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.

$$Y_t = C_t + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + Y_{t-p} + \varepsilon_t + \lambda_1 \varepsilon_{t-1} + \lambda_2 \varepsilon_{t-2} + \dots + \lambda_q \varepsilon_{t-p}$$

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.

2.0.1. Autocorrelation Function (ACF): This is defined as $r_k = \frac{\sum_{t=k+1}^{n} (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n} (Y_t - \bar{Y})^2}$. 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)

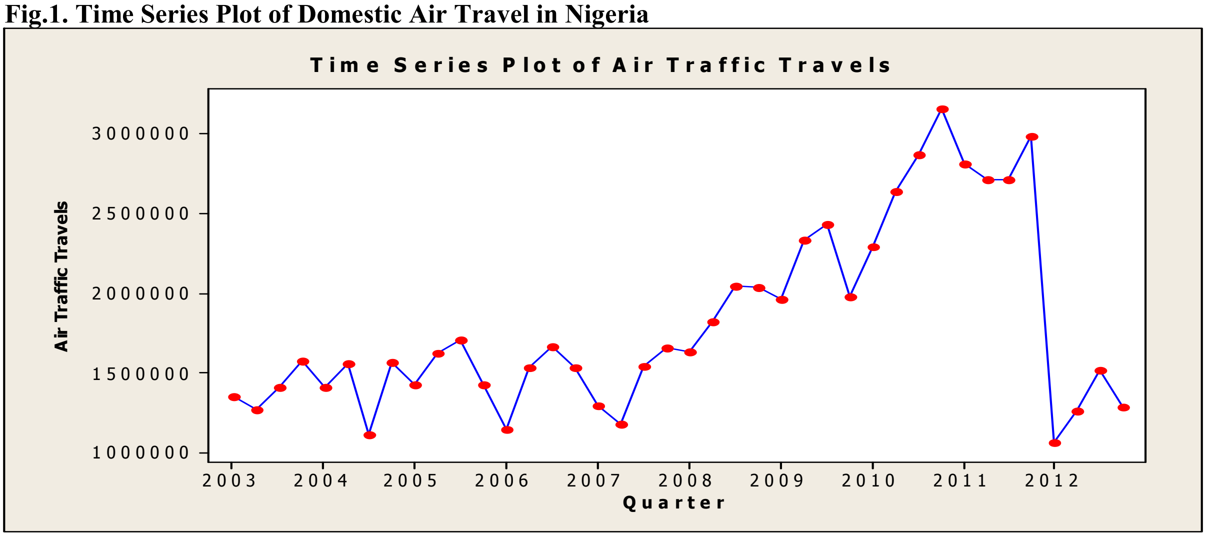
2.0.2. Partial Autocorrelation Function (PACF): This is a measure of correlation that is used to identify the extent of relationship between current values of a variable with earlier values of the same variable (values for various time lags) holding the effect of all other things constant. (Makridakis 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 $Q^{*} = n(n+2)\sum_{k=1}^{h}(n-k)^{-1} r_k^2$. 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

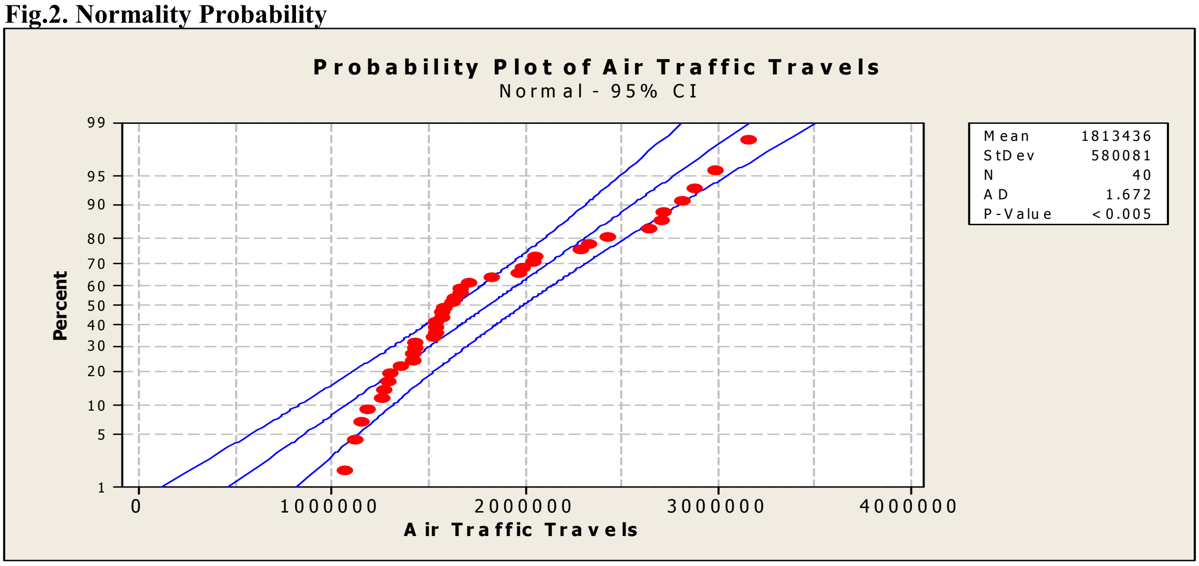


Authour's computation and Minitab 16 output

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

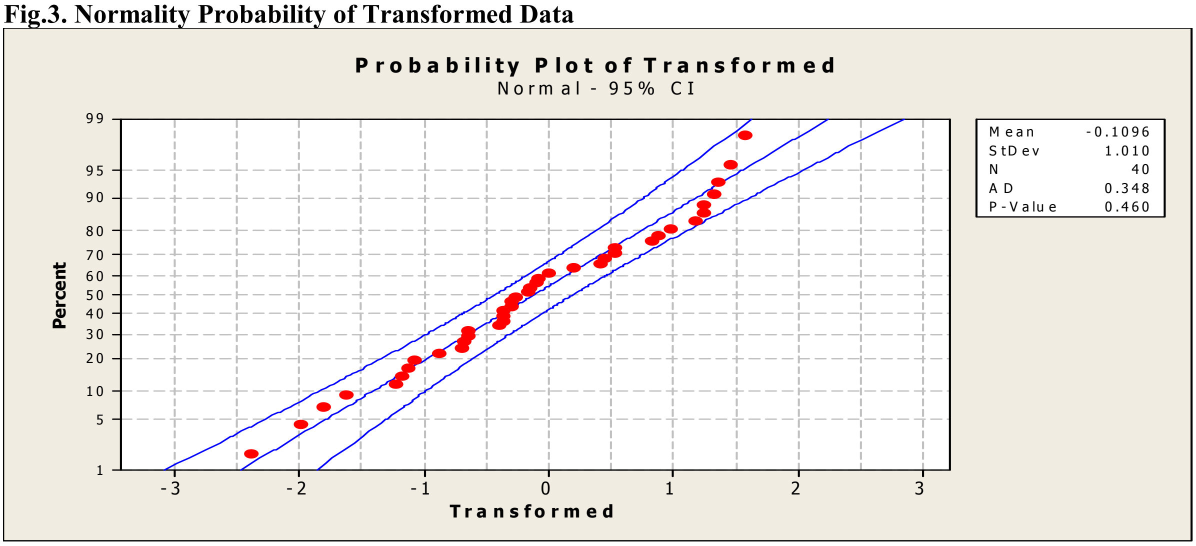


Authour's computation and Minitab 16 output

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



Authour's computation and Minitab 16 output

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:

$H0$: Data are homogeneous

$H1$: There is a date at which there is a change in the data

Decision Rule: We shall reject $H0$ if p-value<α, otherwise, we shall not reject $H0$ if α = 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: There is unit root in the data (is not stationary).

H_1: There is no unit root in the data (is stationary).

Decision Rule: We shall reject H_0 if p-value<α, otherwise, we shall not reject H_0. α = 0.05.

Table 1: ADF Stationarity Test on Transformed Data

Augmented Dickey-Fuller Unit Root Test on Transformed Data				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-3.0404	0.0399
Test critical values:	1% level		-3.6105	
	5% level		-2.9390	
	10% level		-2.6080	
Variable	Coefficient	Std. Error	t-Statistic	Prob
Transformed Data(-1)	-0.4054	0.1333	-3.0404	0.0043
C	-0.0402	0.1334	-0.3014	0.7648
R-squared	0.1999		Mean dependent var	-0.0062
Adjusted R-squared	-0.2002		S.D. dependent var	0.9158
S.E. of regression	0.8302		Akaike info criterion	2.5156
Sum squared resid	25.5012		Schwarz criterion	2.6009
Log likelihood	-47.0543		F-statistic	9.2438
Durbin-Watson stat	2.1356		Prob (F-statistic)	0.0043

Authour'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

$$Y_t = C_t + \varepsilon_t + \lambda_1 \varepsilon_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
MA 1	-0.4780	0.1425	-3.36	0.002
Constant	-0.1241	0.2054	-0.60	0.549
Mean	-0.1241	0.2054		

Number of observations: 40

Residuals: SS = 29.4784 (backforecasts excluded)

MS = 0.7757 DF = 38

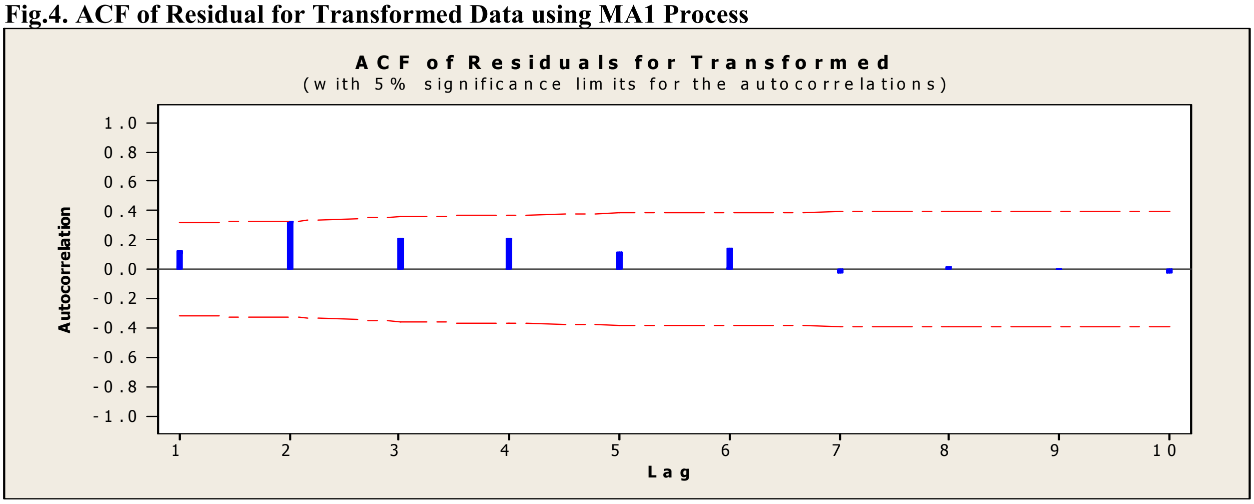
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	11.4	24.6	42.3	*
DF	10	22	34	*
P-Value	0.328	0.319	0.156	*

Interpretation

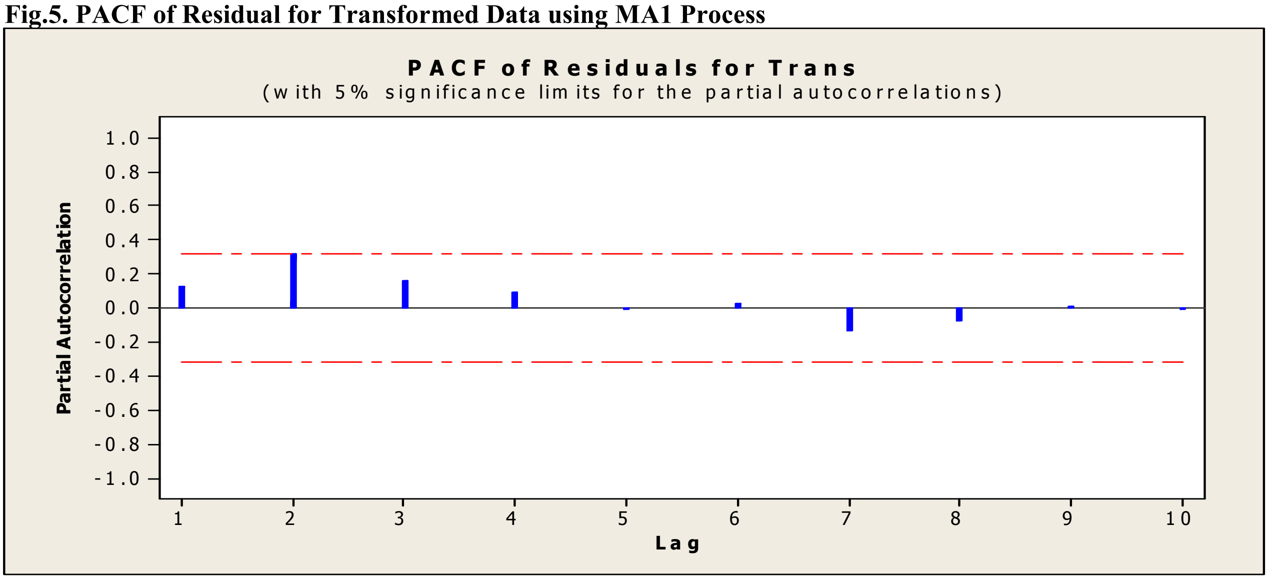
The ARMA (0,0,1) parameter (MA1) had a p-value of 0.002 which indicates that it is significant. The Ljung-Box statistics also gives significant p-values, indicating that the residuals are uncorrelated. This means that the model is adequate, but we shall fit other ARMA models to compare.

Fig.4. ACF of Residual for Transformed Data using MA1 Process



Authour's computation and Minitab 16 output

Fig.5. PACF of Residual for Transformed Data using MA1 Process



Authour,s computation and Minitab 16 output

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.

ARMA (1,0,0)

$$Y_t = C_t + \alpha_1 Y_{t-1} + \varepsilon_t \quad \text{......................(2)}$$

Final Estimates of Parameters

Type	Coef	SE Coef	T	P
AR 1	0.6048	0.1317	4.59	0.000
Constant	-0.0682	0.1305	-0.52	0.605
Mean	-0.1725	0.3302		

Number of observations: 40
Residuals: SS = 25.7347 (backforecasts excluded)
MS = 0.6772 DF = 38

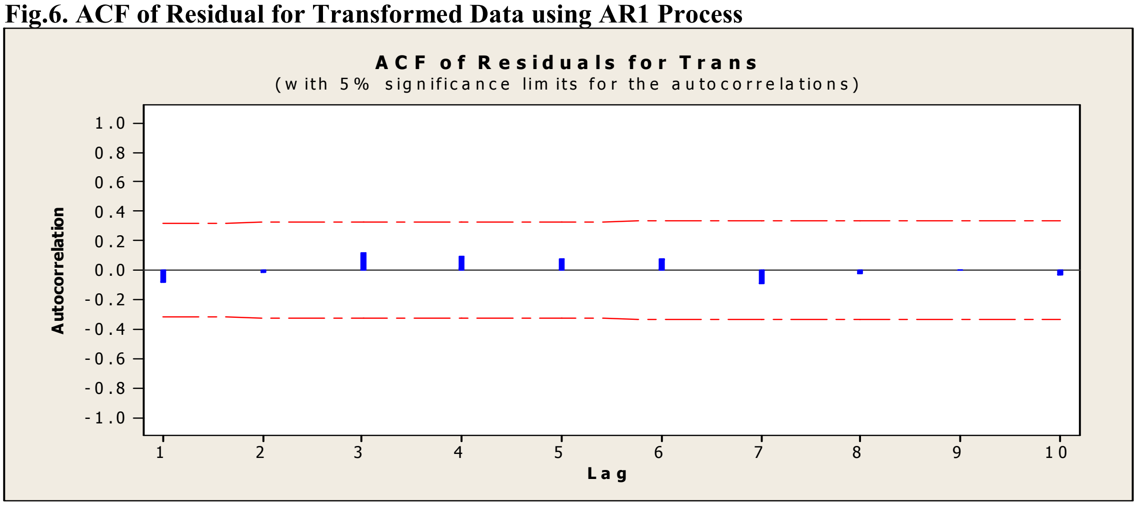
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	2.6	12.5	30.1	*
DF	10	22	34	*
P-Value	0.990	0.945	0.660	*

Interpretation

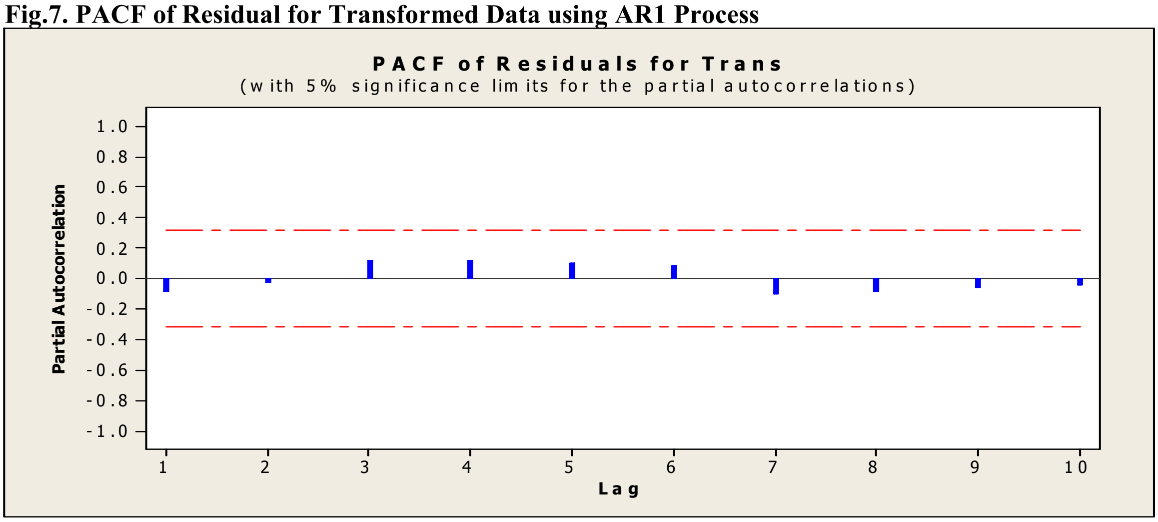
The ARMA (1,0,0) parameter (AR1) had a p-value of 0.000 which indicates that it is significant. The Ljung-Box statistics also gives significant p-values, indicating that the residuals are uncorrelated. This means that the model is adequate, but we shall fit yet another model for comparison.

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



Authour's computation and Minitab 16 output

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



Authour'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)

$$Y_t = C_t + \alpha_1 Y_{t-1} + \varepsilon_t + \lambda_1 \varepsilon_{t-1} \text{..................(3)}$$

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		

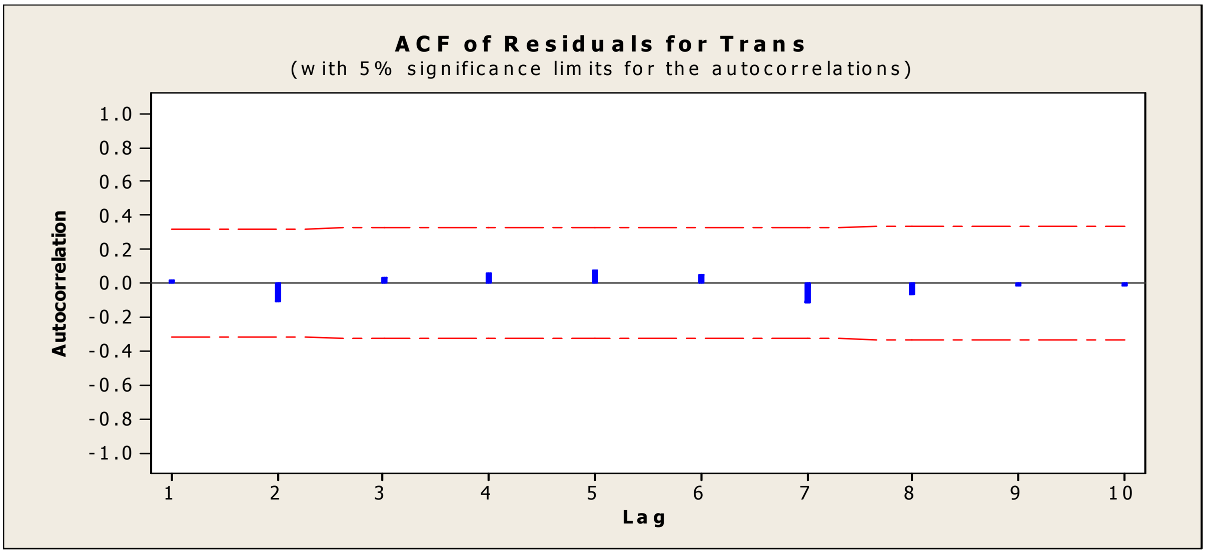
Number of observations: 40
Residuals: SS = 24.8743 (backforecasts excluded)
MS = 0.6723 DF = 37

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	*

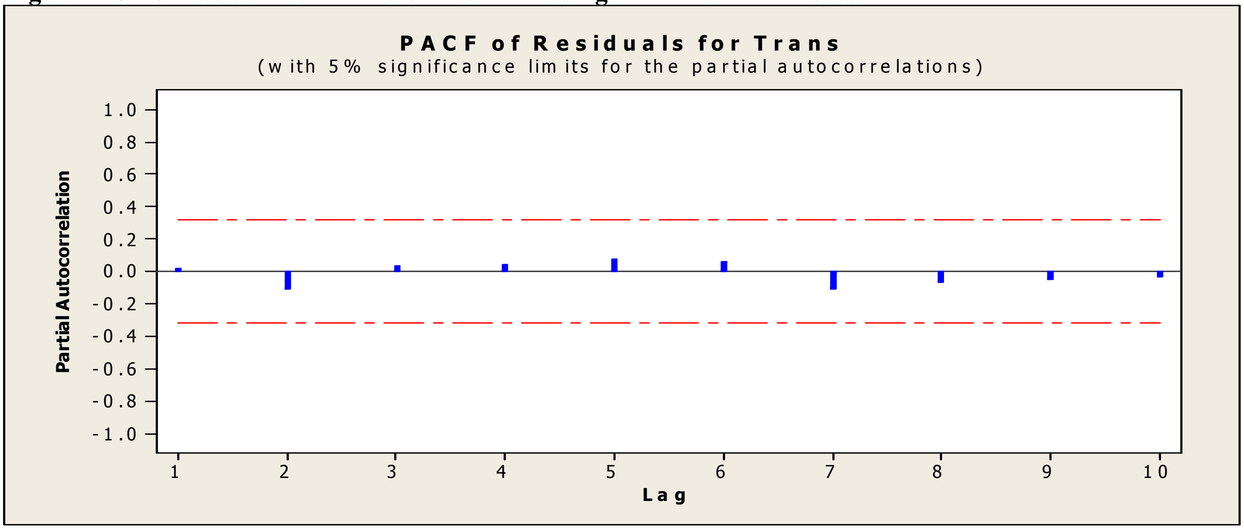
Interpretation
The ARMA(1,0,1) parameters (AR1 & MA1) have p-values of 0.000 & 0.145 respectively, which indicates that AR1 is significant while MA1 is not. However, the Ljung-Box statistics also gives significant p-values, indicating that the residuals are uncorrelated. We shall compare the models, 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



Authour’s computation and Minitab 16 output

Fig.9. PACF of Residual for Transformed Data using AR1 and MA1Process



Authour'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

MODELS	MSE
ARMA(0,0,1)	0.7757
ARMA(1,0,0)	0.6772
ARMA(1,0,1)	0.6723

Authour'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:

$$Y_t = -0.0443 + 0.8221_{y_{t-1}} - 0.3517_{\ell_{t-1}} + \ell_t \qquad \text{.............................(4)}$$

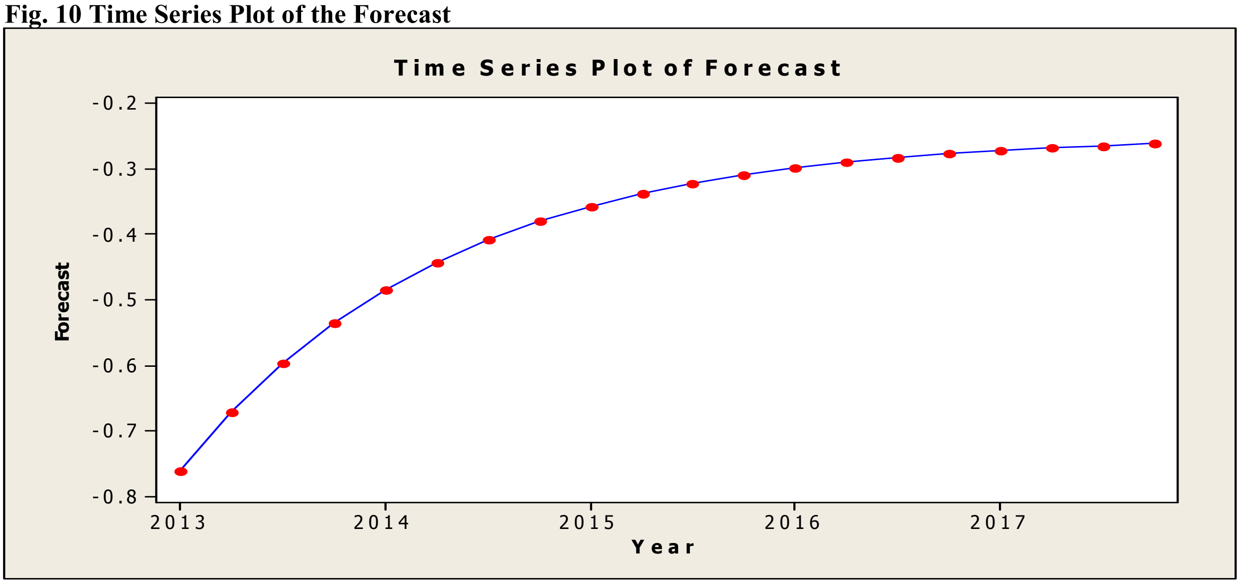
The five year quarterly forecast computed with the model is shown in table 3 b, while fig.10 shows the plot.

Table 3: Five Year Period Quarterly Forecast

Quarter	Forecast	Lower Limit	Higher Limit
2013	-0.7627	-2.3701	0.8447
2013	-0.6713	-2.4476	1.1051
2013	-0.5961	-2.4781	1.2858
2013	-0.5343	-2.4844	1.4158
2014	-0.4835	-2.4784	1.5113
2014	-0.4418	-2.4663	1.5827
2014	-0.4074	-2.4517	1.6369
2014	-0.3792	-2.4368	1.6784
2015	-0.3560	-2.4225	1.7105
2015	-0.3369	-2.4095	1.7356
2015	-0.3213	-2.3978	1.7553
2015	-0.3084	-2.3877	1.7709
2016	-0.2978	-2.3789	1.7834
2016	-0.2891	-2.3715	1.7933
2016	-0.2819	-2.3652	1.8013
2016	-0.2760	-2.3598	1.8078
2017	-0.2712	-2.3554	1.8130
2017	-0.2672	-2.3517	1.8172
2017	-0.2639	-2.3486	1.8207
2017	-0.2612	-2.3460	1.8235

Authour's computation and Minitab 16 output

Fig. 10 Time Series Plot of the Forecast



Authour's computation and Minitab 16 output

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 Authourity of Nigeria (FAAN). The analysis has helped to answer key questions posited 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

[1] Box, G.E.P., Jenkins, G.M. and Rensell, G.C (1994) Time Series Analysis: Forecasting and Control. 3rd Edition, Engle Wood Cliffs, NJ: Prentice Hall
[2] Christ, G. (1951): Time Series Analysis and Econometric Models. McGraw-Hill Inc., New York
[3] Chumbers J. (1971): Time Series Analysis and Projections. UDH Publishers, New Delhi
[4] Egbulonu K. G. (2005): Basic Econometric Methods.2nd Edition. Peace Publishers, Owerri, Nigeria
[5] Lirby M. (1966): Fundamentals of Time Series Analysis. McGraw-Hill Inc., New York
[6]Makridakis, S., Wheelwright, S.C., and Hyndman, R. J. (1998): Forecasting Methods and Applications. 3rd Edition. John Wiley and Sons, Inc, New York
[7] Mobalaji,S.S., and Wilfred, I. U. (2011). Airport Capacity Utilization in Nigeria: A Performance and Efficiency Analysis. African Journal. PP 1-12.
[8] Nwachukwu V.O. (2008): Principles of Statistical Inference. 3rd Edition, Zelon Enterprises, Port-Harcourt, Nigeria
[9] Nwobi N. F. (2003): Introductory Inference. 1st Edition, Supreme Publishers, Owerri, Nigeria
[10] Paul J. (2012): Time Series Models. 3rd Edition. John Wiley and Sons, Inc, New York