Prediction for Some Tech Stock Prices of U.S. Stock Market based on ARIMA Model

Zhuojin Wu*

College of Harbor, and Coastal Engineering (CHCE), Jimei University, Xiamen 361000, China

*Corresponding author: 202121512030@jmu.edu.cn

Abstract. Predicting the future price of stocks is a common practice in the financial industry. Individuals and organizations engage in stock price forecasting for investment decision-making, risk management, portfolio optimization and asset allocation. However, the predicting stock prices is challenging due to the complexity of financial markets, the influence of numerous variables, and the presence of random and unpredictable events, which all make it harder to predict. This paper uses predicted future stock price trends for technology companies in The United States based on ARIMA and breaks it down using APPLE as a specific example for a more in-depth discussion of ARIMA. The raw data is collected from 2001/5/11 to 2021/4/22. Then it predicted the stock prices for the next two weeks (working days) by ARIMA and made a fitted trend curve, and finally compared it with the real data from 2021/4/23 to 2021/5/12 to see if ARIMA's predictions were accurate. The result shows that prediction model is in the form of ARIMA (1,1,1), and the prediction of APPL from 2021/4/23 to 2021/5/12 is trending down, which matches the real trend.

Keywords: ARIMA; predicted value; tendency.

1. Introduction

The U.S. stock market, one of the oldest stock exchange markets in history, attracts a lot of investors to invest and trade. Companies, regardless of size, could and raise capital with the support of investment banks. There is no currency control in the United Stated, and the U.S. dollars are free to enter and exit the country. The policy also encourages foreign companies to participate in investment behavior, making the U.S. the financial capital of the world.

Abundant supply of funds, good liquidity: a huge number of funds of different sizes, institutions and individual investors find different investment objectives, for the stock market of The United Stated provides solid capital base, which makes the stock market trading actively. What's more, the U.S. trading method is flexible: there is no stopping practice, the stock can be bought and sold at any time.

Standard & Poor's provides investors with credit ratings, independent analytical research, and investment advisory services, including the S&P 500 Index, which serves as a benchmark for U.S. portfolio indexes. Since the index is calculated based on the prices of most common stocks in New York Stock Exchange, it can flexibly adjust for price changes caused by subscriptions to new equity, share dividends and stock splits, etc., and the index values are more precise and have good continuity, so it tends to have better representation than the Dow-Jones.

Markowitz introduced Modern Portfolio Theory in 1952 and the Modern Portfolio Theory proposes methods to quantify the relationship between expected returns and potential risks [1]. Then the curve which connects all the combination of expected returns and potential risks is available, so that we can get the highest R/E ratio with the help of efficient frontier. Depending on the tolerance for risk, investors will choose different portfolio on the efficient frontier. As time progressed, there were opposition and criticisms on Markowitz's theory. In Markowitz's theory, the correlation coefficient is always constant and invariably fixed. However, Chesnay and Jondeau introduced in their empirical study on correlation coefficients that the prices of stock market always positively correlated with time during the turbulent period [2]. Moreover, it means that the correlation of any two assets would deviate from their mean historical correlation coefficient. Also, Markowitz

discussed with Fabozzi and Gupta on his Modern Portfolio Theory later their collaborative paper in 2002 [3].

The high degree of freedom in the U.S. stock market also means that it has a high degree of risk, so trading strategies are very important. Historical stock price data is a good reference tool to help us predict the future price of a particular stock or class of stocks, which can greatly help our trading strategy and enable us to get more return on our investment in the stock market. Naturally, ARIMA and ARMA models can help us a lot in this regard. To correctly estimate the ARIMA model, non-stationarity must be identified and eliminated by differentiation. Therefore, the difference between the value and its lagging value needs to be calculated [4]. First time Kalman Filter was introduced by Monirujjaman [5]. Kalman Filter has made significant contributions to statistical applications of econometrics, forecasting of time series data, and other economic attributes [6]. Stoean provided practical examples of time series analysis using R, including ARIMA modeling. Yermal focused on Kalman filtering in his book, which covers some aspects of time series modeling. Li provided a comprehensive introduction to time series analysis and forecasting. What's more, Kunnsch discussed resampling methods in the paper [10].

This thesis primarily uses data from tech stocks because tech stocks have the following benefits. Companies like Adobe, IBM, Apple, Citrix Systems and NVIDIA, Cisco have consistently held the top spots in terms of market capitalization. As a result, movements in their stock prices can heavily influence the overall market indices like the S&P 500 and Nasdaq Composite.

2. Methods

2.1. Data Source

Apple is one of the world's most iconic and influential technology companies. Steve Jobs, Steve Wozniak, and Ronald Wayne founded it on April 1, 1976, which has played a pivotal role in shaping the modern tech landscape. The company is headquartered in Cupertino, California, and is known for its innovation in consumer electronics, software, and services. Apple's stock was primarily exposed to investors on the NASDAQ in the United States, but it also trades on other global exchanges. The company has a diversified product portfolio, which includes hardware, software, services, and wearables. This diversification can help it weather economic downturns and market fluctuation. Figure 1 shows the data and tendency of APPL from 2001/5/11-2021/5/12.

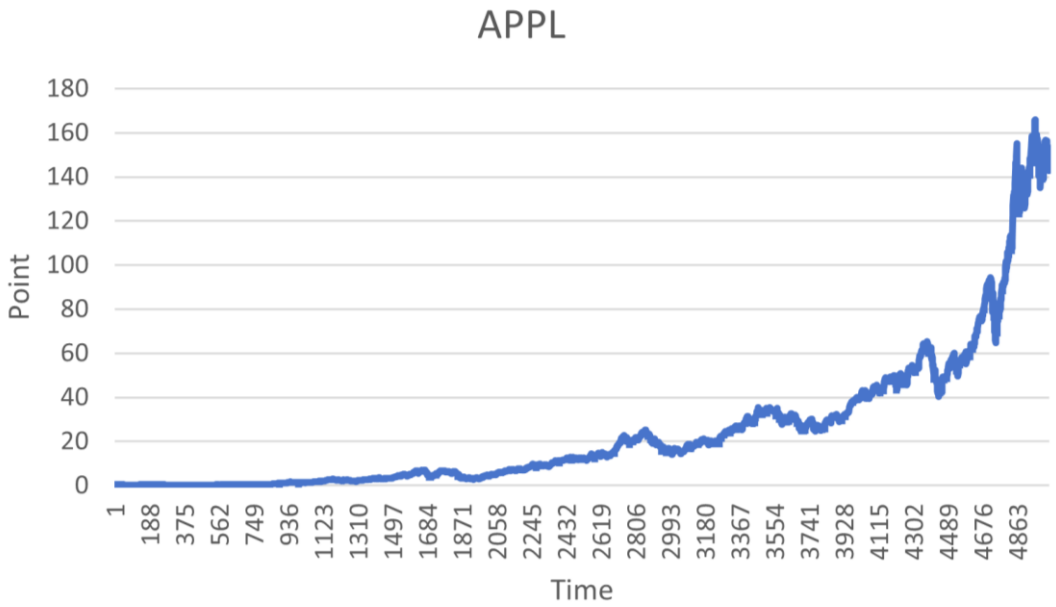

Fig. 1 Point of Stock APPL

2.2. Variable Description

This paper is about forecasting future stock prices using an ARIMA model, so the variables are only time T and stock price P. A stock price is the value of a single share of a company's stock. It represents the market's expectation of the company's financial health, potential for growth, and overall performance.

2.3. Method Introduction

Auto Regressive (AR) Component: It models the relationship between current and past observations.

$$y(t) = \phi_1 * y(t-1) + \phi_2 * y(t-2) + \cdots + \phi_p * y(t-p) + \varepsilon(t) \quad (1)$$

Integrated (I) Component: Differencing the series to make it stationary, removing trends.

$$y'(t) = y(t) - y(t-d) \quad (2)$$

Moving Average (MA) Component: Models the relationship between the observation and past errors.

$$y(t) = \theta^1 * \varepsilon(t-1) + \theta^2 * \varepsilon(t-2) + \ldots + \theta_p * \varepsilon(t-q) + \varepsilon(t) \quad (3)$$

Estimation techniques involve methods like maximum likelihood estimation to find the best-fitting parameters. The goal is to determine parameters that fit the data. The model can then be used for forecasting based on estimated parameters and past data.

3. Results and Discussion

3.1. Second Order Difference Sequence Diagram

Determining the appropriate differencing order (d) in an ARIMA model is a crucial step in time series forecasting. The differencing order helps make the time series data stationary, which is a requirement for ARIMA modeling.

Firstly, parameters need to be determined and the corresponding ARIMA (p, d, q) model should be built. The ADF test showed a statistic of 3.806, and it was greater than the 10% critical value of -2.567. Moreover, the original data had unit roots and were unstable. So, the raw data should be converted to log returns after first-order differencing. Then, the stability of the data set was tested again by ADF test, and the ADF test statistic was that the original data had a unit root and was unstable. Figure 1 shows the undifferenced tendency.

After the first-order difference test, the ADF test statistic is -12.692. It is less than the 1% threshold value of -3.432, and the original data series becomes a smooth set, so the data series is a first-order difference smooth series which means the parameter d=1. Figure 2 shows the tendency of the first differential data series and the image of the first-order difference is significantly more stable.

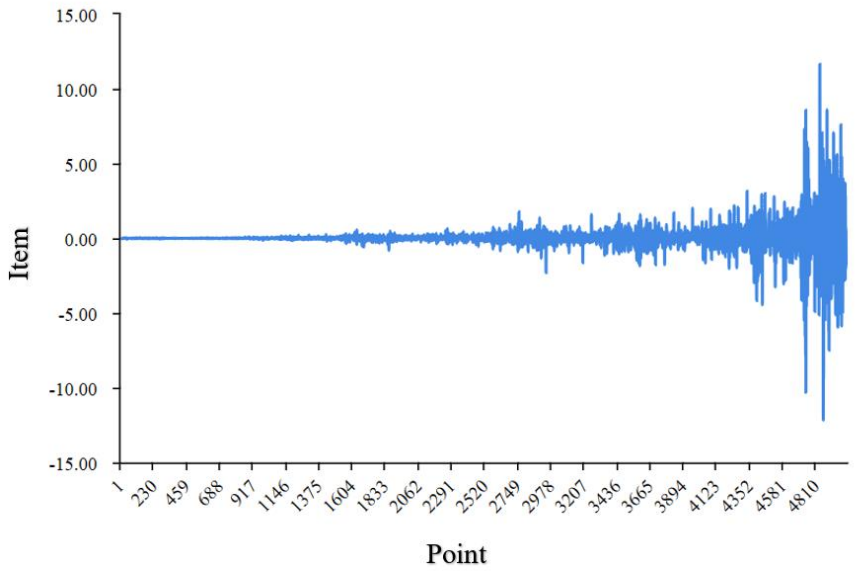

Fig. 2 First-order difference diagram

3.2. ACF and PACF Test

In ARMA model, the parameters p and q are determined with the help of the correlation and partial correlation plots. Since the data are first-order differential smooth series, the value of p and q can be determined from the autocorrelation coefficient (ACF) in Figure 3 and the partial correlation coefficient (PACF) plots in Figure 4 to be 1 and 1, respectively, thus forming the ARIMA (1,1,1) model.

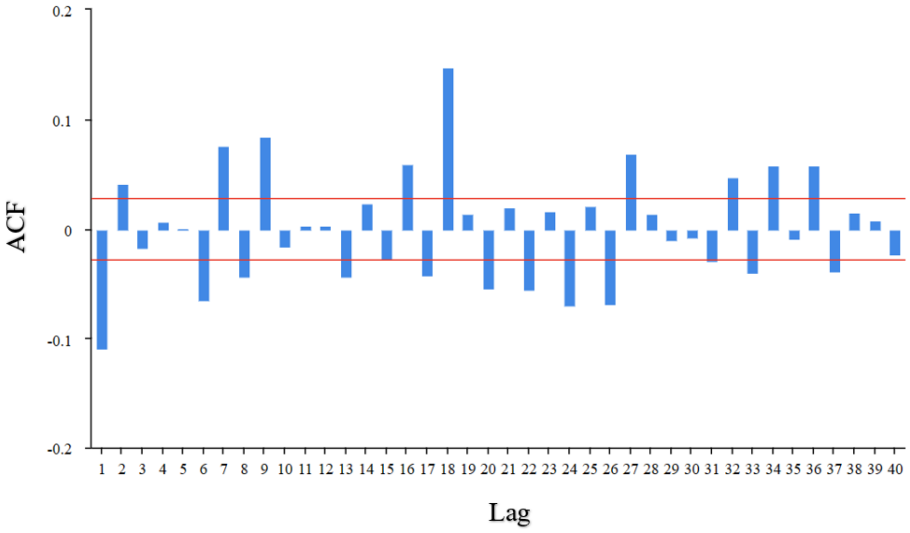

Fig. 3 ACF Chart

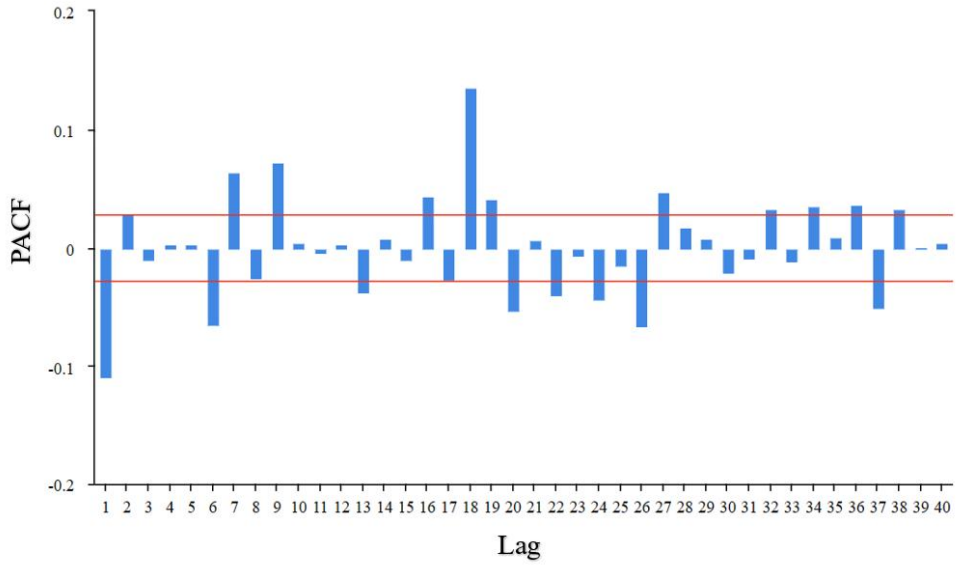

Fig. 4 PACF Chart

3.3. Predicted Results of ARIMA

This is an image fitted using an ARIMA model based on the true values, and the trend of the fitted values gives the predicted trend of the future data (Figure 5). The fitted data includes 14 days of unknown data that need to be predicted, and by using these predicted data, the trend of the stock price can be judged accurately.

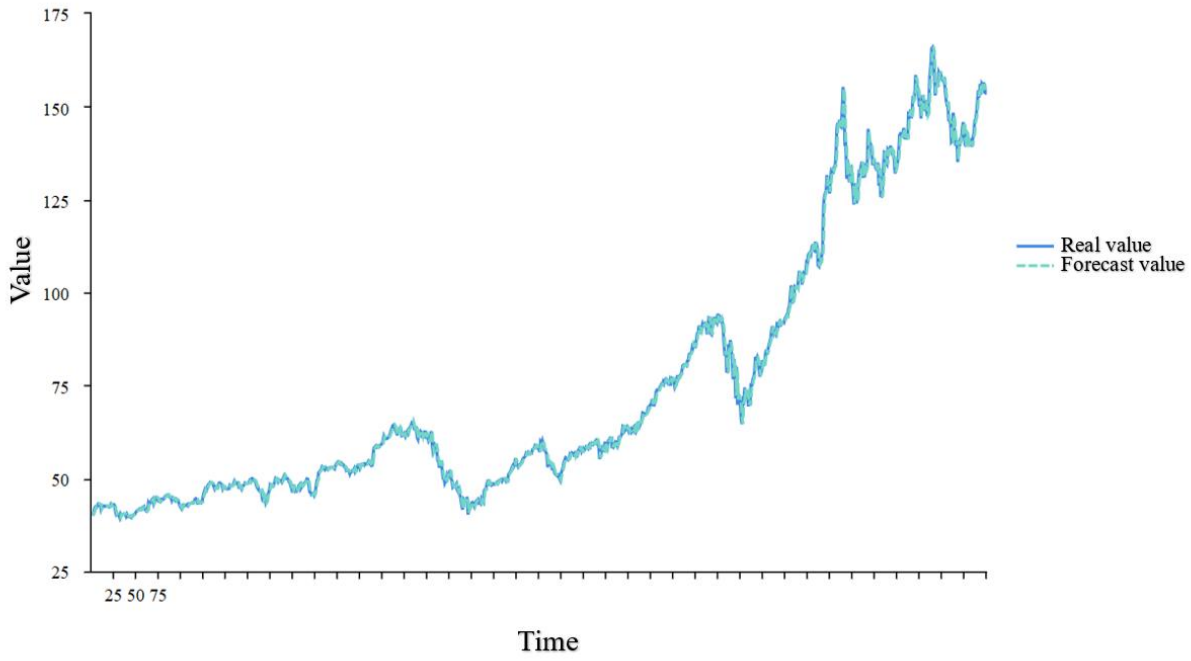

Fig. 5 Prediction and Fitting

3.4. Real Values and Predicted Value

Table 1 shows the ARIMA model's prediction of stock prices for the next two weekly business days (14 days) compared to the stock's true value from business day 2021/4/23 to business day 2021/5/12. Although the ARIMA model is not able to accurately predict the true price of the stock, it is able to predict a downward trend in the stock price for the next 14 days.

Table 1. Results and Comparison

Data	Real Value	Forecast Value
2021/4/23	155.95	153.45
2021/4/26	156.42	153.40
2021/4/27	156.04	153.46
2021/4/28	155.10	153.48
2021/4/29	154.98	153.51
2021/4/30	152.63	153.54
2021/5/3	153.89	153.57
2021/5/4	148.44	153.60
2021/5/5	148.73	153.63
2021/5/6	150.64	153.66
2021/5/7	151.44	153.70
2021/5/10	147.53	153.73
2021/5/11	146.44	153.76
2021/5/12	142.78	153.79

4. Conclusion

This paper proposes a method to predict the future price of a stock by time series. It is theoretically feasible to predict the trend for the next two weeks by looking at nearly 20 years of stock prices. At first, the raw data, which is not smooth and full of randomness, is first order differenced to obtain relatively smooth data. Then draw its PACF plot with ACF plot to determine the parameters p and q. Finally, an ARIMA model in the form of ARIMA (1,1,1) was confirmed and a downward trend was derived. However, some problems also came to the fore during the study. Firstly, the impact of the broader market, such as the spx500, is not considered, and fluctuations in the broader market can also have an impact on the technology sector, which needs to be analyzed in conjunction with the broader market data for this thesis. Moreover, the period of time for prediction is too short. Compared to a twenty-year time series, a two-week forecast trend can be minimal. If the prediction time is replaced with one month, then the predictive effect of the ARIMA model can be better demonstrated. Finally, it may work better if some of the excessively biased stock prices are removed.

References

[1] Markowitz H M. Portfolio selection. The Journal of Finance, 1952, 7: 77–91.

[2] Jondeau E, Chesnay F. Does correlation between stock returns really increase during turbulent periods? Economic Notes, 2001, 30: 53– 80.

[3] Markowitz H M, Fabozzi F J, Gupta F. The legacy of modern portfolio theory. The Journal of Investing, 2002, 11: 7–22.

[4] Vo N, Ślepaczuk R. Applying Hybrid ARIMA-SGARCH in Algorithmic Investment Strategies on S&P500 Index. Entropy, 2022, 24: 158.

[5] Monirujjaman Mohammad, Alam, Md Farabi, Ridoy, Shoumik Mahabub. Comparative Analysis of ARIMA and LSTMM achine Learning Algorithm for Stock PricePrediction. IEEE WORLD AI IOT CONGRESS (AIIOT), 2022, 460-465.

[6] Muhammad Aamir, Ani Shabri. Modelling and forecasting monthly crude oil price of Pakistan: A comparative study of ARIMA, GARCH and ARIMA Kalman model. AIP Conference Proceedings, 2016.

[7] Stoean R, Stoean C, Sandita A. Evolutionary Regressor Selection in ARIMA Model for Stock Price Time Series Forecasting, International decision technologies, 2017, 73: 117-126.

[8] Yermal L, Balasubramanian P. Application of Auto ARIMA Model for Forecasting Returns on Minute Wise Amalgamated Data in NSE. IEEE International Conference on Computational Intelligence and Computing Research, 2017, 745-749.

[9] Li ZJ, Liao YP, Lu YT. A Financial Deep Learning Framework: Predicting the Values of Financial Time Series With ARIMA and LSTM. International Journal of WEB Services Research, 2022, 19.

[10] Kunsch H R. Introduction to Random Signals and Applied Kalman Filtering, Annals of Statistics, 1989, 17(3), 1217-1241.