BAREKENG: Journal of Mathematics and Its Applications
March 2026 Volume 20 Issue 1 Page 0271–0286
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss1pp0271-0286

PREDICTION OF THE INDONESIA COMPOSITE INDEX (ICI) USING THE ARCH GARCH APPROACH AND THE FOURIER SERIES

M. Fariz Fadillah Mardianto[1*], Hanny Valida[2], Farah Fauziah Putri[3], Doni Muhammad Fauzi[4], Elly Pusporani[5]

[1,2,3,4,5]*Statistics Study Program, Faculty of Science and Technology, Universitas Airlangga*
Jln. Dr. Ir. H. Soekarno, Mulyorejo, Surabaya, 60115, Indonesia

*Corresponding author's e-mail: * **m.fariz.fadillah.m@fst.unair.ac.id***

Article Info

Article History:
Received: 8th March 2025
Revised: 13th April 2025
Accepted: 25th July 2025
Available online: 24th November 2025

Keywords:
ARCH;
Fourier series;
GARCH;
Indonesia composite index;
Investor;
Prediction.

ABSTRACT

The Indonesia Composite Index (ICI) is a key indicator of stock market performance in Indonesia, often experiencing high volatility due to various domestic and global economic factors. In recent years, ICI has shown a significant upward trend, influenced by both local and international factors. In 2024, from June to October, the ICI saw a notable increase, reaching its highest value since 2020 at Rp 7,670. Despite fluctuations in stock prices, the rise in ICI reflects a positive outlook for the Indonesian stock market, attracting both domestic and foreign investors. This study aims to predict ICI movements using ARIMA-GARCH and Fourier Series approaches. The ARIMA model is employed to analyze time series data, while the ARCH-GARCH model addresses heteroskedasticity in residual variance. For comparison, the Fourier Series Estimator is applied to capture seasonal patterns in the data. Although ICI volatility is driven by a range of external macroeconomic and geopolitical factors, this study focuses on univariate modeling to evaluate the predictive capability of the index's own historical movements, without involving exogenous variables. The data used comes from Investing.com. Weekly ICI data from March 2020 to June 2024 is used, split into training and testing sets. The analysis results indicate that the ARIMA-GARCH method provides higher accuracy, with a Mean Absolute Percentage Error (MAPE) of 5% (out-sample), compared to the Fourier Series method, which has a MAPE of 8.57%. This suggests that ARIMA-GARCH is more effective in predicting ICI trends, reflecting its ability to account for volatility and market changes more accurately.

How to cite this article:

M. F. F. Mardianto, H. Valida, F. F. Putri, D. M. Fauzi and E. Pusporani., "PREDICTION OF THE INDONESIA COMPOSITE INDEX (ICI) USING THE ARCH GARCH APPROACH AND THE FOURIER SERIES," *BAREKENG: J. Math. & App.*, vol. 20, iss. 1, pp. 0271-0286, Mar, 2026.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

The Indonesia Composite Index (ICI) is a key indicator that reflects the overall stock market performance listed on the Indonesia Stock Exchange (IDX). First introduced in 1983, ICI initially included 13 stocks and now lists 778 stocks [1]. As a major parameter to measure the Indonesian stock market's condition, ICI reflects market sentiment and investment activities in Indonesia, both from domestic and foreign investors [2]. In recent years, ICI has shown a significant upward trend, influenced by various domestic and global economic factors. Although ARCH and GARCH models have been widely applied in modeling and forecasting financial market volatility due to their ability to capture time-varying variance, most studies focus solely on these parametric approaches. Comparisons with alternative nonparametric or semi-parametric methods remain relatively rare, especially in the context of the Indonesian market. One such alternative is the Fourier Series method, which is capable of modeling periodic patterns and capturing complex structures in time series data without assuming a specific distribution form.

According to CNBC, in 2023, the ICI experienced very high fluctuations but ended with a significant upward trend [3]. In March 2020, ICI dropped sharply to Rp 4194.94, its lowest level since 2020. However, it started to rise again from May 2020 through December 2024, reaching its highest value in September 2024 at Rp 7812.13. Despite fluctuations, the overall rise in ICI provides a positive outlook on Indonesia's stock market [4]. This not only attracts domestic investors but also foreign investors, which benefits companies in Indonesia.

The uncertainty in the capital market, particularly in the fluctuations of the Indonesia Composite Index (ICI), has become a major concern for investors, regulators, and economists due to its significant impact on economic stability and financial decision-making. As a result, accurately predicting ICI has become both a challenge and a necessity, especially for investors looking to minimize risk and maximize potential returns.

Traditional approaches to predicting ICI often rely on linear models that cannot fully capture the non-linear dynamics of market fluctuations. The high volatility of stock indices leads to heteroscedasticity in the data [5]. To overcome this limitation, Autoregressive Conditional Heteroskedasticity (ARCH) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models are used [6]. These models excel in capturing dynamic volatility patterns, enabling more accurate predictions that are relevant to unstable market conditions. In addition to the ARCH and GARCH models, a predictive approach using Nonparametric Regression with the Fourier Series method is also employed. Nonparametric regression is flexible in determining its curve estimation patterns [7].

Previous research shows good prediction results for ICI volatility using the GARCH method with a model goodness test (MAPE) of 17.26% for the 2012-2022 period [5]. This is in line with other studies which show that the GARCH model is able to forecast the ICI for the 2016-2021 period [8]. However, these studies did not compare other methods, which could lead to more accurate results. It is important to distinguish between the GARCH method and the GARCH model. The GARCH method refers to the broader statistical technique used to capture volatility that changes over time in time series data. Meanwhile, the GARCH model denotes a specific form of that technique, such as GARCH(1,1), applied in empirical analysis. In short, the method represents the overall approach, while the model reflects its specific mathematical application.

This research distinguishes itself by conducting a comparative analysis between a parametric method (ARCH/GARCH) and a nonparametric method (Fourier Series). In contrast to previous studies that typically focused on a single model, this study assesses the predictive performance of both approaches to determine which yields more accurate forecasts of ICI volatility. Additionally, it utilizes the most recent weekly data from March 10, 2024, to June 23, 2024 capturing post-pandemic market behavior and reflecting the current economic environment. The incorporation of the Fourier Series method into volatility modeling also represents a novel contribution, as this technique has been seldom applied in previous research within the context of the Indonesian stock market.

This study aims to predict the Indonesia Composite Index (ICI) more accurately using the ARCH/GARCH and Fourier Series methods. By utilizing historical weekly ICI data from March 15, 2020, to June 23, 2024, this research seeks to identify and forecast volatility patterns to support more reliable market predictions. The results of this study are expected to provide a positive outlook for investors to invest in Indonesian companies, while the government should implement policies that ensure the continued rise of stocks in Indonesia.

2. RESEARCH METHODS

2.1 Source and Data of Research Variables

This study employs a quantitative approach with a focus on time series analysis. The data utilized consists of Indonesia Composite Index (ICI) stock prices, sourced from the Indonesia Investing.com website [4]. This data set consists of weekly data covering the period from March 15, 2020, to June 23, 2024, comprising 222 data observations. The research is structured into two segments: training data and testing data. The training dataset, covering the period from March 15, 2020, to March 3, 2024, is used to develop the model with a total of 207 training data (93%). Meanwhile, the testing dataset, spanning from March 10, 2024 to June 23, 2024, was used to evaluate the accuracy of the model with a total of 15 testing data (7%).

2.2 Data Analysis Steps

The steps of data analysis used in this research are as follows:

1. Determine research variables
   The selection of variables is based on the characteristics of time series data, where past values of a variable can influence its future values. This approach is commonly used in models such as ARIMA, as it effectively captures historical patterns and market trends. The period from March 2020 to June 2024 was deliberately chosen because it encompasses the crisis period caused by the COVID-19 pandemic as well as the post-pandemic economic recovery phase, providing a relevant context for analyzing the dynamics of the Indonesian stock market. Prior to analysis, the data underwent preprocessing, which included removing duplicates, verifying consistent time intervals, and handling missing values through imputation methods such as forward filling to maintain the continuity of the time series. Additionally, data transformation, such as calculating logarithmic returns, was also considered to stabilize variance and meet the statistical assumptions of the model.
2. Checking data stationary using Augmented Dickey Fuller (ADF) and Box-Cox Transformation
   Non-stationarity in the mean indicates that the average value of the data changes over time, which is usually characterized by an upward or downward trend. To address this, a commonly used method is differencing, which calculates the difference between the current and previous values of the data, in order to remove the trend. On the other hand, non-stationarity in variance means that the degree of dispersion or volatility of the data also changes over time. This is often found in financial data such as stock indices or exchange rates, which tend to show large fluctuations in certain periods. To stabilize the variance, Box-Cox transformation can be used, which is a data transformation technique that aims to homogenize the variance by selecting the optimal lambda (λ) parameter value. After transforming and differencing, a retest of stationarity such as using the Augmented Dickey-Fuller (ADF) test is required to ensure that the data has met the stationary assumptions.
3. Selection of the best ARIMA model
   Identify the ARIMA parameter values, namely p (autoregressive), d (differencing), and q (moving average). The value of d is determined by the amount of differencing required to achieve stationarity. Meanwhile, the p and q values can be estimated from the PACF and ACF graph patterns. Once a candidate model is determined based on a combination of p and q values, it is then estimated and compared using selection criteria such as AIC (Akaike Information Criterion), where the best model is the one with the lowest AIC value. Once a model has been selected, it is important to diagnose the residuals, the remaining model error. A good residual should be white noise, which means it has no pattern, is random, and has a constant variance. In addition, the residual data should be normally distributed.
4. Detecting heteroscedasticity
   Heteroskedasticity was detected using the ARCH-LM test. A significant p-value (< 0.05) indicates ARCH effects in the residuals, justifying the use of ARCH/GARCH modeling.
5. Estimating the ARIMA- ARCH-GARCH model on Training Data
   The ARCH/GARCH method is a continuation of the ARIMA method, provided that the selected ARIMA model has heteroscedasticity assumptions. The Autoregressive Conditional Heteroskedasticity (ARCH) model is an autoregressive model that occurs when the variance is not constant [9]. Fluctuations in the data cause the variance of the residuals to be inconstant and heteroscedasticity. ARCH is used to be alternatively modeled by allowing the conditional variance

of the squared residuals to depend on the previous squared residual values. The ARCH model can be defined in the following Eq. (1) [10].

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \tag{1}$$

With:
σ_t^2 : is the conditional variance of the time series at time t
α_0 : is a constant term and must be positive
α_1 : is the coefficient of the lagged squared residual
ε_{t-1}^2 : is the squared error (shock) from the previous time period $t - 1$
In general, the ARCH (p) model can be represented by the following Eq. (2).

$$\sigma_t^2 = \sigma_0 + \sigma_1 \varepsilon_{t-1}^2 + \sigma_2 \varepsilon_{t-2}^2 + \cdots + \sigma_p \varepsilon_{t-p}^2 \tag{2}$$

With:
σ_t^2 : is the conditional variance of the time series at time t
α_0 : is a constant term and must be positive
α_1 : is the coefficient of the lagged squared residual
ε_{t-p}^2 : is the squared error (shock) from the previous time period $t - p$
Meanwhile, GARCH (Generalized Autoregressive Conditional Heteroscedasticity) is a development of the ARCH model designed to handle heteroscedasticity problems in time series data. While ARCH only uses past squared residuals to estimate the current variance, GARCH enhances it by incorporating both the squared residuals and the variance estimated in the previous period. This approach allows GARCH to capture volatility clustering patterns and long-term dependencies more effectively, thus making it more flexible and accurate, especially in analyzing financial data that often experiences volatility fluctuations in the form of clusters. Therefore, while ARCH is simpler and remains useful in certain cases, GARCH offers broader and more realistic modeling capabilities in describing complex variance behavior in time series. In general, the GARCH (p, q) model, where p indicates the ARCH element and q indicates the GARCH element, can be stated in the following Eq. (3) [10].

$$\sigma_t^2 = \omega + \sum_{j=1}^{q} a_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{3}$$

Where:
σ_t^2 : The conditional variance of the time series at time t.
ω : A constant term (intercept), must be positive: $\omega > 0$
ε_{t-j}^2 : The squared error (shock) from j time steps ago.
a_j : The ARCH coefficient for the lag j; measures impact of past shocks.
σ_{t-j}^2 : The conditional variance from j time steps ago.
β_j : The GARCH coefficient for the lag j; measures persistence of past volatility.
p : The order of the ARCH term, number of past squared errors used.
q : The order of the GARCH term, number of past variances used.

6. Forecasting ICI using ARIMA- ARCH-GARCH on Testing Data
The same modeling procedure was applied to both training and testing data. In the training phase, the ARIMA ARCH-GARCH model was fitted by estimating optimal parameters and testing for heteroskedasticity. The testing phase used the trained model to generate forecasts, which were then evaluated against actual values using metrics such as MAPE and R-squared to assess prediction accuracy and volatility capture.
7. Calculating the MAPE value of the ARCH-GARCH model
Mean Absolute Percentage Error (MAPE) represents the average of the absolute percentage errors, indicating the extent of prediction error in relation to the actual value. A lower MAPE value signifies greater accuracy in the forecasting results. Eq. (4) below is the formula used to calculate MAPE [11].

$$MAPE = \frac{1}{n} \sum_{r=1}^{n} \left| \frac{A_i - F_i}{A_i} \right| \times 100\% \tag{4}$$

With, A_i is the i-th actual data, F_i is the prediction results for the i-th actual data, and n is the number of data.
The resulting MAPE value has an interpretation in Table 1 as follows **[12]**.

Table 1. Interpretation of MAPE Value

MAPE (%)	Interpretation
<10%	High accuracy prediction
10% - 20%	Good prediction
21% - 50%	Prediction is within reason
>50%	Inaccurate prediction

8. Testing using Fourier Series estimation
The Fourier series is a trigonometric polynomial that offers significant flexibility, as it is composed of curves generated by Sine and Cosine functions. In regression analysis, the Fourier series estimator is employed to approximate a function or curve from data with an unknown pattern, particularly for data exhibiting seasonal trends [13]. Data is considered seasonal if the data curve forms certain patterns in each time period. If the observation data is known (t_r, y_r) the general regression model is as follows in Eq. (5) [14].

$$y_r = m(t_r) + \varepsilon_r, \qquad r = 1,2,3,\dots,n \tag{5}$$

The regression function $m(tr)$ is of unknown form and will be estimated with Fourier series. It is assumed that $m(tr) \in L2[a,\ b]$ is contained in Hilbert space as a linear combination of basic elements of $L2[a,\ b]$ which can be expressed in Eq. (6) as follows.

$$m(t_r) = \sum_{j=1}^{\infty} \beta_j x_j(t_r) + \varepsilon_r, \qquad r = 1,2,3,\dots,n \tag{6}$$

Hence, the model becomes.

$$y_r = \sum_{j=1}^{\infty} \beta_j x_j(t_r) + \varepsilon_r, \qquad r = 1,2,3,\dots,n \tag{7}$$

Assume that $t_1, t_2, t_3, \dots, t_n$ are equally spaced over the interval $[a, b]$.
In the estimation of unknown Fourier coefficients, it can be estimated by determining the optimal λ value which can express the number of Fourier coefficients βj that determine the smoothness of the regression curve. A higher λ leads to a smoother curve but may cause overfitting if too large. The optimal λ can be selected using criteria such as cross-validation or minimizing the Generalized Cross-Validation (GCV) score. The Fourier series estimator can be written as Eq. (8) follows [15].

$$\widehat{m}(t_r) = \beta_0 + \sum_{j=1}^{\lambda} \left[a_j \cos\left(\frac{2\pi j(r-1)}{n}\right) + b_j \sin\left(\frac{2\pi j(r-1)}{n}\right) \right] \tag{8}$$

9. Determining the optimal lambda based on the minimum GCV (Generalized Cross Validation) and R-Squared values
Cross validation (CV) and generalized cross validation methods can be used to select the optimal bandwidth value [16]. The equations for the CV and GCV methods are given in the following Eqs. (9) and (10) [16], [17].

$$CV(h) = \frac{1}{n} \sum_{i=1}^{n} [y_i - \hat{g}_{h-i}(x_i)]^2 \tag{9}$$

$$GCV(h) = \frac{\sum_{i=1}^{n} [y_i - \hat{g}_{h-i}(x_i)]^2}{\{n^{-1} tr[\boldsymbol{I} - \boldsymbol{A}(h)]\}^2} \tag{10}$$

Given y_i is the value of the response variable at the i-th observation, $\hat{g}(x_i)$ is the estimated value of the regression function at point x_i, $\hat{g}_{h-i}(x_i)$ is the estimated value of the regression function at point x_i with the i-th observation removed, $\boldsymbol{I}$ is the identity matrix of size $n \times n$, and $\boldsymbol{A}(h)$ is the matrix of size $n \times n$ for each h. The regression function at point x_i is the estimated value of the regression function.

R-Squared or commonly called the coefficient of determination explains how well the model's ability to explain variations in the dependent variable [18]. The range of the coefficient of determination is between 0 and 1. The formula for the coefficient of determination is written as Eq. (11) below [19].

$$R^2 = \frac{(\hat{\boldsymbol{y}} - \overline{\boldsymbol{y}})'(\hat{\boldsymbol{y}} - \overline{\boldsymbol{y}})}{(\boldsymbol{y} - \overline{\boldsymbol{y}})'(\boldsymbol{y} - \overline{\boldsymbol{y}})} \tag{11}$$

Where $\overline{\boldsymbol{y}}$ is a vector containing the mean of the response data. A good model is measured by the largest R^2 value.

10. Determine Fourier Series modeling
11. Calculating the MAPE value of Fourier series modeling
12. Forecasting ICI using the Fourier series model
The Fourier series model forecasts ICI by breaking down historical data into sine and cosine components to capture underlying cyclical patterns. These patterns are then projected forward to estimate future index values based on the repetition of past trends.

3. RESULTS AND DISCUSSION

3.1 Overview of the Democracy Index in Indonesia

In this study, presents descriptive statistics, including the mean, minimum, and maximum values, along with a line plot of the ICI (Indonesia Composite Index) over time. Descriptive statistics are used to summarize the data, while the line plot helps identify the data pattern for prediction. The description of each research variable follows.

Table 2. Descriptive Statistics

Variable	N	Mean	St. Dev	Min	Date (Min)	Max	Date (Max)
ICI (Indonesia Composite Index	223	6441.7	768,89	4194.94	15/03/2020	7228.91	3/3/2024

Based on Table 2, the Indonesia Composite Index (ICI) had an average value of 6441.7 from 223 observations, with a standard deviation of 768.89, indicating significant fluctuations. The lowest value was 4194.94 on March 15, 2020, and the highest was 7228.91 on March 3, 2024, suggesting a period of market recovery or growth. Overall, the index exhibits high volatility.

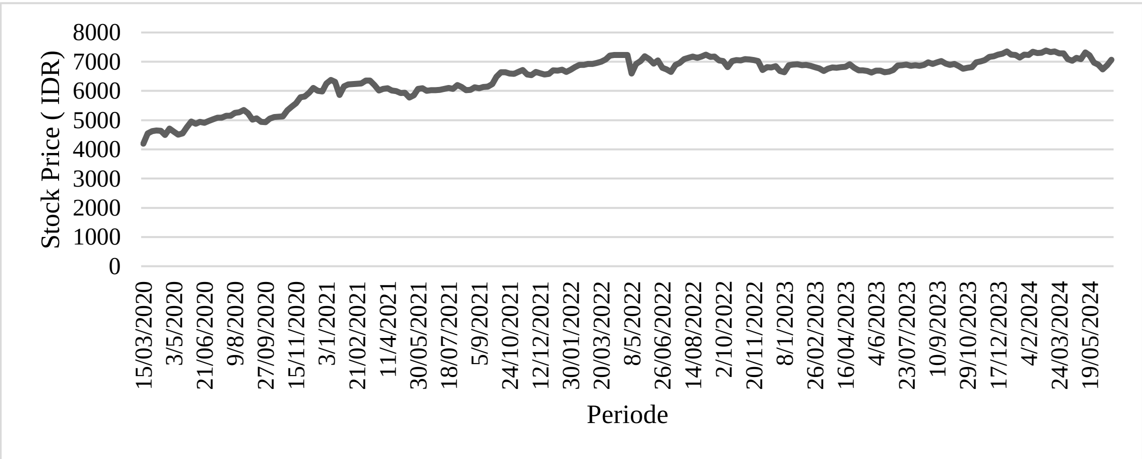


Figure 1. Time Series Plot

Based on Fig. 1, The data exhibits recurring fluctuations over a certain period and shows an upward trend, making the Fourier method suitable for predicting stock prices. Additionally, stock prices tend to demonstrate volatility that changes over time, which makes the ARCH-GARCH method also appropriate for predicting stock prices.

3.2 ARCH- GARCH Forecasting

3.2.1 Data Stationarity

Data stationarity is a mandatory condition in classical time series modeling, such as ARIMA. Data is considered stationary if it remains stable over time with constant mean and covariance. The statistical tests used to test the stationarity of time series data are the Augmented Dickey-Fuller (ADF) test and the Box-Cox transformation. Stationarity occurs if the ADF test shows a p-value less than the determined significance level ($\alpha = 0.05$). The Box-Cox transformation is one of the methods used to address non-stationarity in variance when the transformation parameter is not equal to 1 ($\lambda \neq 1$).

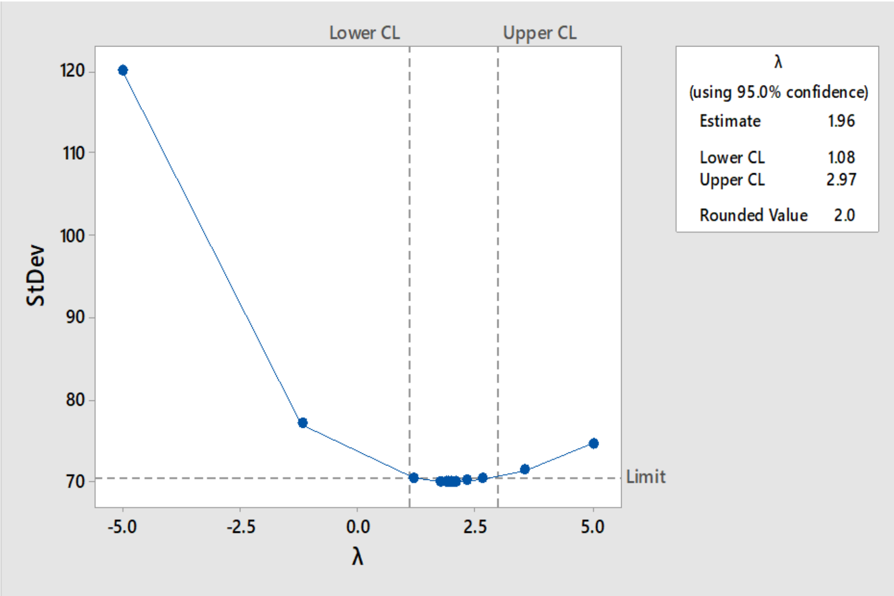


Figure 2. Box-Cox Plot for Original Data

Based on Fig.e 2, the Box-Cox transformation yields an estimated lambda (λ) of 1.74, which is within the 95% confidence interval (1.08 to 2.97) and rounded value or lambda (λ) is 2.0 so that the data still needs to be transformed into a box-cox in the form of Zt^2. After the transformation, the results of Box-Cox have a rounded value or lambda (λ) of 1 so that further testing can be carried out.

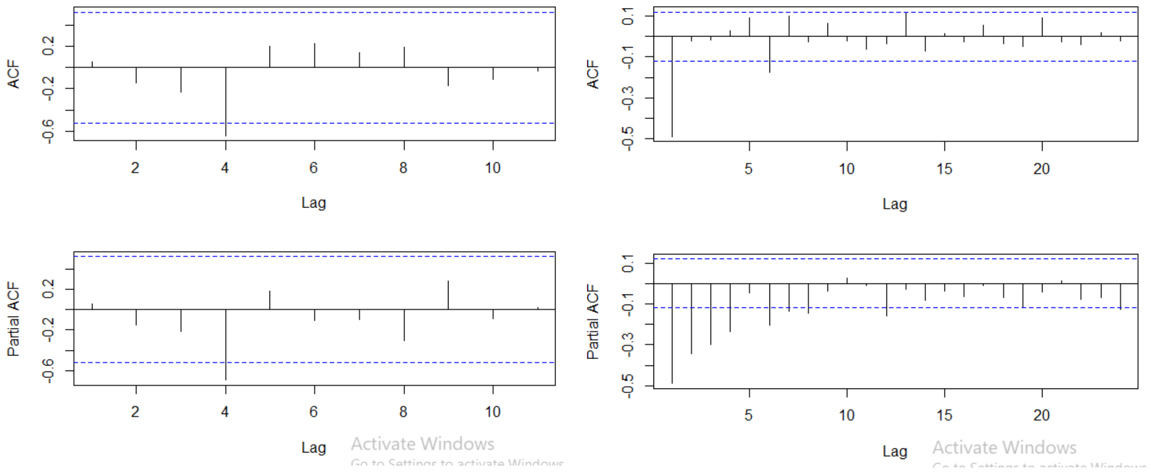
Table 3. ADF Test Results

Variable	*P-value*	Results
ICI	0.4964	Non-stationary data
d(ICI)	0.0100	Stationary data
d(d(ICI))	0.0100	Stationary data

Based on Table 3, the results of the Augmented Dickey-Fuller (ADF) test show that the original ICI data is non-stationary, as indicated by a p-value of 0.4964 (> 0.05). After the first differencing, the p-value becomes 0.0100 (< 0.05), suggesting that the data is already stationary.

However, further examination of the ACF and PACF plots after first differencing Fig. 3 (a) did not show a clear pattern (lags were not significantly distinct). Therefore, a second differencing was performed to better capture the structure of the time series and improve model identification. After second differencing, the ADF test still confirms stationarity ($p - value = 0.0100$), and the ACF and PACF plots now display significant lags Fig. 3 (b), which can be used to select appropriate ARIMA model parameters.

3.2.2 ACF and PACF plots



(a) (b)

Figure 3. ACF and PACF Plots (a) Differencing-1 , (b) Differencing-2

Based on Fig. 3 **(**a) (first differencing), both ACF and PACF show very few significant spikes and no clear cutoff or tailing pattern. In Fig. 3 (b) (second differencing), the ACF plot shows a significant spike at lag 1 and tails off afterward, while the PACF shows multiple significant spikes before decaying. These patterns suggest the potential for ARIMA$(p, 2, q)$ modeling, with further candidate models explored in the next section.

3.2.3 Selection of the Best ARIMA Models

In evaluating ARIMA model candidates, several diagnostic criteria are considered to ensure model validity and accuracy. First, the statistical significance of the parameters (AR and MA terms) is assessed using p-values, where values below 0.05 indicate that the parameters significantly contribute to explaining the variance in the data. Second, the Akaike Information Criterion (AIC) is used to compare model performance. A lower AIC value indicates a better model in terms of balancing fit and complexity, making it a key metric for model selection. Additionally, white noise testing is performed on the residuals to ensure that the model has adequately captured the structure of the data. Residuals that behave like white noise (i.e., have no autocorrelation) indicate a well-fitted model. Finally, residual normality is examined to check whether the model errors are normally distributed, which is a desirable property for statistical inference.

Table 4. Model Diagnostic Test

Model	Significance	*White Noise*	Residual Normality	AIC Values
ARIMA (1, 2, 0)	Yes	No	Yes	3396.95
ARIMA (2, 2, 0)	Yes	No	No	3351.12
ARIMA (3, 2, 0)	Yes	No	Yes	3335.76
ARIMA (4, 2, 0)	Yes	No	No	3316.35
ARIMA (5, 2, 0)	Yes	Yes	No	3311.92
ARIMA (1, 2, 1)	No	Yes	No	3281.05
ARIMA (2, 2, 1)	No	Yes	No	3278.53
ARIMA (3, 2, 1)	No	Yes	No	3280.45
ARIMA (4, 2, 1)	No	Yes	No	3282.28
ARIMA (5, 2, 1)	No	Yes	No	3282.56
ARIMA (0, 2, 1)	Yes	Yes	No	3281.09

Based on Table 4 shows the diagnostic test results for several ARIMA models with second-order differencing. Although none of the models passed all three diagnostic tests, ARIMA(5,2,0) and ARIMA(0,2,1) passed two out of three. A closer look at the ACF and PACF plots supports the selection of ARIMA(0,2,1)**,** as the ACF cuts off sharply after lag 1, and the PACF tails off, which is consistent with the characteristics of an MA(1) model. Therefore, ARIMA(0,2,1) is recommended as the final model due to both statistical support and alignment with theoretical identification patterns.

These findings demonstrate how stationarity tests and the behavior of ACF and PACF plots directly influence the model selection process. The need for second differencing, despite statistical stationarity after first differencing, is justified by the improved clarity in lag structure, which is essential for selecting appropriate AR and MA terms. Further examination of residual normality for this model will be done using its residual histogram.

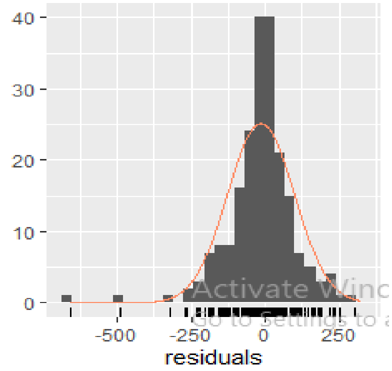


Figure 4. Histogram ARIMA (0,2,1)

As seen in Fig. 4, the residual histogram for the ARIMA(0,2,1) model indicates non-normality, with values densely clustered near zero due to high prediction accuracy. Nonetheless, the model is still interpretable, and residual variance heteroscedasticity testing can proceed.

3.2.4 ARIMA Model Estimation

Table 5 below presents the results of the estimation of the parameters of the ARIMA model (0,2,1), including the coefficient, error standards, p-values, as well as evaluation metrics such as Akaike Information Criterion (AIC) and Mean Squared Error (MSE).

Table 5. ARIMA Model Estimation (0, 2, 1)

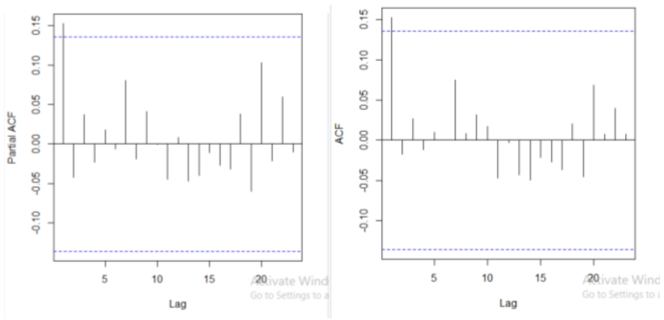
Parameter	Coefficient	Error Standards	P-Value	AIC	MSE
MA (1)	-0.99087	-0.01924	0.000	3281.09	254.3628

Based on Table 5, since the p-value is less than 0.05, the ARIMA model estimate (0, 2, 1) will be used for the ARIMA GARCH estimation. The mathematical equation of the estimated ARIMA model (0, 2, 1) t model estimation in Eq. (12) is as follows.

$$(1 - B)^d Zt^* = \theta q\,(B)\varepsilon_t$$
$$(1 - B)^d Zt^* = (1 - \theta_1 B)\varepsilon_t$$
$$(1 - B)^2 Zt^* = (1 - \theta_1 B)\varepsilon_t$$
$$(1 - 2B + B^2)Zt^* = (1 - \theta_1 B)\varepsilon_t$$
$$Zt^* - 2Z_{t-1}^* + Z_{t-2}^* = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$
$$Zt^* = 2Z_{t-1}^* - Z_{t-2}^* + \varepsilon_t - \theta_1 \varepsilon_{t-1}$$
$$Zt^* = 2Z_{t-1}^* - Z_{t-2}^* + \varepsilon_t - (-0.99087)\varepsilon_{t-1}$$
$$Z_t^* = 2Z_{t-1}^* - Z_{t-2}^* + \varepsilon_t + 0.99087\varepsilon_{t-1} \quad (12)$$

3.2.5 Detection of Heteroscedasticity of Residual Variance

In the ARIMA model, it is assumed that the residuals have a normal distribution with a mean of $\mu = 0$ and homogeneous variance σ^2. However, in economic data such as exchange rates, inflation, and stock prices, high volatility is often observed, which violates the assumption of homogeneous residual variance. To detect this, a heteroscedasticity test is performed on the squared residuals generated by the ARIMA model. In this case, the squared error ϵ^2 from the ARIMA (0,2,1) model is used as an estimator of the residual variance σ^2. Below are the ACF and PACF plots for the squared residuals.



(a) (b)

Figure 5. Residual Quadratic Model ARIMA (0,2,1) (a) Plot ACF, (b) Plot PACF

Based on Fig. 5, lags in both the ACF and PACF plots, indicating autocorrelation and suggesting heteroscedasticity in the residual variance. In addition to the ACF and PACF plots, heteroscedasticity can be tested using the ARCH-LM (Lagrange Multiplier) test, with the null hypothesis of no conditional heteroskedasticity in the ARIMA model residuals.

Table 6. ARCH-LM Test Results

Order	Significsis LM-Test
1	0.01205348
2	0.03445677
3	0.06469904
4	0.1147188
5	0.1885327

The test results in Table 6 show that for five lags, the initial two lags of significance value (p-value) are less than (5%) so the decision is to fail to reject the null hypothesis. Thus, it can be concluded that there is an autocorrelation between the residual squares of the ARIMA model (0,2,1). To overcome this, advanced modeling is needed to deal with heteroscedasticity that occurs using ARCH/GARCH.

3.2.6 Estimation of the ARCH/GARCH Model

In the previous point, it was known that the estimation results of the ARIMA model (0,2,1) experienced heteroscedasticity symptoms in its residuals, so it was necessary to model variance with ARCH. The ARCH/GARCH modeling was based on the square residual ACF and PACF plots with the provision that the ARCH(p) model was determined by the p-second lag that came out in the ACF plot, the variances were ARCH (1) or GARCH (1.0) and ARCH (2) or GARCH (2.0).

Table 7. GARCH Model Estimation

Model	Parameter	Coefficient Estimation	Significance	AIC
GARCH (1,0)	ω	11463.98576	0	12.30645
	α	0.23613	0.026004	
GARCH (2,0)	ω	1.1279E+04	0	12.31652
	α_1	2.0591E-01	0.044909	
	α_1	2.8284E-02	0.049992	

Based on Table 7, The selection of the GARCH(1,0) model was based on two primary considerations, the Akaike Information Criterion (AIC) and the statistical significance of the estimated parameters. As shown in the results, the GARCH(1,0) model yields an AIC value of 12.30645, which is slightly lower than that of the GARCH(2,0) model (12.31652). A lower AIC indicates a better balance between model fit and parsimony, favoring simpler models with strong explanatory power. Furthermore, both parameters in the GARCH(1,0) model are statistically significant at the 5% level. The constant term ω has a p-value of 0, and the lagged error term coefficient α has a p-value of 0.026004, confirming the relevance of past shocks in explaining current volatility. Although the GARCH(2,0) model also produces significant coefficients, the additional parameter α_1 only marginally contributes to the model while increasing complexity. Given this, the GARCH(1,0) model is selected as the final volatility model due to its lower AIC and more efficient structure.Thus, the variance equation of the ARIMA residuals is as follows in Eqs. (13) and (14).

$$\sigma_t^2 = 11463.98576 + 0.23613\varepsilon_{t-1}^2 \quad (13)$$

$$\varepsilon_t = y_t\sigma_t \quad (14)$$

3.2.7 ARIMA-GARCH Model Estimation

The combination of Eqs. (13) and (14) represents the merged equation of the ARIMA mean model and the ARCH variance model, forming ARIMA(0,2,1)-ARCH(1), which can be mathematically expressed as follows in Eq. (17).

$$Z_t^* = 2Z_{t-1}^* - Z_{t-2}^* + y_t\sqrt{11463.98576+0.23613\varepsilon_{t-1}^2} + 0{,}99087y_t\sqrt{11463.98576+0.23613\varepsilon_{t-2}^2} \quad (15)$$

$$Z_t^{*2} = \left(2Z_{t-1}^* - Z_{t-2}^* + y_t\sqrt{11463.98576+0.23613\varepsilon_{t-1}^2} + 0{,}99087y_t\sqrt{11463.98576+0.23613\varepsilon_{t-2}^2}\right)^2 \quad (16)$$

$$\sigma_t^2 = \left(2Z_{t-1}^* - Z_{t-2}^* + y_t\sqrt{11463.98576+0.23613\varepsilon_{t-1}^2} + 0{,}99087y_t\sqrt{11463.98576+0.23613\varepsilon_{t-2}^2}\right)^2 \quad (17)$$

3.2.8 Evaluation of ARIMA (0,2,1) - ARCH(1) Model Performance

The accuracy of the ARIMA (0,2,1) - ARCH(1) model can be assessed by comparing the estimated values with the actual data. Fig. 6 illustrates the comparison between the actual values and the predicted results, showing how well the model captures the observed trend and volatility.

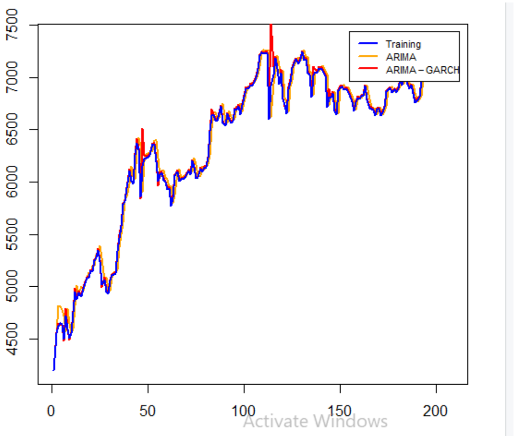


Figure 6. Plot Comparison of Actual Value and Estimated Results

As in Fig. 6, the results of ICI modeling provide a pattern that is quite volatile and has an upward trend according to actual data, The results of the calculation of the goodness of the ARIMA model (0,2,1) for in sample data are given in Table 8 as follows.

Table 8. Results of the Goodness of the ARIMA Model (0, 2, 1) - ARCH(1) on Data In-Sample and Out-Sample

	In-Sample	**Out-Sample**
MAPE	0.3%	5%
R-Square	97%	85.6%

3.2.9 ICI Forecasting with ARIMA-GARCH

Forecasting aims to predict an event for several future periods. Table 9 presents the forecasting results for the ICI over the next 15 periods, from March 10, 2020, to June 23, 2024, as an out-sample prediction using the ARIMA(0,2,1)-ARCH(1) model.

Table 9. Summary of ICI Predictions

Date	Current	Predictions	Upper Limit of Prediction	Lower Limit of Prediction
10/3/2024	7328.05	7245.488	8575.215	6503.762
17/03/2024	7350.15	7200.460	8491.247	6323.672
24/03/2024	7288.81	7378.234	8541.428	6299.041
31/03/2024	7286.88	7345.009	8586.180	6279.838
14/04/2024	7087.32	7185.784	8627.517	6264.051
21/04/2024	7036.08	7145.559	8666.466	6250.651
28/04/2024	7134.72	7067.333	8703.635	6239.032
5/5/2024	7088.79	7189.108	8739.418	6228.798
12/5/2024	7317.24	7245.883	8774.087	6219.679
19/05/2024	7221.04	7324.658	8807.839	6211.477
26/05/2024	6970.74	6945.433	8840.822	6204.043
2/6/2024	6897.95	6876.207	8873.151	6197.264
9/6/2024	6734.83	6894.982	8904.916	6191.048
16/06/2024	6879.98	6856.757	8936.191	6185.323
23/06/2024	7063.58	7145.532	8967.036	6180.028

Based on Table 9, predicted values closely follow actual ICI values during the testing period, with most actual values remaining within the 95% confidence interval. This indicates the model performs well in capturing market trends.

3.3 Fourier Forecasting

3.3.1 Lambda Determination

Fourier series data estimation requires Fourier coefficients that represent the underlying cyclical patterns in the time series. The number of Fourier terms used in the estimation is controlled by a parameter called lambda (λ), also known as the smoothing parameter. A higher value of λ allows the model to capture more variation and detail, but it can also lead to overfitting if too many coefficients are used. Conversely, a λ that is too low may oversimplify the model. Therefore, determining the optimal value of λ is essential, and this is done using the Generalized Cross Validation (GCV) method.

Table 10. GCV Value for Each Lambda

Lambda	GCV
10	59956.11
19	46225.93
25	44548.98
28	40801,95

Based on Table 10, the GCV values are evaluated across various λ values. The GCV score indicates the model's prediction error, and a lower value suggests a better balance between smoothness and accuracy. In this case, the lowest GCV value (39010.12) is obtained when λ = 34, meaning that using 34 sine and cosine terms yields the best-fitting model for the ICI data. Therefore, the final model will use λ = 34 Fourier coefficients in its estimation.

To better visualize the behavior of GCV values across different λ values, a plot of GCV versus λ is presented in Fig. 7. This helps illustrate how the smoothing parameter affects model performance.

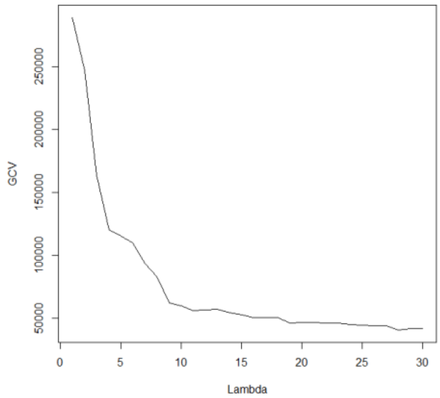


Figure 7. GCV and Lambda Plot

The graph in Fig. 7 shows that the GCV values generally decrease as λ increases, reaching a minimum at $\lambda = 34$. After that point, GCV begins to increase again or flatten, indicating that larger λ values no longer improve model accuracy and may cause overfitting. This confirms that $\lambda = 34$ is the optimal choice for balancing model complexity and predictive accuracy.

3.3.2 Determining Fourier Modeling

Next, determine the value of a_i as the cosine function component and $a_i . b_j$ as the sine Fourier function component presented in the following Table 11 and obtain the estimation model in Eq. (18).

Table 11. Fourier Parameter Value

Lambda	Value a_i	Value b_j	Lambda	Value a_i	Value b_j
1	-463.524	-641.7841	18	22.4228	-32.602
2	-10.89533	-302.4193	19	-16.92024	-84.93953
3	-81.29823	-396.3109	20	-15.5431	-24.95324
4	51.20761	-287.7816	21	-15.3916	-33.02719
5	25.27814	-115.0535	22	-36.55945	-31.69785
6	-45.00541	-105.1539	23	-15.1187	-36.95112
7	-88.00813	-157.9726	24	-39.77752	-32.56711
8	-1.518772	-144.5985	25	1.063712	-38.71983
9	-19.99195	-192.3657	26	-32.91205	-31.18255
10	5.581796	-72.86746	27	-13.13657	-34.16038
11	-25.13977	-86.53688	28	-33.26339	-57.04202
12	0.4737994	-36.63409	29	-5.003212	-7.928584
13	-19.45094	-22.95041	30	-19.41951	-31.87481
14	-45.28438	-62.66026	31	-37.58458	-16.36071
15	11.39632	-64.61948	32	-19.87968	-46.58064
16	-10.10649	-67.25473	33	-16.32611	-31.15632
17	-6.821116	-38.72626	34	-23.85774	-25.06652

$$\begin{aligned}\hat{m}(tr) = 6393.361 &- 641.7841\, cos(2\pi t_r) - 463.524\, sin(2\pi t_r) - 302.4193\, cos(4\pi t_r) \\ &- 10.89533\, sin(4\pi t_r) - \cdots - 23.85774\, cos(68\pi t_r) - 25.06652\, sin(68\pi t_r)\end{aligned} \quad (18)$$

To evaluate the performance of the Fourier model in predicting ICI share prices, a comparison between the estimated and actual values is presented. Fig. 8 illustrates the relationship between the predicted results and the actual data, showing the model's accuracy in capturing market trends.

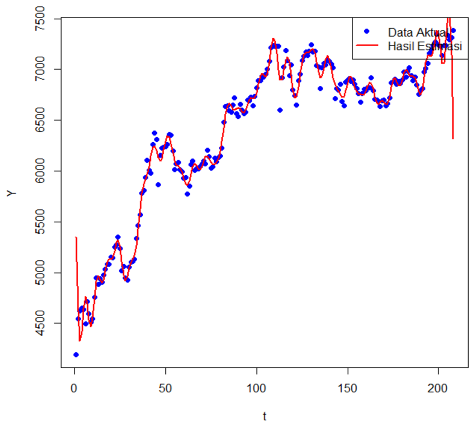


Figure 8. Plot of Estimated and Actual Results

Based on Fig. 8, the estimation results are quite close to the actual value of the ICI share price with the calculation of the goodness test of the out sample and in sample data models as follows.

Table 12. Results of the Virtue of the Fourier Model on In-Sample and Out-Sample Data

	In-Sample	**Out-Sample**
MAPE	1.24%	8.57%
R-Square	97%	85%

Based on Table 12, in the in-sample data, the forecasting results obtained a MAPE value of 1.24% with an r square value of 97%. Meanwhile, the outsample data obtained a MAPE value of 8.75% with an R-square value of 85%. From the MAPE value obtained, it can be said that ICI modeling with the Fourier series has high accurate prediction.

3.3.3 ICI Forecasting with Fourier Series

Forecasting aims to find out the prediction of an event over the next several periods. In Table 13, the data from the ICI forecast for the next 15 periods is presented, namely March 10, 2020 to June 23, 2024 or predictions in an out sample with the Fourier model.

Table 13. Summary of ICI Stock Price Prediction

Date	**Current**	**Predictions**	**Upper Limit of Prediction**	**Lower Limit of Prediction**
10/3/2024	7328.05	7076.34	7500.82	6340.89
17/03/2024	7350.15	6869.02	7450.15	6133.57
24/03/2024	7288.81	6792.93	7289.81	6057.49
31/03/2024	7286.88	6812.84	7200.88	6077.39
14/04/2024	7087.32	6856.05	7200.32	6120.60
21/04/2024	7036.08	6870.29	7267.08	6134.84
28/04/2024	7134.72	6850.78	7634.72	6115.33
5/5/2024	7088.79	6823.97	7456.79	6088.53
12/5/2024	7317.24	6813.07	7000.24	6077.62
19/05/2024	7221.04	6819.44	7003.04	6084.00
26/05/2024	6970.74	6709.45	6831.18	6095.74
2/6/2024	6897.95	67390.0	6840.96	6105.52
9/6/2024	6734.83	6756.96	6849.96	6114.51
16/06/2024	6879.98	6769.48	6854.48	6119.03
23/06/2024	7063.58	6997.87	7005.53	6091.45

The results of the ICI forecast provide a pattern that is quite volatile. as well as actual data. Although there are prediction results that exceed the stock price in the original data. it is still considered reasonable because it is still between the upper and lower limits of the prediction.

The results of this study are in line with previous research according to [8], the study shows that the ARCH-GARCH model is able to predict ICI volatility well with a MAPE of 17.26%. However, unlike previous studies that only used ARCH-GARCH, this study combined ARCH-GARCH with Fourier to capture cycle patterns that may not have been detected by the ARCH-GARCH model alone. The results obtained show that this approach can improve the accuracy of predictions, which can be seen from the lower

MAPE values compared to previous studies. Thus, this study contributes to developing a more comprehensive ICI volatility prediction method.

4. CONCLUSION

This study found ARIMA-GARCH closely matches ICI data, with ARIMA(0,2,1) as the best model. Despite heteroskedasticity, further testing with ARCH achieved excellent accuracy (MAPE 0.3%, R-squared 97% and 85.6%). Fourier series estimation also performed well (MAPE 1.24% in-sample, 8.57% out-sample), but ARIMA-GARCH proved superior due to its lower MAPE. The modeling from this study can be used to predict the next period without having to include the actual data as the modeling is pre-trained.

This study is limited to the use of weekly ICI data up to June 2024, and future research could use actual data and add macroeconomic variables. Future research could explore the integration of macroeconomic indicators, test longer forecasting horizons, and compare with other relevant models to improve the goodness of model.

Based on these findings, investors are encouraged to take advantage of the upward trend in the ICI by increasing exposure to Indonesian equities, particularly in sectors showing strong, consistent growth and low volatility. Investors should also consider diversifying their portfolios to mitigate risk and optimize returns. Policymakers are recommended to maintain economic stability, strengthen investment-friendly regulations, and invest in key infrastructure to support sustained market growth and attract both domestic and international capital.

Author Contributions

M. Fariz Fadillah Mardianto: Conceptualization, Funding Acquisition, Supervision, Validation, Methodology. Hanny Valida: Methodology, Software, Formal Analysis, Investigation, Writing - Original Draft. Farah Fauziah Putri: Formal Analysis, Data Curation, Software, Writing - Original Draft, Writing - Review and Editing. Doni Muhammad Fauzi: Formal Analysis, Data Curation, Writing - Original Draft, Writing - Review and Editing. Elly Pusporani: Methodology, Software. All authors discussed the results and contributed to the final manuscript. All authors discussed the results and contributed to the final manuscript

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declarations

The authors declare that there is no conflict of interest regarding the publication of this paper. All authors confirm that they have no competing interests to disclose.

REFERENCES

[1] I. Widiarma, “KEBIJAKAN PENUTUPAN KODE BROKER DAN TIPE INVESTOR PADA LAYAR BURSA DAPAT MENYEHATKAN PASAR MODAL,” *Jurnal Ilmiah Manajemen, Ekonomi, & Akuntansi (MEA)*, vol. 7, no. 2, pp. 1025–1051, Jul. 2023, doi: https://doi.org/10.31955/mea.v7i2.3109.

[2] H. F. Kusnandar, S. Mulyati, and A. S. Rohana, “ANALISIS PERGERAKAN INDEKS HARGA SAHAM GABUNGAN (IHSG) SEBELUM DAN SESUDAH PANDEMI COVID,” *EKOMBIS REVIEW: Jurnal Ilmiah Ekonomi dan Bisnis*, vol. 11, no. 2, pp. 1113–1120, Jul. 2023, doi: https://doi.org/10.37676/ekombis.v11i2.3834.

[3] R. Andrianto, “IHSG 2023 TAK TERBENDUNG, CATAT REKOR BARU HINGGA SENTUH 7300,” CNBC Indonesia. Accessed: Nov. 28, 2024. [Online]. Available: https://www.cnbcindonesia.com/research/20231230023432-128-501422/ihsg-2023-tak-terbendung-catat-rekor-baru-hingga-sentuh-7300

[4] Investing.id, “IDX COMPOSITE HISTORICAL DATA.,” Investing.id. [Online]. Available: https://id.investing.com/indices/idx composite-historical-data.

[5] B. Jange, “PREDIKSI VOLATILITAS INDEKS HARGA SAHAM GABUNGAN MENGGUNAKAN GARCH,” *ARBITRASE: Journal of Economics and Accounting*, vol. 4, no. 1, pp. 1–6, Jul. 2023, doi: https://doi.org/10.47065/arbitrase.v4i1.1122.

[6] R. S. Nurfajriyah, D. Harjadi, and F. Adzimatinur, "ANALISIS PERBANDINGAN METODE ARCH DAN GARCH DALAM PERAMALAN INDEKS HARGA SAHAM," *Jurnal Ilmiah Manajemen, Ekonomi Dan Bisnis*, vol. 1, no. 2, pp. 148–159, 2024, [Online]. Available: https://journal.feb.uniku.ac.id/jimeb/article/view/70

[7] E. Era, S. W. Rizki, and Y. Yundari, "PEMODELAN INDEKS HARGA SAHAM GABUNGAN DENGAN MODEL REGRESI SEMIPARAMETRIK KERNEL," *Equator: Journal of Mathematical and Statistical Sciences*, vol. 1, no. 1, pp. 8–17, Dec. 2022, doi: https://doi.org/10.26418/ejmss.v1i1.57431.

[8] R. Adisetiawan, N. Nuraini, H. T. Putri, and A. Ahmadi, "PERMODELAN GARCH PADA IHSG DAN INDEKS LQ45," *J-MAS (Jurnal Manajemen dan Sains)*, vol. 6, no. 2, pp. 448–453, Oct. 2021, doi: https://doi.org/10.33087/jmas.v6i2.307.

[9] A. P. Desvina and K. Khairunnisa, "PENERAPAN METODE ARCH / GARCH DALAM MERAMALKAN TRANSAKSI NILAI TUKAR (KURS) JUAL MATA UANG INDONESIA (IDR) TERHADAP MATA UANG EROPA (GBP)," *Jurnal Sains Matematika dan Statistika*, vol. 4, no. 2, pp. 114–123, 2018.

[10] D. M. AL-Najjar, "MODELLING AND ESTIMATION OF VOLATILITY USING ARCH/GARCH MODELS IN JORDAN'S STOCK MARKET," *Asian Journal of Finance & Accounting*, vol. 8, no. 1, pp. 152–167, May 2016, doi: https://doi.org/10.5296/ajfa.v8i1.9129.

[11] N. H. Pajriati, "PENERAPAN METODE AVERAGE BASED FUZZY TIME SERIES LEE UNTUK PERAMALAN HARGA EMAS DI PT. X," *Jurnal Riset Matematika*, vol. 1, no. 1, pp. 73–81, Oct. 2021, doi: https://doi.org/10.29313/jrm.v1i1.221.

[12] I. D. N. A. Manuaba, I. B. G. Manuaba, and M. Sudarma, "KOMPARASI METODE PERAMALAN GREY DAN GREY-MARKOV UNTUK MENGETAHUI PERAMALAN PNBP DI UNIVERSITAS UDAYANA," *Majalah Ilmiah Teknologi Elektro*, vol. 21, no. 1, pp. 83–88, Jul. 2022, doi: https://doi.org/10.24843/MITE.2022.v21i01.P12.

[13] A. T. R. Dani and N. Y. Adrianingsih, "PEMODELAN REGRESI NONPARAMETRIK DENGAN ESTIMATOR SPLINE TRUNCATED VS DERET FOURIER," *Jambura Journal of Mathematics*, vol. 3, no. 1, pp. 26–36, Jan. 2021, doi: https://doi.org/10.34312/jjom.v3i1.7713.

[14] A. T. R. Dani, A. F. Dewi, and L. Ni'matuzzahroh, "STUDI SIMULASI DAN APLIKASI: ESTIMATOR DERET FOURIER PADA PEMODELAN REGRESI NONPARAMETRIK," in *Prosiding Seminar Nasional Matematika, Statistika, dan Aplikasinya*, Samarinda, May 2022, pp. 279–288.

[15] H. Nurcahayani, "ESTIMATOR CAMPURAN SPLINE TRUNCATED DAN DERET FOURIER PADA REGRESI NONPARAMETRIK MULTIRESPON," Doctoral Thesis, Institut Teknologi Sepuluh Nopember, Surabaya, 2022.

[16] N. Chamidah and B. Lestari, *ANALISIS REGRESI NONPARAMETRIK DENGAN PERANGKAT LUNAK R*. Airlangga University Press, 2023.

[17] Y. Wang, *SMOOTHING SPLINES*, 1st ed. New York: Chapman and Hall/CRC, 2011. doi: https://doi.org/10.1201/b10954.

[18] F. Fatmawati and A. S. Lubis, "PENGARUH PERILAKU KEWIRAUSAHAAN TERHADAP KEMAMPUAN MANAJERIAL PADA PEDAGANG PAKAIAN PUSAT PASAR KOTA MEDAN," *Jurnal Muhammadiyah Manajemen Bisnis*, vol. 1, no. 1, pp. 1–9, Apr. 2020, doi: https://doi.org/10.24853/jmmb.1.1.1-10.

[19] M. Mardianto, "KUMPULAN KARYA ILMIYAH DALAM BENTUK ARTIKEL SEBAGAI BAHAN PENILAIAN MATA KULIAH REGRESI NONPARAMETRIK DAN LITERATURE REVIEW," 2017.