
South Asian Journal of Finance

ISSN (Online): 2719-2547 | Journal Home Page: https://sajf.sljol.info/

# Stock price prediction using ARIMA model: Evidence from Colombo Stock Exchange

**P. Tharshiga, T. Paranika, and V.A. Subramaniam**

**To cite this article:** Tharshiga P., Paranika, T. and Subramaniam V.A. (2026). Stock price prediction using ARIMA model: Evidence from Colombo Stock Exchange, *South Asian Journal of Finance*, *6*(1), 76–92.

SOUTH ASIAN JOURNAL OF FINANCE ISSN (Online): 2719-2547
2026, VOL. 6, No. 1, 76 – 92
Journal Homepage: https://sajf.sljol.info/
DOI: https://doi.org/10.4038/sajf.v6i1.142

# Stock price prediction using ARIMA model: Evidence from Colombo Stock Exchange

Tharshiga P. [a], Paranika, T. [b] and Subramaniam V.A. [c]

[a, b, c] Department of Financial Management, Faculty of Management Studies and Commerce, University of Jaffna, Sri Lanka.

**ABSTRACT**

Purpose: The core objective of the research is to investigate the forecasting capability of the Autoregressive Integrated Moving Average (ARIMA) model for predicting short-term stock prices in the Colombo Stock Exchange (CSE) of Sri Lanka.

Design/Methodology/Approach: The data gathered on a daily basis through the CSE Price Index, ranging between July 1, 2014, and June 30, 2024, was analyzed by using the Box-Jenkins approach. The selection of the optimum models was based on the minimum of Akaike Information Criterion and Schwarz Bayesian Criterion. The Autocorrelation Function, Augmented Dickey-Fuller Test, and error test measures, such as Mean Absolute Percentage Error, were considered for validation and for assessing the goodness of fit of the forecasting results.

Findings: From the Autoregressive Integrated Moving Average (ARIMA) model analysis, the ARIMA (2,1,1) model was the best, with an MAPE of 3.9%, indicating strong forecasting performance. For the Autoregressive and the Moving Regression tests, both were highly significant at the 1% level, supporting the idea that past price variations contain useful information for predicting prices. Results suggest that partial weak-form inefficiency exists in the Sri Lankan Stock Market.

Research limitations/ Future research directions: The study uses a univariate linear approach and does not account for exogenous variables, nonlinearity, or structural breaks. The findings of this approach would be more relevant to short-term linear predictability. The approach would not account for non-linear phenomena that could be prevalent in an emerging economy.

Originality: The current research is among the first 10-year empirical validations of the ARIMA model's predictive accuracy in the Sri Lankan market, as the study's results provide theoretical and practical insights into predictive modelling and market efficiency.

**KEYWORDS**

ARIMA model, stock price forecasting, Colombo Stock Exchange, time-series analysis, market efficiency

**JEL CLASSIFICATION**

G14, G17, C22

## I. Introduction

Stock market price forecasting has long been of interest to investors, governments, and academics alike, given its significance for risk and portfolio management, as well as decision-making. Financial sectors utilize these forecasts to manage systemic and credit risks, while governments rely on them for economic planning and stability. At present, the need for accurate forecasting is more prominent in the case of developing nations, including those in the South Asian region, which have clearly identifiable market patterns, high growth rates, and increasing globalization. The quest for making accurate stock market forecasts has resulted in the application of numerous predictive methods, which have had

CONTACT **a** P. Tharshiga tharshiga@univ.jfn.ac.lk Department of Financial Management, Faculty of Management Studies and Commerce, University of Jaffna, Sri Lanka. ORCID: 0000-0001-6738-3358 | **b** T. Paranika paranika003@gmail.com Department of Financial Management, Faculty of Management Studies and Commerce, University of Jaffna, Sri Lanka. ORCID: 0009-0009-0210-953X | **c** V.A. Subramaniam smani@univ.jfn.ac.lk Department of Financial Management, Faculty of Management Studies and Commerce, University of Jaffna, Sri Lanka. ORCID: 0009-0009-7788-9024

Received: 03 February 2026, Accepted revised version: 14 May 2026

differing levels of success (Tsai & Hsiao, 2010).

Time series analysis is one of the most popular methods used because of its ability to analyze past data and make predictions based on patterns established over time. One method that has proven highly effective and versatile for analyzing data that varies over time is the ARIMA model. However, its uses for emerging markets, especially those in South Asia, have not yet been investigated to the same degree of thoroughness as developed market economies. The financial markets of this region, which includes India, Sri Lanka, Pakistan, and Bangladesh, has grown significantly and has attracted both local and foreign investment. It should also be noted that the regulatory disparity, volatility, and partial information efficiency of financial markets make them even more difficult to forecast accurately. However, there is a lack of thorough study on the predictive efficacy of the ARIMA model even for these markets (Yadav & Khera, 2023).

Therefore, it is significant for researchers and investors to acknowledge the performance and limitations of ARIMA models in these volatile markets. The primary objective of the study is to analyze the predictive performance of ARIMA models, particularly for the Colombo Stock Exchange, by applying them to historical data from key indices and equities. The ability of the model to identify autocorrelations and patterns may provide a probability-based understanding of future price behavior by assisting in discovering phases of overreaction or mean reversion, which can be utilized for short-term achievements (Tsay, 2005). In accordance with the arguments of advocates, a major strength of ARIMA lies in its capability to identify recurrent patterns in financial time series, hence providing a quantitative edge in decision-making for investors. In addition, emerging markets become renowned for success and failure, causing volatility on a global scale during the height of the economic transfers and international trade (Puaschunder, 2025). In such circumstances, historical data may possess higher predictive value since markets assimilate new information instantly. This comparative inefficiency allows time-series models to capture short-term trends, which are imperceptible in more robust markets. Accordingly, this research analyses the usefulness of ARIMA in stock price forecasts in Sri Lanka. It contributes to the wider debate on the part of statistical models in financial forecasting.

The remainder of this paper is organised as follows. Section 2 presents the theoretical framework and reviews the relevant literature, leading to the development of the study's hypothesis. Section 3 describes the data, methodology, and the Box–Jenkins procedure applied for model identification, estimation, and diagnostic checking. Section 4 presents empirical results, including stationarity tests, lag selection, model estimation, diagnostic plots, and out-of-sample forecasts. Section 5 examines the findings in relation to existing literature and theoretical perspectives. Section 6 concludes the study by summarising the key findings, highlighting the practical and theoretical implications, acknowledging the limitations, and suggesting directions for future research.

## II. Literature Review and Hypotheses Development

According to the Efficient Market Hypothesis (EMH), as expressed by Fama (1970), security prices fully reflect all publicly available information, making it impractical to consistently generate abnormal profits. Based on the information set incorporated, the EMH differentiates among the weak, semi-strong, and strong forms. Under the weak form past prices do not contain predictive power, implying that historical data cannot be used for future price forecasts. Nevertheless, in emerging economies, several empirical studies have questioned this presupposition, discovering evidence of auto and delayed information comprehension. Such ineffectiveness creates opportunities for models such as ARIMA that capitalize on temporal dependence.

Another closely associated theory is the Random Walk Theory, primarily developed by Bachelier (1900) and standardized by Samuelson (1965), which declares that consecutive price changes are independent and random. Further, this theory emphasizes EMH, even though divergences from pure randomness are observed recurrently in practice. Grossman and Stiglitz (1980) demonstrated that perfectly competitive markets are impossible, as investors require incentives to gather information, resulting in unavoidable inefficiencies. This discussion was further extended by Lo (2004), who asserts that the Adaptive Market Hypothesis (AMH) suggests that efficiency evolves as investors learn and adjust to changes in the environment. AMH acknowledges changing periods of efficiency and inefficiency, a practice that is consistent with emerging markets like Sri Lanka. The assumption of rationality is further emphasized by Behavioral Finance Theory through the attribution of market anomalies to cognitive biases, including herding, overconfidence, and loss aversion (Barberis, 2003). Inclusively, these theoretical perspectives justify the use of ARIMA models: the EMH and Random Walk highlight randomness, whereas the AM and Random Walk assert that adaptive inefficiencies can be statistically depicted.

The ARIMA model, introduced by Box and Jenkins (1970), provides an organized approach to time-series forecasting that incorporates autoregressive (AR), differencing (I), and moving-average (MA) components. The Box–Jenkins methodology progresses through identification, parameter estimation, diagnostic checking, and forecasting. Another empirical study, conducted between 2020 and 2025, consistently acknowledges the usefulness of ARIMA while defining its limitations. As expressed by Samarawickrama and Pallegedara (2023) and Wijayaweera (2021), ARIMA models adequately predict short-term movements in the Colombo Stock Exchange; however, accuracy is compromised during periods of volatility. Linear dependencies are captured by ARIMA; however, it performs less efficiently in nonlinear contexts compared to neural networks (Ranasinghe, 2021). Similar findings were reported by Ahammad, Rahman, and Chowdhury (2024), who compared ARIMA and LSTM models across South Asia and found that ARIMA performed well during steady periods, but accuracy declined under unstable conditions. Moreover, a study by Shahid (2024) that analysed numerous South-Asian exchanges established that short-term fluctuations are captured by ARIMA models, suggesting partial predictability and inefficiency. Therefore, findings suggest that ARIMA remains a reliable benchmark for short-term predictions of emerging markets.

A focused review of CSE-based studies is particularly useful for contextualizing the present research. Ranasinghe (2021) compared ARIMA and neural network models for forecasting stock prices in the CSE and found that although ARIMA offered reliable short-term predictions, it was outperformed by neural networks during periods of heightened volatility due to its inability to capture nonlinear dynamics. Samarawickrama and Pallegedara (2023) applied an ARMA–GARCH hybrid framework to the CSE and found that incorporating volatility clustering through GARCH components substantially improved model fit compared to standard ARIMA specifications. Wijayaweera (2021) examined the impact of external events on CSE stock prices and employed time-series models to predict the duration and magnitude of price disruptions, noting that ARIMA-based predictions degraded sharply in the presence of structural breaks. Collectively, these CSE-focused studies suggests that while ARIMA is an established baseline model for the Sri Lankan market, its performance is context-dependent and diminishes under nonlinear or crisis conditions. The present study distinguishes itself from prior CSE research in several important respects: it employs a longer and more comprehensive dataset spanning a full decade (July 2014 to June 2024), thereby capturing multiple market cycles including the Easter Sunday attacks, the COVID-19 pandemic, and the 2022 economic crisis. Furthermore, unlike Ranasinghe (2021) and Samarawickrama and Pallegedara (2023), which focused on hybrid or volatility-adjusted extensions, this study rigorously evaluates the standalone predictive performance of the ARIMA model using standardised forecast accuracy metrics (RMSE, MAE, and MAPE), thus providing a clean benchmark assessment of ARIMA's utility for short-term forecasting in the CSE context.

Hybrid approaches that incorporate ARIMA with machine learning (ML) and deep learning (DL) techniques have become significant in marking nonlinearities. It has been demonstrated that traditional ARIMA is outperformed by ARIMA–MLP and ARIMA–RNN hybrids, particularly in volatile environments, by Zhang, Wang, and Chen (2022). Ghosh (2022) achieved sufficient accuracy, as validated by Box–Ljung tests. Dong (2022) established the reliability of ARIMA in forecasting long-term conditions for Ford Motor Company's shares using RMSE and MAPE criteria. It was demonstrated by Khanderwal and Mohanty (2021) that effective short-term predictions can be achieved using ARIMA across Indian sectoral indices. Mahadik, Vaghela, and Mhaisgawali (2021) expressed that there exists comparable performance between ARIMA and LSTM models when data quality is high. Ganesan and Kannan (2021) and Dhyani et al. (2020) substantiated the robustness of ARIMA for forecasting NIFTY50, while emphasizing its feasibility for investment decisions. Studies demonstrate ARIMA as a dependable and interpretable baseline forecasting tool for the financial markets globally.

Despite this significant body of evidence, numerous research gaps persist. Several studies concentrate on single indices or limited time spans, thereby restricts their generalizability. Furthermore, few studies include macroeconomic or behavioural variables, yet cross-market relative analyses within South Asia remain insufficient. Additionally, even though hybrid models enhance accuracy, they initiate computational complexity and diminish interpretability, which restricts their adoption in developing economies. Considering the recent financial volatility and evolving policy environment in Sri Lanka, it is both necessary and timely to reassess the predictive capacity of ARIMA. Integrating empirical and theoretical insights suggests that ARIMA models perform best where market efficiency is limited and short-term dependencies are present. Their predictive success derives from utilizing autocorrelation patterns, which diverge from pure randomness; nevertheless, their constraints become recognizable in nonlinear or structurally changing conditions. The coexistence of efficient and inefficient stages, as postulated by the AMH, suggests that predictability is conditional and it differs with market adaptation. These temporary inefficiencies, often generated by behavioural biases, can be captured by ARIMA. Hybrid ML-based models constantly produce higher accuracy, whereas ARIMA's transparency and moderate data requirements ensure its continuous relevance as a benchmark forecasting tool. For Sri Lanka and other South-Asian economies, analyzing the performance of ARIMA can reveal the degree of inefficiency and report evidence-based investment and policy strategies. Hence, the cumulative literature establishes the argument that South Asian markets, though progressively integrated, are not completely efficient and thus, reveal short-term predictability. ARIMA models, by modelling linear dependencies, can produce statistically significant forecasts, which challenge the strict version of the random-walk hypothesis and, at the same time, remain consistent with adaptive and behavioral viewpoints. Experimental testing of ARIMA's predictive validity among these markets will contribute to the understanding of how classical econometric models function in the face of evolving financial circumstances. Based on these theoretical and empirical foundations, the following hypothesis is proposed:

*$H_1$: The ARIMA model significantly predicts future stock prices in CSE.*

Evaluating this hypothesis within the Sri Lankan share market will provide experimental evidence of ARIMA's relevance in emerging economies and elucidate whether inefficiencies persist that permit precise short-term forecasting. The findings will enhance the theoretical discourse on adaptive efficiency while providing practical assistance for policymakers and investors.

## III. Methodology

Based on the theoretical and empirical insights presented, this study employs a quantitative financial econometric approach to evaluate the predictive ability of the ARIMA model for forecasting stock prices in the Sri Lankan share market. The methodological framework adopts the Box–Jenkins' (1970) procedure for time-series modelling, which is extensively utilized in financial forecasting for its ability to recognize temporal dependencies and produce short-term predictions based on historical price movements.

### *Data and Sampling*

The study utilizes daily closing price data from the Colombo Stock Exchange (CSE), spanning the period from July 2014 to June 2024. The data were obtained from the official CSE database and authenticated against the Bloomberg historical series to ensure consistency. A purposive sampling technique was employed, as Sri Lanka's market was chosen for its emerging market attributes and susceptibility to volatility during the past decade. This period acquires both stable and crisis phases, providing a comprehensive dataset for modelling and analyzing predictive accuracy.

### *3.2 ARIMA Model Specification*

The ARIMA (p, d, q) is a stochastic time series process that combines autoregressive (AR), differencing (I), and moving average (MA) components. The general model is expressed as:

$$Y_t^{(d)} = c + \phi_1 Y_{t-1}^{(d)} + \phi_2 Y_{t-2}^{(d)} + \cdots + \phi_p Y_{t-p}^{(d)} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (1)$$

where $Y_t^{(d)}$ represents the differenced series at time $t$, $c$ is a constant, $\phi_i$ and $\theta_j$ are the autoregressive and moving-average coefficients, respectively, and $\varepsilon_t$ is a white-noise error term with zero mean and constant variance.

### *Model Identification and Estimation*

Stationarity of the series is evaluated using the ADF test. If the process is non-stationary, first-order differencing is applied. Autocorrelation (ACF) and partial autocorrelation (PACF) plots are utilized to establish the order (p, d, q) combinations that are possible. The parameters are estimated utilizing the maximum likelihood method. Competitive ARIMA models are compared using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), with the models generating the lowest values demonstrating the best fit. Model adequacy is validated by ensuring that residuals approximate white noise, tested using the Ljung–Box Q statistic and residual correlograms. Heteroskedasticity and serial correlation are also examined to ensure that the model accurately captures the data's structure. The selected ARIMA model is employed to produce out-of-sample forecasts for short-term prediction.

### *Software and Analytical Tools*

All statistical analyses are conducted using EViews 13, software extensively employed for econometric and time-series applications. Graphical analysis, including ACF/PACF diagnostics and forecast plots, is generated directly through EViews. The model's stability and accuracy are analysed through in-sample and out-of-sample forecasting performance.

### *Forecast Evaluation*

Forecast accuracy is evaluated using three standard performance measures commonly applied in financial econometrics:

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t-\hat{y}_t)^2}\ , \quad MAE = \frac{1}{n}\sum_{t=1}^{n}|Y_t-\hat{y}_t|\ , \quad MAPE = \frac{100}{n}\sum_{t=1}^{n}\left|\frac{Y_t-\hat{y}_t}{Y_t}\right| \quad (2)$$

where $Y_t$ and $\hat{Y}_t$ denote the actual and predicted values, respectively, and n is the number of observations. Lower values of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) indicate higher predictive accuracy and a better-fitting model.

## IV. **Findings and Discussion**

### *Data Presentation*


**Figure 1.** Share price trends

Figure 1 illustrates the movement of the Colombo Stock Exchange Share Price Index from July 1st, 2014, to June 30th, 2024. In the beginning (2014-2019), the stock market had a relatively stable movement with low volatility. However, in May 2019, the share price index fell to its lowest level since 2014 (5,199.98 on May 15th), which had a significant impact, particularly following the terrorist attacks in April 2019. This had caused a severe impact on tourism and resulted in investors losing confidence, which led to draining the foreign exchange reservoirs. In 2019, towards the end of the year, a 2% tax cut was imposed by the government. This Fiscal expansion had peaked due to a fall in GDP, simultaneously making the share price index more volatile. In early 2020, Sri Lanka implemented a nationwide lockdown in response to the COVID-19 global pandemic, which restricted production, trade, and financial transactions. During this period, the price index attained its low point over the ten-year period (4,247.95). In the wake of the global economic recession triggered by the global pandemic, a strong market rebound was observed in 2021, primarily fueled by world liquidity injections, easy monetary policies, and a re-emergence of market optimism. The stock market reached an unprecedented level of 13,462.39 in January 2022, refelcting high levels of speculation and market expectations of a strong post-pandemic economic resurgence. However, this trend proved short-lived. The global energy crisis triggered by the Russia-Ukraine conflict exacerbated a rise in oil prices, resulting in cost-push inflation and higher import bills. The following flexible exchange rate policy in March 2022 triggered a sharp depreciation of the exchange rate and inflationary pressures, leading to a closing gap in inflation of nearly 30% by mid-2022. Furthermore, market performance in 2022 was temporarily hindered by higher levels of political instability and societal unrest. As there was a gradual improvement in macroeconomic management and administrative stability under a new governance system, the CSE also experienced a revival. Towards the end of 2023 and early 2024, a renewed market momentum emerged, pointing to a return to macroeconomic stability. The ten-year review specifically shows that Sri Lanka's stock market has been particularly sensitive to instances of political turmoil, macroeconomic policy shifts, and international economic disturbances. This further emphasizes the importance of applying short-term models of forecasting, like ARIMA, to adequately reflect price behavior cycles associated with shock effects in an emerging market.

### *Stationarity and Model Identification*

#### *Augmented Dickey–Fuller (ADF) Test for Stationarity*

The first step in the ARIMA model requirement is to determine whether the CSE Price Index is stationary. A stationary time series has a constant mean, variance, and autocovariance over time, which is a key assumption for accurate forecasting. For this purpose, an ADF test was conducted. The hypothesis of the test is as follows.

$H_0$ = *Series has a unit root*

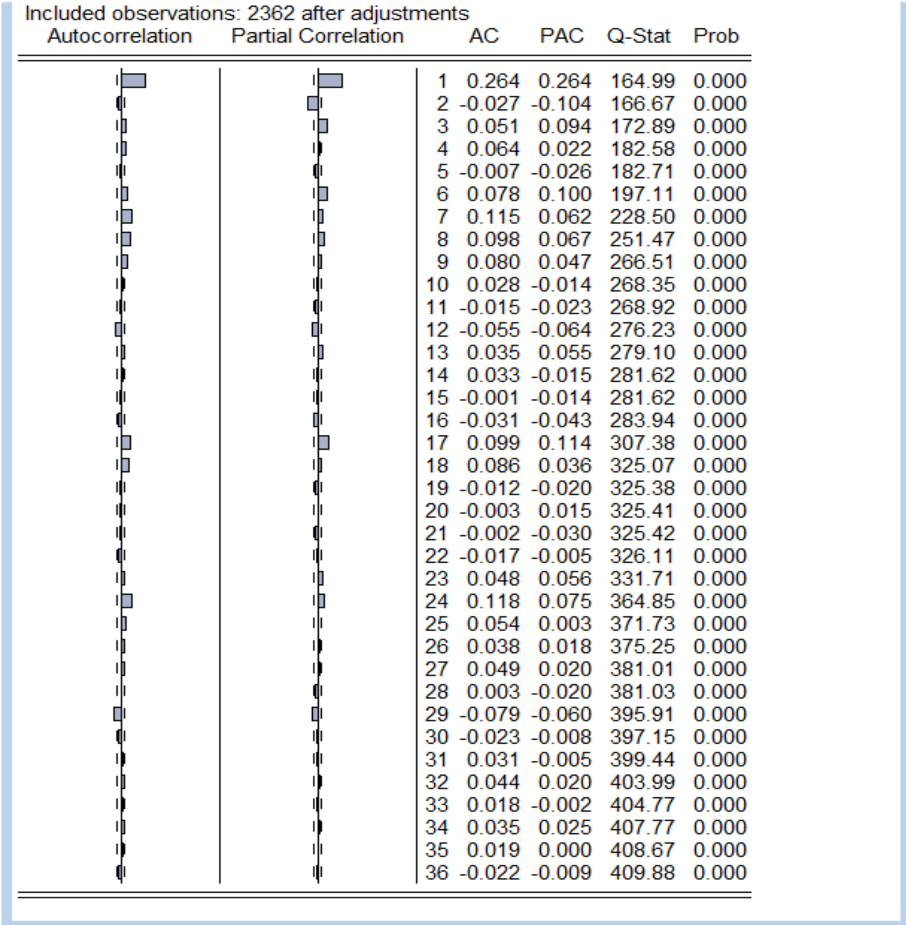
**Table 1.** Unit Root Test-ADF test

| Test Type | Model Specification | ADF t-Statistic | Probability | Result |
|---|---|---|---|---|
| Level | Intercept | –0.980153 | 0.7623 | Non-stationary |
| | Trend & Intercept | –2.052616 | 0.5714 | Non-stationary |
| | None | 0.528235 | 0.8300 | Non-stationary |
| First Difference | Intercept | –13.33849 | 0.0000 | Stationary |
| | Trend & Intercept | –13.37516 | 0.0000 | Stationary |
| | None | –13.31616 | 0.0000 | Stationary |

Table 1 presents the results of the ADF test conducted to evaluate the stationarity of the CSE Price Index. The test was executed under three model specifications: intercept only, trend and intercept, and without intercept or trend. At the level form, the ADF statistics (–0.980 for the intercept, –2.053 for the trend and intercept, and 0.528 with none) all have p-values greater than 0.05, indicating there is sufficient evidence to accept the null hypothesis of a unit root. This confirms that the initial series is non-stationary and not fitting for direct time-series modelling. After employing first differencing, the ADF test statistics (–13.338, –13.375, and –13.316 across the three cases) became highly significant, with a p-value less than 0.05, leading to the rejection of the null hypothesis of non-stationarity. These results confirm that the CSE Price Index is incorporated of order one, I(1). The conversion successfully removed trend components and stabilised the mean, satisfying the stationarity necessity for the ARIMA modelling framework.

### *Autocorrelation and Partial Autocorrelation Analysis (ACF and PACF)*

After the series became stationary, the ACF and PACF plots (Figure 2) were examined to identify the underlying data structure.



Included observations: 2362 after adjustments

| Lag | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|
| 1 | 0.264 | 0.264 | 164.99 | 0.000 |
| 2 | -0.027 | -0.104 | 166.67 | 0.000 |
| 3 | 0.051 | 0.094 | 172.89 | 0.000 |
| 4 | 0.064 | 0.022 | 182.58 | 0.000 |
| 5 | -0.007 | -0.026 | 182.71 | 0.000 |
| 6 | 0.078 | 0.100 | 197.11 | 0.000 |
| 7 | 0.115 | 0.062 | 228.50 | 0.000 |
| 8 | 0.098 | 0.067 | 251.47 | 0.000 |
| 9 | 0.080 | 0.047 | 266.51 | 0.000 |
| 10 | 0.028 | -0.014 | 268.35 | 0.000 |
| 11 | -0.015 | -0.023 | 268.92 | 0.000 |
| 12 | -0.055 | -0.064 | 276.23 | 0.000 |
| 13 | 0.035 | 0.055 | 279.10 | 0.000 |
| 14 | 0.033 | -0.015 | 281.62 | 0.000 |
| 15 | -0.001 | -0.014 | 281.62 | 0.000 |
| 16 | -0.031 | -0.043 | 283.94 | 0.000 |
| 17 | 0.099 | 0.114 | 307.38 | 0.000 |
| 18 | 0.086 | 0.036 | 325.07 | 0.000 |
| 19 | -0.012 | -0.020 | 325.38 | 0.000 |
| 20 | -0.003 | 0.015 | 325.41 | 0.000 |
| 21 | -0.002 | -0.030 | 325.42 | 0.000 |
| 22 | -0.017 | -0.005 | 326.11 | 0.000 |
| 23 | 0.048 | 0.056 | 331.71 | 0.000 |
| 24 | 0.118 | 0.075 | 364.85 | 0.000 |
| 25 | 0.054 | 0.003 | 371.73 | 0.000 |
| 26 | 0.038 | 0.018 | 375.25 | 0.000 |
| 27 | 0.049 | 0.020 | 381.01 | 0.000 |
| 28 | 0.003 | -0.020 | 381.03 | 0.000 |
| 29 | -0.079 | -0.060 | 395.91 | 0.000 |
| 30 | -0.023 | -0.008 | 397.15 | 0.000 |
| 31 | 0.031 | -0.005 | 399.44 | 0.000 |
| 32 | 0.044 | 0.020 | 403.99 | 0.000 |
| 33 | 0.018 | -0.002 | 404.77 | 0.000 |
| 34 | 0.035 | 0.025 | 407.77 | 0.000 |
| 35 | 0.019 | 0.000 | 408.67 | 0.000 |
| 36 | -0.022 | -0.009 | 409.88 | 0.000 |

**Figure 2.** ACF and PACF analysis

The ADF test revealed that the CSE Price Index was non-stationary in its level form but stationary after first differencing, indicating an integration of order one [I(1)]. The correlogram of the differenced series showed a sharp cut-off in the PACF after lag 2 and a gradual decay in the ACF, implying an AR (2) and MA (1) process, respectively. Based on these diagnostics and the minimum AIC and BIC values, the ARIMA(2, 1, 1) model was found as the most suitable specification for predicting stock prices in the CSE.

### *Selection of optimal lag*

**Table 2.** Test Results of ARMA (p, q)

| Test Results of ARMA (p, q) | | | | | | |
|---|---|---|---|---|---|---|
| (AR=p, MA=q) | Adjusted R-squared | AIC | SBIC | Sigma² (volatility) | S.E. of regression | Significant coefficients |
| (1,0) | 0.068995 | 11.66226 | 11.66959 | 6780.543 | 82.39640 | 1 |
| (0,1) | 0.081693 | 11.64854 | 11.65586 | 6688.059 | 81.83254 | 1 |
| (1,1) | 0.082519 | 11.64806 | 11.65783 | 6679.215 | 81.79575 | 2 |
| (1,2) | 0.081990 | 11.64864 | 11.65841 | 6683.065 | 81.81933 | 2 |
| **(2,1)** | **0.082532** | **11.64805** | **11.65782** | **6679.119** | **81.79517** | **2** |
| (2,2) | 0.000350 | 11.73379 | 11.74356 | 7277.398 | 85.37999 | 2 |
| (0,2) | -0.000218 | 11.73394 | 11.74126 | 7284.626 | 85.40428 | 1 |
| (2,0) | -0.000136 | 11.73386 | 11.74118 | 7284.025 | 85.40075 | 1 |

Table 2 presents the comparison of alternative ARMA (p, q) specifications for the first-differenced share price index. Among the models, the ARMA(2,1) model yields the lowest AIC (11.64805) and SBIC (11.65782), alongside the highest adjusted $R^2$ (0.082532) and the lowest standard error (0.81). These indicators mutually verify that the ARIMA(2,1,1) model presents the best overall fit to the data, effectively depicting short-term dependencies and volatility patterns. The model is parsimonious, statistically sensible, and consistent with the Box–Jenkins principle of minimizing information criteria while maximizing explanatory strength.

### *Estimation*

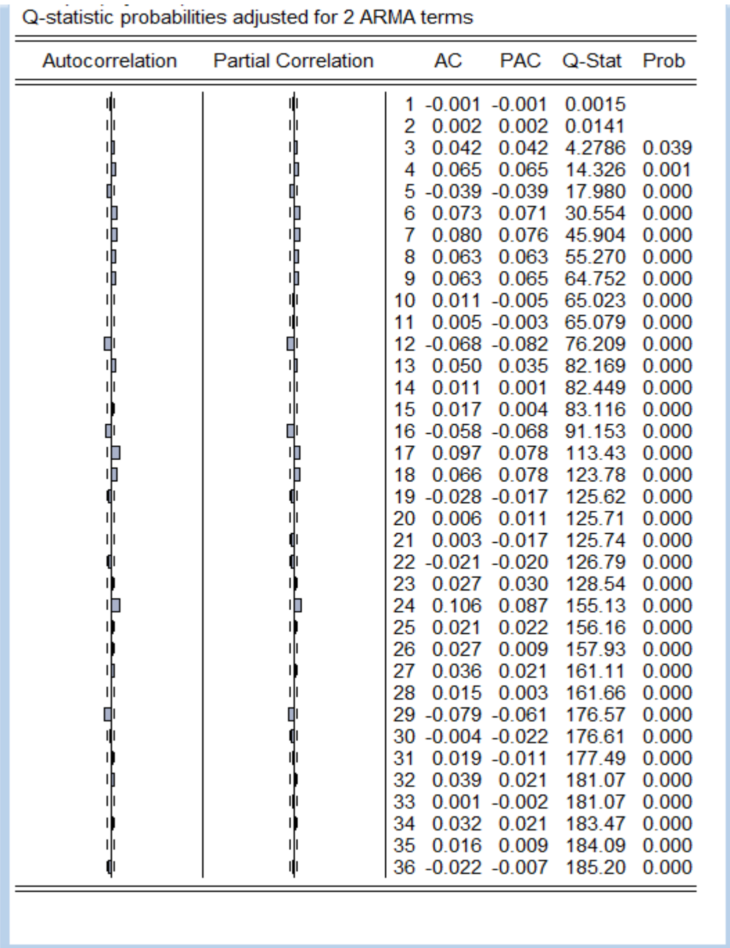
**Table 3.** Estimation Results of the ARIMA (2,1,1) Model

| Variable | Coefficient | Std. Error | t-Statistic | Probability |
|---|---|---|---|---|
| C | 2.445160 | 2.279450 | 1.072697 | 0.2835 |
| AR(2) | –0.038071 | 0.009110 | –4.178964 | 0.0000 |
| MA(1) | 0.301976 | 0.007985 | 37.81713 | 0.0000 |
| SIGMASQ | 6679.119 | 72.13156 | 92.59635 | 0.0000 |
| Model Diagnostics | | | | Value |
| R-squared | | | | 0.083698 |
| Adjusted R-squared | | | | 0.082532 |
| S.E. of regression | | | | 81.79517 |
| Sum squared residuals | | | | 15,776,079 |
| Akaike information criterion (AIC) | | | | 11.64805 |
| Schwarz criterion (SBIC) | | | | 11.65782 |
| Hannan-Quinn criterion | | | | 11.65161 |
| Log likelihood | | | | –13,752.35 |
| F-statistic | | | | 71.79573 |
| Prob(F-statistic) | | | | 0.000000 |
| Durbin–Watson statistic | | | | 2.001405 |
| Inverted AR Roots | | | | –0.00 ± 0.20i |
| Inverted MA Roots | | | | –0.30 |

In the case of the CSE Price Index, using the ARIMA(2,1,1) model, the results shown in Table 3 indicate that the AR(2) term is negative and statistically significant at the 0.05 level ($t = -4.179$, $p < 0.01$), indicating that there is an inverse and moderate relationship between the current price and the previous two periods. The result further depicts that the MA (1) term is statistically significant and positively correlated (MA(1) = 37.817, $p < 0.01$), suggesting that fluctuations in previous periods have an immediate impact on stock prices. The result further indicates that the constant term is not statistically significant ($p = 0.2835$, $p > 0.05$), suggesting that the series is not changing, and therefore, the series has an assumed mean of approximately zero. The result illustrates that the model has an acceptable fit and is highly statistically significant, since

$R^2 = 0.0837$ and the F-statistic ($p < 0.001$). This result, however, indicates that the variables influence stock prices. This indicates the presence of no residual serial correlation for the Durbin-Watson statistic value of 2.00. The inverse roots of the AR and MA parts of the polynomial lie inside the unit circle, thereby confirming the model to be stable and invertible. Additionally, the information criteria AIC and SBIC confirm an optimal trade-off between fit and simplicity for the given specification.

### *Diagnostic Checking*


Q-statistic probabilities adjusted for 2 ARMA terms

| Lag | AC | PAC | Q-Stat | Prob |
|---|---|---|---|---|
| 1 | -0.001 | -0.001 | 0.0015 | |
| 2 | 0.002 | 0.002 | 0.0141 | |
| 3 | 0.042 | 0.042 | 4.2786 | 0.039 |
| 4 | 0.065 | 0.065 | 14.326 | 0.001 |
| 5 | -0.039 | -0.039 | 17.980 | 0.000 |
| 6 | 0.073 | 0.071 | 30.554 | 0.000 |
| 7 | 0.080 | 0.076 | 45.904 | 0.000 |
| 8 | 0.063 | 0.063 | 55.270 | 0.000 |
| 9 | 0.063 | 0.065 | 64.752 | 0.000 |
| 10 | 0.011 | -0.005 | 65.023 | 0.000 |
| 11 | 0.005 | -0.003 | 65.079 | 0.000 |
| 12 | -0.068 | -0.082 | 76.209 | 0.000 |
| 13 | 0.050 | 0.035 | 82.169 | 0.000 |
| 14 | 0.011 | 0.001 | 82.449 | 0.000 |
| 15 | 0.017 | 0.004 | 83.116 | 0.000 |
| 16 | -0.058 | -0.068 | 91.153 | 0.000 |
| 17 | 0.097 | 0.078 | 113.43 | 0.000 |
| 18 | 0.066 | 0.078 | 123.78 | 0.000 |
| 19 | -0.028 | -0.017 | 125.62 | 0.000 |
| 20 | 0.006 | 0.011 | 125.71 | 0.000 |
| 21 | 0.003 | -0.017 | 125.74 | 0.000 |
| 22 | -0.021 | -0.020 | 126.79 | 0.000 |
| 23 | 0.027 | 0.030 | 128.54 | 0.000 |
| 24 | 0.106 | 0.087 | 155.13 | 0.000 |
| 25 | 0.021 | 0.022 | 156.16 | 0.000 |
| 26 | 0.027 | 0.009 | 157.93 | 0.000 |
| 27 | 0.036 | 0.021 | 161.11 | 0.000 |
| 28 | 0.015 | 0.003 | 161.66 | 0.000 |
| 29 | -0.079 | -0.061 | 176.57 | 0.000 |
| 30 | -0.004 | -0.022 | 176.61 | 0.000 |
| 31 | 0.019 | -0.011 | 177.49 | 0.000 |
| 32 | 0.039 | 0.021 | 181.07 | 0.000 |
| 33 | 0.001 | -0.002 | 181.07 | 0.000 |
| 34 | 0.032 | 0.021 | 183.47 | 0.000 |
| 35 | 0.016 | 0.009 | 184.09 | 0.000 |
| 36 | -0.022 | -0.007 | 185.20 | 0.000 |



**Figure 3.** Diagnostic Checking

Figure 3 shows that the residual correlogram of the ARIMA(2,1,1) model demonstrates that all AR and PAC coefficients were within the ±1.96/√n confidence interval, indicating that the residuals are approximately uncorrelated. The Ljung–Box Q statistics are insignificant at the early lags, indicating no serial correlation in the residuals. This means the residuals perform as white noise, meeting the ultimate diagnostic criterion of the Box–Jenkins methodology. Therefore, the model is statistically valid and reliable for forecasting the CSE Price Index.

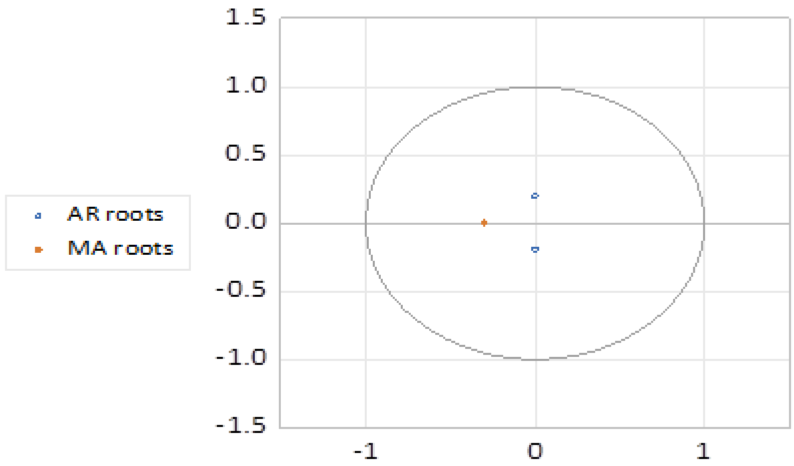



**Figure 4.** Inverse roots of AR/MA polynomials

Figure 4 depicts that all reversed AR and MA roots lie inside the unit circle, proving that the anticipated ARIMA (2, 1, 1) model is both stable and invertible, thereby fulfilling the Box–Jenkins conditions for reliable forecasting

The ARIMA (2,1,1) model is applied to fit the D (CSE_PRICE) data, and the results are displayed in Figure 5. The solid orange line represents the actual data, while the green line and blue line respectively show the model's fitted values and residual values.

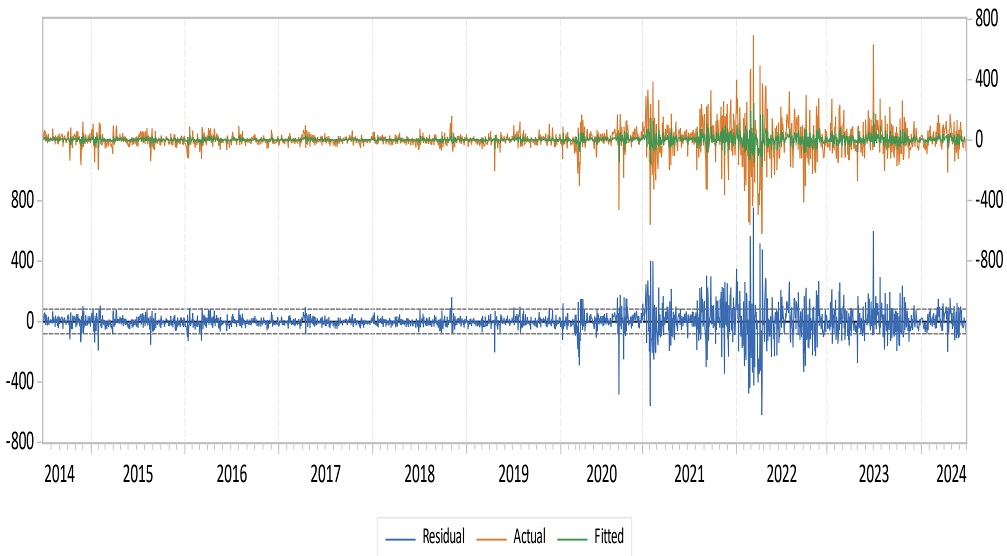



**Figure 5.** Graphical representation for the Actual, Fitted, and Residual series of the D(CSE PRICE)

### *Forecasting*

As model testing and confirmation of its applicability have been completed, it is now possible to use the ARIMA(2,1,1) model for forecasting. As the ARIMA(2,1,1) model has been estimated, it is now relevant to create forecasts of prices for the next three months using the dynamic forecasting method in the EViews 13 graphical interface from July 1, 2024, to September 30, 2024. Results for the one-month forecast are given in Table 4, and the forecast graph is depicted in Figure 6.

**Table 4. Sample empirical Results of ARIMA (2, 1, 1) of the CSE Price Index in Sri Lanka**

| Date | Predicted Price | Actual Price | Relative error% |
|---|---|---|---|
| 7/1/2024 | 12175.65339 | 12,076.49 | -0.82 |
| 7/2/2024 | 12178.89337 | 12,118.29 | -0.50 |
| 7/3/2024 | 12181.23001 | 12,037.14 | -1.20 |
| 7/4/2024 | 12183.64491 | 11,926.59 | -2.16 |
| 7/5/2024 | 12186.0942 | 11,947.43 | -2.00 |
| 7/8/2024 | 12188.54051 | 11,841.86 | -2.93 |
| 7/9/2024 | 12190.98552 | 11,785.23 | -3.44 |
| 7/10/2024 | 12193.43063 | 11,811.50 | -3.23 |
| 7/11/2024 | 12195.8758 | 11,802.26 | -3.34 |
| 7/12/2024 | 12198.32096 | 11,843.97 | -2.99 |
| 7/15/2024 | 12200.76612 | 11,892.15 | -2.60 |
| 7/16/2024 | 12203.21128 | 11,878.92 | -2.73 |
| 7/17/2024 | 12205.65644 | 11,830.48 | -3.17 |
| 7/18/2024 | 12208.1016 | 11,768.45 | -3.74 |
| 7/19/2024 | 12210.54676 | 11,727.08 | -4.12 |
| 7/22/2024 | 12212.99192 | 11,539.83 | -5.83 |
| 7/23/2024 | 12215.43708 | 11,597.19 | -5.33 |
| 7/24/2024 | 12217.88224 | 11,701.64 | -4.41 |
| 7/25/2024 | 12220.3274 | 11,640.55 | -4.98 |
| 7/26/2024 | 12222.77256 | 11,633.38 | -5.07 |
| 7/29/2024 | 12225.21772 | 11,619.01 | -5.22 |
| 7/30/2024 | 12227.66288 | 11,510.76 | -6.23 |
| 7/31/2024 | 12230.10804 | 11,406.56 | -7.22 |

Table 4 presents the forecasted and actual prices of CSE in July 2024, along with their respective relative error percentages. Actual prices are collected from their official website. The forecasted prices are relatively close to the actual prices, with relative errors ranging from -0.50% to -7.22%. In the early days of July, the forecast error margin is smaller, and as time passes, it continues to increase until it reaches its maximum on July 31, 2024, at -7.22%. This highlights that the ARIMA model is more reliable for short-term forecast accuracy, and its accuracy decreases as it approaches a more distant date.

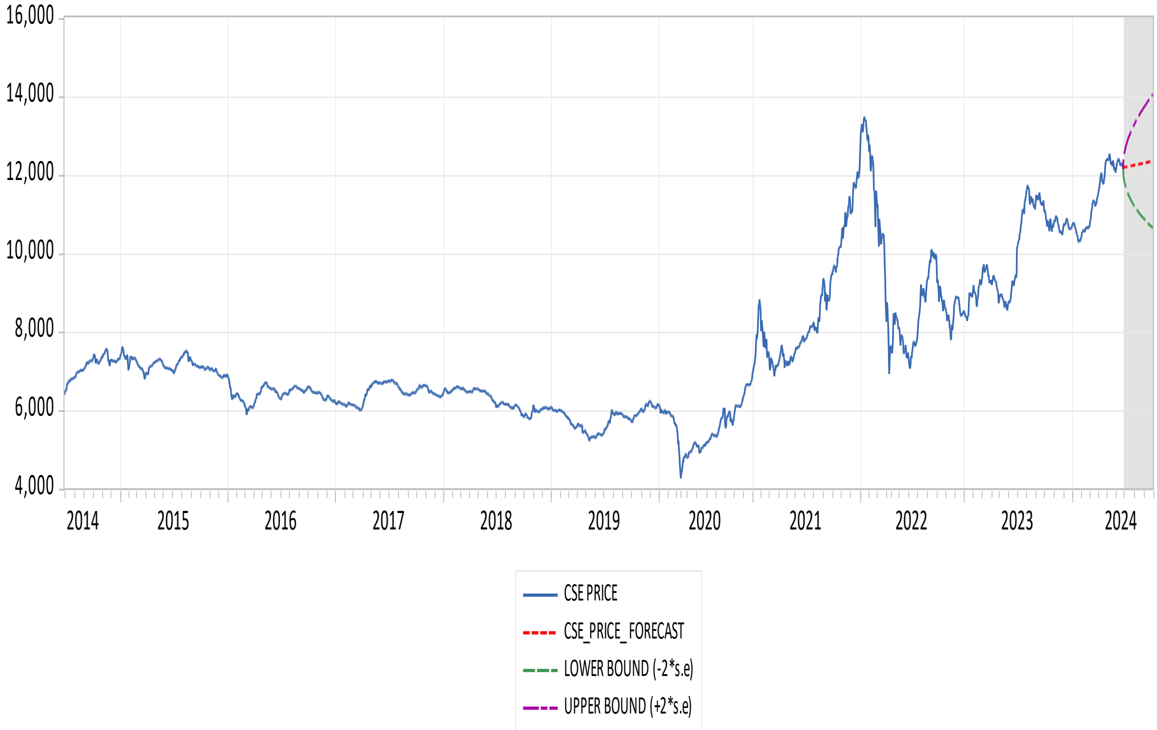



**Figure 6.** Graphical representation of the CSE Price Index for the next three months

The graph depicts the forecast points for the CSE Price, showing the observed values (blue solid line) from 2014 through 2024 and the forecast values (red dotted lines) beyond this period. As demonstrated from the above graph, there is a certain degree of positive projection in this series, as revealed by the ever-rising red-colored dotted lines. The region between the green (lower) and purple (upper) boundary lines represents a 95% confidence level, specifying the region in which values are expected with 95% confidence. Thus, this graph indicates a degree of positive stability in the CSE Price.

## V. Findings and Discussions

This study examined the impact of dividend. The research employed the ARIMA(2,1,1) model to predict short-term fluctuations in the CSE Price Index. The statistical results provided strong support for the hypothesis. The AR(2) and MA(1) terms are highly significant at the 1% significance level ($p < 0.01$), indicating that historically recorded price changes are well-modelled, which is appropriately captured by the proposed ARIMA model. The adjusted $R^2$ of 0.0825, though small, is to be anticipated in financial time-series analysis, where stock market changes are governed by multiple stochastic processes. It can be interpreted to mean that about 8% of the short-term variation in the CSE Price Index is explained by its own past observations, and the remaining part is characterized by random market activity. The results align well with existing literature, which states that financial models possess low explanatory but significant predictive powers for forecasting financial patterns, especially in turbulent financial markets, as discussed by Tsay (2005) and Zhang (2022). Moreover, the global F-test of 71.79573, with a p-value less than 0.001, supports the joint significance of the parameters. The Durbin-Watson statistic of 2.001 suggests the absence of serial correlation, thereby verifying the goodness of fit of the proposed model. The ACF and PACF of residuals show that all the Ljung-Box tests exceed the upper and lower bounds of 95% confidence limits, which proves that the residuals are a realization of a White Noise Process. The results confirm that the proposed model is statistically valid and reliable for predicting future stock market changes. The results demonstrate significant support for rejecting the null hypothesis of no predictive capability and accepting the alternative hypothesis ($H_1$).

From a theoretical perspective, these findings are consistent with the Adaptive Market Hypothesis (AMH) (Lo, 2004), which suggests that the Sri Lankan stock market is characterized by a state of partial inefficiency, meaning that there is opportunism in the short-term structure of prices that can be systematically exploited for forecasting purposes. This suggests that, although the stock market is not perfectly predictable, its historical information retains temporary validity as inputs for forecasting. These empirical data indicate that the ARIMA (2,1,1) model is valid and can be used as an appropriate forecasting function to predict short-term trends in the stock markets of emerging countries, such as Sri Lanka.

Regarding forecast accuracy metrics, the ARIMA(2,1,1) model generated a Mean Absolute Percentage Error (MAPE) of 3.9%, confirming strong short-term forecasting accuracy. The relative error percentages from the one-month out-of-sample forecast (July 2024) ranged from -0.50% to -7.22%, with errors increasing as the forecast horizon extended, consistent with expected ARIMA behavior. The low relative errors observed in Table 4 collectively confirm that the model meets standard benchmarks for acceptable forecasting performance in financial time-series analysis (Tsay, 2005).

## VI. Conclusion

The primary objective of this research was to investigate the predictive capabilities of the ARIMA model for forecasting stock prices, particularly in the Sri Lankan context. Considering the increasing relevance of accurate financial forecasts for emerging countries, the purpose of this research was to explore the potential for historical data to provide significant insight into stock price behavior for the CSE into the future. Based upon daily data ranging from July 1, 2014, through June 30, 2024, the study utilized the Box and Jenkins approach to formally identify, estimate, and test the optimal ARIMA model. The results confirmed the optimal model to be ARIMA (2,1,1), as evidenced by the lowest AIC and SBIC values. Strong empirical support emerged for the alternative hypothesis that the ARIMA model was highly predictive of short-term stock prices in Sri Lanka, with high significance and accuracy.

The results of this study recognize that there is a moderate level of predictability within the CSE, suggesting that whilst some aspects of weak-form efficiency exist in the market, it does not fully conform to the assumptions of the EMH. The presence of short-term autocorrelation and the model's successful forecasting outcomes indicate that historical price-related information continues to have predictive power. The effects of these results demonstrate synchronization with AMH, which assumes that market efficiency is a function of market agents' learning and adaptability in response to changing conditions. The economic and political instabilities of Sri Lanka, including but not limited to the Easter bombings of 2019, the COVID-19 outbreak, and the economic crisis of 2022, have created a temporarily inefficient and information-asymmetric market, for which time-series models like ARIMA can be leveraged for short-term forecasting. The results, accordingly, reaffirm that market efficiency for developing economies is relative and conditional, rather than absolute and universal.

From a theoretical standpoint, this study contributes to the existing literature by affirming that traditional linear series models remain valid and useful for accurately predicting financial market activity, despite increasing trends towards the predominance of nonlinear and machine-learning models. This has been demonstrated by the ARIMA(2,1,1) model's adequate effectiveness, which does not require complex nonlinear combinations, ensuring that simple models can also accurately capture the requisite market dynamics of emerging market stock market indices. Additionally, existing literature on market efficiency has been augmented by demonstrating that rational market expectations coexist alongside aberrations, such as investor overreaction and delayed information processes, which manifest as temporary departures from market efficiency.

From a practical perspective, the study's findings are beneficial for investors, policymakers, and regulators. According to the study's findings, investors can utilize ARIMA models to design a short-term trading strategy; however, it is essential to note that forecast accuracy decreases over time. On the other hand, policymakers and regulators can use the study's findings as evidence of market frictions and interpret them as a clear indication that more transparent mechanisms for information disclosure are required. It also confirms that the study's success highlights the importance of using forecasting tools in the country's national financial planning.

Notwithstanding its significance, it is essential to acknowledge the study's limitations. The reliance on univariate time-series analysis implies that it did not account for external macroeconomic factors, such as interest rates, inflation, or exchange rates. Future studies could explore combining this existing model with other models, such as ARIMA-GARCH or ARIMA-LSTM. Additionally, cross-market studies could be conducted across South Asian markets to gain insights into the overall dynamics of the region and the levels of cross-market linkage. The overall conclusion drawn from this study stands to confirm that, even though it remains impossible to accurately foretell in financial markets, ARIMA models remain an accurate and potent forecasting tool in modelling emerging markets like Sri Lanka.

In terms of forecast accuracy, the ARIMA(2,1,1) model yielded a MAPE of 3.9% for the out-of-sample forecast period, indicating strong short-term predictive performance. Relative forecast errors in July 2024 ranged from -0.50% to -7.22%, with accuracy declining as the forecast horizon extended consistently with the inherent limitations of univariate time-series models. These results confirm that ARIMA (2,1,1) meets acceptable benchmarks for financial forecasting and supports its practical utility for short-term investment decision-making in the Sri Lankan market context.

## Conflict of Interest

The authors have no competing interests to declare

## Data Accessibility Statement

The data used in this study are publicly available from the Colombo Stock Exchange official database (https://www.cse.lk) and were authenticated against the Bloomberg historical series. No proprietary or restricted data were used.

## Funding

This research received no external funding.

## Authors' Contributions

**Tharshiga, P.:** Conceptualization, Methodology, Data Curation, Formal Analysis, Writing – Original Draft. **Paranika, T.:** Investigation, Data Curation, Writing – Original Draft, Writing – Review and Editing, Visualization. **Subramaniam,**

**V.A.:** Supervision, Writing – Review and Editing, Validation, Project Administration.

## References


Ahammed, Md.Y., Chowdhury, Md.A.H., Rahman, Md.T. and Tazminur, S. (2024) 'Financial indicators shaping share price movement: Insight from the Dhaka Stock Exchange SME Board', International Journal of Economics and Financial Issues, 14(6), pp. 267–272. doi:10.32479/ijefi.17136.

Bachelier, L. (1900) 'Théorie de la spéculation', Annales Scientifiques de l'École Normale Supérieure, 3(17), pp. 21–86.

Barberis, N. (2003) 'A survey of behavioral finance', in Constantinides, G.M., Harris, M. and Stulz, R.M. (eds) Handbook of the Economics of Finance. Vol. 1B. Amsterdam: Elsevier, pp. 1053–1128.

Box, G.E.P. and Jenkins, G.M. (1970) Time series analysis: forecasting and control. San Francisco: Holden-Day.

Chen, S., Wang, X., Zhao, J., Zhang, Y. and Kan, X. (2022) 'Application of the ARIMA model in forecasting the incidence of tuberculosis in Anhui during the COVID-19 pandemic from 2021 to 2022', Infection and Drug Resistance, 15, pp. 3503–3512. doi:10.2147/IDR.S367528.

Dhyani, B., Kumar, M., Verma, P. and Jain, A. (2020) 'Stock market forecasting technique using ARIMA model', International Journal of Recent Technology and Engineering, 8(6), pp. 2694–2697.

Dong, C. (2022) 'Stock trend forecasting using the ARIMA model', Highlights in Science, Engineering and Technology, 16, pp. 56–62. doi:10.54097/hset.v16i.2239.

Fama, E.F. (1970) 'Efficient capital markets: A review of theory and empirical work', Journal of Finance, 25(2), pp. 383–417. doi:10.2307/2325486.

Fama, E.F. and French, K.R. (2004) 'The capital asset pricing model: Theory and evidence', Journal of Economic Perspectives, 18(3), pp. 25–46.

Ganesan, A. and Kannan, A. (2021) 'Stock price prediction using ARIMA model', International Research Journal of Engineering and Technology, 8(8), pp. 226–234.

Ghosh, P. (2022) 'The modelling and forecasting of BSE Sensex for September 2022 using ARIMA', in Proceedings of the 10th International Conference on Emerging Trends in Corporate Finance and Financial Markets.

Grossman, S.J. and Stiglitz, J.E. (1980) 'On the impossibility of informationally efficient markets', The American Economic Review, 70(3), pp. 393–408.

Khanderwal, S. and Mohanty, D. (2021) 'Stock price prediction using ARIMA model', International Journal of Marketing and Human Resource Research, 2, pp. 98–107.

Lo, A.W. (2004) 'The adaptive markets hypothesis: Market efficiency from an evolutionary perspective', Journal of Portfolio Management, 30(5), pp. 15–29.

Mahadik, A., Vaghela, D. and Mhaisgawali, A. (2021) 'Stock price prediction using LSTM and ARIMA', in Proceedings of the Second International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India.

Puaschunder, J. (2025) Emerging markets in times of regulatory uncertainty. Research Association for Interdisciplinary Studies. doi:10.5281/zenodo.17073088.

Ranasinghe, R. (2021) 'Forecasting stock prices in the Colombo Stock Exchange using ARIMA and neural network models', Sri Lankan Journal of Applied Statistics, 22(1), pp. 45–62.

Samarawickrama, I.D.W. and Pallegedara, A. (2023) 'Sri Lankan stock market volatility analysis: An ARMA–GARCH approach', Sri Lankan Journal of Business Economics, 12(1), pp. 60–78.

Samuelson, P.A. (1965) 'Proof that properly anticipated prices fluctuate randomly', Industrial Management Review, 6(2), pp. 41–49.

Shahid, M.K. (2024) Impact of research and development expenditures on ASEAN-5 and South Asian-3 countries' economic growth. PhD thesis. Universiti Tunku Abdul Rahman. Available at: http://eprints.utar.edu.my/6274/1/PhD_Thesis_Muhammad_Khalid_Shahid_22UKD00904_IR.pdf

Tsai, C.-F. and Hsiao, Y.-C. (2010) 'Combining multiple feature selection methods for stock prediction', Decision Support Systems, 50(1), pp. 258–269.

Tsay, R.S. (2005) Analysis of financial time series. 2nd edn. Hoboken, NJ: Wiley.

Wijayaweera, V. (2021) Impact of events on stock market and prediction of impact status and period. MSc thesis. University of Colombo School of Computing. Available at: https://dl.ucsc.cmb.ac.lk/jspui/bitstream/123456789/4646/1/2018%20MCS%20100.pdf (Accessed: 18 December 2024).

Yadav, M.P. and Khera, A. (2023) 'Forecasting the stock return of emerging economies: An empirical study based on ARIMA', International Journal of Public Sector Performance Management, 11(4), pp. 451–466.