FORECASTING THE FREQUENCY OF DOMESTIC AIR PASSENGERS AT JUANDA AIRPORT USING ARIMA AND TRANSFER FUNCTION AS A BASIS FOR FUTURE DEVELOPMENT OF AIRPORT SCENARIO

Ary Miftakhul Huda, Heri Kuswanto, Suhartono
Commodity of Futures Trading Regulatory Agency (CoFTRA) – Ministry of Trade
Kramat Raya Street No. 172, Jakarta10430 and (021) 31924744
e-mail: ary.huda@gmail.com

ABSTRACT

In this modern age, the increase of human population is proportional to the need for airport development services, including air transportation. In Indonesia, the development of air transportation services can be seen from the increase of both passenger numbers and goods transported during the last 5 years, including at Juanda Airport in East Java. The aim of this research was to create an airport scenario based on the number of forecasted domestic air passengers. Globally, air passenger forecasting has been one of the most important policy tools used by decision makers. The growth in air passenger numbers in the future can be forecasted based on a time series analytical approach such as multivariate time series, for example the transfer function model as multivariate time series, and ARIMA Box Jenkins as univariate time series. In transfer function method, the input series is the number of aircrafts, inflation, oil prices, while the output series is the number of passengers. Based on Outsample criteria, the best model to forecast the number of domestic passengers at Juanda airport is univariate time series models, ARIMA(1,1,0)(1,0,0)\textsubscript{12} due to its minimum RMSE values. The best model obtained implied that the forecasted number of domestic passengers at Juanda airport this month was related to the number of passengers at 1, 2, 12, 13, 14 months earlier. The result of this research indicated that the forecasted value of domestic air passengers at Juanda airport for 5 years with ARIMA(1,1,0)(1,0,0)\textsubscript{12}, ie from January 2012 to December 2016, would increase gradually each year. Therefore, PT Angkasa Pura 1 should make a scenario to expand the capacity of airports in order to accommodate the increase of air passengers in the future, especially at Juanda Airport.

Keywords: passengers, aircraft, inflation, oil prices, transfer function

INTRODUCTION

Indonesia’s 2012 population, there has been an increment in the population, about 20 million people based on the previous census data. In 2010 population census stated that the population of Indonesia ranges is about 237.556.363 souls that occupy an area which has an area of approximately 1.904.569 km\textsuperscript{2} and the current (in 2012) population in Indonesia has reached 257.516.167 people (Central Bureau of Statistics, 2012). The increase of human population is proportional to the demand of transportation, including air transportation. The development of air transportation services in Indonesia can be seen from the increase in both number of passengers and goods transported during the last 5 years, especially at Juanda Airport.

The increase of air passengers in the future can be forecasted with time series analysis
approaches such as univariate time series (ARIMA) and multivariate time series (Transfer Function). Several studies have been conducted to forecast air passenger. Suryani (2009) et al. conducted a research related to how to develop a model to forecast air passenger needs and to evaluate some policy scenarios related with runway and passenger terminal capacity expansion to meet the future demand using regression methods. In those studies, it was found that plane tickets, service levels, GDPs, population, number of flights per day are influential in determining the volume of air passengers. Furthermore, Insanil (2010) forecasted the number of passengers and aircraft which are very important to be used as reference in the world of transportation and tourism development using the ARIMA Box-Jenkins and ARIMAX involving dummy variable widths as input. Other studies about air passenger have been developed such as in studies indexed [D], [E], [F], and [G].

Transfer Function Model is a model that describes the predictive value of the future from time series (called series or output Yt) that is based on past values of the time series itself and based also on one or more related time series (called input series or Xt) with output series. Several studies using the Transfer Function method have been carried out in Indonesia. Dwi (2010) conducted modeling and forecasting gas sales in Surabaya. Data from the sales of the output variables are estimated to have the effect of input variables and the foreign exchange value of the number of customers which can be analyzed by using multiple input transfer function. By using this method, obtained from the forecasting model whose input variables is Foreign Exchange Rates and the output variable is the monthly sales results.

Nurani (2011) forecasted the stock price index PT HM Sampoerna Tbk by using multiple input Transfer Function. In this study, the stock price index forecasting model will be determined by the input variables that affect the exchange rate, the SBI rate, inflation rate, interest rate deposits, time deposits and positions in rupiah. The best forecasting model will be selected by the model that has all significant parameters, residual white noise, and the normal distribution, as well as having the smallest AIC. The results of the forecasting model in Transfer Function Multi Input value error evaluation remains at certain intervals. Interval evaluation is below 20%, which means the model is good for forecasting data. Other studies about Transfer Function in Indonesia has applied such as in studies indexed [J] and [K].

The development of research using the Transfer Function method has been performed on a variety of research around the world. Nelson and Paulo (2003) conducted a study to explain the numerical linear algebra algorithms for calculating the variable the dominant power in the Multi Input Multi Output (MIMO) using a Transfer Function. Joe Kniss et al. (2003) conducted a study using a Gaussian Transfer Function because it is considered suitable to model and classify variable multivariate multidimensional data structures with numerical and graphical approach. Other studies have also applied this method, such as studies indexed [N], [O], and [P].

RESEARCH METHOD

Air traffic statistical data, namely SLALU, were obtained from PT Angkasa Pura 1, an Indonesian airport operator. The data includes data of air passenger numbers and monthly domestic flight numbers within the period of January 2000 to December 2011. Monthly inflation and oil price data were obtained from the website of BPS-Statistics Indonesia, starting from January 2000 to December 2011 with monthly.

As for the analyses, this research employed two forecasting models known as ARIMA and Transfer Function. In the Transfer Function model, the input series is the number of aircrafts, inflation, oil prices, while the output series is the number of air passengers. The first procedure performed in the data analysis was the separation of input data series i.e., the number of aircrafts, global oil prices, and inflation rates; and output data series i.e., two passenger numbers data, namely the data in-sample and out-sample of data. Data in-sample totaled 132 from 2000 to 2011 and the data out-sample amounted to 12 i.e January to December of 2011.

Then, the air passengers were forecasted using two approaches, ARIMA and Transfer Function. Box and Jenkins (1976) explained that there are four procedures in ARIMA which are model identification, parameter estimation, diagnostic check, and forecasting. Bowerman (1993) explained that there are three procedures for building a Transfer Function model. Those are:

1. Identification of a model describing x_t and prewhitening of x_t and y_t.
2. Calculation of the sample cross-correlation function between the α and β, values, and identification of a preliminary transfer function model.
3. Identification of a model describing and of a final transfer function model.

RESULT AND DISCUSSION

Annual average of passenger numbers from 2000 to 2011 at Juanda Airport showed a continuous growth. This shows an increase in air transportation demand. The value of standard deviation was 254,340 passengers, meaning that the number of passengers per month displays a high variation. The highest number of passengers recorded by PT Angkasa Pura from 2000 to 2011 are as many as 1,190,385 people, while the lowest number of passengers are as many as 152,139 people.

ARIMA Box Jenkins (Univariate Time Series)

In determining the forecasting model with univariate time series, it can be seen from the time series plot of the data on the number of passengers to arbitrate the patterns and characteristics of the data. The time series plot of air passengers showed that the data of numbers of passengers is not stationary in the mean, so it is necessary to process non-seasonal differencing order 1. After undertaking the process, data on the number of passengers becomes stationary, then the estimation parameter is tested using t test statistic. ARIMA model must be checked whether the residual assumption is white noise with L-Jung Box Test. Then, the testing residual which is having a normal distribution or not with Kolmogorov Smirnov. Based on the results of the Kolmogorov Smirnov test, three models, such as ARIMA (0,1,1) [12], ARIMA (0,1,1)(1,0,0) [12], are not normally distributed because the p-value of the residual value is less than the value of α = 0.05. However, the model ARIMA (1,1,0)(1,0,0) [12] had normal distribution because p-value of the residual has greater value than the value of α = 0.05.

<table>
<thead>
<tr>
<th>Model ARIMA</th>
<th>AIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,1,1)[12]</td>
<td>171,2007</td>
</tr>
<tr>
<td>(1,1,0)[12]</td>
<td>164,1818</td>
</tr>
<tr>
<td>(0,1,1)[12]</td>
<td>171,6642</td>
</tr>
</tbody>
</table>

The best model in the number of passengers is ARIMA (1,1,0)(1,0,0)12 for the model parameters as well as meeting the assumptions of significant residual white noise and normally distributed, and has minimum AIC value than other models. ARIMA model equations for forecasting the number of passengers above, explain that forecasting the numbers of passengers at the Juanda airport this month is related to the number of passengers at 1, 2, 12, 13, 14 months earlier.

Transfer Function (Multivariate Time Series)

To forecast the number of passengers with transfer function method, the considered best ARIMA model must be selected in the input sequence, such as the number of aircrafts, global oil prices, and inflation rates. ARIMA model equations are considered important in the formation of the transfer function model. ARIMA (1,1,0)(1,0,0) [12] showed that the number of aircrafts at the current time period is related to the number of aircraft in period 1, 12, 13 months earlier.

The best ARIMA model for global oil price model is ARIMA(1,1,0). Equation in global oil prices meant that forecasting oil prices this month associated with global oil prices at 1, 2, 6 months earlier.

The best model which is used to model the inflation model is ARIMA (0,1,1) for the model parameters as well as meeting the assumptions of significant residual white noise and normally distributed, and has the minimum AIC value compared to other models. The equations is taken from ARIMA model for the input series of inflation, namely ARIMA(0,1,1) which explains that the value of forecasting inflation this month is related to the inflation rate in the next 1 month earlier.

The next procedure in modeling the transfer function is doing prewhitening to sequence the input and output series. Then, the transfer function model for each input sequence will be formulated. Noise series model for input sequence in number of aircraft is ARIMA(0,0,1)[12,23]). The relationship between input series with series output, occur at lag-0 and lag-12. As the final Transfer Function model is obtained, it can be explained that the pattern of domestic passenger numbers at Juanda airport this month has dependencies with domestic passenger numbers 1 month earlier and has a linear correlation with the number of aircraft 12 and 13 months earlier. Therefore, the interpretation of the noise series model which is ARIMA(0,0,1)[12,23]) with linkages occur at lag-0 and lag-12, can be used to describe a linear relationship between the number of aircraft on the number of domestic passengers at Juanda airport.

The Transfer Function model on the input sequence in global oil prices is formulated as ARIMA(1)[12],0,0) with those orde is b= 2, r=0; s=0. The interpretation of the final model of do-
mestic passenger’s pattern at Juanda airport has dependencies in the number of domestic passengers at Juanda airport in months of 1, 2, 12, 13 months earlier and has a linear correlation with oil prices at 2, 3, 4, 14, 15 months earlier.

Meanwhile, the Transfer Function model for the Inflation rate is ARIMA(0,0,2 [12,24]) of those orde is b=16; r=0; s=0. It can be seen that the number of domestic passengers at Juanda airport this month has dependencies with the number of domestic passengers at Juanda airport in one month in advance and has a linear correlation with inflation at 16 and 17 months earlier.

Moreover, the establishment of Multi Input Transfer Function Models is done by firstly testing the significance of parameters using a statistical t test with significance level $\alpha = 0.05$ to ARIMA (0,0,1 [12]). It is then known that all the ARIMA model parameters are statistically significant seen from its p-value less than $\alpha=0.05$ but the orde of parameter (b, r, s) from the output series, such as global oil price ($\omega_2$), inflation rate ($\omega_16$) that are not statistically significant because p-value is greater than $\alpha=0.05$. Input sequence which is only statistically significant by number of aircrafts with p-value less than $\alpha=0.05$. Therefore, the forecasting model of air passengers using the approach of Multivariate Time Series is Single Input Transfer Function with an input sequence of the number of the aircrafts.

**Best Model in Forecasting Domestic Air Passengers**

As the forecasting model for domestic air passengers with ARIMA (univariate time series) and transfer function (multivariate time series) being obtained, the next stage is to compare both best models to forecast domestic air passengers at Juanda airport with the out-sample criteria using the criteria of RMSE.

**Table 2. Comparison between ARIMA and Transfer Function**

<table>
<thead>
<tr>
<th>Model</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>0.600891</td>
</tr>
<tr>
<td>Transfer Function</td>
<td>0.9255667</td>
</tr>
</tbody>
</table>

Based on the Out-sample criteria, the best model to forecast number of domestic air passengers at Juanda airport is the univariate time series models, ARIMA (1,1,0)(1,0,0)$^{12}$ because it has a minimum RMSE values.

From Fig. 1. (a) it is understood that the forecasting value and the actual value of domestic passenger numbers at the Juanda airport using ARIMA Box Jenkins are still in the lower and upper bound forecasting with 95% confidence level. This suggests that ARIMA(1,1,0)(1,0,0)$^{12}$ is capable of forecasting domestic air passengers. From Fig. 1. (b), the forecasted the number of domestic passengers at Juanda airport for 5 years, explained that the number of passengers from January 2012 to December 2016 is well within the upper and lower limit prediction. It suggests that the value of forecasted number of domestic passengers at Juanda airport in January 2012 to December 2016 is worth using.
In Table 3, the forecasted of the number of domestic air passengers at Juanda airport for 5 years, i.e from January 2012 to December 2016, shows an increasing trend annually. It is therefore suggested that PT Angkasa Pura 1 should make a planning to face this condition.

**CONCLUSION**

In this research, Univariate Time Series and Multivariate Time Series were used. Based on the ARIMA Box Jenkins analysis, the best model is ARIMA(1,1,0)(1,0,0)\(^2\). ARIMA model equations for forecasting the number of passengers explained that forecasting the number of passengers at the airport Juanda this month is related to the number of passengers at 1, 2, 12, 13, 14 months earlier.

Furthermore, it is also concluded that the best model for forecasting domestic air passengers using multivariate time series is Single input transfer function. The generated model explained that the pattern of the number of domestic passengers at Juanda airport has dependencies with the number of domestic passengers at Juanda airport in one month in advance and have a linear correlation with the number of aircraft on 12 and 13 the previous month. Therefore, the interpretation of the model noise series ARIMA (0,0,1 [12,23]) with linkages occur at lag-0 and lag-12 can be used to describe a linear relationship between the number of aircraft on the number of domestic passengers at airport Juanda.

The best model among univariate time series and multivariate time series is choosen by an out sample criteria. The result showed that the best model for forecasting domestic air passengers is univariate time series with ARIMA(1,1,0)(1,0,0)\(^1\) because it has minimum value of Root Mean Squared Error (RMSE). Then, the ARIMA model is capable of forecasting domestic air passengers in 5 years because the forecast value is still in lower and upper 95% confidence Interval. In the forecast value, we know that the domestic air passengers is likely to increase each year. So, based on this information, PT Angkasa Pura 1 should expand his capacity in order to face the increase of domestic air passengers in the future.

**ACKNOWLEDGMENT**

Authors would like to express thank to the reviewer for giving helpful comments. Authors would also like thank PT. Angkasa Pura 1 for providing air passengers and aircraft data from 2000 until 2011. This research was funded by the Department of Statistics, Mathematics and Science Faculty, Institute of Teknologi Sepuluh Nopember (ITS).

**DAFTAR PUSTAKA**


---

**Table 3. Forecasting domestic air passengers for 5 years**

<table>
<thead>
<tr>
<th>Month</th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
<th>2015</th>
<th>2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>January</td>
<td>11,242.762</td>
<td>12,279.305</td>
<td>13,058.033</td>
<td>13,643.140</td>
<td>14,082.768</td>
</tr>
<tr>
<td>February</td>
<td>10,611.070</td>
<td>11,804.525</td>
<td>12,701.301</td>
<td>13,375.105</td>
<td>13,881.376</td>
</tr>
<tr>
<td>March</td>
<td>11,232.006</td>
<td>12,271.166</td>
<td>13,051.917</td>
<td>13,638.545</td>
<td>14,079.315</td>
</tr>
<tr>
<td>April</td>
<td>10,897.924</td>
<td>12,020.093</td>
<td>12,863.270</td>
<td>13,496.802</td>
<td>13,972.815</td>
</tr>
<tr>
<td>May</td>
<td>11,229.997</td>
<td>12,269.635</td>
<td>13,050.767</td>
<td>13,637.681</td>
<td>14,078.666</td>
</tr>
<tr>
<td>June</td>
<td>11,545.782</td>
<td>12,506.882</td>
<td>13,229.026</td>
<td>13,771.618</td>
<td>14,179.301</td>
</tr>
<tr>
<td>July</td>
<td>12,319.675</td>
<td>13,088.370</td>
<td>13,665.935</td>
<td>14,099.895</td>
<td>14,425.957</td>
</tr>
<tr>
<td>August</td>
<td>10,324.480</td>
<td>11,589.246</td>
<td>12,539.548</td>
<td>13,253.570</td>
<td>13,790.059</td>
</tr>
<tr>
<td>September</td>
<td>12,776.622</td>
<td>13,431.700</td>
<td>13,923.900</td>
<td>14,293.720</td>
<td>14,571.590</td>
</tr>
<tr>
<td>October</td>
<td>11,933.663</td>
<td>12,798.329</td>
<td>13,448.008</td>
<td>13,936.153</td>
<td>14,302.927</td>
</tr>
<tr>
<td>November</td>
<td>12,315.779</td>
<td>13,085.438</td>
<td>13,663.731</td>
<td>14,098.239</td>
<td>14,424.713</td>
</tr>
<tr>
<td>December</td>
<td>12,475.947</td>
<td>13,205.782</td>
<td>13,754.153</td>
<td>14,166.179</td>
<td>14,475.760</td>
</tr>
</tbody>
</table>
FORECASTING THE FREQUENCY OF DOMESTIC AIR PASSENGERS AT JUANDA AIRPORT USING ARIMA AND TRANSFER FUNCTION AS A BASIS FOR FUTURE DEVELOPMENT OF AIRPORT SCENARIO


