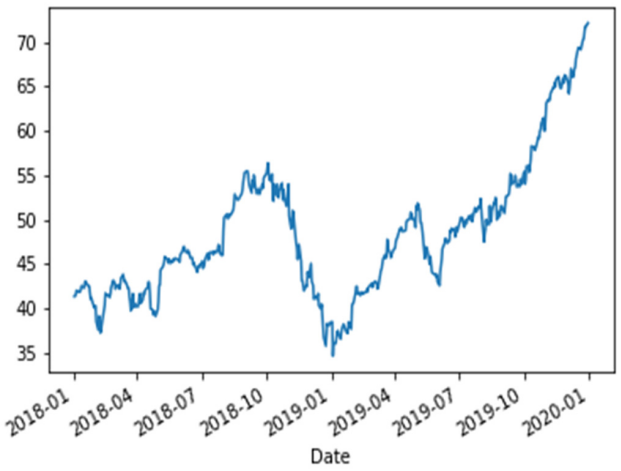

Figure 1. Stock prices of the considered asset

By observing the figure above, we can initially judge that the data sequence is not smooth. Then we need to further perform ADF inspection [2] on the original data, and the results are as below:

TABLE 1 THE RESULT OF ADF INSPECTION OF ORIGINAL DATA

	t-Statistic	p-value
ADF test statistic	0.6842	0.9895
1% level	-3.4435	
5% level	-2.8673	
10% level	-2.5699	

By calculation, we got the ADF test statistic. It is 0.6842, and the absolute value is greater than the critical value at 10% significance level, which means the null hypothesis that the data sequence has a unit root is not rejected. Therefore, the sequence is not stationary [3].

As the original sequence is not smooth, we will take the data's logarithm and perform first-order difference processing on the data to get a new sequence. Then ADF inspection will be performed again to examine if the new sequence is smooth [2]. The results of the processing are as below:

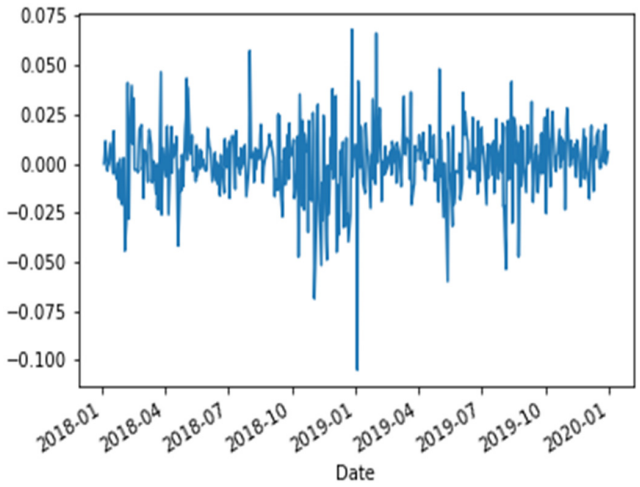

Figure 2. The stock price after taking the logarithm and making the first-order difference

TABLE 2 THE RESULT OF ADF INSPECTION OF PROCESSED DATA

	t-Statistic	p-value
ADF test statistic	-22.4638	0.0001
1% level	-3.4435	
5% level	-2.8673	
10% level	-2.5699	

By calculation, we got the new ADF test statistic. It is -22.4638, and the absolute value is smaller than the critical value at 1% significance level, which means the null hypothesis that the data sequence has a unit root can be rejected with a 99% probability. Up till now, we have finished the smooth processing of raw data.

B. Identify the Specific Form of the Model

After processing the data sequence's stationarity, we need to identify the model form by using ACF and PACF graphs. At the same time, information criteria will be used to figure out the lag order p and q.

Since economic variables are generally ARMA models with no more than 5th order, we can use Python to write a loop code to find out the values of p and q with the minimum AIC [4].

After calculation by Python, we picked the best order with the minimum AIC. Then we applied the Ljung-Box test to get the LB-value and p-value. Notice that the p-value is greater than 0.05, which indicates that the residuals are independent, known as white noise, at the 95% level, and thus an ARMA (4, 4) model provides a good model fit.

The results are as follows:

TABLE 3 THE BEST ORDER AND ITS RELATED INDICATORS

Best Order of (p,q)	Minimum AIC	LB-value	P-value
(4,4)	-2635.0341	13.3798	0.8605

Notice that the p-value is greater than 0.05, which indicates that the residuals are independent, known as white noise, at the 95% level, and thus an ARMA (4, 4) model provides a good model fit.

To better understand the ARMA model, we also drew the ACF graph, PACF graph, QQ plot, and Probability plot as below:

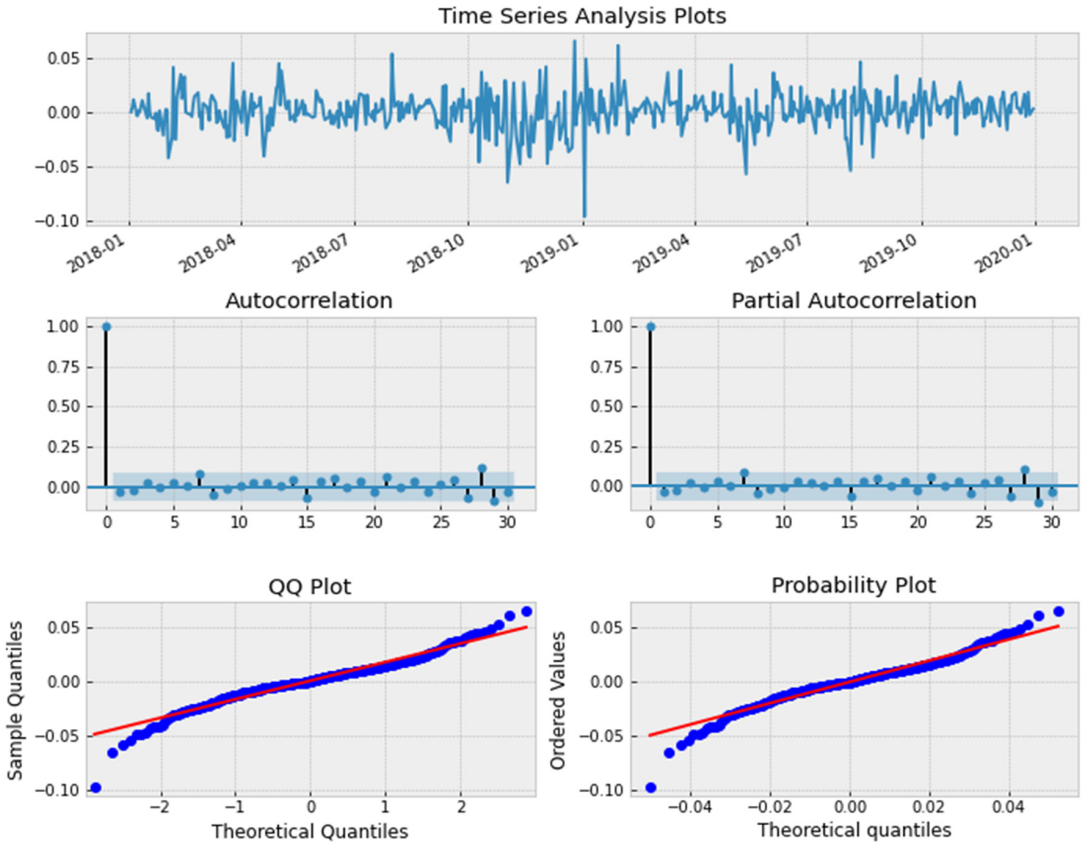

Figure 3.Several Plots of ARMA Analysis

C. *Forecasting Analysis of Adjusted Closing Prices*

We can use Python to calculate the coefficients of ARMA(3, 2): α1, α2, α3, α4 and β1, β2, β3, β4:

TABLE 4 Results of the Parameter

α_1	α_2	α_3	α_4	β_1	β_2	β_3	β_4
-0.3487	0.0042	0.3568	0.9875	0.3316	0.0044	-0.3403	-0.9936

The adjusted closing price in the next five days can be acquired according to the equation below:

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \alpha_3 x_{t-3} + \alpha_{t-4} x_{t-4} + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \beta_3 \varepsilon_{t-1} + \beta_4 \varepsilon_{t-4} \quad (1)$$

Then we will compare the results with the real data:

TABLE 5 The Comparison between the Predictive Value and Actual Value

	Predictive Value	Actual Value	Error Rate
2019-12-31	69.8176	72.6753	0.0393
2020-01-02	70.2525	74.3335	0.0549
2020-01-03	72.3808	73.6108	0.0167
2020-01-06	72.2909	74.1973	0.0257
2020-01-07	71.2543	73.8484	0.0351

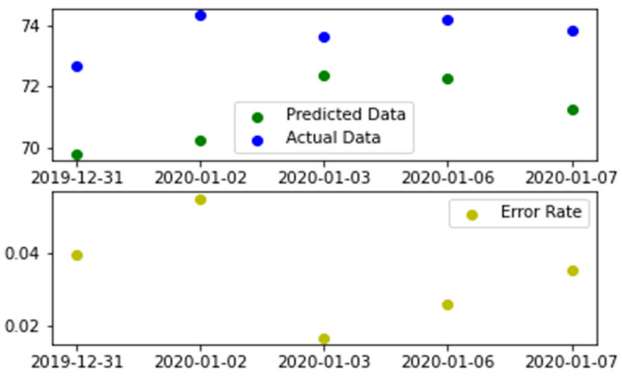

Figure 4. The graph of data and error rate

According to the table above, the adjusted closing prices predicted by the model are close to the real prices, and the errors are small, indicating that the ARMA (4, 4) model can be used to predict stock prices. However, due to the stocks' large volatility, this model is only suitable for short-term forecasting instead of long-term forecasting [5].

D. Backwards of ARMA Model

In the processing of the data, it will lead to certain errors when modeling the sequence due to the certain time lag and lack of data. Therefore, the model fitting in this paper is only a simple analysis of the attributes of the time series itself, without taking other factors into consideration [6].

E. Improvement of ARMA Model

In order to further improve the accuracy and effectiveness of the model, the ARMA-GARCH model can be used to further test it, or the CARR model can be used to predict stock volatility [7]. Comparing different models, the closest prediction value to the real one is the model with the smallest prediction error, which is the best model [8].

F. Suggestions for the Stock Markets

First, we should improve the information disclosure mechanism. There is a very typical herd effect in the stock market, and a large number of investment methods exist. That is to say, stockholders buy and sell stocks based on inside information collected from various channels. If the scale of relying on inside information is too large, it will cause great market volatility, which will lead to the loss of confidence by stockholders. Then the stockholders' loss of confidence in the market will inevitably affect the changes in household capital structure, and domestic demand will also be affected without stimulation. Therefore, a sound information disclosure mechanism can prevent the deterioration of the stock market environment and promote rational investment by investors [9].

Second, it is indispensable to strengthen the marketization of securities. As is known to all, the government's intervention in the securities market cannot be too great. Meanwhile, reducing the occurrence of speculative behavior is also necessary in the current actual situation. In addition, strengthening the rational allocation of resources, giving full play to the role of the market, and strengthening the construction of marketization play an important role in stabilizing the market.

Finally, it is of vital importance to strengthen rational investment education for investors. The majority of investors in the stock market are retail investors, which has led to the majority of stock market investments following the trend and short-term investments, virtually increasing market risks. Therefore, there is no doubt to educate investors on rational investment. When buying and selling securities, they must start based on the actual situation of the market, and they should not blindly follow the trend, only in which way can they cultivate their own investment ability and implement a wise investment [10].

IV. Conclusion

This article uses the ARMA model to empirically analyze stock prices and predict short-term stock prices. First of all, this article selects real sample data and carries out a stationarity test, that is, ADF inspection. When it is found that the original data is not stable, we perform the first-order difference data processing to make it stable. Then we determine the form of the model and also perform a residual test on it to prove the rationality of the model. Last but not least, we predict the adjusted closing prices in the next five days by using the ARMA model. It turns out that the predicted stock prices are close to the real prices with a low error rate, which means that the ARMA model can be applied to the predicting of the short-term financial time series. Therefore, the decision-makers can use this kind of model and write relative Python codes to suggest the final decision.

References

[1] C. Xu, H. Fang. Empirical Study on Stock Price Forecast by ARMA Model. 2019, (31): 77-82.

[2] Z. Xiao, Peter C.B. Phillips. An ADF coefficient test for a unit root in ARMA models of unknown order with empirical applications to the US economy. 1998, 1(2): 27-43.

[3] Menelaos Karanasos. Prediction in ARMA Models with GARCH in Mean Effects. Journal of Time Series Analysis, 2001, 22(5).

[4] H. Liu, Y. Wang. Research on the Volatility of Stock Market——An Empirical Analysis Based on ARMA-TGARCH-M Model. Journal of Beihang University (Social Science Edition), 2017, 30(04): 56- 66.

[5] Harsha Saxena, Ainapurapu Venkata Anurag, Nihal Chirayath, et al. Stock Prediction using ARMA. 2018, 8(2): 1-4.

[6] Z. Ma. The application of ARMA model in short-term stock forecasting[J]. Guangxi Quality Supervision Herald, 2020(05): 152-153.

[7] S. Xu, X. Liang. Research on stock price prediction based on ARIMA-GARCH model. Journal of Henan Institute of Education (Natural Science Edition), 2019, 28(04): 20-24.

[8] F. Javier Trivez, Beatriz Catalan. Detecting level shifts in ARMA-GARCH (1,1) Models. Journal of Applied Statistics, 2009, 36(6).

[9] Neha Chhabra Roy, Basava Rajesh Beera. Investment Risk Identification and its Assessment: A Study of Indian Pharmaceutical Sector. International Journal of Banking, Risk & Insurance, 2017.

[10] L. Wang. System Simulation Research on the Influence of High Frequency Trading on the Stability of Stock Market [D]. Nanjing University, 2020.