Quantitative Economics and Management Studies (QEMS) *ISSN: 2722-6247 (online)*
Vol. 4 No. 5 (2023)
https://doi.org/10.35877/454RI.qems1040

Forecasting the Export Value of Oil and Gas in Indonesia using Autoregressive Integrated Moving Average (ARIMA)

Ansari Saleh Ahmar[a,*], Abdul Rahman[b], Parkhimenko Vladimir Anatolievich[c], Rusli[b], & Sitti Masyitah Meliyana R.[a]

[a]*Department of Statistics, Universitas Negeri Makassar, Makassar, 90223, Indonesia*
[b]*Faculty of Engineering and Economics, Belarusian State University of Informatics and Radioelectronics (BSUIR), Belarus*
[c]*Department of Statistics, Universitas Negeri Makassar, Makassar, 90223, Indonesia*

Abstract

This study aims to utilize the ARIMA method to predict the value of Indonesia's oil and gas exports. As quantitative research, it employs secondary data sourced from the Central Bureau of Statistics of the Republic of Indonesia's website. The data spans January 2010 to March 2022 and are presented on a monthly basis. Through the results and discussion, three ARIMA models were established, namely ARIMA (1,1,0), ARIMA (0,1,1), and ARIMA (1,1,1). Among these models, the ARIMA (0,1,1) model with an AIC value of 2047.65 was found to be the most suitable for forecasting Indonesia's oil and gas exports. The forecasted values for the next five periods were 1254.124 (April 2022), 1309.678 (May 2022), 1289.236 (June 2022), 1296.758 (July 2022), and 1293.990 (August 2022).

Keywords: forecasting; ARIMA; oil and gas.

Received: 4 April 2023 Revised: 13 July 2023 Accepted: 21 September 2023

1. Introduction

A country's economic growth can be gauged by its productivity success rate. Both developed and developing nations recognize exports as a means of achieving economic growth. Exports sell products from within a country to foreign markets (Farina & Husaini, 2017). Engaging in export activities has numerous advantages such as increasing state revenue in the form of foreign exchange, creating numerous domestic jobs, and stimulating economic growth by expanding the market (Sihombing et al., 2021).

Generally, Indonesia's export products are divided into two categories: oil and gas exports and non-oil and gas exports. According to Indonesian Law No. 22 of 2021, oil and gas are crucial commodities that significantly contribute to the nation's economy, and their management must be optimized for the well-being of people. Unfortunately, over the past decade, from 2012 to 2021, there has been a shortfall in Indonesia's oil and gas export trade balance, indicating that the country's expenditures have exceeded its income.

The deficit in oil and gas exports has a detrimental effect on national development, as there is a lack of funds available for investment. To address this issue, the government implemented a foreign debt policy to obtain funding. However, relying heavily on foreign debt can reduce foreign exchange reserves. A decrease in foreign exchange reserves can negatively affect economic growth, causing the value of the national currency to depreciate.

To address the persistent oil and gas export deficit in Indonesia, careful planning and improvement measures are essential. Both short-term and long-term planning can serve as a guide for actions that must be taken to prevent future oil and gas export shortages. The planning process can be initiated by gathering information on the projected value of future oil and gas exports (Hayati et al. 2021). To accurately estimate the future value of oil and gas exports, a reliable

[*] Corresponding author.
E-mail address: ansarisaleh@unm.ac.id

forecasting method is required, which can be utilized by the National Export Development Agency (BPEN) and the Ministry of Trade of Indonesia (Yulisa et al., 2023) to formulate effective strategies.

Predicting future events involves both art and science, and the results of these predictions can be utilized by policymakers to create strategic policies (Utami & Darsyah, 2016; Singh, et.al, 2021). The time-series method involves examining the connections between the variables measured over time (Ahmar, et. al., 2022). A commonly used time series analysis method is the Box-Jenkins model, also known as the Autoregressive Integrated Moving Average (ARIMA) model, which is a type of time-series forecasting model that is based on the behavior of recorded variable data.

2. Methods

The objective of this quantitative study was to use secondary data as the primary source of information. Specifically, data were obtained from the website of the Central Bureau of Statistics of the Republic of Indonesia, which provides monthly data on the value of oil and gas exports in Indonesia from January 2010 to March 2022. To achieve the research objectives, the ARIMA model was employed for forecasting.

3. Result and Discussions

To initiate the forecasting process, it is crucial to visualize the data and identify their patterns and characteristics. This can be achieved by plotting the data, as shown in Fig. 1.

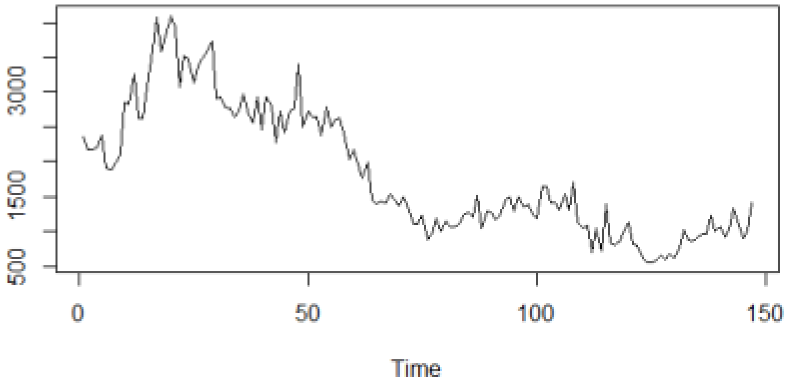

Figure 1. Time-series plot of data on the value of oil and gas exports in Indonesia.

From Figure 1, it is evident that the data pattern has not yet achieved stationarity on average. This conclusion is further supported by the Augmented Dickey-Fuller test results presented below.

```
> adf.test(data_migas.ts)

        Augmented Dickey-Fuller Test

data:  data_migas.ts
Dickey-Fuller = -2.548, Lag order = 5, p-value = 0.3482
alternative hypothesis: stationary
```

The Augmented Dickey-Fuller test suggests that the p-value of 0.3482, which is greater than the significance level of $\alpha = 0.05$, indicates that the data are not stationary on average.

The next step involves conducting a Box-Cox test to determine whether the variance of the data is stationary. The results of the test are as follows.

```
> BoxCox.lambda(data_migas.ts)
[1] 0.3679229
```

The Box-Cox test revealed that the value of λ is 0.3679229, implying that the variance of the data is not yet stationary; therefore, transformation is required. The specific types of transformations employed are shown in Table 1.

Table 1. Type of Transformation Based on λ Value

From Table 1, it can be seen that the type of transformation performed was $\sqrt{Z_t}$. In the next stage, the data were transformed as follows:

```
> transzt <- sqrt(data_migas.ts)
> transzt
          Jan      Feb      Mar      Apr      May      Jun      Jul      Aug
2010 48.42417 46.64011 46.56823 46.95317 48.67546 43.60619 43.37511 44.64863
2011 51.13707 51.11262 55.33353 60.23537 63.81849 59.92495 61.66441 63.96562
2012 56.05890 57.92668 59.04320 59.67160 61.03196 53.84886 54.03425 52.75415
2013 51.51408 50.67149 54.11377 49.51868 54.09436 52.91881 47.77656 52.15841
2014 50.01700 52.24079 51.39358 51.49175 48.74115 52.78257 49.96299 50.97254
2015 44.26059 41.87362 44.59709 38.18639 37.31890 37.94601 37.70676 39.12672
2016 33.28663 33.36615 35.20369 29.86135 30.95158 34.45867 31.60063 33.74315
2017 35.75752 34.76492 38.93841 32.19006 35.97777 35.72534 34.13210 35.12264
2018 36.64287 37.26661 35.44150 34.33366 40.41163 40.57955 37.63642 37.73195
2019 33.63480 32.41605 32.82377 26.23166 32.46845 26.72265 37.42325 29.03274
2020 28.55346 28.37605 24.84754 23.70865 23.68333 23.82016 25.69825 24.48673
2021 29.72877 29.33598 30.13138 31.02257 31.11913 35.10128 31.77420 32.66190
2022 30.01999 31.54045 37.48466
          Sep      Oct      Nov      Dec
2010 45.63880 53.30947 53.06977 57.09028
2011 62.69769 55.34167 59.35318 59.03389
2012 52.63554 51.48398 52.12581 54.46926
2013 49.13960 52.10758 52.60038 58.35323
2014 51.21133 49.12433 45.11541 46.56179
2015 38.12611 37.14297 38.69108 36.04858
2016 32.58067 32.49461 33.21144 35.35817
2017 38.14446 38.57720 35.99722 38.68462
2018 36.33456 39.31030 36.23396 41.31344
2019 28.33725 29.32576 32.15121 33.66452
2020 25.83215 24.78911 27.60797 31.91865
2021 30.54177 32.02031 36.50205 33.06660
2022
```

Following data transformation, the Box-Cox test was conducted once more to determine whether the variance of the data was stationary. The results of the Box-Cox test are presented below.

```
> BoxCox.lambda(transzt)
[1] 0.7096078
```

The Box-Cox test indicates that the value of λ moves closer to 1, suggesting that the variance of the data has become stationary.

The next step involves stabilizing the data that have not been stationary, which is achieved through the process of differencing. The resulting difference data are presented below:

```
> diff1 <- diff(transzt, differences = 1)
> diff1
             Jan         Feb         Mar         Apr         May         Jun
2010             -1.78405603 -0.07188198  0.38493864  1.72228767 -5.06926361
2011 -5.95320868 -0.02444995  4.22091339  4.90183753  3.58312095 -3.89353955
2012 -2.97499097  1.86778019  1.11652673  0.62839677  1.36035756 -7.18309627
2013 -2.95518164 -0.84258475  3.44227804 -4.59508570  4.57567866 -1.17555627
2014 -8.33623759  2.22379155 -0.84720930  0.09816755 -2.75059387  4.04141984
2015 -2.30119495 -2.38697257  2.72346573 -6.41070010 -0.86748899  0.62711827
2016 -2.76194438  0.07951656  1.83754248 -5.34234674  1.09022883  3.50709582
2017  0.39934932 -0.99259184  4.17348806 -6.74835233  3.78771034 -0.25242977
2018 -2.04175120  0.62373597 -1.82510594 -1.10784414  6.07797503  0.16791951
2019 -7.67863267 -1.21875799  0.40772646 -6.59211161  6.23678596 -5.74579665
2020 -5.11106275 -0.17741240 -3.52851115 -1.13888688 -0.02532074  0.13683206
2021 -2.18987265 -0.39278921  0.79539426  0.89119343  0.09655352  3.98215608
2022 -3.04660612  1.52045688  5.94421331
             Jul         Aug         Sep         Oct         Nov         Dec
2010 -0.23108415  1.27352013  0.99017108  7.67067455 -0.23970843  4.02051398
2011  1.73946131  2.30120139 -1.26792841 -7.35602132  4.01151416 -0.31929163
2012  0.18538583 -1.28010164 -0.11860743 -1.15156123  0.64183126  2.34344804
```

```
2013 -5.14224244  4.38184926 -3.01881548  2.96798395  0.49279920  5.75285447
2014 -2.81958658  1.00955533  0.23878524 -2.08699482 -4.00892449  1.44637934
2015 -0.23925096  1.41995390 -1.00061067 -0.98313951  1.54811721 -2.64250591
2016 -2.85803799  2.14251455 -1.16247834 -0.08605418  0.71682887  2.14672356
2017 -1.59324481  0.99054594  3.02181998  0.43273309 -2.57997322  2.68740015
2018 -2.94313297  0.09553079 -1.39739277  2.97574744 -3.07634202  5.07947410
2019 10.70060523 -8.39051466 -0.69548551  0.98850197  2.82544867  1.51331612
2020  1.87808944 -1.21151787  1.34541941 -1.04304000  2.81885935  4.31067672
2021 -3.32707865  0.88770104 -2.12012974  1.47853139  4.48174868 -3.43545527
2022
```

After performing the differencing process, the Augmented Dickey-Fuller test was repeated to assess whether the data had become stationary, on average. Otherwise, the differencing process is applied once more. The outcomes of the Augmented Dickey-Fuller test are as follows.

```
> adf.test(diff1)

        Augmented Dickey-Fuller Test

data:  diff1
Dickey-Fuller = -5.3466, Lag order = 5, p-value = 0.01
alternative hypothesis: stationary
```

In the Augmented Dickey-Fuller test, the p-value is 0.01, which is smaller than the significance level of $\alpha = 0.05$, indicating that the data has been stationary on average. This finding is supported by the time-series plot for differencing presented in Figure 2, which reveals that the data has stabilized.

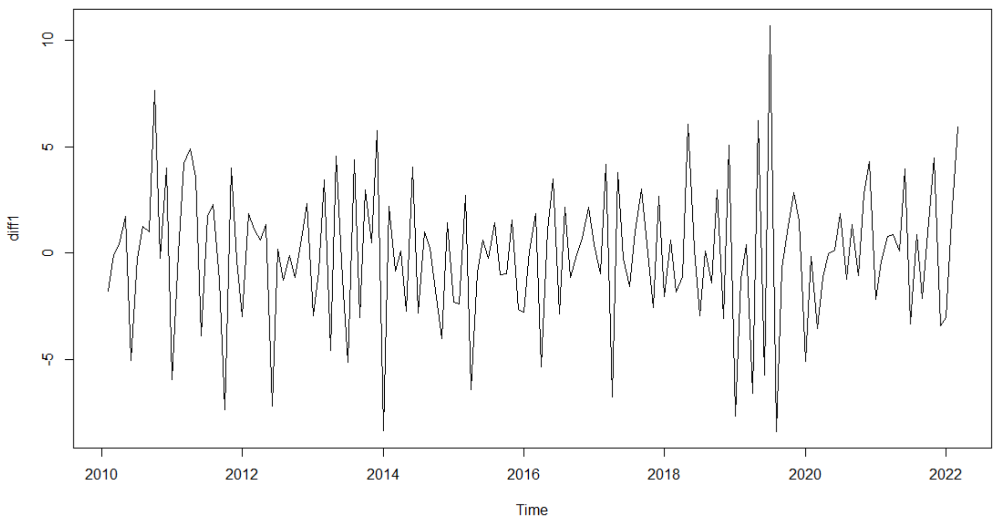

Figure 2. Time series plot for diff1

To identify the appropriate ARIMA model, we examined the ACF and PACF plots presented in Figures 3 and 4.

By examining figures 3 and 4, it is evident that the ACF plot exhibits a decline after the first lag, coinciding with the autocorrelation value at that lag, while the PACF plot diminishes and cuts off after the first lag. Consequently, the ARIMA models that have been established are ARIMA (1,1,0), ARIMA (0,1,1), and ARIMA (1,1,1).

The next step involves parameter estimation and conducting diagnostic checks on the model.

a. Model ARIMA(1,1,0)

```
> fit1 <- arima(data_migas.ts, orde = c(1,1,0), method = "ML")
> fit1

Call:
arima(x = data_migas.ts, order = c(1, 1, 0), method = "ML")

Coefficients:
```

```
          ar1
      -0.3680
s.e.   0.0773

sigma^2 estimated as 71231:  log likelihood = -1022.92,  aic = 2049.83

> coeftest(fit1)

z test of coefficients:

     Estimate Std. Error z value  Pr(>|z|)
ar1 -0.367964   0.077285 -4.7611 1.925e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

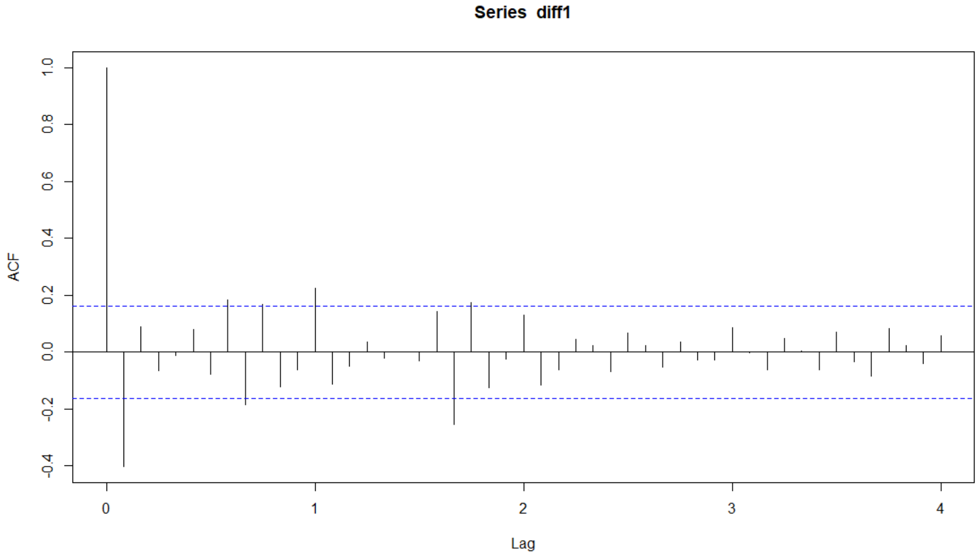

Figure 3. ACF plot of diff1

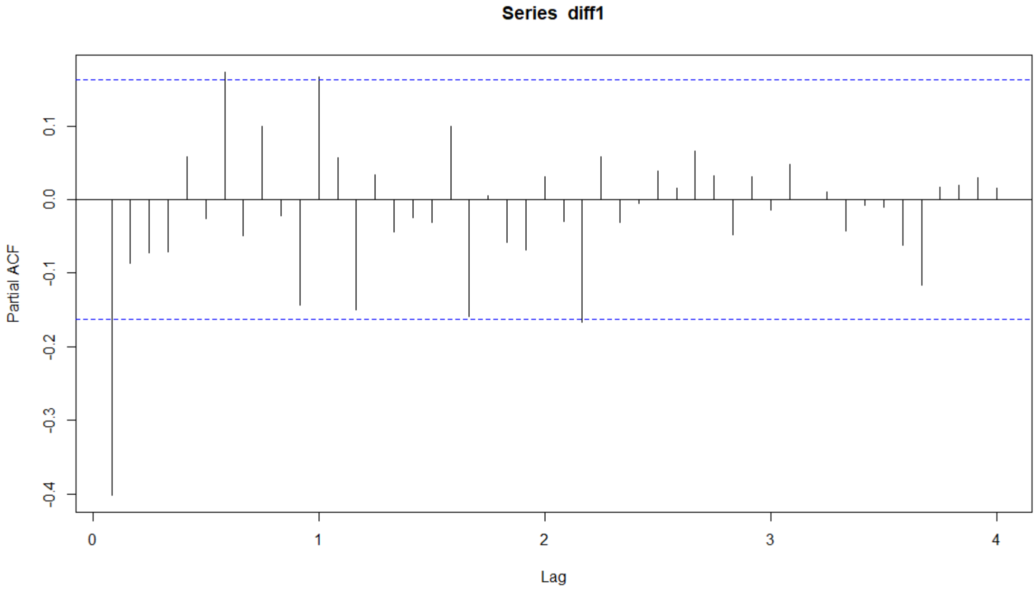

Figure 4. PACF plot of diff1

The ARIMA(1,1,0) analysis results reveal that the p-value or Pr(>|z|) of the AR(1) parameter is less than 0.05, which indicates that the parameter is statistically significant, with a value of -0.367965. Additionally, the residuals of this model were examined using the Ljung–Box test to determine whether they fulfilled the white noise requirement.

```
> Box.test(fit1$residuals, type = "Ljung")

        Box-Ljung test

data:  fit1$residuals
X-squared = 0.27113, df = 1, p-value = 0.6026
```

According to the results of the Ljung-Box test, the p-value is 0.6026, which is larger than the significance level of α = 0.05. Therefore, we was not reject H0, indicating that the residuals were white noise.

The following test assesses whether the residuals exhibit a normal distribution by employing the Kolmogorov-Smirnov test.

```
> lillie.test(fit3$residuals)

       Lilliefors (Kolmogorov-Smirnov) normality test

data: fit3$residuals
D = 0.054951, p-value = 0.3412
```

According to the Kolmogorov-Smirnov test, the p-value of 0.3412 is greater than the significance level α = 0.05, indicating that the null hypothesis H0 cannot be rejected. This suggests that the residuals followed a normal distribution.

b. Model ARIMA(0,1,1)

```
> fit2 <- arima(data_migas.ts, orde = c(0,1,1), method = "ML")
> fit2

Call:
arima(x = data_migas.ts, order = c(0, 1, 1), method = "ML")

Coefficients:
          ma1
      -0.4064
s.e.   0.0735

sigma^2 estimated as 70157:  log likelihood = -1021.83,  aic = 2047.65

> coeftest(fit2)

z test of coefficients:

     Estimate Std. Error z value  Pr(>|z|)
ma1 -0.406430   0.073525 -5.5278 3.243e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The ARIMA(0,1,1) analysis results revealed that the p-value or Pr(>|z|) for the MA(1) parameter was less than the significance level of 0.05, indicating that the parameter had a significant value of -0.406430. Additionally, the Ljung-Box test is employed to examine whether the residuals of this model adhere to the white noise criterion.

```
> Box.test(fit2$residuals, type = "Ljung")

        Box-Ljung test

data:  fit2$residuals
X-squared = 0.039669, df = 1, p-value = 0.8421
```

The Ljung-Box test results indicate that the p-value is 0.8421, which is greater than the significance level of α = 0.05.

Therefore, the null hypothesis H0 cannot be rejected, suggesting that the residuals are white noise.

The next step is to determine whether the residuals from the model follow a normal distribution using the Kolmogorov-Smirnov test.

```
> lillie.test(fit2$residuals)

        Lilliefors (Kolmogorov-Smirnov) normality test

data:  fit2$residuals
D = 0.067499, p-value = 0.09808
```

The Kolmogorov-Smirnov test yielded a p-value of 0.09808, which was larger than the significance level of α = 0.05. Consequently, we do not have sufficient evidence to reject the null hypothesis, which posits that the residuals are normally distributed.

c. Model ARIMA(1,1,1)

```
> fit3 <- arima(data_migas.ts, orde = c(1,1,1), method = "ML")
> fit3

Call:
arima(x = data_migas.ts, order = c(1, 1, 1), method = "ML")

Coefficients:
          ar1      ma1
      -0.0695  -0.3487
s.e.   0.2045   0.1924

sigma^2 estimated as 70100:  log likelihood = -1021.77,  aic = 2049.53

> coeftest(fit3)

z test of coefficients:

     Estimate Std. Error z value Pr(>|z|)  
ar1 -0.069495   0.204514 -0.3398  0.73400  
ma1 -0.348655   0.192395 -1.8122  0.06996 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

The ARIMA(1,1,1) analysis results showed that the p-value or Pr(>|z|) for the AR(1) and MA(1) parameters was greater than α = 0.05, indicating that the parameters were not statistically significant.

Table 2 presents a comprehensive overview of the outcomes of the parameter estimation and model diagnostic assessments for every conceivable model.

Table 2. Summary of parameter estimation results and model diagnostic checks

Model	Assumptions Test			AIC
	Parameter Significance	White Noise (Residual)	Normally Distributed Residuals	
ARIMA(1,1,0)	Significantly	Significantly	Significantly	2049.83
ARIMA(0,1,1)	Significantly	Significantly	Significantly	2047.65
ARIMA(1,1,1)	Not significantly	-	-	-

To evaluate the models presented in Table 2, it is evident that only two meet the ARIMA (1,1,0) and ARIMA (0,1,1) criteria. To determine the best model among these two, the AIC value was compared with the model with the smallest AIC being considered the best. As shown in Table 2, the most suitable and superior model is ARIMA (0,1,1).

The final step in the process is forecasting, which involves utilizing data from five subsequent periods to generate predictions. The forecast results are then presented in an easily understandable format.

```
> ramal <- forecast(data_migas.ts, model = fit1, h=5)
> ramal
```

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Apr 2022	1254.124	912.0887	1596.160	731.0259	1777.223
May 2022	1309.678	905.0525	1714.303	690.8568	1928.499
Jun 2022	1289.236	806.9266	1771.546	551.6074	2026.865
Jul 2022	1296.758	755.5855	1837.931	469.1060	2124.410
Aug 2022	1293.990	697.1399	1890.841	381.1864	2206.794

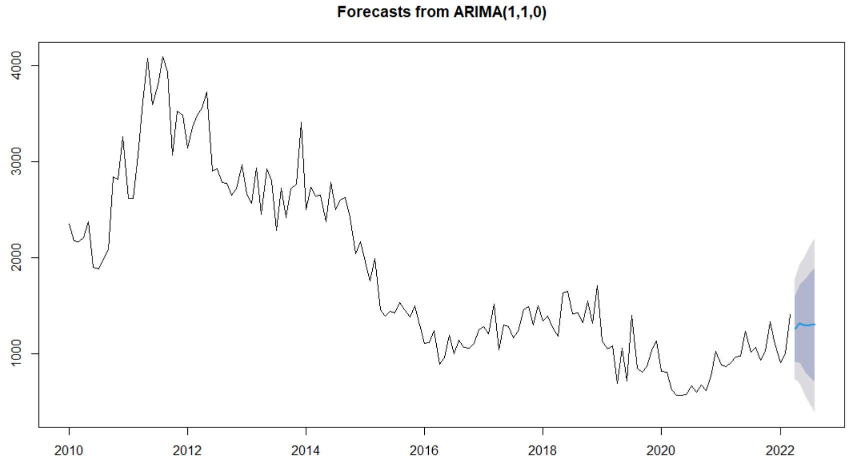

Figure 5. Forecasting results for the next 5 periods.

4. Conclusions

Based on the analysis and discussion, it can be concluded that there are three ARIMA models that have been formed, namely ARIMA (1,1,0), ARIMA (0,1,1), and ARIMA (1,1,1). Among these three models, the ARIMA (0,1,1) model is the most suitable for forecasting the value of oil and gas exports in Indonesia, as indicated by its AIC value of 2047.65. The forecasted values for the next five periods were 1254.124 (April 2022), 1309.678 (May 2022), 1289.236 (June 2022), 1296.758 (July 2022), and 1293.990 (August 2022).

References

Ahmar, A. S., Singh, P. K., Ruliana, R., Pandey, A. K., & Gupta, S. (2023). Comparison of ARIMA, SutteARIMA, and Holt-Winters, and NNAR Models to Predict Food Grain in India. *Forecasting*, *5*(1), 138-152.

Ahmar, A. S., Boj del Val, E., El Safty, M. A., AlZahrani, S., & El-Khawaga, H. (2022). SutteARIMA: A novel method for forecasting the infant mortality rate in Indonesia. *CMC-Computers Materials & Continua*, 70(3), 6007-6022.

Anityaloka, R. N., & Ambarwati, A. N. (2013). Peramalan Saham Jakarta Islamic Index Menggunakan Metode ARIMA Bulan Mei-Juli 2010. *Jurnal Statistika Universitas Muhammadiyah Semarang*, 1(1), 1–5.

Farina, F., & Husaini, A. (2017). Pengaruh Dampak Perkembangan Tingkat Ekspor dan Impor terhadap Nilai Tukar Negara ASEAN Per Dollar Amerika Serikat (Studi pada International Trade Center Periode Tahun 2013-2015). *Jurnal Administrasi Bisnis S1 Universitas Brawijaya*, 50(6), 44-50.

Hayati, F. N., Nurlaily, D., & Pusporani, E. (2021). Peramalan Data Ekspor Non Migas Provinsi Kalimantan Timur Menggunakan Univariate Time Series. *J Statistika: Jurnal Ilmiah Teori dan Aplikasi Statistika*, *14*(2), 59-66.

Kertayuga, D., Santoso, E., & Hidayat, N. (2021). Prediksi Nilai Ekspor Impor Migas Dan Non-Migas Indonesia Menggunakan Extreme Learning Machine (ELM). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, *5*(6), 2792–2800.

Singh, P. K., Chouhan, A., Bhatt, R. K., Kiran, R., & Ahmar, A. S. (2021). Implementation of the SutteARIMA method to predict short-term cases of stock market and COVID-19 pandemic in USA. *Quality & Quantity*, 1-11.

Sihombing, M, Sihotang, J, and Purba M. L. (2021). Analisis Pengaruh Ekspor Migas, Ekspor Non Migas dan Penanaman Modal Asing Terhadap Pertumbuhan Ekonomi Indonesia Tahun 2000-2019, *Journal of Economics and Business*, 2(2), 40–51

Utami, T.W dan Darsyah, M. (2015). Peramalan Data Saham Dengan Model Winter's. *Jurnal Statistika*, 3(2), 2–5.

Yulisa, P.N., Haris, M.A., dan Arum, P.R. 2023. Peramalan Nilai Ekspor Migas di Indonesia dengan Model Long Short Term Memory (LSTM) dan Gated recurrent Unit (GRU). *Jurnal Statistika: 16(1),* 324-341