Time Series Modelling of Air Passenger Traffic Flow at Murtala Mohammed International Airport Lagos

A. Ibrahim[1]; L. Y. Ibrahim[2]

[1]Department of Statistics,
Nasarawa State University, Keffi, Nigeria.

[2]Department of Statistics,
College of Agriculture, Science and Technology, Lafia, Nigeria.
e-mail: ibrahimloko@nsuk.edu.ng[1]

***Abstract* - Over the years, there have been remarkable influence of aviation on the Nigerian economy. Murtala Muhammad International Airport (MMIA) Lagos is the busiest airport in Nigeria, accounting for over 60% of the total air passenger and aircraft movement in the country. In such an increasingly competitive aviation sector, it is imperative to make fairly accurate forecasts so as to enable long-term planning, short-term planning and decisions regarding infrastructure development, flight networks and effective management. In this study, Artificial Neural Network (ANN), Seasonal Auto-Regressive Integrated Moving Average (SARIMA) and Holt-Winters Exponential Smoothing (HWES) models were used to model air passenger traffic flow in MMIA. The performances of these proposed models were compared in-sample and out-of-sample performance by employing static forecast procedure over January 2007 to December 2019 forecast horizon. The best models from the [SARIMA, ANN and HWES] were selected by employing some performance metrics comprising, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) and residual diagnostics. The selected models forecasting performances were compared using the statistical loss functions, Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) for the measurement of forecast accuracy. Results show that ANN outperforms the other models in the domestic sector, while the HWES had the best performance in the international sector even though it was outperformed by SARIMA in the domestic sector. ANN yielded the best in-sample performance for domestic and international air passenger traffic. It was concluded that the ANN, which represents a class of non-linear time series model was very efficient in mimicking time series patterns and giving good forecasts.**

Keywords: *SARIMA, ARIMA, Artificial Neural Network, Akaike Information Criterion, Bayesian Information Criterion.*

I. Introduction

Some decades ago, Aviation in Nigeria has been recognized as being a key sector of the economy. It is one of the indices for measuring the development of a country. The importance of this sector to the economy of Nigeria cannot be overemphasized. The effect of this which is typically characterized by business cycles or alternation between periods of economic growth and downturns, brings about our data series exhibiting cyclical patterns or seasonality. Furthermore, every economy is highly susceptible to a variety of shocks of different nature (economic, political, climatic etc), which are likely to modify the past trends and the volatility in the data (Bougas, 2013). The International Air Transport Association (IATA) released the IATA airline forecast for 2013-2017 showing that airline expect to see a 31% increase in passengers flow between 2012 and 2017. By 2017 the total passenger numbers are expected to rise to 3.91billion- an increase of 930 million passengers over the 2.98billion carried out in 2012.

The IATA airline industry forecast 2013-2017 is a consensus outlook for system-wide passenger growth. Demand is expected to expand by an average of 5.4% Compound Annual Growth Rate (CAGR) between 2013 and 2017. By comparison, global passenger growth expanded by 4.3% CAGR between 2008 and 2012, largely reflecting the negative impact of the 2008 global financial crisis and recession. Of the new passengers, approximately 292 million were carried on international routes and 638 million on domestic routes.

The report further stated that, emerging economies of the Middle-East and Asia Pacific will see the strongest international passenger growth with CAGR of 6.3% and 5.7% followed by Africa and Latin America with CAGR of 5.3% and 4.5%. (IATA, 2013).

Predicting future air passenger traffic flow is important as it allows air transport authorities to adapt necessary infrastructures and offers airline companies the capacity to match the increasing passenger demand for air transportation. There are a great number of studies, particularly since the 1990's that use time series models to forecast air passenger numbers. The forecasting performance of each model varies depending on the country of the passengers, the type of flight considered (domestic or international), the performance measure and the forecasting horizon (Emiray and Rodriguez, 2003). No methodological approach has been found to always dominate another in terms of forecasting performance (Shen *et al.,* 2011). However, one model that has been successfully used in its various applications is the ANN. The ANN model is a mathematical model inspired by the function of the human brain and its use is mainly motivated by its capability of approximating any Borel-measurable function to any degree of accuracy (Hornik*et al.,*1989). This study provides an evaluation of three time series forecasting models performance based on air passenger traffic movement. These comprise an Autoregressive Moving Average Model known as Seasonal Auto – Regressive Integrated Moving Average (SARIMA), ANN model and the HWES model.

Lagos Airport also known as MMIA (renamed in the mid 1970s, after a former Nigerian Military Head of state, General Murtala Muhammed) is an international airport providing world class flight services both in Nigeria and elsewhere. MMIA consists of an international and a domestic terminal located at about one kilometer apart from each other. The international terminal was modeled after Amsterdam's Airport Schiphol and was opened officially on 15thMarch 1979. A new domestic privately funded terminal known as MMA2 has been constructed and was commissioned on 7th April 2007. In 2009 alone the airport served 5,644,572 passengers.

The airport houses the headquarters of the Federal Airport Authority of Nigeria (FAAN) as well as the Head Office of the Accident Investigation Bureau (AIB). The Lagos office of the Nigerian Civil Aviation Authority (NCAA) is located in Aviation House within the airport. MMIAis believed to be the busiest airport in Nigeria rendering services internationally and locally. In recent years, the domestic and overseas passenger traffic has risen steadily at an average of 10% per annum and being the Nation's main gateway, it accounts for over 60% of the total passenger and aircraft movement in the country.

Over the years, the demand for air transportation in Nigeria has greatly increased. This is indicated by increase in air passenger traffic for both domestic and international flights, emergence of more airlines, creation and expansion of terminals. Also, with the rate of increase in passenger traffic and increasing influence of Aviation on the Nigerian economy, it is important that a time series models are accurately selected for proper management in the airlines, airports and the Nigerian Airspace. Many researchers adopt the more common time series approach in predicting air passenger traffic. In some cases, salient characteristics in the time series data might not be captured due to the simplicity of the adopted models. However, there is a need to apply the conventional methods and compare it with a more robust approach in modelling the air passenger traffic so as to improve forecast accuracy.

Presently, there are many time series models used in forecasting, some of which include; ANN, Harmonic Regression (HR), HWES as well as the class of model known as the Auto-Regressive Moving Average (ARMA) model, which consist of the Autoregressive Integrated Moving Average (ARIMA) model, the Seasonal Autoregressive Integrated Moving Average (SARIMA) model and Grey Model among others; the Autoregressive Conditional Heteroscedasticity(ARCH) model (Engle,1982) and all of its extensions is often applied in forecasting of financial time series. As for multivariate models, the Vector Auto-Regressive (VAR) model and the Error Correction Model (ECM) are also common (Artis and Zhang1990).

Forecasting performance of all these time series models vary with the forecasting horizon and depend on the adequate detection of seasonal roots.

In view of this, this study aimed at comparing the forecast performances of time series models in approximately predicting the air passenger traffic flow in MMIA Lagos, Nigeria. This is achieved through by Fitting three-time series models (ANN, SARIMA and HWES); Evaluating the accuracy of the proposed models using some performance criteria comprising Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC) using monthly forecast of air passenger flow at MMIA; Evaluate the out-of-sample forecast performance of these models using some classical loss functions such as MAPE and RMSE and be able to say which among them is better than the others.

The empirical study only covers thirteen years, between January 2009 and December 2019, this is due to insufficient data from the primary source to cover the desired number of years for the research work. Only few models will be used for the forecast. Most factors responsible for air passenger traffic flow such as population and economic activities will not be measured hence the analyses will be based on the available data to be obtained from the airport. The study does not exclusively check the evaluation of the models' performances on each airline.

II. METHODOLOGY

This study adopts a descriptive research method and using time series analysis to make adequate model selection and predictions. Box-Jenkings Seasonal Auto-Regressive Integrated Moving Average Model, the Holt-Winters Exponential Smoothing Model and the Artificial Neural Network Approach was used in modelling the air passenger traffic flow in Murtala Muhammad International Airport for domestic and international flights.

The population of the study comprises of all air passenger traffic flow for international and domestic flights in Nigeria over the past 5 decades. The research used a convenient sampling to select the monthly air passenger traffic flow for international and domestic flights over the period 13 years from January 2007 to December 2019 in Murtala Muhammad International Airport Lagos. This was obtained from the Nigerian Airspace Management Agency (NAMA).

The data collected for this study is from a secondary source. The data used in this study are the monthly air passenger traffic flow for international and domestic flights over the period of 13 years from January 2007 to December 2019. This was obtained from the Nigerian Airspace Management Agency (NAMA).

The domestic air passenger traffic comprises of local passenger arrival and departure to and from Murtala Muhammed International Airport whereas the international air passenger traffic comprises of air passengers from international flights from and to other countries other than Lagos, Nigeria. The data set consists of 156 monthly observations of air passenger traffic flow for domestic flights and 156 monthly observations for international flights.

i. Technique for Data Analysis

The method of data analysis used for this research study is time series analysis. Time series data is defined as a collection of values of a variable that differs over time. The intervals between observations of a time series can vary. However, the range of the intervals should be consistent throughout the observed period e.g. daily, weekly, monthly etc. In general, the time series is assumed to be stationary in empirical work based on time series (Gujarati & Porter 2008).

ii. Concept of Time Series Model

Time series data is defined as a collection of values of a variable that differs over time. The intervals between observations of a time series can vary. However, the range of the intervals should be consistent throughout the observed period e.g. daily, weekly, monthly etc. In general, the time series is assumed to be stationary in empirical work based on time series (Gujarati & Porter 2008).

iii. Stochastic Processes

A process is said to be stochastic, or random, if the collection of a variable is gathered over a sequence of time. A stochastic process can be either stationary or nonstationary (Gujarati & Porter 2008).

iv. Autoregressive Model

An autoregressive model is a model where the dependent variable is regressed on at least one lagged period of itself. If an autoregressive model includes one lagged period of itself, it follows a first-order autoregressive stochastic process, denoted AR(1). Furthermore, if the model includes p number of lagged periods of the dependent variable, it follows a pth-order autoregressive process, denoted AR(p) (Gujarati & Porter 2008).

v. Stationary Process

There are different types of stationarity. Second order stationary, commonly known as weakly stationary, is considered to be sufficient in most empirical works. A stochastic process is weakly stationary if it has constant mean and variance and the covariance is time invariant, i.e. the statistics do not change over time (Gujarati & Porter 2008).

A white noise process is a special type of stationary stochastic process. A stochastic process is considered to be white noise if the mean is equal to zero, the variance is constant, and the observations are serially uncorrelated (Gujarati & Porter 2008).

vi. Nonstationary Process

A stochastic process that has a time-varying mean, variance, or covariance is said to be nonstationary. Financial data usually follows a random walk which is a type of nonstationary stochastic process. A random walk is either with or without drift, indicating the presence of an intercept, and is an AR(1) process. Regressing Y_t on Y_{t-1} estimates the following:

$$Y_t = \rho Y_{t-1} + u_t \quad (3.1)$$

and if ρ equals 1, the model becomes what is known as a random walk (Gujarati & Porter 2008).

A random walk without drift is a process where the dependent variable can be estimated on one lagged period of itself plus an error term, assumed to be white noise, known as a random shock. The formula for a random walk without drift excludes the intercept. The mean is constant over time in a random walk without drift, however, the variance is increasing indefinitely over time, making it a nonstationary stochastic process (Gujarati & Porter 2008).

Random walk without drift:

$$Y_t = Y_{t-1} + u_t \qquad (3.2)$$

Similar to a random walk without drift, a random walk with drift is a process where the variable is dependent on its own lagged values and a random shock. However, the model that may be used to estimate a random walk with drift includes an intercept known as the drift parameter, denoted by δ. This parameter indicates if the time series is trending upwards or downwards, depending on whether δ is positive or negative. A random walk with drift is a nonstationary stochastic process since the mean and variance are increasing over time (Gujarati & Porter 2008).

Random walk with drift:

$$Y_t = \delta + Y_{t-1} + u_t \qquad (3.3)$$

The preceding random walks have infinite memory which means that the effects of random shocks persist throughout the whole time period. The random walks are known as difference stationary processes, meaning that even though the stochastic processes are nonstationary, they become stationary through the first order difference (Gujarati & Porter 2008).

vii. **Integrated Process**

A nonstationary stochastic process that has to be differenced one time to become stationary, is said to be integrated of the first order, denoted $I(1)$. Likewise, a nonstationary stochastic process that has to be differenced twice to become stationary, is said to be integrated of the second order, denoted $I(2)$. Furthermore, this means that a nonstationary stochastic process that has to be differenced d times, is said to be integrated of order d, denoted $Y_t \sim I(d)$. A time series that is stationary without any differencing is integrated of order zero, denoted $Y_t \sim I(0)$ (Gujarati & Porter 2008).

viii. **Deterministic Trend**

A time series that is deterministic can be perfectly forecasted. However, most time series are partially deterministic and partially stochastic, making them impossible to predict perfectly due to the probability distribution of future values (Chatfield 2003).

If a variable is dependent on its past values and a time variable, it is estimated by the following;

$$Y_t = \beta_1 + \beta_2 t + y_t + u_t \qquad (3.4)$$

where t is a variable that measures time chronologically and u_t is an error term, assumed to be white noise. The equation is known as a random walk with drift and deterministic trend and is stochastic but also partially deterministic, due to the time trend t (Gujarati & Porter 2008).

ix. **Modeling of Time Series Data**

When working with forecasting of time series data, the underlying time series is assumed to be stationary. Assuming stationarity, there are several different approaches to construct forecasting models, for example an autoregressive process, a moving average process, an autoregressive and moving average process, and an autoregressive integrated moving average process (Gujarati & Porter 2008).

x. **Autoregressive Process**

An autoregressive process may be used to forecast a time series. As mentioned earlier, a first-order autoregressive model is denoted AR(1) and is Y_t regressed on Y_{t-1}. An autoregressive model of the pth-order is denoted AR(p) and takes the form of

$$Y_t = \delta_1 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + u_t \qquad (3.5)$$

where the constant is denoted by δ and u_t is white noise (Gujarati & Porter 2008).

xi. **Moving Average Process**

In a moving average process, the dependent variable is regressed on current and lagged error terms and is therefore estimated through a constant and a moving average of the error terms. If the dependent variable is regressed on the current and one lagged error term, it follows a first-order moving average process, denoted MA(1). Moreover, a model that includes q number of error terms follows a qth-order moving average process, denoted MA(q). A MA(q) process is defined as

$$Y_t = \theta + \alpha_1 y_{t-1} + \beta_0 u_t + \beta_1 u_{t-1} \qquad (3.6)$$

where the error terms u are assumed to be white noise and μ is the constant (Gujarati & Porter 2008). In a MA model the error terms are usually scaled to make β_0 equal to one (Chatfield 2003).

xii. **Autoregressive and Moving Average Process**

It is possible to combine an autoregressive process and a moving average process since the dependent variable often possess characteristics of both. This is called an autoregressive and moving average process, or ARMA. If both of the underlying AR and MA models are of the first-order, the model is denoted ARMA(1, 1) and defined as

$$Y_t = \theta + \alpha_1 y_{t-1} + \beta_0 u_t + \beta_1 u_{t-1} \qquad (3.7)$$

where θ is the constant. If the underlying autoregressive model is of order p and the moving average model is of order q, the ARMA process is denoted by ARMA(p, q) (Gujarati & Porter 2008).

xiii. **Autoregressive Integrated Moving Average Process**

If the time series of an ARMA model has to be differenced a certain number of times to become stationary, the model becomes what is known as an autoregressive integrated moving average model, or an ARIMA model. As mentioned previously, a time series which has to be differenced d number of times in order to become stationary, is integrated of order d, denoted $I(d)$. In its general form, the ARIMA model is denoted ARIMA(p, d, *q*) which means that the AR is of the *p*th-order, the time series is integrated *d* number of times, and the moving average is of the *q*th-order. This further means that if the underlying AR and MA models are of the first-order, and the time series is stationary at the first difference, the ARIMA model is denoted ARIMA(1, 1, 1). It is important to note that an ARIMA model is not derived from any economic theory, that is, it is an atheoretic model. The Box-Jenkins methodology can be followed to determine p, d, and q and estimate an ARIMA model (Gujarati & Porter 2008).

Autoregressive AR(p) Models A pure AR(p) process may be represented as follows, where X_t is modelled as lagged Values of itself plus a 'white noise' error term

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \ldots + \varphi_p X_{t-p} + \varepsilon_t \quad (3.8)$$

This may be alternatively written as

$$\varphi(B)\, X_t = \varepsilon_t \quad (3.9)$$

Where $\emptyset(B)$ is a p-order polynomial in the backshift operator, which is equal to:

$$(1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^2) \quad (3.10)$$

and B is the backshift operator, such that

$$B^0 = X_t, B^1 X_t = X_{t-1}, B^2 X_t = X_{t-2}, \ldots \quad (3.11)$$

A useful way of gaining insight into univariate processes is to consider their autocorrelation and partial autocorrelation functions (ACF and PACF). The ACF measures the ratio of the covariance between observations k lags apart and the geometric average of the variance of observations (i.e., the variance of the process if the process is stationary, as:

$$V(X_t) = V(X_{t-k}) \quad (3.12)$$

$$\rho_k = \frac{cov(X_t, X_{t-k})}{V(X_t).V(X_{t-k})^{1/2}} \quad (3.13)$$

However, some of the observed autocorrelation between X_t and X_{t-k} could be due to both being correlated with intervening lags. The PACF seeks to measure the autocorrelation between X_t and X_{t-k}correcting for the correlation with intervening lags. For example, consider an AR(1) process of the form $X_t = 0.8X_{t-1} + \epsilon_t$. The first order autocorrelation coefficient is the autocorrelation coefficient for the second lag is 0.64, although the partial autocorrelation coefficient for the second lag is zero, as the process is an AR (1) process. In other words, the autocorrelation between observations two lags apart is due only to the correlation between observations one lag apart which feeds through into the second lag. As the lag length increases the autocorrelation coefficient declines (at lag length k the autocorrelation coefficient is $(0.8)^k$ the PACF is calculated as the partial regression coefficient, φ_{kk} in the kth order Autoregression

$$X_t = \varphi_{k1} X_{t-1} + \varphi_{k2} X_{t-2} + \ldots + \varphi_{kk} X_t + \varepsilon_t \quad (3.14)$$

Thus, for an AR(p) process, $\varphi_{kk} = 0, \forall k > p$ (3.15)

Some general properties of the ACF and PACF for AR processes can be observed by considering a simple AR(1) process given as

$$X_t = \varphi_1 X_{t-1} + \varepsilon_t \quad (3.16)$$

Note that the AR(1) model can be written as an infinite length MA process, providing $\emptyset_1 < 1$ Denote the AR(1) series as,

$$(1 - \varphi_1 B)\, X_t = \varepsilon_t \quad (3.17)$$

which gives,

$$X_t = (1 - \varphi_1 B)^{-1} \varepsilon_t \quad (3.18)$$

which upon expansion and providing $\varphi_1 < 1$ yields

$$\text{x}X_t = \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_1^2 \epsilon_{t-2} + \ldots + \quad (3.19)$$

This result holds more generally so that any finite order stationary AR process may be expressed as an infinite order MA process. This duality between AR and MA processes is an important property which can often be exploited when attempting to identify ARMA models.

For the AR(1) process the value of the ACF at lag k is given by φ_1^k

The value of the autoregressive coefficient can yield some insight into the underlying data generating process.

For example, higher values of $\emptyset_1$ indicate a higher degree of persistence in the series. A negative autoregressive component indicates a process which oscillates around its mean value. For more general AR(p) models, the behavior of the process is determined by the solution to the p-order polynomial $(1 - \varphi_1 B - \varphi_2 B^2 - \ldots - \varphi_p B^p)$ given by

$$\varphi(B) = (1 - g_1 B)(1 - g_2 B)\ldots(1 - g_p B) = 0 \quad (3.20)$$

For the process to be stationary it is a necessary and sufficient condition for the roots of the p-order polynomial to lie outside the unit circle.

Moving Average MA(q) Models

An MA(q) process may be represented as follows, where X_t is modelled as the weighted average of a 'white noise' series,

$$X_t = \varepsilon_t + \varphi_1 \varepsilon_{t-1} + \varphi_2 \varepsilon_{t-2} + \ldots + \varphi_q \varepsilon_{t-q} \quad (3.21)$$

or alternatively

$$X_t = \theta(B)\, \varepsilon_t \quad (3.22)$$

where ϵ_t is a white noise process,

$$\varphi B = 1 + \varphi_1 B + \varphi_2 B^2 + \ldots + \varphi_q B^q \quad (3.23)$$

is a q-order polynomial in the backshift operator

Note that the expected value of X_t equals zero. Furthermore, the autocorrelation between X_t and X_{t+k} equals zero for k greater than q. Thus the order of the MA process, q, Indicates the 'memory' of the process. All MA processes are stationary, regardless of the Coefficients of the model. However, to ensure inevitability of the model (i.e., that the finite Order MA process can be written in terms of a stationary infinite order AR process) the roots Of the MA polynomial must lie outside the unit circle. MA models can be particularly useful for representing some economic time series as they can handle random shocks such as strikes, Weather patterns, etc.

The Box-Jenkin's Modelling Procedure

The stages in Box-Jenkins procedure are:

xiv. **Identification:** the series is differenced if necessary to make it stationary. Thus, the sample ACF (Autocorrelation Function) and PACF (partial autocorrelation function) are calculated.

The behavior of ACF and PACF determine the number of autoregression parameter (p)

xv. **Estimation:** least square estimate of the process parameters are generated.

xvi. **Diagnostic checking**: the residual from the estimated model should be like a random series failing that further analysis of the residual leads to a re-specification of the model.

xvii. **Forecasting:** the fitted model having first been integrated if necessary is used to forecast the y_t^s

Identification Phase

The following brief summary is based on practical recommendations of Pankratz (1983)

1. One autoregressive (p) parameter: ACF – exponential decay; PACF – spike at lag 1
2. Two autoregressive (p) parameter: ACF – a sine-wave

shape pattern or a set of exponential decays; PACF – spikes at lags 1 and 2

3. One moving average (q) parameter: ACF – spikes at lag 1, PACF – damps out exponentially.
4. Two moving average (q) parameters: ACF – spikes at lags 1 and 2, PACF – a sine-wave shape pattern or a set of exponential decays
5. One autoregressive (p) and one moving average (q) parameter: ACF – exponential decay starting at lag 1, PACF- exponential decay starting at lag 1 (Peter E. O, 2016)

Autoregressive Integrated Moving Average ARIMA (p,q) Models

An ARIMA (p,q) series may be represented as

$$X_t - \varphi_1 X_{t-1} - \varphi_2 X_{t-2} - \dots - \varphi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad (3.24)$$

or alternatively

$$\sum_{i=0}^{p} \varphi_1 X_{t=i} \sum_{j=0}^{q} \varphi_{j\varepsilon_{t-j}} \quad (3.25)$$

where, $\emptyset_0$ and $\emptyset_0$ are equal to 1, or more compactly

$$\varphi(B)X_t = \varphi(B)\varepsilon_t \quad (3.26)$$

Using mixed ARMA models can be useful as it should usually be possible to represent a time series satisfactorily using fewer parameters than might be required with a pure AR or pure MA models.

Seasonal ARMA Models

Seasonal data may be also modelled. The numbers of seasonal AR and MA terms are usually denoted by P and Q respectively. Thus, a general seasonal ARMA model may be

Represented as;

$$\varphi(B)\phi(B)X_t = \theta(B)\Theta(B)\varepsilon_t \quad (3.27)$$

where:

$$\varphi(B) = 1 - \phi_{1s}B^{1s} - \phi_{2s}B^{2s} - \dots - \phi_{ps}B^{ps}, \quad (3.28)$$

$$\Theta(B) = 1 + \theta_{1s}B^{1s} + \theta_{2s}B^{2s} + \dots + \theta_{Qs}B^{Qs}, \quad (3.29)$$

and S is the seasonal span, hence quarterly data S = 4 and for monthly data S = 12.

xviii. **Artificial Neural Networks (ANNs)**

Artificial neural networks (ANNs) approach has been suggested as an alternative technique to time series forecasting and it gained immense popularity in last few years. The basic objective of ANNs was to construct a model for mimicking the intelligence of human brain into machine Similar to the work of a human brain, ANNs try to recognize regularities and patterns in the input data, learn from experience and then provide generalized results based on their known previous knowledge. Although the development of ANNs was mainly biologically motivated, but afterwards they have been applied in many different areas, especially for forecasting and classification purposes.

First, ANNs are data-driven and self-adaptive in nature. There is no need to specify a particular model form or to make any *a priori* assumption about the statistical distribution of the data; the desired model is adaptively formed based on the features presented from the data. This approach is quite useful for many practical situations, where no theoretical guidance is available for an appropriate data generation process. Second, ANNs are inherently non-linear, which makes them more practical and accurate in modeling complex data patterns, as opposed to various traditional linear approaches, such as ARIMA methods. There are many instances, which suggest that ANNs made quite better analysis and forecasting than various linear models.

Finally, as suggested by Hornik and Stinchcombe (1987), ANNs are universal functional approximators.

They have shown that a network can approximate any continuous function to any desired accuracy. ANNs use parallel processing of the information from the data to approximate a large class of functions with a high degree of accuracy. Further, they can deal with situation, where the input data are erroneous, incomplete or fuzzy.

xix. **The ANN Architecture**

The most widely used ANNs in forecasting problems are multi-layer perceptrons (MLPs), which use a single hidden layer feed forward network (FNN) The model is characterized by a network of three layers, viz. input, hidden and output layer, connected by acyclic links. There may be more than one hidden layer. The nodes in various layers are also known as processing elements. The three-layer feed forward architecture of ANN models can be diagrammatically depicted as below:

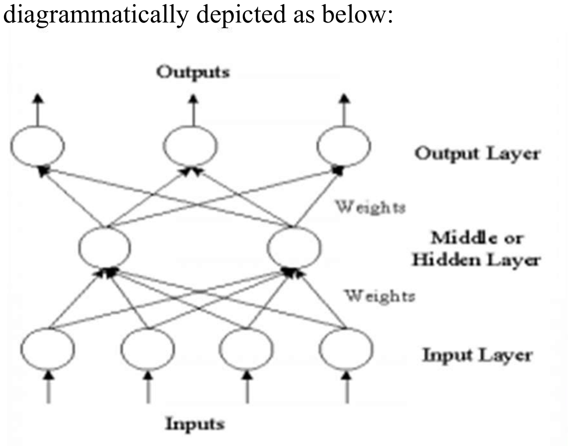


Fig. 3.1 The three-layer feed-forward ANN architecture

The output of the model is computed using the following mathematical expression

$$Y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j\, g\left[\beta_{0j} + \sum_{i=1}^{p} \beta_{ij}\, y_{t-1}\right] + \varepsilon_t \; \forall_t \qquad (3.30)$$

Here y_{t-i} (i =1,2,..., p) are the p inputs and y_t is the output. The integers p, q are the number of input and hidden nodes respectively. α_j (j = 0,1,2,...,)q and β_{ij} (i = 0,1,2,..., p; j = 0,1,2,...,)q are the connection weights and ε_t is the random shock; α_0 and β_{0j} are the bias terms. Usually, the logistic sigmoid function $g(x) = \frac{1}{1+e^{-x}}$ is applied as the nonlinear activation function. Other activation functions, such as linear, hyperbolic tangent, Gaussian, etc. can also be used. The feed forward ANN model in fact performs a non-linear functional mapping from the past observations of the time series to the future value, i.e. $y_t = f(y_{t-1}, y_{t-2}, \ldots . y_{t-p}, \mathbf{w}) + \varepsilon_t$, where **w** is a vector of all parameters and f is a function determined by the network structure and connection weights. To estimate the connection weights, non-linear least square procedures are used, which are based on the minimization of the error function.

$$f(\Psi) = \sum_t e_t^2 = \sum_t (y_t - \hat{y}_t)^2 \qquad (3.31)$$

Here Ψ is the space of all connection weights.

The optimization techniques used for minimizing the error function are referred as *Learning Rules*. The best-known learning rule in literature is the *Backpropagation* or *Generalized Delta Rule*.

Holt-Winters Exponential Smoothing Model

The data used in this study consist of trend and seasonal component. It is however appropriate to apply necessary smoothing technique to model the data used. Smoothing can be seen as a technique to separate the signal and the noise as much as possible and in that a smoother acts as a filter to obtain an "estimate" for the signal (Montgomery *et al*., 2008).

Actual signal Actual data
Estimated

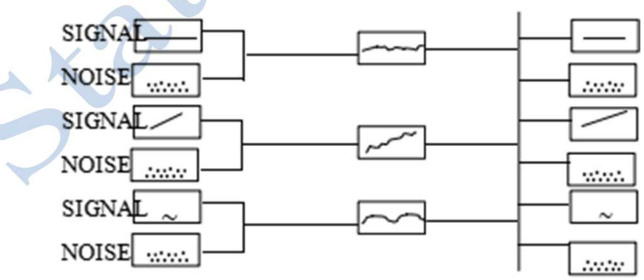


Fig. 3.2 Presence of signal and noise

The fig. 3.2 above shows the presence of signal and noise in the actual data. The signal represents any pattern caused by the intrinsic dynamics of the process from which the data is collected and it can assume various forms. Exponential Smoothing could be Single Exponential Smoothing, Double Exponential Smoothing and Triple Exponential Smoothing.

xx. **Single Exponential Smoothing**

The single exponential smoothing is also referred to as simple exponential smoothing. It assumes that the data fluctuates around a reasonably stable mean. The model is given below;

$$S_{t+1} = \alpha y_t + (1-\alpha) S_t \;\; 0 < \alpha \leq 1, t > 0 \qquad (3.32)$$

Each successive observation in the series that the above is applied to gives each new smoothed value computed as the weighted average of the current observation and the previous smoothed observation. The weights being applied to get each smoothed value decrease exponentially depending on the value of the parameter ∝. New forecast is previous plus an error adjustment; this can be written as:

$$S_{t+1} = S_t + \alpha\varepsilon_t \quad (3.33)$$

Where ε_t is the forecast error for period *t*. However single exponential smoothing is not effective when there is a trend. The single parameter ∝ does not accommodate this.

Double Exponential Smoothing

The single exponential smoothing has only one constant,∝ as indicated in the equation above, which brings about the limitation in handling the presence of trend. However, this situation is improved in the double exponential smoothing by the introduction of another equation with additional parameter, a second constant is shown in the equations below

$$S_t = \alpha y_t + (1-\alpha)(S_{t-1} + b_{t-1})\ 0 < \alpha < 1 \quad (3.35)$$

$$b_t = y(S_t - b_{t-1})\ 0 < \alpha < 1 \quad (3.36)$$

The current value of the series is used to calculate its smoothed value replacement in double exponential smoothing. There are several methods of setting the initial values for S_t and b_t $S_1$is in general set to y_1. For b_1 the following could be adopted

$$b_2 = y_2 - y_1 \quad (3.37)$$

$$b_1 = \frac{[(y_2-y_1)+(y_3-y_2)+(y_5-y_4)]}{3} \quad (3.38)$$

$$b_1 = \frac{y_n-y_1}{3} \quad (3.39)$$

III. Results and Discussion

3.1 Descriptive Statistics

It is essential, before estimating the models that an exploration of the data has to be carried out so as to identify the statistical properties of the data and the unforeseen outliers which could impede a good analysis. The Jarque Bera test is used to test for normality.

Table 4.3 Descriptive Statistics for Domestic and International air passenger traffic.

	Domestic	*International*
Mean	264046.6346	193869.4359
Standard Error	6502.076071	4910.678026
Median	274695	189003
Mode	338857	257722
Standard Deviation	81210.9041	61334.34889
Sample Variance	6595210945	3761902354
Kurtosis	-1.333351207	-1.09850549
Skewness	-0.043720811	0.076022309
Range	341102	226354
Minimum	100335	85372
Maximum	441437	311726
Sum	41191275	30243632
Count	156	156
Confidence Level(95.0%)	12844.1174	9700.490178

Table 4.3 shows a record of raw data on the air passenger traffic for domestic and international flight. The highest recorded passenger movement was obtained from the domestic flight while the lowest air passenger flow record was from international flight. Also, it could be observed from the table that averagely, over the years there have been more air passenger flow domestically than internationally.

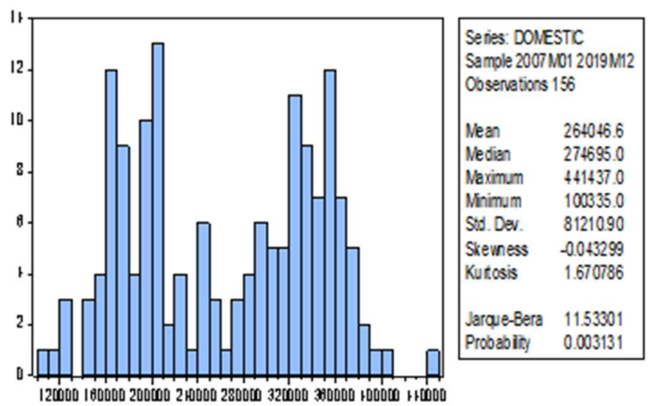


Figure 1: Domestic Air Passenger Traffic Histogram

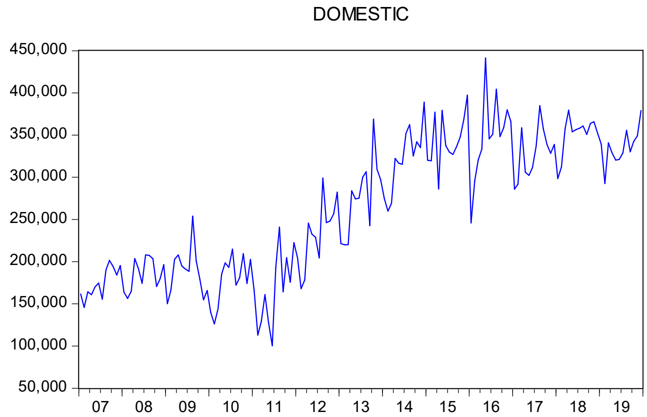


Figure 2 : *Time Plot of Domestic Air Passenger Traffic*

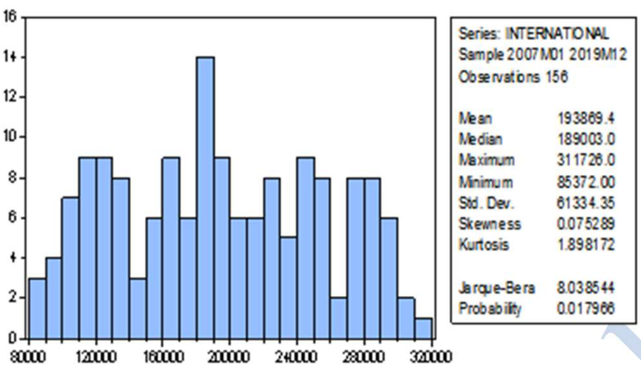


Figure 3: *International Air Passenger Traffic Histogram*

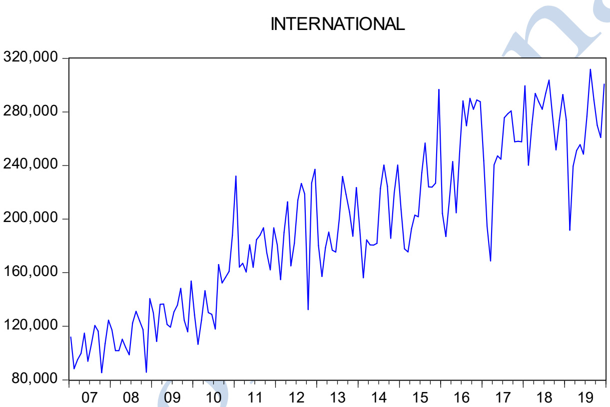
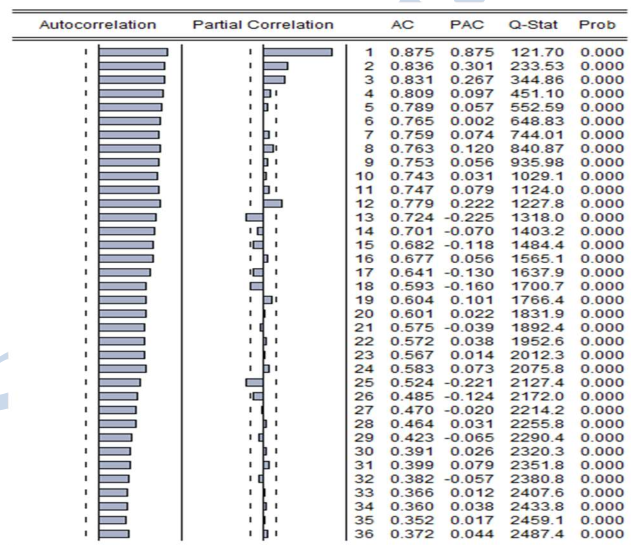


Figure 4: *Time Plot of International Air Passenger Traffic in Lagos*

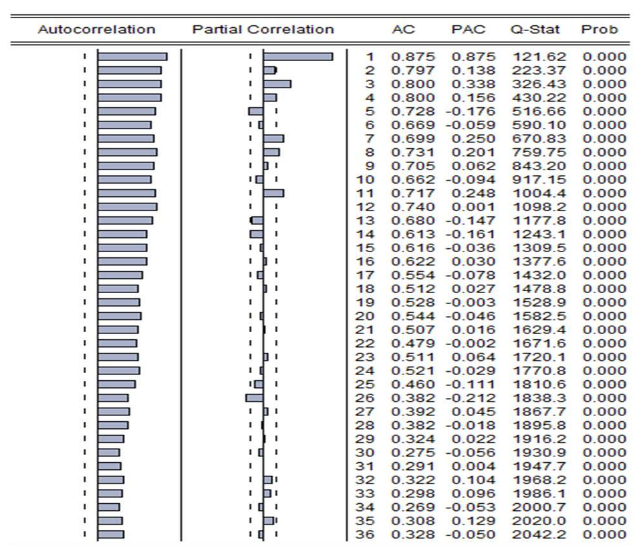
The time plots of the domestic air passenger traffic and international air passenger traffic above show that the time series data sets are not stationary. By visual observation, there is evidence of an upward trend in the air passenger traffic in both cases. It would be necessary to achieve stationarity, through differencing, before embarking on the Box Jenkins methodology.

Looking at the graph for domestic air passenger traffic, a dip in passenger traffic flow could be observed between the year 2009 and 2011. This is a very sad period for the Nigerian aviation sector because it was marked with series of major commercial air crashes and disaster, hence the loss of trust in the sector. The aviation sector experienced major overhaul which immensely affected airlines. After that period it has been a steady increase in air passenger traffic both internationally and locally. Airlines were healthier and new airlines emerged, bringing about healthy competition.



	AC	PAC	Q-Stat	Prob
1	0.875	0.875	121.70	0.000
2	0.836	0.301	233.53	0.000
3	0.831	0.267	344.86	0.000
4	0.809	0.097	451.10	0.000
5	0.789	0.057	552.59	0.000
6	0.765	0.002	648.83	0.000
7	0.759	0.074	744.01	0.000
8	0.763	0.120	840.87	0.000
9	0.753	0.056	935.98	0.000
10	0.743	0.031	1029.1	0.000
11	0.747	0.079	1124.0	0.000
12	0.779	0.222	1227.8	0.000
13	0.724	-0.225	1318.0	0.000
14	0.701	-0.070	1403.2	0.000
15	0.682	-0.118	1484.4	0.000
16	0.677	0.056	1565.1	0.000
17	0.641	-0.130	1637.9	0.000
18	0.593	-0.160	1700.7	0.000
19	0.604	0.101	1766.4	0.000
20	0.601	0.022	1831.9	0.000
21	0.575	-0.039	1892.4	0.000
22	0.572	0.038	1952.6	0.000
23	0.567	0.014	2012.3	0.000
24	0.583	0.073	2075.8	0.000
25	0.524	-0.221	2127.4	0.000
26	0.485	-0.124	2172.0	0.000
27	0.470	-0.020	2214.2	0.000
28	0.464	0.031	2255.8	0.000
29	0.423	-0.065	2290.4	0.000
30	0.391	0.026	2320.3	0.000
31	0.399	0.079	2351.8	0.000
32	0.382	-0.057	2380.8	0.000
33	0.366	0.012	2407.6	0.000
34	0.360	0.038	2433.8	0.000
35	0.352	0.017	2459.1	0.000
36	0.372	0.044	2487.4	0.000

Figure 5*: Correlogram of the Domestic Air Passenger Traffic*



	AC	PAC	Q-Stat	Prob
1	0.875	0.875	121.62	0.000
2	0.797	0.138	223.37	0.000
3	0.800	0.338	326.43	0.000
4	0.800	0.156	430.22	0.000
5	0.728	-0.176	516.66	0.000
6	0.669	-0.059	590.10	0.000
7	0.699	0.250	670.83	0.000
8	0.731	0.201	759.75	0.000
9	0.705	0.062	843.20	0.000
10	0.662	-0.094	917.15	0.000
11	0.717	0.248	1004.4	0.000
12	0.740	0.001	1098.2	0.000
13	0.680	-0.147	1177.8	0.000
14	0.613	-0.161	1243.1	0.000
15	0.616	-0.036	1309.5	0.000
16	0.622	0.030	1377.6	0.000
17	0.554	-0.078	1432.0	0.000
18	0.512	0.027	1478.8	0.000
19	0.528	-0.003	1528.9	0.000
20	0.544	-0.046	1582.5	0.000
21	0.507	0.016	1629.4	0.000
22	0.479	-0.002	1671.6	0.000
23	0.511	0.064	1720.1	0.000
24	0.521	-0.029	1770.8	0.000
25	0.460	-0.111	1810.6	0.000
26	0.382	-0.212	1838.3	0.000
27	0.392	0.045	1867.7	0.000
28	0.382	-0.018	1895.8	0.000
29	0.324	0.022	1916.2	0.000
30	0.275	-0.056	1930.9	0.000
31	0.291	0.004	1947.7	0.000
32	0.322	0.104	1968.2	0.000
33	0.298	0.096	1986.1	0.000
34	0.269	-0.053	2000.7	0.000
35	0.308	0.129	2020.0	0.000
36	0.328	-0.050	2042.2	0.000

Figure 6: *Correlogram of the international Air Passenger Traffic*

Table 4.4: Unit root test for domestic air passenger first difference

Null Hypothesis: D(DOMESTIC) has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on SIC, maxlag=13)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-14.24854	0.0000
Test critical values:	1% level	-3.473382	
	5% level	-2.880336	
	10% level	-2.576871	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DOMESTIC,2)
Method: Least Squares
Date: 09/20/22 Time: 21:08
Sample (adjusted): 2007M04 2019M12
Included observations: 153 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(DOMESTIC(-1))	-1.826154	0.128164	-14.24854	0.0000
D(DOMESTIC(-1),2)	0.328644	0.077244	4.254634	0.0000
C	2454.688	2786.284	0.880990	0.3797

R-squared	0.720183	Mean dependent var	74.94771
Adjusted R-squared	0.716452	S.D. dependent var	64614.60
S.E. of regression	34406.79	Akaike info criterion	23.74931
Sum squared resid	1.78E+11	Schwarz criterion	23.80873
Log likelihood	-1813.822	Hannan-Quinn criter.	23.77345
F-statistic	193.0319	Durbin-Watson stat	2.090868
Prob(F-statistic)	0.000000		

Table 4.5: Unit root test for International air passenger first difference

Null Hypothesis: D(INTERNATIONAL) has a unit root
Exogenous: Constant
Lag Length: 10 (Automatic - based on SIC, maxlag=13)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.933113	0.0000
Test critical values:	1% level	-3.476143	
	5% level	-2.881541	
	10% level	-2.577514	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(INTERNATIONAL,2)
Method: Least Squares
Date: 09/20/22 Time: 21:09
Sample (adjusted): 2008M01 2019M12
Included observations: 144 after adjustments

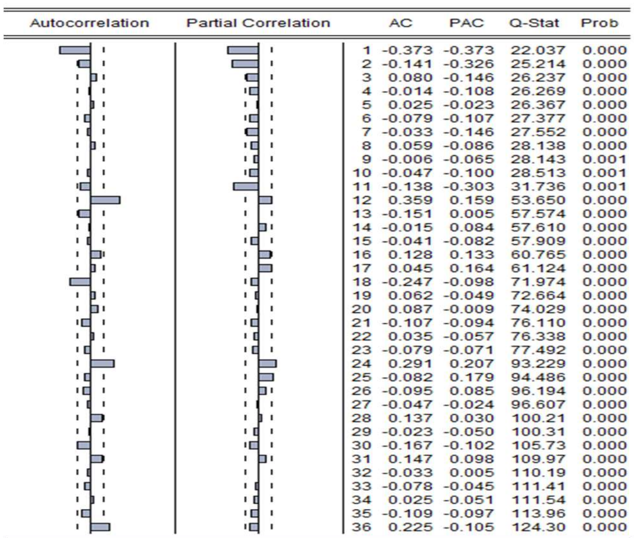
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(INTERNATIONAL(-1))	-5.715777	0.720496	-7.933113	0.0000
D(INTERNATIONAL(-1),2)	4.156492	0.675429	6.153854	0.0000
D(INTERNATIONAL(-2),2)	3.583654	0.624739	5.736239	0.0000
D(INTERNATIONAL(-3),2)	3.129090	0.558603	5.601632	0.0000
D(INTERNATIONAL(-4),2)	2.805320	0.488142	5.746939	0.0000
D(INTERNATIONAL(-5),2)	2.400018	0.425213	5.644272	0.0000
D(INTERNATIONAL(-6),2)	1.837445	0.369951	4.966719	0.0000
D(INTERNATIONAL(-7),2)	1.336793	0.307351	4.349402	0.0000
D(INTERNATIONAL(-8),2)	0.923840	0.232730	3.969584	0.0001
D(INTERNATIONAL(-9),2)	0.645754	0.154897	4.168930	0.0001
D(INTERNATIONAL(-10),2)	0.197993	0.088371	2.240480	0.0267
C	6360.099	1985.163	3.203817	0.0017

R-squared	0.788818	Mean dependent var	155.7847
Adjusted R-squared	0.771220	S.D. dependent var	45686.29
S.E. of regression	21852.19	Akaike info criterion	22.90165
Sum squared resid	6.30E+10	Schwarz criterion	23.14913
Log likelihood	-1636.918	Hannan-Quinn criter.	23.00221
F-statistic	44.82315	Durbin-Watson stat	1.987980
Prob(F-statistic)	0.000000		

Table 4.6: Unit Root Test using ADF (1979)

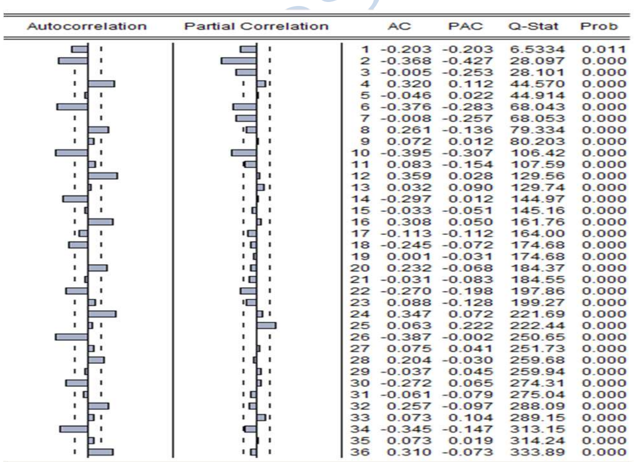
Variables	t-statistic	1% level	5% level	10% level	Remark
Domestic	-1.602388	-3.48162	-2.88393	-2.57879	Non-stationary
International	-1.498611	-3.48162	-2.88393	-2.57879	Non-stationary
DDomestic	-14.24	-3.4787	-2.8822	-2.5711	Stationary (First Difference)
DInternational	-7.93	-3.4712	-2.8843	-2.5732	Stationary (First Difference)

Table 4.7: ACF and PACF for Domestic Air passenger traffic.



	AC	PAC	Q-Stat	Prob
1	-0.373	-0.373	22.037	0.000
2	-0.141	-0.326	25.214	0.000
3	0.080	-0.146	26.237	0.000
4	-0.014	-0.108	26.269	0.000
5	0.025	-0.023	26.367	0.000
6	-0.079	-0.107	27.377	0.000
7	-0.033	-0.146	27.552	0.000
8	0.059	-0.086	28.138	0.000
9	-0.006	-0.065	28.143	0.001
10	-0.047	-0.100	28.513	0.001
11	-0.138	-0.303	31.736	0.001
12	0.359	0.159	53.650	0.000
13	-0.151	0.005	57.574	0.000
14	-0.015	0.084	57.610	0.000
15	-0.041	-0.082	57.909	0.000
16	0.128	0.133	60.765	0.000
17	0.045	0.164	61.124	0.000
18	-0.247	-0.098	71.974	0.000
19	0.062	-0.049	72.664	0.000
20	0.087	-0.009	74.029	0.000
21	-0.107	-0.094	76.110	0.000
22	0.035	-0.057	76.338	0.000
23	-0.079	-0.071	77.492	0.000
24	0.291	0.207	93.229	0.000
25	-0.082	0.179	94.486	0.000
26	-0.095	0.085	96.194	0.000
27	-0.047	-0.024	96.607	0.000
28	0.137	0.030	100.21	0.000
29	-0.023	-0.050	100.31	0.000
30	-0.167	-0.102	105.73	0.000
31	0.147	0.098	109.97	0.000
32	-0.033	0.005	110.19	0.000
33	-0.078	-0.045	111.41	0.000
34	0.025	-0.051	111.54	0.000
35	-0.109	-0.097	113.96	0.000
36	0.225	-0.105	124.30	0.000

Table 4.8: ACF and PACF for International Air passenger traffic.



	AC	PAC	Q-Stat	Prob
1	-0.203	-0.203	6.5334	0.011
2	-0.368	-0.427	28.097	0.000
3	-0.005	-0.253	28.101	0.000
4	0.320	0.112	44.570	0.000
5	-0.046	0.022	44.914	0.000
6	-0.376	-0.283	68.043	0.000
7	-0.008	-0.257	68.053	0.000
8	0.261	-0.136	79.334	0.000
9	0.072	0.012	80.203	0.000
10	-0.395	-0.307	106.42	0.000
11	0.083	-0.154	107.59	0.000
12	0.359	0.028	129.56	0.000
13	0.032	0.090	129.74	0.000
14	-0.297	0.012	144.97	0.000
15	-0.033	-0.051	145.16	0.000
16	0.308	0.050	161.76	0.000
17	-0.113	-0.112	164.00	0.000
18	-0.245	-0.072	174.68	0.000
19	0.001	-0.031	174.68	0.000
20	0.232	-0.068	184.37	0.000
21	-0.031	-0.083	184.55	0.000
22	-0.270	-0.198	197.86	0.000
23	0.088	-0.128	199.27	0.000
24	0.347	0.072	221.69	0.000
25	0.063	0.222	222.44	0.000
26	-0.387	-0.002	250.65	0.000
27	0.075	0.041	251.73	0.000
28	0.204	-0.030	259.68	0.000
29	-0.037	0.045	259.94	0.000
30	-0.272	0.065	274.31	0.000
31	-0.061	-0.079	275.04	0.000
32	0.257	-0.097	288.09	0.000
33	0.073	0.104	289.15	0.000
34	-0.345	-0.147	313.15	0.000
35	0.073	0.019	314.24	0.000
36	0.310	-0.073	333.89	0.000

Domestic represents domestic air passenger traffic, International (International air passenger traffic), DDomestic (first difference of Domestic) and DInternational (first difference of International)

The results in table above show that the original time series data for international and domestic air passenger traffic are not stationary but become trend and intercept stationary after first differencing. The time series data for the air passenger traffic were also seasonally differenced to take care of the presence of seasonal effect.

Model Identification

In identifying the appropriate seasonal ARIMA model for the time series data, it would be necessary to decide on the non-seasonal and seasonal autoregressive terms, seasonal and non-seasonal differencing and the moving average terms. From the time plots, a non-seasonal differencing would be needed, the ADF test in table 4.5 however confirmed that the time series data attained stationarity at order 1. A glance at the time plots and ACF plots show seasonal pattern. This forms a basis in determining the autoregressive and moving average orders by employing the AIC and BIC information criterion

Tables 4.7 and 4,8 show the plots of ACF and PACF for both domestic and international air passenger traffic respectively. This gives ample knowledge in identifying the order p, q, P and Q parameters of the proposed SARIMA models by visual inspection, it is observed that the ACF cuts at q=2 and Q=1 which suggests a moving average parameter of order 2 and a seasonal moving average parameter of order 1. The PACF plot by visual observation suggests AR parameter of order1 and 2 and a Seasonal AR of order 1. Similarly, from Figure 4.8, by visual inspection, it is observed that the ACF cuts at q=1and Q=1, while the PACF slightly cuts at p=3(AR parameter of order 3) and P=1(Seasonal AR parameter of order 1). From this inference the best SARIMA models are selected based on error diagnostics, AIC, BIC, RMSE and also keeping in mind parsimony. Based on this, SARIMA models were considered each for domestic and international air passenger traffic.

From Table 4.9, considering the models performance in terms of AIC, BIC and RMSE $SARIMA(2,1,1)(0,1,1)_{12}$ and $SARIMA(1,1,1)(0,1,1)_{12}$ with lower AIC and BIC values performed competitively better than the other SARIMA models. The L-jung Box portmanteau test for residual autocorrelation shows that the residuals are not serially correlated and the Jarque-Bera normality test confirms that the residuals are normal. Both models should be adequate in modelling the domestic air passenger traffic. The $SARIMA(1,1,1)(0,1,1)_{12}$ is the selected model since it is more parsimonious.

Table 4.9 Postulated SARIMA models for Domestic Air Passenger Traffic

Model	AIC	BIC	RMSE	Normality Test		Serial Correlation	
				JB Test	P-value	Q-statistic	P-value
$SARIMA(2,1,1)(0,1,1)_{12}$	-1.222	-1.104	0.129	2.269	0.322	26.027	0.762
$SARIMA(2,1,1)(1,1,1)_{12}$	1.157	1.004	0.132	1.066	0.587	23.667	0.587
$SARIMA(1,1,1)(1,1,1)_{12}$	1.123	0.997	0.135	1.587	0.452	34.654	0.342
$SARIMA(1,1,1)(0,1,1)_{12}$	1.203	1.109	0.13	2.102	0.35	32.94	0.47
$SARIMA(1,1,2)(1,1,1)_{12}$	1.135	0.984	0.133	1.172	0.557	29.161	0.561

Table 4.10 Postulated SARIMA models for International Air Passenger Traffic

Model	AIC	BIC	RMSE	Normality Test		Serial Correlation	
				JB Test	P-value	Q-statistic	P-value
$SARIMA(1,1,2)(1,1,1)_{12}$	-1.558	-1.408	0.108	32.6	0	23.989	0.811
$SARIMA(3,1,1)(0,1,2)_{12}$	1.527	1.361	0.11	13.923	0.009	27.1	0.618
$SARIMA(2,1,2)(2,1,2)_{12}$	-1.841	-1.596	0.092	1.302	0.522	31.495	0.296
$SARIMA(1,1,2)(0,1,1)_{12}$	-1.553	-1.435	0.109	27.881	0	25.362	0.791
$SARIMA(3,1,1)(2,1,2)_{12}$	-1.968	1.721	0.086	4.352	0.114	16.82	0.952

From Table 4.9, considering the models performance in terms of AIC, BIC and RMSE $SARIMA(2,1,1)(0,1,1)_{12}$ and $SARIMA(1,1,1)(0,1,1)_{12}$ with lower AIC and BIC values performed competitively better than the other SARIMA models. The L-jung Box portmanteau test for residual autocorrelation shows that the residuals are not serially correlated and the Jarque-Bera normality test confirms that the residuals are normal. Both models should be adequate in modelling the domestic air passenger traffic. The $SARIMA(1,1,1)(0,1,1)_{12}$ is the selected model since it is more parsimonious.

From Table 4.10, $SARIMA(3,1,1)(2,1,2)_{12}$ was selected to be the best model for the international air passenger traffic. Compared to the other proposed SARIMA models, $SARIMA(3,1,1)(2,1,2)_{12}$ has the lowest AIC and BIC, the errors are not serially correlated, errors are normally distributed. The $SARIMA(2,1,2)(2,1,2)_{12}$ is the SARIMA model that performs closely to the selected model $SARIMA(3,1,1)(2,1,2)_{12}$ but the selected model performs generally better. The other models fail the Jarque-Bera test for normality.

Estimation of the SARIMA model

Table 4.11: Parameter of Estimated SARIMA Models

Domestic air passenger traffic			International Air passenger traffic		
$SARIMA(1,1,1)(0,1,1)_{12}$			$SARIMA(3,1,1)(2,1,2)_{12}$		
Variables	Estimates	P-value	Variables	Estimates	P-value
AR(1)	0.133	0.364	*AR*(1)	-0.39	0.085
MA(1)	-0.709	0	*AR*(2)	-0.324	0.079
SMA(12)	0.906	0	*AR*(3)	0.176	0.242
JB Test	2.102	0.35	*SAR*(12)	-0.195	0.029
Q-Test	32.94	0.47	*SAR*(24)	0.275	0.007
AIC	1.203		*MA*(1)	-0.459	0.04
BIC	-1.109		*SMA*(12)	0.665	0
			SMA(24)	0.872	0
			JB Test	4.352	0.113
			Q-Test	16.82	0.952
			AIC	-1.968	
			BIC	-1.721	

Table 4.12 Estimates of the Holt-Winters Exponential Smoothing Parameters

Variables	Model	α (Level)	β (Trend)	γ (Seasonal)	BIC	AIC
Domestic	Multiplicative	0.2174	0.0000	0.2032	192.4516	146.3267
Domestic	Additive	0.3300	0.0000	0.0335	166.4400	120.3152
International	Multiplicative	0.3034	0.0253	0.2447	140.8886	94.7637
International	Additive	0.2676	0.0000	0.000	116.9094	70.7846

Table 4.11 presents the estimated parameters of the seasonal models for both domestic and international air passenger traffic. The diagnostics checks performed indicate that the models should be adequate in modelling the air passenger traffic. The performances of these models were evaluated in comparison with the Holt-winters Exponential smoothing and the performance of the artificial neural network.

Estimation of the Holt-Winter's Exponential Smoothing Model

The time series data obviously indicate the presence of seasonality as shown in the time plot for both the domestic air passenger and international passenger traffic. The multiplicative and additive Holt-winters exponential smoothing models would be examined on the variables to see which models fit best.

Estimates of the parameters in the Table 4.12 show that the additive Holt-winters model is best suited for modelling the air passenger traffic for both domestic and the international flight based on AIC and BIC criteria. The parameters are estimated objectively rather than subjectively, by choosing the values that best minimize the sum of squared errors. The value of the parameter β being zero in the models for both domestic and international air passenger traffic, indicates that the slope is relatively constant over time. The values of α and γ show that emphasis is only fairly placed on the recent observation. The residuals of the best of these models would be examined for presence of non-zero autocorrelations, using the Ljung-Box test and also the forecast performance characteristics.

Table 4.13: Holt-Winters Model Performance Evaluation

Variables	Model	AIC	BIC	Q-statistic	P-value
Domestic	Additive	120.3152	166.4400	29.7100	0.1946
International	Additive	70.7846	116.9094	20.212	0.6847

From Table 4.13 having subjected the models to diagnostics (AIC, BIC & ACF)it could be observed that the Holt-Winters additive models the air passenger traffic better than the multiplicative model does for both domestic and international air passenger traffic. The Ljung-Box test reveals absence of autocorrelation in the residuals of the additive models.

Estimation of the Artificial Neural Network Models

In estimating the ANN model for the air passenger traffic flow, the Feed Forward Neural Network(FFN) is implemented by the back propagation algorithm. Four models, with 3 layers (input layer, hidden layer and output) as expected of most standard neural network architecture were considered. The models with 12 neurons in the input layer, 4 neurons in the hidden layer and one neuron in the output layer (12-4-1), 12 neurons in the input layer, 6 neurons in the hidden layer and 1 neuron in the output layer (12-6-1), 12 neurons in the input layer, 8 neurons in the hidden layer and one neuron in the output layer (12-8-1) and 12 input neurons,12 hidden layer neurons, an input neuron (12-12-1) were examined.

The Sigmoid and Bipolar Sigmoid activation functions were used in the learning process so as to examine which of these activation functions models the time series data better by observing the mean squared error yielded in each process.

Table 4.14 Performance of the ANN models for Domestic Air Passenger

Model	Activation function	Iterations	MAE	MSE
12-4-1	Sigmoid	20000	0.10098	0.01740
	Bipolar Sigmoid	30000	0.05980	0.00645
12-6-1	Sigmoid	20000	0.10135	0.01754
	Bipolar sigmoid	30000	0.04687	0.00410
12-8-1	Sigmoid	20000	0.10015	0.01733
	Bipolar Sigmoid	30000	0.04020	0.00324
12-12-1	Sigmoid	20000	0.10049	0.01724
	Bipolar Sigmoid	40000	0.01689	0.00079

Table 4.15: Performance of the ANN models for International Air Passenger

Model	Activation function	Iterations	MAE	MSE
12-4-1	Sigmoid	20000	0.08080	0.01154
	Bipolar Sigmoid	20000	0.05694	0.00538
12-6-1	Sigmoid	20000	0.08173	0.01177
	Bipolar sigmoid	20000	0.03962	0.00277
12-8-1	Sigmoid	20000	0.07998	0.01143
	Bipolar Sigmoid	30000	0.02933	0.00160
12-12-1	Sigmoid	20000	0.07994	0.01150
	Bipolar Sigmoid	30000	0.01028	0.00028

In Tables 4.14 and 4.15 it was observed that in the 4 ANN models that were considered, the best results were obtained when the bipolar sigmoid activation function was used. This was deduced by the measure of accuracy. In each of the cases better results were obtained with the increase in number of neurons in the hidden layer, with the model 12-12-1 using bipolar sigmoid function yielding the best results based on the in-sample measure of error. However, the ANN 12-12-1 model for the international air passenger traffic was not as effective, when checking the out sample performance separately, it yielded $RMSE = 0.32836$ compared with the other ANN models with fewer neurons in the hidden layer. ANN 12-4-1 tends to perform averagely better with better out-sample, forecast performance, $RMSE = 0.16185$. For the domestic air passenger traffic, ANN 12-4-1 which has fewer neurons (fewer parameters) and sufficiently good forecast accuracy based on the measure of error would be employed in comparison with the best SARIMA and HWES model.

The input neurons are the lagged series, $y_{t-1}, y_{t-2} \cdots y_{t-112}$ of the time series data, these are weighted randomly and linearly combined and the results are modified by the bipolar sigmoid function to serve as input to the next layer.

The performance of the ANN best model, 12-4-1, for domestic and international air passenger using the bipolar sigmoid function was compared with the best SARIMA and Holt-Winters exponential smoothing models for the in-sample and out-sample forecast, so as to examine their relative performances based on MAPE and RMSE.

Performance Comparison of the Models

In evaluating the forecast performance accuracy in each of the models, the training set of data and a test set were used. The training data set contains data from 2007 to 2019, a total of 156 monthly observations in each case, domestic and international air passenger traffic. The test data set contains data from 2007 to 2019, a total of 24 monthly observations in each case. The static out of-sample forecast is employed in this study. One step ahead forecast, the forecast for the month of January 2014, is made by using the training data set, this estimate is added to the training data set which in turn is used to make forecast for the month of February 2014. The forecast estimate for the month of February 2014 is added to the new training set so as to obtain forecast for March 2014.

The iteration continues until the estimate for the last month, December 2015. This estimates from the static forecast are compared with the actual test data set, using statistical loss functions. Based on this, the values of the MAPE and RMSE for the respective proposed models are compared so as to select the model that has the best out-of-sample performance for the forecast horizon.

Table 4.16: Performance Comparisons of Models for Air Passenger Traffic

Model	Domestic air traffic passenger			
	In-Sample		Out-Sample	
	MAE	RMSE	MAPE	RMSE
SARIMA$(1,1,1)(0,1,1)_{12}$	0.0983	0.1304	0.5291	0.0808
Holt-Winters	0.09183	0.12162	0.6000	0.0932
ANN(12 – 4 - 1)	0.05896	0.00803	0.40918	0.06364
	International air passenger traffic			
Model	**In-Sample**		**Out-Sample**	
	MAE	**RMSE**	**MAPE**	**RMSE**
SARIMA$(3,1,1)(2,1,2)_{12}$	0.0632	0.0864	0.9605	0.1453
Holt-Winters	0.07291	0.10081	0.6435	0.0979
ANN(12 - 4 - 1)	0.05694	0.07335	0.98190	0.16185

Table 4.16 gives the empirical results of the forecasting performance of the models, ANN, SARIMA and Holt-Winters. All the examined models produced good forecast in the sectors considered, since the MAPE and RMSE are generally low. In modelling the air passenger traffic for both domestic and international flights, the ANN models performed very well in the in-sample forecast. ANN (12 -

4 - 1) outperforms the Holt-Winters and SARIMA models for in-sample and out-sample forecast in the domestic sector. It is observed that the values of MAE, MAPE and RMSE are smaller than the other models.

The ANN(12 - 4 - 1) model for international air passenger traffic, by comparing the MAPE, MAE and RMSE has a better in-sample performance than the other models but has the least out-sample forecast performance. The effectiveness of the ANN models in this dissertation corroborates emphases made on the flexibility and excellent function approximation capability of the ANN in previous studies by White (1989). Though the ANN(12 - 4 -1), in modeling the international air passenger traffic, has the least out sample performance, observing the MAE and RMSE for the training set it is evident that the ANN has a great learning and pattern recognition ability. From the time plot of the international air traffic passenger, the exhibition of a more stable characteristics was obtained compared to the time plot for domestic air passenger traffic. This could be the reason for the better performance of the ANN in modelling the domestic air passenger traffic, since the ANN does very well in capturing some hidden salient characteristics and non-linearity in data.

The Holt-Winters Exponential Smoothing model shows better post forecast accuracy, for international air passenger traffic, than SARIMA $(3,1,1)(2,1,2)_{12}$ and ANN, based on the MAE, RMSE and MAPE. In modelling the domestic air passenger traffic, the ANN had the best in sample and out-of-sample performance, though SARIMA $(1,1,1)(0,1,1)_{12}$ performed competitively with ANN.

IV. Discussion

The conclusion and recommendations based on the inference from the results obtained from the 3 models, ANN, SARIMA and Holt-Winters exponential smoothing are discussed. In this paper, two sets of time series data from 2007 to 2019, monthly air passenger traffic for domestic and international flight were obtained. Three-time series model were employed in order to achieve the set objectives aimed at employing time series forecasting models to approximately predict air passenger traffic flow in Murtala Muhammad International Airport Lagos, Nigeria.

The data were divided into two sets, training and test sets. The time series models considered were evaluated on these data set so as to measure the accuracy of the models for in-sample and out-sample performance. The time series data were differenced in order to stabilize the variance, so as to obtain more desirable results in the models considered.The time plots of the air passenger traffic reveal non-stationarity, presence of trend, seasonality and noise. After first differencing, the Augmented Dickey Fuller test shows that the time series data attained stationarity. The Jarque-Bera test for normality however reveals that the time series data are not normally distributed. The Box-Jenkins methodology was employed in building the best SARIMA models for obtaining a fairly good forecast for air passenger traffic flow in Murtala Muhammad International Airport.

After the necessary diagnostics using AIC, BIC, error measurements based on in-sample and out sample forecast performance, SARIMA$(1,1,1)(0,1,1)_{12}$ and SARIMA$(3,1,1)(2,1,2)_{12}$ were selected, as the best SARIMA models for modelling the air passenger traffic for domestic and international respectively. The summary of these models show that, they gave better fit for the air passenger traffic. This was also compared with the other proposed models.

In selecting the best Holt-Winters model, the additive and multiplicative models were considered so as to select the best models. The appropriate smoothing parameters, α, β and γ which best minimize the sum of squared errors for the level, trend and seasonality were used in determining the best Holt-Winters model. In modelling the air passenger traffic, additive model appeared to be the better model based on the AIC and BIC.

The Holt-Winters gave a generally good performance for in and out of sample forecast, it outperforms the SARIMA and ANN in the out-sample forecast in the international sector based on the MAPE and RMSE.

The feed forward neural network, using the back propagation algorithm was used to model the time series data. A three layered feed forward neural network was used, with the number of neurons in the hidden layer varied to obtain the neural network architecture that best models the time series data.

The ANN(12-4-1), ANN(12-6-1), ANN(12-8-1) and ANN(12-12-1) were considered using the sigmoid and bipolar sigmoid functions. ANN(12-12-1) in both domestic and international sector produced the best in-sample performance. However, the out-sample performance was least, relative to the other ANN models. ANN(12-4-1) appeared to be the best model for modelling the domestic air passenger traffic, considering the in-sample and out-sample errors. For the international air passenger traffic, ANN(12-4-1) had the least out-sample results.

V. Conclusion

Generally, in achieving the aim of this dissertation, three-time series models were considered, the seasonal auto-regressive integrated moving average (SARIMA), Holt-Winters Exponential Smoothing model and the Artificial Neural Network. These time series models were tested on the international air passenger and domestic traffic, to evaluate their performances and test the predictive capability of the artificial neural network relative to the other models.

The models were estimated on the training data set which covers the period from January 2003 to December 2013. The test set which covers the period from January 2014 to 2015 December was used to obtain the forecast accuracy measure of these models based on RMSE(Root Mean Squared Error) and MAPE(Mean Absolute Percentage Error).Empirical results show that all the models provide good forecasts of the air passenger traffic for international and domestic.

Comparing results across the models, it was observed that no model completely outperforms the other in all the sectors. However, the ANN model was found to be very efficient and had the best in-sample performance across the two sectors. In modelling the domestic air passenger, the ANN model was seen to be significantly dominant, while for the international air passenger traffic the ANN also gave the best in-sample accuracy performance but the least out-sample performance. The Holt-Winters exponential smoothing and SARIMA both yielded good results. The Holt-Winters exponential smoothing out-performed the SARIMA and ANN for the out-of-sample forecast of international air passenger traffic while the SARIMA was more dominant than Holt-Winters in the domestic sector.

Conclusively, this study has been able to establish the effectiveness of the ANN in modeling and forecasting time series data and also select time series models that gave fairly accurate forecast of air passenger traffic in Murtala Muhammed International Airport Lagos, Nigeria.

References

Abed, S. Y., Ba-Fail, A. O., and Jasimud in, S. M. (2001). An econometric analysis of international air travel demand in Saudi Arabia. *Journal of Air Transport Management*, *7*(3), 143-148.

Andreoni, A. and Postorino, M. (2006). A multivariate ARIMA model to forecast air transport demand. *Proceedings of the Association for European Transport and Contributors*, pages 1–14.

Artis, M. J., and Zhang, W. (1990). BVAR forecasts for the G-7. *International Journal of Forecasting*, *6* (3), 349-362.

Bao, Y., Xiong, T., and Hu, Z. (2012). Forecasting air passenger traffic by support vector machines with ensemble empirical mode decomposition and slope-based method. *Discrete Dynamics in Nature and Society*, 2012.

Basheer, I. A., and Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. *Journal of microbiological methods*, *43*(1), 3-31.

Bates, J. and Granger, C. (1969). The combination of forecasts. *Operational Research Quarterly*, 20(4):451–468.

Bjørnland, H., Gerdrup, K., Jore, A., Smith, C., and Thorsrud, L. (2011). Does forecast combination improve norges bank inflation forecasts? *Oxford Bulletin of Economics and Statistics*, 74(2):163– 179.

Bougas, C. (2013). Forecasting air passenger traffic flows in Canada: an evaluation of time series models and combination methods. *Master's thesis, Laval University*.

Bourbonnais, R. and Terraza, M. (2004). *Analyse des sériestemporelles: applications à l'économieet à la gestion*, volume 3. Dunod.

Box, G. E. P and Jenkins, GM (1970). *Time series analysis, forecasting and control*.

Box, G. E., and Jenkins, G. M. (1976). *Time series analysis: forecasting and control, revised ed*. Holden-Day.

Bresson, G. and Pirotte, A. (1995). *Économétrie des sériestemporelles: Théorieet applications*. Presses Universitaires de France, first edition.

Brown, B. and Murphy, A. (1996). Improving forecasting performance by combining forecasts: The example of road-surface temperature forecasts. *Meteorological Applications*, 3 (3) :257–265

Canova, F., and Hansen, B. E. (1995). Are seasonal patterns constant over time? A test for seasonal stability. *Journal of Business & Economic Statistics*, *13*(3), 237-252.

Chan, Y. L., Stock, J. H., and Watson, M. W. (1999). A dynamic factor model framework for forecast combination. *Spanish Economic Review*, *1*(2), 91-121.

Chen, C., Chang, Y., and Chang, Y. (2009). Seasonal ARIMA forecasting of inbound air travel arrivals to taiwan. *Transport metrica*, 5(2):125–140.

Chen, S.-C., Kuo, S.-Y., Chang, K.-W., and Wang, Y.-T. (2012). Improving the forecasting accuracy of air passenger and air cargo demand: the application of back-propagation neural networks. *Transportation Planning and Technology*, 35(3):373–392.

Coshall, J. (2006). Time series analyses of UK outbound travel by air. *Journal of Travel Research*, 44(3):335–347.

Coshall, J. (2009). Combining volatility and smoothing forecasts of UK demand for international tourism. *Tourism Management*, 30(4):495–511.

Çuhadar, M. (2014). Modelling and Forecasting Inbound Tourism Demand to Istanbul: A Comparative Analysis. *European Journal of Business and Social Sciences*, *2*(12), 101-119.

Dickey, D. A., and Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, 1057-1072.

Doguwa, S. I., &Alade, S. O. (2015). On time series modeling of Nigeria's external reserves. *CBN Journal of Applied Statistics*, *6*(1), 1-28.

Elliott, D. L. (1993). A better activation function for artificial neural networks.

Emiray, E., & Rodríguez, G. (2003). *Evaluating Time Series Models in Short and Long-Term Forecasting of Canadian Air Passenger Data* (No. 0306E). University of Ottawa, Department of Economics.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of the united Kingdom Inflation. *Econometric*, 50 (4): 987-1008.

Engle, R. F., and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, 251-276.

Fildes, R., Wei, Y., and Ismail, S. (2011). Evaluating the forecasting performance of econometric models of air passenger traffic flows using multiple error measures. *International Journal of Forecasting*, 27(3):902–922.

Grosche, T., Rothlauf, F., and Heinzl, A. (2007). Gravity models for airline passenger volume estimation. *Journal of Air Transport Management*, 13(4):175–183.

Hornik, K., Stinchcombe, M., and White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural networks*, *2*(5), 359-366.

Hsu, C. C., and Chen, C. Y. (2003). Applications of improved grey prediction model for power demand forecasting. *Energy Conversion and Management*, *44*(14), 2241-2249.

Hsu, C. I., and Wen, Y. H. (1998). Improved grey prediction models for the trans-pacific air passenger market. *Transportation planning and Technology*, *22*(2), 87-107.

Hu, L. (2002). Estimation of a censored dynamic panel data model. *Econometrica*, *70*(6), 2499-2517.

IATA (2010). www.iata.org/policy/Documents

IATA (2013). www.iata.org/pressroom/Pr/pages/2013-12-10-01.

Jarque, C. M., and Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, *6*(3), 255-259.

Kaastra, I., and Boyd, M. (1996). Designing a neural network for forecasting financial and economic time series. *Neurocomputing*, *10*(3), 215-236.

Kalekar, P. S. (2004). Time series forecasting using holt-winters exponential smoothing. *Kanwal Rekhi School of Information Technology*,4329008, 1-13.

Koizumi, K. (1999). An objective method to modify numerical model forecasts with newly given weather data using an artificial neural network. *Weather and forecasting*, *14*(1), 109-118.

Kulendran, N., and Shan, J. (2002). Forecasting China's monthly inbound travel demand. *Journal of Travel & Tourism Marketing*, *13*(1-2), 5-19.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, *54*(1), 159-178.

Larose, D. T. (2005). Decision Trees. *Discovering Knowledge in Data: An Introduction to Data Mining*, 107-127.

Law, R. (2000). Back-propagation learning in improving the accuracy of neural network-based tourism demand forecasting. *Tourism Management*, *21*(4), 331-340.

Ljung, G. and Box, G. (1978). On a measure of lack of fit in time series models. *Biometrika*,

Louvieris, P. (2002). Forecasting international tourism demand for Greece: A contingency approach. *Journal of Travel & Tourism Marketing*, *13*(1-2), 21-40.

Matthews, L. (1995). Forecasting peak passenger flows at airports. *Transportation*, *22*(1), 55-72.

McClelland, J. L., and Rumelhart, D. E. (1986, January). A distributed model of human learning and memory. In *Parallel distributed processing* (pp. 170-215). Mit Press.

Montgomery, D. C., Cheryl, L. J., and Murat, K. (2008). *Introduction to Time Series Analysis and Forecasting*. John Wiley and Sons, Inc.

National Bureau of Statistics (2014). www.nigerianstat.gov.ng.

Poore, J. W. (1993). Forecasting the demand for air transportation services. *Journal of Transportation Engineering*, *19*(5), 22-34.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1988). Learning representations by back propagating errors. *Cognitive modeling*, *5*(3), 1.

Shen, S., Li, G., and Song, H. (2011). Combination forecasts of international tourism demand. *Annals of Tourism Research*, 38(1):72–89.

Timmermann, A. (2006). Forecast combinations. *Handbook of economic forecasting*, 1:135 – 196.

Tsui, W. Balli, H., and Gower, H. (2011). Forecasting airport passenger traffic: the case of hongkong international airport. *Aviation Education and Research Proceedings*, 2011:54–62.

White, H. (1989). Neural-network learning and statistics. *AI expert*, *4*(12), 48-52.

Wong, K. K., Song, H., Witt, S. F., and Wu, D. C. (2007). Tourism forecasting: to combine or not to combine? *Tourism Management*, *28*(4), 1068-1078.

Zealand, C. M., Burn, D. H., and Simonovic, S. P. (1999). Short term streamflow forecasting using artificial neural networks. *Journal of hydrology*, *214*(1), 32-48.

Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. N*eurocomputing* 50: 159-175.