Inflation Forecasting In East Java Using Autoregressive Integrated Moving Average Method

[1*]Imam Mukhtar Syarifudin, [2]Ardhi Sanwidi, [3]M. Nur Haqqul Qomarudin
[1,2,3]Matematika, Universitas Nahdlatul Ulama Blitar
E-mail: [1] mukhtar.syariefudin@gmail.com, [2]ardhisanwidi@gmail.com, [3]haqquludin@gmail.com
*Corresponding Author

Abstract— Inflation is an economic event that often occurs even though it is not wanted. Based on data from Badan Pusat Statistik in 2015-2020 inflation in East Java was 3.08%, 2.74%, 4.04%, 2.86%, 2.12%, 1.44%. From these data, it can be seen that inflation data is fluctuating. Therefore it is necessary to control inflation because high and unstable inflation can have a negative impact on the socio-economic conditions of the community. In addition, it also makes it difficult for the government to determine future policies. Seeing the importance of controlling inflation, it is necessary to study to predict the inflation rate in the future. One of the studies/methods to predict that is often used is the Autoregressive Integrated Moving Average (ARIMA) method or also known as the Box-Jenkins method. The ARIMA method is a method that is easy to use because it is flexible in following existing data patterns and has high accuracy and tends to have a small error value because of the detailed process. From the analysis results, the best ARIMA (p,d,q) model is the ARIMA model (2,1,1) with an AIC value of 76.77. The results of forecasting with the ARIMA model (2,1,1) respectively are 0.2593698, 0.1892990, 0.1340639, 0.1368309, 0.1572021, 0.1642381, 0.1598897, 0.1557251, 0.1556074, 0.1570151, 0.1576092, 0.1573511, 0.1570423, and 0.1570111.

Keywords— Forecasting; Inflation; Arima

Corresponding Author:
Imam Mukhtar Syarifudin,
Matematika,
Universitas Nahdlatul Ulama Blitar,
Email: mukhtar.syariefudin@gmail.com

I. INTRODUCTION

Inflation is an economic event that often occurs even though it is not wanted. Inflation tends to occur in developing countries, especially Indonesia. Inflation should not only be the responsibility of the state, but also the provincial government, including the province of East Java. Based on data from the Central Statistics Agency (BPS), East Java inflation in 2015-2020 was 3.08%, 2.74%, 4.04%, 2.86%, 2.12%, 1.44%. From these data, it can be seen that inflation data is fluctuating. The biggest inflation occurred in 2017 up from the previous year, but decreased in the following year.

Controlling the inflation rate is very important to do to stabilize prices in the community. According to Sutawijaya (2012) one of the reasons for the government to control inflation is because inflation can worsen the distribution of income (become unbalanced). Another consideration is that high and unstable inflation can have a negative impact on the socio-economic conditions of the community. First, high inflation will cause people's real income to decline. Second, unstable inflation will create uncertainty for economic actors in making decisions. This will have an impact on decreasing economic growth (www.bi.go.id). Inflation instability from time to time will make it difficult for the government of East Java to determine policy.

Seeing the importance of controlling inflation, it is necessary to study to predict the inflation rate in the future. According to Hartati (2017), to get an information on inflation in the future, it is not enough with the data that already exists at this time. Inflation data in the past is also very necessary. From this data, we can create a method that is able to describe the pattern and nature of the changes in inflation. One of the forecasting methods that is often used is the Autoregressive Integrated Moving Average (ARIMA) method or also known as the Box-Jenkins method. ARIMA is a statistic that is suitable for predicting a number of variables quickly, simply, cheaply, and accurately because it only requires variable data to be forecasted. Noviyanto et al (2015) say that the advantages of this method are that it can accept all types of data models, although in the process it must be stationary first. Stationarity means that there is no drastic change in the data.

Hutasuhut, et al (2014) said that the ARIMA method is a flexible method because it follows existing data patterns and has high accuracy and tends to have a small error value because of the detailed process. In line with that statement, previous research with the title "Perbandingan Metode Peramalan Inflasi: Ordinary Least Square (OLS), Exponential Smoothing dan ARIMA" by Friska Zehan and Iman Sugena it is found that ARIMA is the best method for forecasting the inflation rate because it has the smallest Mean Absolute Error (MAE), Mean Square Error (MSE), and Mean Absolute Precentage Error (MAPE). In addition, another study "Perbandingan Metode Triple Exponential Smoothing Dan Metode Seasonal Arima Untuk Peramalan Inflasi Di Kota Tanjung

Pandan" by Putri Choirunisa and Kariyam Kariyam also obtained that ARIMA forecasting is the best method based on the results of Mean Square Eror (MSE) dan Root Mean Square Eror (RMSE).

Based on the description above and the absence of a study on inflation forecasting in East Java using the ARIMA method, this study intends to predict monthly inflation in East Java until the end of 2022 using the ARIMA forecasting method.

II. RESEARCH METHOD

1. DATA SOURCE

The data obtained is sourced from the official website of the Kediri Regency BPS at www.kedirikab.bps.go.id. The data used is data from January 2014 to October 2021 with consideration of the updated method of calculating inflation

2. ARIMA

Model *Autoregressive Integrated Moving Average* (ARIMA) used based on the assumption that the time series data used must be stationary. In overcoming the non-stationary data, a differencing process is carried out so that the data becomes stationary. Since of the model *Autoregressive* (AR), *Moving Average* (MA), *Autoregressive Moving Average* (ARMA) unable to explain the meaning of differencing, a mixture called *Autoregressive Integrated Moving Average* (ARIMA) or ARIMA (p,d,q) is used so that it becomes more effective in explaining the differencing process. The general form of the ARIMA model is:

$$\Phi_p (B) D^d Z_t = \mu + \theta_q (B) a_t$$

where,

Φ_p=parameter coefficient *Autoregressive* p

θ_q=parameter coefficient *Moving Average* q

B= operator backshift

D= differencing

μ= constant

a_t= residual at t

p= *Autoregressive* degree

d= differencing degree

q= *Moving Average* degree

2. Differencing

The concept of differencing serves to overcome modeling problems if there is a non-stationary process. What is meant by differencing is to calculate changes or differences in the value of observations. Diferencing order 1 of a time series data Xt is defined by the following equation:

$$\Delta Z_t = (1-B)\,Z_t = Z_t - Z_{(t-1)}$$

where

$Z_{(t)}$ = value variable Z at t

$Z_{(t-1)}$ = value variable Z at t-1

B = backward shift

The differencing value obtained is checked based on the p-value of . If it is not stationary, then differencing is done again. The data is stationary if the p-value < α.

3. Augmented Dickey-Fuller Test

According to Juanda and Junaidi (2012) the model on the ADF test for testing the serial correlation between residuals can be expressed in the general form of an *Autoregressive* process as follows:

$$\Delta Z_t = \beta_1 + \beta_2 t + \delta Z_{t-1} + \alpha_1 \Delta Z_{t-1} + \alpha_2 \Delta Z_{t-2} + \cdots \alpha_1 \Delta Z_{t-2} + a_t$$

where ε_t is white noise process is normally distributed $N(0,\sigma^2)$ and $\Delta Z_{t-1} = (Z_{t-1} - Z_{t-2})$

hypothesis:

$H_0 : \delta = 0$ (not stasioner)

$H_1 : \delta < 0$ (stasioner)

Statistic test

$$t = \frac{\hat{\delta}}{SE(\hat{\delta})}$$

criteria test:

H_0 rejected if ADF < critical value ADF or p-value < α.

4. Autocorrelation Function (ACF)

The autocorrelation function is used to see if there is a *Moving Average* (MA) of a time series, which in the ARIMA equation is represented by the quantity q. If there is an MA trait, q is generally 1 or 2, it is very rare to find a model with a q value of more than 2 (Rohmah, 2018). The ACF value at lag-k can be formulated as follows:

$$\rho_k = \frac{\sum_{t=1}^{N-k}(Z_t - \bar{Z})\,(Z_{t+k} - \bar{Z})}{\sum_{t=1}^{N}(Z_t - \bar{Z})^2}$$

with ρ_k = autokorelasi function

Z_t = time data at t

$\bar{Z}$ = average data at t

Z_{t+k} = time data at $t + k$

5. Partial Autocorrelation function (PACF)

PACF is generally used to identify of AR (*Autoregressive*) which are denoted by the magnitude of p (Rohmah, 2018). The PACF value at lag-k can be determined as follows:

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,j}\rho_j}$$

with

ϕ_{kk} = autokorelasi parsial function

ρ_k = autokorelasi function

6. Uji Asumsi *white noise*.

Diagnostic checking is divided into two parts, namely parameter significance test and model suitability test. Soejoeti (in Pitaloka 2019) explained that diagnosis or verification is intended to check whether the estimated model matches the existing data. To see the randomness of the error value, a test is carried out on the value of the autocorrelation coefficient of the error, using the Q-Ljung-Box test as follows:

$$Q = n(n + 2) \sum_{k=1}^{m} (n - k)^{-1} \hat{\rho}_k^2$$

Q distribution $\chi^2(\alpha, K - m)$

with n = datas

$\hat{\rho}_k$ = autokorelasi error lag k

$$m = p + q$$

Significance level: α

Hyphotesis:

$H_0 : \rho_1 = \rho_2 = \cdots = \rho_k = 0$

H_1 : at least one $\rho_k \neq 0$

H_0 reject if $Q > \chi^2(\alpha, K - m)$ or $p - value < \alpha$

Furthermore, the residual normality test was carried out to determine whether the residues were normally distributed or not. Testing can be done with normal probability plot graph analysis.

7. Akaike's Information Criterion (AIC)

The main requirement of a model is said to be good if it has the smallest AIC value (Simbolon, 2013). The AIC equation in the model selection is as follows:

$$AIC = log\hat{\sigma}^2 + \frac{2k}{n}$$

where :

$log\hat{\sigma}^2$ = ukuran likelihood

k = number of parameters n = observations

III. RESULT AND DISCUSSION

The earliest process in the analysis is to perform a stationarity test which is the main requirement of ARIMA analysis. The following is a graph of the East Java inflation time-series from January 2014 – July 2021:

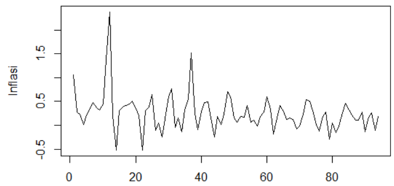

Figure 1 Inflation graph East Java

From Figure 1, the time-series graph shows that the data moves so fluctuating that it cannot be believed that the data is stationary, it is necessary to do an ADF test to make sure the data is stationary or not. The following are the results of the ADF test on East Java inflation time-series data:

```
        Augmented Dickey-Fuller Test

data:  data_inflasi
Dickey-Fuller = -3.3207, Lag order = 12, p-value = 0.0725
alternative hypothesis: stationary
```

Figure 2 ADF test

Through Figure 2, the ADF test shows that the p-value = 0.0725. Based on the ADF test with p-value > α = 5% it means that the hypothesis H0 is accepted where there is a unit root so that the data is not stationary. The non-stationary data needs to be stationary by doing differencing. The following is a graph of the first differencing process on inflation data:

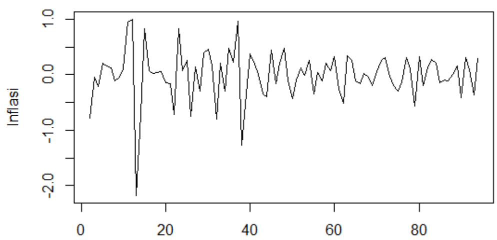

Figure 3 First differencing result

From Figure 3, it can be seen that the data is so fluctuating again that it is not sure believed that the data is stationary, so the ADF test is carried out again to ensure whether the results of the first differencing are stationary or not. The following are the results of the ADF test from the first differencing:

```
        Augmented Dickey-Fuller Test

data:  data_inflasi_d1
Dickey-Fuller = -4.4335, Lag order = 12, p-value = 0.01
alternative hypothesis: stationary
```

Figure 4 ADF test

From Figure 4, the ADF test above shows that the p-value = 0.01. With p-value < α where 0.01 < 0.05. This means that based on the initial hypothesis, H0 is rejected where there is no unit root so it can be said that the data is stationary.

The next process is forming the ARIMA model. The ARIMA model is formed based on the degree of Autoregressive/AR (p), Integrated (d), and Moving Average/MA (q). The degree of d has been known through the differencing process so as to determine the degree/value of p from AR and the degree/value of q from MA. The degree/value of the MA can be determined through the ACF plot, while the degree/value of the AR is obtained through the PACF plot.

The following is a plot of the results of the ACF analysis based on time-series inflation data for East Java which has been stationary/differencing:

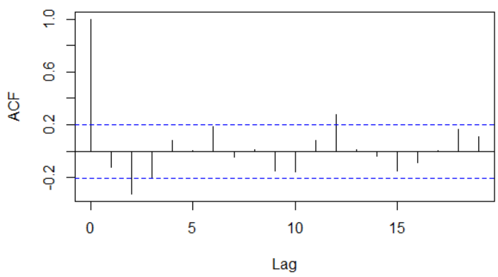

Figure 5 ACF

From Figure 5, it can be seen that the data lag that lies outside the confidence interval is in the 0th lag and cuts again on the 2nd lag. In the ACF plot above, the significant data lag after lag 0 is the 2 nd lag. In the 3rd lag, it can be seen that the line does not come out of the confidence interval, which means that the lag cannot be included in the parameter estimation. So the possible p value in AR is 0 to 2.

Next is a plot of the results of the PACF analysis of East Java inflation time-series data that has been stationary/differencing:

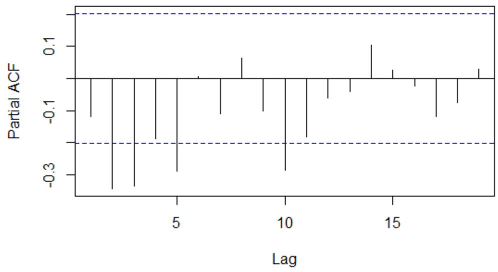

Figure 6 PACF

In Figure 6 above, it can be seen that the lag line that lies outside the confidence interval is in the 2nd lag and the 3rd lag and the fourth lag are already within the confidence interval, so the possible p-value in AR is 2 or 3.

Furthermore, an examination is carried out to see whether the model is in accordance with the data and has a normal distribution. The model is said to be good and has fulfilled the diagnostic test if the p-value > α. This examination was carried out using the L-Jung Box test method. The following are the results of the diagnostic examination of each ARIMA model (p, d, q) that have been obtained:

Tabel 1 L-Jung Box Test

ARIMA (p,d,q)	p-value	Conclusion
ARIMA (2,1,0)	3,548e-0,5	Reject H_0
ARIMA (2,1,1)	0,1297	Accept H_0
ARIMA (2,1,2)	0,3023	Accept H_0
ARIMA (3,1,0)	0,005476	Reject H_0
ARIMA (3,1,1)	0,3369	Accept H_0
ARIMA (3,1,2)	0,432	Accept H_0

From table 1 above, the results of the analysis show that there are 4 out of 6 ARIMA model estimates (p,d,q) that suitable the hypothesis and can be used for the forecasting process. Next is the residual normality test to prove whether the data are normally distributed. The normality test uses an ordinary histogram where the data is normally distributed. The models tested are 4 models that have passed the L-jung Box Test. The following is the normality test on the ARIMA model (p,d,q):

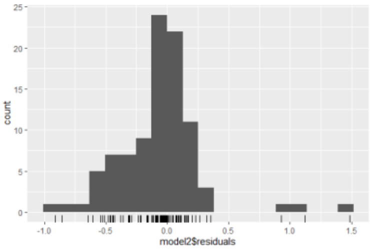

Figure 7 Normality test ARIMA (2,1,1)

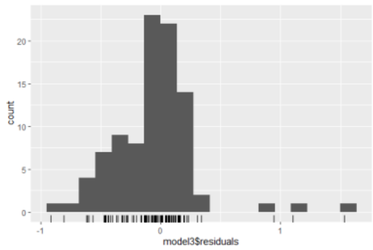

Figure 8 Normality test ARIMA (2,1,2)

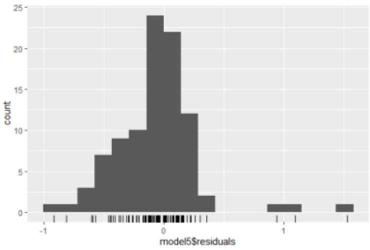

Figure 9 Normality test ARIMA (3,1,1)

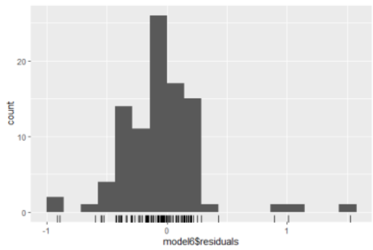

Figure 10 Normality test ARIMA (3,1,2)

From the results of the normality test on the ARIMA model estimation (p, d, q) above, it can be seen that the four models have a similar distribution of data. Through the diagnostic examination, the best model for forecasting has not been obtained, so another method is used to select the best model, namely through Akaike's Information Criterion (AIC). To see the AIC value, it can be seen through the output as follows:

Table 2 AIC value

ARIMA (p,d,q)	AIC value
ARIMA (2,1,1)	76,73
ARIMA (2,1,2)	77,54
ARIMA (3,1,1)	77,4
ARIMA (3,1,2)	77,58

The main requirement of a model is said to be good if it has the smallest AIC value. In the output above, it can be seen that the smallest AIC value is 76.77 which is in ARIMA (2,1,1). Next is the forecasting process using the ARIMA (2,1,1) model. Because the latest data used is inflation in October 2021, the forecast to be made is inflation forecasting for the next 14 months until the end of 2022. Thus, the inflation forecast for 2022 is as follows:

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
95	0.2593698	-0.1836566	0.7023963	-0.4181806	0.9369203
96	0.1892990	-0.2970494	0.6756475	-0.5545067	0.9331048
97	0.1340639	-0.3547693	0.6228972	-0.6135420	0.8816698
98	0.1368309	-0.3560149	0.6296767	-0.6169117	0.8905735
99	0.1572021	-0.3360160	0.6504202	-0.5971099	0.9115141
100	0.1642381	-0.3321567	0.6606329	-0.5949322	0.9234084
101	0.1598897	-0.3382756	0.6580550	-0.6019884	0.9217678
102	0.1557251	-0.3429959	0.6544462	-0.6070029	0.9184532
103	0.1556074	-0.3435651	0.6547798	-0.6078110	0.9190258
104	0.1570151	-0.3428735	0.6569037	-0.6074985	0.9215288
105	0.1576092	-0.3431624	0.6583809	-0.6082549	0.9234734
106	0.1573511	-0.3442576	0.6589598	-0.6097933	0.9244954
107	0.1570423	-0.3453268	0.6594114	-0.6112649	0.9253495
108	0.1570111	-0.3461029	0.6601251	-0.6124354	0.9264576

Figure 11 Forecasting ARIMA (2,1,1)

If the data is figured it will look like the following:

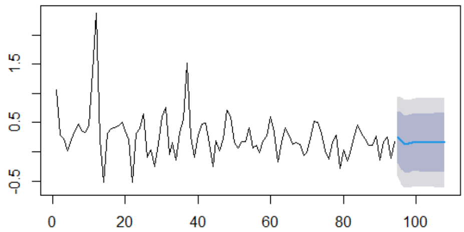

Figure 12 Forecast graph ARIMA (2,1,1)

From Figure 11 the forecasting results are shown by Point Forecast. From the depiction of the inflation forecast in Figure 12, it can be seen that the forecasting results look like a blue line.

IV. CONCLUSION

From the results of the analysis of the East Java inflation rate forecast above, it is found that there are 4 models that suitable and are adequate to be used in the forecasting process. Of the four models, the best model to predict the inflation rate in East Java is the ARIMA model (2,1,1) which has the smallest Akaike's Information Criterion (AIC) value with a value of 76.77. In the forecasting process with the ARIMA (2,1,1) model, forecasts are obtained for the next 14 months until the end of 2022. The forecasting results with the ARIMA (2,1,1) model are 0.2593698 , 0.1892990 , 0.1340639 , 0.1368309 , 0.1572021 , 0.1642381 , 0.1598897 , 0.1557251 , 0.1556074 , 0.1570151 , 0.1576092 , 0.1573511 , 0.1570423 , and 0.1570111.

ACKNOWLEDGMENTS

On this occasion, the author would like to thank all those who have supported this research either directly or indirectly. Thank you to the Badan Pusat Statistik Kediri, Thank you to the Mathematics Lecturer at Nahdlatul Ulama University Blitar who helped the author with discussions and suggestions given during the research.

REFERENCES

Hartati. (2017, Maret). Penggunaan Metode Arima Dalam Meramal Pergerakan Inflasi. Jurnal Matematika, Sains, dan Teknologi, 18.

Hutasuhut, A., Anggraeni, W., & Tyasnurita, R. (t.thn.). Pembuatab Aplikasi Pendukung Keputusan Untuk Peramalan Persediaan Bahan Baku Produksi Plastik Blowing dan Inject menggunakan Metode ARIMA (Autoregressive Integratd Moving Average) Di CV. Asia. Jurnal Teknik Pomits, 3.

Juanda, B., & Junaidi. (t.thn.). Ekonometrika Deret Waktu Teori Dan Aplikasi. Bogor: IPB Press.

Noviyanto, A., Nugroho, R. A., & Kartini, D. (2015). Peramalan Permintaan Paving Blok Dengan Metode Arima. Konferensi Nasional Sistem & Informatika. Bali.

Pitaloka, R. A., Sugito, S., & Rahmawati, R. (2019). Perbandingan Metode Arima Box-Jenkins Dengan Arima Ensemble Pada Peramalan Nilai Impor Provinsi Jawa Tengah. Jurnal Gaussian. Jurnal Gaussian, 8(2), 194-207.

Rohmah, S. (2018). Implementasi Metode Autoregressive Integrated Moving Average (Arima) Dalam Peramalan Jangkapendek (Short Term Forecasting) Terhadap Jumlah Penumpang Kapal Dengan Eviews.

Rosadi, D., & Kurniawan, N. W. (2011). Analisis Ekonometrika & Runtun Waktu Terapan Dengan R. Yogyakarta: Andi.

Simbolon, C. D., Sujatmoko, B., & Sebayang, M. (2013). Analisa Statistika Untuk Curah Hujan Harian Pada Das Kampar Berdasarkan AIC (Akaike Information Criterion).

Statistik, B. P. (2021). Inflasi Provinsi Jawa Timur dan Kota Kediri 2014-2021. (Badan Pusat Statistik)

Sutawijaya, A. (2012). Pengaruh Faktor-Faktor Ekonomi Terhadap Inflasi Di Indonesia. Jurnal Organisasi Dan Manajemen, 8, 85-101.