# Hybrid ARIMA Modeling with Stochastic Volatility for Forecasting the Value of Non-Oil and Gas Exports in Indonesia

Evatia Suryatin, Mustika Hadijati, Zulhan Widya Baskara

***Abstract*—Export activities consist of oil and gas exports and non-oil and gas exports. Non-oil and gas exports are one of the sectors that provide the largest foreign exchange contribution to Indonesia, and the movement of non-oil and gas export values has an impact on economic growth. Therefore, the purpose of this research is to create a model used to predict future non-oil and gas export values. One mathematical model that can be to predict Indonesia's non-oil and gas export values is the combination of the ARIMA model and the stochastic volatility model, also known as Hybrid ARIMA with stochastic volatility. The Hybrid ARIMA with stochastic volatility modeling has advantages in creating models for data with high volatility and is capable of combining linear patterned data and nonlinear patterned data. In this study, the best ARIMA (1,1,1) model was obtained with a MAPE value of 13.2082%. From the residuals of the ARIMA (1,1,1) model, there were signs of heteroscedasticity, so the GARCH model with the best GARCH (0,1) model was used. In the GARCH (0,1) model, it was found that there was an asymmetric influence, so the EGARCH and GJR-GARCH models were used. The comparison of EGARCH and GJR-GARCH models was carried out to address the asymmetric residual data pattern. Based on the research results, the best model used for prediction is the hybrid ARIMA (1,1,1) with EGARCH (1,1) model, with a MAPE value of 9.35158%.**

***Index Terms*—ARIMA, Hybrid ARIMA, MAPE, Non-oil gas exports,Stochastic Volatility**

## I. Introduction

Economic growth is defined as a long-term increase in a country's capacity to provide various economic goods to its population [1] . It serves as evidence of a nation's level of success. According to Bank Indonesia, 2018 [2], one of the activities that contribute to economic growth is exportation.

The expenditure of goods from a country's domestic circulation, which are sent abroad under government regulations in exchange for foreign currency, is referred to as exports [3]. Indonesia is among the countries impacted by the performance of export activities across various sectors, including non-oil and gas exports. Non-oil and gas exports contribute significantly to Indonesia's foreign exchange earnings. The cumulative value of Indonesia's exports decreased by 6.94%, and non-oil and gas exports experienced a decline of 4.82% [4]. This decline in non-oil and gas exports has consequences for economic performance and growth. As a result, non-oil and gas export values are a crucial factor in Indonesia's economic growth, necessitating the modeling of these values.

E. Suryatin, M. Hadijati, Z. W. Baskara are with the Universitas Mataram, Mataram, 83125 Indonesia e-mail: zulhan.wb@unram.ac.id

Manuscript received October 30, 2023; accepted January 30, 2024.

Modeling Indonesia's non-oil and gas export values can be employed to make forecasts that will inform economic policy decisions in the country. Forecasting is the process of estimating the magnitude or quantity of something in the future, based on past data, and is analyzed scientifically using statistical methods [5]. Time series analysis is a statistical method that can be used for forecasting, making it applicable to the modeling of Indonesia's non-oil and gas export values.

A time series is a collection of data observations from a fixed source that occur sequentially based on a time index t, with precise time intervals [6]. Time series analysis uniquely records economic behavior over time. According to Wei [7], there are two classifications of time series based on their form or function: linear models and nonlinear models. Autoregressive Integrated Moving Average (ARIMA) is used for linear data modeling.

Autoregressive Integrated Moving Average (ARIMA) is a suitable forecasting method, as it requires only the variable to be forecasted, and provides simple, accurate, and rapid results [8]. The ARIMA method combines Autoregressive (AR) and Moving Average (MA) models. However, financial sector data, such as Indonesia's non-oil and gas export values, which exhibit random behavior, high volatility, and non-constant variance, indicate data variation (conditional variance) resulting in heteroscedasticity [9]. As a result, the ARIMA model is insufficient for data exhibiting heteroscedasticity, necessitating the use of a hybrid model.

A hybrid model is an artificial intelligence-based method designed to address difficulties with data containing both linear and nonlinear models [10]. Several models can overcome the limitations of the ARIMA model and are capable of modeling the volatility in Indonesia's non-oil and gas export values, including Generalized Autoregressive Conditional Heteroscedasticity (GARCH), Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH), and the Glosten, Jaganathan, and Runkle Generalized Autoregressive Conditional Heteroscedasticity (GJR-GARCH) model. The consideration of these three models for integration with the ARIMA model is based on the patterns generated from the ARIMA residuals when there is an asymmetric effect on the squared ARIMA residuals with GARCH residuals. Therefore, this study will undertake Hybrid ARIMA Modeling with Stochastic Volatility for Forecasting Non-Oil and Gas Export Values.

## II. METHODOLOGY

The data used in this study are the values of Indonesia's non-oil and gas exports, obtained from the official BPS website (https://www.bps.go.id), which consist of monthly data of Indonesia's non-oil and gas export values as variable $(Y_t)$. This study aims to develop a hybrid ARIMA model with stochastic volatility for forecasting Indonesia's non-oil and gas export values. The method used is time series analysis.

### A. Autoregressive Integrated Moving Average (ARIMA)

Autoregressive Integrated Moving Average (ARIMA) is one of the appropriate methods used for forecasting because it only requires the variable to be forecasted in a simple, accurate, and fast manner [8]. In the 1970s, the ARIMA model gained popularity for its ability to solve problems in various broad situations in the field of econometrics. Model Autoregressive Integrated Moving Average (ARIMA) is a short-term forecasting model that has good forecasting accuracy. The difference between the ARIMA method and the ARMA method lies in stationarity, where the ARMA model requires the data to be stationary and does not consider the presence of non-stationary processes. However, the ARIMA method is capable of addressing non-stationary data issues through transformation or differencing processes. Therefore, the equation for the ARIMA model $(p,d,q)$ is as follows:

$$\phi_p(B)(1-B)^d Y_t = \theta_q(B)$$

Where

| | | |
|---|---|---|
| $Y_t$ | = | Observation data at time-$t$ |
| $\phi$ | = | Parameter (AR) |
| $\theta$ | = | Parameter (MA) |
| $(1-B)^d$ | = | Differencing Operator of orde-$d$ |
| $p$ | = | Orde AR |
| $q$ | = | Orde MA |

To validate the best ARIMA model, the Akaike Information Criteria (AIC) is utilized. Model selection through the AIC method involves choosing the lowest AIC value, indicating the suitable model for a given estimated equation, enabling further measurement of forecasting accuracy using MAPE. The goodness-of-fit testing of the obtained ARIMA model is based on whether there is evidence of heteroskedasticity. This is done to ensure that the usage of the ARIMA model is suitable for predictions or whether it's necessary to employ a combined model to reduce errors from the ARIMA model. In time series data, the presence of heteroskedasticity can be determined by observing residual patterns and also by conducting the ARCH-LM test. The ARCH-LM test can detect whether time series data exhibit signs of heteroskedasticity by examining the result values. If there is evidence of heteroskedasticity, it can be proceeded by constructing a model of the residuals that indicate the presence of heteroskedasticity using the ARCH model. If the ARCH model is insufficient in addressing the heteroskedasticity issue in the time series data, the process is extended by utilizing the GARCH model.

### B. GARCH Model

High volatility results in a non-constant moving variance. In time series modeling, the assumption is that the variance should be constant (homoscedasticity). To address this issue, [11] introduced the Autoregressive Conditional Heteroscedasticity (ARCH) model. The ARCH method can handle serial correlation and cases of heteroscedasticity.

For the ARCH model, heteroscedasticity occurs because there is high volatility in the time series data. The ARCH model (p) can be expressed in the form of two equations, as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_p \varepsilon_{t-p}^2$$

or

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2$$

Determining the ARCH model requires large values to obtain an appropriate model for time series data. The use of the Generalized Autoregressive Condition-al Heteroscedasticity (GARCH) model is employed to address the issue of large values in determining the ARCH model [12]. The GARCH model is an extension of the ARCH model. The GARCH equation can be expressed in the form of two equations, as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_p \varepsilon_{t-p}^2 + \beta_1 \sigma_1^2 + \ldots + \beta_q \sigma_{t-q}^2$$

or

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-i}^2$$

Where

| | | |
|---|---|---|
| $\sigma_t^2$ | = | Variance at time-$t$ |
| $\alpha$ | = | Parameter ARCH |
| $\beta$ | = | Parameter GARCH |
| $\varepsilon_t$ | = | Error at time-$t$ |

The GARCH model can be used for data that exhibits the same volatility characteristics, or in other words, equal positive and negative shocks or spikes. This means that it can capture asymmetry in financial data resulting from changes in volatility [13]. To test for the presence of asymmetry effects in the data, a Cross-correlation test or cross-correlation is employed by examining the correlation between the squared standard residuals from the ARIMA model and the residuals from the GARCH model. If the correlation value is not equal to zero, it can be concluded that there is an asymmetry effect in the data [14].

### C. Exponential GARCH Model

Exponential GARCH or EGARCH is a modification of the GARCH model first introduced by Nelson (1991). The Exponential GARCH model, or EGARCH, is a modification of the GARCH model first introduced by Nelson (1991). The advantage of the EGARCH model is that it does not require the data to be stationary, unlike the ARCH, GARCH, and GJR-GARCH models which have stationarity requirements. The equation for the EGARCH model is as follows:

$$\ln\left(\sigma_t^2\right) = \alpha_0 + \alpha_1 \left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| + \gamma_1 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \ldots + \alpha_p \left|\frac{\varepsilon_{t-p}}{\sigma_{t-p}}\right| + \gamma_q \frac{\varepsilon_{t-q}}{\sigma_{t-q}} + \beta_1 \ln(\sigma_{t-1}) + \ldots + \beta_r \ln(\sigma_{r-1})$$

$$\ln\left(\sigma_t^2\right) = \alpha_0 + \sum_{i=1}^{p} \alpha_i \left|\frac{\varepsilon_{t-i}}{\sigma_{t-i}}\right| + \sum_{j=1}^{q} \gamma_j \frac{\varepsilon_{t-j}}{\sigma_{t-i}} + \sum_{k=1}^{r} \beta_k \ln\left(\sigma_t^2\right)$$

From the equation above, it can be seen that there is a ln conditional variance on the left-hand side, indicating that the asymmetric effect is exponential. In the EGARCH equation, the use of ln indicates the non-negativity of the variance. If the value of $\gamma_j \neq 0$, it signifies the presence of asymmetry in the data, indicating that $\gamma_j$ is significant. The presence of a leverage effect is marked by the hypothesis $\gamma < 0$.

### D. GJR-GARCH Model

Glosten, jahathan, and runkle GARCH or another name are GJR-GARCH and TARCH. That is one of model that was first time known as TARCH (Zakoin;1990). Glosten, Jahathan, and Runkle (1993) make this model more specific that modified by GARCH model which the stationary must be fulfilled. The similarity of the GJR-GARCH model is:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_i \varepsilon_{t-i}^2 + \gamma_i \varepsilon_{t-1}^2 I_{t-1}^- + \ldots + \gamma_j \varepsilon_{t-j}^2 I_{t-j}^- + \beta_1 \sigma_{t-1}^2 + \ldots + \beta_{t-k} \sigma_{t-k}^2$$

or

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \gamma_j \varepsilon_{t-j}^2 I_{t-j}^- + \sum_{k=1}^{r} \beta_k \sigma_{t-k}^2$$

Where $I$ is a dummy variable with the condition that $I_{t-1} = 1$ if $\varepsilon_{t-1} < 0$, and $I_{t-1} = 0$ if $\varepsilon_{t-1} > 0$ [15]. For the GJR-GARCH model, the asymmetry effect occurs when $Y_i \neq 0$, indicating that $\gamma$ is significantly different from zero.

### E. Hybrid ARIMA with Stochastic Volatility

A hybrid model in research combines linear and nonlinear models for a given dataset. The purpose of using hybrids is to incorporate nonlinear models without discarding linear ones. One application of hybrid models involves combining the ARIMA model with stochastic volatility models. This leads to the emergence of the Stochastic Volatility Hybrid ARIMA equation as follows:

$$\hat{Y}_t^{hybrid} = \hat{L}_t^{ARIMA} + E_t^{VS}$$

Where

$\hat{L}_t^{ARIMA}$ = $\phi_p(B)(1-B)^d Y_t = \theta_q(B)$
$E_t^{VS}$ = $\sigma_t^2$, (GARCH, EGARCH, or GJR-GARCH)

The equation above shows that represents the linear component, while represents the nonlinear component. After performing predictions and obtaining results using the Stochastic Volatility Hybrid ARIMA, an evaluation of forecasting accuracy will be conducted.

### F. Forecast Accuracy

To select the best model, forecasting accuracy is crucial. The accuracy of a forecasting method is evaluated based on the forecast errors. In this study, error measures and relative measures are used, including Mean Absolute Percentage Error (MAPE). MAPE is used to indicate the magnitude of errors in forecasting by comparing them to the previous values of the series [16]. The equation for calculating MAPE is as follows: $PE_t$ (Percentage Error) It is the percentage error, and n represents the number of errors.

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} |PE_t|$$

$PE_t$ (Percentage Error) It is the percentage error, and $n$ represents the number of errors

$$PE_t = \left(\frac{Y_t - F_t}{Y_t}\right) 100\%$$

where

$Y_t$ = Observation on period of-$t$
$F_t$ = forecast on period of-$t$

For more details, the research was carried out using research steps as follows Figure 1

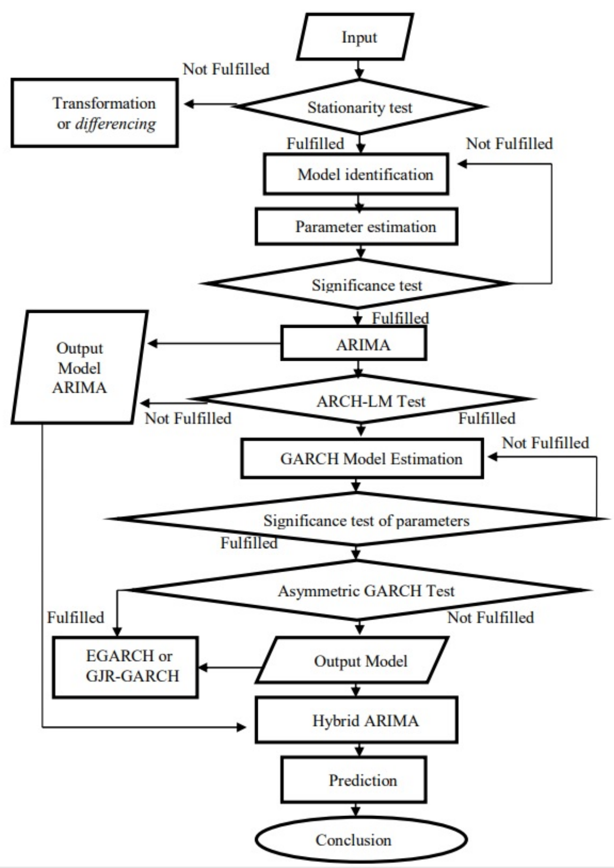


Fig. 1: flowchart

## III. Finding and Discussion

### A. ARIMA Modelling

By using the in-sample data, a time series plot of Indonesia's non-oil and gas export values $(Y_t)$ is created, which is the first step to be taken before developing an ARIMA model, to understand the patterns formed in the data, resulting in the following time series plot:

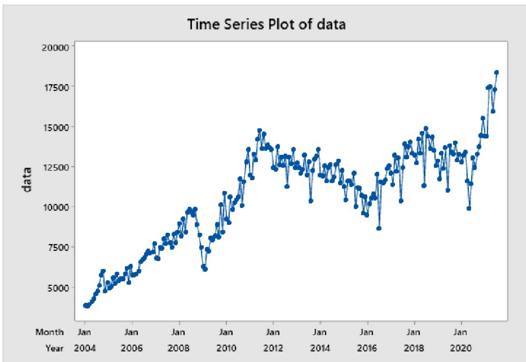


Fig. 2: The Non-oil and Gas Export Values

From Figure 2, it can be observed that the non-oil and gas export values data show an increasing trend or non-constant fluctuations. Additionally, the fluctuations in Indonesia's non-oil and gas export values occur due to unavoidable volatility in the financial market, which is sensitive to changes. From the time series plot above, it can be concluded that the data is not stationary concerning the mean and variance. Through using Box-Cox as result that the rounded value $(\lambda)$ is obtained at 0.00 using a 95% confidence level with an upper limit value of -0.41 and a lower limit of 0.34, so that the data is not stationary for the variance, so a natural logarithmic transformation is performed. The natural logarithmic transformation of the data looks again at the Box-Cox using $ln\, Y_t$.

According to Wei [7] if a data obtains a rounded value $(\lambda)$ of 0.50, it is necessary to carry out a root transformation process $(\sqrt{\ln Y_t})$. For data cases like this, it is used $\ln Y_t$ as initial data so that only one transformation is used in making a linear model from the ARIMA method. Seeing the changes obtained after the transformation of the resulting root $(\sqrt{\ln Y_t})$ data is stationary concerning the variance, then proceed with analyzing whether the data is stationary concerning the mean. For stationary concerning the mean ADF test carried out. From ADF test the stationary data of average is done by a process of differencing. ARIMA model can be made through an analysis ACF and PACF plots:

From Figure 3. plot $\pm\frac{2}{\sqrt{n}}$ the red line is the critical value for knowing autocorrelation and the blue line is interval time (lag). it can be seen that there is *a cut off* at first lag, because it is calculated from the initial five lags only one lag is significant. In the ACF plot,it can be seen that after the initial five lags there are several significant lags but do not form a seasonal pattern so it is sufficient to use the ARIMA process. After viewing the ACF plot it continues by analyzing the pattern of PACF plot.

From Figures 3 and 4, the ACF and PACF plots, it can be seen that the generated data pattern for ACF experiences a cut-

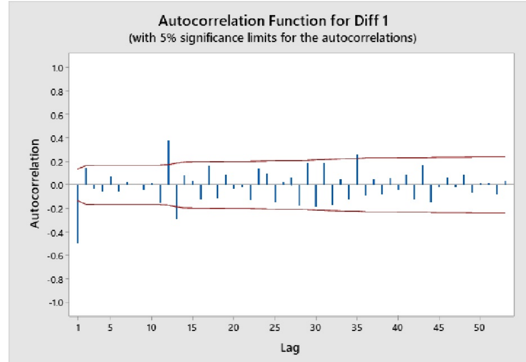


Fig. 3: ACF plot after first differencing

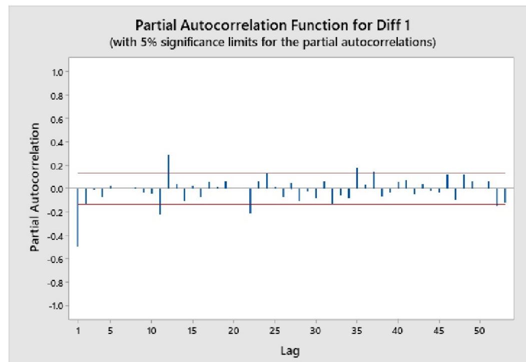


Fig. 4: PACF plot after first differencing

off at the first lag, and for PACF experiences a cut-off at the first lag as well. Therefore, it can be inferred that the pattern contains ARIMA (1,1,0), ARIMA (0,1,1), and ARIMA (1,1,1) models. The selection of the best ARIMA model is performed using the smallest AIC value.

TABLE I: ARIMA Model Selection

| Model | Estimation | *P-Value* | Description | AIC |
|---|---|---|---|---|
| ARIMA (0,1,1) | | | Significant | 0,0075 |
| $\theta_1$ | 0,4982 | 0,0000 | | |
| ARIMA (1,1,0) | | | Significant | 0,0075 |
| $\phi_1$ | -0,4925 | 0,0000 | | |
| ARIMA (1,1,1) | | | | |
| $\theta_1$ | -0,294 | 0,0022 | Significant | 0,0074 |
| $\phi_1$ | 0,266 | 0,041 | | |

In Table I above, it can be seen that ARIMA (1,1,1) is a significant model and has the smallest AIC value among the ARIMA (0,1,1) and ARIMA (1,1,0) models, even though all three models tested are significant. After obtaining the best model, the prediction of the ARIMA model is then carried out.

### B. Prediction of ARIMA

The ARIMA prediction is carried out after obtaining the best ARIMA model. For instance, predicting the value of Indonesia's non-oil and gas exports from August 2021 $(\hat{Y}_{212})$ with the number of time origins used $(n)$=211, by substituting the value of $t$ with $t = 212$ as follows:

$$\ln Y_{212} = \ln Y_{211} + \phi_1 \ln Y_{211} - \phi_1 \ln Y_{210} + e + \theta e_{211}$$

$$\ln\ Y_t = \ln Y_{t-1} - 0,294 \ln Y_{t-1} + 0,294 \ln Y_{t-2} + e_t + 0,266 e_{t-1}$$

By using the same steps for the period from September 2021 to July 2022, a comparison between the actual data and prediction results yields a MAPE value of 13.20824%. Using the residuals from the ARIMA (1,1,1) model, we can test whether there are signs of heteroskedasticity.

### C. ARCH-LM Test

The test used to determine whether there is an ARCH effect or not is the Lagrange Multiplier test, With the statistic test:

$$LM = NR^2$$

From the LM test result, the value obtained is 24,991 with a $\chi^2_{(a;1)}$ value is 3,841, where=0,05. This means that the LM value $> \chi^2_{(a;1)}$, which implies that there are signs of heteroskedasticity in the residuals of the ARIMA (1,1,1) model. In conclusion, the residuals from the ARIMA (1,1,1) model can be modeled using the ARCH or GARCH method.

### D. GARCH Modelling

In the squared residuals of the ARIMA model from the ARCH-LM test, there are signs of heteroscedasticity. Therefore, the ARIMA residuals are used to create a GARCH model with the following possible models:

TABLE II: GARCH Model Selection

| Model | Parameter | Estimation | *P-Value* | Description |
|---|---|---|---|---|
| GARCH (1,0) | $\alpha_0$ | 0,0005759 | 0,00000 | Significant |
| | $\alpha_1$ | 0,210620 | 0,007217 | Not Significant |
| GARCH (0,1) | $\alpha_0$ | 0,000010 | 0,000736 | Significant |
| | $\alpha_1$ | 0,999000 | 0,000000 | Significant |
| GARCH (1,1) | $\alpha_0$ | 0,000010 | 0,000736 | Significant |
| | $\alpha_1$ | 0,000000 | 1,000000 | Not Significant |
| | $\beta_1$ | 0,999000 | 1,000000 | Significant |
| GARCH (2,0) | $\alpha_0$ | 0,005408 | 0,0000 | Significant |
| | $\alpha_1$ | 0,189039 | 0,10289 | Not Significant |
| | $\alpha_2$ | 0,068394 | 0,28035 | Not Significant |
| GARCH (0,2) | $\alpha_0$ | 0,000013 | 0,00000 | Significant |
| | $\beta_1$ | 0,154832 | 0,00000 | Significant |
| | $\beta_2$ | 0,844168 | 0,00000 | Significant |
| GARCH (1,2) | $\alpha_0$ | 0,000013 | 0,34368 | Not Significant |
| | $\alpha_1$ | 0,000000 | 1,00000 | Not Significant |
| | $\beta_1$ | 0,157811 | 0,00000 | Significant |
| | $\beta_2$ | 0,841189 | 0,00000 | Significant |
| GARCH (2,1) | $\alpha_0$ | 0,000011 | 0,00000 | Significant |
| | $\alpha_1$ | 0,000000 | 1,00000 | Not Significant |
| | $\beta_1$ | 0,000000 | 1,00000 | Not Significant |
| | $\beta_2$ | 0,999000 | 0,00000 | Significant |
| GARCH (2,2) | $\alpha_0$ | 0,000013 | 0,39201 | Not Significant |
| | $\alpha_1$ | 0,000000 | 1,00000 | Not Significant |
| | $\alpha_2$ | 0,000000 | 1,00000 | Not Significant |
| | $\beta_1$ | 0,143370 | 0,00000 | Significant |
| | $\beta_2$ | 0,855630 | 0,00000 | Significant |

Through the significance test of the parameters in the table 2, it is found that the GARCH (0,1) and GARCH (0,2) models are significant. Therefore, the best model needs to be selected using the smallest AIC value as follows:

TABLE III: AIC Value of Garch

| No. | Model | AIC |
|---|---|---|
| 1 | GARCH(0,1) | -2,6874 |
| 2 | GARCH(0,2) | -2,6872 |

From Table 3, it can be seen that a model GARCH (0.1) has the smallest AIC value, so it can be used for further testing. In the follow-up tests, it will identify if there are asymmetric effects. There is or not an asymmetric effect characterized by a correlation that is not equal to zero.

### E. Asymmetric Testing

In financial data, an asymmetric effect occurs because there is a positive or negative correlation between the present time value and future volatility values, which can result in the GARCH model needing to be more precise in modeling. To test whether there is an asymmetrical effect on the data, the Correlation test is used with the correlation between the standard squared residuals from ARIMA and the residuals from the GARCH model, where if the correlation value is not equal to zero, then it can be concluded that there is an asymmetric effect on the data. In asymmetric testing, a cross-correlation test is used, where the squared residuals of ARIMA (1,1,1) are correlated with the residuals of GARCH (0,1).

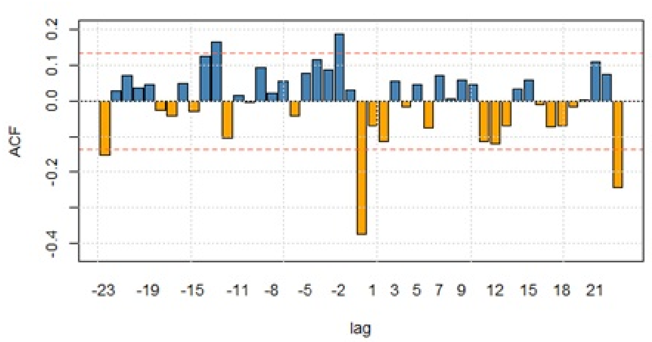


Fig. 5: CCF Plot

From Figure 5 above, it is known that the squared ARIMA (1,1,1) residual and GARCH (0,1) residual models contain asymmetric effects, where the cross-correlation plot produces cross-correlation results. Therefore, it is necessary to use modified methods of GARCH, such as EGARCH and GJR-GARCH, to create models for data experiencing asymmetric effects.

### F. EGARCH Model

The use of the EGARCH model is one method that can model the asymmetric effects that occur in time series data, so several possible EGARCH models can be used as follows:

TABLE IV: EGARCH Model Selection

| Model | Parameter | Estimation | *P-Value* | Description |
|---|---|---|---|---|
| EGARCH (1,1) | $\alpha_0$ | -0,1229 | 0,000736 | Significant |
| | $\alpha_1$ | -0,0514 | 0,000000 | Significant |
| | $\beta_1$ | 0,9753 | 0,000000 | Significant |
| | $\gamma_1$ | -0,1836 | 0,000000 | Significant |
| EGARCH (1,2) | $\alpha_0$ | -0,209637 | 0,000000 | Significant |
| | $\alpha_1$ | -0,50654 | 0,000000 | Significant |
| | $\beta_1$ | 0,601031 | 0,000000 | Significant |
| | $\beta_2$ | 0,356302 | 0,000000 | Significant |
| | $\gamma_1$ | -0,206116 | 0,000000 | Significant |
| EGARCH (2,1) | $\alpha_0$ | -0,223122 | 0,000000 | Significant |
| | $\alpha_1$ | 0,019252 | 0,000000 | Significant |
| | $\alpha_2$ | -0,067247 | 0,000000 | Significant |
| | $\beta_1$ | 0,955972 | 0,000000 | Significant |
| | $\gamma_1$ | 0,073507 | 0,000000 | Significant |
| | $\gamma_2$ | -0,275930 | 0,000000 | Significant |
| EGARCH (2,2) | $\alpha_0$ | -0,5130599 | 0,110073 | Not Significant |
| | $\alpha_1$ | 0,049999 | 0,453136 | Not Significant |
| | $\alpha_2$ | -0,152500 | 0,034917 | Significant |
| | $\beta_1$ | 0,184208 | 0,000033 | Significant |
| | $\beta_2$ | 0,712357 | 0,000000 | Significant |
| | $\gamma_1$ | 0,200551 | 0,211283 | Not Significant |
| | $\gamma_2$ | -0,117625 | 0,190278 | Not Significant |

In Table 4, it can be seen that EGARCH (1,1), EGARCH (1,2), and EGARCH (2,1) have significant estimation values. Before concluding the best EGARCH model, observations on the GJR-GARCH model were carried out.

### *G. GJR-GARCH Model*

GJR-GARCH model is one of the same model as EGARCH that is used for data that have asymmetric effect, but the GJR-GARCH method must qualified to be stationary in the process, which is different from EGARCH. For this data case, GJR-GARCH can use the:

TABLE V: GJR-GARCH Model

| Model | Parameter | Estimation | *P-Value* | Description |
|---|---|---|---|---|
| GJR-EGARCH (1,1) | $\alpha_0$ | 0,000385 | 0,009239 | Significant |
| | $\alpha_1$ | 0,000000 | 1,0000 | Not Significant |
| | $\beta_1$ | 0,925486 | 0,000000 | Significant |
| | $\gamma_1$ | 0,036367 | 0,273499 | Not Significant |
| GJR-EGARCH (1,2) | $\alpha_0$ | 0,000480 | 0,010653 | Significant |
| | $\alpha_1$ | 0,000000 | 1,0000 | Not Significant |
| | $\beta_1$ | 0,590296 | 0,213128 | Not Significant |
| | $\beta_2$ | 0,312080 | 0,5168199 | Not Significant |
| | $\gamma_1$ | 0,053621 | 0,262351 | Not Significant |
| GJR-EGARCH (2,1) | $\alpha_0$ | 0,0001399 | 0,080724 | Not Significant |
| | $\alpha_1$ | 0,000000 | 1,00000 | Not Significant |
| | $\alpha_2$ | 0,000000 | 1,00000 | Not Significant |
| | $\beta_1$ | 0,964584 | 0,000000 | Significant |
| | $\gamma_1$ | 0,213853 | 0,063915 | Not Significant |
| | $\gamma_2$ | -0,184678 | 0,107568 | Not Significant |
| GJR-EGARCH (2,2) | $\alpha_0$ | 0,000139 | 0,092877 | Not Significant |
| | $\alpha_1$ | 0,00000 | 1,00000 | Not Significant |
| | $\alpha_2$ | 0,00000 | 1,00000 | Not Significant |
| | $\beta_1$ | 0,964584 | 0,00000 | Significant |
| | $\beta_2$ | 0,0000 | 0,999991 | Not Significant |
| | $\gamma_1$ | 0,213853 | 0,064062 | Not Significant |
| | $\gamma_2$ | -0,184678 | 0,107676 | Not Significant |

In Table 5. for the estimation of the GJR-GARCH model it can be seen that there is no significant model, so the model that can be used to make predictions is the EGARCH model. Next, using the Akaike Information Criteria (AIC), the AIC results for EGARCH (1,1) are obtained as shown below:

TABLE VI: AIC Value of EGARCH

| No. | Model | AIC |
|---|---|---|
| 1 | EGARCH(1,1) | -0,07277 |
| 2 | EGARCH(1,2) | 1,28440 |
| 3 | EGARCH(2,1) | 1,27730 |

From Table 6, it can be seen that the EGARCH (1,1) model has the smallest AIC value of -0.07277, which means that the EGARCH (1,1) model indicates the appropriate model to be used for making predictions.

### *H. EGARCH Prediction*

The prediction results for the EGARCH (1,1) model, using the t-value replaced with t in the following equation:

$$\ln \sigma_t^2 = -0,1229 - 0,0514 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| - 0,1836 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + 0,9753 \ln\left(\sigma_{t-1}^2\right) \quad (1)$$

They have an impacted sign effect of -0,1836. A sign effect means a negative impact on volatility, resulting in a decline of 0,1836. However, side effects are affected by utility impacts, which means that volatility affects the $t-1$. Therefore, the prediction results are obtained for the EGARCH model for the next 12 periods. Then, a hybrid process is performed to achieve the optimal prediction results.

### *I. Hybrid ARIMA-EGARCH*

After obtaining the prediction results using the ARIMA method and the EGARCH method, the models are combined as follows:

$$\begin{aligned}\hat{Y}_t^{hybrid} &= \ln\ Y_{t-1} - 0,294 \ln\ Y_{t-1} + 0,294 \ln\ Y_{t-2} \\ &+ e_t + 0,266 e_{t-1} - 0,1229 \\ &- 0,0514 \left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right| \\ &- 0,1836 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + 0,9753 \ln\left(\sigma_{t-1}^2\right)\end{aligned}$$

the forecast results for 12 periods from August 2021 to July 2022 are as follows:

TABLE VII: Hybrid ARIMA-EGARCH Prediction

| Period | Actual Data | Prediction | Residual |
|---|---|---|---|
| August 2021 | 20360,3 | 19175 | 1185,34 |
| September 2021 | 19672,8 | 19544,7 | 128,12 |
| October 2021 | 21004,4 | 19630,9 | 1373,53 |
| November 2021 | 21512 | 19799,1 | 1712,94 |
| December 2021 | 21266,1 | 19944,2 | 1321,89 |
| January 2022 | 18272,5 | 20097,6 | -1825,1 |
| February 2022 | 9469,2 | 20249,9 | -780,67 |
| March 2022 | 25092,4 | 20403,7 | 4688,69 |
| April 2022 | 25889 | 20558,7 | 5330,27 |
| May 2022 | 20013,7 | 20714,6 | -700,87 |
| June 2022 | 24600,8 | 20871,6 | 3729,22 |
| July 2022 | 24198,4 | 21029,8 | 3168,62 |

Table 7 above, obtained prediction results with the hybrid ARIMA-EGARCH method for export grade data Indonesian oil and gas from August 2021 until July 2022. the prediction result of the hybrid ARIMA-EGARCH method, we can calculate of MAPE grade to know the accuracy grade of the prediction result.

### J. Hybrid ARIMA-EGARCH Forecasting Accuracy

Next, the forecasting accuracy calculation is performed using the MAPE (Mean Absolute Percentage Error) as follows:

$$MAPE = 9,35158\%$$

Because the forecasting accuracy obtained is very accurate, being less than 10%, it shows that it can reduce the initial ARIMA error, which was 13.20824%, so the model can be used for predicting non-oil export values.

## IV. CONCLUSION

Based on the results of research for data on the fluctuating value of Indonesian non-oil and gas exports, a combined Arima model was used. the use of a hybrid model can reduce the errors that exist in the ARIMA model by considering asymmetric effects to choose the best model that can be combined with the ARIMA model. In this study, the best model was obtained, Hybrid ARIMA-EGARCH with a MAPE value of 9,35158% which shows that the prediction results obtained are very accurate and can later be used to predict the value of non-oil and gas exports in the future.

## REFERENCES


[1] M. Jhingan, *Ekonomi Pembangunan dan Perencanaan*. Rajawali Pers, Jakarta, 2013.

[2] B. Indonesia, "Laporan Perekonomian Indonesia 2018," https://www.bi.go.id/id/publikasi/laporan/Documents/14_LPI2018, 2018, last accessed 22/07/2022.

[3] M. T. Amir, "Inovasi Pendidikan Melalui Problem Based Learning," *Kencana Prenada Media Group*, 2009.

[4] B. P. Statistik, "Statistika Perdagangan Luar Negeri Indonesia Ekspor Menurut Kode ISIC ,2019-2020," https://www.bps.go.id/publication/2021/07/06/9f3d9054c2f29bc478d56cd1, 2019, last accessed 25/03/2022.

[5] Sudjana, *Metode Statistika*. Tarsito, Bandung, (1989).

[6] J. Cryer, *Time Series Analysis*. PWS-Kent Publishing Co, Boston, 1986.

[7] W. S. Wei, *Time Series Analysis: Univariate and Multivariate Methods, 2nd Edition*. Addison Wesley, Boston, 2006.

[8] H. Hartati, "Penggunaan metode ARIMA dalam meramal pergerakan Inflasi," *Jurnal Matematika Sains Dan Teknologi*, vol. 18, no. 1, pp. 1–10, 2017.

[9] W. Y. Eliyawati, R. R. Hidayat, and D. F. Azizah, "Penerapan Model GARCH (Generalized Autoregressive Conditional Heteroscedasticity) Untuk Menguji Pasar Modal Efisien di Indonesia," *Administrasi Bisnis (JAB)*, 2014.

[10] G. P. Zhang, "Time series forecasting using a hybrid arima and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[11] R. F. Engle, "Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation," *Econometrica: Journal of the econometric society*, pp. 987–1007, 1982.

[12] T. Bollerslev, "Generalized autoregressive conditional heteroskedasticity," *Journal of econometrics*, vol. 31, no. 3, pp. 307–327, 1986.

[13] I. S. Mubarokah, A. Fitrianto, and F. M. Affendi, "Perbandingan Model GARCH Simetris dan Asimetris pada Data Kurs Harian," *Indonesian Journal of Statistics and Its Applications*, vol. 4, no. 4, pp. 627–637, 2020.

[14] P. Abiyani, "Peramalan data saham s&p 500 index menggunakan model threshold autoregressive conditional heteroscedastic (tarch)," Ph.D. dissertation, Universitas Negeri Malang, 2013.

[15] M. Ekananda, *Analisis Data Time Series*. Mitra Wacana Media, Jakarta, 2014.

[16] M. A. Maricar, "Analisa perbandingan nilai akurasi moving average dan exponential smoothing untuk sistem peramalan pendapatan pada perusahaan xyz," *Jurnal Sistem dan Informatika (JSI)*, vol. 13, no. 2, pp. 36–45, 2019.