

THE FORECASTING OF MONTHLY INFLATION IN MALANG CITY USING AN AUTOREGRESSIVE INTEGRATED MOVING AVERAGE

Eni Farida¹⁾ Mohamad As'ad^{2*)}

STMIK PPKIA Pradnya Paramita Malang, Jl. Laksda Adi Sucipto 249A Blimbing, Malang, Indonesia^{1,2}

E-mail: eni@stimata.ac.id¹⁾; asad@stimata.ac.id^{2*)}

* : corresponding author

Abstract: Malang is known as a student city since there are a lot of schools and universities that can be found in Malang Indonesia. Malang is also an attractive tourist place with many tourist attractions in the city of Malang. Public transportation in the city of Malang is also very varied, ranging from conventional and based online. Access to the city of Malang is varied, namely trains, buses, and planes. Thus economic growth in the city of Malang is getting better, this can be seen from the economic activity in the increasingly crowded city of Malang. A good economy is usually followed by stable inflation. For this reason, it is necessary to examine how the monthly inflation rate in Malang city. This study aims to forecast inflation in the coming periods using the Autoregressive Integrated Moving Average (ARIMA) model. Secondary monthly inflation data is obtained from BPS Malang. From this research, the ARIMA model (2,0,3) is obtained. The accuracy model is used in this research namely root means square error (RMSE), mean absolute error (MAE), and mean absolute square error (MASE). The accuracy value is RMSE equal 0.2645467, MAE equal 0.2013898, and MASE equal 0.6047399.

Keywords: *Monthly inflation forecasting, BPS Malang city, ARIMA model.*

1. Introduction

Malang is the second-largest city in East Java after Surabaya. Students crowded activities on weekends and weekdays potentially enhance the economic sectors in Malang. According to Bawono (2019), the relationship between inflation and economic growth was negative, which means that when there was economic growth, inflation did not occur (inflation does not rise or stable) and vice versa if there was an economic decline, inflation will rise. The link between economic growth and inflation rates also seems to occur in Malang. This economic activity shows that the city of Malang was experiencing economic growth. The sectors that influence the economic growth (according to the Central Bureau of Statistics or BPS) were food and beverage (restaurants, cafes, stalls, malls, retail, etc.), accommodation (hotels, boarding houses, etc.), electricity, fuel gas, clothing, health, education, recreation, sports, transportation, and financial services. All these sectors affect the level of inflation in the city of Malang. The economy can

develop well on which the inflation is not too high or deflation is not so low. According to BPS, inflation is defined as the tendency for rising prices of goods and services in general to continue. If the price of products and services in the country increases, inflation increases. Rising costs of products and services cause a decrease in the value of money. Thus, inflation can be interpreted as the decreased value of money towards the cost of products and services (bps).

Uncontrolled high Inflation causes the declinable power of purchasing. The effect of the aforementioned issue is the economic wheels cannot work well. On the other hand, economic growth is going to be depressed if inflation is too low or also called deflation. However, stable inflation is expected to be achieved, so, the economy can develop properly (Boediono, 2001). BPS calculates the inflation rate with the variable consumer price index (CPI). The movement of CPI from time to time shows changes in the prices of goods and services consumed by the community. The price determination of products and services was done by BPS by surveying the cost of living in the city. According to BPS, other inflation indicators based on international best prices were wholesale price index (WPI), consumer price indicators (CPI), gross domestic product deflator (GDP), and asset price index (API). Inflation was calculated by BPS.

To predict future inflation, according to the knowledge of researchers, it has never been done and published by BPS. The prediction of future inflation also needs to be done to support the planning of economic activities that have to do with inflation, whether planned by the government or the private sector. From this, it was necessary to predict the inflation for future periods with time series statistical models that were suitable for the conditions of data in the field. In econometrics, inflation models often contain autoregressive factors and also include heteroscedasticity (Gujarati, 2009). According to the aforementioned issues, the researchers conduct this study to predict the value of inflation in the next few periods by using a model that matches the data in the field. The following models that might be used were Autoregressive integrated moving average (ARIMA) model, Autoregressive Conditional Heteroscedasticity (ARCH), or a combination of the two models (HYBRID).

Several studies on inflation prediction have been carried out in several big cities in a country or inflation at the country level. Iqbal and Naveed (2016) have researched quarterly inflation predictions in Pakistan by using the Autoregressive integrated moving average (ARIMA) model. The best model used for forecasting was ARIMA ([4,10], 0,4) which means that inflation is influenced by the data of the 4th period, the data of the 10th period, and also the average residual four (1-4) periods in the past. This study uses quarterly data from 1970 to 2006. Popoola (2017) researched to predict the inflation in Nigeria by using the ARIMA model. The results obtained by the best ARIMA model was ARIMA (0,1,1). It means that the data was not stationary and was stationary by doing differencing. After differencing, the moving average (MA) model was conducted for a period. The data used were monthly inflation data in Nigeria from January 2006 to December 2015. The other researchers used the ARIMA model was also carried out by Abdulrahman (2018) to predict annual inflation that occurred in Sudan. The data used for ARIMA modeling was taken from yearly inflation data from 1970-2018. The results of this study were the ARIMA model (1,2,1) which was used to predict the annual inflation rate from 2017-2026. ARIMA (1,2,1) means that the data used was not stationary so that it was done differencing with lag 2. Then, the data was modeled into ARIMA (1,0,1) in other words the

current data (data that has been differentiated) depends on the data one the previous period and the moving average of the previous period. From the three aforementioned studies, it can be seen that the inflation data used did not contain heteroscedasticity or the three ARIMA models above, the error had a white noise condition.

Besides the ARIMA model, other researchers want to compare Holt's exponential smoothing model. This study used monthly inflation data in Zambia from May 2010 to May 2014. The results of this study were the ARIMA model ([12], 1,0) that best matches the data analyzed. This study showed that inflation data did not contain volatility, so the model did not contain heteroscedasticity (Jere & Sianga, 2016). Another study that used the ARIMA model in combination with General Autoregressive Conditional Heteroscedasticity (GARCH) was Uwilingiyimana et al. (2016). These researchers predicted inflation rates in Kenya with monthly inflation data taken from 2000 to 2014. The combination of the ARIMA-GARCH model was a hybrid model. From the research results obtained by ARIMA (1,1,12) and GARCH (1,2). Inflation data in this study was the ARIMA model and the residual was a GARCH model. Other researchers who used the GARCH and ARIMA models were Osarumwense and Waziri (2016). This researcher performed monthly inflation forecasting in Nigeria. This study used monthly inflation data from January 1995 to December 2011. The final results of this study got the GARCH (1,0) and ARIMA (1,0,0) models. Furthermore, this model was made to predict inflation data that will occur from January 2012 to December 2013.

2. Research Method

In this study, we have used a time series forecasting model that was suitable for the conditions of the field data and which had a high accuracy value. To choose which model that can be used in stages of scientific analysis as had been done by other researchers. Data plots would be conducted to find out whether the data was stationary or not, as done by Iqbal and Naveed (2016), Popoola (2017), Abdulrahman (2018), Jere & Sianga (2016), Uw Surroundiyimana (2016), and Osarumwense and Waziri (2016). The next stage will be plotted by using Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). From this initial identification, the monthly inflation data of Malang city has an autoregressive (AR) model only or contains (AR) and heteroscedasticity can be estimated.

This research aims to predict the monthly inflation in Malang for several periods in the future. Inflation data is secondary data obtained from the Central Bureau of Statistics (BPS) of Malang City. The data used in this research is monthly inflation from January 2015 to June 2019. Before performed data analysis, it is plotted to graphically show whether the data contains trend, seasonal, trend, and seasonal elements or contains volatility. If it seems to include a trend element, a data stationarity test is performed with the Dicky Fuller Test. If the data includes the trend is not stationary, the differencing process will be carried out. If it does not contain directions or stationary, it will be identified earlier. Initial identification by plotting ACF and PACF. ACF to determine the demand of the MA model, and PACF to determine the order of the AR, (Wei, 2006).

After the AR (p) and MA (q), orders are obtained, the next step is to estimate the ARMA model parameters (p, q). At this stage, the ARMA model coefficient test (p, q) is performed with

the t-test. If the parameters tested are not significant, then the model is changed so that all the coefficients are significant (As'ad, 2012).

The next step after all significant model coefficients is to test the error of the model must be white noise. The model error is white noise, which means the fault is not autocorrelated and is normally distributed. To check the non-autocorrelation error, the Ljung Box test is used, whereas to check for normality it uses the Shapiro-Wilk test. If one of the requirements cannot be reached, the error is not white noise. One of the ways to overcome the incident is by analyzing whether the error might occur heteroskedasticity or also not stationary variance. If heteroscedasticity occurs, it can be modeled again with the ARCH or GARCH models. If the variance is not stationary, data must be transformed by Box-Cox. After the transformation of Box-Cox and the data are stationary in the variance, the next step is using the ARIMA model (Dritsaki, 2018).

After testing the white noise condition of the residual for the ARIMA model, the next step is to choose the best ARIMA model. According to Wei (2006), to select the best ARIMA model by looking at the lowest Akaike Information Criteria (AIC) value. In addition to the minimum AIC amount, the best ARIMA model is also seen from the lowest amount of the accuracy for the model, including; root mean square error (RMSE), mean absolute error (MAE), and mean absolute square error (MASE). The best ARIMA models use the parsimony principle. The principle of parsimony is essentially choosing a simple model. Based on the three determinations, the best ARIMA model can be selected.

In the ARIMA model, the residual is not white noise caused by autocorrelation or heteroscedasticity. Residual from such ARIMA models must be done by using ARCH or GARCH models. To test the heteroscedasticity effects, a squared residual test with the Lagrange Multiplier test or the ARCH-LM Test (Bollerslev, 1994) is performed. If there is an ARCH or GARCH effect on the residuals, then the ARCH or GARCH models are performed. Furthermore, from the best ARIMA models and ARCH / GARCH models, they are combined into Hybrid models (Dritsaki, 2018).

3. Results and Discussion

This research uses secondary monthly inflation data obtained from BPS Malang. The data used was monthly inflation from January 2015 to June 2019. Furthermore, the data is plotted to find a picture of the data (figure 1).

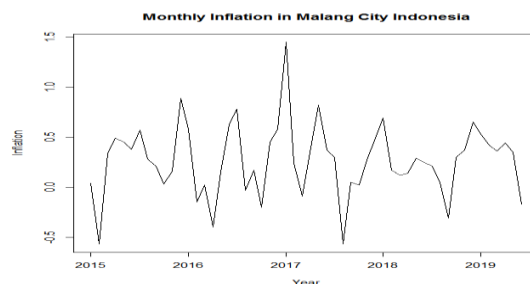


Figure 1.

The plot of monthly inflation data for the Malang city

At first glance from Figure 1 does not indicate a trend. A database test is performed by using the Dicky Fuller test to ensure that the result is correct. The test results are as follows (tabel 1) :

Tabel 1.

Augmented Dickey-Fuller (ADF) Test

Lag	no drift no trend		with drift no trend		with drift and trend	
	ADF	p-value	ADF	p-value	ADF	p-value
1	0 -3.99	0.0100	0 -5.45	0.01	0 -5.37	0.01
2	1 -3.54	0.0100	1 -5.94	0.01	1 -5.84	0.01
3	2 -3.23	0.0100	2 -6.32	0.01	2 -6.24	0.01
4	3 -1.99	0.0466	3 -4.36	0.01	3 -4.31	0.01

Note: in fact, p.value = 0.01 means p.value <= 0.01

The p-value <= 0.01 when compared to α (5% = 0.05) is smaller, that means accept H1 (stationary data). Furthermore, plotting ACF and PACF are done in this test. The results are in figure 2 and figure 3 as follows:

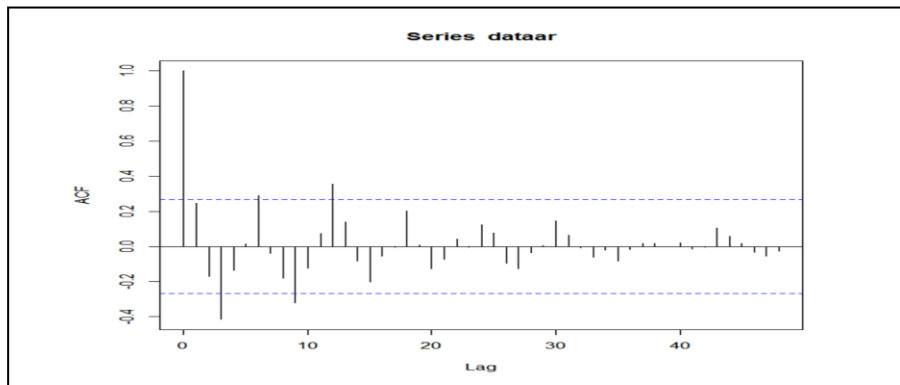


Figure 2.

ACF plot of monthly inflation data in Malang

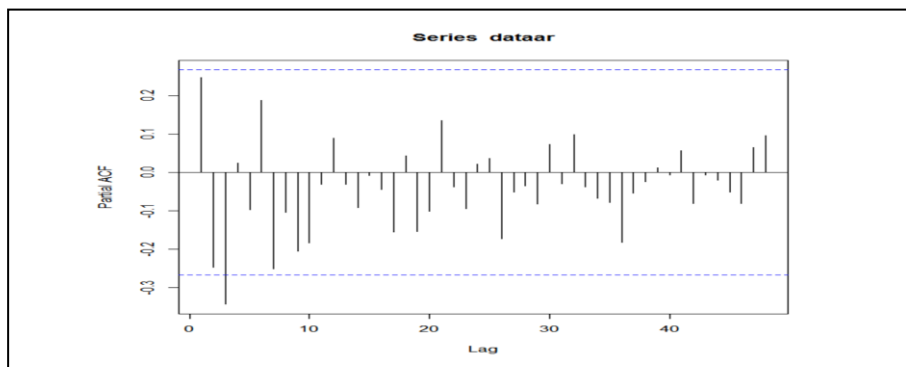


Figure 3.

PACF plot of monthly inflation data in Malang

The tentative model from the ACF and PACF plots is ARIMA (3,0,3). The results of the parameter estimation for ARIMA (3,0,3) are in table 3 and table 4 as follows:

Tabel 3.

Parameter Estimation for ARIMA (3,0,3)

Call: arima(x = dataar, order = c(3, 0, 3))

Coefficients:

	ar1	ar2	ar3	ma1	ma2	ma3
intercept	-0.1516	0.4664	-0.1957	0.3489	-0.8207	-0.5282
0.2752						
s.e.	0.2887	0.1555	0.2782	0.3044	0.1281	0.3364
0.0074						

sigma^2 estimated as 0.07842: log likelihood = -9.74, aic = 35.49

The results of the t-test (it was approached with the Z test) are as follows:

Tabel 4.

Z-test Parameter for ARIMA (3,0,3)

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z)
ar1	-0.1515762	0.2887106	-0.5250	0.599576
ar2	0.4663857	0.1554602	3.0000	0.002699 **
ar3	-0.1956526	0.2782382	-0.7032	0.481941
ma1	0.3489444	0.3043750	1.1464	0.251618
ma2	-0.8207154	0.1280930	-6.4072	1.482e-10 ***
ma3	-0.5282278	0.3363554	-1.5704	0.116312
intercept	0.2752124	0.0073885	37.2489	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

It appears that ar3 is not significant ($0.481941 > 0.05 (\alpha)$) which means this model is not suitable. Then ar3 is eliminated and the model becomes ARIMA (2,0,3). The results of the parameter estimation for ARIMA (2,0,3) are in table 5 and table 6 as follows:

Tabel 5.

Parameter Estimation for ARIMA (2,0,3)

Call: arima(x = dataar, order = c(2, 0, 3))

Coefficients:

	ar1	ar2	ma1	ma2	ma3
intercept	-0.4112	0.3731	0.7608	-0.7608	-0.9999

0.2749
 s.e. 0.1313 0.1360 0.1113 0.1018 0.1195
 0.0076

sigma² estimated as 0.06998: log likelihood = -9.26, aic = 32.52

The results of the t-test (it was approached with the Z test) are as follows:

Tabel 6.
 Z-test Parameter for ARIMA (2,0,3)

z test of coefficients:				
	Estimate	Std. Error	z value	Pr(> z)
ar1	-0.411174	0.131305	-3.1314	0.001740 **
ar2	0.373115	0.135996	2.7436	0.006077 **
ma1	0.760847	0.111334	6.8339	8.264e-12 ***
ma2	-0.760846	0.101832	-7.4716	7.923e-14 ***
ma3	-0.999897	0.119490	-8.3681	< 2.2e-16 ***
intercept	0.274948	0.007628	36.0448	< 2.2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

It appears that the parameters ar1, ar2, ma1, ma2, ma3, and the intercept are all significant (there are asterisks or p-values <0.05). Then, the data is compared by using the ARIMA model (2,0,2). The results are in table 7 and table 8 as follows:

Tabel 7.
 Parameter Estimation for ARIMA (2,0,2)

Call: arima(x = dataar, order = c(2, 0, 2))					
Coefficients:					
	ar1	ar2	ma1	ma2	intercept
	0.9082	-0.8413	-0.7061	0.5560	0.2718
s.e.	0.1386	0.1381	0.2395	0.2348	0.0390

sigma² estimated as 0.09631: log likelihood = -13.78, aic = 39.56

Next test the coefficient significance:

Tabel 8.
 Z-test Parameter for ARIMA (2,0,2)

z test of coefficients:				
	Estimate	Std. Error	z value	Pr(> z)
ar1	0.908191	0.138590	6.5531	5.637e-11 ***
ar2	-0.841282	0.138084	-6.0926	1.111e-09 ***

ma1	-0.706146	0.239509	-2.9483	0.003195 **
ma2	0.556004	0.234760	2.3684	0.017866 *
intercept	0.271843	0.038952	6.9790	2.973e-12 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

It appears that the parameters ar1, ar2, ma1, ma2, and intercept are all significant (there are asterisks or p-values <0.05).

The ARIMA (2,0,3) and ARIMA (2,0,2) models all have significant parameters, but, the researcher will take the best model. The selection of the best model based on the Akaike Information Criteria (AIC) in table 3. Table 5, Table 7, obtained the best ARIMA model (2,0,3) because it has the smallest AIC value (32.52 <35.49 <39.56). In addition to the AIC value for selecting the best ARIMA model, forecasting accuracy values such as; RMSE, MAE, and MASE are used. Comparison of accuracy values can be seen in the following table 9:

Tabel 9.
The accuracy value of the three ARIMA models

Model / accuracy	RMSE	MAE	MASE
ARIMA(3,0,3)	0.2800393	0.2023801	0.6077137
ARIMA(2,0,3)	0.2645467	0.2013898	0.6047399
ARIMA(2,0,2)	0.3103419	0.233071	0.6998733

By comparing the AIC value, forecasting accuracy, and also the parsimony model in selecting the best model is ARIMA (2,0,3).

Next will be examined the assumption of white noise from the residual of the ARIMA model (2,0,3).

The first is the assumption of autocorrelation. With the Ljung-Box test the rest of the ARIMA models (2,0,3) are obtained in table 10 as follows:

Tabel 10.
Ljung-Box test for residual of the ARIMA model (2,0,3).

Box-Ljung test data: residuals(model1)
X-squared = 8.2006, df = 12, p-value = 0.7693

Chi-square value with the Ljung-Box test of 8,2006 with a lag of 12 obtained a p-value of 0.7693 > 0.05 which means that the residual is random or random without autocorrelation.

Then the second residual assumption test is the residual normality test with the Shapiro-Wilk test. The results are in table 11 as follows:

Tabel 11.
Shapiro-Wilk test for residual of the ARIMA model (2,0,3).

Shapiro-Wilk normality test : residuals(model1)
W = 0.98006, p-value = 0.5029

The Shapiro Wilk test statistic value of 0.98006 with p-value = 0.5029 > 0.05 (\square) means that the H_0 test data is normally distributed. With this model residual test, the white noise traits are fulfilled; they are free autocorrelation and normal distribution. Thus the best model for forecasting monthly inflation in Malang in this study is ARIMA (2,0,3).

Such econometric models usually contain high volatility, or the remainder of the model is still heteroscedasticity, for this reason, the ARCH / GARCH effect test will be performed both on monthly inflation data and on the ARIMA (2,0,3) model residual data. This test uses the Lagrange Multiplier test or the ARCH-LM Test.

In the monthly inflation data (dataar) test for the ARCH / GARCH model, the ARCH LM-test results are obtained in table 12 as follows:

Tabel 12.

ARCH LM-test for data Monthly Inflation		
ARCH LM-test; Null hypothesis: no ARCH effects		
data: dataar		
Chi-squared = 9.5414,	df = 12,	p-value = 0.6561

From the ARCH LM-test results for monthly inflation data, the Chi-Squared value = 9.5414 with a p-value of 0.6561 ($P\text{-value} > \alpha$ or $0.6561 > 0.05$), which means that H_0 is accepted (the data does not contain the ARCH model). Whereas in the residual of the ARIMA model (2,0,3) data test for the ARCH / GARCH effect, the ARCH LM-test results are obtained in table 13 as follows:

Tabel 13.

ARCH LM-test for data residual of the ARIMA (2,0,3) model.		
ARCH LM-test; Null hypothesis: no ARCH effects		
data: residuals(model1)		
Chi-squared = 9.6425,	df = 12,	p-value = 0.6473

The ARCH LM-test results for the residual of the ARIMA model (2,0,3) obtained Chi-Squared value = 9.6425 with a p-value of 0.6473 ($P\text{-value} > \alpha$ or $0.6473 > 0.05$), this means that H_0 (the rest does not contain effects ARCH). From the series of tests above, the best model for predicting monthly inflation in Malang is the ARIMA model (2,0,3). The data plot of the inflation month and forecast results from January 2015 to June 2019 are in figure 4 as follows:

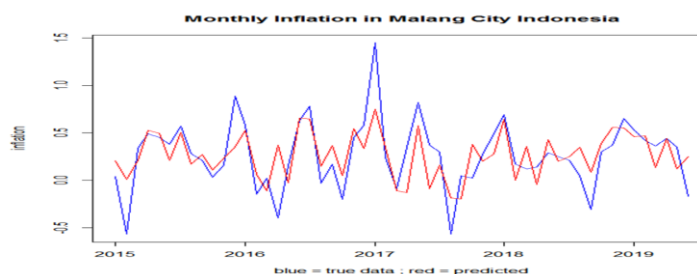


Figure 4.

Plots of monthly inflation data and forecast results

Furthermore, the next three months' forecasting period for July, August, and September 2019 can be seen in table 14.

Table 14.

Inflation forecast for July-September 2019

Month	July	August	September
Forecasting	-0.007629137	0.317746401	0.559477606

From the ARIMA model (2,0,3) it can be interpreted that inflation occurring in the current month is determined by inflation last month and also two months ago and then added a moving average from the residual of the autoregressive model three periods ago (MA (3)). From table 13 we can read the inflation forecast value that occurred in July amounted to -0.007629137, which slightly increased from June inflation of -0.17. In August there was an increase in inflation which reached point 0.317746401, this was possible because it coincided with the Idul Adha holiday and the commemoration of Augustus. In September, there is an increase in inflation at the point 0.559477606. If seen from the historical graph in Figure 4, the possibility of inflation in September is the peak and will experience deflation in the following month. To find out the inflation in October, the researchers do a remodel by adding the latest data in July and also August. This remodeling is considered a significant action since the latest data can provide even more accurate information in building models for forecasting in the next one or two months.

4. Conclusion

From this study, it can be concluded that the best ARIMA model for predicting monthly inflation in Malang is ARIMA (2,0,3). The ARIMA model (2,0,3) means that the current month inflation in Malang is influenced by the two past monthly inflation periods. It is also influenced by the moving average error (the difference between the actual data and its forecast value) three periods ago. Accuracy of forecasting ARIMA (2,0,3) model for root means square error (RMSE) is 0.2645467, mean absolute error (MAE) is 0.2013898 and mean absolute square error (MASE) is 0.6047399. To maintain the accuracy of forecasting, the latest data should be inputted so that the model becomes renewable and also forecasting the future period is not too long, maybe one to three periods only. The forecast value for the next three periods from July to September seems to be inflation, so this prediction needs to be considered as one of the references in planning for the next three months that has to do with inflation in the city of Malang.

Reference

- Abdurrahman. BMA., Ahmed AYA., & Abdellah AEY., (2018), Forecasting of Sudan Inflation Rates using ARIMA Model, *International Journal of Economics and Financial*, Vol. 8 no. 3, p.17-22.
- As'ad, M. (2012), Finding the Best ARIMA Model to Forecast Daily Peak Electricity Demand, Proceedings of the Fifth Annual ASEARC Conference - Looking to the future - Programme and Proceedings, 2 - 3 February 2012, University of Wollongong, Australia.
- Bawono, A. (2019), Factors Influencing The Inflation Of Indonesia In Islamic Perspectives, *Jurnal Ilmiah Ekonomi Islam*, Vol. 5 No. 2 p.81-88

- Boediono (2001), *Ekonomi Makro*, BPFE, Yogyakarta.
- Bollerslev, T., Engle, RF., & Nelson DB. (1994), *Hand Book Econometrics-Chapter 49*, Elsevier Science B.V. All rights reserved.
- Distsarki, C., (2018), The Performance of Hybrid ARIMA-GARCH Modeling and Forecasting Oil Price, *International Journal of Energy Economics and Policy*, Vol 8(3), p.14-21.
- Gujarati, D.N. 2009. *Basics Econometrics*, McGraw-Hill Irwin, Boston England.
- <https://www.bps.go.id/subject/3/inflasi.html#subjekViewTab1>. Accessed on 11 January 2021.
- Iqbal. M., & Naveed. A., (2016), Forecasting Inflation: Autoregressive Integrated Moving Average Model, *European Scientific Journal*, vol. 12, no. 1.
- Jere, S., & Sianga, M., (2016), Forecasting Inflation Rate of Zambia Using Holt's Exponential Smoothing, *Open Journal of Statistics*, vol. 6, p.363-372
- Osarumwense, OI., & Waziri, EI., (2016). Modeling Monthly Inflation Rate Volatility, using Generalized Autoregressive Conditionally Heteroscedastic (GARCH) models: Evidence from Nigeria, *Australian Journal of Basic and Applied Sciences*, Vol.7(7): p.991-998.
- Popoola, OP., Ayanrinde AW., Rafiu AA., & Oddusina MT., (2017), Time Series Analysis to Model and Forecasting Inflation Rate in Nigeria, *Anale. Seria Informatica*. Vol. XV fasc. 1.
- Uwilingiyimana, C., Mung'tu J., & Harerimana, JDD., (2016), Forecasting Inflation in Kenya Using Arima - Garch Models, *International Journal of Management and Commerce Innovations*, Vol. 3, Issue 2, p. 15-27.
- Wei, W.W.S., 1990. *Time-series Analysis: Univariate and Multivariate Methods*. Addison-Wesley Publishing Co., USA.