Time Series Analysis of Demand for Domestic Air Travel in Nigeria

Eke, Charles N.      Okwuosha, Kenneth C.      Egwim, Kenneth C.      Nkwazema, Oluchukwu A.
Department of Mathematics and Statistics, Federal Polytechnic Nekede, Owerri, Imo State, Nigeria
E-mail: ngomeeke@gmail.com

Abstract
This work fitted three time series models namely, AR (1) model, MA (1) model and ARMA (1,0,1) model on demand for domestic air travel in Nigeria using quarterly data from 2003 to 2012. It was found out that the ARMA (1,0,1) model had the least MSE and fitted the data appropriately. The ARMA (1,0,1) model was used to make a five year quarterly forecast for demand for domestic air travel in Nigeria, which showed that the country’s demand for domestic air travel will be on the rise within the next five years.

Keywords: Model, Forecast, Time Series and Transportation

1.0 Introduction
Air travel is the fastest means of transporting passengers and cargo in Nigeria. This is because of its efficiency and value when long distances are involved. Therefore, this has led to an increase in demand for it by local passengers. In Nigeria, air transportation has contributed significantly towards economic growth and domestic movement. Consequently, this increase in demand calls for proper planning on the part of managers of air transportation. Thus, forecasting the future demand pattern for domestic air travel is very important to enable planning. In order to forecast the future demand pattern, data on demand for domestic air travel for the period (2003 – 2012) was obtained from Federal Airports Authority of Nigeria (FAAN). The aim of this study is to determine a statistical time series model of demand for domestic air travel in Nigeria within the stated period and use the obtained model to forecast future demand for domestic air travel in the country. In this regard, the research questions are what is the trend of domestic air travel in Nigeria? Secondly, is the future demand for domestic air travel in Nigeria increasing or decreasing?

2.0 Method
The Autoregressive Moving Average (ARMA) process was applied to estimate and forecast the future demand for domestic air travel in Nigeria. The general ARMA model applied to this work is for a univariate time series data. The model is given as ARMA (p, q), which is Autoregressive process of order p and Moving Average process of order q.

\[ Y_t = C_t + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots + \alpha_p Y_{t-p} + \epsilon_t + \lambda_1 \epsilon_{t-1} + \lambda_2 \epsilon_{t-2} + \ldots + \lambda_q \epsilon_{t-q} \]

In addition, the Box and Jenkins three step approach was also adopted. The three steps are:

a) Identification Stage: This stage is used to specify the response series and identify candidate ARMA models for it. This process reads time series that are to be used in estimation and forecasting stages, possibly differencing it and compute the autocorrelation function and partial autocorrelation function.

b) Estimation/Diagnostic Checking Stage: At this level, ARMA model is specified to fit the variable of interest that has been previously identified in the first stage. This will enable estimation of the model parameters and selecting the best model using suitable criterion.

c) Forecasting Stage: This stage is to forecast future values of the time series and to generate confidence intervals for the forecast values from the ARMA model.

\[ r_k = \frac{\sum_{t=k+1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2} \]

2.0.1. Autocorrelation Function (ACF): This is defined as . It describes the pattern of autocorrelation for lags 1, 2… A plot of the ACF against the lag is known as the correlogram. It is also usually used to identify whether or not seasonality is present in a given time series. (Makridakis et al, 1998)

2.0.2. Partial Autocorrelation Function (PACF): This is a measure of correlation that is used to identify the extent of relationship between current values of a variable with earlier values of the same variable (values for various time lags) holding the effect of all other things constant. (Makridakis et al, 1998).

2.0.3. Ljung-Box Test: This is a test for auto-correlated errors. It is an improved version of Box-Pierce test and
defined as 
\[ Q^* = n(n+2) \sum_{k=1}^{k} (n-k)^{-1} \] 
It has a distribution closer to the Chi-square distribution. (Makridakis et al, 1998)

3.0 Results
3.0.1 Graph of Domestic Air Travel in Nigeria
Fig.1. Time Series Plot of Domestic Air Travel in Nigeria

Author’s computation and Minitab 16 output
The graph of domestic air travel in Nigeria as shown in fig.1 shows that there is a trend and seasonal variation in the data.

3.0.2 Anderson Darling Normality Test
Fig.2. Normality Probability

Author’s computation and Minitab 16 output
The Anderson Darling normality test was applied to verify if the data was from a normally distributed population. The plot in fig. 2 shows that the data points are outside the bounds, which means the data is not from a normally distributed population. This is confirmed by AD of 1.672 with P-Value < 0.005. Therefore the data need a transformation to make it a normally distributed data.
3.0.3 Johnson’s Transformation (Natural Logarithm Transformation)
Fig.3. Normality Probability of Transformed Data

The original data was transformed using the Johnson’s transformation by converting the original values of the data to its natural logarithm values. The transformation of the original data shows a normally distributed data in fig.3 with all the data points within bounds. The AD of 0.348 with P-Value of 0.460 is a test confirmation.

3.0.4 Test for Homogeneity of Variance
Hypothesis:

\[ H_0 : \text{Data are homogeneous} \]
\[ H_1 : \text{There is a date at which there is a change in the data} \]

Decision Rule: We shall reject \( H_0 \) if p-value<\( \alpha \), otherwise, we shall not reject \( H_0 \) if \( \alpha = 0.05 \).

Bartlett’s Test (Normal Distribution)
Test statistic = 5.23, p-value = 0.814
Conclusion: The p-value (0.814) of the Bartlett’s Test leads us not to reject the null hypothesis, thereby concluding that the data is homogeneous.

3.0.5 Test for Stationarity
Augmented Dickey-Fuller Unit Root Test
Hypothesis:

\[ H_0 : \text{There is unit root in the data (is not stationary).} \]
\[ H_1 : \text{There is no unit root in the data (is stationary).} \]

Decision Rule: We shall reject \( H_0 \) if p-value<\( \alpha \), otherwise, we shall not reject \( H_0 \). \( \alpha = 0.05 \).
126

Table 1: ADF Stationarity Test on Transformed Data

<table>
<thead>
<tr>
<th>Augmented Dickey-Fuller Unit Root Test on Transformed Data</th>
<th>t-Statistic</th>
<th>Prob.*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Augmented Dickey-Fuller test statistic</td>
<td>-3.0404</td>
<td>0.0399</td>
</tr>
<tr>
<td>Test critical values:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1% level</td>
<td>-3.6105</td>
<td></td>
</tr>
<tr>
<td>5% level</td>
<td>-2.9390</td>
<td></td>
</tr>
<tr>
<td>10% level</td>
<td>-2.6080</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformed Data(-1)</td>
<td>-0.4054</td>
<td>0.1333</td>
<td>-3.0404</td>
<td>0.0043</td>
</tr>
<tr>
<td>C</td>
<td>-0.0402</td>
<td>0.1334</td>
<td>-0.3014</td>
<td>0.7648</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA 1</td>
<td>-0.4780</td>
<td>0.1425</td>
<td>-3.36</td>
<td>0.002</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.1241</td>
<td>0.2054</td>
<td>-0.60</td>
<td>0.549</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.1241</td>
<td>0.2054</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Mean dependent var: -0.1241
S.D. dependent var: 0.2054
Akaike info criterion: 2.5156
Schwarz criterion: 2.6009

Author’s computation and Minitab 16 output

Conclusion: Since the ADF statistic of -3.0404 shows that there is no unit root at both 5% and 10%, we reject the null hypothesis and conclude that there is no unit root in the data – it is stationary.

3.0.6 Fitting ARIMA Model to the Differenced Data

ARMA (0,0,1) Model

\[ Y_t = C_t + \epsilon_t + \lambda_1 \epsilon_{t-1} \]  

(1)

Final Estimates of Parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA 1</td>
<td>-0.4780</td>
<td>0.1425</td>
<td>-3.36</td>
<td>0.002</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.1241</td>
<td>0.2054</td>
<td>-0.60</td>
<td>0.549</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.1241</td>
<td>0.2054</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of observations: 40
Residuals: SS = 29.4784 (backforecasts excluded)
MS = 0.7757 DF = 38

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

<table>
<thead>
<tr>
<th>Lag</th>
<th>Chi-Square</th>
<th>DF</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>11.4</td>
<td>22</td>
<td>0.328</td>
</tr>
<tr>
<td>24</td>
<td>24.6</td>
<td>34</td>
<td>0.156</td>
</tr>
</tbody>
</table>

Interpretation

The ARMA (0,0,1) parameter (MA1) had a p-value of 0.002 which indicates that it is significant. The Ljung-Box statistics also gives significant p-values, indicating that the residuals are uncorrelated. This means that the model is adequate, but we shall fit other ARMA models to compare.
As the spikes die down, the ACF and PACF of the residuals in fig.4 and fig.5 respectively corroborate with our interpretation above, affirming that the residuals are uncorrelated. The ARMA (0,0,1) model appears to fit well, but we shall fit more models to compare.

\[ Y_t = C_t + \alpha \epsilon_{t-1} + \epsilon_t \quad \ldots \quad (2) \]

Final Estimates of Parameters

<table>
<thead>
<tr>
<th>Type</th>
<th>Coef</th>
<th>SE Coef</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR 1</td>
<td>0.6048</td>
<td>0.1317</td>
<td>4.59</td>
<td>0.000</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0682</td>
<td>0.1305</td>
<td>-0.52</td>
<td>0.605</td>
</tr>
<tr>
<td>Mean</td>
<td>-0.1725</td>
<td>0.3302</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of observations: 40
Residuals: \( SS = 25.7347 \) (backforecasts excluded)
\( MS = 0.6772 \) DF = 38
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

<table>
<thead>
<tr>
<th>Lag</th>
<th>12</th>
<th>24</th>
<th>36</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-Square</td>
<td>2.6</td>
<td>12.5</td>
<td>30.1</td>
<td>*</td>
</tr>
<tr>
<td>DF</td>
<td>10</td>
<td>22</td>
<td>34</td>
<td>*</td>
</tr>
<tr>
<td>P-Value</td>
<td>0.990</td>
<td>0.945</td>
<td>0.660</td>
<td>*</td>
</tr>
</tbody>
</table>

**Interpretation**

The ARMA (1,0,0) parameter (AR1) had a p-value of 0.000 which indicates that it is significant. The Ljung-Box statistics also gives significant p-values, indicating that the residuals are uncorrelated. This means that the model is adequate, but we shall fit yet another model for comparison.

**Fig.6. ACF of Residual for Transformed Data using AR1 Process**

![ACF of Residuals for Transformed Data](image)

Author’s computation and Minitab 16 output

**Fig.7. PACF of Residual for Transformed Data using AR1 Process**

![PACF of Residuals for Transformed Data](image)

Author’s computation and Minitab 16 output

The ACF and PACF spikes are all within the lines and also die down as shown in fig.6 and fig.7 respectively; the residuals confirm our interpretation above, affirming that the residuals are uncorrelated.

ARMA (1,0,1)

\[ Y_t = C_t + \alpha Y_{t-1} + \varepsilon_t + \lambda \varepsilon_{t-1} \]  

……………….(3)

Final Estimates of Parameters
Type          Coef  SE Coef      T      P
AR 1       0.8221   0.1487   5.53  0.000
MA 1       0.3517   0.2363   1.49  0.145
Constant  -0.04427  0.08569  -0.52  0.608
Mean       -0.2488   0.4816

Number of observations: 40
Residuals: SS = 24.8743 (backforecasts excluded)
            MS = 0.6723  DF = 37

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag            12     24     36   48
Chi-Square    2.3   13.5   27.9   *
DF              9     21     33   *
P-Value     0.986  0.888  0.720   *

Interpretation
The ARMA(1,0,1) parameters (AR1 & MA1) have p-values of 0.000 & 0.145 respectively, which indicates that AR1 is significant while MA1 is not. However, the Ljung-Box statistics also gives significant p-values, indicating that the residuals are uncorrelated. We shall compare the models, and then pick the best to forecast future demand for domestic air travel in Nigeria.

Fig.8. ACF of Residual for Transformed Data using AR1 and MA1 Process

Author’s computation and Minitab 16 output
Fig.9. PACF of Residual for Transformed Data using AR1 and MA1Process

Author’s computation and Minitab 16 output

As the spikes decay, the ACF and PACF spikes of the residuals in fig.8 and fig.9 corroborate our interpretation above, affirming that the residuals are uncorrelated. The ARMA(1,0,1) model appears to fit well, but we shall fit compare the fitted model and choose the best for forecasting the future demand of domestic travel in Nigeria.

3.0.7 Comparative Analysis of the Models
Table 2: Mean Square Error of the Models

<table>
<thead>
<tr>
<th>MODELS</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMA(0,0,1)</td>
<td>0.7757</td>
</tr>
<tr>
<td>ARMA(1,0,0)</td>
<td>0.6772</td>
</tr>
<tr>
<td>ARMA(1,0,1)</td>
<td>0.6723</td>
</tr>
</tbody>
</table>

Author’s computation and Minitab 16 output

The table 2 above shows the Mean Square Error (MSE) of the estimated model fitted to the data. From the table, it shows that the Mean Square Error of ARMA(1,0,1) has the smallest Mean Square Error of 0.6723. This means that ARMA(1,0,1) is the most fit for the data and adequate for forecasting future demand for domestic air travel in Nigeria.

3.0.7 Forecast

The ARMA(1,0,1) model will be used as the forecasting model since it has the least Mean Square Error value when compared with other estimated models. The model is as shown in equation 4:

\[ Y_t = -0.0443 + 0.8221 Y_{t-4} - 0.3517 Y_{t-1} + \ell_t \]

The five year quarterly forecast computed with the model is shown in table 3 b, while fig.10 shows the plot.
The forecast table and plot show that demand for the domestic air travel in Nigeria is on the increase and will continue to be on the increase for the next five years.

4.0 Conclusion
The time series analysis for the demand of domestic air travel in Nigeria was evaluated using the available data obtained from Federal Airports Authority of Nigeria (FAAN). The analysis has helped to answer key questions posed at the beginning of this study. It could be extracted that in Nigeria, that demand for domestic air travel has an increasing trend and, more so, within the next five years there will be a surge in the demand for air travel domestically in Nigeria. This means that air transportation managers in the country should effectively plan for this increase in demand by providing all the necessary facilities that will aid in smooth operations in our local airports. This also, calls for our aviation policy makers to be proactive in their decisions towards air transportation in Nigeria.
References
The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digital Library, NewJour, Google Scholar