

# FORECASTING THE VALUE OF INDONESIA'S OIL AND GAS IMPORTS USING SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

Vera Maya Santi<sup>1\*</sup>, Rahadian Wahyu<sup>2</sup>, Ibnu Hadi<sup>3</sup>

<sup>1</sup>Statistics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Negeri Jakarta

<sup>2,3</sup>Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Negeri Jakarta  
Rawamangun Muka St., Pulo Gadung, East Jakarta, 13440, Indonesia

Corresponding author's e-mail: \*[vmsanti@unj.ac.id](mailto:vmsanti@unj.ac.id)

## ABSTRACT

### Article History:

Received: 18<sup>th</sup> June 2023

Revised: 20<sup>th</sup> September 2023

Accepted: 19<sup>th</sup> October 2023

### Keywords:

Indonesian Oil and Gas;

MAPE Value;

SARIMA;

Time Series.

The value of Indonesia's oil and gas imports is a combination of the value of crude oil (petroleum), oil, and natural gas products. Throughout 2021, the value of Indonesia's oil and gas imports reached US\$ 25.53 billion or the equivalent of 382.95 trillion rupiah (estimated at US\$ 1 = Rp. 15,000.00). The high demand for petroleum in Indonesia is due to the fact that petroleum is the main source of energy for daily life needs, especially for industrial, transportation, and household needs. The requirement for oil imports is expected to increase along with the growth in Indonesia's population. Therefore, a step is needed to prevent an increase in the value of oil and gas imports in the coming period. One method of analysis that can be used is forecasting using the time series method with the Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The SARIMA model is a time series method with data that has a seasonal pattern and the forecasting results will get a pattern similar to the previous data. The data used is data on the monthly value of oil and gas imports from January 2005 to December 2022 with totaling 216 data. This research aims to find the best model and predict the value of Indonesia's oil and gas imports in the next 12 periods with data test in 4 periods (January to April 2023). The best model for the results of this research is  $(2, 1, 0)(0, 1, 1)^{43}$  with a MAPE value of 13.90%. Based on the accuracy of the MAPE value, this percentage has good quality forecasting results.



This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution-ShareAlike 4.0 International License.

### How to cite this article:

V. M. Santi, R. Wahyu and I. Hadi., "FORECASTING THE VALUE OF INDONESIA'S OIL AND GAS IMPORTS USING SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL," *BAREKENG: J. Math. & App.*, vol. 17, iss. 4, pp. 2047-2058, December, 2023.

Copyright © 2023 Author(s)

Journal homepage: <https://ojs3.unpatti.ac.id/index.php/barekeng/>

Journal e-mail: [barekeng.math@yahoo.com](mailto:barekeng.math@yahoo.com); [barekeng.journal@mail.unpatti.ac.id](mailto:barekeng.journal@mail.unpatti.ac.id)

Research Article · Open Access

## 1. INTRODUCTION

The 1945 Constitution, in chapter 33 article (3), explains that "Earth, water, and the natural resources contained therein are controlled by the state and used as much as possible for the prosperity of the people" [1]. Natural wealth (commonly called natural resources) is the most important asset in creating and maintaining the life axis of a country. Ideally, the more natural resources a country has, the easier it will be for the country to become a country that makes its people prosperous. However, in reality, this is not the case because there are many factors that can influence the success of a country in exploiting its natural resources.

Natural resources are wealth that exists on earth, both in the form of inanimate objects and living things that can be used to meet the needs of human life [2]. The natural resources contained in a country will be one of the greatest potentials for the country's progress. What's more, natural resources are also the main component of the country in meeting human needs, such as oil and gas. Oil and gas are composed of several main elements, and if based on the value of oil and gas is a combination of the value of crude oil (petroleum), the value of oil products, and the value of gas [3].

The value of Indonesia's oil and gas exports throughout 2021 reached US\$ 12.28 billion or the equivalent of 184.2 trillion rupiah (estimated at US\$ 1 = Rp. 15,000.00), while Indonesia's oil and gas import value reached US\$ 25.53 billion or the equivalent of 382.95 trillion rupiah (estimated at US\$ 1 = Rp. 15,000.00) [4]. Besides that, Indonesia also ranks 24th out of 61 countries as oil-producing countries in the world with a magnitude of 692 thousand barrels per day, and on the other hand, Indonesia has a need for petroleum reaching 1.47 million barrels per day [5]. If viewed from the value of exports, these statistics give Indonesia the opportunity to make petroleum an asset in increasing the country's valuation. However, in other situation, Indonesia also has a higher import value. This creates a dilemma that requires the attention of the Indonesian government.

The high demand for petroleum in Indonesia is due to the fact that petroleum is the main source of energy in everyday life [6]. This causes concern that the value of Indonesia's oil and gas imports will increase so that it will affect the prices of cooking fuel, vehicle fuel and other daily materials. Seen in the past time series, the value of oil and gas imports has experienced repeated increases and decreases in several periods. This resulted in the value of oil and gas imports having a seasonal pattern in the data series. Therefore, a step is needed to prevent the increase in these prices, that is by estimating the value of oil and gas imports in the future. Because the data is seasonal, so the steps can be taken are forecasting using the time series method with the Seasonal Autoregressive Integrated Moving Average (SARIMA) model.

SARIMA is a time series model using observational behavior on data that has a seasonal pattern [7] [8][9]. There are several previous researchs that discuss the value of oil and the SARIMA model, including the results of forecasting the number of passengers of PT. Angkasa Pura I (Persero) Yogyakarta Adisutjipto Airport branch office [10]. This research proves that the SARIMA model has more accurate forecasting results where the Mean Squared Deviation (MSD) value is smaller than the Winter's Exponential Smoothing method. In another research, forecasting the number of flood disasters in Indonesia using the SARIMA model also proved that the SARIMA model has a Mean Absolute Percentage Error (MAPE) value of 8.7%, where from the criteria for the accuracy of forecasting the model has a percentage of <10% so that the forecast has very good quality [11]. In addition, in the research on forecasting the Kelayang drought index using the SARIMA and SPI methods it has a mean squared value of 4,111.51 [12]. In research [13] has applied SARIMA to forecast the production of distillate fuel oil refinery and propane blender net, it is also proven that the time series algorithm or SARIMA can be used to forecast the amount of fuel oil production. Based on that analysis, the purpose of this study is to forecast and predict the value of Indonesia's oil and gas import using SARIMA model with analysis spectral to find the periodic in this data.

## 2. RESEARCH METHODS

### 2.1 Research Data

The data used in this research is secondary data, with the data source used is monthly data on the value of Indonesia's oil and gas imports of 216 data from January 2005 to December 2022 collected from the official of Central Statistics Agency (BPS) [14].

## 2.2 Research Stages

The stages of research conducted in this article are as follows:

1. Collect data from the official website of the BPS.
2. Input oil and gas import value data with the help of R software.
3. Display a data plot of the oil and gas import value against time.
4. Looking for seasonal periods contained in the data with spectral analysis.

Spectral analysis is one of methods for analyzing time series in the frequency domain. This method is part of a statistical analysis based on the concept of frequency, which is described in the form of a spectrum (range) [15][16]. Spectral analysis is also a form of Fourier transform, namely a form that converts a time series into a set of sine or cosine waves at various frequency conditions [17][18]. It can be used to find hidden periodicity in time series data. The basic concept and purpose of spectral analysis is to calculate the periodogram and display the power spectral lines.

The periodogram is a function of the spectrum of power to frequency. The application of the periodogram in time series data is to see the hidden periodicity of the time series data, that is, by looking at the frequencies that are paired with the peak points of the spectral lines. The period that has been obtained is then used as the period of the time series model [19].

$$I(\omega_x) = \frac{n(a_x^2 + b_x^2)}{2} \quad (1)$$

where  $a_x$  and  $b_x$  are the Fourier coefficients which are written as follows.

$$a_x = \frac{1}{n} \left[ \sum_{t=1}^n Z_t \cos \frac{2\pi xt}{n} \right]; x = 1, 2, \dots, \left[ \frac{n}{2} \right] \text{ (even number of data)}$$

$$a_x = \frac{2}{n} \left[ \sum_{t=1}^n Z_t \cos \frac{2\pi xt}{n} \right]; x = 1, 2, \dots, \left[ \frac{n-1}{2} \right] \text{ (odd number of data)}$$

$$b_x = \frac{2}{n} \left[ \sum_{t=1}^n Z_t \sin \frac{2\pi xt}{n} \right]; x = 1, 2, \dots, \left[ \frac{n}{2} \right]; x = 1, 2, \dots, \left[ \frac{n-1}{2} \right]$$

$$\omega_x = \frac{2\pi x}{n} \text{ dan } T = \frac{2\pi}{\omega}$$

where  $I(\omega_x)$  = x-th Periodogram,  $\omega_x$  = x-frequency,  $a_x$  = Parameter a fourier coefficient,  $b_x$  = Parameter b fourier coefficient,  $n$  = Lots of data,  $Z_t$  = Actual data in the t-period,  $\pi = 3.14$  or  $\frac{22}{7}$ ,  $t$  = Observation time, and  $T$  = Seasonal period.

5. Testing the stationarity of the data against the variance and mean with Box-Cox Transformation and Augmented Dickey Fuller Test.
6. Carry out the box-cox transformation and differencing process (if the data is not stationary) and re-test the data stationarity.
7. Display a plot of data on oil and gas import values that are already stationary.
8. Specify temporary SARIMA models with general model .

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_Q(B^s)\alpha_t$$

9. Estimate parameter of SARIMA models used maximum likelihood estimation.
10. Test of significance parameters model for AR and MA.

The parameter significance test hypothesis is as follows.

$$H_0: \phi_p = 0 \text{ atau } \theta_q = 0 \text{ (the model parameters are not significant)}$$

$$H_1: \phi_p \neq 0 \text{ atau } \theta_q \neq 0 \text{ (the model parameters are significant)}$$

Significance criterion test, that is reject  $H_0$  if the result  $p - value < \alpha$  or  $|t_{hit}| > t_{\frac{\alpha}{2}, (n-k)}$  where  $k=2$ .

$$|t_{hit}| > t_{\frac{\alpha}{2}, (n-k)}$$

$$|t_{hit}| > t_{\frac{0.05}{2}; (216-2)}$$

$$|t_{hit}| > t_{0.025; (214)}$$

$$|t_{hit}| > 1.971$$

For parameters  $p\text{-value} < 0.05$  or  $|t_{hit}| > 1.971$ , then these parameters are significant model parameters so that they can be used for the feasibility test of the next stage of the model.

11. Carry out residual diagnostics with the Ljung-Box Test whether it meets the White Noise assumption.

The hypothesis test for the assumption of residual white noise is as follows.

$H_0$  : there is no residual correlation between lags (independent)

$H_1$  : there is a residual correlation between lags (not independent)

Accept  $H_0$  or the model has met the assumption of white noise, if  $p\text{-value} > \alpha$ .

12. Carry out residual diagnostics with normal distribution test. The hypothesis test for the assumption of residual is as follows.

$H_0$  : there is a normal distribution of the data

$H_1$  : there is no normal distribution of the data

Accept  $H_0$  or the model has met the assumption of white noise, if  $Q$  (Chi - Square)  $< X^2_{(\alpha, df)}$  with degrees of freedom  $(df) = k - p - q$ .

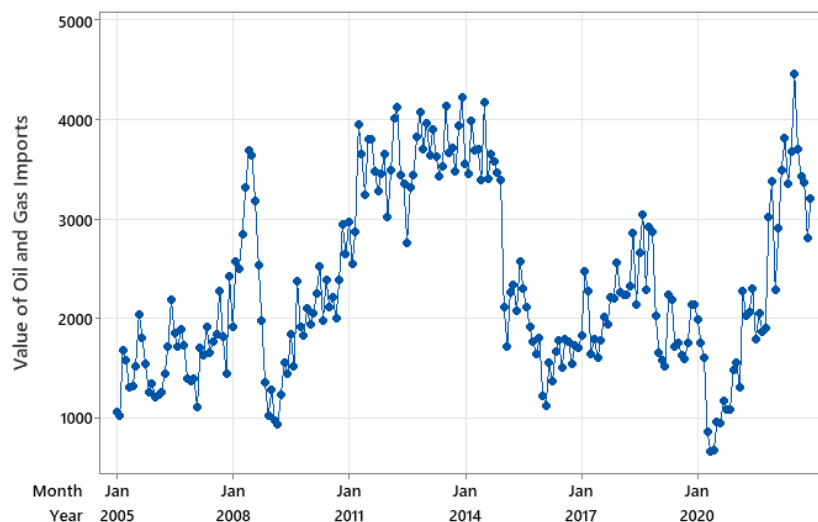
13. Determine the best model from several SARIMA models while then forecasting with the best model.

14. Testing the accuracy of forecasting results using the MAPE value.

### 3. RESULTS AND DISCUSSION

#### 3.1 Data on Indonesia's Oil and Gas Import Value

Data on oil and gas import values used in this article are oil and gas import values from January 2005 to December 2022. The data is monthly data as much as 216 data. The value of oil and gas imports will be displayed in the form of a graphic plot that aims to analyze the data characteristics of the data.



**Figure 1. Time Series Plot of Indonesia's Oil and Gas Imports Values**

**Figure 1** shows a process of repeated increases and decreases in the value of oil and gas imports. The highest value was in July 2022 and the lowest value was in May 2020. The repeated increase and decrease in value indicates that the data contains a seasonal pattern. In addition, it can be seen that the data is not stationary because the data fluctuations are not in a fixed mean range. Therefore, it is necessary to perform spectral analysis to determine the seasonal period in the data and stationarity test to test the stationarity of the data.

### 3.2 Spectral Analysis and Periodogram

By using the equation above, the periodogram value is obtained from the data on Indonesia's oil and gas import values as follows.

**Table 1. Periodogram Value of Indonesia's Oil and Gas Import Value**

$x$	$a_x$	$b_x$	$\omega_x$	$I(\omega_x)$	$T_x$
1	-304.61	217.56	0.029	15132917.4	216
2	229.56	-364.05	0.058	20004861.2	108
3	-33.96	201.52	0.087	4510467.9	72
4	158.98	-307.24	0.116	12924474.3	54
5	59.09	-452.48	0.145	22488816.1	43.2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
105	18	-21.44	3.053	84636.7	2.057
106	1.36	-8	3.082	7111.8	2.038
107	9.46	-10.45	3.111	21459.0	2.019
108	1.54	-4.16	3.14	2125.1	2

Based on **Table 1**, it can be seen that the highest periodogram value is at the 5th frequency with the periodogram value ( $I(\omega_5)$ ) of 22,488,816.1 with value  $\omega_5 = \frac{2\pi(5)}{216} = 0.145$  so we get the seasonal period  $T_5 = \frac{2\pi}{\omega} = 43.2 = 43$ . Based on this, it can be seen that the seasonal period in the value of Indonesian oil and gas imports is 43 months or  $s = 43$ .

### 3.3 Identification of Stationary Data

#### 3.3.1 Stationary Test of Variance

Stationary to variance means that time series data has a constant variance value over time. In addition, this stationary variant also uses a box-cox transformation in the testing process. Following are the results of the stationary test of variance with box-cox test.

**Table 2. Box-cox Value of Indonesia's Oil and Gas Imports**

Test	Value of Lambda
Box-Cox Test	0.689

**Table 2** shows the data on the value of oil and gas imports is not stationary with respect to the variance because the rounded value is 0.689 while a data is said to be stationary if it has a lambda of 1. Therefore, data on the value of oil and gas imports needs to be transformed using a formula  $Z_t^{0.689}$ . After the data is transformed, the data is tested again for stationarity with box-cox.

**Table 3. Box-cox Value of Indonesia's Oil and Gas Imports is Stationary**

Test	Value of Lambda
Box-Cox Test	1.022

**Table 3** shows that the data on the value of oil and gas imports has been said to be stationary with respect to the variant because it has a lambda of 1. Next, the data will be tested for stationarity with respect to the average.

#### 3.3.2 Stationary Test of Mean

Data that is stationarity with respect to the mean can be seen p-value of Augmented Dickey Fuller (ADF) test. On ADF test, if p-value  $< 0.05$  so the data is stationary with respect to the mean. The following ADF values can be seen in **Table 4** below.

**Table 4.** ADF Value of Indonesia's Oil and Gas Imports

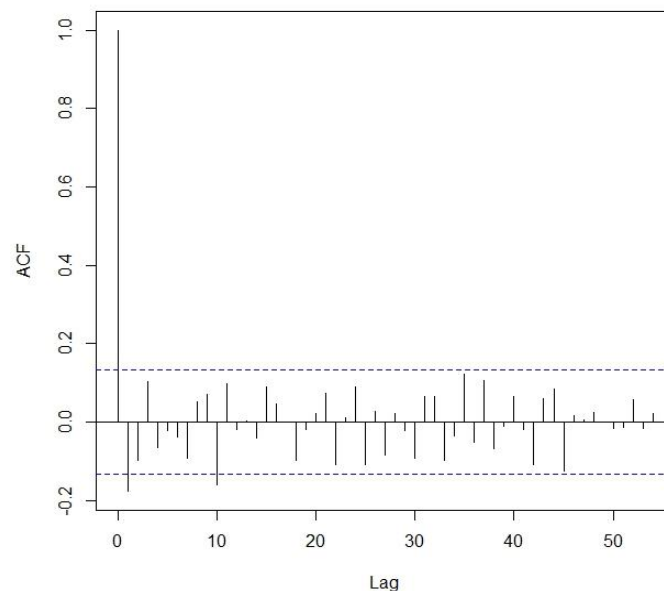
Test	p-value
ADF Test	0.317

**Table 4** shows results of the p-value is  $0.317 > 0.05$  so that the data not yet stationary with respect to the average. Next, the differencing process needs to be carried out to satisfy the stationary test of the mean. The differencing process is carried out by calculating the difference between  $Z_t$  and  $Z_{t-1}$ . Furthermore, the data that has been differentiated will be searched for values again the ADF is to test whether the data is stationary with respect to the average.

**Table 5.** ADF Value of Indonesia's Oil and Gas Imports is Stationary

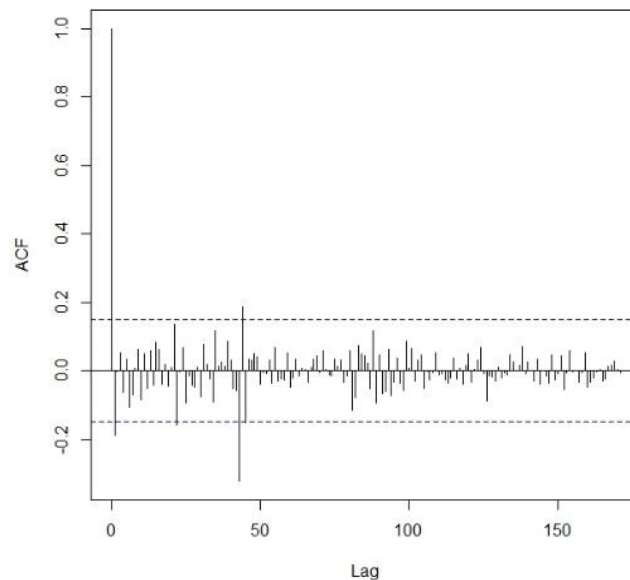
Test	p-value
ADF Test	0.01

In **Table 5** is obtained that the p-value is  $0.01 < 0.05$  so that the data is already stationary with respect to the average. In the next stage, calculations and plotting are carried out on the ACF value to find out the lag that is outside the interval limit to define a provisional model.

**Figure 2.** ACF Graph after Difference 1

In **Figure 2**, the ACF value after differencing results does not appear to decrease continuously. Apart from that, there are also less than 3 lags coming out of confidence intervals. Thus, the data is stationary with respect to the average.

In addition, ACF graphs can also be used to see stationarity on seasonal model of data. Based on Table 1, the seasonal period is obtained at  $s = 43$  so that you can see the ACF graph in **Figure 2** for the lags of seasonal period that is significantly different from zero. On the graph of the ACF it seems that there is no lag outside the confidence interval so that the seasonal model cannot be determined. Therefore, it is necessary to carry out the process of differencing 1 on the seasonal model, that is, on the 43rd lag.



**Figure 3.** ACF Value of Indonesia's Oil and Gas Imports is Stationary to Seasonality

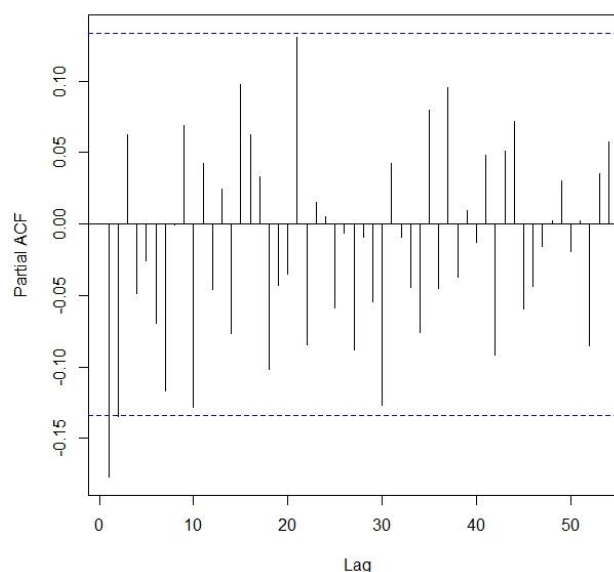
Based on **Figure 3**, the results of seasonal differencing in ACF do not appear to be decreasing continuously. In addition, there is also a seasonal lag outside the confidence interval. Thus, the data is said to be stationary with respect to the average in the seasonal model.

### 3.4 SARIMA Model Identification

The general form of the SARIMA model is  $(p, d, q)(P, D, Q)^s$  [20][21][22]. Order  $s$  is the period of the seasonality, that is  $s = 43$ . This model is a combination of two models, namely non-seasonal and seasonal. Both models are known from the number of lags that come out of the confidence interval limits. The number of values in the ACF and PACF lags that are outside the interval limit or significantly different from zero, will be used as an order in the temporary SARIMA model.

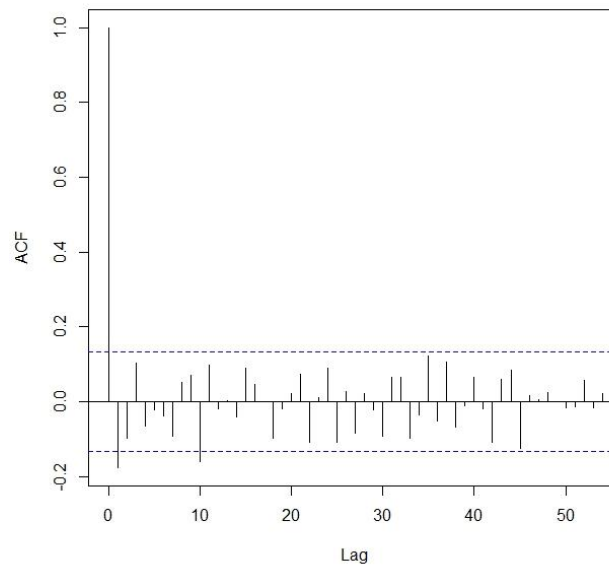
#### 3.4.1 Non-Seasonal Model

In the non-seasonal model, it can be seen in the order  $(p, d, q)$ . The  $p$  value is the order of Autoregressive (AR) on the non-seasonal PACF chart that goes outside the interval limit, the  $d$  value is the number of differencing on the non-seasonal, and the  $q$  value is the order of the Moving Average (MA) on the non-seasonal ACF chart that goes out of the interval limit.



**Figure 4.** PACF Graph for Non-seasonal Modeling



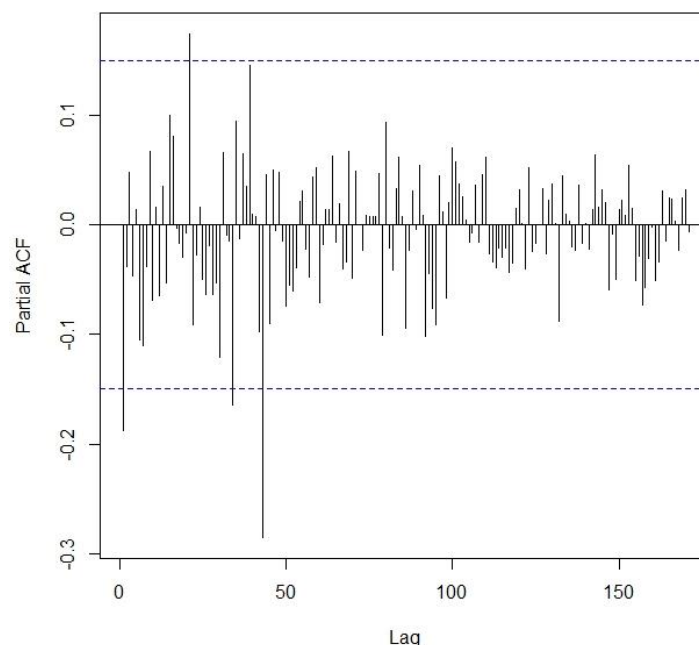


**Figure 5. ACF Graph for Non-Seasonal Modeling**

From **Figure 4**, PACF values that come out of the interval are lag-1 and lag-2, so it can be seen that there are 2 lags that come out of the interval limit on the PACF graph, based on [23] the order of  $p = 2$  atau AR(2). From **Figure 5**, ACF values that come out of the confidence interval are lag-1 and lag-10, so it can be seen that there are 2 lags that come out of the interval limit on the ACF graph, based on [23] the order of  $q = 2$  atau MA(2). In this stationary test process, one differencing is performed, so the value is  $d = 1$ .

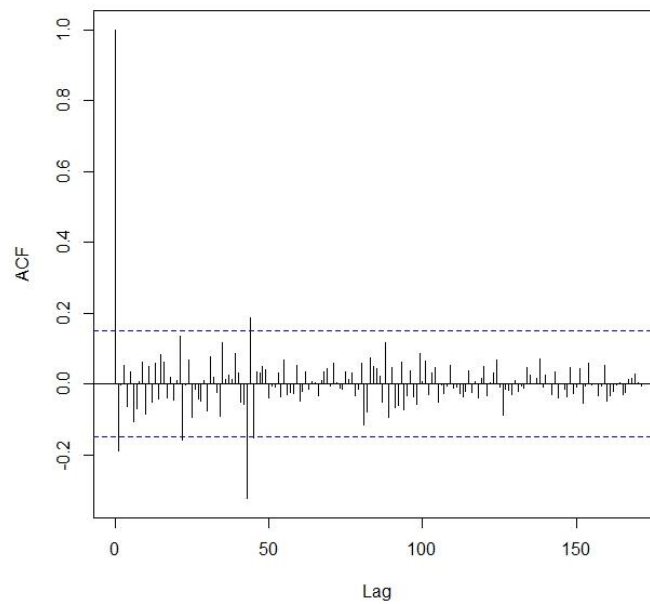
### 3.4.2 Seasonal Model

In the non-seasonal model, it can be seen in the order  $(P, D, Q)$ . The P value is the order of Autoregressive (AR) on the seasonal PACF chart that exits the interval limit, the D value represents the number of seasonal differencing, and the Q value is the order of the Moving Average (MA) on the seasonal ACF chart that exits the interval limit.



**Figure 6. PACF graph for seasonal modeling**





**Figure 7. ACF Graph for Seasonal Modeling**

Based on **Figure 6** and **Figure 7**, the PACF lag that comes out of the interval is lag-43, so the order of  $P = 1$ . The ACF lag that comes out of the confidence interval is lag-43, so the order of  $Q = 1$ . In this stationary test process, differencing once is done for seasonality, so the value of  $D = 1$ .

### 3.4.3 SARIMA Model Determination

Based on the lags that come out at the interval limit above, 15 models can be determined  $(p, d, q)(P, D, Q)^S$ , they are  $(2,1,1)(1,1,1)^{43}$ ,  $(1,1,1)(1,1,1)^{43}$ ,  $(0,1,1)(1,1,1)^{43}$ ,  $(2,1,0)(1,1,1)^{43}$ ,  $(1,1,0)(1,1,1)^{43}$ ,  $(2,1,1)(1,1,0)^{43}$ ,  $(1,1,1)(1,1,0)^{43}$ ,  $(0,1,1)(1,1,0)^{43}$ ,  $(2,1,0)(1,1,0)^{43}$ ,  $(1,1,0)(1,1,0)^{43}$ ,  $(2,1,1)(0,1,1)^{43}$ ,  $(1,1,1)(0,1,1)^{43}$ ,  $(0,1,1)(0,1,1)^{43}$ ,  $(2,1,0)(0,1,1)^{43}$ , dan  $(1,1,0)(0,1,1)^{43}$ .

### 3.5 Estimation and Test of Significance of Parameters of the SARIMA Model

At this stage, parameter estimation is carried out for the temporary SARIMA model. From 15 existing models, only 6 models met the significance test according to **Table 6**.

**Table 6. SARIMA Model Parameter Estimation Results**

Models	Parameters	p-value	Information	Decision
$(1,1,0)(1,1,1)^{43}$	AR(1)	$2.2 \times 10^{-16}$	$p - value < 0.05$	Significance
	SAR(43)	$4.7 \times 10^{-14}$	$p - value < 0.05$	Significance
	SMA(43)	$2.2 \times 10^{-16}$	$p - value < 0.05$	Significance
$(2,1,0)(1,1,0)^{43}$	AR(1)	0.002	$p - value < 0.05$	Significance
	AR(2)	0.033	$p - value < 0.05$	Significance
	SAR(43)	$4.7 \times 10^{-10}$	$p - value < 0.05$	Significance
$(0,1,1)(1,1,0)^{43}$	MA(1)	0.001	$p - value < 0.05$	Significance
	SAR(43)	$4.8 \times 10^{-9}$	$p - value < 0.05$	Significance
$(2,1,0)(0,1,1)^{43}$	AR(1)	0.001	$p - value < 0.05$	Significance
	AR(2)	0.031	$p - value < 0.05$	Significance
	SMA(43)	$3.4 \times 10^{-6}$	$p - value < 0.05$	Significance
$(1,1,0)(0,1,1)^{43}$	AR(1)	0.008	$p - value < 0.05$	Significance
	SMA(43)	$1.2 \times 10^{-6}$	$p - value < 0.05$	Significance
$(0,1,1)(0,1,1)^{43}$	MA(1)	0.001	$p - value < 0.05$	Significance
	SMA(43)	$1.5 \times 10^{-6}$	$p - value < 0.05$	Significance

### 3.6 Test the Residual Assumptions of the SARIMA Model

A model can be said to be a feasible model if it meets the residual white noise assumption and normal distribution test. At this stage, the Ljung-Box test will be carried out to find out whether the model meets the residual white noise assumption and Chi-Square Test to meet the normal distribution test. From 6 models that

met the significance test criteria, all of models met the residual white noise assumption and normal distribution test as shown in the table below.

**Table 7. White Noise Assumption Test Results**

Models	p-value	Information	Decision
(1,1,0)(1,1,1) <sup>43</sup>	0.614	$p - value > 0.05$	Independent
(2,1,0)(1,1,0) <sup>43</sup>	0.967	$p - value > 0.05$	Independent
(0,1,1)(1,1,0) <sup>43</sup>	0.898	$p - value > 0.05$	Independent
(2,1,0)(0,1,1) <sup>43</sup>	0.977	$p - value > 0.05$	Independent
(1,1,0)(0,1,1) <sup>43</sup>	0.560	$p - value > 0.05$	Independent
(0,1,1)(0,1,1) <sup>43</sup>	0.967	$p - value > 0.05$	Independent

Based on **Table 7**, it can be seen that there all of models that meet the assumption of white noise.

**Table 8. Normal Distribution Test Results**

Models	Type	12th Lag	24th Lag	36th Lag	48th Lag	Result
(1,1,0)(1,1,1) <sup>43</sup>	Df	8	20	32	44	Not normally distribute data
	Chi-Square	26.47	43.33	66.13	91.68	
	$\chi^2_{(0.05; df)}$	15.51	31.41	46.19	60.48	
(2,1,0)(1,1,0) <sup>43</sup>	Df	8	20	32	44	Normally distribute data
	Chi-Square	9.37	22.29	35.16	49.52	
	$\chi^2_{(0.05; df)}$	15.51	31.41	46.19	60.48	
(0,1,1)(1,1,0) <sup>43</sup>	Df	9	21	33	45	Not normally distribute data
	Chi-Square	20.42	37.64	55.12	79.19	
	$\chi^2_{(0.05; df)}$	16.92	32.67	47.40	61.66	
(2,1,1)(0,1,1) <sup>43</sup>	Df	8	20	32	44	Normally distribute data
	Chi-Square	5.97	18.77	31.83	48.94	
	$\chi^2_{(0.05; df)}$	15.51	31.41	46.19	60.48	
(1,1,0)(0,1,1) <sup>43</sup>	Df	9	21	33	45	Not normally distribute data
	Chi-Square	16.78	33.00	49.10	70.72	
	$\chi^2_{(0.05; df)}$	16.92	32.67	47.40	61.66	
(0,1,1)(0,1,1) <sup>43</sup>	Df	9	21	33	45	Normally distribute data
	Chi-Square	11.74	26.35	41.33	60.77	
	$\chi^2_{(0.05; df)}$	16.92	32.67	47.40	61.66	

Based on **Table 8**, only 3 models meet normal distribution test.

### 3.7 Best Model Selection

After identifying the model, estimating parameters, and testing the model on the value of Indonesia's oil and gas imports, some of the best models are obtained and are suitable for forecasting. Next, the selection of the best model is carried out by looking at the smallest Mean Absolute Error (MAE) value in the several models. Below are the MAE values for some SARIMA models.

**Table 9. Best Model Selection**

SARIMA Model	MAE Value
(2,1,0)(1,1,0) <sup>43</sup> ,	304.54
(2,1,0)(0,1,1) <sup>43</sup>	286.15
(0,1,1)(0,1,1) <sup>43</sup>	286.71

**Table 9** shows that the SARIMA model (2,1,0)(0,1,1)<sup>43</sup> is the best model and feasible to use for forecasting. This model was chosen because it has the smallest MAE value.

### 3.8 Forecasting Results

After getting the SARIMA model (2,1,0)(0,1,1)<sup>43</sup> as the best model where the general form of the SARIMA model is as follows.

$$\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_Q(B^s)\alpha_t$$

Furthermore, from **Table 9**, the results of forecasting the value of Indonesia's oil and gas imports are obtained as in **Table 10** below.

**Table 10. SARIMA Model Forecasting Results (2, 1, 0)(0, 1, 1)<sup>43</sup>**

Period	Forecasting Results
January 2023	2978.13
February 2023	2826.88
March 2023	2554.76
April 2023	2352.19
May 2023	2636.55
June 2023	2726.38
July 2023	2919.59
August 2023	2949.81
September 2023	2751.60
October 2023	2770.15
November 2023	2507.37
December 2023	2374.20

### 3.9 Forecasting Accuracy Rate

The final stage in this forecasting is evaluation by comparing the actual value with the results of forecasting the value of Indonesia's oil and gas imports. For the test with compared data from January 2023 to April 2023. The purpose of this evaluation is to compare the error rate of the MAPE value prediction results. The following is the evaluation of forecasting the value of Indonesia's oil and gas imports using the SARIMA model (2,1,0)(0,1,1)<sup>43</sup>.

**Table 11. MAPE Value of SARIMA Model (2, 1, 0)(0, 1, 1)<sup>43</sup>**

Data to	Period	Actual Value	Forecasting Value	MAPE Value
217	January 2023	2906.1	2978.13	2.48 %
218	February 2023	2406.5	2826.44	17.47 %
219	March 2023	3014.8	2554.76	15.26 %
220	April 2023	2955.0	2352.19	20.40 %
Overall MAPE Score				13.90 %

According to **Table 11**, the MAPE value is obtained from forecasting the value of oil and gas imports using the SARIMA model (2,1,0)(0,1,1)<sup>43</sup> with a MAPE value of 13.90%. From the results of the evaluation, the MAPE value can be seen to be between 10-20%, which means it has good forecasting abilities.

## 4. CONCLUSIONS

Based on the results and discussion regarding forecasting the value of Indonesia's oil and gas imports using the SARIMA model, the following conclusions can be drawn :

1. The best SARIMA model formed from data on Indonesia's oil and gas import values from January 2005 to December 2022 is the SARIMA model (2,1,0)(0,1,1)<sup>43</sup>.
2. The forecasting results obtained using the SARIMA model produce a MAPE value of 13.90% which means the quality of good forecasting results.
3. SARIMA Model (2,1,0)(0,1,1)<sup>43</sup> used to predict the value of Indonesia's oil and gas imports from January 2023 to December 2023 obtained 2978.13; 2826.88; 2554.76; 2352.19; 2636.55; 2726.38; 2919.59; 2949.81; 2751.60; 2770.15; 2507.37 and 2374.20 (million US\$).

## ACKNOWLEDGMENT

Thank you to Faculty of Mathematics and Natural Sciences Jakarta State University for funding and support this research.

## REFERENCES

- [1] A. Redi, "Dinamika Konsepsi Penguasaan Negara Atas Sumber Daya Alam," *J. Konstitusi*, vol. 12, no. 2, p. 401, 2016.
- [2] Surtani, "Peran Serta Masyarakat dalam Pemanfaatan Sumber Daya Alam Secara Efektif dan Efisien Sumber Daya Alam Secara Efektif dan Efisien," *Pros. Semin. Nas. Geogr. 2016*, p. 320, 2016, [Online]. Available: [http://repository.unp.ac.id/653/1/SURTANI 18.pdf](http://repository.unp.ac.id/653/1/SURTANI%2018.pdf)
- [3] Kementerian ESDM, "Minyak dan Gas Bumi Semester I 2021," *Miny. dan Gas Bumi Semester I 2021*, p. 104, 2021.
- [4] V. B. Kusnandar, "Defisit Neraca Perdagangan Migas Indonesia Catat Rekor Terdalam pada Desember 2021," <https://databoks.katadata.co.id/datapublish/2022/01/19/defisit-neraca-perdagangan-migas-indonesia-catat-rekor-terdalam-pada-Desember-2021>, 2022.
- [5] T. Purwanti, "Ini 5 Negara dengan Cadangan Minyak Terbesar di Dunia," *CNBC Indonesia*, 2022. <https://www.cnbcindonesia.com/market/20220912082219-17-371169/ini-5-negara-dengan-cadangan-minyak-terbesar-di-dunia>
- [6] E. M. Tamboesai, "Kajian Korelasi Genetika Geokimia Molekular Minyak Bumi Cekungan Sumatera Tengah, Riau," *J. Ind. Che. Acta*, vol. 3, no. 1, pp. 5–10, 2012, [Online]. Available: <https://ica.ejournal.unri.ac.id/index.php/ICA/article/view/913/906>
- [7] H. Feng, G. Duan, R. Zhang, and W. Zhang, "Time series analysis of hand-foot-mouth disease hospitalization in Zhengzhou: Establishment of forecasting models using climate variables as predictors," *PLoS One*, vol. 9, no. 1, pp. 1–10, 2014.
- [8] M. Giovani, I. Anggriani, and S. A. W. D. Simatupang, "Comparison in Predicting the Short-Term Using the Sarima, Dsarima and Tsarima Methods," *BAREKENG J. Ilmu Mat. dan Terap.*, vol. 16, no. 4, pp. 1487–1496, 2022.
- [9] T. C. Mills, "Time series modelling of temperatures: An example from Kefalonia," *Meteorol. Appl.*, vol. 21, no. 3, pp. 578–584, 2014.
- [10] A. N. Munawaroh, "Peramalan Jumlah Penumpang pada PT. Angkasa Pura I (Persero) Kantor Cabang Bandar Udara Internasional Adisutjipto Yogyakarta dengan Metode Winter's Exponential Smoothing dan Seasonal ARIMA," Universitas Negeri Yogyakarta, 2010.
- [11] N. AZIZAH, *PERAMALAN JUMLAH BENCANA BANJIR DI INDONESIA MENGGUNAKAN METODE SEASONAL AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (SARIMA)*, vol. 2, no. 8.5.2017. 2022.
- [12] W. Rahmalina and Novreta, "Peramalan Indeks Kekeringan Kelayang Menggunakan Metode Sarima dan SPI," *Potensi J. Sipil Politek.*, vol. 22, no. 1, pp. 64–75, 2020.
- [13] A. Mishra, R. Morisetty, and R. Sarawagi, "Forecasting the production of Distillate Fuel Oil Refinery and Propane Blender net production by using Time Series Algorithms," pp. 1–21, 2022, [Online]. Available: <https://arxiv.org/ftp/arxiv/papers/2208/2208.05964.pdf>
- [14] B. P. Statistik, "Nilai Impor Migas-Non Migas (Juta US\$), 2005-2022," *Badan Pusat Statistik*. <https://www.bps.go.id/indicator/8/1754/2/nilai-impor-migas-nonmigas.html> (accessed Jun. 10, 2023).
- [15] M. Fiecas, C. Leng, W. Liu, and Y. Yu, "Spectral analysis of high-dimensional time series," *Electron. J. Stat.*, vol. 13, no. 2, pp. 4079–4101, 2019.
- [16] D. Grzesica and P. Więcek, "Advanced Forecasting Methods Based on Spectral Analysis," *Procedia Eng.*, vol. 161, pp. 253–258, 2016.
- [17] E. Ghaderpour, S. D. Pagiatakis, and Q. K. Hassan, "A survey on change detection and time series analysis with applications," *Appl. Sci.*, vol. 11, no. 13, 2021.
- [18] P. Bloomfield, *Fourier Analysis of Time : An Introduction*, Second Edi. Canada: A Wiley-Interscience Publication, 2004. [Online]. Available: <https://books.google.co.id/books?hl=id&lr=&id=zQsupRg5rrAC&oi=fnd&pg=PR11&dq=Spectral+analysis+is+also+a+form+of+Fourier+transform,+namely+a+form+that+converts+a+time+series+into+a+set+of+sine+or+cosine+waves+at+various+frequency+conditions.+It+can+be+use>
- [19] J. D. Scargle, "Studies in astronomical time series analysis. II-Statistical aspects of spectral analysis of unevenly spaced data," *Astrophys. J.*, vol. 263, pp. 835–853, [Online]. Available: <https://adsabs.harvard.edu/full/1982ApJ...263..835S7>
- [20] A. H. Adineh, Z. Narimani, and S. C. Satapathy, "Importance of data preprocessing in time series prediction using SARIMA: A case study," *Int. J. Knowledge-Based Intell. Eng. Syst.*, vol. 24, no. 4, pp. 331–342, 2021.
- [21] M. Othman, R. Indawati, A. A. Suleiman, M. B. Qomaruddin, and R. Sokkalingam, "Model Forecasting Development for Dengue Fever Incidence in Surabaya City Using Time Series Analysis," *Processes*, vol. 10, no. 11, 2022.
- [22] B. SENNAROĞLU and W. ZAYAT, "Performance Comparison of Holt-Winters and SARIMA Models for Tourism Forecasting in Turkey," *Doğuş Üniversitesi Derg.*, vol. 21, no. 2, pp. 63–77, 2020.
- [23] Zulhamidi and R. Hardianto, "PERAMALAN PENJUALAN TEH HIJAU DENGAN METODE ARIMA (STUDI KASUS PADA PT. MK)," *J. PASTI*, vol. XI, no. 3, pp. 231–244, 2017.