doi: 10.21608/esju.2025.338704.1053

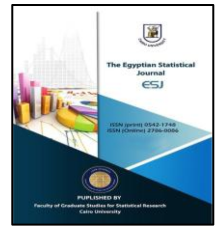
Research Article

The Egyptian Statistical Journal (ESJ), 69 (1): 30-48

Homepage: https://esju.journals.ekb.eg/

The Egyptian Statistical Journal

Print ISSN 0542-1748– Online ISSN 2786-0086



Enhancing the Efficiency of Time Series Forecasting by Hybrid Univariate Box Jenkins–GARCH Models

Mona Nazih Abdel Bary *

Received 23 November 2024; revised 23 January 2025; accepted 28 January 2025

Keywords

Box-Jenkins models; GARCH models; Hybrid models; Maximum likelihood estimator; Saudi general index.

Abstract

Due to the high non-linearity and volatility of the data, financial time series forecasting has been classified as a standard problem. The current study presents a method for modeling stationary, non-stationary, non-linear, and high volatility time series using a combined model of statistical methods. This study focuses on the performance of univariate Box Jenkins and the generalized autoregressive conditionally heteroscedasticity (GARCH) models in predicting financial time series and their volatility, and it presents an approach for forecasting financial time series that outperforms the performance of univariate Box Jenkins or GARCH models separately. According to the study, the performance of univariate Box Jenkins models can be improved by using the GARCH model of residuals of highly skewed data. The study's findings show that the SARIMA model is adequate for modeling the monthly Saudi General Index, with the best model being SARIMA (2, 2, 0) (2, 1, 1)4-GARCH (1, 1), with MAE, RMSE, and MAPE values of 38.2284, 57.35, and 4.247. The performance of Hybrid univariate Box Jenkins-GARCH Models shows that hybrid SARIMA-GARCH models fit financial time series and are highly accurate for short-term forecasting.

1. Introduction

The autoregressive integrated moving average (ARIMA) model and the generalized autoregressive conditionally heteroscedasticity (GARCH) model are commonly used in financial time series analysis. ARIMA can show the conditional mean of a time series. With the implicit assumption of homoskedasticity, GARCH is completely efficient in studying the volatility characteristics of time series. Therefore, the combination of ARIMA and GARCH models can be an effective way to overcome the limitations of each component model while also improving forecasting accuracy. In the current study, GARCH is used to build a hybrid model that overcomes the linear limitations of ARIMA models by including volatility in the forecast model, resulting in excellent forecast results (Nelson, 1991; Makridakis, Wheelwright, & Hyndman, 1998; Wang, Huang, & Wang, 2012; Garai et al., 2023).

One of the limitations of Box-Jenkins approach is the assumption of stationarity. It is not always possible to make a time-series stationary by differencing or by some other means. Additionally, it

✉ Corresponding author*: mona_abdelbary@commerce.suez.edu.eg
Department of Statistics and Insurance, Faculty of Commerce. Suez Canal University, Al Ismailia, Egypt.

is not capable of modelling those data sets that depict volatility. Additionally, The Box-Jenkins approach assumes that future values of a time series have a linear relationship with current and previous values, as well as with white noise; thus, autoregressive integrated moving average (ARIMA) models may be insufficient for complex nonlinear issues. Furthermore, ARIMA models require a large amount of historical data to achieve the desired results; therefore, ARIMA cannot handle the volatility and nonlinearity of the data series. Forecasting financial time series has proven to be difficult due to their nonlinearity and volatility. Statistical methods are based on market assumptions, do not account for all market variables, and may fail to detect nonlinearities.

This paper addresses the following issues; using Box-Jenkins SARIMA models for financial time series data. In addition, GARCH models are used to investigate and analyze the volatility in financial time series. Additionally, improves the forecast accuracy of Box-Jenkins and GARCH models. Introducing a new framework for identifying and estimating the mean and variance components of financial time series data using combined Box-Jenkins and GARCH models, which explains the volatility structure of the residuals obtained using the best mean models for the time series.

The remainder of this paper is organized as follows: Literature review discussed in Section (2). Section (3) outlines the Methodology. The proposal frame is introduced in Section (4). Section (5) introduces the results and discussion. Section (6) incorporates the conclusion and remarks.

2. Literature Review

Hybrid models have received significant attention in recent years, such as ARIMA-GARCH models, reflecting their growing importance in forecasting and analysis of time series. Boudrioua and Boudrioua (2020) introduce the ARIMA models to predict the Algerian Stock Exchange. The results indicate that the Algerian Stock Exchange could be efficiently modeled and predicted using the Box-Jenkins approach. Shahraki and Alimardani (2020) apply ARIMA models to model and forecast agricultural production. The study indicated that ARIMA could provide forecasts to help farmers and policymakers plan for harvests and suggested that incorporating climatic data into the ARIMA model could improve its forecasting accuracy. Qasim, Ali, Malik, and Liaquat (2021) combined GARCH to improve the performance of the ARIMA model in predicting inflation volatility, and the results show the combined ARIMA-GARCH provides a more effective forecast than traditional models. Hong et al. (2023) combine GARCH models with graph structures to model the dependencies and interactions between multiple time series, enhancing forecasting accuracy by capturing both temporal and spatial correlations. The proposed method shows its superior performance in handling complex and high-dimensional time series forecasting tasks. Adewole (2024) uses a hybrid of ARIMA and GARCH models in modeling volatilities in the Nigeria Stock Exchange. This study shows that the hybrid ARIMA-GARCH model performs better in predicting volatility compared to individual ARIMA or GARCH models. The findings provide valuable insights for investors and financial analysts in understanding and managing stock market risks in Nigeria.

3. Methodology

3.2 Univariate Box-Jenkins Approach

George Box and Gwilym Jenkins introduced the Autoregressive Integrated Moving Average (ARIMA) model family, which involves four iterative steps for model identification, parameter

estimation, diagnostic checking, and forecasting (Akaike, 1974; Box et al., 2015; Bollerslev & Wooldridge, 1992; Magnus & Fosu, 2006).
Equation (1) introduces AR(p), an autoregressive model of order.

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t, \quad (1)$$

where $\phi_1, \ldots, \phi_p$ are autocorrelation coefficients, and ε_t is a residual error term. The AR(p) model forecasts a variable based on a linear combination of previous values, while ARMA models combine AR and MA to create a composite time series model.

In Equation (2), the model is MA(q); then y_t is a function in the errors ε_t, so it is a linear function in current and previous errors; (Hillmer & Tiao, 1982).

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}, \quad (2)$$

where $\theta_1, \ldots, \theta_q$ are moving average coefficients. Mixed autoregressive moving average (ARMA) with order (p, q). The form is described as follows by Equation (3):

$$y_t = \mu + \sum_{i=1}^{p} \emptyset_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t. \quad (3)$$

For nonstationary and nonseasonal data series, the autoregressive integrated moving average model of order p and q, ARIMA(p,d,q), is recommended. Equation (4) illustrates the general form of ARIMA(p,d,q), which generates a time series with a mean μ.

$$\left(1 - \sum_{i=1}^{p} \emptyset_i B^i\right)(1 - B)^d y_t = \mu + \left(1 + \sum_{j=1}^{q} \theta_j B^j\right)\varepsilon_t, \quad (4)$$

where $\emptyset_i$ are coefficients of the AR model, θ_j are coefficients of the MA model, d is the order of differencing, and B is the backward shift operator.

Before applying an ARIMA model, the data series undergoes transformation and differencing to stabilize variance and eliminate trends. Box and Cox (1964) proposed a family of power transformations to deal with skewness in data (Box & Cox, 1964; Girish, 2016). The form is described as follows by Equation (5):

$$y_t^* = \begin{cases} \dfrac{y_t^\lambda - 1}{\lambda} & If\ \lambda \neq 0 \\ \ln(y_t) & If\ \lambda = 0 \end{cases}. \quad (5)$$

In this equation, y_t represents the actual data at time t. y_t^* represents the transformed data at time t, and λ represents the minimum mean square error of residuals. The above equation's transformation applies only to positive values of the time series $y_t > 0$. If the values of the time series also include negative values, the transformation will take the following form:

$$y_t^*(\lambda) = \begin{cases} \dfrac{(y_t + \lambda_2)^{\lambda_1} - 1}{\lambda_1} & If\ \lambda_1 \neq 0 \\ \ln(y_t + \lambda_2) & If\ \lambda_1 = 0 \end{cases}, \quad (6)$$

where λ_1 is the transformation parameter and λ_2 is selected so that $y_t > -\lambda_2$.

According to Equation (7), the seasonal Box-Jenkins models are represented by SARIMA(p,d,q)(P,D,Q)$_s$, where p, d, q stand for short-term components and P, D, Q for seasonal components with length of seasonality s and a white noise sequence is ε_t,

$$\phi_p\,(B)\,\Phi_P\,(B^s)(1-B)^d(1-B^s)^D y_t = \theta_q(B)\Theta_Q(B^s)\varepsilon_t. \quad (7)$$

The autocorrelation function (ACF) and partial autocorrelation function (PACF) of the data are employed in Box-Jenkins modeling to determine the time series model's order, as indicated in Table (1).

Table 1. Selecting the Order of the Polynomials Model using ACF and PACF Figures

Model	ACF	PACF
AR(p)	Dies down (large spikes)	Cut off after lag p
MA(q)	Cut off after lag q	Dies down (large spikes)
ARMA (p, q)	Dies down (large spikes)	Dies down (large spikes)
$AR(P)_S$	Dies down (large spikes)	Cuts off after lag Ps
$MA(Q)_S$	Cuts off after lag Qs	Dies down (large spikes)
$ARMA\ (P, Q)_S$	Cuts off after lag Qs	Cuts off after lag Ps

The following actions can be taken regarding the Univariate Box Jenkins model:

- Stationarity Test: Use the ADF test to determine whether or not the time series is stationary. If it's not stationary, it must be made stationary by taking differences, seasonal differences, or transformations.
- ACF and PACF for the non-seasonal stationary data: The autoregressive (AR), integrated (I), and moving average (MA) terms in the SARIMA model are arranged in these plots.
- Seasonality: Select the appropriate seasonal period by looking for any seasonal patterns in the data. The seasonal part of the stationary data's ACF and PACF: The order of the seasonal autoregressive (SAR) and seasonal moving average (SMA) terms in the SARIMA model is determined by these plots (Hyndman & Athanasopoulos, 2018; Deretić et al., 2022).

3.2 Univariate GARCH Models

Engle (1982) introduced the Autoregressive Conditional Heteroskedasticity (ARCH) model, which has been used to model a change in variance in time series with multiple constraints so that the model can accurately estimate volatility. Equation (8) illustrates the ARCH(p) process in its general form.

$$\sigma_t^2 = \alpha + \sum_{i=1}^{p} \beta_i \varepsilon_{t-1}^2. \quad (8)$$

One of the weaknesses of ARCH models is that they assume that volatility is affected by both positive and negative shocks in the same way and that volatility only lasts for a short period of time unless *p* is large (Boudrioua & Boudrioua, 2020; Engle, 1982; Kumari & Tan, 2018).

Generalized Autoregressive Conditional Heteroskedasticity (GARCH) is an extension of this method that explains how to estimate changes in time-dependent volatility. Bollerslev (1986) suggests changing the ARCH model to the GARCH family, which has volatility that varies at random (Angelidis et al., 2004; Bollerslev, 1986; Shahraki & Alimardani, 2020).

Equation (9) illustrates the GARCH(p, q) model in its general form. Past volatility that could impact the present is incorporated into this model.

$$\sigma_t = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-1}^2 + \sum_{i=1}^{q} \beta_i \sigma_{t-1}^2. \quad (9)$$

The coefficients of ARCH and GARCH are α_i and β_i, where $\alpha > 0$, and $0 < \beta < 1$. The residuals' ACF and PACF aid in defining the GARCH orders, p and q. GARCH(p, q) = ARCH(p) if $q = 0$.

Dickey & Fuller (1979) determines whether $\phi = 0$ in the autoregressive (AR) time series model, as indicated by Equation (10), and tests the null hypothesis that a unit root exists in the model.

$$\Delta y_t = (y_t - y_{t-1}) = \alpha + \beta t + \gamma y_{t-1} + \varepsilon_t \quad (10)$$

The study variable is y_t, and the first difference operator is Δ. If $\gamma = 0$, then a random walk process is in place. If $-1 < \gamma + 1 < 1$, the process is stationary. However, the Augmented Dickey-Fuller test is applied to higher-order autoregressive processes by including Δy_{t-p} in the model. The null hypothesis asserts that the data are not stationary for both tests.

$$\Delta y_t = (y_t - y_{t-1}) = \alpha + \beta t + \gamma y_{t-1} + \delta_1 \Delta y_{t-1} + \delta_2 \Delta y_{t-2} + \cdots. \quad (11)$$

The normality of the time series data is examined using the Jarque-Bera test of skewness and kurtosis. There is never a negative test statistic. It indicates that the data does not have a normal distribution if it is far from zero. The definition of the test statistic is:

$$\frac{4NS^2 + 6N(K-3)^2}{24} \sim \chi^2. \quad (12)$$

where the numbers S, K, and N stand for skewness, kurtosis, and the number of observations, respectively. It is possible to model the variance change using an autoregressive process, like ARCH, if it can be correlated over time (Jarque & Bera, 1980, 1981, 1987).

The general procedure for a GARCH model can be summed up in three steps: estimating a best-fitting autoregressive model, calculating the error term's autocorrelations, and testing for significance.

GARCH models have limitations even though they are helpful in a variety of applications (Gourieroux, 1997; Magnus & Fosu, 2006):

- The best conditions for GARCH models, which are parametric specifications, are time series that are reasonably stable. The purpose of GARCH is to model conditional variances that change over time. GARCH models, however, frequently fall short in describing extremely irregular phenomena. These include erratic market swings and other unforeseen circumstances that have the potential to cause major structural change.
- The fat tails seen in asset return series are frequently not adequately captured by GARCH models. A portion of the fat-tail behavior can be explained by heteroscedasticity, but not all of it. Fat-tailed distributions, like Student's t, have been used in GARCH modeling to make up for this restriction.

According to the GARCH model, the conditional variance at time t is dependent on the series' historical squared errors (ε_2) and conditional variances (σ^2). The impact of previous squared errors and conditional variances on the current conditional variance is indicated by the α and β parameters, respectively. The series' long-term average level of variance is denoted by the ω term.

The steps that follow can be taken for the GARCH model:

- Remove any serial correlation from the data and create an ARIMA model for the stationary data. The ARCH effect can be verified using the model's residual series. To determine whether the data is conditionally heteroscedastic, the Box-Pierce (Ljung-Box) test is employed.
- ACF and PACF of the ARIMA model's squared residuals: The purpose of these plots is to determine the GARCH orders, r and s, respectively.

3.3 Hybrid Box-Jenkins–GARCH Models

A linear ARIMA model and the conditional variance of a GARCH model are combined to create the hybrid univariate Box-Jenkins–GARCH model, a nonlinear time series model. The maximum likelihood approach is the foundation of the GARCH and ARIMA models' estimation processes. The nonlinear Marquardt's algorithm is used to estimate parameters in the logarithmic likelihood function (Chen et al., 2011; Marquardt, 1963; Schwarz, 1978).

The proposed model combines the non-linear GARCH model with the linear time series ARIMA model. The suggested hybrid model of Box-Jenkins and GARCH has a two-stage process (Liu & Shi, 2013; Liu et al., 2013; Miswan, Ping, & Ahmad, 2013).

- The linear data of time series is modeled in the first stage using the best univariate Box-Jenkins model; the residual of this linear model will only include the nonlinear data.
- The nonlinear patterns of the residuals are modeled in the second stage using the GARCH.

In this process, it is said that the ARIMA model's error term follows a GARCH process of orders p and q. Equation (13) provides the following description of the form:

$$\sigma_t{}^2 = \alpha_0 + \sum_{i=1}^{q} \varphi_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \mu_i \sigma_{t-i}^2 \qquad (13)$$

Before beginning to estimate the GARCH model, it is necessary to investigate whether heteroscedasticity exists in time series data. The absence of serial correlation will show that GARCH can be used to estimate the conditional variance for the errors.

These two models are combined in the hybrid model to capture the time series data's volatility clustering and seasonal patterns. The time series' mean is modeled by the univariate Box-Jenkins component, and its volatility is modeled by the GARCH component.

The variance of the univariate Box-Jenkins model's error term follows a GARCH process in the SARIMA-GARCH model. Equation (14) can be used to express the model. The following is the standard expression for the GARCH formula:

$$S_t = \mu + \epsilon_t, \quad \epsilon_t = \sigma_t \varepsilon_t, \; \sigma_t^2 = \omega + \sum_{i=1}^{r} \alpha_i \, \varepsilon_{t-i}^2 + \sum_{j=1}^{S} \beta_j \, \sigma_{t-j}^2 \qquad (14)$$

S_t and ϵ_t represent the stationary time series data and random error at time t, respectively, and μ represents the conditional mean of S_t. The conditional variance of the error term at time t is denoted by σ_t^2, while the long-term average variance of the error term is represented by the constant term t , ω. The residual error at time t, denoted by ϵ_t, has a continuous distribution with zero-mean I.I.D. The lagged squared error terms' coefficients are α_i, where $\alpha_0 = 0$, and $\alpha_i \geq 0$ for $i = 1, 2, \ldots, r$. The coefficients for the lagged conditional variance terms are β_j, where $\beta_0 = 0$ and $\beta_j \geq 0$ for $j = 1, 2, \ldots, s$.

These two models are combined in the hybrid model to capture the time series data's volatility clustering and seasonal patterns. The time series' mean is modeled by the SARIMA component, and its volatility is modeled by the GARCH component. The variance of the SARIMA model's error term follows a GARCH process in the SARIMA-GARCH model. Equation (15), and (16) can be used to express the model.

$$\Phi_P(B^S)\varphi_p(B)(1-B)^d(1-B^s)^D \acute{y}_t = \Theta_Q(B^S)\,\theta_q(B)\varepsilon_t \tag{15}$$

$$\epsilon_t = \sigma_t\varepsilon_t,\ \ \sigma_t^2 = \omega + \sum_{i=1}^{r}\alpha_i\ \varepsilon_{t-i}^2 + \sum_{j=1}^{S}\beta_j\,\sigma_{t-j}^2 \tag{16}$$

4. The Proposal Frame

4.1 Data Set and Software

- Data Set: the study uses the monthly Saudi General Index for the period January 2015 to June 2024, https://sa.investing.com. The database is divided into 109 observations in sample and 6 observations out of sample data.
- Software: the study uses SPSS, MINITAB, and R, which are currently used on applications.

4.2 The Suggested Frame

The study integrates the GARCH and SARIMA models to propose a framework for modeling financial time series. The univariate series is analyzed, and the approximation values are predicted using this hybrid model, which combines a SARIMA model with GARCH error components. The ARIMA model's error term ε_t is said to follow a GARCH process of orders r and s in this process.

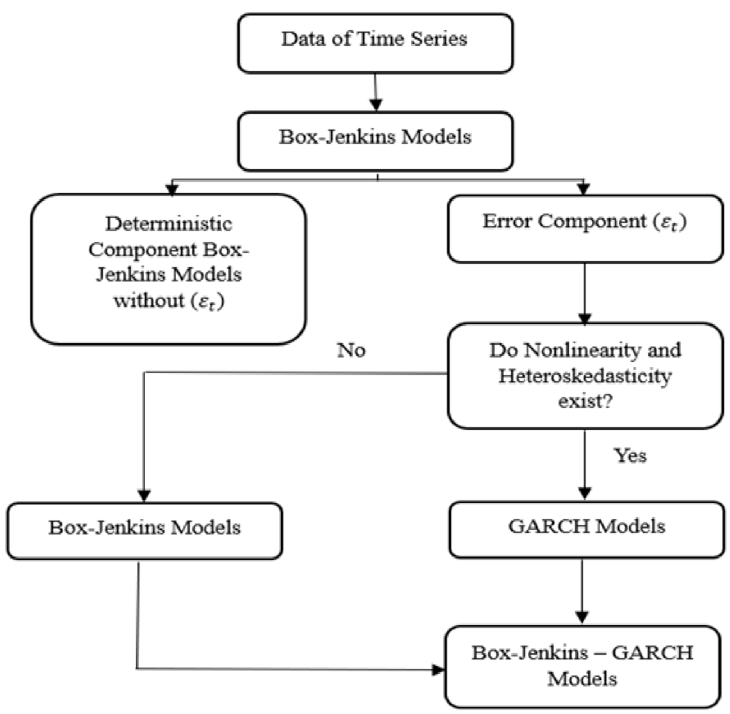


Figure 1. Frame of Hybridization of Box-Jenkins and GARCH Models

The suggested hybrid model combines GARCH and SARIMA in a two-phase process.

1. The first step involves modeling the linear data of time series using the best SARIMA models; only the nonlinear data will be included in the linear model's residual.
2. The second step involves modeling the residuals' nonlinear patterns using the GARCH.

The following steps make up the parameter estimation procedure for the SARIMA-GARCH model:

1. Analysis of Stationarity and Seasonality: To ascertain whether the time series is stationary and shows any seasonal patterns, apply statistical tests such as the ADF test and the Seasonal Decomposition of the Time Series method, respectively.

2. SARIMA Parameter Estimation: Use MLE to estimate the parameters of the model's SARIMA component. This entails choosing the proper sequences for the model's AR, MA, SAR, and SMA components.

3. Finding the best models involves comparing the performance of suggested models using MSE, MAE, MAPE, AIC, and BIC.

4. Estimating GARCH Parameters: Use MLE to estimate the parameters of the model's GARCH component. To do this, the model's ARCH and GARCH components must be chosen in the proper order.

Model Selection: Select the appropriate SARIMA-GARCH model based on the AIC and BIC as proposed by Equation (17) and (18), respectively. The AIC or BIC with the lowest value is preferred in the model selection criteria. Model selection is done using the Bayesian information criterion (BIC) and the Akaike information criterion (AIC). An assessment of a model's fit to the data it was created from is done mathematically using the AIC.

$$AIC = 2K - 2ln(L) \tag{17}$$

The number of independent variables used is denoted by K, and the log-likelihood estimate, or the probability that the model could have generated your observed y-values, is determined by L.

$$BIC = -2\, ln\,(L) + K\, ln(N) \tag{18}$$

The best model for forecasting future observations can be found using the AIC, whereas the BIC is more helpful in choosing the right model. A crucial step in the modeling process is diagnostic checking, which enables us to assess the model's suitability and spot any possible flaws.

Diagnostic checking is crucial for SARIMA-GARCH models to ensure accurate representation of time series and volatility components, incorporating serial correlation, heteroscedasticity, and normality tests to ensure white noise behavior. The model residuals' ACF and PACF can be plotted, and the Box-Pierce (Ljung-Box) Chi-Square statistic test can be used to determine whether the residuals exhibit serial correlation (Garai et al., 2023; Hillmer & Tiao, 1982).

Forecast accuracy, which is frequently represented as a percentage or numerical score, gauges how well a model predicts future events. It needs to be tested and updated frequently to stay relevant. Equations (19), (20), and (21) provide the MAE (Mean Absolute Error), RMSE (Root Mean Squared Error), and MAPE (Mean Absolute Percentage Error), which are common indicators of forecast accuracy.

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \tag{19}$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{n}} \tag{20}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \tag{21}$$

The number of out-of-sample data is denoted by n, while y_t and t represent the observed and predicted values at time t, respectively. The model that produces the smallest prediction error is the most effective forecasting model.

5. Results and discussion

The first step in the identification process is to see if the Saudi General Index time series movement exhibits seasonality and an upward or downward trend, as illustrated in Figure (1).

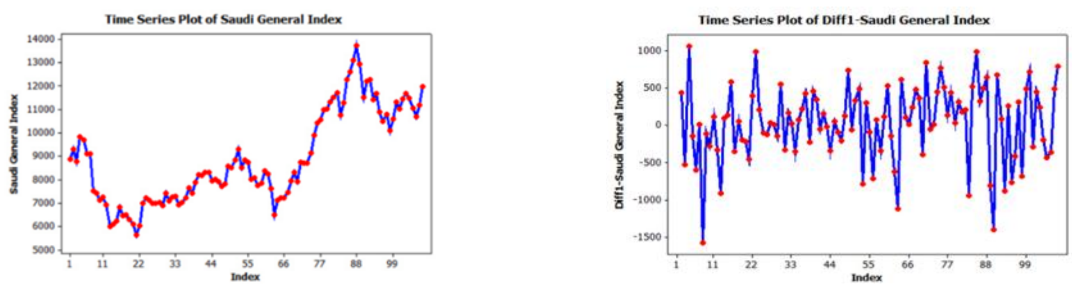


Figure 2. (a) Plot of the time series of monthly Saudi General Index (b) Plot of the first different of monthly Saudi General Index

The Saudi General Index series does not fluctuate around a fixed level, as seen in Figure (2, a), indicating that its mean and seasonal trend are non-stationary. Following a one-lag difference to the transformed series, the series becomes stationary. As seen in Figure (2, b), the transformed series is visible.

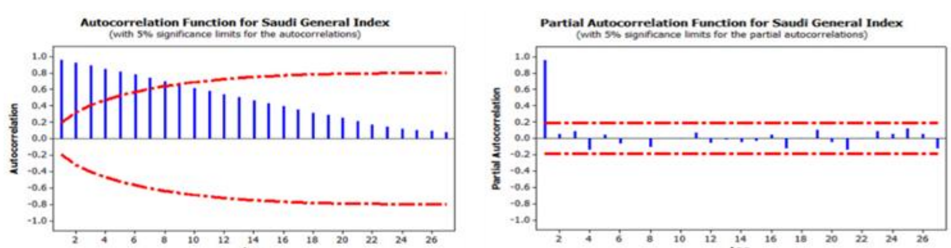


Figure 3. ACF and PACF for the Time Series of Monthly Saudi General Index

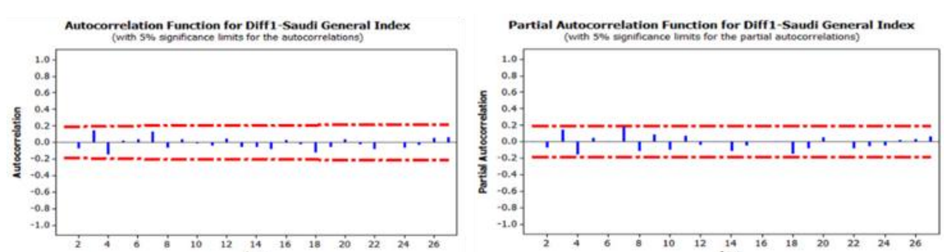


Figure 4. ACF and PACF for the First Differenced Transformed Series

Since the first difference of the time series data passed the Shapiro-Wilk normality test with W = 0.9798 and a p-value of 0.08513, we are unable to rule out the null hypothesis that the sample was drawn from a normal distribution.

Table 2. Stationarity Tests in Time Series Data

Box-Pierce test		
X-Squared	D.F.	P-Value
105.8	1	2.2e-16
Augmented Dickey-Fuller test		
Dicky-Fuller	Lag Order	P-Value
-3.4606	4	0.04889
Phillips-Perron Unit Root Test		
Dicky-Fuller Z (alpha)	Truncation Lag Parameter	P-Value
-111.413	4	0.4553

Phillips-Perron Unit Root, Box-Pierce, and Augmented Dickey-Fuller tests are displayed in Table (2). Data differencing is required because these tests show that the null hypothesis is not rejected, indicating that the unit root is present in the data series. Figures (3) and (4) show that **the seasonal trend is confirmed by the ACF and PACF spikes on one side.**

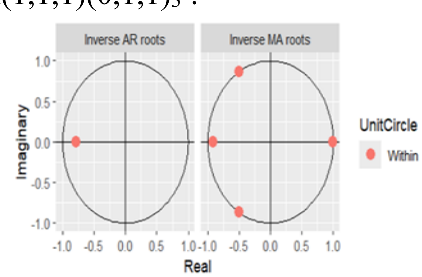
5.1 Univariate Box Jenkins Models

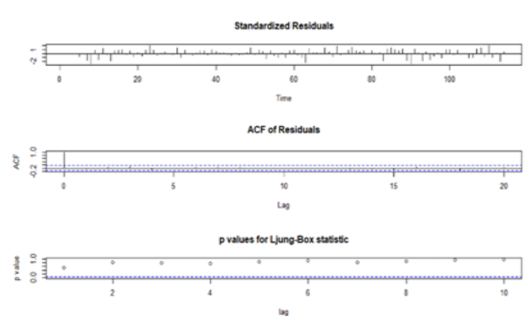
Table (3) displays the estimation results using the ordinary least squares method at a significance level of 0.05 with values for the Akaike and Schwarz information criteria. SARIMA$(1,1,1)(0,1,1)_3$, SARIMA$(1,1,1)(1,2,1)_6$, and SARIMA$(2,2,0)(2,1,1)_4$ are the best SARIMA models for the observed series, as they have the lowest forecasting error metrics.

Table 3. Forecasting Error Metrics of Significant SARIMA Models

Model	ME	RMSE	MAE	MPE	MAPE	MASE	AIC
SARIMA$(1,1,1)(0,1,1)_3$	32.103	477.5	374.26	0.3325	4.275	0.965	1688.5
SARIMA$(1,1,1)(1,2,1)_3$	20.406	553.0	401.75	0.4070	4.630	1.036	1691.4
SARIMA$(1,1,1)(2,1,0)_3$	24.218	532.8	399.85	0.2531	4.577	1.031	1711.1
SARIMA$(1,1,1)(1,2,1)_6$	3.7931	553.1	406.22	0.1604	4.675	4.675	1614.1
SARIMA$(1,2,0)(2,1,1)_4$	9.2376	587.9	469.21	0.12131	5.320	1.210	1717.9
SARIMA$(2,2,0)(2,1,1)_4$	16.909	524.6	409.51	0.25044	4.762	1.056	1697.2
SARIMA$(1,2,0)(1,1,1)_4$	11.004	585.9	460.06	0.13881	5.238	1.187	1719.0
SARIMA$(2,2,0)(1,1,1)_4$	19.863	522.7	397.57	0.28152	4.630	1.026	1697.6

SARIMA$(1,1,1)(0,1,1)_3$:





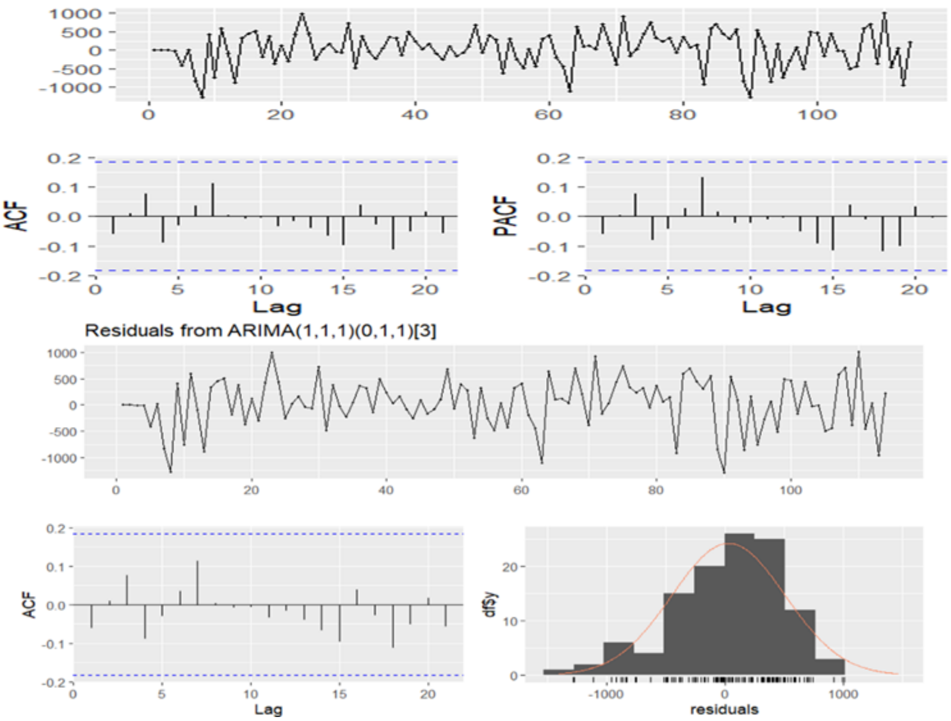


Figure 5. The Characteristic Roots and ACF and PACF Residuals of the for SARIMA(1,1,1)(0,1,1)$_3$

The residuals plots for SARIMA(1,1,1)(0,1,1)$_3$ are displayed in Figure 5, along with the ACF and PACF of residuals. These spikes are not significant, suggesting that there is no correlation between the residuals.

Table 4. Final Estimates of Parameters of SARIMA(1,1,1)(0,1,1)$_3$ Model

Type	Coef	SE Coef	T	P
AR	-0.7521	0.0903	-8.30	0.000
MA	-0.9407	0.0903	-25.18	0.000
SMA	0.9534	0.0452	21.11	0.000

The final estimates for the SARIMA(1,1,1)(0,1,1)$_3$ parameter are shown in Table (4); the parameter's P-value is less than 0.05. Thus, the model fits the data well, as evidenced by the parameters' large deviation from zero.

Table 5. Modified Box-Pierce (Ljung-Box) Chi-Square statistic for testing the conditional heteroscedasticity SARIMA(1,1,1)(0,1,1)$_3$ Model

Modified Box-Pierce test				
lag	12	24	36	48
Chi-Square	2.5	13	18.2	29.7
DF	9	21	33	45
P-Value	0.98	0.90	0.98	0.96
Ljung-Box test				
	Q* = 3.8968	d.f = 7	p-value = 0.7916	

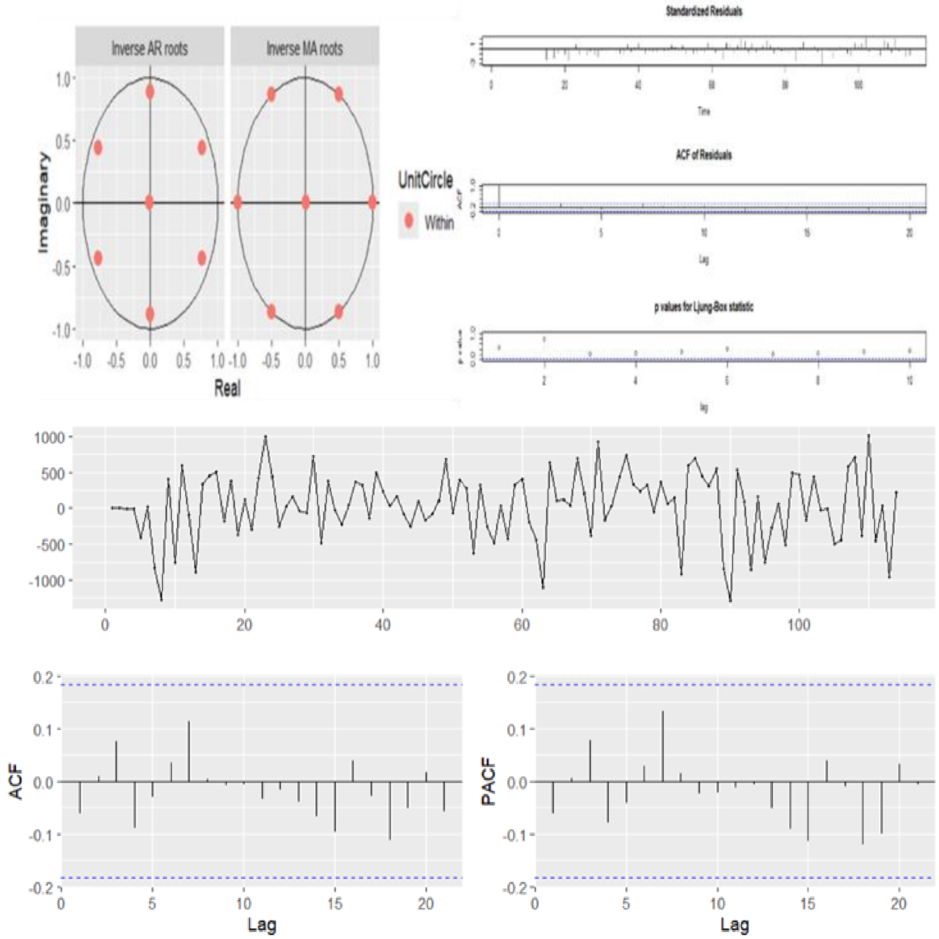
Table (5) displays the SARIMA(1,1,1)(0,1,1)$_3$ model's modified Box-Pierce (Ljung-Box) Chi-Square statistic. With a P-value greater than 0.05, the Box-Pierce (Ljung-Box) Chi-Square test is not significant. The residuals thus seem to be uncorrelated. Since the model's parameters deviate significantly from zero and the residuals show no correlation, the SARIMA(1,1,1)(0,1,1)$_3$ model seems to fit well and can be used to forecast. LM-test for ARCH; null hypothesis: no ARCH effects. df = 1, p-value = 0.2493, chi-squared = 1.3275. Thus, the null hypothesis that there is no ARCH effect cannot be rejected.

Table 6. The comparison between actual Saudi General Index and forecast price for an out-sample period by SARIMA$(1,1,1)(0,1,1)_3$

Month	Actual Index	Forecast Index
01-2024	11796.63	12272.5
02-2024	12630.86	12068.2
03-2024	12401.56	12166.1
04-2024	12394.91	12353.9
05-2024	11503.49	12237.5
06-2024	11498.93	12269.5

The results of forecasting the Saudi General Index from January to June 2024 using SARIMA$(1,1,1)(0,1,1)_3$ are displayed in Table (6).

SARIMA$(1,1,1)(1,2,1)_6$



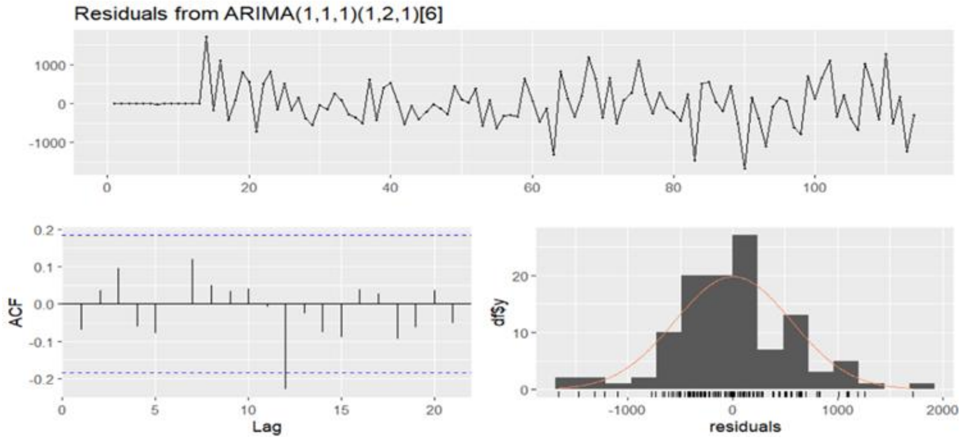


Figure 6. The Characteristic Roots and ACF and PACF Residuals of the for SARIMA(1,1,1)(1,2,1)$_6$

SARIMA(1,1,1)(1,2,1)$_6$ and ACF and PACF residual plots are shown in Figure (6), where nonsignificant spikes signify uncorrelated residuals.

Table 7. Final Estimates of Parameters of the SARIMA(1,1,1)(1,2,1)$_6$

Type	Coef	SE Coef	T	P
AR1	-0.6929	0.0917	-7.55	0.000
SAR6	-0.5460	0.0915	-5.97	0.000
MA1	-0.9488	0.0461	-20.58	0.000
SMA6	0.8994	0.0792	11.35	0.000

With uncorrelated residuals and significant parameter deviations, the SARIMA(1,1,1)(1,2,1)$_6$ model has a P-value less than 0.05, as shown in Table (7).

Table 8. Modified Box-Pierce (Ljung-Box) Chi-Square statistic of the check the conditional heteroscedasticity SARIMA(1,1,1)(1,2,1)$_6$

lag	12	24	36	48
Chi-Square	11.5	22.7	31.4	42.9
DF	8	20	32	44
P-Value	0.175	0.302	0.496	0.518
Ljung-Box test				
Q* = 5.4705	d.f = 6		p-value = 0.485	
ARCH LM-test				
Chi-squared = 0.58498	d.f = 1		p-value = 0.4444	

Nonsignificant Ljung-Box statistics, which show uncorrelated residuals as discrete white noise, are shown in Table (8). The null hypothesis is supported by the p-values of the Ljung-Box and Lagrange multiplier tests, which show that the data series does not exhibit an ARCH effect.

Table 9. The comparison between actual price and forecast price for an out-sample period by SARIMA(1,1,1)(1,2,1)$_6$

Month	Actual Index	Forecast Index
01-2024	11796.6	12113
02-2024	12630.9	11685.7
03-2024	12401.6	11709.3
04-2024	12394.9	11928.7
05-2024	11503.5	12036
06-2024	11498.9	12853.9

SARIMA(2,2, 0)$(2,1,1)_4$

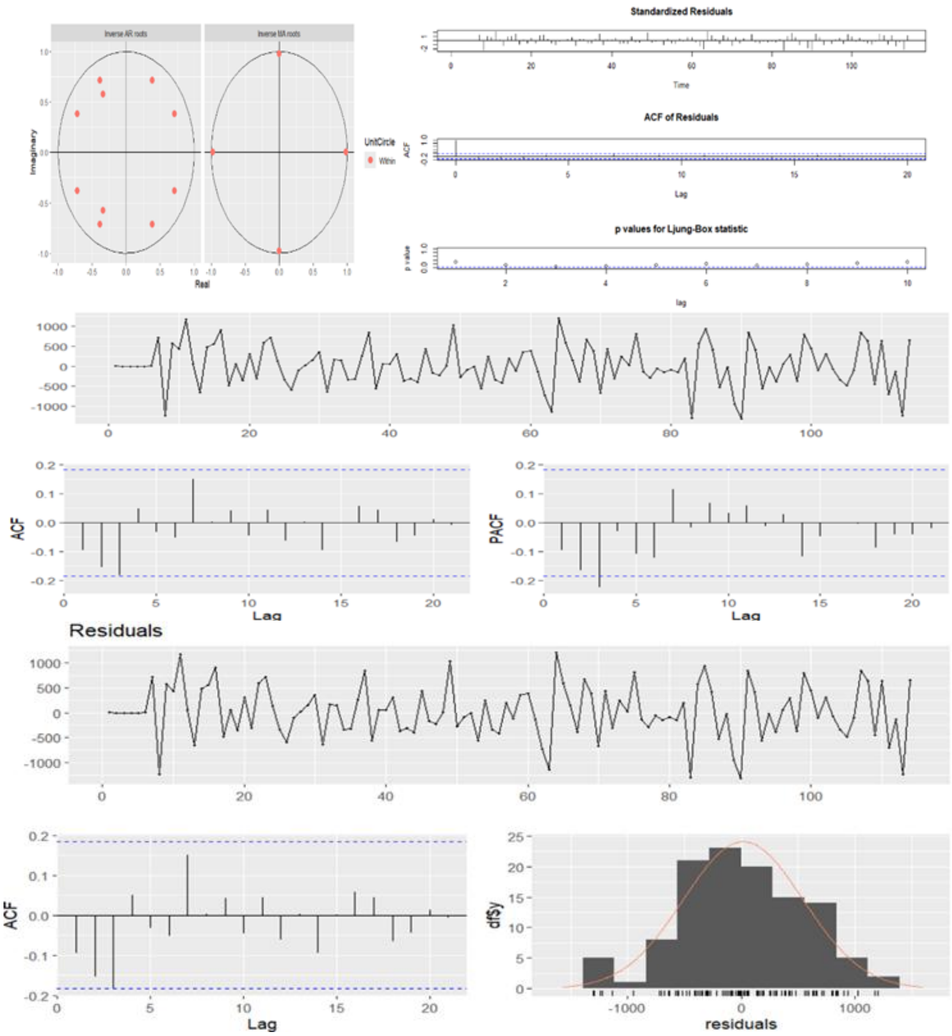


Figure 7. The Characteristic Roots and ACF and PACF Residuals of the for SARIMA(2,2,0)$(2,1,1)_4$

Table 10. Final Estimates of Parameters of the SARIMA(2,2,0)$(2,1,1)_4$

Type	Coef	SE Coef	T	P
AR1	-0.6346	0.0920	-6.90	0.000
AR2	-0.4962	0.0882	-5.62	0.000
SAR4	-0.4356	0.1034	-4.21	0.000
SAR8	-0.2952	0.1035	-2.85	0.005
SMA4	0.9234	0.0558	16.54	0.000

Table (10) displays the results of OLS estimation at 0.05 significance using the Schwarz information criterion and Akaike information criterion, which indicate that the SARIMA model is significant (SARIMA(2,2,0)$(2,1,1)_4$).

Table 11. Modified Box-Pierce (Ljung-Box) Chi-Square statistic of the check the conditional heteroscedasticity SARIMA(2,2,0)$(2,1,1)_4$

lag	12	24	36	48
Chi-Square	12.9	17.3	32.8	41.6
DF	7	19	31	43

P-Value	0.074	0.572	0.377	0.531

Ljung-Box test		
Q* = 11.881	d.f = 10	p-value = 0.293
ARCH LM-test		
Chi-squared = 4.3051	d.f = 1	p-value = 0.038

Furthermore, Table (11) demonstrates that the Ljung-Box statistics yield nonsignificant results (P-value > 0.05), suggesting that the residuals are uncorrelated and represent discrete white noise. Null hypothesis: no ARCH effects; ARCH LM-test Since chi-squared = 4.3051, d.f. = 1, and p-value = 0.038, the null hypothesis is rejected. When conditional variance in a data series is not constant over time, the ARCH effect takes place.

The Box-Pierce (Ljung-Box) test indicates that the SARIMA(2,2,0)(2,1,1)$_4$ model is heteroscedastic, as shown in Table (11). The model needs to be used in conjunction with the GARCH model in order to control the volatility in the data series. The Jarque-Bera test is used for the normalcy test. The assumption of normalcy has been satisfied when the p-value is greater than 0.05. However, because the model's p-value is less than 0.005, the assumption of normalcy is not satisfied. The hybrid ARIMA-GARCH approach was used to compare the residuals of the ARIMA models with those of the GARCH models. To model the variance behavior in our study, we used the conventional GARCH (1,1) model.

5.2 Combined Box-Jenkins and GARCH Models

Univariate Box Jenkins models, like ARIMA models, explain the linear portion of the time series, while GARCH models, which are derived from the residual series of an ARIMA model, explain the nonlinear features. The ARIMA model under consideration was unable to address the heteroscedasticity present in the data series, and the ARCH effect was observed in data series where conditional variance was not constant over time. The time series data was first fitted using Box-Jenkins models. If there is an ARCH effect in the residuals, GARCH is fitted.

Table12. Estimation results for SARIMA(2,2,0)(2,1,1)$_4$-GARCH(1, 1)

Parameter	Value	Std Error	T-Statistic	Pr(>|t|)
mu	19.0010	24.10481	3.268675	0.430539
AR1	0.56255	0.172104	3.268675	0.001081
MA1	-0.7827	0.126555	-6.184658	0.000000
Omega	278.068	7312.340	0.038027	0.969666
Alpha1	0.00000	0.029981	0.000001	1.000000
Beta1	0.99900	0.004298	232.422998	0.000000
Information Criteria	**Akaike**	**Bayes**	**Shibata**	**Hannan-Quinn**
	15.399	15.543	15.394	15.458

The results of estimating SARIMA(2,2,0)(2,1,1)$_4$-GARCH(1,1) using maximum likelihood estimation are displayed in Table (12). For all symmetric GARCH models taken into consideration, the results were found to be statistically significant and preferred at the 0.05 significance level.

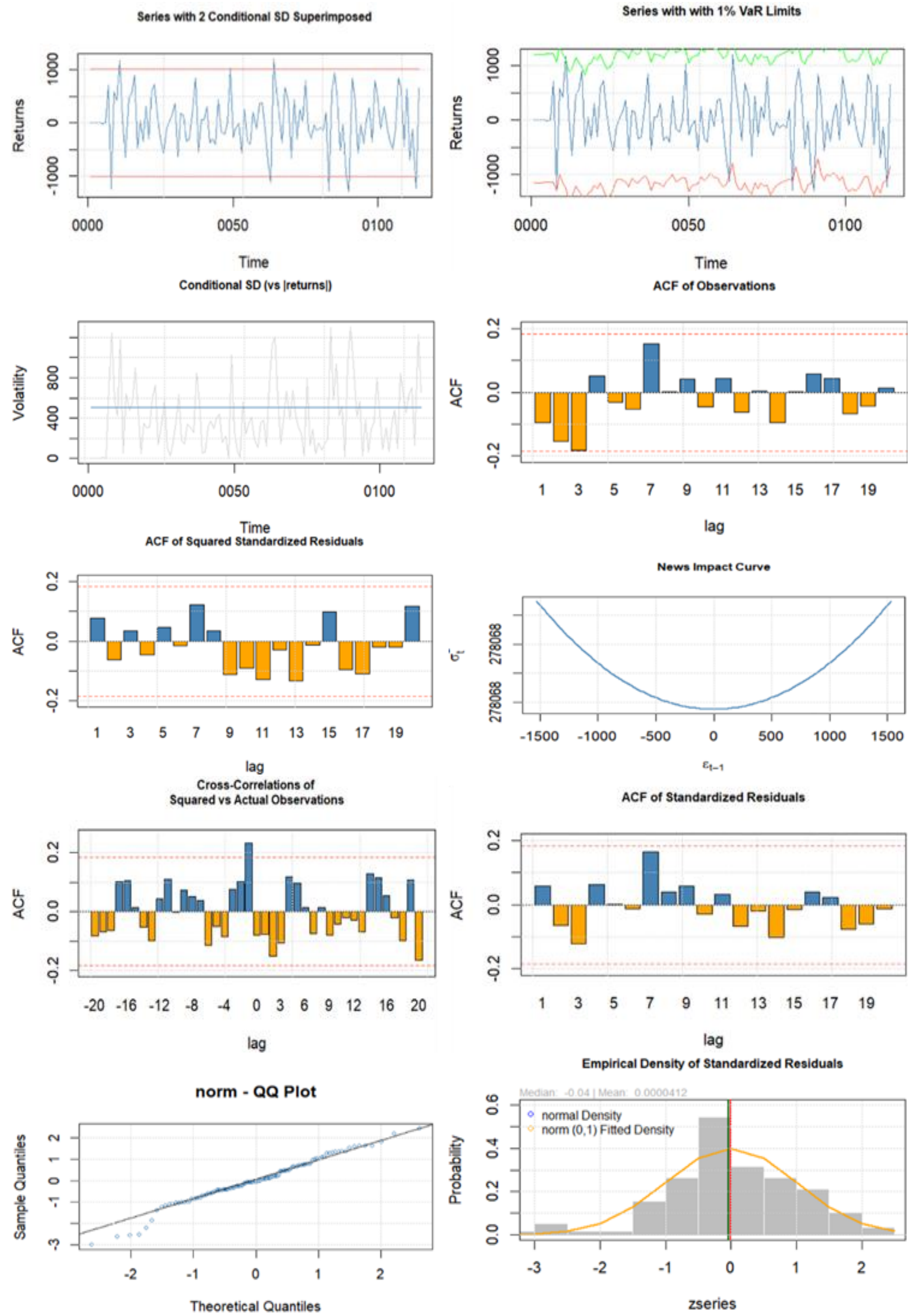


Figure 8. Analysis of Financial Time Series and Residuals: Volatility, Autocorrelations, and Distributional Properties

Table 13. Forecasting Error Metrics of Significant SARIMA Models

Model	ME	RMSE	MAE	MPE	MAPE	MASE
SARIMA(1,1,1)(0,1,1)$_3$	32.103	477.55	374.26	0.3325	4.2756	0.9659
SARIMA(1,1,1)(1,2,1)$_6$	3.7931	553.14	406.22	0.3604	4.6751	4.6751
SARIMA(2,2,0)(2,1,1)$_4$	16.909	524.64	409.51	0.2504	4.7627	1.0569
SARIMA(2,2,0)(2,1,1)$_4$-GARCH(1,1)	1.5769	57.351	38.228	0.2062	4.2474	1.0001

Table 14. The comparison between actual Saudi General Index and forecast price for out-sample period by Hybrid Model.

Month	Actual Index	Forecast Index
01-2024	11796.63	11967.39
02-2024	12630.86	12796.63
03-2024	12401.56	12630.86
04-2024	12394.91	12401.56
05-2024	11503.49	11994.91
06-2024	11498.93	11669.05

The forecast results are based on three evaluation criteria that have been widely used in prior literature: mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE). The series of out-of-sample transformed data, which comprises six observations, is used in the forecasting stage. MAE, RMSE, and MAPE forecast evaluations are 38.2284, 57.35, and 4.247, respectively.

6. Conclusion and Remark

The goal of this study is to improve the effectiveness of financial time series forecasting by combining univariate Box-Jenkins and GARCH models. The empirical findings show that the linear limitation of the SARIMA and GARCH models can be overcome by the combining model of SARIMA(2,2,0)(2,1,1)$_4$-GARCH(1,1). The Box-Jenkins-GARCH combination is a novel method for financial time series forecasting since it combines the ability of GARCH to handle volatility with the ability of the Box-Jenkins model to generate forecasts based on historical patterns. Initially, the ARIMA model fits the time series data. GARCH is fitted if the residuals exhibit the ARCH effect. The final prediction is derived from the aforementioned condition.

Funding: This study did not receive any funding

Declaration of interest: The authors declare that they have no conflict of interest.

References

Akaike, H. (1974). A new look at the statistical model identification, *EEE Transactions on Automatic Control,* 19(6): 716-723. https://doi.org/10.1109/TAC.1974.1100705

Angelidis, T., Koutoupis, A., & Milonas, N. (2004). The use of GARCH models in VaR estimation. *Statistical Methodology*, 1(1-2): 105-128. https://doi.org/10.1016/j.stamet.2004.08.004

Adewole, A. I. (2024). On the hybrid of ARIMA and GARCH model in modelling volatilities in Nigeria Stock Exchange. *Bima Journal of Science and Technology, 8*(2A): 169-180. https://doi.org/10.56892/bima.v8i1.601

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations, *Journal of the Royal Statistical Society*, Series B, 26: 211–252. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). Time Series Analysis: *Forecasting and Control* (5th ed.). Wiley. https://doi.org/10.1002/9781118619193

Boudrioua, M. S., & Boudrioua, A. (2020). Modeling and forecasting the Algerian stock exchange using the Box-Jenkins methodology. *Journal of Economics, Finance and Accounting Studies,* 2(1): 1–15. https://doi.org/10.5281/zenodo.3903241

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31(3): 307–328. https://doi.org/10.1016/0304-4076(86)90063-1

Bollerslev, T. & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Reviews*, 11(2): 143–172. https://doi.org/10.1080/07474939208800229

Chen, C., Hu, J., Meng, Q. & Zhang, Y. (2011). Short-time traffic flow prediction with ARIMA-GARCH model. *2011 IEEE Intelligent Vehicles Symposium (IV)*, Baden-Baden, Germany, June 5-9. https://doi.org/ 10.1109/IVS.2011.5940418

Deretić, N., Stanimirović, D., Awadh, M. A., Vujanović, N., & Djukić, A. (2022). SARIMA modelling approach for forecasting of traffic accidents. *Sustainability, 14*(8), 4403. https://doi.org/10.3390/su14084403

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366): 427–431. https://doi.org/10.2307/2286348

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of variance of U.K. inflation. *Econometrica,* 50: 987–1008. https://doi.org/10.1016/0304-4076(86)90063-1

Gourieroux, C. (1997). *ARCH models and financial applications*. Springer-Verlag. https://doi.org/10.1007/978-1-4612-1860-9.

Girish, G. P. (2016). Spot electricity price forecasting in Indian electricity market using autoregressive- GARCH models. *Energy Strategy Reviews, 11*(12): 52–57. https://doi.org/10.1016/j.esr.2016.06.005

Garai, S., Paul, R. K., Rakshit, D., Yeasin, M., Paul, A. K., Roy, H. S., Barman, S., & Manjunatha, B. (2023). An MRA based MLR model for forecasting Indian annual rainfall using large scale climate indices. *International Journal of Environment and Climate Change,* 13(5): 137–150. https://doi.org/10.9734/IJECC/2023/v13i51755

Hillmer, S. C., & Tiao, G. C. (1982). An ARIMA-model based approach to seasonal adjustment. *Journal of the American Statistical Association,* 77(377): 63–70. https://doi.org/10.2307/2287770

Hong, J., Yan, Y., Kuruoglu, E. E., & Chan, W. K. (2023). Multivariate time series forecasting with GARCH models on graphs. *IEEE Transactions on Signal and Information Processing Over Network*s. https://doi.org/10.1109/TKDE.2022.3221989

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. (2nd ed.) OTexts. https://otexts.org/fpp2/

Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity, and serial independence of regression residuals. *Economics Letters,* 6(3): 255–259. https://doi.org/10.1016/0165-1765(80)90024-5

Jarque, C. M., & Bera, A. K. (1981). Efficient tests for normality, homoscedasticity, and serial independence of regression residuals: Monte Carlo evidence. *Economics Letters,* 7(4): 313–318. https://doi.org/10.1016/0165-1765(80)90024-5

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistical Review, 55*(2): 163–172. https://doi.org/10.2307/1403192

Kumari, S. N., & Tan, A. (2018). Modeling and forecasting volatility series: With reference to gold price. *Thai Journal of Statistical Science,* 16(1): 77–93. https://doi.org/10.1088/1757-899X/548/1/012023

Liu, H., & Shi, J. (2013). Applying ARMA-GARCH approaches to forecasting short-term electricity prices. *Energy Economics,* 37: 152–166. https://doi.org/10.1016/j.eneco.2013.02.006

Liu, H., Shi, J., & Qu, X. (2013). Empirical investigation on using wind speed volatility to estimate the operation probability and power output of wind turbines. *Energy Conversion and Management,* 67: 8–17. https://doi.org/10.1016/j.enconman.2012.10.016

Magnus, F. J., & Fosu, O. (2006). Modeling and forecasting volatility of returns on the Ghana Stock Exchange using GARCH models. *Journal of Applied Sciences, 3*(3): 2042–2048. https://doi.org/10.2139/ssrn.927168

Miswan, N. H., Ping, P. Y., & Ahmad, M. H. (2013). On parameter estimation for Malaysian gold prices modelling and forecasting. *International Journal of Mathematical Analysis,* 7(20): 1059–1068. https://doi.org/10.12988/ijma.2013.13106

- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and applications*. John Wiley & Sons. https://doi.org/10.1057/jors.1984.11

Marquardt, D. W. (1963). An algorithm for least squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics, 11*(2): 431–441. https://doi.org/10.1137/0111030

Nelson, D. (1991). Conditional Heteroscedasticity in Asset Return: A New Approach, Econometrica, 59: 347–370. https://doi.org/10.2307/2938260

Qasim, T. B., Ali, H., Malik, N. and Liaquat, M. (2021). Forecasting Inflation Applying ARIMA Model with GARCH Innovation: The Case of Pakistan. *Journal of Accounting and Finance in Emerging Economies*, 7(2): 313–332. https://doi.org/10.26710/JAFEE.V7I2.1681

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics,* 6(2): 461–464. https://doi.org/10.1214/aos/1176344136

Shahraki, N., & Alimardani, M. (2020). Application of ARIMA models in forecasting agricultural production. *Journal of Agricultural Economics, 51*(2): 256–267. https://doi.org/10.1214/aos/1176344136

Wang, B., Huang, H., & Wang, X. (2012). A novel text mining approach to financial time series forecasting. *Neurocomputing*, 83: 136–145. https://doi.org/10.1016/j.neucom.2012.04.004