International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Forecasting Analysis of the Number of Visits of Foreign Tourists to Indonesia Using the ARIMA and Seasonal ARIMA Models

Nur Fatzria Sudrajat[1], Riza Adrianti Supono[2]
[1]Master of Information Systems Management Department, Business Information System, Gunadarma University, Depok, West Java, Indonesia-16424
[2]Secretary of the Information Systems Management Department, Gunadarma University, Depok, West Java, Indonesia-16424
Email Address: [1]nurfatzria@gmail.com, [2]adrianti@staff.gunadarma.ac.id

Abstract**—***Since 1995 the tourism sector has become one of Indonesia's mainstay sectors in generating foreign exchange. But unfortunately, according to data from the Central Statistics Agency for 2020, the number of foreign tourists coming to Indonesia has decreased due to the pandemic covid-19. This resulted in a decrease in the country's foreign exchange earnings from the tourism sector. Therefore a model is needed to make forecasts related to the number of foreign tourists coming to Indonesia, which can be useful for the government and related parties in making decisions and policies. In forecasting, this study uses the method time series by using 2 models, namely the ARIMA model and the SARIMA model (Seasonal ARIMA). Based on the results of the research conducted, it can be concluded that this research can produce forecasts of the number of foreign tourist visits to Indonesia using the ARIMA model where the results of the best testing model are ARIMA (1,1,1). Meanwhile, for the SARIMA model, the results of the best test model were SARIMA (1,1,1,1,1,12). The results of the research show that the ARIMA model has a value error that is smaller than the value error of the SARIMA model is 13.86208 (13%). As well as value forecast with the ARIMA model closer to the actual value of the number of foreign tourist visits to Indonesia. So that it can be said that the ARIMA model is more suitable for predicting the number of foreign tourist visits to Indonesia*.

Keywords**—** *ARIMA, Forecasting, International Tourists, Seasonal ARIMA, Tourism.*

I. INTRODUCTION

Since 1995 the tourism sector has become one of Indonesia's mainstay sectors in generating foreign exchange for development where in that year the tourism sector ranks third after oil and gas and textiles in generating foreign exchange for the country's development (Nur Djakaria M, 2008). But unfortunately, according to data from the Central Statistics Agency for 2020, the number of foreign tourists coming to Indonesia has decreased due to the pandemic covid-19. The total number of foreign tourist visits to Indonesia in 2020 is 4.02 million visits. When compared to 2019, the number of foreign tourists decreased by 75.03 percent. According to data from the Central Statistics Agency for 2020, the pandemic threatens 13 million workers in the tourism sector and 32.5 million workers who are indirectly related to the tourism sector.

Therefore a model is needed to make forecasts related to the number of foreign tourists coming to Indonesia. This forecasting model is expected to be useful for the government and related parties in making decisions and policies related to the number of foreign tourists coming to Indonesia so that the tourism sector in Indonesia can increase.

Forecasting is a method used to estimate predictive information to determine the direction of future decisions for both companies and government agencies, using historical data or references. Forecasting is also a science that can predict events in the future using past data (Heizer., 2019).

In forecasting, this study uses the method time series by using 2 models, namely the ARIMA (Autoregressive Integrated Moving Average) and the SARIMA model (Seasonal Autoregressive Integrated Moving Average). Where In this study, the researcher proposes research to predict the number of foreign tourist visits to Indonesia using the ARIMA (Autoregressive Integrated Moving Average) model data set the number of foreign tourist arrivals starts from 2017 to 2022. The research phase starts with identifying the testing model, estimating the ARIMA (Autoregressive Integrated Moving Average) and SARIMA (Seasonal Autoregressive Integrated Moving Average) models, estimating model parameters using ACF (Auto Correlation Function) and PACF (Partial Auto Correlation Function) parameters, testing model parameters, examining diagnostic which will then select the best model for the ARIMA and SARIMA models, and finally forecasting. The expected results in this study are to obtain the best ARIMA and SARIMA models so that they can predict the number of foreign tourist visits to Indonesia during 1 year, from January 2023 to December 2023. As well as determine which model is better between the ARIMA and SARIMA models to determine the number of foreign tourist visits to Indonesia by comparing the MAPE error values (Mean Absolute Percentage Error) as well as comparing the actual data released by the Central Agency Statistics for January 2023 to March 2023.

II. LITERATURE REVIEW

A. Traveller

Based on RI Law NO. 10 of 2009, it is stated that tourists are people who travel. Whereas tourists according to (Yoeti, 1996) are visitors who stay temporarily in a place for at least 24 hours in the city or country they visit with the motivation of traveling only for holidays, fun, health, study, religion, sports, family visits, conferences and so on. certain mission.

According to (Sugiama, 2011), tourists are individuals or groups who travel to rest, do business, do medical treatment, or make religious visits and study trips. By traveling and leaving his place of residence for a while, he can be said to be a tourist. Apart from that, in traveling, a tourist has intentions and goals such as resting, doing business, and other things in his tourist destination.

B. ARIMA (Autoregressive Integrated Moving Average)

Model ARIMA (Autoregressive Integrated Moving Average) is one of the popular time series models used in research so quite a lot of researchers use this model to make predictions with this model. Model Autoregresive Integrated Moving Average (ARIMA) was developed by George E. P. Box and Gwilyam M. Jenkins. Model identification can be seen from the ACF results (Auto Correlation Function) which is the Determination of p and q with the help of autocorrelation correlogram and PACF (Partial Auto Correlation Function) which is a measure of the correlation between observations with a k-th lag and by controlling for the correlation between two observations with a lag of less than k. ARIMA is very good in short-term forecasting accuracy, but less precise in long-term forecasting. ARIMA is a model that ignores independent variables in making a forecast and a model that assumes data must be stationary (Wei, 1990).

Forecasting with the ARIMA Box-Jenkins method will generally give better results than other forecasting methods because this method does not ignore the rules of time series data (Mulyana, 2004).

The general form of the ARIMA model is ARIMA (p, d, q) where p is ordo autoregressive, d is ordo Integrated and q is ordo moving average (Rahayu Wiwin Sri, 2019). The general formula for the ARIMA model (p, d, q) is as follows (Wei, 2016).

$$\emptyset_p(B)(1-B)^d Z_t = \emptyset_q(B)a_t$$

With:

$\emptyset_p(B) = \left(1-\emptyset_1 B - \cdots - \emptyset_p B^P\right)$, AR (p)

$\emptyset_Q(B) = \left(1-\emptyset_1 B - \cdots - \emptyset_Q B^Q\right)$, MA (q)

$(1-B)^d$: Differencing orde d

a_t : Nilai residual when t

C. SARIMA (Seasonal Autoregressive Integrated Moving Average)

According to (Suhartono 2011), forecasting using univariate time series data which was first introduced by Box and Jenkins in 1976 was Autoregressive Integrated Moving Average (ARIMA) and is still the most popular forecasting model. This model is derived from the model Autoregressive (AR), model Moving Average (MA) and a combination of AR and MA, in the ARMA model. In cases where there is a seasonal component in the model, this model is referred to as the SARIMA model (Suhartono, 2011)

The general form of the SARIMA model is SARIMA (p, d, q) (P, D, Q) S (Tadesse and Dinka, 2017). is a component Autoregressive which is used to model the autocorrelation contained in the time series by carrying out a regression on the lag variable of p, orde d is orde differencing to make non-stationary data stationary, orde q is orde Moving Average to model lag error as much as q, orde P is orde Seasonal Autoregressive, D orde differencing in the seasonal period, and Q is orde Seasonal Moving Average. The general formula for the ARIMA model (p, d, q) is as follows (Wei, 2016).

$$\emptyset_p B^S \emptyset_P(B)(1-B)^d(1-B^S)^D Z_t = \emptyset_q(B)\theta_q(B^S)a_t$$

With:

$\emptyset_p(B)$: AR nonseasonal

$\emptyset_p B^S$: AR seasonal

$(1-B)^d$: Differencing nonseasonal

$(1-B^S)^D$: Differencing seasonal

$\theta_q(B)$: MA nonseasonal

$\theta_Q(B^S)$: MA seasonal

The SARIMA method is known as the seasonal ARIMA method which was studied in depth by George Box and Gwilym Jenkins and consists of four stages, namely identification, estimation, diagnostic examination, and forecasting.

III. Result and Discussion

A. Research Stage

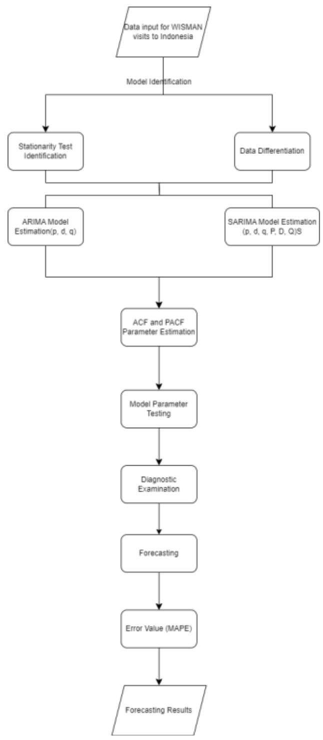


Fig. 1. Stages of the Research Model

The initial stages of this research began with inputting data on the number of international tourist visits to Indonesia, the data used came from the Central Statistics Agency (BPS) (www.bps.go.id) from January 2017 to December 2022, where the research data amounted to 72 record. Test data that has been processed forecasting will generate data forecasting for the number of tourist visits from January 2023 to December 2023. Then the next step is to identify the model using stationarity tests and data differentiation to obtain estimates of the ARIMA model (p, d, q) and the SARIMA model $(p, d, q, P, D, Q)^S$. After obtaining the suspected ARIMA and SARIMA models, the next step is to estimate the model parameters using ACF and PACF.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Then proceed with testing the model parameters by entering the formula to analyze the value of each model that has been obtained before. After getting the results of testing the model parameters, it continued with the diagnostic examination stage. This inspection stage is carried out by testing white noise, testing white noise said to be good and can be used as ARIMA and SARIMA modeling if the ACF and PACF plots have probabilities (p-value) > 0.05. After the testing and examination phase of the model estimation, the ARIMA and SARIMA models which have good accuracy values are obtained, then the data can be used for the forecasting process. Next, a comparison is made on which model is better to use following the forecasting results. The following is an explanation of the stages in this study.

B. Input Data on The Number of WISMAN Visits to Indonesia

Before testing the ARIMA dan SARIMA model, the first thing to do is perform input data training form file in Excel format (xls or xlsx) into the eviews 12 application by doing copy-paste into the eviews application.

C. Model Identification

After the data input process, the next step is to identify the model, which aims to find out what kind of ARIMA and SARIMA models match the data used in the research. The process of identifying this model also aims to find out whether the data used in the study are included in the stationary data or not. Stationary data itself is data that shows a value mean, variant, and car variant (in lag variations) remains the same at whatever time the data is formed, so it can be said that the data used is a stationary model time series that can be said to be more stable.

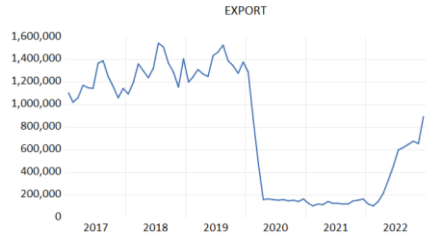
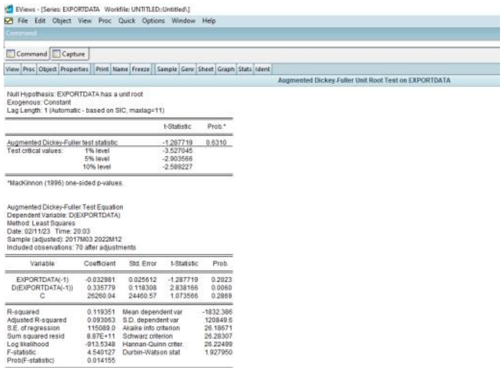


Fig. 2. Graph of Test Data

It can be seen in Figure 2 where the test graph shows that the data is not stationary because the average variance is inconsistent. It can be seen in the graph that the number of foreign tourist visits to Indonesia decreased sharply at the beginning of 2020 due to the pandemic Covid-19.

To make the data used stationary, stationary testing can be carried out on the test data. In this study, stationary testing uses a model Augmented Dickey-Fuller (ADF), through unit tests root as well as testing with the level of differentiation. In differentiation testing, 3 criteria are carried out, namely Level (Level 0),1st Difference (Level 1), and 2nd Difference (Level 2). If the data is stationary, then the differentiation test is stopped at that level (Rinaldo Isnawan Prasetyono and Dyah Anggraini, 2021).



EViews - [Series: EXPORTDATA Workfile: UNTITLED::Untitled\]

File Edit Object View Proc Quick Options Window Help

Augmented Dickey-Fuller Unit Root Test on EXPORTDATA

Null Hypothesis: EXPORTDATA has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.287719	0.6310
Test critical values:	1% level	-3.527045	
	5% level	-2.903566	
	10% level	-2.589227	

*MacKinnon (1996) one-sided p-values.

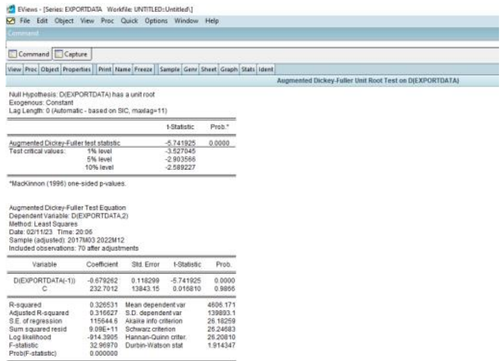
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(EXPORTDATA)
Method: Least Squares
Date: 02/11/23 Time: 20:03
Sample (adjusted): 2017M03 2022M12
Included observations: 70 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
EXPORTDATA(-1)	-0.032981	0.025612	-1.287719	0.2023
D(EXPORTDATA(-1))	0.335779	0.118308	2.838166	0.0060
C	26260.04	24460.57	1.073566	0.2869

R-squared	0.119351	Mean dependent var	-1832.386
Adjusted R-squared	0.093063	S.D. dependent var	120849.6
S.E. of regression	115089.0	Akaike info criterion	26.18671
Sum squared resid	8.87E+11	Schwarz criterion	26.28307
Log likelihood	-913.5348	Hannan-Quinn criter.	26.22499
F-statistic	4.540127	Durbin-Watson stat	1.927950
Prob(F-statistic)	0.014155		

Fig. 3. Level 0 Differentiation Stationary Testing ARIMA Model

In Figure 3 the stationarity test is carried out with differentiation at level 0 with the result obtained is that the probability value p-values H0 are 0.6310 > 0.05 and value t-statistics of -1.287719 > -3.527045 so that it indicates that the data above is not stationary when using level 0 differentiation, so it can be said that the H0 value is rejected.



EViews - [Series: EXPORTDATA Workfile: UNTITLED::Untitled\]

File Edit Object View Proc Quick Options Window Help

Augmented Dickey-Fuller Unit Root Test on D(EXPORTDATA)

Null Hypothesis: D(EXPORTDATA) has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.741925	0.0000
Test critical values:	1% level	-3.527045	
	5% level	-2.903566	
	10% level	-2.589227	

*MacKinnon (1996) one-sided p-values.

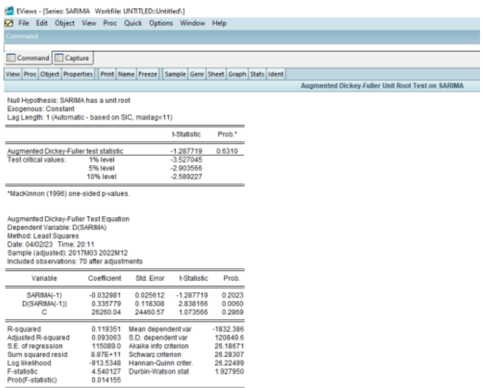
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(EXPORTDATA,2)
Method: Least Squares
Date: 02/11/23 Time: 20:06
Sample (adjusted): 2017M03 2022M12
Included observations: 70 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(EXPORTDATA(-1))	-0.679262	0.118299	-5.741925	0.0000
C	232.7012	13843.15	0.016810	0.9866

R-squared	0.326531	Mean dependent var	4606.171
Adjusted R-squared	0.316627	S.D. dependent var	139893.1
S.E. of regression	115644.6	Akaike info criterion	26.18259
Sum squared resid	9.09E+11	Schwarz criterion	26.24683
Log likelihood	-914.3905	Hannan-Quinn criter.	26.20810
F-statistic	32.96970	Durbin-Watson stat	1.914347
Prob(F-statistic)	0.000000		

Fig. 4. Level 1 Differentiation Stationary Testing ARIMA Model

Then in Figure 4 the stationarity test is carried out with differentiation at level 1 with the result obtained that the probability value p-values H0 are 0.0000 < 0.05 and value t-statistics of -5.741925 < -2.903566 indicating that the above data is stationary when using level 1 differentiation, so it can be said that the H1 value is accepted.



EViews - [Series: SARIMA Workfile: UNTITLED::Untitled\]

File Edit Object View Proc Quick Options Window Help

Augmented Dickey-Fuller Unit Root Test on SARIMA

Null Hypothesis: SARIMA has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.287719	0.6310
Test critical values:	1% level	-3.527045	
	5% level	-2.903566	
	10% level	-2.589227	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(SARIMA)
Method: Least Squares
Date: 04/02/23 Time: 20:11
Sample (adjusted): 2017M03 2022M12
Included observations: 70 after adjustments

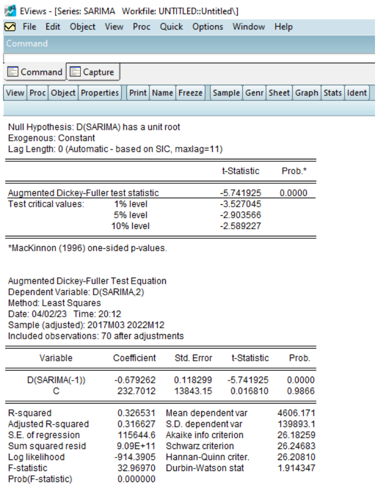
Variable	Coefficient	Std. Error	t-Statistic	Prob.
SARIMA(-1)	-0.032981	0.025612	-1.287719	0.2023
D(SARIMA(-1))	0.335779	0.118308	2.838166	0.0060
C	26260.04	24460.57	1.073566	0.2869

R-squared	0.119351	Mean dependent var	-1832.386
Adjusted R-squared	0.093063	S.D. dependent var	120849.6
S.E. of regression	115089.0	Akaike info criterion	26.18671
Sum squared resid	8.87E+11	Schwarz criterion	26.28307
Log likelihood	-913.5348	Hannan-Quinn criter.	26.22499
F-statistic	4.540127	Durbin-Watson stat	1.927950
Prob(F-statistic)	0.014155		

Fig. 5. Level 0 Differentiation Stationary Testing of the SARIMA Model

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

In Figure 5 the stationarity test is carried out with differentiation at level 0 with the result obtained that the probability value p-values H0 is 0.6310 > 0.05 and value t-statistics of -1.287719 > -3.527045 so that it indicates that the data above is not stationary when using level 0 differentiation, so it can be said that the H0 value is rejected.


EViews - [Series: SARIMA Workfile: UNTITLED::Untitled\]

Null Hypothesis: D(SARIMA) has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=11)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.741925	0.0000
Test critical values:	1% level	-3.527045	
	5% level	-2.903566	
	10% level	-2.589227	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(SARIMA,2)
Method: Least Squares
Date: 04/02/23 Time: 20:12
Sample (adjusted): 2017M03 2022M12
Included observations: 70 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SARIMA(-1))	-0.679262	0.118299	-5.741925	0.0000
C	232.7012	13843.15	0.016810	0.9866

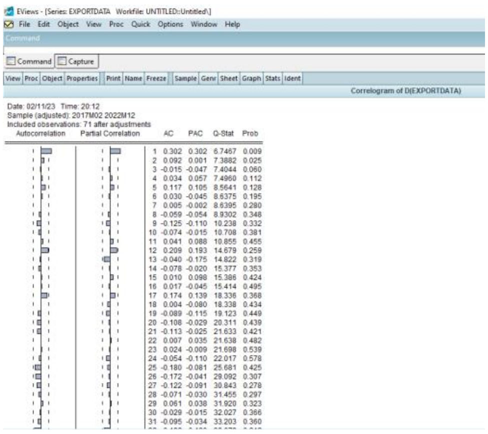
R-squared	0.326531	Mean dependent var	4606.171
Adjusted R-squared	0.316627	S.D. dependent var	139893.1
S.E. of regression	115644.6	Akaike info criterion	26.18259
Sum squared resid	9.09E+11	Schwarz criterion	26.24683
Log likelihood	-914.3905	Hannan-Quinn criter.	26.20810
F-statistic	32.96970	Durbin-Watson stat	1.914347
Prob(F-statistic)	0.000000		

Fig. 6. Level 1 Differentiation Stationary Testing of the SARIMA Model

Then in Figure 6 the stationarity test is carried out with differentiation at level 1 with the result obtained that the value probability p-valuesH0 is 0.0000 < 0.05 and value t-statistics of -5.741925 < -2.903566 indicating that the above data is stationary when using level 1 differentiation, so it can be said that the H1 value is accepted.

D. ACF and PACF Parameter Estimation

Furthermore, after testing the stationarity and differentiation tests, the next step is to test the ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) to get the parameters p and q. The results for testing ACF and PACF can be seen in Figure 7 & 8 as follows


EViews - [Series: EXPORTDATA Workfile: UNTITLED::Untitled\]

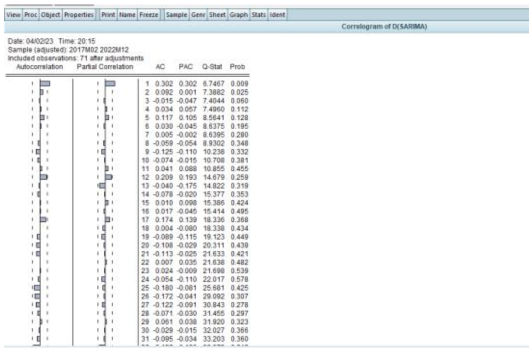
Correlogram of D(EXPORTDATA)

Date: 02/11/23 Time: 20:12
Sample (adjusted): 2017M02 2022M12
Included observations: 71 after adjustments

	AC	PAC	Q-Stat	Prob
1	0.302	0.302	6.7467	0.009
2	0.092	0.001	7.3882	0.025
3	-0.015	-0.047	7.4044	0.060
4	0.034	0.057	7.4960	0.112
5	0.117	0.105	8.5641	0.128
6	0.030	-0.045	8.6375	0.195
7	0.005	-0.002	8.6395	0.280
8	-0.059	-0.054	8.9302	0.348
9	-0.125	-0.110	10.238	0.332
10	-0.074	-0.015	10.708	0.381
11	0.041	0.088	10.855	0.455
12	0.209	0.193	14.679	0.259
13	-0.040	-0.175	14.822	0.319
14	-0.078	-0.020	15.377	0.353
15	0.010	0.098	15.386	0.424
16	0.017	-0.045	15.414	0.495
17	0.174	0.139	18.336	0.368
18	0.004	-0.080	18.338	0.434
19	-0.089	-0.115	19.123	0.449
20	-0.108	-0.029	20.311	0.439
21	-0.113	-0.025	21.633	0.421
22	0.007	0.035	21.638	0.482
23	0.024	-0.009	21.698	0.539
24	-0.054	-0.110	22.017	0.578
25	-0.180	-0.081	25.681	0.425
26	-0.172	-0.041	29.092	0.307
27	-0.122	-0.091	30.843	0.278
28	-0.071	-0.030	31.455	0.297
29	0.061	0.038	31.920	0.323
30	-0.029	-0.015	32.027	0.366
31	-0.095	-0.034	33.203	0.360

Fig. 7. ARIMA Model PACF and ACF Plot Results

It can be seen in Figure 7 The PACF plot occurs cut off on lag 1 is similar to the ACF plot where there is cut off in lag 1, so it can be determined that in testing data on foreign tourist visits to Indonesia, the estimation of the ARIMA model obtained is only 1 model, namely Autoregressive with a value of AR (1) and Autoregressive with a value of MA (1), so it can be said that the estimation of the ARIMA model on foreign tourist visit data to Indonesia has only 1 model, namely ARIMA (1,1,1).


Correlogram of D(SARIMA)

Date: 04/02/23 Time: 20:15
Sample (adjusted): 2017M02 2022M12
Included observations: 71 after adjustments

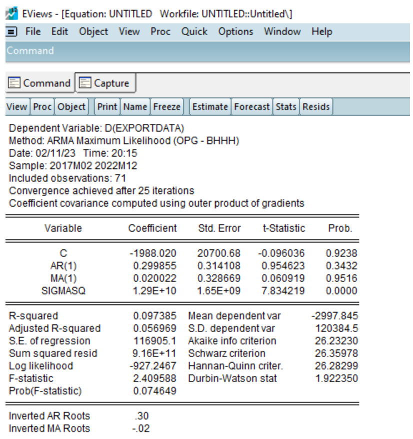
	AC	PAC	Q-Stat	Prob
1	0.302	0.302	6.7467	0.009
2	0.092	0.001	7.3882	0.025
3	-0.015	-0.047	7.4044	0.060
4	0.034	0.057	7.4960	0.112
5	0.117	0.105	8.5641	0.128
6	0.030	-0.045	8.6375	0.195
7	0.005	-0.002	8.6395	0.280
8	-0.059	-0.054	8.9302	0.348
9	-0.125	-0.110	10.238	0.332
10	-0.074	-0.015	10.708	0.381
11	0.041	0.088	10.855	0.455
12	0.209	0.193	14.679	0.259
13	-0.040	-0.175	14.822	0.319
14	-0.078	-0.020	15.377	0.353
15	0.010	0.098	15.386	0.424
16	0.017	-0.045	15.414	0.495
17	0.174	0.139	18.336	0.368
18	0.004	-0.080	18.338	0.434
19	-0.089	-0.115	19.123	0.449
20	-0.108	-0.029	20.311	0.439
21	-0.113	-0.025	21.633	0.421
22	0.007	0.035	21.638	0.482
23	0.024	-0.009	21.698	0.539
24	-0.054	-0.110	22.017	0.578
25	-0.180	-0.081	25.681	0.425
26	-0.172	-0.041	29.092	0.307
27	-0.122	-0.091	30.843	0.278
28	-0.071	-0.030	31.455	0.297
29	0.061	0.038	31.920	0.323
30	-0.029	-0.015	32.027	0.366
31	-0.095	-0.034	33.203	0.360

Fig. 8. PACF and ACF Plot Results for the SARIMA Model

In contrast to the ARIMA model which only uses p, d, q values, the SARIMA model uses $(p, d, q, P, D, Q)^S$ values, where there is additional values seasonal/season on models. It can be seen in Figure 8 that the PACF plot occurs cut off at lag 1 as well as the ACF plot where there is cut off in lag 1 and there was an increase cut off at lag 12 which is a value addition seasonal. So that in testing data on foreign tourist visits to Indonesia using the SARIMA model, there is only 1 model, namely SARIMA (1, 1, 1, 1, 1, 1)[12].

E. Model Parameter Testing

After obtaining an estimation of the ARIMA model from each test data that has been carried out based on the ACF and PACF tests, the next step is to test the parameters of the ARIMA model obtained. In this study to test the parameters of the ARIMA model were obtained using 2 criteria, namely AIC (Akaike Info Criterion) and SBC (Schwarz Info Criterion). If there is more than 1 model obtained, it is necessary to compare the AIC values (Akaike Info Criterion) and SBC (Schwarz Info Criterion). However, because this study only used 1 prediction model, no comparison was made with the AIC criterion value (Akaike Info Criterion) and SBC (Schwarz Info Criterion) other.


EViews - [Equation: UNTITLED Workfile: UNTITLED::Untitled\]

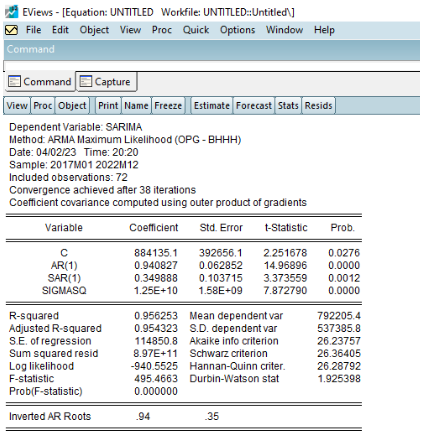
Dependent Variable: D(EXPORTDATA)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 02/11/23 Time: 20:15
Sample: 2017M02 2022M12
Included observations: 71
Convergence achieved after 25 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1988.020	20700.68	-0.096036	0.9238
AR(1)	0.299855	0.314108	0.954623	0.3432
MA(1)	0.020022	0.328669	0.060919	0.9516
SIGMASQ	1.29E+10	1.65E+09	7.834219	0.0000

R-squared	0.097385	Mean dependent var	-2997.845
Adjusted R-squared	0.056969	S.D. dependent var	120384.5
S.E. of regression	116905.1	Akaike info criterion	26.23230
Sum squared resid	9.16E+11	Schwarz criterion	26.35978
Log likelihood	-927.2467	Hannan-Quinn criter.	26.28299
F-statistic	2.409588	Durbin-Watson stat	1.922350
Prob(F-statistic)	0.074649		
Inverted AR Roots	.30		
Inverted MA Roots	-.02		

Fig. 9. ARIMA Model Test Results

Nur Fatzria Sudrajat, Riza Adrianti Supono, "Forecasting Analysis of the Number of Visits of Foreign Tourist to Indonesia Using the ARIMA and Seasonal ARIMA Models," *International Research Journal of Advanced Engineering and Science*, Volume 8, Issue 3, pp. 39-45, 2023.



EViews - [Equation: UNTITLED Workfile: UNTITLED::Untitled\]
File Edit Object View Proc Quick Options Window Help

Dependent Variable: SARIMA
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 04/02/23 Time: 20:20
Sample: 2017M01 2022M12
Included observations: 72
Convergence achieved after 38 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	884135.1	392656.1	2.251678	0.0276
AR(1)	0.940827	0.062852	14.96896	0.0000
SAR(1)	0.349888	0.103715	3.373559	0.0012
SIGMASQ	1.25E+10	1.58E+09	7.872790	0.0000

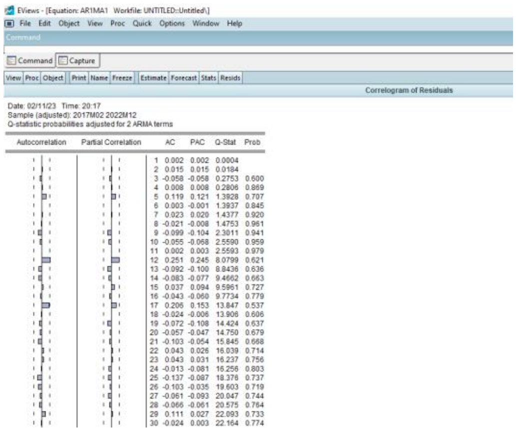
R-squared	0.956253	Mean dependent var	792205.4
Adjusted R-squared	0.954323	S.D. dependent var	537385.8
S.E. of regression	114850.8	Akaike info criterion	26.23757
Sum squared resid	8.97E+11	Schwarz criterion	26.36405
Log likelihood	-940.5525	Hannan-Quinn criter.	26.28792
F-statistic	495.4663	Durbin-Watson stat	1.925398
Prob(F-statistic)	0.000000		

Inverted AR Roots	.94	.35

Fig. 10. SARIMA Model Test Results

F. Diagnostic Examination

Diagnostic Checks are carried out to find out whether the modeling that has been obtained temporarily produces a significant value or not. The model is said to be significant if the probability (p-values) < 0.05. Testing carried out at this stage is by testing white noise. Testing with white noise is said to be good and can be used as a model if the ACF and PACF plots have probabilities (p-value) > 0.05. One way to see whether white noise can be tested is through correlogram ACF and PACF of residual. If ACF and PACF are not significant, this indicates a residual white noise which means the model is suitable, if not then the model is not suitable. In addition, testing can also be done by looking at ARMA structure, by looking at whether white noise on roots, with AR indicator roots in MA roots be on the unit circle.



EViews - [Equation: AR1MA1 Workfile: UNTITLED::Untitled\]
File Edit Object View Proc Quick Options Window Help
Correlogram of Residuals

Date: 02/11/23 Time: 20:17
Sample (adjusted): 2017M02 2022M12
Q-statistic probabilities adjusted for 2 ARMA terms

Lag	AC	PAC	Q-Stat	Prob
1	0.002	0.002	0.0004	
2	0.015	0.015	0.0184	
3	-0.058	-0.058	0.2753	0.600
4	0.008	0.008	0.2806	0.869
5	0.119	0.121	1.3928	0.707
6	0.003	-0.001	1.3937	0.845
7	0.023	0.020	1.4377	0.920
8	-0.021	-0.008	1.4753	0.961
9	-0.099	-0.104	2.3011	0.941
10	-0.055	-0.068	2.5590	0.959
11	0.002	0.003	2.5593	0.979
12	0.251	0.245	8.0799	0.621
13	-0.092	-0.100	8.8436	0.636
14	-0.083	-0.077	9.4662	0.663
15	0.037	0.094	9.5961	0.727
16	-0.043	-0.060	9.7734	0.779
17	0.206	0.153	13.847	0.537
18	-0.024	-0.006	13.906	0.606
19	-0.072	-0.108	14.424	0.637
20	-0.057	-0.047	14.750	0.679
21	-0.103	-0.054	15.845	0.668
22	0.043	0.026	16.039	0.714
23	0.043	0.031	16.237	0.756
24	-0.013	-0.081	16.256	0.803
25	-0.137	-0.087	18.376	0.737
26	-0.103	-0.035	19.603	0.719
27	-0.061	-0.093	20.047	0.744
28	-0.066	-0.061	20.575	0.764
29	0.111	0.027	22.093	0.733
30	-0.024	0.003	22.164	0.774

Fig. 11. ARIMA Model Diagnostic Examination

Can be seen in the picture where it is not there cut off on ACF and PACF so that it can be identified that the ARIMA model used meets the qualifications for forecasting.

In addition to using testing white noise, testing can also be done by looking at ARMA structure, by looking at whether white noise on roots, with AR indicator roots in MA roots, is on the unit circle.

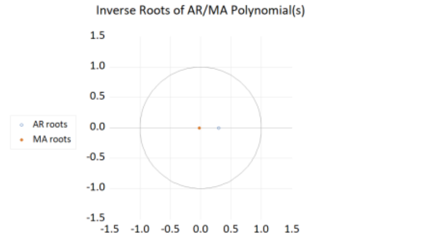
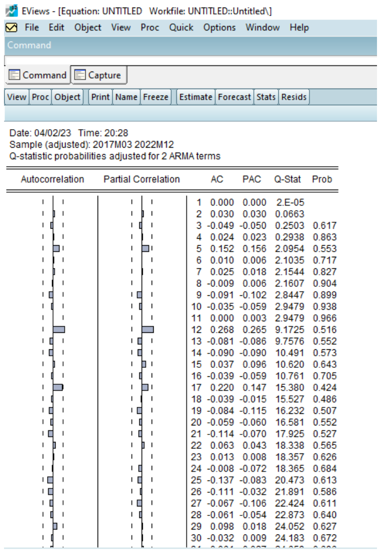


Fig. 12. ARIMA Model Diagnostic Examination (ARMA Structure)

It can be seen in the image that there are no AR points roots in MA roots which is an outside circle so it can be said that there is no AR and MA points white noise.



EViews - [Equation: UNTITLED Workfile: UNTITLED::Untitled\]
File Edit Object View Proc Quick Options Window Help

Date: 04/02/23 Time: 20:28
Sample (adjusted): 2017M03 2022M12
Q-statistic probabilities adjusted for 2 ARMA terms

Lag	AC	PAC	Q-Stat	Prob
1	0.000	0.000	2.E-05	
2	0.030	0.030	0.0663	
3	-0.049	-0.050	0.2503	0.617
4	0.024	0.023	0.2938	0.863
5	0.152	0.156	2.0954	0.553
6	0.010	0.006	2.1035	0.717
7	0.025	0.018	2.1544	0.827
8	-0.009	0.006	2.1607	0.904
9	-0.091	-0.102	2.8447	0.899
10	-0.035	-0.059	2.9479	0.938
11	0.000	0.003	2.9479	0.966
12	0.268	0.265	9.1725	0.516
13	-0.081	-0.086	9.7576	0.552
14	-0.090	-0.090	10.491	0.573
15	0.037	0.096	10.620	0.643
16	-0.039	-0.059	10.761	0.705
17	0.220	0.147	15.380	0.424
18	-0.039	-0.015	15.527	0.486
19	-0.084	-0.115	16.232	0.507
20	-0.059	-0.060	16.581	0.552
21	-0.114	-0.070	17.925	0.527
22	0.063	0.043	18.338	0.565
23	0.013	0.008	18.357	0.626
24	-0.008	-0.072	18.365	0.684
25	-0.137	-0.083	20.473	0.613
26	-0.111	-0.032	21.891	0.586
27	-0.067	-0.106	22.424	0.611
28	-0.061	-0.054	22.873	0.640
29	0.098	0.018	24.052	0.627
30	-0.032	0.009	24.183	0.672

Fig. 13. SARIMA Model Diagnostic Examination

Can be seen in the picture where it is not there cut off on ACF and PACF so that it can be identified that the SARIMA model used meets the qualifications for forecasting.

In addition to using testing white noise, testing can also be done by looking at ARMA structure, by looking at whether white noise on roots, with AR indicator roots in MA roots, is on the unit circle.

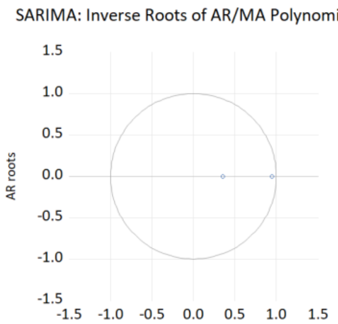


Fig. 14. SARIMA Model Diagnostic Examination (ARMA Structure)

G. Forecasting Result

At this stage, MAPE error value of each model can be seen, where the ARIMA model has a value error of 13.80355 (13%) while the SARIMA model has a value error or MA p-value of

16.54849 (16%) for a value error of both models are still in the good category because they are below 20% value error.

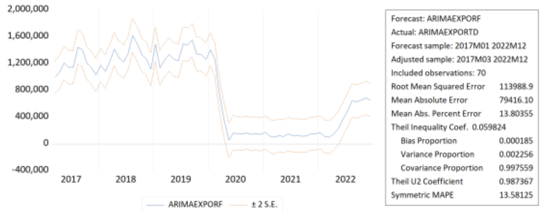


Fig. 15. ARIMA Forecasting Results and Datasets

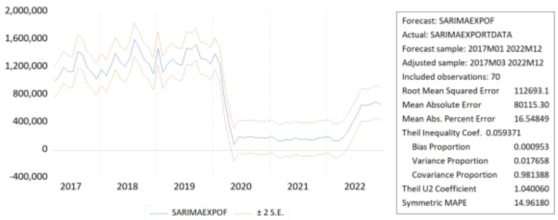


Fig. 16. SARIMA Forecasting Results and Datasets

The following are the forecasting results from each test with the ARIMA and SARIMA models that have been carried out

TABLE I. Forecasting Result.

Period	**Forecasting Result**	
	ARIMA	**SARIMA**
January 2023	969982.2	971724.7
February 2023	991037.7	991658.1
March 2023	995959.4	991905.4
April 2023	996043.3	985534.5
May 2023	994674.6	977151.3
June 2023	992874.9	968369.0
July 2023	990942.7	959742.3
August 2023	988971.4	951452.4
September 2023	986988.4	943549.0
October 2023	985001.9	936036.0
November 2023	983014.4	928901.9
December 2023	98102605	922130.3

Apart from seeing the value error to see which model is better, this research can also be seen with the actual values that occur where these values can be seen from the official website of the Central Statistics Agency (BPS) regarding the number of foreign tourist visits to Indonesia, where on the official website the Central Statistics Agency (BPS) has updated regarding the number of foreign tourist arrivals to Indonesia for the 2023 period from January to March with the following comparison.

TABLE II. Comparison of the Number of Foreign Tourist Visits to Indonesia.

Period	**Forecasting Result Comparison**		
	ARIMA	**SARIMA**	**Current**
January 2023	969982.2	971724.7	735947
February 2023	991037.7	991658.1	701931
March 2023	995959.4	991905.4	809959

IV. Conclusion and Suggestion

A. Conclusion

Based on the results of the research conducted, several conclusions can be drawn.

This research can produce forecasts of the number of foreign tourist visits to Indonesia using the ARIMA model where the results of the best testing model are ARIMA (1, 1, 1). This research can also produce predictions of the number of foreign tourist arrivals to Indonesia using the SARIMA model where the results of the best testing model are SARIMA (1, 1, 1, 1, 1, 1)12. Judging from the value error Both models and parameter tests that have been carried out in this study are included in the good category.

The results of research using the ARIMA model are better than those of the SARIMA model, where the ARIMA model has a value error that is smaller than the value error of the SARIMA model (MAPE) is 13.80355 (13%). As well as value forecast obtained with the ARIMA model is closer to the actual value of the number of foreign tourist arrivals to Indonesia updated by the Central Statistics Agency (BPS). So that it can be said that the ARIMA model is more suitable for predicting the number of foreign tourist visits to Indonesia.

Based on the research results, the number of foreign tourist visits to Indonesia has increased, as seen from the value of the number of foreign tourist visits to Indonesia which has increased compared to the previous period.

B. Suggestion

Further development related to data testing of foreign tourist visits to Indonesia can be carried out, among others.

1. In the next test can use the amount dataset with the number record more so that it can produce better and more accurate forecasting values.
2. In the next test, it is recommended to use a forecasting model other than the ARIMA and SARIMA models such as the model Generalized Seasonal Autoregressive Integrated Moving Average (GSARIMA), Autoregressive Integrated Moving Average with Exogenous (ARIMAX), or others. So that it can be used as a comparison of which model has a better accuracy value in forecasting the number of foreign tourist visits to Indonesia

References

[1] Agus Widarjono. (2013). Ekonometrika: Pengantar dan aplikasinya, Ekonosia, Jakarta: Publisher UPP STIM YKPN Yogyakarta.
[2] Azmiyati, S dan Widya N.T. (2017). Peramalan Jumlah Tandan Buah Segar (TBS) Kelapa Sawit Dengan Metode Fuzzy Time Series Chen dan Algoritma Ruey Chyn Tsaur, 8(1), 36-44.
[3] Badan Pusat Statistika. (2023). Jumlah Kunjungan Wisatawan Mancanegara per bulan Ke Menurut Kebangsaan (Kunjungan), 2023.
[4] Bernard Davis., et al. (2012). Food and Beverage Management. UK: Publisher Butterworth-Heinemann Elsevier Ltd.
[5] D. C. Montgomery, C. L., Jennings dan M. Kulahci. (2015). Introduction to Time Series Analysis and Forecasting, Second Edition ed. New Jersey: Publisher John Wiley & Sons, Inc.
[6] Djakaria M Nur. (2008). Otonomi Daerah Dalam Pengembangan Sektor Pariwisata, 8(1).
[7] Draper, N.R. dan Smith, H. (1992). Analisis Regresi Terapan, Second Edition. Terjemahan Bambang Sumantri. Jakarta: Publisher PT Gramedia Pustaka Utama.
[8] Egsaugm. (2021). Pariwisata Indonesia di Tengah Pandemi, 2021.
[9] Francq, C. and Zakoian, J.M. (2010). GARCH Models: Structure, Statistical Inference and Financial Applications. John Wiley & Sons Ltd. Chichester.
[10] Gerlach, R., & Wang, C. (2016). Forecasting risk via realized GARCH, incorporating the realized range. Quantitative Finance, 16(4), 501–511.

[11] Gitosudarmo, Indriyo dan Mohamad Najmudin. (2003). Anggaran Perusahaan, First Edition. Yogyakarta: Publisher BPFE-Yogyakarta.
[12] H. Moh. Ali Ramdhani. (2014). Manajemen Operasi. Bandung: Publisher CV Pustaka Setia.
[13] Hanke, J. E., dan Reitsch, Arthur G. (1998). Business Forecasting. Sixth Edition. New Jersey: Publisher Prentice Hall.
[14] Heizer, Jay and Barry Render. (2015). Manajemen Operasi Keberlangsungan dan Rantai Pasokan, Eleventh Edition. Jakarta: Publisher Selemba Empat.
[15] Heizer, Jay, & Barry, R. (2019). Operations Management. New Jersey: Publisher Prentice Hall International.
[16] Hendikawati, P. (2015). Peramalan Data Runtun Waktu: Metode dan Aplikasinya dengan Minitab dan Eviews. Semarang: Publisher FMIPA Universitas Negeri Semarang.
[17] Herlin Fransiska, Pepi Novianti, Dian Agustina. (2019). Permodelan Curah Hujan Bulanan Di Kota Bengkulu Dengan Seasonal Autoregressive Integrated Moving Average (SARIMA), 2019(1), 390-395.
[18] Husein Umar. (2013). Metode Penelitian untuk Skripsi dan Tesis Bisnis, Second Edition. Jakarta: Publisher Rajawali Pers
[19] Indriantoro Nur dan Bambang Supono. (2013). Metodologi Penelitian Bisnis Untuk Akuntansi dan Manajemen. Yogyakarta: Publisher FEB Universitas Gajah Mada.
[20] Kotler, Philip. R., et al. (2009). Marketing for Hospitality and Tourism Sixth Edition. International Edition. Pearson
[21] Lewis, C.D. (1982). International And Business Forecasting. London: Publisher Butterworths
[22] Makridakis, Spyros G., et al. (1999). Forecasting Methods and Applications Third Edition. Terjemahan Untung S. Andriyanto dan Abdul Basith. Jakarta: Publisher Erlangga.
[23] McLeod, R. dan Schell, G.P. (2007). Management Information System, Tenth Edition. New Jersey: Publisher Pearson Prentice Hall.
[24] Mokhamad Hilmi Pamungkas. (2016). Estimasi Parameter Model ARIMA Menggunakan Kalman Filter Untuk Peramalan Permintaah Darah (Studi Kasus: UTD PMI Surabaya). Skripsi. Institut Teknologi Sepuluh November Surabaya.
[25] Mulyana. (2004). Analist Data Deret Waktu (Buku Ajar). Jawa Barat: Publisher Universitas Padjajaran
[26] P. Ramos, N. Santos dan R. Rebelo. (2015). Performance of State Space and ARIMA Models for Consumer Retail, Robotics and Computer-Integrated Manufacturing, vol. 34, 151–163.
[27] Rafi Al Amin dan Edwin Agung Wibowo. (2021). Pengaruh Kelengkapan Data, Ketelitian, Kecepatan Terhadap Kepuasan Konsumen Pada PT. Federal International Finance (FIF) Cabang Batam, 1(1), 21-29.
[28] Republika.co.id. (2021). Kunjungan Wisatawan Bergantung Penanganan Covid-19.
[29] Rinaldo Isnawan Prasetyono dan Dyah Anggraini. (2021). Analisis Peramalan Tingkat Kemiskinan Di Indonesia Dengan Model ARIMA, 26(2), 95-110.
[30] Sugiama, A. G. (2011). Ecotourism: Pengembangan Pariwisata berbasis konservasi alam. Bandung: Publisher Guardaya Intimarta
[31] Sugiyono. (2018). Metode Penelitian Kuantitatif. Bandung: Publisher Alfabeta.
[32] Suhartono. (2011). Time Series Forecasting by using Seasonal Autoregressive Integrated Moving Average.
[33] Tadesse, K. B. and Dinka, M. O. (2017). Application of SARIMA model to forecasting monthly flows in Waterval River, South Africa, Journal of Water and Land Development 2017; 35(10-12): 229-236.
[34] Wing Wahyu Winarno. (2011). Analisis Ekonometrika dan Statistika dengan Eviews, Third Edition. Yogyakarta: Publisher Unit Penerbit dan Percetakan (UPP STIM YKPN).
[35] Wiliam, W. S. Wei. (2016). Time Series Analysis, Univariate and Multivariate Methods. Second Edition ed. Pennsylvania: Publisher Pearson Education Inc.
[36] Williams dan Sawyer. (2007). Analisis Tekhnologi Informasi.
[37] Wiwin Sri Rahayu, et al. (2019). Analisis Prediksi Debit Sungai Amprong Dengan Model ARIMA (Autoregressive Integrated Moving Average) Sebagai Dasar Penyusunan Pola Tata Tanam, 10(2), 110-119.
[38] Yoeti, Oka. (1996). Pengantar ilmu pariwisata. Bandung: Publisher Angkasa.