Online at www.e-iph.co.uk
Indexed in Clarivate Analytics WoS,

E-B Environment - Behaviour Proceedings Journal

ICLT2023Padang
https://iclt.uitm.edu.my/index.php/about/iclt-2023

e-IPH e-International Publishing House Ltd., United Kingdom

2nd International Conference on Logistics & Transportation 2023
Convention Hall, Universitas Andalas, Padang, Indonesia, 20 - 22 Nov 2023

Organised by: Research Nexus UiTM (ReNeU), Universiti Teknologi MARA

Forecasting Stock Prices of Three Industries using ARIMA and ARFIMA Model

Siti Aida Sheikh Hussin[1*], Muhammad Fakhri Syaabani[2], Harizatul Nadia Abdul Rahman[3], Zalina Zahid[4]
**Corresponding Author*

Faculty of Computer and Mathematical Sciences,
Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

sitiaida842@uitm.edu.my, muhdfakhrisyaabani198@gmail.com,harizatuln18@gmail.com,zalina@tmsk.uitm.edu.my
Tel:+60355435321

Abstract
Stock prices reflect a company's market value. This study aims to model the stock prices of three industries using the autoregressive integrated moving average (ARIMA) and the autoregressive fractionally integrated moving average (ARFIMA) models, and to forecast stock prices using the best model. Data for three companies were obtained from Yahoo Finance to represent the telecommunications, financial, and construction sectors. The forecast results show that all daily stock values were within 99% of the expected ranges for both models. According to the evaluation results, the ARFIMA model outperformed the ARIMA model.

Keywords: ARFIMA; ARIMA; stock prices; time series;

DOI: https://doi.org/10.21834/e-bpj.v10iSI41.7728

1.0 Introduction

The Singapore Stockbrokers Association registered as an institution in Malaya in 1930, marking the beginning of Bursa Malaysia. The development of the Malaysian capital market is significantly aided by Bursa Malaysia, which provides services and infrastructure that facilitate the adoption of globally recognised standards and the expansion of a competitive global marketplace.

The primary index and market indicator in Malaysia is the Kuala Lumpur Composite Index. Most academics conclude that a variety of interconnected environmental factors influence the variation in stock market returns and stock prices. Experts have previously determined that the Malaysian stock market is very responsive to the country's external environment. As a result, investors may predict or estimate the future performance of enterprises by analysing the level of surrounding activity in the nation.

In 2020, due to the impact of COVID-19 on global financial markets and Malaysia's economy, there was significant volatility in the Kuala Lumpur Composite Index. Global threats, such as COVID-19, are harming businesses and the nation's wealth. Investors and individuals are interested in the forecast value of share prices, especially during an economic crisis, to plan their portfolios and achieve optimal returns.

Time series forecasting and analysis help researchers use numerous models to analyse sequence data points that change over time and obtain meaningful statistics. The ARIMA model is one of the most widely used models for analysing time series data. The ARIMA model would estimate future stock prices based on prior periods or predict them based on prior performance. However, in some

eISSN: 2398-4287 © 2025. The Authors. Published for AMER by e-International Publishing House, Ltd., UK. This is an open access article under the CC BY-NC-ND license (http://c5eativecommons.org/licenses/by-nc-nd/4.0/). Peer–review under responsibility of AMER (Association of Malaysian Environment-Behaviour Researchers
DOI: https://doi.org/10.21834/e-bpj.v10iSI41.7728

cases, the time series data also exhibited long-memory correlation; in such situations, an alternative time series analysis method is the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model. The ARFIMA model can handle non-stationary data and has better forecast accuracy for long-memory data. This study will focus on three main sectors: telecommunication, construction, and finance. Yahoo Finance and the Bursa Malaysia website provided the data for the top-gaining organisation selected to represent the industry. Axiata Berhad, Gamuda, and Maybank are the companies (top gainers for each sector) that have been chosen.

This study aims to evaluate the impact of the COVID-19 pandemic on Malaysian stock prices. Yahoo Finance and the Bursa Malaysia website provided the data for the top-gaining organisation selected to represent the industry. Axiata Berhad, Gamuda, and Maybank are the companies (top gainers for each sector) that have been chosen.

The main objective is to model the stock prices using the autoregressive integrated moving average (ARIMA) and the autoregressive fractionally integrated moving average (ARFIMA) models. The evaluation of forecast accuracy was measured using mean absolute percentage error (MAPE) and root mean square error (RMSE). We forecast the future value of stock prices using the best model. The period chosen is during the COVID-19 pandemic. Although ARIMA and ARFIMA models are considered simple univariate models when compared to advanced machine learning and hybrid models, both models are chosen due to superior accuracy, the ability to handle both short and long memory and the results are easily interpreted and understood.

Lastly, it is hoped that by conducting this research, individuals and investors can maximise financial returns from buying stocks and mitigate risks, especially during bad economic conditions. The results will be beneficial to investors in planning and strategising their investments by identifying the best time to buy and sell stocks once the forecast values of stock prices are analysed analytically.

2.0 Literature Review

2.1 Construction Industry

The Malaysian construction industry contributes between 3 percent to 5 percent to the Malaysian economy. According to Mordor Intelligence, the Malaysian construction industry has been one of the sectors severely affected by COVID-19. Construction activity was halted across many parts of the country due to containment measures. In June 2020, the Construction Industry Development Board inspected 6,750 construction sites. Of these, 5,000 were not yet operational. From RM146.4 billion in 2019 to RM117.9 billion in 2020, the value of construction work decreased by 19.4%. The government allocates some budget as part of an initiative to grow the construction industry. GAMUDA Berhad, a prominent Malaysian infrastructure company that has established itself as a leader in the industry by consistently offering innovative and groundbreaking solutions on a global scale, has been chosen to represent the construction industry. GAMUDA has established several collaborative partnerships with other firms, including CIDB, Huawei, and Schneider Electric, to ensure the organisation's long-term sustainability.

2.2 Telecommunication Industry

The telecommunications industry in Malaysia has seen a significant transformation. The industry is projected to sustain growth through 2025. The expansion of industry is due to the growing urban population and the upgrade of mobile phones from 3G to 5G services. The telecommunications sector is projected to experience substantial growth in the coming years, driven by the expanding use of the Internet of Things (IoT) across industries, which enables connectivity via both wired and wireless networks. In 2018, the telecommunications sector recorded revenue growth of 3.7%, reaching RM35.86 billion. Mobile operators contributed 64% of this income. Axiata Berhad is well recognised as a prominent telecommunications corporation. Axiata Berhad, founded in Malaysia on June 12th, 1992, is regarded as one of the premier telecommunications companies in Asia. Axiata Group Berhad holds a significant ownership stake in several subsidiaries operating in diverse markets across South and Southeast Asia. Based on Axiata Berhad's 2015 annual report, the company reported revenue of RM19.9 billion. In 2020, Axiata's operational revenue reached around RM24 billion, marking a consistent upward trend since 2015.

2.3 Financial Industry

The financial services industry plays a crucial role as one of the main catalysts for the nation's economy. Considering the pandemic, Malaysia's business environment has undergone massive transformations as companies sought to reassess and enhance their corporate portfolios for recovery. The fundamental components of the financial services industry in Malaysia consist of commercial banks, investment banks, and Islamic banks. Maybank Berhad, the recipient of the prestigious Global Finance's World's Best Consumer Digital Banks in Asia Pacific award for two consecutive years in 2018 and 2019, has reported a total revenue of RM5.6 billion. This represents an 18.18% increase compared to the previous year. The net operating income increased by 2.8% to RM25.45 billion, in line with Malaysia's overall economic growth, estimated at 3.1% for 2021. Therefore, Maybank Berhad is chosen to represent the financial industry.

2.4 Forecasting Stock Prices

The forecast of stock prices is beneficial, especially for professionals involved in the finance department to manage their investments effectively. Stock price forecasting is a well-studied and crucial topic in both business and academic studies. Time series forecasting is widely recognised as the most common and indispensable approach for doing this study. The prediction of stock returns holds significant implications for market efficiency. Forecasting future fluctuations in stock prices has long been a challenge for scholars in the field.

Indeed, there is significant investor interest in research on stock price prediction. When predicting stock prices, the model considers the disparities within a sequence rather than directly assessing absolute values.

The autoregressive fractional integral moving average (ARFIMA) model is a type of time series model used to model time series data, extending the ARIMA model by allowing non-integer differencing parameters. ARFIMA can be categorised as a fractional order signal processing approach that generalises the conventional integer order models of both autoregressive integral moving average (ARIMA) and autoregressive moving average (ARMA) models. The ARFIMA (p,d,q) model belongs to the long memory model family that aims to achieve the goal of explicitly accounting for persistence in long-term correlations in the series.

2.5 Autoregressive Integrated Moving Average (ARIMA)

The autoregressive integrated moving average (ARIMA) model is among the most widely used statistically recognised process of time series model prediction tools. In an ARIMA model, it is postulated that the future values of a variable may be expressed as a linear combination of past values and current errors. The ARIMA model is predicated on the assumption of stationarity in the data. According to Chen et al. (2017), a stationary series can also be characterised as a white noise series or a series exhibiting cyclic activity. The ARIMA model is widely recognised as a crucial tool for generating precise forecasts across several domains. A manufacturing organisation uses an ARIMA model to inform and guide its business planning, procurement, and production objectives. The occurrence of forecast inaccuracies has the potential to result in significant disruptions within the company's supply chain and manufacturing operations. The use of precise forecasting techniques contributes to the reduction of expenditures and enhanced fulfillment of client expectations in a more effective manner.

Another field of study that uses the ARIMA model is the study by Taghreed Alghamdi et al. (2019). The goal of the study is to investigate the factors that have a substantial impact on traffic congestion rates using ARIMA-based modelling. This research investigates and analyses short-term time series over three months, considering the non-normality of the data distribution in the time domain, and tests the traffic data behavior for accurate prediction using various ARIMA models, also known as the Box-Jenkins techniques. The prediction performance of the research conventional ARIMA model may be examined in further depth using the findings obtained. While the ARIMA model is useful for time series forecasting, it has limitations. As linear models, they struggle with nonlinear relationships and sudden shifts, such as economic shocks. ARIMA is not suitable for long-term or real-time forecasting.

2.6 Autoregressive Fractionally Integrated Moving Average (ARFIMA)

The autoregressive fractional integral moving average (ARFIMA) model extends the usage of ARIMA models by allowing non-integer differencing parameter values. ARFIMA can be categorised as a fractional order signal processing approach that generalises the conventional integer order models of ARIMA. Among the past research on ARFIMA models is the study conducted by D. Safitri from Diponegoro University, focused on gold prices in Indonesia. The gold price data in Indonesia exhibits characteristics of long-term memory or time series data, indicating a persistent dependence across time. Research was done in 2021 by Nur Atiqah Ismail et al. The objective of the study was to compare the forecasting performance of ARIMA and ARFIMA models for predicting Kijang Emas (gold) prices in Malaysia. The evaluation of the models' forecasting accuracy was based on three metrics: mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE). The results show that the ARFIMA model is superior pe in predicting the prices of Kijang Emas in Malaysia than the ARIMA model. This is because of the lower values of MAE, RMSE, and MAPE.

2.7 Research gap

This research will extend insufficient long memory modelling during the crisis period (COVID) and forecast the future stock prices to evaluate if the results are aligned with previous research (non-crisis period). Furthermore, we would also like to investigate whether the best models are the same across the three industries chosen.

3.0 Methodology

3.1 Data

The data for three distinct businesses, specifically telecommunications (Axiata Berhad), finance (Maybank), and construction (Gamuda), has been obtained from Bursa Malaysia and Yahoo Finance. The dataset consists of daily stock prices spanning one year, specifically from 1st December 2021 to 1st December 2022. The data has been forecasted up to December 5th, 2022. R Studio was used for both models' estimates. The data undergoes a mathematical transformation in which the closing price is turned into a stock return price by computing the logarithmic difference in returns. The formula may be expressed in the following manner:

$$log\ log\ return{:}ln\ ln\ (b\ )\ \ -ln\ (a)\ (1)$$

a is the value of the close price at time t

b is the value of the close price at time t+1

3.2 ARIMA

In the ARIMA model, a variable's future value is a linear mixture of its previous values, represented as follows:

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \cdots + \Phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (2)$$

Where, Yt is the closing price,
Φ and θ is the coefficient of the model,
p q ε is the random error at time t
P and Q are the parameters chosen based on ACF and PACF.
P and q represent AR and MA, respectively, to model ARIMA(p,d,q).
ε is a t sequence of uncorrelated random variables with zero mean and constant variance.

3.3 ARFIMA
Granger, Joyeux, and Hoskings devised the Autoregressive Fractionally Integrated Moving Average (ARFIMA) model, which is suitable for time series data with long-term memory. The expression for ARFIMA (p, d, q) is as follows:

d Φ $(B)(1 - B)$ Y = θ (B)ε (3)

Where:
p t q t Φ (B) is an AR polynomial equation of order p
θ (B) is an MA polynomial equation of order q
$(1 - B)$ It is a fractional difference operator

3.4 Model Comparison
To get accurate predictions, we conducted a comparative analysis of the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) values for both models. This investigation is to identify the superior model, which is the model with the lowest RMSE and MAPE values.

4.0 Findings

4.1 Model
The equations representing three distinct businesses' stock prices, are as follows:

Maybank Berhad

Yt = -0.1554Y(t-1) -0.9856ε(t-1) (4)

AXIATA Berhad

Yt = -0.5693Y(t-1) -0.7221Y(t-2) -0.2332Y(t-3) -0.5677ε(t-1)+0.3617ε(t-2)-0.7940ε(t-3) (5)

GAMUDA Berhad

Yt = -1.3667Y(t-1) -1.0291Y(t-2) -0.1227(t-3) +0.324ε(t-1) -0.3241ε(t-2) -0.9999ε(t-3) (6)

4.2 Model Evaluation
This section aims to evaluate the predictive performance of the ARIMA and ARFIMA models, both of which successfully predicted stock prices within the 99% forecast intervals for the period from December 2nd, 2022, to December 6th, 2022. In the present investigation, a more robust prediction model would be suggested by reduced forecast errors, as evidenced by the values of RMSE and MAPE. Based on the findings shown in Table 1, it can be observed that the ARFIMA model consistently demonstrates superior accuracy compared to the ARIMA model across all firms. This is due to the ARFIMA model's capability in handling non-stationary data effectively. GAMUDA demonstrated a higher level of accuracy. The table presents the anticipated values of the ARFIMA (1, 0.03, 1) model, which is regarded as the optimal model for predicting the stock return of Gamuda. The value of the forecasted 5-day value for the three organisations based on the best representation of the ARFIMA model is tabulated in Tables 2, 3, and 4.

Table1. Model Comparison

COMPANY	MODEL	RMSE	MAPE
MAYBANK	ARIMA(1,1,1)	0.9952368	25.330
	ARFIMA(1,0.38,1)	0.0171446	13.998
AXIATA	ARIMA(3,1,3)	0.9575167	33.241
	ARFIMA(1,0.1,1)	0.01648723	15.614
GAMUDA	ARIMA(3,1,3)	0.9420125	45.221
	ARFIMA(1,0.3,1)	0.01652375	11.822

Table 2. Forecast for Maybank using ARFIMA (1, 0.38, 1)

Date	Point forecast	Low 99	High 99
2 December 2022	1.381594	-1.245540	4.008728
3 December 2022	1.289808	-1.740655	4.320271
4 December 2022	1.254888	-1.912703	4.422478
5 December 2022	1.234149	-2.009587	4.477884
6 December 2022	1.2184	-2.076829	4.513629

Table 3. Forecast for AXIATA using ARFIMA (1, 0.01, 1)

Date	Point forecast	Low 99	High 99
2 December 2022	0.7940277	-1.710029	3.298085
3 December 2022	0.6092386	-2.195867	3.414344
4 December 2022	0.5102245	-2.389802	3.410251
5 December 2022	0.4271317	-2.476851	3.391115
6 December 2022	0.4284121	-2.519315	3.376139

Table 4. Forecast for GAMUDA using ARFIMA (1,0.03,1)

Date	Point forecast	Low 99	High 99
2 December 2022	0.9940052	-1.519970	3.507980
3 December 2022	0.8090257	-2.082204	3.700256
4 December 2022	0.7408460	-2.218951	3.700643
5 December 2022	0.6781826	-2.338241	3.694606
6 December 2022	0.6318214	-2.419051	3.682693

The accuracy of ARFIMA in forecasting stock return data is found to be greater than that of ARIMA, as indicated by the respective accuracy (%) values for each model. The ability to disentangle time series data into its fundamental components (trend and noise), followed by the reconstruction of the separated elements. The output of each component from the ARFIMA model can enhance the accuracy of forecasting. On the given day, characterized by its unremarkable nature, the empirical evidence (in the form of the actual forecast) suggests that the sales volume of stock returns during the preceding five days did not display any noticeable trend. Therefore, to make optimal future investment decisions from stock prices, applying the ARFIMA model will be a good move.

5.0 Conclusion and Recommendations

Through the application and comparative analysis of the ARFIMA and ARIMA models, we can draw key conclusions. Firstly, forecasting should be applied to comprehend the pattern of stock prices for the three sectors. However, the existing models will no longer be suitable when unanticipated events, such as the COVID-19 pandemic, occur. Updating the models when new data becomes available is necessary. To examine the impact of abrupt occurrences on time series data for future research, we would suggest considering an intervention time series model such as Holt's model.

MAPE and RMSE results indicate that the ARFIMA model has an advantage in terms of accuracy. However, each model has its own strengths and weaknesses in different aspects, and the choice of model should be made based on specific needs and objectives.

Acknowledgements

The authors would like to thank other researchers, lecturers, and friends for their ideas and discussions. We extend our special gratitude to the College of Computing, Informatics and Mathematics, Universiti Teknologi MARA, for supporting the publication of this paper.

Paper Contribution to the Related Field of Study

Modelling real data using ARIMA and ARFIMA time series models to forecast stock price returns for three industries and evaluate models to identify the optimal forecast values for stock prices during crisis period (COVID). By conducting this research individuals and investors

can maximize financial returns from buying stocks and mitigate risks especially during bad economic conditions.

References

Alghamdi, T., Elgazzar, K., Bayoumi, M., Sharaf, T., & Shah, S. (2019). Forecasting traffic congestion using ARIMA Modeling. *15th International Wireless Communications & Mobile Computing Conference (IWCMC*),1227-1232. https://doi:10.1109/IWCMC.2019.8766698

Al Wadi, S., Sadam, J., Aladwan, M., & Bouqaleh, B. (2022). Forecasting earnings of firms listed in ASE using the ARIMA model. *Italian Journal of Pure and Applied Mathematics*.

Bao, G., Xie, Z., Bao, J., Guo, Y., Xie, W., & Zhang. (2020). Model for stock price trend prediction based on LSTM and GA. *Computer Science*, *47*, 467–473.

Chen, K., Zhou, Y., & Dai, F. A. (2015). LSTM-based method for stock returns prediction: A case study of the China stock market. In *the 2015 IEEE International Conference on Big Data*. IEEE.

Choudhary, R., & Gianey, H. K. (2017). Comprehensive review on supervised machine learning algorithms. In the *2017 International Conference on Machine Learning and Data Science (MLDS)* (pp. 37–43). IEEE.

Fu, Y., & Li, F. (2012). Application of ARMA model in sports stock price prediction in China. *Statistics and Decision*, *21*(3), 101–103.

Jessica, L., Keogh, E., Lonardi, S., & Chu, B. (2003). A symbolic representation of time series, with implications for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*. ACM.

Khashei, M., & Bijari, M. (2010). An artificial neural network (p,d,q) model for time series forecasting. *Expert Systems with Applications*, *37*(1), 479–489. https://doi.org/10.1016/j.eswa.2009.05.044

Lawrence, R. (1997). *Using neural networks to forecast stock market prices*. https://people.ok.ubc.ca/rlawrenc/research/Papers/nn.pdf

Nur, A., Azan, A., Faizatul, N., Mohd, A., Mototo, Z., Jin, P., & Mah, W. (2021). The comparison between ARIMA and ARFIMA model to forecast Kijang Emas (Gold) prices in Malaysia using MAE, RMSE, and MAPE highlights. *Journal of Computing Research and Innovation*, *6*(3), 22–33.

Safitri, D., Mustafid, Ispriyanti, D., & Sugito, (2019). Gold price modelling in Indonesia using ARFIMA method. *Journal of Physics: Conference Series*,1217.https://doi.org/10.1088/1742-6596/1217/1/012087

Shah, D., Campbell, W., & Zulkernine, F. A. (2018). Comparative study of LSTM and DNN for stock market forecasting. In *2018 IEEE International Conference on Big Data*. IEEE.

Yen, G., & Lee, C. F. (2008). Efficient market hypothesis (EMH): Past, present, and future. *Review of Pacific Basin Financial Markets and Policies*, *11*(2), 305–329. https://doi.org/10.1142/S0219091508001354