Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)
Vol. 9, No. 3, September 2023, pp. 720-730
ISSN: 2338-3070, DOI: 10.26555/jiteki.v9i3.22286

Autoregressive Integrated Moving Average (ARIMA) Model for Forecasting Indonesian Crude Oil Price

Masna Wati, Haviluddin, Akhmad Masyudi, Anindita Septiarini, Heliza Rahmania Hatta
Informatika, Universitas Mulawarman, Jl. Sambaliung Kampus Gg. Kelua, Samarinda 75119, Indonesia

ARTICLE INFO

Article history:

Received November 23, 2021
Revised July 29, 2023
Published August 07, 2023

Keywords:

ARIMA;
Crude oil;
Forecasting;
ICP

ABSTRACT

Crude oil is the main commodity of the global economy because oil is used as an ingredient for many industries globally and is the price base used in the state budget. Indonesian Crude Price (ICP) fluctuates following developments in world crude oil prices. A significant increase in crude oil prices will certainly disrupt the economy. Thus, the movement or fluctuation of ICP is essential for business players in the energy market, especially domestically. Therefore, crude oil price forecasting is needed to assist business people in making decisions related to the energy market. This study aims to find a suitable forecasting model for Indonesian crude oil prices using the Autoregressive Integrated Moving Average (ARIMA) method. The forecasting process used ICP time-series data per month for 50 types of crude oil within five years or 63 months. Based on the experimental results, it was found that the most fit ARIMA models were (0,1,1), (1,1,0), (0,1,0), and (1,2,1). The test results for April to September 2020 have a good and proper interpretation, except the type of BRC oil indicates inaccurate forecasts. The ARIMA error rate is very dependent on the value of the data before it is predicted and external factors, the more unstable the data value every month, the higher the error rate.

Corresponding Author:

Masna Wati, Universitas Mulawarman, Jl. Sambaliung Kampus Gg. Kelua, Samarinda 75119, Indonesia
Email: masnawati@fkti.unmul.ac..id

1. INTRODUCTION

One of the most active trades in the world is the crude oil trade. Oil has a strategic role for both oil-exporting and oil-importing countries because of the large volume of trade [1], [2]. Crude oil is the primary commodity of the global economy [3]. In fact, crude oil is an essential component of economic development and growth in industrialized and developing countries [4]. Crude oil is one source of energy that is very important and needed for human life. Oil is used as a material for various large and small industries, causing high demand for oil from various countries worldwide [5], [6].

The Indonesian Crude Price (ICP) fluctuates following developments in world crude oil prices. A significant rise in the price of crude oil is certainly not expected by many governments worldwide, mainly the crude oil-producing countries, including Indonesia [1], [3], [4]. For countries that act as exporters, an increase in oil prices will trigger a decline in demand in the future. Meanwhile, for importing countries, this increase will disrupt their economic growth due to high inflation increases [7].

The ICP is the basis for crude oil prices used in the State Budget. Thus, the movement or fluctuation of ICP is essential for business people in the energy markets, especially in our country [8]. Therefore, crude oil price forecasting is needed to assist business people in making decisions related to the energy market and further reduce the impact of price fluctuations. Forecasting crude oil prices can utilize time-series data, which is forecasting with observations of an ordered set of data on Indonesian crude oil prices in a certain past period. One of the time series methods is the Autoregressive Integrated Moving Average (ARIMA) method or the Box-Jenkins method.

The ARIMA method is appropriate to overcome the complexity of time series and other forecasting situations. The ARIMA method could be used to estimate historical data with conditions that are difficult to understand, and its impact on the data is technically very accurate for short-term forecasts [9]–[11]. In addition, Gaetano Perone's research on the ARIMA model to estimate the spread and final size of the COVID-19 epidemic in Italy shows that the ARIMA model could be seen as a fast and straightforward tool for programming health monitoring systems at the national and local levels. Interestingly, the ARIMA forecasting approach is simple to utilize and interpreted [12]. In contrast, it is sensitive to outliers in the data and does not account for noise. For this reason, this model is considered a good model for short-term forecasting, but the results should be interpreted prudently [13].

Hence, it would be interesting to be done a study to describe fluctuations in Indonesian oil prices and forecast oil prices for some time to come using the ARIMA method. It could be expected that this research has contributed to making decisions in problems related to the price of Indonesian crude oil.

2. METHOD

The research implementation stage is carried out in this study is described by a diagram. That is, it explains all processes executed to predict the ICP using the ARIMA method. The following is the research implementation stage in Fig. 1.

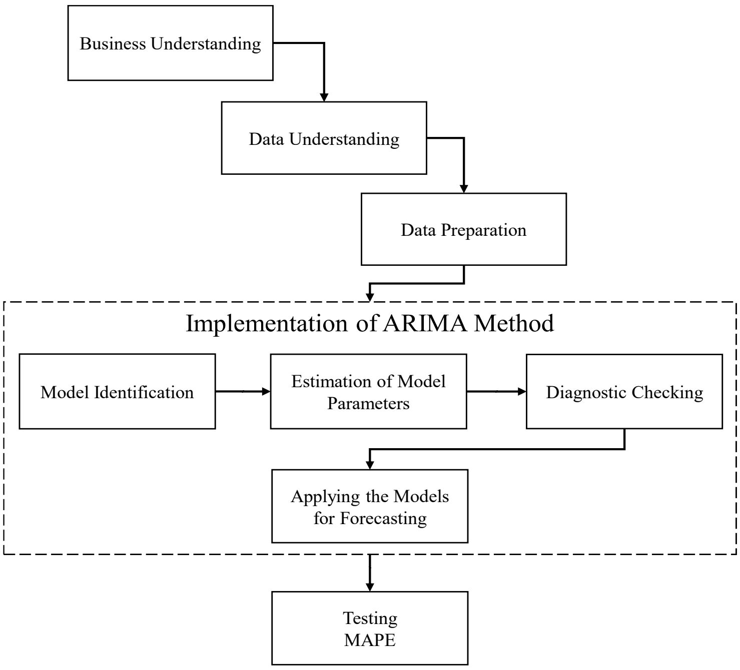


Fig. 1. Research Stages of ICP Forecasting using ARIMA Method

The phase of forecasting Indonesian crude oil prices in Fig. 1 are as follows:

a. Business understanding: The problem identification step is executed to find the focus of the problem to be analyzed. This study aims to describe fluctuations in Indonesian oil prices and forecasts for the future period. Forecasting results can be taken into consideration in making decisions related to Indonesian oil prices.

b. Data understanding: This stage is the process of understanding and collecting data needed in forecasting Indonesian crude oil prices.

c. Data preparation: The dataset used for forecasting crude oil prices is time-series data on Indonesian crude oil prices per month for 50 types of crude oil over five years or 63 months, namely January 2015 to March 2020.

d. Application of the ARIMA Method: The method will be applied the ARIMA method since it is fast, simple, and cheap. Further reasonably, it only requires variable data to be predicted. In addition, the ARIMA method is known to have good accuracy for short-term forecasting.

e. Testing: The model performance test is accomplished to know the accuracy rate of the forecasting results and determine the best ARIMA model for forecasting the object of this research. The test will use the Mean Absolute Percentage Error (MAPE) parameter.

2.1. ARIMA Method

The time series is a series of data in observation values measured over a certain period, based on time with equal intervals [9], [14], [15]. Time series analysis is basically used to perform data analysis that considers the effect of time. Data is collected periodically based on time in units of hours, days, weeks, months, quarters, or years. Time series analysis can be done to assist in planning for the future [16].

There are two components to the time series forecasting technique. First, forecasting models based on statistical and mathematical approaches: moving average (MA), exponential smoothing (ES), regression, and ARIMA [17]–[20]. Second, artificial intelligence-based forecasting techniques: neural networks (NN), genetic algorithms (AG), simulated annealing (SA), genetic programming (GP), classification, and hybrids [21], [22].

The ARIMA method, commonly referred to as the Box-Jenkins method, is intensively elaborated by George Box and Gwilym Jenkins (1976). It is a new development in economic forecasting methods, does not aim to form a structural model (single equation or simultaneous equation) that based on economic theory and logic but to analyze the probabilistic or stochastic analysis of time series data by holding the philosophy of "let the data speak for themselves" [17], [23]. The type of linear time series models included in this method includes Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), and ARIMA [24], [25].

Successively implementing the ARIMA method are specification or model identification, estimating model parameters, diagnostic checking, and forecasting. The following is a description of the steps of the ARIMA method [11], [25]–[29].

1. Model Identification

   At this step, model identification is made by:

   a. Data Description: plots a graph to see data values as the identification process to know whether the data shape is seasonal or non-seasonal.
   b. Augmented Dickey-Fuller (ADF) test: determine the *p-value* to determine the data stationarity because the data used in the ARIMA method must be stationary.
   c. Differencing process: carried out if the data is not stationary. After the differencing is done, the ADF value will be tested back. The differencing process will continue until the data are stationary if they are still not [30].
   d. Identification of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) graphs: plotting ACF and PACF graphs to assist the process of estimating p and q parameter values in the ARIMA model (p, d, q). Determination of p and q values by observing the pattern of the ACF and PACF functions of the series is analyzed, based on Table 1 [31], [32].

Table 1. Formula of Autocorrelation and Partial Autocorrelation

Autocorrelation (ACF)	Partial autocorrelation (PACF)	ARIMA tentatif
Going to zero after $lag\ q$	Gradually descending/fluctuating.	ARIMA $(0, d, q)$
Gradually descending/fluctuating	Going to zero after $lag\ q$	ARIMA $(p, d, 0)$
Gradually descending/fluctuating (as long as $lag\ q$ not zero)	Gradually descending/fluctuating (as long as $lag\ q$ not zero)	ARIMA (p, d, q)

2. Estimation of Model Parameters

   At this stage, selecting the best model estimate to be used is determined by searching for the smallest Akaike Information Criterion (AIC) value. The AIC is a criterion for selecting the best model proposed by Akaike in 1973 based on considering the many parameters in the model [33]. The AIC value is expressed in (1) [26], [27], [34].

$$AIC\ (k) = N\ ln(\hat{\sigma}_a^2) + 2k \quad (1)$$

where k is number of parameters in the model, N is number of data, and $\hat{\sigma}_a^2$ is maximum likelihood estimation of σ_a^2.

3. Diagnostic Checking

   While obtaining the ARIMA estimator, a diagnostic test is carried out to check whether the suspected model meets the requirements based on the residual graph, KDE histogram, Q-Q, and correlogram.

4. Applying the Models for Forecasting

   While the appropriate model is obtained, the next step is forecasting the future period on training data and forecasting data using the model.

2.2. Measurement of Accuracy Rate

The MAPE parameter indicates how much the forecast error is compared to the actual value [35]. A forecast has outstanding performance if a MAPE value is below 10% and good performance if the MAPE value ranges from 10% to 30%. The general form of this model is expressed in (2) [36]–[40].

$$MAPE = \frac{\sum_{t=1}^{n} \left| \left(\frac{A_t - F_t}{A_t} \right) 100 \right|}{n} \tag{2}$$

where n is the value of the period; A_t is the actual value in the t-th period; F_t is the forecast value in period-t. The interpretation of MAPE values is presented in Table 2 [41]–[43].

Table 2. Interpretation of MAPE Values

Range MAPE	Interpretation
< 10%	The highly accurate forecasting performance
10% - 20%	Good accurate forecasting performance
20% - 50%	Decent forecasting performance
> 50%	Inaccurate forecasting performance

3. RESULTS AND DISCUSSION

3.1. Data Processing

The data used in forecasting oil prices are in the form of ICP data for 50 types of crude oil produced in Indonesia from January 2015 to March 2020. It is consists of 66 data for each crude oil type, so the total data used is 3,300 crude oil price data. The ICP data is displayed in graphical form using the panda's library and Matplotlib to describe the time series data used in the forecast process. ICP data from one type of crude oil is presented in Fig. 2.

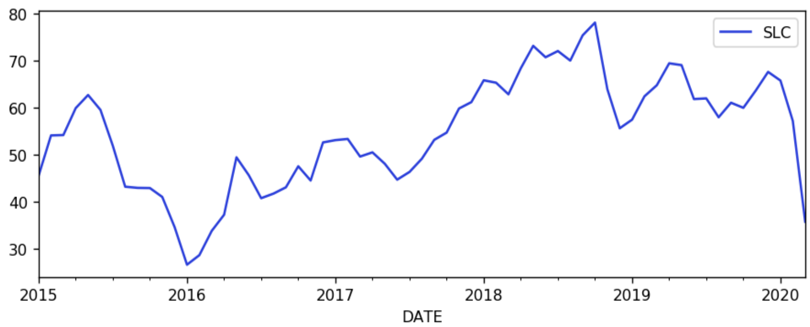


Fig. 2. Price of SLC oil (US Dollar/Barrel) in 2015-2020

Based on Fig. 2, at the beginning of 2015, SLC oil prices gradually rose and declined in the middle to the end of the year. In 2016, the SLC price rose initially and then fell in the middle of the year. However, the price rose again at the end of the year. Opposite, in 2017, it gradually decreased and then increased from the middle to the end of the year. In 2018, it tended to rise and then fell at the end of the year. In 2019, the price of SLC oil fluctuated throughout the year, and in early 2020 the price of oil decreased.

Based on the observation of oil prices for the last five years, SLC oil prices are not seasonal data. Fluctuations in other types of Indonesian crude oil experienced the same thing.

3.2. Implementation of ARIMA Method

The implementation of the ARIMA method in forecasting Indonesian oil prices consists of the ADF testing (stationarity test), differencing process, identification of ACF and PACF charts, estimation of ARIMA model parameters, diagnostic testing, distribution of training data, forecasting of training data, and the application of ARIMA to forecast data.

1. The ADF Testing

This stage is implemented using the Statsmodels library. The P-value from the ADF test results is used to see if the time series data is stationary or not. A differencing process is required if the data is not stationary (p-value is more than 0.05). The results of the ADF testing process are presented in Table 3.

Table 3. The Result of ADF ICP Testing

Types of Indonesian Crude Oil	P-Value
SLC	0.26119290
Arjuna	0.31691161
Attaka	0.32015443
Cinta	0.27446755
Duri	0.17991271
⋮	⋮
Walio Mix	0.32900943

2. Differentiating Process

The differencing process uses the help of Matplotlib's library. Based on the ADF test results' p-values in each differencing step, it is possible to determine whether the data is stationary. If the data is still not stationary, then the differencing process is needed again until it becomes stationary. Table 4 shows the result of differencing results.

Table 4. The Result of ADF Testing after Differencing

Types of Indonesian Crude Oil	P-Value Diff 1	P-Value Diff 2
SLC	0.00007612	-
Arjuna	0.00052180	-
⋮	⋮	⋮
Bontang Return Condensate (BRC)	0.00033203	0.00004055
Komplek Palembang Selatan (KPS)	0.00006035	0.00002330
⋮	⋮	⋮
Walio Mix	0.00033203	-

Based on the ADF test after the differencing process in Table 4, almost all types of crude oil only require a one-time differencing process. The type of crude oil that requires two times differencing for stationary is the type of Bontang Return Condensate (BRC) and the South Palembang Complex (KPS).

3. Identifying ACF and PACF Charts

This stage uses the help of the Matplotlib library. ACF and PACF graphs are formed based on the number of times the differencing process is performed. ACF and PACF graphs are used to determine a tentative model in estimating the ARIMA model parameters.

The ACF and PACF graphs of SLC oil in Fig. 3 show that the ICP data of SLC oil that underwent a differencing process resulted in a stationary data model on the average ICP of SLC oil. In addition to the 0th lag, which definitely crosses the significant limit, the first lag line crosses the significant limit (blue area) on both the ACF and PACF charts. ARIMA tentative models that can be proposed for SLC oil types are 1,1,1), (1,1,0), (0,1,1), or (0,1,0). Likewise, the tentative model for oil types other than BRC, Klamono, and KPS.

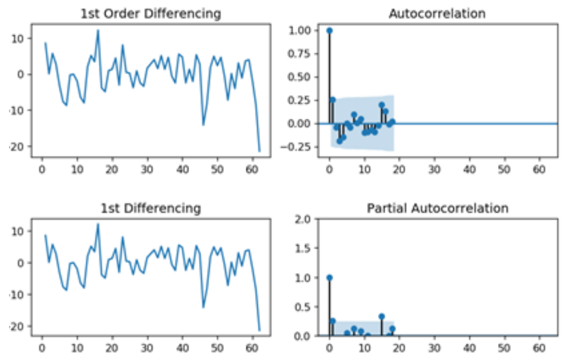


Fig. 3. The ACF and PACF of SLC oil

According to the ACF and PACF charts of BRC oil in Fig. 4, the ACF chart for BRC oil type 8th lag line crosses its significant limit, while the third lag line on the PACF chart crosses the significant limit so that the ARIMA tentative model proposed for BRC oil type is (3,2,1), (2,2,1), (3,2,0), (2,2,0), (1,2,1), (1,2,0), (0, 2.1), or (0.2.0). Fig. 5 shows the ACF and PACF charts for Klamono oil, where on the ACF and PACF charts, only the 0th lag crosses the significant limit, so the tentative model that can be proposed for the Klamono oil type is ARIMA(0,1,0). Based on Fig. 6, the first lag line on the ACF chart for KPS oil types crosses a significant limit, while the second lag line on the PACF chart crosses a significant limit so that the ARIMA tentative model that can be proposed for KPS oil types is (2,2,1), (2,2,0), (1,2,1), (1,2,0), (0,2,1), or (0,2,0).

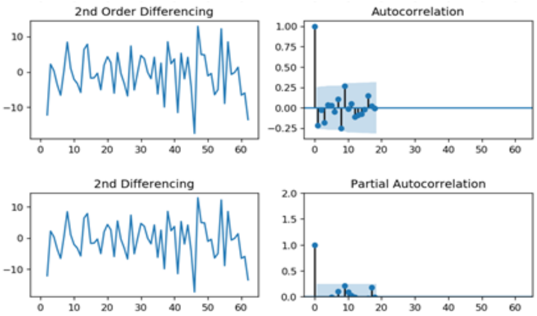


Fig. 4. The ACF and PACF of BRC oil

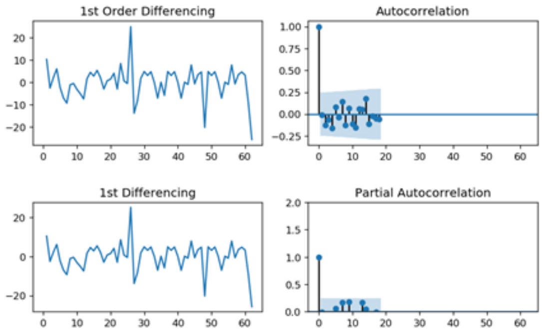


Fig. 5. The ACF and PACF of Klamono oil

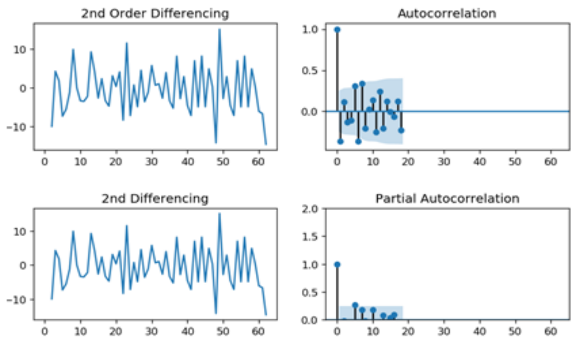


Fig. 6. The ACF and PACF of KPS oil

4. Parameter Estimation of ARIMA Model

At the parameter estimation stage of the ARIMA Model, a search (either stepwise or parallel) is carried out on the possible models within the provided constraints and selects the parameters that minimize the given metrics. The estimation of ARIMA model parameters based on the smallest AIC value in Table 5.

Table 5. The Result of ARIMA Parameter Estimation

Types of Indonesian Crude Oil	ARIMA (p, d, q)
SLC	(0,1,1)
Arjuna	(0,1,1)
Attaka	(0,1,1)
Cinta	(0,1,1)
Duri	(0,1,1)
⋮	⋮
Walio Mix	(0,1,1)

5. Diagnostic Testing

The diagnostic testing phase utilized the help of the Statsmodels plot_diagnostics library. The graphs formed are residual, correlograms, histograms, and normal Q-Q. The residual graph is checked based on unclear seasonality patterns and fluctuates around an average of zero.

The correlogram graph has a low correlation with the lag of the ACF graph marked by 95% lag whose value is more than one should not be significant. On the KDE histogram chart, the curve lines follow the normal distribution curve of the KDE histogram chart. On the normal Q-Q chart, most of the points should be perfectly parallel to the red line. The results of the SLC oil diagnostic test are presented in Fig. 7.

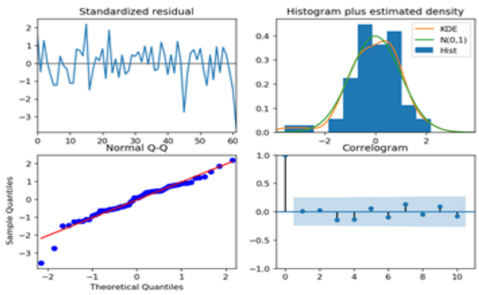


Fig. 7. The Result of ARIMA model's diagnostic testing for SLC oil

Appropriate to the diagnostic test results of the SLC oil ARIMA model in Fig. 7, it can be seen that the residual graph has a value that fluctuates around an average of zero and is not fixed. The correlogram graph for the lag line that exceeds one is not significant (passes the blue zone). In addition, on the KDE histogram graph, the shape of the curve line follows the normal distribution curve of the KDE histogram graph. Lastly, most of the blue dots are almost perfectly aligned with the red lines on the normal Q-Q chart.

6. Training Data Ratio

This stage is the process of separating training data and test data. This stage is implemented using the panda's library and the iloc function.

7. Forecasting on Training Data

This stage is implemented using the Statsmodels pmdarima's library. The amount of data on the train variable or training data is 60 data in the period January 2015 to September 2019 data for each crude oil type or a total of 3,000 data. The amount of data on the test variable or test data is 6 data in the period October 2019 to March 2020 for each crude oil type or a total of 300 data. The forecasting outcomes from the training data are shown in Table 6.

According to Table 6, it has been found that the results of forecasting the training data, the ICP value from October 2019 to March 2020 tends to increase. The actual data also tended to increase until December 2019 and then decrease from January to March 2020. This decline was due to external factors, which the COVID-19 pandemic significantly affected crude oil demand worldwide.

Table 6. The Forcasting Result of Data Training

Types of Indonesian Crude Oil	Oct/2019	Nov/2019	...	Mar/2020
SLC	62.41716	62.75162	...	64.08944
Arjuna	63.35777	63.72267	...	65.18228
Attaka	64.87574	65.22896	...	66.64187
Cinta	61.49420	61.81359	...	63.09115
Duri	70.72576	71.23714	...	73.28267
Widuri	61.62518	61.95586	...	63.27855
Belida	64.01594	64.45043	...	65.72353
Senipah Condensate	58.99595	59.23095	...	60.17093
ANOA	65.27574	65.62896	...	67.04187
Arun Condensate	58.99595	59.23095	...	60.17093
Bekapai	64.87574	65.22896	...	66.64187
⋮	⋮	⋮	⋮	⋮
Walio Mix	61.32356	61.68915	...	63.15151

8. Evaluating of Forecasting result on Testing Data

This stage is implemented using the Numpy library, where the test parameter used is the MAPE value. Based on (1), the following results are obtained. Testing the results of forecasting training data using MAPE is presented in Table 7.

1) SLC

$$MAPE\ Arjuna = \frac{115.06414646}{6} = 19.17735774$$

2) Arjuna

$$MAPE\ SLC = \frac{106.34229023}{6} = 17.72371503$$

50) Walio Mix

$$MAPE\ WM = \frac{138.89092888}{6} = 23.14848814$$

Table 7. Evaluation of Forecasting result on Testing Data

No.	Types of Indonesian Crude Oil	MAPE (%)
1	SLC	17.7237150394625
2	Arjuna	19.1773577435903
3	Attaka	20.1076032624674
4	Cinta	18.6994720576033
5	Duri	17.7237150394625
⋮	⋮	⋮
50	Walio Mix	23.1484881471690

Based on Table 7, the forecasting results have a MAPE value of 15-23%. In the MAPE interpretation in Table 2, the forecasting results are categorized as good and proper forecasting (acceptable).

9. Forecasting using Forecasting Dataset

The ARIMA model that has been tested is applied for ICP forecasting in the period April to September 2020. The results of forecasting forecast data are presented in Table 8.

Appropriate to the results of forecasting training data and forecasting data, ARIMA's accuracy rate is highly dependent on the previous period's ICP data and external factors that can affect such as extraordinary events of the Covid-19 pandemic. The more unstable the data value is each month, the higher the MAPE value.

Table 8. Forecasting result of ICP on April-September 2020

Types of Indonesian Crude Oil	Apr/2020	Mei/2020	...	Sep/2020
SLC	29.09194	28.88167	...	28.04058
Arjuna	27.90949	27.69297	...	26.82691
Attaka	26.49403	26.22042	...	25.126
Cinta	26.68656	26.43667	...	25.4371
Duri	34.26134	34.16911	...	33.80023
Widuri	27.31916	27.08711	...	26.15889
Belida	24.83039	21.2588	...	17.86737
Senipah Condensate	25.58528	25.28069	...	24.06232
ANOA	26.89403	26.62042	...	25.52599
Arun Condensate	25.58528	25.28069	...	24.06232
Bekapai	26.49403	26.22042	...	25.126
⋮	⋮	⋮	⋮	⋮
Walio Mix	22.01958	21.74417	...	20.64254

4. CONCLUSION

According to the research that has been done, based on the results of parameter estimation and diagnostic tests for 50 types of Indonesian crude oil, the fittest ARIMA models are (0,1,1), (1,1,0), (1,1,1), and (1, 2.1).

ICP forecasting using the ARIMA method in the period October 2019 to March 2020 has a good and proper interpretation (acceptable) with an average forecasting error rate (MAPE) o,f 15-23%. This study can be developed for further research to compare with other economic forecasting methods based on time series, and combining the ARIMA method with other methods or supporting variables improves the accuracy of the forecast results.

REFERENCES

[1] Z. Cen and J. Wang, "Crude oil price prediction model with long short term memory deep learning based on prior knowledge data transfer," *Energy*, vol. 169, pp. 160–171, 2019, https://doi.org/10.1016/j.energy.2018.12.016.

[2] A. Wanto *et al.*, "Forecasting the export and import volume of crude oil, oil products and gas using ANN," in *Journal of Physics: Conference Series*, pp. 1–6, 2019, https://doi.org/10.1088/1742-6596/1255/1/012016.

[3] H. Millia *et al.*, "The influence of crude oil prices volatility, the internet and exchange rate on the number of foreign tourist arrivals in Indonesia," *International Journal of Energy Economics and Policy*, vol. 10, no. 6, pp. 280–287, 2020, https://doi.org/10.32479/ijeep.10083.

[4] S. Herawati and A. Djunaidy, "Implementing Method of Empirical Mode Decomposition based on Artificial Neural Networks and Genetic Algorithms for Crude Oil Price Forecasting," in *Journal of Physics: Conference Series*, pp. 1–6, 2020, https://doi.org/10.1088/1742-6596/1569/2/022075.

[5] A. Hanafie and A. R. Hakim, "Forecasting with Feed Forward Neural Network model and adaptive simulated annealing algorithm (Case: world crude oil prices that was published by OPEC)," in *Journal of Physics: Conference Series*, pp. 1–11, 2019, https://doi.org/10.1088/1742-6596/1217/1/012079.

[6] I. Dawar, A. Dutta, E. Bouri, and T. Saeed, "Crude Oil Prices and Clean Energy Stock Indices: Lagged and Asymmetric Effects with Quantile Regression," *Renew Energy*, vol. 163, pp. 288–299, 2021, https://doi.org/10.1016/j.renene.2020.08.162.

[7] Sukono et al., "Application of ARIMA-GARCH Model for Prediction of Indonesian Crude Oil Prices," *Operations Research: International Conference Series*, vol. 1, no. 1, pp. 25–33, 2020, https://doi.org/10.47194/orics.v1i1.21.

[8] C. Tjahjaprijadi, *APBN 2019: Sehat, Adil dan Mandiri*, 2018, https://fiskal.kemenkeu.go.id/baca/2018/11/15/090302647864240-apbn-2019-anggaran-pemerintah-yang-sehat-adil-dan-mandiri.

[9] A. K. Rachmawati, "Peramalan Penyebaran Jumlah Kasus Covid19 Provinsi Jawa Tengah dengan Metode ARIMA," *Zeta - Math Journal*, vol. 6, no. 1, pp. 11–16, 2020, https://doi.org/10.31102/zeta.2021.6.1.11-16.

[10] J. Guo, "Oil price forecast using deep learning and ARIMA," in *2019 International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, pp. 241–247, 2019, https://doi.org/10.1109/MLBDBI48998.2019.00054.

[11] C. Dritsaki, D. Niklis, and P. Stamatiou, "Oil consumption forecasting using ARIMA models: An empirical study for Greece," *International Journal of Energy Economics and Policy*, vol. 11, no. 4, pp. 214–224, 2021, https://doi.org/10.32479/ijeep.11231.

[12] N. Darapaneni, D. Reddy, A. R. Paduri, P. Acharya, and H. S. Nithin, "Forecasting of COVID-19 in India Using ARIMA Model," in *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, pp. 894–899, 2020, https://doi.org/10.1109/UEMCON51285.2020.9298045.

[13] G. Perone, "An ARIMA Model to Forecast the Spread and the Final Size of COVID-2019 Epidemic in Italy," *MedRxiv*, 2020, https://doi.org/10.1101/2020.04.27.20081539.

[14] V. K. R. Chimmula and L. Zhang, "Time Series Forecasting of Covid-19 Transmission in Canada Using LSTM Networks," *Chaos Solitons Fractals*, vol. 135, pp. 1– 6, 2020, https://doi.org/10.1016/j.chaos.2020.109864.

[15] J. F. Torres, D. Hadjout, A. Sebaa, F. Martínez-Álvarez, and A. Troncoso, "Deep Learning for Time Series Forecasting: A Survey," *Big Data*, vol. 9, no. 1, pp. 3–21, 2021, https://doi.org/10.1089/big.2020.0159.

[16] R. Krispin, *Hands-On Time Series Analysis with R: Perform time series analysis and forecasting using R*. Packt Publishing Ltd, 2019, https://books.google.co.id/books?hl=id&lr=&id=tTmbDwAAQBAJ.

[17] H. Yonar, A. Yonar, M. A. Tekindal, and M. Tekindal, "Modeling and Forecasting for the number of cases of the COVID-19 pandemic with the Curve Estimation Models, the Box-Jenkins and Exponential Smoothing Methods," *EJMO*, vol. 4, no. 2, pp. 160–165, 2020, https://doi.org/10.14744/ejmo.2020.28273.

[18] C. Katris, "A time series-based statistical approach for outbreak spread forecasting: Application of COVID-19 in Greece," *Expert Syst Appl*, vol. 166, pp. 1–9, 2021, https://doi.org/10.1016/j.eswa.2020.114077.

[19] D. Bello-Angulo, C. Mantilla-Duarte, E. Montes-Paez, and C. Guerrero-Martin, "Box–Jenkins Methodology Application to Improve Crude Oil Production Forecasting: Case Study in a Colombian Field," *Arab J Sci Eng*, vol. 47, 2021, https://doi.org/10.1007/s13369-021-05997-7.

[20] A. R. S. Parmezan, V. M. A. Souza, and G. E. Batista, "Evaluation of Statistical and Machine Learning Models for Time Series Prediction: Identifying The State-Of-The-Art and The Best Conditions for The Use of Each Model," *Inf Sci (N Y)*, vol. 484, pp. 302–337, 2019, https://doi.org/10.1016/j.ins.2019.01.076.

[21] X. Qiu, Y. Ren, P. N. Suganthan, and G. A. J. Amaratunga, "Empirical Mode Decomposition based ensemble deep learning for load demand time series forecasting," *Appl Soft Comput*, vol. 54, pp. 246–255, 2017, https://doi.org/10.1016/j.asoc.2017.01.015.

[22] F. Prado, M. C. Minutolo, and W. Kristjanpoller, "Forecasting Based on An Ensemble Autoregressive Moving Average-Adaptive Neuro-Fuzzy Inference System–Neural Network-Genetic Algorithm Framework," *Energy*, vol. 197, pp. 1–13, 2020, https://doi.org/10.1016/j.energy.2020.117159.

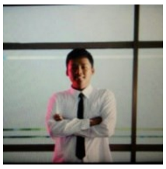
[23] J. Mboso, "The Autoregressive Integrated Moving Average Models of The Classical Box-Jenkins Methods of Time Series Analysis," *American Journal of Statistics and Actuarial Sciences*, vol. 4, no. 1, pp. 18–34, 2022, https://www.ajpojournals.org/journals/index.php/AJSAS/article/view/1294.
[24] A. Hernandez-Matamoros, H. Fujita, T. Hayashi, and H. Perez-Meana, "Forecasting of COVID19 per regions using ARIMA models and polynomial functions," *Appl Soft Comput*, vol. 96, pp. 1–14, 2020, https://doi.org/10.1016/j.asoc.2020.106610.
[25] E. A. Irhami and F. Farizal, "Forecasting the Number of Vehicles in Indonesia Using Auto Regressive Integrative Moving Average (ARIMA) Method," in *Journal of Physics: Conference Series*, pp. 1–7, 2021, https://doi.org/10.1088/1742-6596/1845/1/012024.
[26] S. Atique, S. Noureen, V. Roy, V. Subburaj, S. Bayne, and J. Macfie, "Forecasting of total daily solar energy generation using ARIMA: A case study," in *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 114–119, 2019, https://doi.org/10.1109/CCWC.2019.8666481.
[27] T. Alghamdi, K. Elgazzar, M. Bayoumi, T. Sharaf, and S. Shah, "Forecasting Traffic Congestion Using ARIMA Modeling," in *2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 1227–1232, 2019, https://doi.org/10.1109/IWCMC.2019.8766698.
[28] E. Sartika and S. Murniati, "Peramalan Distribusi Kedatangan Turis Mancanegara Melalui Pintu Masuk Bandara Soekarno Hatta Menggunakan ARIMA," *Jurnal Saintika Unpam: Jurnal Sains dan Matematika Unpam*, vol. 4, no. 1, pp. 52–70, 2021, https://garuda.kemdikbud.go.id/documents/detail/2433043.
[29] M. A. Omar, S. E. Shahin, and M. Roshdy, "Forecasting Inflation in Egypt (2019-2022) by using AutoRegressive Integrated Moving Average (ARIMA) Models," *J Adv Vet Res*, vol. 12, no. 6, pp. 670–676, 2022, https://www.advetresearch.com/index.php/AVR/article/view/1099/606.
[30] B. Nath, D. S. Dhakre, and D. Bhattacharya, "Forecasting Wheat Production in India: An ARIMA Modelling Approach," *J Pharmacogn Phytochem*, vol. 8, no. 1, pp. 2158–2165, 2019, https://www.phytojournal.com/archives/2019/vol8issue1/PartAJ/7-6-238-422.pdf.
[31] I. M. Wirawan, T. Widiyaningtyas, and M. M. Hasan, "Short Term Prediction on Bitcoin Price Using ARIMA Method," in *2019 International Seminar on Application for Technology of Information and Communication (iSemantic)*, IEEE, pp. 260–265, 2019, https://doi.org/10.1109/ISEMANTIC.2019.8884257.
[32] A. Deka and N. G. Resatoglu, "Forecasting Foreign Exchange Rate and Consumer Price Index with ARIMA model: The Case of Turkey," *International Journal of Scientific Research and Management (IJSRM)*, vol. 7, no. 8, pp. 1254–1275, 2019, https://doi.org/10.18535/ijsrm/v7i8.em01.
[33] R. Patil, D. M. Nagaraj, B. S. Polisgowdar, and S. Rathod, "Forecasting Potential Evapotranspiration for Raichur District Using Seasonal ARIMA Model," *Mausam*, vol. 73, no. 2, pp. 433–440, 2022, https://doi.org/10.54302/mausam.v73i2.5488.
[34] M. As'ad, S. S. Wibowo, and E. Sophia, "Peramalan Jumlah Mahasiswa Baru Dengan Model Autoregressive Integrated Moving Average (Arima)," *JIMP-Jurnal Informatika Merdeka Pasuruan*, vol. 2, no. 3, pp. 20–33, 2017, https://doi.org/10.37438/jimp.v2i3.77.
[35] M. Wang *et al.*, "ARIMA and ARIMA-ERNN Models For Prediction Of Pertussis Incidence In Mainland China from 2004 to 2021," *BMC Public Health*, vol. 22, no. 1, p. 1447, 2022, https://doi.org/10.1186/s12889-022-13872-9.
[36] K. D. Hartomo, S. Y. Prasetyo, and R. A. Suharjo, "Prediksi Stok dan Pengaturan Tata Letak Barang Menggunakan Kombinasi Algoritma Triple Exponential Smoothing dan FP-Growth," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 5, p. 869, 2020, https://doi.org/10.25126/jtiik.2020751863.
[37] M. Matyjaszek, P. Riesgo Fernández, A. Krzemień, K. Wodarski, and G. Fidalgo Valverde, "Forecasting coking coal prices by means of ARIMA models and neural networks, considering the transgenic time series theory," *Resources Policy*, vol. 61, pp. 283–292, 2019, https://doi.org/10.1016/j.resourpol.2019.02.017.
[38] M. A. A. Amin and M. A. Hoque, "Comparison of ARIMA and SVM for Short-term Load Forecasting," in *2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON)*, pp. 1–6, 2019, https://doi.org/10.1109/IEMECONX.2019.8877077.
[39] W. I. A. Wan Mohamad Nawi *et al.*, "Developing Forecasting Model for Future Pandemic Applications Based On Covid-19 Data 2020–2022," *PLoS One*, vol. 18, no. 5, pp. 11–26, 2023, https://doi.org/10.1371/journal.pone.0285407
[40] B. Dey, B. Roy, S. Datta, and T. S. Ustun, "Forecasting Ethanol Demand in India to Meet Future Blending Targets: A Comparison Of ARIMA And Various Regression Models," *Energy Reports*, vol. 9, pp. 411–418, 2023, https://doi.org/10.1016/j.egyr.2022.11.038.
[41] A. Apriliani, H. Zainuddin, A. Agussalim, and Z. Hasanuddin, "Peramalan Tren Penjualan Menu Restoran Menggunakan Metode Single Moving Average," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 6, pp. 1161–1168, 2020, https://doi.org/10.25126/jtiik.2020722732.
[42] I. A. S. A. Amra and A. Y. A. Maghari, "Forecasting Groundwater Production and Rain Amounts Using ARIMA-Hybrid ARIMA: Case Study of Deir El-Balah City in GAZA," in *2018 International Conference on Promising Electronic Technologies (ICPET)*, pp. 135–140, 2018, https://doi.org/10.1109/ICPET.2018.00031.
[43] T. Falatouri, F. Darbanian, P. Brandtner, and C. Udokwu, "Predictive Analytics for Demand Forecasting – A Comparison of SARIMA and LSTM in Retail SCM," *Procedia Comput Sci*, vol. 200, pp. 993–1003, 2022, https://doi.org/10.1016/j.procs.2022.01.298.

Vol. 9, No. 3, September 2023, pp. 720-730

BIOGRAPHY OF AUTHORS

Masna Wati, Lecture at Informatika Department in Mulawarman University. Master of informatics Engineering, Hasanuddin University. Deep research on Machine Learning, and Decision Support System. Email: masnawati@fkti.unmul.ac.id.

Haviluddin, Lecture at Informatika Department in Mulawarman University. Doctor of Computer Science, Universiti Malaysia Sabah. Concern in Artificial Intelligence, Data Analytics and Time Series Prediction. Email: haviluddin@unmul.ac.id.



Akhmad Masyudi, Bachelor's degree in the Informatics department at Mulawarman University. He is interested in studying machine learning. Email: akhmadmasyudii@gmail.com.

Anindita Septiarini, Lecture at Informatika Department in Mulawarman University. Doctor of Computer Science, Gadjah Mada University. Concern in Machine Learning and Computer Vision. Email: anindita@unmul.ac.id.

Heliza Rahmania Hatta, Lecture at Informatika Department in Mulawarman University. She obtained master's degrees from Sepuluh Nopember Institute of Technology. Her research focuses on Artificial Intelligence, and Decision Support System. Email: heliza_rahmania@yahoo.com.