POSITIVE SCIENCE

Positive Science

Positive Science - Economics

Time Series Modeling and Forecasting of NEPSE Index: An ARIMA/SARIMA Approach with Market Efficiency Evidence

Nischal Shrestha

Publication Information

Published: May 11, 2026

Publisher location: Strasbourg, France

Time Series Modeling and Forecasting of the NEPSE Index using SARIMA: Evidence from Rolling Forecast Evaluation

Nischal Shrestha
Independent Researcher, Nepal

ABSTRACT

Forecasting stock market behavior remains an important issue for investors, policymakers, and financial researchers, particularly in emerging markets where market volatility and informational inefficiencies may influence investment decisions. Despite growing interest in forecasting the Nepal Stock Exchange (NEPSE) index, existing studies have primarily focused on conventional ARIMA-based approaches without sufficiently incorporating rolling-window out-of-sample evaluation and benchmark comparison within the context of weak-form market efficiency. Addressing this gap, the present study examines the forecasting performance of the NEPSE daily closing index using an ARIMA/SARIMA-based time series framework and compares the selected model with a naïve random walk benchmark. The study utilized 1,157 daily observations of the NEPSE index and applied descriptive statistics, Augmented Dickey–Fuller stationarity testing, autocorrelation analysis, SARIMA model selection procedures, residual diagnostic testing, and rolling-window one-step-ahead forecasting techniques. Based on model selection criteria, seasonal structure, and diagnostic adequacy, SARIMA (3,1,**0) (**2,0,**0) [**5] was selected as the final forecasting model. The findings revealed that the SARIMA model generated slightly lower forecasting errors than the naïve benchmark during the rolling-window evaluation period, although the forecasting gains remained economically modest. Consequently, the results are broadly consistent with weak-form market efficiency, suggesting that historical price information provides only a limited predictive advantage in forecasting the NEPSE index. Residual diagnostic analysis further indicated the presence of remaining volatility clustering, implying that future studies may benefit from integrating GARCH-type volatility models. Overall, the study contributes to the growing NEPSE forecasting literature by incorporating rolling-window out-of-sample forecasting and benchmark comparison within a seasonal time series framework.

Keywords: NEPSE, SARIMA, Stock Market Forecasting, Time Series Analysis, Rolling-Window Forecasting, Weak-Form Market Efficiency.

INTRODUCTION

Forecasting stock market movements has remained a central issue in financial economics due to its implications for investment decision-making, portfolio management, risk assessment, and financial market efficiency. In particular, the ability of time series models to extract predictive information from historical price movements continues to attract significant academic and practical attention. However, financial time series are characterized by non-stationarity, volatility clustering, structural fluctuations, and dynamic dependence patterns, making accurate forecasting inherently challenging.

The Efficient Market Hypothesis (EMH), introduced by Fama (1970), provides an important theoretical foundation for evaluating forecasting performance in financial markets. Under weak-form efficiency, current asset prices fully incorporate all historical price information, implying that future price movements cannot be systematically predicted using past prices alone. Consequently, if sophisticated forecasting models fail to generate economically meaningful improvements over a naïve random walk benchmark, the findings may be interpreted as evidence broadly consistent with weak-form market efficiency.

Despite these theoretical implications, empirical studies have reported mixed evidence regarding stock market predictability, particularly in emerging markets where information inefficiencies, lower market depth, and evolving institutional structures may create temporary forecasting opportunities. In this context, time series econometric models such as the Autoregressive

Integrated Moving Average (ARIMA) and Seasonal ARIMA (SARIMA) frameworks have been widely employed to model and forecast stock market indices. The Box-Jenkins methodology developed by Box (1970), remains one of the most widely applied approaches for capturing linear dependence structures in financial time series data.

The Nepal Stock Exchange (NEPSE), the sole organized securities exchange in Nepal, has experienced substantial structural and behavioral change in recent years. The market has witnessed increased investor participation, expansion of the dematerialized trading system, rapid growth in trading volume, and heightened market volatility. Such developments have intensified interest in understanding the statistical behavior and forecasting dynamics of the NEPSE index. At the same time, the Nepalese stock market remains relatively underexplored in empirical financial econometrics compared to more developed financial markets.

Previous studies on NEPSE forecasting have primarily focused on the application of ARIMA-type models to evaluate short-term predictive performance (Gaire, 2017). Existing literature has also documented the presence of volatility clustering and conditional heteroskedasticity within the Nepalese stock market (G.C., 2009). However, several important limitations remain in the current body of research. First, prior studies have largely emphasized in-sample model fit rather than rigorous out-of-sample forecasting evaluation. Second, relatively limited attention has been paid to rolling-window forecasting procedures, which have not been systematically compared against a naïve random walk benchmark in terms of forecasting performance within the framework of weak-form market efficiency.

Accordingly, the present study seeks to address these gaps by applying ARIMA/SARIMA-based forecasting techniques to the daily NEPSE closing index using recent market data covering the period from 2001 to 2006. The study employs a rolling-window out-of-sample forecasting framework and evaluates predictive performance relative to a naïve random walk benchmark using standard forecast accuracy measures. In addition, the findings are interpreted within the context of weak-form market efficiency to assess whether historical price information provides a meaningful predictive advantage.

The contribution of this study is threefold. First, it provides updated empirical evidence on NEPSE forecasting using recent post-pandemic market data. Second, it adopts a rolling-window forecasting framework that improves methodological rigor and reduces look-ahead bias relative to conventional in-sample evaluation. Thirdly, the study links forecasting performance with market efficiency interpretations by comparing SARIMA forecasts against a naïve benchmark model.

Literature Review

Efficient Market Hypothesis and Random Walk Behavior

The Efficient Market Hypothesis (EMH), proposed by Eugene F. Fama, remains one of the foundational theories in financial economics. According to the weak-form version of the hypothesis, current asset prices fully reflect all historical market information, implying that future price movements cannot be systematically predicted using past price behavior alone. Under such conditions, forecasting models based solely on historical prices are expected to generate only limited predictive superiority over naïve benchmark models (Fama, 1970)

The random walk theory further supports this perspective by suggesting that stock price changes follow an unpredictable stochastic process in which successive movements are largely independent over time. Consequently, future market behavior should not be consistently forecasted through historical price analysis alone. However, subsequent empirical evidence has challenged the strict assumption of random walk behavior. Lo and MacKinlay (1998) documented serial dependence in stock returns, providing evidence against pure random walk behavior in financial markets. Similarly, Jegadeesh and Titman (1993) identified momentum effects in stock returns, suggesting that certain short-term price continuation patterns may exist within financial markets.

Despite these findings, the predictive forecasting value of historical price information remains highly debated. Malkiel (2003) argued that although markets may not be perfectly efficient, achieving economically meaningful and consistently exploitable forecasting gains remains extremely difficult in practice. Brown (2020) further emphasized that the Efficient Market Hypothesis continues to play a central role in understanding investment behavior, market efficiency, and financial forecasting limitations.

The debate surrounding market predictability is particularly relevant in the context of emerging financial markets, where information inefficiencies, lower market depth, and evolving institutional structures may create temporary forecasting opportunities. Nevertheless, even in such markets, forecasting superiority over naïve benchmark models frequently remains economically limited, reinforcing the continuing relevance of weak-form efficiency arguments.

Time Series Forecasting and Financial Market Dynamics

Time series econometric models have been widely applied in financial forecasting due to their ability to capture temporal dependence structures in sequential data. Among these approaches, the Box-Jenkins ARIMA framework developed by Box (1970) remains one of the most extensively used methodologies for modeling and forecasting non-stationary financial time series. ARIMA models combine autoregressive and moving average processes with different differencing procedures to capture underlying stochastic structures and generate short-term forecasts.

Financial time series, however, it exhibits several complex statistical properties that complicate forecasting performance. These characteristics include non-stationarity, volatility clustering, structural fluctuations, and time-varying variance. Such features often reduce the stability and forecasting accuracy of traditional linear models. Mandelbrot (1963) highlighted that speculative price movements frequently display irregular fluctuations and non-normal behavior, challenging conventional assumptions regarding financial data.

To address volatility behavior in the financial market, Engle (1982) introduced the Autoregressive Conditional Heteroskedasticity (ARCH) framework, which models time-varying variance in financial series. Subsequent developments by Bollerslev et al. (1992) further emphasized the importance of volatility modeling in financial econometrics and demonstrated that financial series commonly exhibit volatility clustering and conditional heteroskedasticity. These findings imply that while ARIMA-type models may adequately capture mean dynamics, volatility processes often remain incompletely modeled within simple linear forecasting frameworks.

Hence, forecasting stock market behavior remains inherently challenging. Even when econometric models identify short-term dependence structures, the presence of volatility persistence and rapidly changing market conditions often limits the economic significance of forecasting improvements.

ARIMA/ SARIMA Applications in Stock Market Forecasting

ARIMA and SARIMA models have been extensively applied in stock market forecasting across both developed and emerging financial markets. Numerous studies report that ARIMA-type models are capable of capturing short-term linear dependence structures and producing relatively accurate short-horizon forecasts. These models are particularly useful in financial series exhibiting recurring trading patterns and periodic structures.

Nevertheless, empirical evidence regarding the forecasting superiority of ARIMA-based models remains mixed. While several studies report improvements in forecast accuracy relative to simple benchmark models, others conclude that the forecasting gains are often economically modest and inconsistent across different evaluation measures. This inconsistency has contributed to ongoing debates regarding the practical usefulness of time series forecasting models in financial markets.

Recent forecasting literature increasingly emphasizes the importance of out-of-sample evaluation and rolling-window forecasting procedures for obtaining a realistic predictive assessment. Bergmeir et al. (2018) argued that appropriate forecasting validation frameworks are essential for avoiding overfitting and obtaining reliable predictive evaluations in autoregressive time series models. Rolling-window forecasting procedures are therefore considered methodologically superior to purely in-sample evaluation because they better approximate a real-world forecasting environment and reduce look-ahead bias.

In addition to statistical forecasting accuracy, several researchers emphasize the importance of evaluating whether predictive improvements are economically meaningful. Leitch and Tanner (1991) argued that conventional forecast error measures alone may not fully capture the practical usefulness of forecasting models in financial decision-making. These developments suggest that forecasting evaluation should not focus solely on statistical fit, but also on whether forecasting models meaningfully outperform naïve benchmark specifications under realistic out-of-sample conditions.

Empirical Evidence from Emerging Markets and NEPSE

Empirical evidence from emerging stock markets suggests that forecasting behavior may differ from that observed in highly developed financial markets due to lower informational efficiency, evolving regulatory systems, and changing investor behavior. Emerging markets frequently exhibit greater volatility, speculative trading activity, and temporary market inefficiencies, which may create limited short-term forecasting opportunities.

However, several studies also report that although forecasting models may capture short-run dependence structures within emerging markets, forecasting improvements over naïve benchmark models often remain economically limited. Such findings imply that even in relatively less efficient markets, consistently outperforming random walk benchmarks remains difficult.

Within the Nepalese context, empirical studies examining the Nepal Stock Exchange (NEPSE) remain comparatively limited. Gaire (2017) applied ARIMA and GARCH models to forecast the NEPSE index and documented the presence of volatility dynamics within the Nepalese stock market. Similarly, Maskey (2022) employed ARIMA-based forecasting techniques and reported acceptable short-term predictive performance for the NEPSE index. Paudel (2024) also applied ARIMA modeling to NEPSE forecasting and provided recent evidence regarding time series forecasting behavior in the Nepalese market.

Although these studies contribute valuable empirical evidence, several limitations remain within the existing literature. First, previous NEPSE forecasting studies have primarily focused on in-sample model estimation rather than rigorous rolling-window out-of-sample forecasting evaluation. Second, relatively limited attention has been devoted to systematic comparison between ARIMA/SARIMA models and naïve random walk benchmark models within the framework of weak-form market efficiency. Third, existing studies frequently emphasize statistical model fit while providing comparatively limited discussion regarding the economic significance of forecasting improvements.

Studies examining volatility behavior in the Nepalese stock market further indicate the presence of heteroskedasticity and changing market variance. G.C. (2009) documented significant volatility dynamics within the Nepalese stock market, suggesting that volatility persistence remains an important feature of NEPSE behavior. Similarly, Shrestha and Kayastha (2024) reported variations in weekday risk behavior within the NEPSE index, indicating the presence of short-term market irregularities and temporal dependence patterns.

Overall, the existing literature suggests that although ARIMA-type models may capture certain short-term dependence structures within the NEPSE index, the broader question of whether such models generate economically meaningful forecasting superiority relative to naïve benchmark models remains insufficiently explored.

Methodology

Data Description

This study employs daily closing values of the Nepal Stock Exchange (NEPSE) index as the primary variable of analysis. The dataset covers the period from March 23, 2021, to March 23, 2026, and consists of 1157 trading observations. The data are secondary in nature and were obtained from official records of the Nepal Stock Exchange.

Daily closing prices are widely utilized in financial econometric research because they represent the aggregate market valuation and investor expectations at the close of each trading session. Compared with lower-frequency observation, daily data provides greater sensitivity to short-term market dynamics and improves the ability of time series models to capture temporal dependence structures.

The NEPSE index was selected due to its importance as the benchmark indicator of Nepal's capital market performance. In recent years, the Nepalese market has experienced substantial fluctuations, increased investor participation, and rapid changes in market activity, making it an appropriate context for evaluating forecasting performance and market efficiency.

Research Design

The study adopts a quantitative time series econometric research design to examine the forecasting behavior of the NEPSE index and evaluate whether historical price information contains meaningful predictive value. The analysis is conducted within the theoretical framework of the Efficient Market Hypothesis (EMH), particularly weak-form efficiency, which suggests that the current market price fully reflects past price information.

Under weak-form efficiency, future price movements should not be systematically predictable using historical market data alone. Consequently, the forecasting performance of the SARIMA Model is evaluated relative to a naïve random walk benchmark economically limited relative to the benchmark; the findings may be interpreted as broadly consistent with weak-form market efficiency. The empirical framework of the study consists of the following stages:

i. Descriptive statistical analysis of the NEPSE index,
ii. Stationarity testing using the Augmented Dickey-Fuller (ADF) test,

iii. Model identification using autocorrelation and partial autocorrelation analysis,
iv. Estimation of ARIMA/SARIMA models,
v. Residual diagnostic testing,
vi. Rolling-window out-of-sample forecasting,
vii. Comparative forecast evaluation against a naïve random walk benchmark.

Model specification

The study applies the Box-Jenkins time series methodology to model the behavior of the NEPSE index using ARIMA and SARIMA specifications. The general ARIMA framework is represented as:

ARIMA (p, d, q) (1)

Where,

P = Autoregressive Order,

d = Degree if differencing is required for stationarity,

q = Moving Average Order.

Because stock market series may exhibit recurring short-term trading patterns, seasonal components were additionally considered through the Seasonal ARIMA Framework:

SARIMA (P, D, Q) [S] (2)

Where,

P = Seasonal Autoregressive

D = Differencing

Q = Moving Average Component

S= Seasonal Period

This inclusion of seasonal structure is economically meaningful in the context of the NEPSE Index due to the five-day weekly trading cycle of the Nepalese stock market.

Stationarity Testing

Stationarity is a fundamental requirement in ARIMA-based time series modeling because non-stationary series may produce misleading statistical inferences and unstable forecasting performance. To evaluate the stationarity properties of the NEPSE index, the Augmented Dickey-Fuller (ADF) unit root test was employed.

The null hypothesis of the ADF test states that the series contains a unit root and is therefore non-stationary, while the alternative hypothesis suggests stationarity. Where necessary, differencing procedures hypothesis suggests stationarity. Where necessary, differencing procedures were applied to transform the series into a stationary process before model estimation.

Model Identification and Selection

Model identification was conducted using a combination of graphical diagnostics and information criteria. The autocorrelation function (ACF) and partial autocorrelation function (PACF) plots were analyzed to identify potential autoregressive and moving average structures.
The auto. arima () procedure in R was employed to estimate competing ARIMA and SARIMA specifications. Model selection was based primarily on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), where lower values indicate comparatively superior model fit while maintaining model.

In addition to information criteria, residual diagnostic performance and forecasting behavior were considered during final model selection. Seasonal components were evaluated using a five-day seasonal structure corresponding to the weekly trading cycle of the Nepal Stock Exchange.

Residual Diagnostic Testing

Following model estimation, diagnostic tests were conducted to evaluate the adequacy of the fitted SARIMA specification.

The Ljung-Box test was employed to assess the presence of residual autocorrelation. The null hypothesis of the test states that the residuals are independently distributed and exhibit no serial correlation. Failure to reject the null hypothesis suggests that the model adequately captured the linear dependency structure of the series.

In addition, the Autoregressive Conditional Heteroskedasticity Lagrange Multiplier (ARCH-LM) test was applied to examine whether the residuals exhibit time-varying volatility. Significant ARCH effects indicate the presence of conditional heteroskedasticity and suggest that volatility dynamics remain unmodeled within the conditional mean specification.

Forecast Evaluation Procedure

To evaluate forecasting performance under realistic market conditions, the study employs a rolling-window out-of-sample forecasting framework. This approach is better because it better approximates real-world forecasting environments and reduces look-ahead bias.

The forecasting exercise was conducted using a 120-day rolling forecast horizon. An initial estimation sample was first used to estimate the SARIMA model, after which one-step-ahead forecasts were recursively generated for the subsequent trading observation. Following each forecast iteration, the estimation window was expanded by one observation, and the model was re-estimated using the updated information set.

Forecast performance was then compared with a naïve random walk benchmark model, where the forecasted value for the next trading day equals the current observed value of the index. The naïve benchmark serves as an important reference model within the framework of weak-form market efficiency because it represents the hypothesis that future price movements cannot be systematically predicted from historical information.

Forecast Accuracy Measures

Forecasting accuracy was evaluated using three commonly applied statistical performance measures: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). MAE measures the average magnitude of forecasting errors without considering direction, while RMSE places relatively greater weight on larger forecast deviations, and MAPE expresses forecast error as a percentage and facilitates comparison across different forecasting models. Lower values of MAE, RMSE, and MAPE indicate superior forecasting performance. Comparative evaluation of these measures between the SARIMA model and the naïve random walk benchmark provides insight into whether the SARIMA specification generates economically meaningful predictive improvement.

Results

Descriptive Statistics

Table 1. The NEPSE daily closing index for the time period March 23, 2021, to March 23, 2026, based on 1157 trading day observations.

Statistic	**Value**
N	1,157
Mean	2394.799
Median	2540.570
Maximum	3198.190
Minimum	1815.130
Range	1383.060
Standard Deviation	370.431
Variance	137218.900
Coefficient of Variation (%)	15.468
Jarque-Bera Statistic	100.721
Jarque-Bera p-value	.000

Table 1 presents the descriptive statistics of the daily NEPSE index series used in the study. The dataset consists of 1157 daily observations, with a mean index value of 2394.799 and a median of 2540.570, indicating that the distribution of the series is moderately asymmetric. The NEPSE index recorded a maximum value of 3198.190 and a minimum value of 1815.130 during the study period, producing a total range of 1383.060 points.

The standard deviation of 370.431 indicates considerable variability in the index movements over the sample period, reflecting the volatile nature of the Nepalese stock market. Similarly, the coefficient of variation of 15.468% suggests moderate relative dispersion in the index values.

The Jarque-Bera statistic of 100.721 with a p-value below 0.05, indicates the rejection of the null hypothesis of normal distribution. These findings suggest that the NEPSE index series does not follow a normal distribution and exhibits characteristics commonly observed in financial time series data. Such behavior is consistent with the findings of Mandelbrot (1963), who argued that financial market series frequently display irregular fluctuations and non-normal distributional properties.

Table 2. NEPSE Log Returns

Statistic	**Value**
N	1,157
Mean	0.0001292653
Standard Deviation	0.0138427
Skewness	0.489847
Kurtosis	5.047341

Table 2 presents the descriptive statistics of NEPSE log returns, indicating a mean daily return of 0.000129 and a standard deviation of 0.0138, suggesting relatively small average returns with a noticeable short-term volatility. The return distribution exhibits positive skewness of 0.490, indicating a slight asymmetry toward positive returns. Similarly, the kurtosis value of 5.047 exceeds the normal benchmark value of 3, implying leptokurtic behavior and the presence of heavier tails in the distribution.

These findings indicate that the NEPSE return series does not follow a normal distribution and exhibits characteristics commonly observed in financial time series data (Mandelbrot, 1963).

Time Series Visualization of NEPSE Index

Figure 1. Time series visualization of NEPSE index

Figure 1 presents the daily movement of the NEPSE index during the study period. The series exhibits substantial fluctuations and changing trend patterns over time, reflecting the volatile nature of the Nepalese stock market. The index experienced periods of rapid growth followed by sharp declines and subsequent recovery phases, indicating considerable variations in market behavior across the same period. Visual inspection of the series suggests the presence of non-stationarity, as both the mean level and variability appear to change over time. Such characteristics are commonly observed in financial time series data and provide preliminary justification for conducting formal stationarity testing before estimating the forecasting model.

Stationarity and Unit Root Testing

Augmented Dickey Fuller (ADF) Test

Table 3. Augmented Dickey–Fuller (ADF) Unit Root Test for NEPSE Index Series

```
        Augmented Dickey-Fuller Test

data:  df$Close
Dickey-Fuller = -1.7841, Lag order = 10, p-value = 0.6697
Alternative hypothesis: stationary
```

Table 3 reports that the ADF test produced a statistic of -1.784 with a p-value of 0.6697. Since the p-value exceeds the 5% significance level, the null hypothesis of a unit root cannot be rejected, indicating that the NEPSE index series is non-stationary in its level form. Hence, differencing was applied to achieve stationarity before model estimation.

Table 4. Augmented Dickey–Fuller (ADF) Unit Root Test for NEPSE Index Series

Augmented Dickey-Fuller Test
data: nepse_diff
Dickey-Fuller = -8.9425, Lag order = 10, p-value = 0.01
Alternative hypothesis: stationary

Table 4 reports that the ADF test for the first differenced series produced a test statistic of -8.943 with a p-value below 0.05. Therefore, the null hypothesis of a unit root was rejected, indicating that the differenced NEPSE series is stationary. This confirms that first-order differencing was sufficient to achieve stationarity for SARIMA modeling.

Model Identification and Selection

Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF)

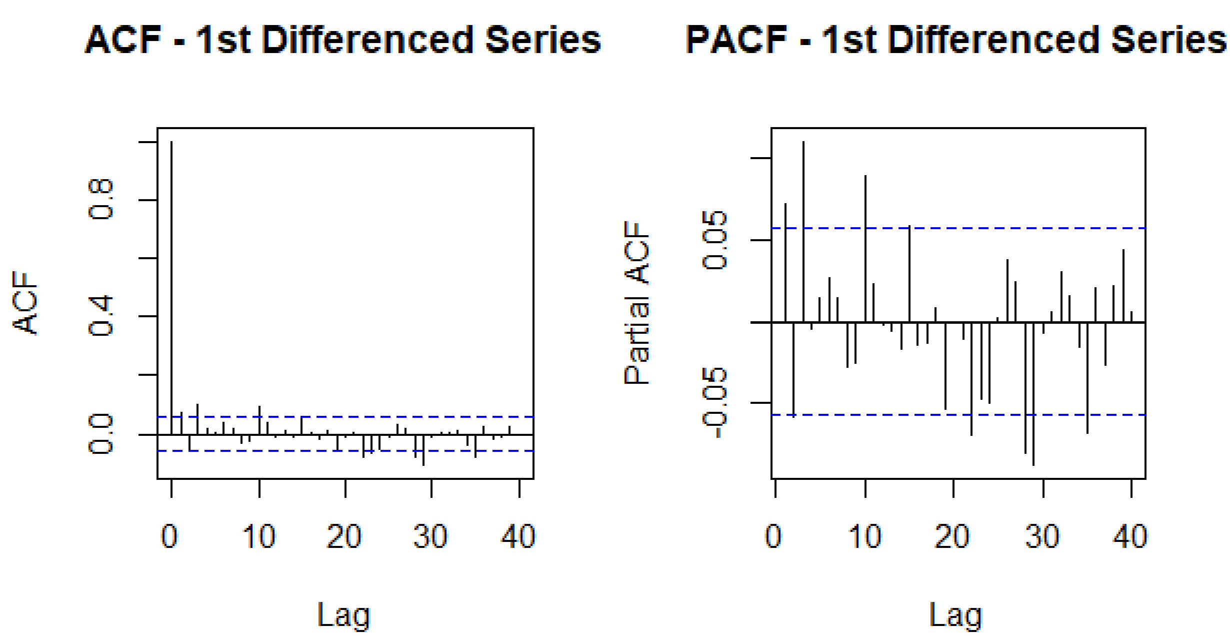

Figure 2. ACF and PACF Plots of the First Differenced NEPSE Series

Figure 2 presents the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the first differenced series, which were examined to identify appropriate model specification. The ACF plot showed a gradual decline with most autocorrelation remaining within the significant bounds after differencing, suggesting that the non-stationarity present in the original series has been substantially removed.

Similarly, the PACF plot displayed several significant spikes at lower lags, indicating the possible presence of autoregressive components in the series. In addition, repeating lag patterns around the five-trading-day interval suggested the existence of weekly seasonal dependence within the NEPSE index.

SARIMA Model Estimation

Table 5. Estimated SARIMA Model Parameters

Series: nepse_ts

ARIMA (3,1,0) (2,0,0) [5]

Coefficients:

	ar1	ar2	ar3	sar1	sar2
	0.0826	-0.0644	0.1116	0.0145	0.0896
s. e	0.0292	0.0295	0.0294	0.0296	0.0297

sigma^2 = 1113: log likelihood = -5692.58

AIC=11397.15 AICc=11397.23 BIC=11427.47

AIC: 11397.15

BIC: 11427.47

Table 5 reports that based on the model identification process, SARIMA (3,1,0) (2,0,0) [5] was selected as the final forecasting specification. The model incorporates both non-seasonal autoregressive terms and seasonal autoregressive components with a five-day trading cycle.

The estimated coefficients indicate the presence of short-term dependence within the NEPSE index series. Among the estimated parameters, the non-seasonal autoregressive terms demonstrated relatively stronger contributions to the model, while the seasonal autoregressive components exhibited comparatively smaller effects. In particular, the first seasonal autoregressive coefficient was relatively weak compared to its standard error, suggesting limited seasonal influence at that lag.

The selected model produced an AIC value of 11397.15 and a BIC value of 11427.47, indicating an acceptable balance between model fit and parsimony. Overall, the SARIMA specification was considered appropriate for capturing the short-term and seasonal dynamics present in the NEPSE index series.

Candidate Model Comparison

Table 6. Comparison of ARIMA Model Selection Criteria

Model	AIC	BIC
ARIMA (1,1,1)	11408.1	11423.26
ARIMA (2,1,2)	11412.1	11437.37
ARIMA (0,1,2)	11412.33	11427.49
ARIMA (2,1,0)	11414.9	11430.06
ARIMA (0,1,1)	11415.86	11425.97
ARIMA (1,1,0)	11416.76	11426.87

Table 6 reports that to ensure model adequacy and avoid overreliance on automated model selection, several alternative ARIMA specifications were additionally compared using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). As presented in Table 6, the competing models produced relatively close information criterion values. However, the selected SARIMA (3,1,0) (2,0,0) [5] specification generated lower information criterion values while also accounting for the weekly seasonal structure observed in the NEPSE series.

The inclusion of seasonal autoregressive terms was considered economically reasonable given the five-day trading cycle of the Nepalese stock market. Therefore, the SARIMA specification was retained as the final forecasting model based on overall model fit, seasonal structure, and diagnostic adequacy.

Residual Diagnostic Testing

Ljung-Box Test

Table 7. Ljung–Box Test for Residual Autocorrelation

Box-Ljung test
data: residuals_model
X-squared = 10.315, df = 20, p-value = 0.9621

Table 7 reports that the Ljung–Box test was conducted to evaluate the presence of residual autocorrelation in the selected SARIMA model. The test produced a chi-square statistic of 10.315 with a p-value of 0.9621. Since the p-value exceeds the 5% significance level, the null hypothesis of no residual autocorrelation could not be rejected.

This result suggests that the SARIMA model adequately captured the linear dependence structure in the conditional mean of the NEPSE index series. Consequently, the residuals behave approximately as white noise, indicating that no significant autocorrelation remained after model estimation.

ARCH-LM Test

Table 8. ARCH–LM Test for Conditional Heteroskedasticity

ARCH LM-test; Null hypothesis: no ARCH effects
data: residuals_model
Chi-squared = 77.725, df = 10, p-value = 1.399e-12

Table 8 reports that the ARCH–LM test was performed to examine the presence of conditional heteroskedasticity in the residuals of the selected SARIMA model. The test produced a Chi-square statistic of 77.725 with a p-value of 1.399e-12. Since the p-value is significantly below the 5% significance level, the null hypothesis of no ARCH effects was rejected.

The result indicates the presence of remaining volatility clustering and conditional heteroskedasticity in the residual series. This suggests that although the SARIMA model adequately captured the linear dependence structure in the conditional mean, it did not fully model the volatility dynamics of the NEPSE index. Therefore, more advanced volatility models such as GARCH-type specifications may provide a more comprehensive treatment of the series.

Residuals Diagnostic Analysis

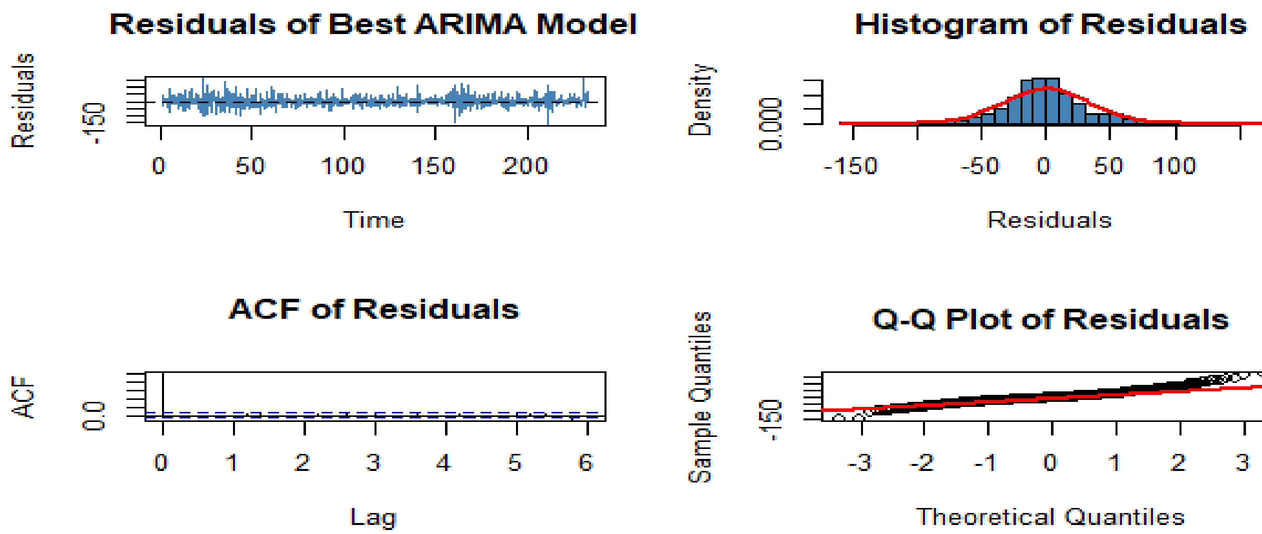

Figure 3. Residual Diagnostic Plots of the Selected SARIMA Model

Figure 3 presents the residual diagnostic plots of the selected SARIMA model. The residual series fluctuates randomly around zero without any clear systematic pattern, suggesting that the model adequately captured the linear dependence structure of the NEPSE index series. Similarly, the residual autocorrelation function (ACF) plot indicates the absence of substantial remaining autocorrelation, which is consistent with the Ljung–Box test results.

The histogram and Q–Q plot further suggest that the residuals are approximately centered around a normal distribution, although slight deviations in the tails remain visible. Such deviations are common in financial time series data due to volatility clustering and non-normal market behavior. Overall, the diagnostic results support the adequacy of the SARIMA model for modeling the conditional mean of the series, while the ARCH–LM test indicates that conditional heteroskedasticity and volatility dynamics remain present in the residuals.

Rolling-Window Forecasting Evaluation

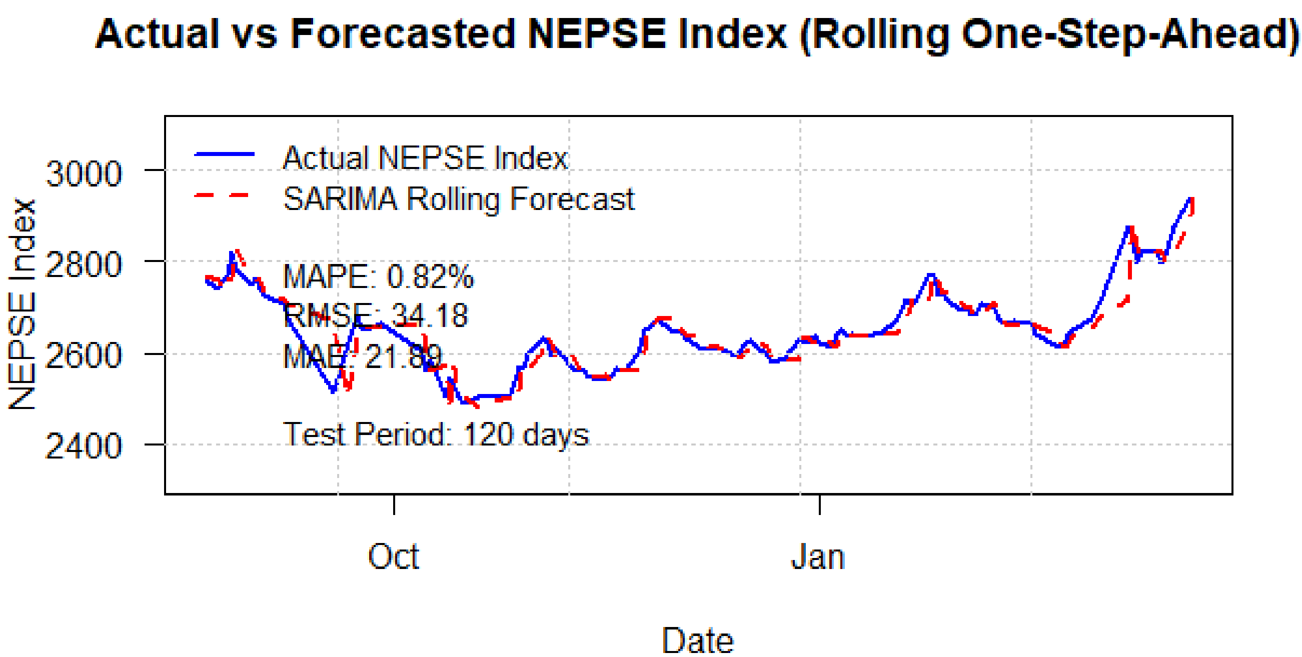

Figure 4. Actual vs. Forecasted NEPSE Index Using Rolling One-Step-Ahead SARIMA Forecasting

Figure 4 presents an evaluation of out-of-sample forecasting performance using a 120-day rolling one-step-ahead forecasting procedure that was implemented using the selected SARIMA (3,1,0)(2,0,0) [5] model. The rolling-window approach was adopted to simulate a realistic forecasting environment by generating sequential forecasts over the test period.

Figure 4 compares the actual NEPSE index values with the rolling SARIMA forecasts during the evaluation period. The forecasted values generally followed the movement of the actual series, indicating that the model was able to capture short-term market dynamics reasonably well.

The forecasting performance measures produced a Mean Absolute Percentage Error (MAPE) of 0.82%, a Mean Absolute Error (MAE) of 21.89, and a Root Mean Squared Error (RMSE) of 34.18. These results suggest that the SARIMA model generated relatively small forecast errors during the test period. However, the forecasting performance should be interpreted cautiously and comparatively against benchmark alternatives rather than solely based on individual error measures.

Discussion

The empirical findings indicate that the NEPSE index exhibits several characteristics commonly associated with financial time series data, including non-stationarity, short-term dependence, and volatility clustering. The Augmented Dickey–Fuller test confirmed that the original index series was non-stationary, while first-order differencing successfully transformed the series into a stationary form appropriate for SARIMA modeling. This finding is consistent with the Box–Jenkins time series framework, which emphasizes the importance of stationarity in ARIMA/SARIMA forecasting models (Box & Jenkins, 1970).

The selected SARIMA (3,1,0) (2,0,0) [5] specification demonstrated reasonable forecasting capability during the rolling-window evaluation period. The inclusion of the five-day seasonal structure was economically meaningful given the weekly trading cycle of the Nepalese stock market. Similar applications of ARIMA-based forecasting models within the NEPSE context were also reported by Gaire (2017), Maskey (2022), and Paudel (2024), who found that time series models could capture certain short-term market dynamics.

However, the comparison with the naïve random walk benchmark revealed that the forecasting gains of the SARIMA model remained relatively limited. Although the SARIMA model generated slightly lower forecasting errors across MAPE, RMSE, and MAE measures, the differences were economically modest. This finding suggests that historical price information provides only a limited predictive advantage in forecasting the NEPSE index. Such results are broadly consistent with the Efficient Market Hypothesis proposed by Fama (1970) and later discussed by Malkiel (2003), who argued that consistently outperforming simple benchmark forecasting approaches using historical price information alone remains difficult in financial markets.

The residual diagnostic analysis further indicated that the selected SARIMA model adequately captured the linear dependence structure in the conditional mean of the series, as supported by the Ljung–Box test results. Nevertheless, the ARCH–LM test revealed significant remaining heteroskedasticity, indicating the presence of volatility clustering within the residuals. This finding aligns with the volatility literature developed by Robert Engle (1982) and Bollerslev, Chou, and Kroner (1992), which emphasizes that financial time series frequently exhibit changing variance behavior that may require GARCH-type modeling frameworks for more comprehensive analysis.

Compared with previous NEPSE forecasting studies, the present study extends the literature by incorporating rolling-window out-of-sample forecasting together with benchmark comparison against a naïve random walk model. This provides a more rigorous evaluation of forecasting performance and contributes additional evidence regarding the predictability and efficiency characteristics of the Nepalese stock market.

Conclusion

This study examined the forecasting performance of the NEPSE daily closing index using an ARIMA/SARIMA-based time series framework and evaluated the predictive capability of the selected model against a naïve random walk benchmark. The empirical analysis identified SARIMA (3,1,0)(2,0,0) [5] as the most suitable forecasting specification based on stationarity transformation, seasonal structure, model selection criteria, and diagnostic adequacy.

The rolling-window forecasting evaluation revealed that the SARIMA model generated only marginal improvements over the naïve benchmark across the forecasting error measures. Although the model demonstrated limited short-term predictive capability, the forecasting gains remained economically modest. Consequently, the findings are broadly consistent with weak-form market efficiency, where historical price information provides limited forecasting advantage.

The study further identified the presence of remaining volatility clustering within the residual series, indicating that future studies may benefit from integrating ARIMA/SARIMA models with GARCH-type volatility frameworks for improved modeling of market variance behavior.

Overall, the study contributes to the growing literature on NEPSE forecasting by applying rolling-window out-of-sample forecasting and benchmark comparison within a seasonal time series framework. The findings provide useful insights into the forecasting behavior of the Nepalese stock market and may serve as a foundation for future forecasting and volatility modeling research.

References

Bergmeir, C., Hyndman, R. J., & Koo, B. (2018). A note on the validity of cross-validation for evaluating autoregressive time series prediction. *Computational Statistics & Data Analysis*, *120*, 70–83. https://doi.org/10.1016/j.csda.2017.11.003

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance. *Journal of Econometrics*, *52*(1–2), 5–59. https://doi.org/10.1016/0304-4076(92)90064-X

Box, G. (1970). Box and Jenkins: Time Series Analysis, Forecasting and Control. In T. C. Mills, *A Very British Affair* (pp. 161–215). Palgrave Macmillan UK. https://doi.org/10.1057/9781137291264_6

Brown, S. J. (2020). The Efficient Market Hypothesis, the *Financial Analysts Journal*, and the Professional Status of Investment Management. *Financial Analysts Journal*, *76*(2), 5–14. https://doi.org/10.1080/0015198X.2020.1734375

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*(4), 987. https://doi.org/10.2307/1912773

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, *25*(2), 383. https://doi.org/10.2307/2325486

Gaire, H. N. (2017). Forecasting NEPSE Index: An ARIMA and GARCH Approach. *NRB Economic Review*, *29*(1), 54–68. https://doi.org/10.3126/nrber.v29i1.52530

G.C., S. B. (2009). Volatility Analysis of Nepalese Stock Market. *Journal of Nepalese Business Studies*, *5*(1), 76–84. https://doi.org/10.3126/jnbs.v5i1.2085

Jegadeesh, N., & Titman, S. (1993). *Returns to buying winners and selling losers: Implications for stock market efficiency* (pp. 65–91). The Journal of Finance. https://doi.org/https://doi.org/10.1111/j.1540-6261.1993.tb04702.x

Leitch, G., & Ernest Tanner, J. (1991). Economic Forecast Evaluation: Profit Versus the Conventional Error Measures. *International Journal of Forecasting*, *8*(2), 279–282. https://doi.org/https://doi.org/10.1016/0169-2070(92)90130-2

Lo, A. W., & MacKinlay, A. C. (1988). Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specific. *The Review of Financial Studies*, *1*(1), 41–66. https://doi.org/https://doi.org/10.1093/rfs/1.1.41

Malkiel, B. G. (2003). The Efficient Market Hypothesis and its Critics. *Journal of Economic Perspectives*, *17*(1), 59–82. https://doi.org/https://doi.org/10.1257/089533003321164958

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *The Journal of Business*, *36*(4), 394. https://doi.org/10.1086/294632

Maskey, A. (2022). Predicting NEPSE Index Using ARIMA Model. *International Research Journal of Innovations in Engineering and Technology*, *06*(02), 80–85. https://doi.org/10.47001/IRJIET/2022.602014

Paudel, S. (2024). Predicting NEPSE Index: An ARIMA-Based Model. *The Lumbini Journal of Business and Economics*, *12*(1), 16–29. https://doi.org/10.3126/ljbe.v12i1.70318

Shrestha, K. K., & Kayastha, R. (2024). Risk Behavior of Different Weekdays in NEPSE Index. *Nepalese Journal of Statistics*, *8*, 1–16. https://doi.org/10.3126/njs.v8i1.73159