Forecasting the Value of Oil and Gas Exports in Indonesia using ARIMA Box-Jenkins

Ansari Saleh Ahmar[a,*], Miguel Botto-Tobar[b,c], Abdul Rahman[d], & Rahmat Hidayat[e]

[a]Department of Statistics, Universitas Negeri Makassar, Makassar, 90223, Indonesia
[b]Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands
[c]Research Group in Artificial Intelligence and Information Technology, University of Guayaquil, 090510, Guayaquil, Ecuador
[d]Department of Mathematics, Universitas Negeri Makassar, Makassar, 90223, Indonesia
[e]Department of Information Technology, Politeknik Negeri Padang, Limau Manis, Padang, 25164, Indonesia

Abstract

The objective of the study was to forecast the value of oil and gas exports in Indonesia using the ARIMA Box-Jenkins. With this prediction, it is hoped that it can be a study for future policy making. This oil and gas export data is obtained from the Indonesian Central Bureau of Statistics (BPS) website, in raw data from January 2010 to March 2022. This data is predicted using the ARIMA method with the help of R software. The stages of data analysis with ARIMA include: data stationary test, build the model indication, parameter estimation and significance test, and residual diagnostic test of the model. The results of data analysis conducted in this study show that there are 3 indications of models that were generated, namely ARIMA(1,1,0); ARIMA(0,1,1); and ARIMA(1,1,0). From these 3 model indications, the best model was ARIMA(0,1,1) with AIC value of 2047.65.

Keywords: ARIMA; forecasting; oil and gas export.

1. Introduction

Indonesia is a country rich in natural resources. These natural resources are essential for human life. Indonesia's natural resources are diverse, both renewable natural resources and non-renewable natural resources. One of the natural resources in Indonesia is oil and gas. This makes Indonesia one of the largest oil and gas producers in the world. The Ministry of Energy and Mineral Resources explains that Indonesia is number 5 in the world in the portion of oil and gas revenue sharing for the state (ESDM, 2013).

Furthermore, ESDM (2013) said that in 2012, Indonesia's oil reserves amounted to 3.6 billion barrels or around 0.2% of the total oil reserves in the world, while Indonesia's gas reserves amounted to 104.25 trillion cubic feet or around 1.7% of the world's total gas reserves.

In 2022, based on BP Statistical Review 2022, Indonesia ranks as the 24th largest oil producing country in the world with a total of 692 thousand barrels per day or contributing to 0.8% of world oil production (Andrianto, 2022).

Oil and natural gas (MIGAS) is the most important thing in life in Indonesia because it is very useful for human livelihoods in living their daily lives. In addition to being used as fuel oil for transportation (motorcycles, cars, airplanes, and vehicles that use fuel), oil and gas are also used for household purposes, such as cooking oil and LPG.

In addition, MIGAS is not only used for daily life but can also be used as foreign exchange for the continuation of the country's development. This makes MIGAS one of the sectors of the Indonesian economy because it makes a very large contribution to state revenue.

This is supported by Salsabila (2021) who said that oil and gas exports have a positive effect on Economic Growth. To see more about Indonesia's oil and gas exports in the future, one way that can be done is by forecasting or predicting the value of oil and gas exports.

---

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

2. Methods

The data used in this study are data on the export value of oil and gas in Indonesia every month from January 2010 to March 2022 (in Million US$). This data consists of 147 data. The research data comes from secondary data, namely data from the BPS website.

This data is forecasted using the ARIMA Box-Jenkins approach with R software as the analysis tool. The step of forecasting using ARIMA are (Kasyok, 2013; Ahmar et. al., 2018):

- **Step 1:** Plot the time series and check the stationarity of the data. A time series data must initially fulfill stationary and there is no trend before forecasting using ARIMA.
- **Step 2:** Transformation and/or Difference. If in step 1 non-stationary data is obtained, then (a) data transformation is carried out to overcome data non-stationarity in variance and (b) difference to overcome data non-stationarity in mean.
- **Step 3:** Build the model indication. After step 2 has been completed, model formation is carried out using the ACF and PACF plots of the data.
- **Step 4:** Model validation. To see the goodness of the model from step 3, model validation process is carried out which includes: (a) parameter estimation and significance test; and (b) residual diagnostic test of the model. If there is more than one model that eligible with this validation, then determining the best model can be seen from the smallest value of AIC or MSE or RMSE or MAPE.

3. Result and Discussion

3.1. Data Stationarity Test

The first stage of ARIMA forecasting is to check the stationarity of the data. The way to check data stationarity is to plot the time series data, check the Augmented Dickey-Fuller (ADF) value, Box-Cox value, and plot the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the data.
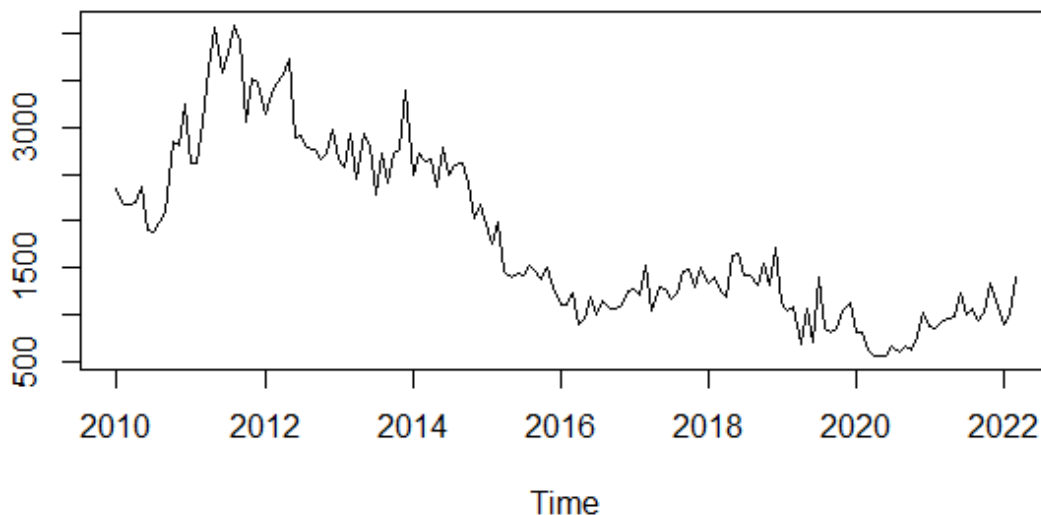

Figure 1. Time series plot of export value of oil and gas in Indonesia

```
Augmented Dickey-Fuller Test
data: migas
Dickey-Fuller = -2.548, Lag order = 5, p-value = 0.3482
alternative hypothesis: stationary

BoxCox.lambda(migas)
[1] 0.5638694
```

Based on the ADF Test results, it can be seen that the p-value obtained is 0.3482. This value is greater than the alpha of 5% which means that accepting H_0 or in other words, the data is not stationary (Mushtaq, 2011). Next, the process of checking the stationarity of the data variance is carried out whether it is stationary in variance or not by looking at the lambda value in the Box-Cox Test. From the Box-Cox test results, it is obtained that from the above results a value of 0.5638694 is obtained which explains that the data is not yet stationary in variance.

To be more confident about the stationarity of the data, the ACF and PACF plots of the data are presented which can be seen in Figure 1. From the ACF plot, it can be seen that the value drops slowly. This indicates that the data is not stationary. Based on the results of the analysis, it can be said that the data is not stationary in mean and variance.

To overcome the problem of data non-stationarity in variance, one way is to perform lambda transformation. Based on the Box-Cox value, the lambda value is 0.5638694 so that the suitable transformation is the root of the data ($T(X_t)$ = $\sqrt{X_t}$) (Osborne, 2010).
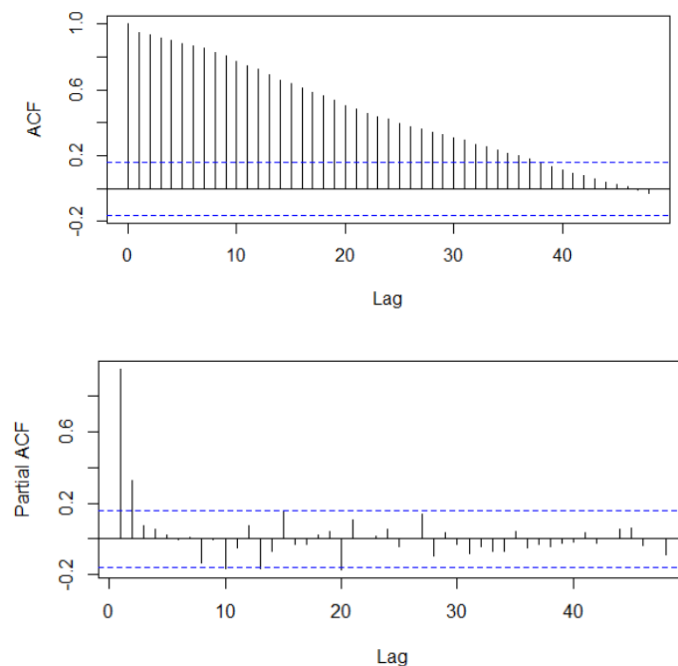
Figure 2. ACF and PACF plots

Box-Cox test results using lambda transformation are obtained:

```
> BoxCox.lambda(translambda)
[1] 1.16589
```

Based on the results of this lambda, which is 1.16589 and close to the value of 1, it can be said that the data has been stationary in variance. The next step is to perform data differencing to overcome the problem of non-stationarity in the mean.

```
> diff1 <- diff(translambda, differences = 1)
> ts.plot(diff1)
```

From Figure 3, it can be seen that the data is stationary in mean. And to be more convincing, the ADF test and ACF and PACF plots were conducted.

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

Based on the ADF results, it is found that the p-value is below alpha 0.5 which indicates that the data is stationary in mean.
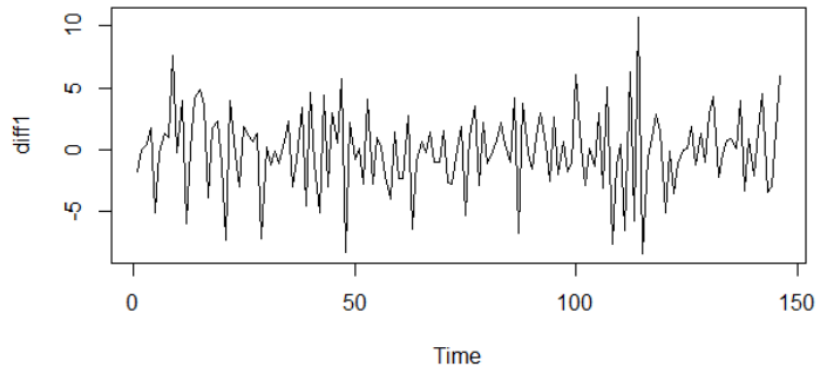

Figure 3. Time Series Plot of 1st Differencing Data (diff1)

3.2. ARIMA Model

To see the ARIMA model formed from oil and gas export data in Indonesia, the ACF and PACF plots of 1st differencing (diff1) data can be seen in Figure 4. Based on Figure 4, the ACF plot of the results of 1st differencing (diff1) shows that the autocorrelation value is cut after lag 1, and the PACF plot dies down and cuts after lag 1. Based on these results, three possible ARIMA models are ARIMA(1,1,0); ARIMA(0,1,1); and ARIMA(1,1,1).

3.2.1. ARIMA(1,1,0)

a. Parameter Estimation

Parameter estimation is done using the maximum likelihood method.

```
> fit1 <- arima(translambda, orde = c(1,1,0), method = "ML")
> fit1
Call:
arima(x = translambda, 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
```

b. Diagnostic Checks

Diagnostic checks include parameter significance tests, white-noise assumption tests and residual normality tests.

```
> 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
```

From the parameter significance test results, it can be seen that the ar1 coefficient value has a p-value that is smaller than $\alpha = 5\%$ or in other words, the ar1 parameter is significant.

To see the residuals of a model whether it has met the white noise assumption or not is done with the Ljung-Box test.

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

        Box-Ljung test
```

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

From the Ljung-Box test results, it can be seen that the p-value of 0.6026 is greater than $\alpha = 5\%$, which means that the residuals are white noise.
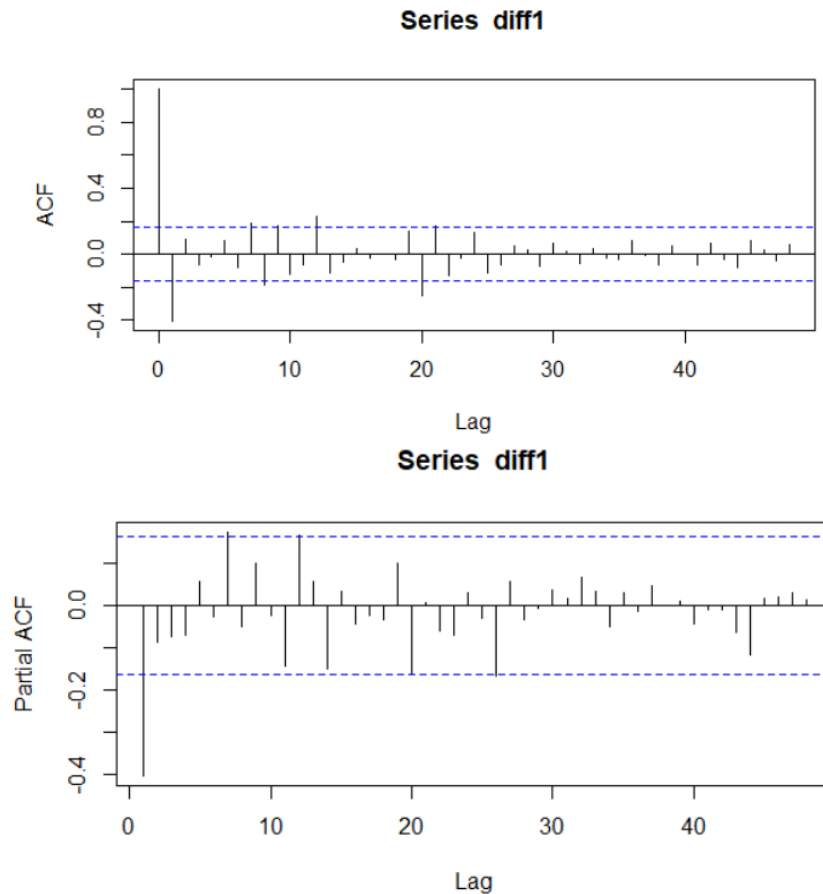
Series diff1

Series diff1

Figure 4. ACF and PACF plots of 1st Differencing (diff1)

The last stage of the diagnostic check is the normality test of the residual data, whether normally distributed or not using the Kolmogorov-Smirnov normality test.

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

        Lilliefors (Kolmogorov-Smirnov) normality test

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

The Kolmogorov-Smirnov test results show that the p-value of 0.3412 is greater than $\alpha = 5\%$, which means that the residuals of the model are normally distributed.

Since all assumption tests have been met, the ARIMA(1,1,0) model is eligible for forecasting.

3.2.2. ARIMA(0,1,1)

a.   Parameter Estimation

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

```
Call:
arima(x = translambda, 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
```

b. *Diagnostic Checks*

```
> 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
```

From the parameter significance test results, it can be seen that the ma1 coefficient value has a p-value that is smaller than $\alpha = 5\%$ or in other words, the ar1 parameter is significant.

To see the residuals of a model whether it has met the white noise assumption or not is done with the Ljung-Box test.

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

        Box-Ljung test

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

From the Ljung-Box test results, it can be seen that the p-value of 0.8421 is greater than $\alpha = 5\%$, which means that the residuals are white noise.

The last stage of the diagnostic check is the normality test of the residual data, whether normally distributed or not using the Kolmogorov-Smirnov normality test.

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

        Lilliefors (Kolmogorov-Smirnov) normality test

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

The Kolmogorov-Smirnov test results show that the p-value of 0.09808 is greater than $\alpha = 5\%$, which means that the residuals of the model are normally distributed.

Since all assumption tests have been met, the ARIMA(0,1,1) model is eligible for forecasting.

3.2.3. ARIMA(1,1,1)

a. *Parameter Estimation*

```
> fit3 <- arima(translambda, orde = c(1,1,1), method = "ML")
> fit3
Call:
arima(x = translambda, 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
```

b. Diagnostic Checks

```
> coeftest(fit1)
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
```

From the parameter significance test results, it can be seen that the ar1 and ma1 coefficient values have a p-value greater than α = 5% so that the parameters are not significant and it can be said that the ARIMA(1,1,1) model does not eligible the assumptions so it is not suitable for use in forecasting.

In general, the results of the Box-Jenkins assumption testing are presented in table 1.

Table 1. Box-Jenkins Assumption Testing Results

Model	Significance of Parameters	Residual of White Noise	Normally Distributed	AIC
ARIMA(1,1,0)	Eligible	Eligible	Eligible	2049.53
ARIMA(0,1,1)	Eligible	Eligible	Eligible	2047.65
ARIMA(1,1,1)	Not Eligible	-	-	-

Table 2. Prediction Results for the next 6 months (April - September 2022) with ARIMA(0,1,1)

Month	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
April 2022	1238.719	899.271	1578.166	719.579	1757.859
May 2022	1238.719	843.977	163.460	635.014	1842.424
June 2022	1238.719	795.529	1681.909	560.918	1916.519
July 2022	1238.719	751.878	1725.559	494.160	1983.277
August 2022	1238.719	711.831	1765.606	432.914	2044.524
September 2022	1238.719	674.621	1802.817	376.005	2101.433

Since there are 2 ARIMA models that are eligible in forecasting oil and gas export data, the best model will be selected. The selection of the best model is based on the smallest AIC value of the model. Based on Table 1, the best ARIMA model in forecasting oil and gas export data in Indonesia is the ARIMA(0,1,1) model.

The ARIMA(0,1,1) model will be used to predict oil and gas export data in Indonesia for the next 6 months (April 2022 - September 2022). The prediction results are presented in Table 2 and the forecasting graph is presented in Figure 5.

4. Conclusion

Based on the results of data analysis and discussion that has been carried out, it can be concluded that the best and appropriate model in forecasting oil and gas export data in Indonesia is the ARIMA(0,1,1) model with an AIC of 2047.65. From the forecasting results, it is predicted that the value of Indonesia's oil and gas exports is stable until September 2022 and there may be an increase and or decrease depending on the oil and gas export policy that will occur in the future, this is also supported by the existing forecasting range both in the 80% and 95% forecasting ranges.
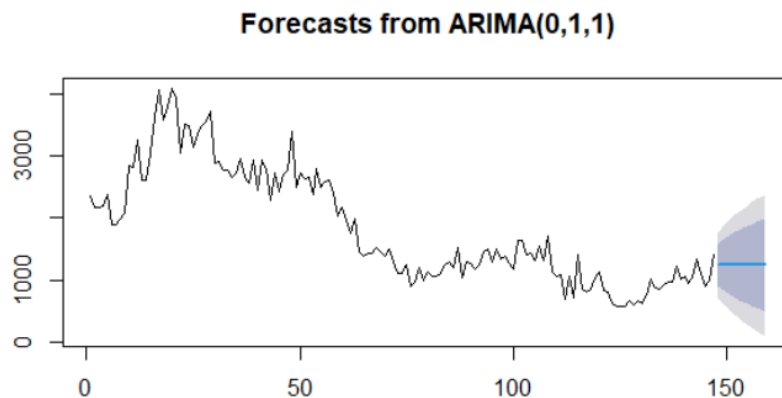
Forecasts from ARIMA(0,1,1)

Figure 5. Prediction Results for the next 6 months (April - September 2022) with ARIMA(0,1,1)

References

Ahmar, A. S., Guritno, S., Rahman, A., Minggi, I., Tiro, M. A., Aidid, M. K., ... & Arifin, A. N. M. (2018). Modeling data containing outliers using ARIMA additive outlier (ARIMA-AO). In *Journal of Physics: Conference Series* (Vol. 954, No. 1, p. 012010). IOP Publishing. https://doi.org/10.1088/1742-6596/954/1/012010

Andrianto, R. (2022). 10 Negara Penghasil Minyak Bumi Terbesar, Ada Indonesia? [The 10 biggest petroleum producing countries, is there Indonesia?]. Retrieved from https://www.cnbcindonesia.com/market/20220906162442-17-369783/10-negara-penghasil-minyak-bumi-terbesar-ada-indonesia

Badan Pusat Statistik. (2022). Nilai Ekspor Migas-NonMigas (Juta US$), 2022. Retrieved from https://bps.go.id/indicator/8/1753/1/nilai-ekspor-migas-nonmigas.html

ESDM. (2013). Indonesia Nomor 5 Dunia Dalam Porsi Bagi Hasil Migas Untuk Negara [Indonesia is the 5th in the world in the allocation of oil and gas revenue to the state]. Retrieved from https://www.esdm.go.id/id/media-center/arsip-berita/indonesia-nomor-5-dunia-dalam-porsi-bagi-hasil-migas-untuk-negara

Kasyok, A. (2013). Simple steps for fitting ARIMA model to time series data for forecasting using R. *International Journal of Science and Research (IJSR), 4(3), 318-321.*

Mushtaq, R. (2011). Augmented Dickey Fuller Test. Available at SSRN: https://ssrn.com/abstract=1911068 or http://dx.doi.org/10.2139/ssrn.1911068

Osborne, J. (2010). Improving your data transformations: Applying the Box-Cox transformation. *Practical Assessment, Research, and Evaluation*, *15*(12), *1-8.*

Salsabila, D. (2021). Analisis Pengaruh Ekspor Migas dan Non Migas terhadap Pertumbuhan Ekonomi Indonesia. *Jurnal Akuntansi Dan Manajemen*, *18*(01), *01-08.* https://doi.org/10.36406/jam.v18i01.374