INTERVENTION MODEL ANALYSIS THE NUMBER OF DOMESTIC PASSENGERS AT SULTAN HASANUDDIN AIRPORTS

Andi Ferosita Sustrisno\(^1\), Rais\(^2\), Iman Setiawan\(^3\)

\(^{1,2,3}\)Statistics Department, Faculty of Mathematics and Natural Sciences, Tadulako University

*e-mail: andiferosit@gmail.com*

ABSTRACT

Domestic passengers are objects whose travel / flight transportation services only cover the domestic area. The increase or decrease in the number of domestic passengers is usually influenced by the occurrence of intervention. This research uses the intervention analysis. Intervention analysis is the time series analysis to model data that is determined by the presence of an intervention. Intervention analysis is one of the time series analysis to model data that are affected by the occurrence of a particular event in a short period of time, such as accidents, natural disasters, and promotions. This research is used to establish intervention model with pulse function of passengers of domestic Sultan Hasanuddin Airport. The result of the research were obtained the model Seasonal ARIMA \((1,1,2)(1,0,1)\)\(^{12}\). There were 6 intervention times during 2006 - 2018, by entering the intervention order \(b = 0\), \(s = 0\), and \(r = 1\) based on the smallest \(\text{AIC}\) value is -303.66 with \(\text{MAPE}\) value is 6.1023.

Keywords: Domestic passanger, Seasonal ARIMA, MAPE, Intervention analysis, pulse function

INTRODUCTION

Air transportation is a means of moving people or goods from a place to a destination by using an aerial vehicle that is driven by humans or machines as a support for human activities by using fast and modern airplanes. In modern times, human needs with the coefficient of time and speed of activity is very important to support all human activities, both in terms of activities for traveling and for shipping goods. The development of air transportation to date is very fast and varied. The development of air transportation to date is very fast and varied. The intended variable is that many private parties establish an airline business that is developed with a variety of prices that are relatively affordable by the public. However, each service and facility provided by the airline varies according to the class or level of the price requested by the consumer (Adisasmita, 2010).

Indonesia has 5 main airports, one of which is Sultan Hasanuddin International Airport. It is the only busiest airport on the island of Sulawesi, located in Makassar City, South Sulawesi Province. The airport is operated by PT. Angkasa Pura I was recorded in the last 5 years domestic aircraft passengers have reached up to 19,856,814 passengers (BPS, 2019).

In the last 13 years, it was found that there were several times during the intervention of Sultan Hasanuddin Airport, including in September 2008 and December 2009 due to the new runway still being worked on causing a decrease in the number of aircraft passengers, then in 2015 there was a quite drastic increase caused by the airport Ngurah Rai air closed, so passengers were dropped off at Sultan Hasanuddin.

Intervention analysis is one of the time series methods that are usually used to explain the effect of interventions from external / internal factors on a time series that affects data patterns (Suhartono, 2007). The specific event referred to here is the existence of an intervention both external and internal. The intervention model in time series data was first used by Box and Tiao in 1975 in the case of the effect of enacting engine design laws, which are thought to have an influence on the level of oxidant pollution in the Los Angeles area.

Several previous studies conducted by Ekayanti (2014) on the analysis of the step function intervention model for forecasting increases in electricity tariffs. The results obtained indicate the amount of usage in May 2014 - April 2015 decreased when the impact of the intervention occurred at the time of the intervention. Other research conducted by Novidian (2015) on the analysis of
interventions to increase the price of fuel oil (BBM), the results of which is the intervention model formed in the intervention to increase subsidized fuel prices in November 2014 is the ARIMA model \((1.0.0)\) with the first intervention order \(b = 1, s = 2, r = 0\) assuming there is no change in the subsidized fuel price policy of 0.3531.

Intervention analysis has the advantage of being able to know whether an event has an effect on the data series so that this method can be applied in modeling the intervention on data on the number of passengers of domestic aircraft by determining the intervention model starting with determining whether the model forms the ARIMA model or the seasonal model.

MATERIALS AND METHODS

1. Research Location

The location of this research was conducted at the Central Statistics Agency and the place of research at the Statistics Laboratory of Statistics Study Program, Department of Mathematics, Faculty of Mathematics and Natural Sciences, Tadulako University.

2. Data

The population used in this study is the number of domestic passengers in 5 major Indonesian airports and the sample used in this study is the number of domestic passengers at Sultan Hasanuddin International Airport as many as 156 observations.

3. Method of Analysis

Data analysis in this study used Intervention Analysis with the help of R. software. The stages of the analysis were as follows:

1. Exploration of data on the number of domestic aircraft passengers at Sultan Hasanuddin Airport.
2. Check the data whether the data is stationary or not, using the Bartlett test to check whether the data is stationary in variety and use the Unit Root test to check whether the data is stationary in the mean.
3. Make ACF and PACF plots to see whether there are seasonal effects or not, and also to measure the close relationship between observations of a time series.
4. Determine either ARIMA or SARIMA model based on ACF and PACF plot. Then evaluate the intervention model use MAPE (Mean Absolute Percentage Error).
5. The best model of step or pulse function intervention.
6. Determine the intervention model based on the calculated AIC value.
7. Interpretation and make conclusions.

4. Time Series Box-Jenkins

The Autoregressive Integrated Moving Average (ARIMA) models have been studied in depth by George Box and Gwilym Jenkins (1976) and their names are synonymous with the ARIMA process which is applied to periodic series processes, forecasting and control or control. Box and Jenkins were successful in determining the relevant information needed to understand and use ARIMA models for periodic series. The basis of the three stages of their approach results indication, assessment and testing, and application. The basic concept of the Box-Jenkins method is based on whether or not there is certainty in a situation in the future, then the time series is divided into two groups, namely time series which is a deterministic process and a stochastic process (Box, 2008).

5. Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF).

The Box-Jenkins forecasting model is identified by observing the behavior of the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF). Basically, both ACF and PACF have the same function, namely to show a linear relationship between time series data separated by lag-k time units. Autocorrelation Function (ACF) Autocorrelation is the relationship that occurs between members of time series data through observations made. Partial Autocorrelation Function (PACF) if there is no autocorrelation in the data, it can be said that the data is random or does not have a pattern. The partial autocorrelation function shows the degree of closeness between \(Y_t\) and \(Y_{t+k}\) provided that it removes the effect of lag 1,2, and so on until k-1.
6. Autoregressive Integreted Moving Average Model (ARIMA)

The Autoregressive Integrated Moving Average (ARIMA) model is based on available data. To identify the ARIMA model, an ACF plot and a PACF plot are made first (Pankratz, 1991). The Autoregressive Integrated Moving Average (ARIMA) model is a non-stationary ARMA (p,q) model. In the non-stationary ARMA (p, q) model, the differentiation process is carried out to make it stationary. After the ARMA model undergoes a differentiation process d times until it is stationary, the ARMA model (p, q) becomes the ARIMA model (p, d, q).

7. Model Seasonal Autoregressive Integrated Moving Average (SARIMA)

Seasonal Autoregressive Integrated Moving Average (SARIMA) is a development of the Autoregressive Integrated Moving Average (ARIMA) model on time series data that has a seasonal pattern. The SARIMA model consists of two types, namely the seasonal model only or ARIMA (P, D, Q) s and the seasonal multiplicative ARIMA model (Montgomery et al. 2008). Seasonality can be determined by identifying autocorrelation coefficients of more than two or three timelags that are significantly different from zero.

8. Intervention Analysis

Intervention analysis is a time series method that is usually used to explain the effect of intervention from external / internal factors on a time series that affects data patterns. Time series data in the intervention model are influenced by other events such as government policies, natural disasters, promotions, and others. Time series data which is influenced by external events with a known time of occurrence can be analyzed using intervention analysis.

In principle, intervention analysis is known by two kinds of intervention function variables, namely the step function intervention variable and the pulse function intervention variable. Step function is a form of intervention that occurs over a long period of time, while pulse function is a form of intervention that occurs only in a certain time or only in one period.

In the intervention model, if the pulse function intervention $P_t$ event occurs during the period $t = T$, it is defined as follows:

$$X_t = P_t = \begin{cases} 
0, & t \neq T \\
1, & t = T 
\end{cases}$$

(1)

RESULTS AND DISCUSSION

1. Data Exploration

Increasing the Number of Passenger Domestic aircraft at Sultan Hasanuddin Airport can be seen between 2006 and 2018. It can be seen that the data patterns form an upward trend and there are seasonal patterns where the upward and downward patterns occur at the same time. On the domestic plane passenger plot that occurred at Sultan Hasanuddin Airport, it can be seen that there were some changes in the time that had a significant influence on the data. The intended effect is a drastic increase and / or decrease in the number of passengers caused by the incident at the time of the intervention. The following graph shows the number of airplane passengers shown in Figure 1:

![Figure 1. Increasing the number of domestic airplane passengers](image-url)
Descriptive analysis is used to provide an overview of the data on the number of domestic aircraft passengers at Sultan Hasanuddin Airport, the description of the data in question can be seen through the magnitude of the variety values, as well as the maximum and minimum values based on data consisting of January 2006 to December 2018.

The factors that influence the increase or decrease in the number of passengers of domestic aircraft at the Airport are clearly known. The occurrence of interventions occurred at six different times. The intervening events at Sultan Hasanuddin Airport are presented in Table 1 as follows:

| Time of Intervention | Label |
|----------------------|-------|
| September, 2008 (T=33) | The decline in the number of airplane passengers was caused by runways. (Khairunnisa, 2017) |
| Desember, 2009 (T=48) | The decline in the number of airplane passengers was caused by runways. (Khairunnisa, 2017) |
| Agustus, 2011 (T=68) | The increase in the number of passengers is due to the decreasing percentage of passengers who cancel flights. (Pakan, 2012) |
| Juli, 2015 (T=115) | The increase in passenger numbers was caused by diverted passenger aircraft from Ngurah Rai Airport which was temporarily closed due to the eruption of Gunung Raung. (Hasanuddin Airport - Angkasa Pura I) |
| Januari, 2016 (T=121) | The increase in the number of aircraft passengers is due to the addition of routes and flight schedules. (Hasanuddin Airport-Angkasa Pura I) |
| Juli, 2017 (T=139) | The increase in the number of airplane passengers is due to the backflow of passengers after the holidays and the many promo airplane tickets. (BUMN- Angkasa Pura Airport) |

2. Data Stationarity

a. Stationarity of variance

Stationarity to variance was carried out to see if there were drastic changes in the data. The fluctuation of the data is in the vicinity of that range value. The non-stationary data will result in a less good estimate of the model. In addition, if the data used in the model is not stationary, then the validity and stability of the data are reconsidered. The detection of stationarity variance is performed during the Bartlett Test with the help of software R.3.4.3, the results in Table 2 are as follows:

| Bartlett’s K-squared | P-value |
|----------------------|---------|
| 30.499               | 0.002348 |

Based on Table 2 the results show that the p-value is 0.002348. The p-value is smaller than 0.05, which means that it is stationary to the variance and data transformation will be performed. Furthermore, after the data is transformed to the value λ = 0, the results obtained in Table 3 are as follows:

| Bartlett’s K-squared | P-value |
|----------------------|---------|
| 17.11                | 0.1455  |

Based on Table 3, the results show that the p-value is 0.1455. The p-value is greater than 0.05, which means the data has been stationary to the variance.

b. Stationarity of Mean

The stationarity test of the mean can be done through the Unit Root test or the Dickey-Fuller Test, which is a formality test for the stationarity of data against the mean.
In stationarity testing against the average using the unit root test and software assistance R.3.4.3 obtained the results in Table 4 as follows:

| Dickey-Fuller Test | P-value |
|--------------------|---------|
| -2.9115            | 0.1964  |

Based on Table 4 the results obtained that the p-value obtained is 0.1964 which is greater than 0.05 (0.1964 > 0.05) which means the data is not stationary to the mean, so that data differencing must be done as much as \( d = 1 \) After differencing data as much as \( d = 1 \), the results obtained in Table 5 are as follows:

| Dickey-Fuller Test | P-value |
|--------------------|---------|
| -6.9618            | 0.01    |

Based on Table 5 the results obtained that the p-value obtained is 0.01 which is smaller than 0.05 (0.01 < 0.05) then accept \( H_1 \), which means the data is stationary to the mean.

3. Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF)

Autocorrelation is the relationship that occurs between members of time series data through observations made. Data that has been stationary with respect to variety and mean can determine the ARIMA or Seasonal ARIMA model which is formed in the data on the number of aircraft passengers by looking at the lag in the ACF graph. The ACF plot is used to get the order in the moving average (MA), while the PACF plot itself is used to get the order \( p \) in a model. This function is performed to determine whether the PACF plot can form a model from the autoregressive (AR) model.

![ACF and PACF plots](image)

Figure 2. (a) Plot ACF and (b) Plot PACF

In Figure 2 it can be seen that the ACF plot of Order \( q \) on the moving average (MA) formed is MA (1) and MA (2) and at lag 12 and lag 36 across the line of trust which means the pattern has a seasonal pattern and a seasonal moving order is obtained average (SMA) are SMA (1)\(^{12}\), SMA (3)\(^{12}\). The Order of Autoregressive (AR) formed is AR (1) and AR (2). Furthermore, for the seasonal pattern formed, which is lag 12, where the lag crosses the line of trust, this study occurs in a seasonal pattern and a seasonal autoregressive order (SAR) is SAR(1)\(^{12}\).

4. Estimation of Seasonal Arima Parameters

The next step after identifying the intervention model is to estimate the parameters that correspond to the model. Estimates for parameters that correspond to the model. With the help of software R.3.4.3, the parameter results for each model and the corresponding MAPE values are shown in Table 6 as follows:
Based on Table 6 the value of the model to be used is seen in the smallest MAPE value. The smallest MAPE value occurs in the ARIMA model ARIMA $(2,1,2)(1,0,1)_{12}$ with a MAPE value of 7.407051 notated as follows:

$$y_t = \frac{e_t + 1.1182e_{t-1} - 0.2259e_{t-2} + 0.9960e_{t-12} + 1.1137e_{t-13} - 0.2249e_{t-14}}{0.4255y_{t-1} - 0.2053y_{t-2} - 0.9984y_{t-12} + 0.4248y_{t-13} + 0.2049y_{t-14}}$$ \hspace{1cm} (2)

In Figure 3 is the actual data plot and the estimated value of the ARIMA model $(2,1,2)(1,0,1)_{12}$ to find out the time of the intervention on the number of domestic aircraft passengers at Sultan Hasanuddin Airport as follows:
5. Identification of Intervention Response

The intervention response identification is done by observing the intervention response plot around the time of the intervention. The intervention response plot in (Appendix 13) shows an image of the intervention response plot that occurred at time $T_1 = 33$, $T_2 = 48$, $T_3 = 68$, $T_4 = 115$, $T_5 = 121$, dan $T_6 = 139$. Determination of the best intervention response will be done in combination with the order $b$, $s$, $r$. Order $b = 0$, $s = 0$ while for order $r$ a combination of order is done with $r = 0$, $r = 1$ and $r = 2$ at $T$ which occurs at the time of intervention. The following results from the intervention responses are shown in Table 7:

| Time- | 33  | 48  | 68  | 115 | 121 | 139 |
|-------|-----|-----|-----|-----|-----|-----|
| order | $b$ | 0   | 0   | 0   | 0   | 0   |
|       | $s$ | 0   | 0   | 0   | 0   | 0   |
|       | $r$ | 1   | 0   | 0   | 0   | 1   |
| AIC   |     | -303.66 |     |     |     |     |

Based on Table 7 Identification of intervention responses in order $r$ which is 83 responses and the best intervention response will be chosen by looking at the smallest AIC value that is equal to -303.66. The effect of the intervention on the sixth time is a pulse function because the time the intervention event occurred only one period since the occurrence of the intervention event Order $b$ and $s$ must have a value of 0 which is known that the variables owned are categorical. As for order $r$ which has a value of 1 because this order explains the number of coefficients $\{\delta_r\}$ in the AR function that are determined at the time the residuals of the data form a clear pattern.

The ARIMA intervention model $(2,1,2)(1,0,1)^{12}$ with the order $b = 0$, $s = 0$, $r = 1$ plot of the actual data passenger plane with the expected value that can be seen in Figure 3.3 as follows:

![Figure 3. Initial data plot and expected values of the ARIMA model $(2,1,2)(1,0,1)^{12}$](image)

![Figure 4. Initial data plot and the estimated value of the ARIMA intervention model ARIMA $(2,1,2)(1,0,1)^{12}$ with order $b = 0$, $s = 0$, and $r = 1$.](image)
Based on Figure 4 shows that the ARIMA (2,1,2) (1,0,1)\(^2\) conjecture model by entering the intervention order \(b = 0\), \(s = 0\), and \(r = 1\) is a good model of guessing. It is shown that the pattern of data between actual passenger data and the estimated value is already good by producing a MAPE value of 6.1023, this value is smaller than the MAPE value before intervention. This is indicated by the graph between the actual data and the conjecture almost coincide.

Order \(b = 0\) shows that the effect of intervention on the number of passengers of domestic aircraft immediately responded at the same time. Order \(s = 0\) shows that the rise and fall in the number of passengers of domestic aircraft at Sultan Hasanuddin Airport only took place at the time of the intervention, after which the number of passengers returned to normal.

### 6. Estimated Intervention Parameters

Estimated intervention parameter is done by least square method, parameter estimation in and calculated using software R.3.4.3 so that the result for the Intervention at the time \(T_1 = 33\), \(T_2 = 48\), \(T_3 = 68\), \(T_4 = 115\), \(T_5 = 121\), \(T_6 = 139\). The corresponding parameter values are as in Table 8 as follows:

| Parameter | Parameter Value |
|-----------|-----------------|
| \(\phi_1\) | -1.4218         |
| \(\phi_2\) | -0.5192         |
| \(\theta_1\) | 0.8410         |
| \(\theta_2\) | 0.0164         |
| \(\Phi_1\) | 0.1956         |
| \(\Theta_1\) | 0.5484         |
| \(\delta_{1(33)}\) | 0.6389         |
| \(\omega_{(33)}\) | -0.2975         |
| \(\omega_{(48)}\) | -0.2888         |
| \(\omega_{(68)}\) | 0.3984         |
| \(\omega_{(115)}\) | 0.1364         |
| \(\omega_{(121)}\) | 0.1757         |
| \(\delta_{1(139)}\) | 1.0186         |
| \(\omega_{(139)}\) | 0.1009         |

Based on Table 8 an intervention model will be formed as follows:

\[
y_t = \frac{-0.2975 X_{t-1}^{(33)} + 0.3030 X_{t-2}^{(33)} + 0.2888 X_{t-1}^{(48)} + 0.4787 X_{t-2}^{(48)} + 0.1880 X_{t-3}^{(48)} - 0.3984 X_{t-1}^{(68)} + 0.6604 X_{t-2}^{(68)} - 0.2593 X_{t-3}^{(68)} - 0.1364 X_{t-1}^{(115)} + 0.493 X_{t-2}^{(115)} - 0.0888 X_{t-3}^{(115)} - 0.1757 X_{t-4}^{(121)} - 0.2913 X_{t-1}^{(121)} + 0.1144 X_{t-2}^{(121)} - 0.1009 X_{t-3}^{(121)} + 0.0645 X_{t-4}^{(121)} + \delta_{1(33)}}{(1 - 1.6575 B - 0.6508 B^2)} + \omega_{(33)}
\]

\[
y_t = \frac{-0.2975 X_{t-1}^{(33)} + 0.3030 X_{t-2}^{(33)} + 0.2888 X_{t-1}^{(48)} + 0.4787 X_{t-2}^{(48)} + 0.1880 X_{t-3}^{(48)} - 0.3984 X_{t-1}^{(68)} + 0.6604 X_{t-2}^{(68)} - 0.2593 X_{t-3}^{(68)} - 0.1364 X_{t-1}^{(115)} + 0.493 X_{t-2}^{(115)} - 0.0888 X_{t-3}^{(115)} - 0.1757 X_{t-4}^{(121)} - 0.2913 X_{t-1}^{(121)} + 0.1144 X_{t-2}^{(121)} - 0.1009 X_{t-3}^{(121)} + 0.0645 X_{t-4}^{(121)} + \delta_{1(33)}}{(1 - 1.6575 B - 0.6508 B^2)} + \omega_{(33)}
\]

\[
y_t = \frac{-0.2975 X_{t-1}^{(33)} + 0.3030 X_{t-2}^{(33)} + 0.2888 X_{t-1}^{(48)} + 0.4787 X_{t-2}^{(48)} + 0.1880 X_{t-3}^{(48)} - 0.3984 X_{t-1}^{(68)} + 0.6604 X_{t-2}^{(68)} - 0.2593 X_{t-3}^{(68)} - 0.1364 X_{t-1}^{(115)} + 0.493 X_{t-2}^{(115)} - 0.0888 X_{t-3}^{(115)} - 0.1757 X_{t-4}^{(121)} - 0.2913 X_{t-1}^{(121)} + 0.1144 X_{t-2}^{(121)} - 0.1009 X_{t-3}^{(121)} + 0.0645 X_{t-4}^{(121)} + \delta_{1(33)}}{(1 - 1.6575 B - 0.6508 B^2)} + \omega_{(33)}
\]

With \(X_t\) the pulse function interventions are caused by only one period of time the intervention effect occurs. From these equations it can be explained that the effects caused by the rise and fall of the number of passengers of domestic aircraft at Sultan Hasanuddin Airport, the time of intervention that occurs is as much as 6 times the equation produced using the Seals onal ARIMA model without intervention and with intervention.
CONCLUSION

Based on the results and discussions that have been done previously, it can be determined the number of domestic aircraft Sultan Sultanuddin Airport using the factors that receive and obtained the ARIMA model $(2,1,2)\times(1,0,1)^{12}$. The model is the best model and has MAPE value.

The best intervention model for pulse function occurs in several interventions known to be the occurrence factor that is 6 times, by entering the order interventions $b = 0, s = 0$, and $r = 1$ is a good guessed model based on the smallest AIC value of -303.66 and also obtained values MAPE before intervention was 7.4070 while for MAPE value after intervention it was smaller that was 6.1023. Intervention models obtained are:

$$y_t = \frac{-0.2975x_{t-1}^{(33)} + 0.3030x_{t-2}^{(33)} + 0.2888x_{t-1}^{(48)} + 0.4787x_{t-2}^{(48)} + 0.1880x_{t-1}^{(48)} + 0.3984x_{t-2}^{(48)} - 0.6604x_{t-3}^{(48)} - 0.2593x_{t-3}^{(48)} - 0.1364x_{t-1}^{(115)} - 0.493x_{t-2}^{(115)} - 0.0888x_{t-3}^{(115)} - 0.1757x_{t-4}^{(115)} - 0.2913x_{t-2}^{(121)} - 0.1144x_{t-3}^{(121)} - 0.1099x_{t-1}^{(139)} - 0.0645x_{t-2}^{(139)} + (1 - 0.8410B - 0.0164B^2 - 0.5484B^{12} + 0.0461B^{13} + 0.00891B^{14})e_t}{(1 - 1.6575B - 0.6508B^2 - 0.5484B^{12} + 0.0461B^{13} + 0.00891B^{14})(1-B)}$$

REFERENCES

Adisasmita, R. (2010). Analisis Kebutuhan Transportasi. Makassar: Universitas Hasanuddin.

Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis Forecasting and Control. San Francisco: Holden-Day.

Box, J. R. (2008). Time Series Analysis: Forecasting and Control (4th ed.). Canada: John Wiley & Sons, Inc.

Badan Pusat Statistik. (2019). Transportasi pada jumlah penumpang pesawat domestik, Badan Pusat Statistik, Jakarta.

Ekayanti. (2014). Analisis Model Intervensi Fungsi Step Untuk Peramalan Kenaikan Tarif Dasar Listrik (Tdl) Terhadap Besarnya Pemakaian Listrik. Pontianak:Universitas Tanjungpura.

Novidian. (2015). Analisis Intervensi Kenaikan Harga Bbm Bersubsidi Pada Data Inflasi Kota Semarang. Semarang:Universitas Diponegoro.

Pankratz, A. (1991). Forecasting with Dynamic Regression Models. John Wiley and Sons, Inc., Canada.

Suhartono & Nuvitasari. (2007). Evaluasi Dampak Krisis Moneter, Bom Bali I dan II terhadap Jumlah Kunjungan Wisatawan ke Bali dengan Model Intervensi Multi Input. Jurnal Ilmiah MatStat.