Forecasting Indonesian Coffee Exports using ARIMA

Veldelen Yaphira[1*], Agus Sofian Eka Hidayat[2]

[1,2] *Study Program of Actuarial Science, School of Business, President University, 17550, Indonesia*

[*] *Corresponding author: veldelen.yaphira@student.president.ac.id*

Abstract— Coffee is one of Indonesia's leading export commodities in the agricultural sector. Good coffee quality makes Indonesia one of the largest coffee exporting countries in the world. The overall volume of Indonesian coffee exports in the world has decreased from year to year. Therefore, forecasting the volume of Indonesia's coffee exports within the next few years with the ARIMA Box-Jenkins method will greatly assist in a better export planning function for various parties. In this paper, we found that by using ARIMA (2,2,1) model, we can get the best prediction with a MAPE of 4124,64293%. With this predicted data, various parties can use this overview of the projected performance of coffee commodities in the short and medium term to maintain competitiveness amid the dynamics of the coffee commodity.

Keywords— Indonesian Coffee Exports; ARIMA Box-Jenkins; Forecasting; MAPE.

I. Introduction

Coffee was introduced to the archipelago by the Dutch who initially planted coffee trees around their territory in Batavia but then quickly expanded coffee production to the Bogor and Sukabumi areas from the 17th century to the 18th century. Indonesia proved to have an almost ideal climate for coffee production and so plantations were soon established in other areas of Java, Sumatra, and Sulawesi as well. Indonesia is one of the largest coffee-producing and exporting countries in the world [1]. Indonesia was the 4th largest coffee producer in the world in 2017, but in terms of exports, it was ranked 7th. On the other hand, this is due to domestic consumption increasing rapidly, namely 43% of the total coffee production in 2017 compared to only 36.5% in 2010 [2]. The value of Indonesian coffee exports fluctuates. Fluctuations in export value are more influenced by changes in coffee prices compared to changes in export volume [3].

Many previous researchers have discussed coffee exports in Indonesia. According to Alexander and Nadapdap (2019) the condition of Indonesian coffee bean exports has strong competitiveness, besides that Indonesian coffee bean commodities can seize market share for coffee beans in the global market and the export trend of Indonesian coffee beans has a stronger trend from 2002 to 2017 [4]. Further, Ramadhani (2018) hoped that the government and relevant agencies will be able to maintain the existing market through trade relations with other countries [5].

The analysis of the number of coffee exported should be considered. One way that can be done is through the forecasting method. Forecasting methods of coffee have been done in many ways such as the ARIMA method. Previous research applied the Autoregressive Integrated Moving Average (ARIMA) method in forecasting Indonesian coffee export prices [6]. ARIMA is commonly used for forecasting in finance such as gold price [7], forecasting the number of crimes [8] electricity using the historical approach. To give great results in forecasting and avoid too much error, many data from previous years need to be considered. It can be seen from previous coffee export data that the overall volume of Indonesian coffee exports in the world has decreased from year to year. From 1990-2015, coffee exports fluctuated and it became difficult to predict future export performance. Therefore, forecasting the volume of Indonesia's coffee exports in the next few years will greatly assist in a better export planning function for various parties. This paper is prepared to provide an overview of the projected performance of coffee commodities in the short and medium term to maintain competitiveness in the dynamics of the coffee commodity. Export trends can be predicted using the ARIMA model. The data used is time series data on the annual coffee export volume from 1990 to 2015 published by UN Comtrade. The purpose of this study is to analyze the trade of Indonesian coffee exports in the international market. This research is expected to be useful for the government to serve as information material in policy-making to realize the development of the competitive Indonesian coffee trade. For external parties, this study is expected to be beneficial as a reference and consideration in evaluating and making plans related to the development of the Indonesian coffee trade in the international market.

II. Literature Review

A. Introduction to Time Series

Time series is a series of observations on a variable taken from time to time and recorded sequentially according to the time sequence of occurrence with fixed time intervals. Time series can also be interpreted as a series of data obtained based on observations of an event in the order in which it occurred. The time of occurrence can be a period in units of seconds, minutes, hours, days, months, years, and other periods, all of which are a series of observational data based on the time of the incident with a certain time interval, which is better known as the time series. Time series analysis is used to obtain time series data patterns in the past which will be used to predict a variable value or condition in the future.

B. Stochastic Process

The stochastic process $\{Y_t; t = T\}$ is a set of random variables where T is the index T assigned to all random variables $Y_t, t \epsilon T$, defined in the same sample. If the index set by T shows time, then the stochastic process is referred to as a time series. The stochastic process properties are as follows:

1) *Mean*

$$\mu = E(Y_t) \; for \; t = 0, \pm 1, \pm 2, \ldots \tag{1}$$

2) *Autocovariance*

$$\gamma_{t,s} = Cov(Y_t, Y_s) \; for \; t, s = 0, \pm 1, \pm 2, \ldots \tag{2}$$

where

$$\begin{aligned} Cov(Y_t, Y_s) &= E[(Y_t - \mu_t)(Y_s - \mu_s)] \\ &= E(Y_t Y_s) - \mu_t \mu_s \end{aligned} \tag{3}$$

3) *Autocorrelation*

$$\rho_{t,s} = Corr(Y_t, Y_s) \; for \; t = 0, \pm 1, \pm 2, \ldots. \tag{4}$$

and

$$Corr(Y_t, Y_s) = \frac{Cov(Y_t, Y_s)}{\sqrt{Var(Y_t)}\sqrt{Var(Y_s)}} = \frac{\gamma_{t,s}}{\sqrt{\gamma_{t,t}}\sqrt{\gamma_{s,s}}} \tag{5}$$

C. Stationarity

The stationary time series is related to the consistency of the movement of the time series data. Data can be said to be non-stationary if the average value and variance vary over time, so the data is said to be stationary if the data moves stable and converges around its average value without any fluctuations in the movement of positive or negative trends. The stationary data properties are as follows:

1) Mean

$$E(X_t) = \mu, \forall t \epsilon T \; (constant) \tag{6}$$

2) Variance

$$Var(X_t) = E(X_t - \mu)^2 = \sigma^2 \leq 0 \tag{7}$$

3) Autocovariance

$$Cov(X_t, X_s) = (X_{t-k}, X_{s-k}) \; for \; all \; t, k, dan \; s \tag{8}$$

For k=s:

$$\begin{aligned} Cov(X_t, X_s) &= Cov(X_{t-s}, X_{s-s}) \\ &= Cov(X_{t-s}, X_0) \end{aligned} \tag{9}$$

For k=t;

$$\begin{aligned} Cov(X_t, X_s) &= Cov(X_{t-t}, X_{s-t}) \\ &= Cov(X_{t-t}, X_0) \end{aligned} \tag{10}$$

So we get:

$$\begin{aligned}\gamma_{t,s} &= Cov(X_t, X_s)\\ &= Cov(X_{t-s}, X_0)\\ &= Cov(C, X_{s-t})\\ &= Cov(C, X_{s-t})\\ &= Cov\big(X_0, X_{|t-s|}\big)\\ &= \gamma_0, X_{|t-s|}\\ &= \gamma_{0,k}\,; where\ k = |t-s|\end{aligned} \tag{11}$$

X_t and X_s at the time difference $|t-s|$ not the original t and s. So, the autocovariance function becomes:

$$\gamma_k = Cov(X_t, X_{t-k}) \tag{12}$$

$$\rho_k = Corr((X_t, X_{t-k}) = \frac{\gamma_k}{\gamma_0} \tag{13}$$

Stationarity means that the data does not grow or decrease. Data can be said to be stationary if the pattern is in equilibrium around a constant average value and the variance around the average is constant for a certain time. Weak stationer indicates that the data plot gives a T value on data $\{Y_t; t = 1, 2, \ldots, T\}$ which fluctuates with constant variance around a fixed value. In other words, time series data is said to be stationary if the mean, variance, and covariance in each lag are the same at all times. If the time series data does not meet these criteria, then the data is said to be non-stationary. Time series data is said to be non-stationary if the mean and variance are not constant, changing over time. A process is called a white noise process if the series consists of uncorrelated random variables and is normally distributed with a constant mean $\mu_t = E(e_t) = \mu$, constant variance $Var(e_t) = \sigma_t^2$ and $\gamma_k = Cov(e_t, e_{t-k})$ for k=0 and 0 for k≠0. Thus, a stationary white noise process with an autocovariance function:

$$\gamma_k = \begin{cases}\sigma_t^2, & k = 0\\ 0, & k \neq 0\end{cases} \tag{14}$$

Autocorrelation function:

$$\rho_k = \begin{cases}1, & k = 0\\ 0, & k \neq 0\end{cases} \tag{15}$$

So, it can be concluded that the covariance depends on k, not t. Therefore, white noise process is stationary.

D. Time Series Analysis Transformation

Several data transformations useful in time series analysis are provided:

- Generate date variables to establish periodicity and to differentiate between historical, validation, and forecasting periods.
- Create a new time series variable as a function of an existing time series variable.
- Replace missing values from the system and users with estimates based on one of several methods.

E. Autocorrelation Function (ACF)

The autocorrelation function, ρ_k is a measure of the correlation between two values X_t and X_{t+k} with a distance of k parts or is called the correlation coefficient at lag k. For a stationary X_t there is a mean value of $E(X_t) = \mu$ and the variance of $Var(X_t) = E(X_t - \mu)^2 = \sigma^2$ is constant. The autocovariance between X_t and X_{t+k} is as follows:

$$\gamma_k = Cov(X_t, X_{t-k}) = E(X_t - \mu)(X_{t+k} - \mu) \tag{16}$$

And the correlation between

$$\rho_k = Cov(X_t, X_{t+k}) = \frac{Cov(X_t,\ X_{t+k})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t+k})}} \tag{17}$$

or

$$\rho_k = \left(\frac{\gamma_k}{\gamma_0}\right) \text{where } Var(X_t) = Var(X_{t+k}) = \gamma_0 \tag{18}$$

In time series analysis γ_k is called an autocovariance function and ρ_k called an autocorrelation function which is a measure of the closeness between X_t and X_{t+k} from the same process and is only separated by the k-th time interval [6]. The stationary conditions of autocovariance and autocorrelation functions can be expressed with the following conditions:

1) $\gamma_0 = Var(\bar{X})\ and\ \rho_0 = 1$
2) $|\gamma_k| \leq \gamma_0\ and\ |\rho_k| \leq 1$
3) $\gamma_k = \gamma_{-k}\ and\ \rho_k = \rho_{-k}$

F. Partial Autocorrelation Function (PACF)

The partial autocorrelation function (PACF) is used to measure the degree of closeness of the relationship between X_t and X_{t+k} after removing the influence of linear dependencies in the variables $X_{t+1}, X_{t+2}, \dots, X_{t+k-1}$, so the PACF function can be expressed as:

$$X_{t+k} = \phi_{k1}X_{t+k-1}, \phi_{k2}X_{t+k-2}, \dots, \phi_{kk}X_t + e_{t+k} \tag{19}$$

where:
ϕ_{k+i} = i-th regression parameter
e_{t+k} = error values that are not related to X_{t+k-j}, for j = 1, 2, …, k.

Partial autocorrelation function:

$$\phi_{kk} = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \tag{20}$$

TABLE 1
ACF AND PACF FOR AR, MA, AND ARMA MODELS

Model	ACF	PACF
AR(p)	Tails off	Cut off after lag p
MA(q)	Cut off after lag q	Tails off
ARMA(p,q)	Tails off	Tails off

A. ARIMA Model

Forecasting methods commonly used are Auto-Regressive Integrated Moving Average (ARIMA) or Box-Jenkins. The Autoregressive Integrated Moving Average (ARIMA) methods have been studied in depth by George Box and Gwilym Jenkins (1976), and their name is often synonymous with the ARIMA process which is applied to time series analysis, forecasting, and control.

The process of forming the ARIMA model includes several stages, namely identification, estimation, diagnostic checks, and forecasting. The identification stage is carried out by observing the ACF and PACF plots from the data which is then used to obtain the appropriate ARIMA model prediction. The next stage is to estimate and test the significance of the parameters and whether the estimated provisional model is sufficient by the time series data.

The identification of the ARIMA model (p, d, q) was carried out after the data was stationary. If the data is not differencing, then d is 0 and if the data is stationary after the 1st differencing, then d is 1 and so on. The ARIMA Box Jenkins model has the following forms:

1) *Autoregressive (AR)*

A linear equation is said to be an autoregressive model if the model shows Y_t as a linear function of the actual number of Y_t in the previous period together with the current error. A time series model that has a unit root is a non-stationary time series model, but the opposite is not true [9]. The form of the model with order p:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t \tag{21}$$

where:
Y_t = time series data as response variable at time t
$Y_{t-1}, Y_{t-2} + \cdots + Y_{t-p}$= time series data to t-1, ..., t-p
$\phi_1, \phi_2, \dots, \phi_p$ = autoregressive parameters
e_t= error value at time t

2) *Moving Average (MA)*

Moving average model shows the Y_t value based on a combination of past linear errors or called lag. The form of the model with order q:

$$Y_t = e_t - \theta_1 e_{t-1} + \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \tag{22}$$

where:
$\theta_1, \theta_2, \dots, \theta_q$ = moving average parameter
$e_1, e_2, \dots, e_{t-q}$ = error value at time series t, t-1, …, t-q

3) *Autoregressive Integrated Moving Average (ARIMA)*

ARIMA is a time series model with stationary data through differencing. If the data is stationary in the differencing process d times, with the basic ARMA model (p, q) then the model formed becomes ARIMA (p, 0, q) where p is the order of R, d is the level of differencing, and q indicates the order of MA. ARIMA model (p, d, q) in general:

$$W_t = \nabla^d Y_t \tag{23}$$

where:
W_t = ARMA model (p, q)
Y_t = ARIMA model (p, d, q)

B. *Forecasting*

Forecasting methods commonly used are Auto-Regressive Integrated Moving Average (ARIMA) or Box-Jenkins. The Autoregressive Integrated Moving Average (ARIMA) methods have been studied by George Box and Gwilym Jenkins (1976), and their name is often synonymous with the ARIMA process which is applied to time series analysis, forecasting and control.

Forecasting models, in general, can be expressed as, Y_t = pattern + error. Data is divided into identifiable components (patterns) and non-identifiable components (errors). So, the use of the forecasting method is to identify a forecasting model in such a way that the error is minimized. The use of forecasting techniques begins with exploring the conditions (data patterns) in the past to develop a model that fits the data pattern by using the assumption that the data pattern in the past will repeat itself in the future.

III. Research Method

The data used in this study is "Indonesian Coffee Exports in the World" data from 1990 to 2015”. The method used in this research is the study of literature and is carried out systematically obtained from documentation data in the form of texts, books, or other media to obtain as much information as possible, then perform simulations as applications to explain the theory that has been obtained. The steps taken in the research are as follows:

1. Test the stationary of the data by looking at the data plot (time series plot) and the ACF (Autocorrelation Function) plot.
2. If the data is not stationary yet, do power transforming differencing until it is stationary.
3. The data from the transformation and differentiation results are then tested for stationarity by looking at the data plot (time series plot) and identifying the minimum lag value in the ACF plot.
4. Determine the order of the model by looking at the ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots.
5. Establishing a provisional conjecture model based on a pre-determined model order.
6. Make parameter estimation for the ARIMA models to check the significance.
7. Perform a residual model to check whether it has met the white noise process.
8. Choose the best-suited model.
9. Do the forecasting.

The flowchart can be seen below:

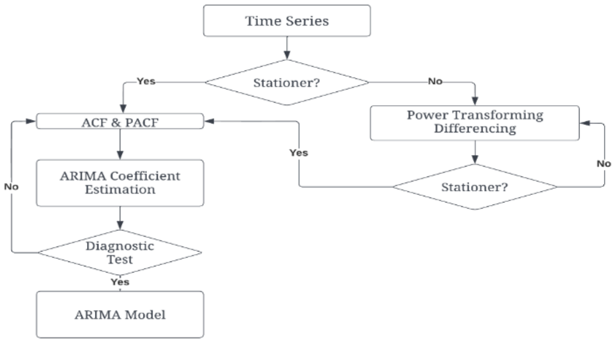

Figure. 1 ARIMA Box-Jenkins Flowchart

IV. Result and Discussion

A. Data Preparation

The data used is Indonesian Coffee Exports in the World" data from 1990 to 2015".

TABLE 2
INDONESIAN COFFEE EXPORT DATA, 1990-2020

Year	Volume (in kilogram)
1990	422162336
1991	380667104
1992	269349152
1993	349916352
1994	289303840
1995	230198624
1996	366603040
1997	313117536
1998	357550048
1999	352761828
2000	339200502
2001	250817493
2002	325009917
2003	323903645
2004	344076860
2005	445929794
2006	414105384
2007	321404023
2008	468749533
2009	510898385
2010	433590626
2011	346492592
2012	448590626
2013	534025073
2014	384827677
2015	502020679

B. Stationarity Check

To check the stationarity of the data, an ADF test will be performed in R Studio. If the p-value is more than 0.01, it can be concluded that the data is not stationary. Before differencing process, the results of the ADF test give a p-value is 0.3095>0.01, which can be concluded that the data is not stationary. Therefore, differencing will be carried out until the data is stationary. From the 1st differencing, and 2nd differencing it gives the p-value of 0.02172 and 0.01 respectively. As the p-value is less than equal to 0.01 after 2nd differencing it means that the data is stationer. It is also shown from the time series plot, autocorrelation function (ACF), and partial autocorrelation function (PACF). To determine p and q, we can use PACF for p values and ACF for q values. The p-value is 3 and the q-value is 1.

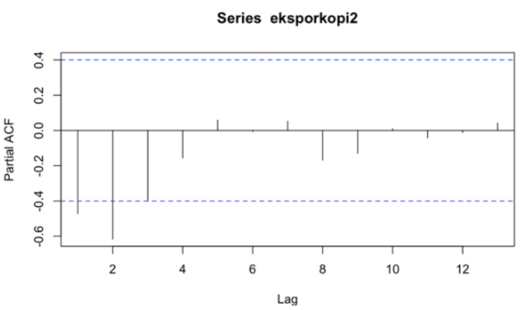

Figure 2. PACF after 2nd differencing

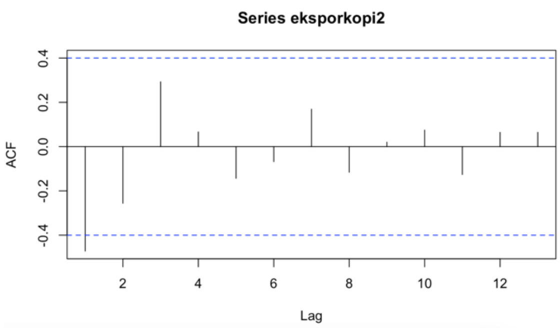

Figure 3. ACF after 2nd differencing

C. Model Specification

Based on the model specification for the ARIMA model, where p=3, d=2, and q=1. The following is a table of parameter estimation for all ARIMA models. Before that, the ARIMA model must be known, and then the estimated coefficients consisting of AR (1), AR (2), AR (3) and MA (1), Log-likelihood, and AIC will be considered for forecasting. Then, parameter estimates can be determined.

TABLE 2
PARAMETER ESTIMATION FOR ARIMA MODEL

Model ARIMA	Coefficient of Estimation Result					
	AR(1)	AR(2)	AR(3)	MA(1)	Log likelihood	AIC
ARIMA (3,2,1)	-0.5859	-0.6269	-0.0518	-0.9057	-468.71	945.42
ARIMA (2,2,1)	-0.5526	-0.5981		-0.9197	-468.73	943.47
ARIMA (1,2,1)	-0.3403			-1.0000	-472.64	949.29
ARIMA (0,2,1)				-1.0000	-474.01	950.03
ARIMA (3,2,0)	-1.1252	-1.0725	-0.4706		-470.96	947.92
ARIMA (2,2,0)	-0.8238	-0.7257			-473.45	950.9
ARIMA (1,2,0)	-0.5369				-480.41	962.83
ARIMA (0,2,0)					-483.95	967.89

D. Residual Analysis

Further analysis to determine the best model is through residual analysis using the Shapiro test and Ljung's test. The model can be said to be the best model if it passes both tests, where the p-value must be more than 0.05.

TABLE 3
ARIMA MODEL RESIDUAL ANALYSIS RESULTS

Model ARIMA	Saphiro Test	Ljung Test		AIC
		Lag = 4	Lag = 8	
ARIMA(3,2,1)	0.4987	0.8823	0.9525	945.42
ARIMA(2,2,1)	0.4585	0.8534	0.9475	943.47
ARIMA(1,2,1)	0.5436	0.03551	0.09987	949.29
ARIMA(0,2,1)	0.6232	0.03646	0.09256	950.03
ARIMA(3,2,0)	0.7521	0.6949	0.9357	947.92

ARIMA(2,2,0)	0.1917	0.2713	0.6283	950.9
ARIMA(1,2,0)	0.5082	0.00539	0.01857	962.83
ARIMA(0,2,0)	0.2641	0.02294	0.08401	967.89

From these results, all models meet the Shapiro test, while for the Ljung-Box test, several models failed to satisfy the test. So, it can be concluded that the models that pass the Shapiro test and Ljung-Box test are Model 1(3,2,1), Model 2(2,2,1), Model 4(0,2,1), Model 5(3,2, 0), and Model 6(2,2,0). By comparing the AIC of the five models, it can be found that Model 2(2,2,1) is the best because it has the smallest AIC compared to the three models. After testing Model 2(2,2,1), from the forecasting results, it can be seen which model is closest to the original data, and for this case the model that is closest to the data is Model 2, namely ARIMA (2,2,1) with

$$Y_t = -0.5526Y_{t-1} - 0.5981Y_{t-2} - 0.9197e_{t-1} + e_t \tag{24}$$

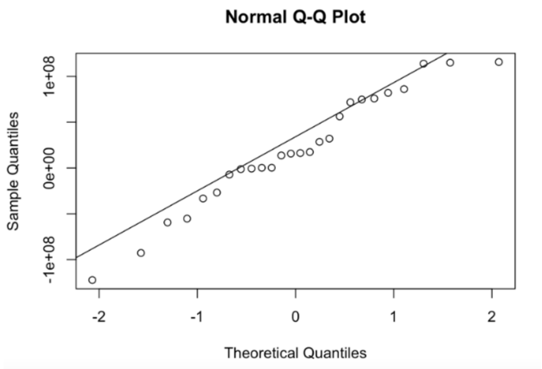

Figure. 4 Q-Q plot

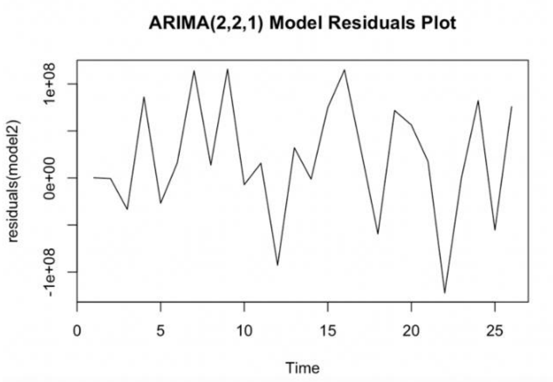

Figure. 5 Plot for residual ARIMA (2,2,1) model

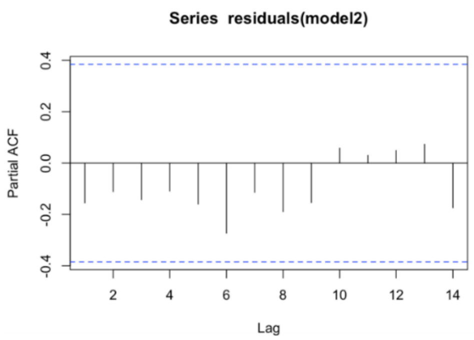

Figure. 7 PACF for residual ARIMA (2,2,1) model

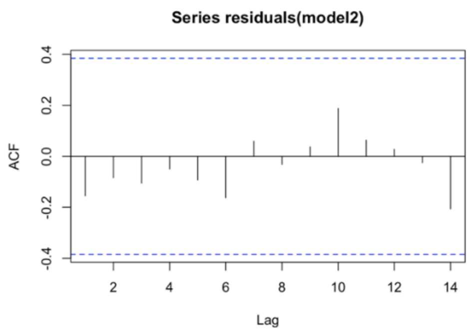

Figure. 8 ACF for residual ARIMA (2,2,1) model

Q-Q plots are used to assess the distribution assumptions of the model. If it is assumed to be normally distributed, the standardized residual plot has no serial correlation, no heteroscedasticity, or other linear dependence. In the figures described above, on the Q-Q plot all points lead to the normal line, it can be assumed that the data is normally distributed and can be used for forecasting. The statement is also supported by the residual histogram figure above.

E. Forecasting and Comparison

Forecasting the number of Indonesian coffee exports in the world for the next 5 years from 2016 to 2020 with ARIMA (2,2,1) using a 95-confidence interval. The black line is the actual data, and the blue line is the predictive data. The following table is a comparison between the actual data and the predicted data from 2016 to 2020. The actual data is still in the forecasting interval with the ARIMA (2,2,1) model for 2016, 2017, 2019, and 2020. However, in 2018 coffee exports in Indonesia experienced a significant decline because domestic coffee production was not good, so it did not enter the forecast interval with this ARIMA model.

TABLE 4

FORECASTING RESULTS WITH ARIMA (2,2,1) MODEL

Year	Lower Bound	Upper Bound	Prediction	Actual
2016	410806266	674539990	542673128	414651150
2017	317048198	615540161	466294179	467797006
2018	346841196	653902946	500372071	279960851
2019	358649609	728158851	543404230	359053322
2020	306770040	724074529	515422284	359053322

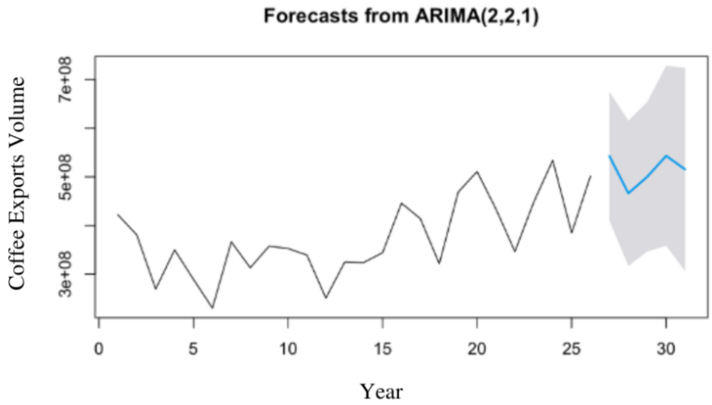

Figure. 8 Forecast Plot with ARIMA (2,2,1)

Based on the table of error values for the ARIMA (2,1,1) model, the actual data and the estimated data obtained are much different. Calculations using other ARIMA models also produce close error values because these models have AIC values with very small differences in each model.

TABLE 5

ERROR VALUE FOR ARIMA (2,2,1) MODEL

Metode	Nilai
MSE	2,4935E+16
RMSE	56460210,63
MAE	139197159,2
MAPE	4124,64293%

V. CONCLUSION

Based on the results of the analysis of coffee exports in Indonesia from 1990 to 2015 using the Box-Jenkins ARIMA method, it can be concluded as follows. In general, the condition of the volume of coffee exports in Indonesia has a significant downward trend every year. Especially in 2018, it can be seen from the forecasting and actual tables that the forecasting results are not accurate. This is caused by various internal and external factors, such as the unfavorable conditions in coffee plantations so that the number of imports also rises. The result of the best forecasting model with data on "Indonesian Coffee Exports in the World" from 1990-2015 is the ARIMA (2,2,1) model with an error value of MAPE 4124,64293%. Forecasting Indonesian coffee exports using the

ARIMA method is considered not accurate enough by only using previous data. The actual data and the predicted data for 2016 and 2017 are still in the interval, but for the next three years, the values obtained are far from the actual data. Therefore, it is necessary to consider other variables in predicting Indonesian coffee exports to obtain accurate forecasting results.

REFERENCES

[1] Kementerian Luar Negeri Indonesia, “Kopi Indonesia.” Accessed on: April 18, 2022, available: https://kemlu.go.id/chicago/id/read/kopi-indonesia/4484/etc-menu

[2] Kementerian Perdagangan Indonesia, “Model for Stationary Time Series.” Accessed on: April 18, 2022, available: https://djpen.kemendag.go.id/app_frontend/admin/docs/publication/9321548126511.pdf

[3] AEKI and PT Mitraconprima, “Analisis Potensi, Strategi Pengembangan, dan Road Map Perkopian Indonesia,” PT MITRACONPRIMA, Jakarta, 2010.

[4] Alexander, Ivan, and J. N. Hendrik, "Analisis Daya Saing Ekspor Biji Kopi Indonesia di Pasar Global Tahun 2002-2017," *Jurnal Sosial Ekonomi Pertanian (J-SEP)*, pp.1-16, 2019.

[5] R. Riska, “Analisis Ekspor Kopi Indonesia”, publication I, 2018.

[6] F. Emlan, E. P. Wawan, I. Andi, and B. A Herlen, “Pendugaan Model Peramalan Harga Ekspor Kopi Indonesia”, *Jurnal AGRITEPA*, 2020.

[7] R. J. Djami and S. J. Latupeirissa, “Peramalan Harga Emas di Indonesia Tahun 2014-2019 Dengan Metode Variance,” *Journal of Statistic and Its Application*, 2020.

[8] P. Chen, H. Yuan, and X. Shu, “Forecasting crime using the arima model,” In *2008 fifth international conference on fuzzy systems and knowledge discovery*, vol. 5, IEEE, 2008, pp. 627-630.

[9] Jing, “Unit Root and ARIMA Models,” 2014. Accessed on: Aug. 12, 2022, available: http://www.fsb.miamioh.edu/lij14/672_2014_S6.pdf