FITTING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE WITH EXOGENOUS VARIABLES MODEL ASSUMING LOGNORMAL ERROR TERM

ANDREW OJUTOMORI BELLO

(MATRIC. NO. 113832)

AUGUST 2021

FITTING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE WITH EXOGENOUS VARIABLES MODEL ASSUMING LOGNORMAL ERROR TERM

By

ANDREW OJUTOMORI BELLO

Matric. No. 113832

B. Sc. Statistics (Ilorin), M. Sc. Statistics (Ibadan)

A dissertation in the Department of Statistics,
Submitted to the Faculty of Science.
In partial fulfillment of the requirements for the Degree of

MASTER OF PHILOSOPHY

of the

UNIVERSITY OF IBADAN.

AUGUST 2021

CERTIFICATION

I certify that this work was carried out by Andrew Ojutomori **Bello** in the Department of Statistics, University of Ibadan, Nigeria under my supervision.

……..………………………………………………………………….

Supervisor

Professor O. I. Shittu,

P. Dipl, (Stat.), B. Sc., M. Sc,. M. Phil., Ph. D. (Statistics) (Ibadan)

Professor of Statistics

Department of Statistics,

University of Ibadan. Nigeria.

DEDICATION

I dedicated the research work to Almighty God and Authorities of Statistics.

ACKNOWLEDGEMENTS

I gave honour, adoration and thanks to the maker of Heavens and the Earth; I bowed before your thrown without Him the study could not have been achieved, God's Grace, Mercy and Favour saw me through the programme. My appreciation goes to my supervisor Professor Shittu for his guidance, positive criticism, giving me audience even when not convenient to him and supporting me towards the success of this work. I appreciate the contribution during my presentation of Prof. Shodipo, Prof. Amaha, Prof. Olubusoye, Dr. Odumgbosu (PG Cord), Dr. (Mrs.) Alaba, Dr. (Mrs.) Chukwu, Dr. Oyemaki, Dr. Obiseso, Head of Department, Lecturers and supporting staff of the Department of Statistics, University of Ibadan, for their contribution to my stay as student of the department and to the success of this work. I appreciate all authors whose materials I have consulted. I stood on your shoulders to see further in Time Series Analysis.

Thanks to Dr. Alaba for corrective suggestions and encouragement. I appreciate Dr. Wale, Mr. Olawale and Mr. Ojeniran for their contribution towards some mathematical derivations and data analysis. I won't forget the families of Dr. Adekunle, my friend and in-law Rasheed Olalekon Solomon, Dr. Igie Alex and Mr. Biola for your hospitality, encouragement and support during this programme.

I appreciate the effort of my beloved wife Mrs. Bunmi Sesi Bello, and God given children; Seyido Omokagbo Bello (Son), (Twins) Mercy Akosomime Bello, Melody Esomime Bello and Priscilla Irete Bello for their encouragement, prayers, support, forgoing some comforts and bearing my absence from home for the success of this programme. Thank you so much. I love you all!!!

I say thank you Dr. Patrick, Mr. Bukoye and Mr. Shaib for their contribution towards the development of the R-package code and guide on Eview package software used in data analyses required in this study.

I appreciate Dr. Osowole who had in time made provision of accommodation. I thank Abdul Fatai and Mohammed my ex-students of Auchi Polytechnic, who now supported me with accommodation, encouragement during my round-up of this programme. I thank you the family of Daddy Omodebo for your one time hospitality.

I will not fail to appreciate the Ibadan Drivers of Angle 90 Motor Park, Auchi whose services I patronized often during my trips to and from Ibadan. Most especially, my gratitude goes to God for granting the enabling safe journey mercies during the period.

I appreciate the Management of Auchi Polytechnic, Dean (SICT), Head of Department and Staff of Statistics Department, Auchi Polytechnic for been supportive and holding brief for me during my trips to University of Ibadan for this programme. I am grateful to my administrative office staff; Secretaries, Admin Officer, Clerk of Department of Statistics, Auchi Polytechnic who were so supportive during my absence for this programme.

To all of you who did one thing or the other to make this dream a reality I pray you will never be stranded in your life in Jesus Name. Amen.

Bello, Andrew Ojutomori

August 2021

ABSTRACT

The conventional Autoregressive Integrated Moving Average with Exogenous Variables (arimax) model with Normal Error term and Multiple Linear Regression (MLR) require stringent assumptions of normality of error term and stationarity of the series. These models have found widespread application in multidimensional relationships among economic variables; when these assumptions are often violated in practice leading to spurious regression model with poor forecast performance. Thus, this study was designed to develop an arimax model with Lognormal Error term capable of analysing time series data even when the assumptions were violated with reasonable forecast performance.

The conventional arimax (1, 0, 1) with normal error term defined as:

$\varepsilon_t = \dfrac{(1-\phi_1 B)y_t - \beta_0 - \beta_1 x_1}{(1+\theta_1)}$; where the lag operator $B = y_{t-1}$; the parameter ϕ_1 was the coefficient of the Autoregressive model (AR), θ_1 was the coefficient of Moving Average (MA), β_0 was the intercept and β_1 was the slope of the Regression part of the model. The proposed model was estimated by modifying the arimax (p, d, q) with lognormal error term where p is order of AR part, d is order of difference and q is order of MA part of the mixed model. The parameters were estimated using the maximum likelihood method. The choice of lognormal error term was based on the asymmetric property which overcomes non normality, the long tail and positive limit values properties overcome non stationarity. The dataset used were monthly External Reserves (Million USD), Official Exchange Rate (Naira to USD), Crude Oil Export (Million Barrel per Day) and Crude Oil Price (USD per Barrel). One hundred and twenty (120) observations were used for the modeling process. The proposed arimax (1, 0, 1) with lognormal error term ameliorate the non-normal and non-stationary assumptions. The proposed model performance was compared with conventional arimax (1, 1, 1) with normal error term and MLR model. Box-Jenkins Time Series procedure was used to model arimax (1, 1, 1) with normal error and Least Squares Estimator (LSE) technique for modeling MLR. The performance of proposed model was tested using Akaike Information Criteria (AIC), Mean Square Forecast Error (MSFE) and Loglikelihood (Loglik) values.

The non normal error function was obtained as:

$$f(\varepsilon_t) = \frac{1}{y_t\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left|\frac{\text{In}\left[(1-\phi_1 B)y_t - \beta_0 - \beta_1 x_1\right] - \text{In}(1+\theta_1)}{\sigma}\right|^2\right.$$

while the loglikelihood function was:

$$InL(\varepsilon_t) =$$

$$\frac{n}{2} In(2\pi\sigma^2) - \sum(Iny_t) - \frac{1}{2\sigma^2}\sum\left|\text{In}\left[y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1\right] - \text{In}(1+\theta_1)\right|^2$$

where σ^2 is variance. All the series were found to be non-stationary and non-normally distributed. The Loglik values of MLR, conventional arimax (1, 1, 1) with normal error and proposed arimax (1, 0, 1) with lognormal error term were -317.41, -240.23 and 1344.47; AIC values were 5.36, 490.45 and -0.41 while MSFE values were 12.41, 12.48 and 1.77. The proposed model has the highest Loglik value, smallest AIC and smallest MSFE values when compared with conventional arimax (1, 1, 1) with normal error and MLR model. Hence, the proposed model was considered better.

The autoregressive integrated moving average with exogenous variables assuming lognormal error term improved the capability of modeling time series data with better forecast performance even when the assumptions of normality of error term and stationarity of series were violated.

Keywords: Arimax, Log-normal error, Exogenous variables.

Word Count: 484

TABLE OF CONTENTS

CHAPTER THREE: METHODOLOGY

CHAPTER FOUR: RESULTS AND DISCUSSION

CHAPTER FIVE: SUMMARY AND CONCLUSION

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ARIMAX – Autoregressive Integrated Moving Average with Exogenous Variables

MLR –Multiple Linear Regression

ACF – Autocorrelation Function

PACF – Partial Autocorrelation Function

AIC – Akaike Information Criteria

MSFE – Mean Square Forecast Error

Loglik – Loglikelihood

CHAPTER ONE

INTRODUCTION

1.1 Background to the study

Time Series (TS): A type of stochastic process that can be indexed by time as well as other dimensions such as space, volume, and frequency. The collection of statistical observations made over a period of time. Because TS allowed for the observation of a time variable at any moment, the temporal gaps between subsequent members of the series do not have to be the same Shangodoyin, (2002).

The property of stationarity allowed group parameters to be estimated using the corresponding temporal averages of a single realization. Time series observations are intended to occur at regular intervals in both practice and theory. For example, firm earnings over a period of years, stock prices over a period of days, export totals over a period of months, air temperature over a period of hours, and so on.

Various mathematical models had been developed to capture the effect of time variation and to generate accurate predictions of future values based on the premise that previous behavior (pattern) would continue into the future. The core idea behind univariate forecasting; series' future values were mathematical function of its past values. Classic Time Series Models frequently assumed series stationarity. However, most macroeconomic and financial time series data are nonstationary (i.e., the mean, variance, and covariance change with time) and trend upward. Furthermore, some exogenous variables (external factors) were contributing aspects that times series models must capture in order to accomplish evidence-based informative decision-making. Classical Regression Analysis: a versatile data-analytic system that could be used whenever a quantitative variable (the dependent variable) was to be studied as a function of, or in relation to, any factor(s) of interest (expressed as independent variables). However, because of autocorrelation and time dependence, it could not been directly applied to time series data.

External reserves, exchange rate, crude oil export, and crude oil price economic variables empowered public sector foreign assets generation under the control monetary authorities in financial transactions of payment imbalances. Nigeria's foreign reserves were derived from the production and sales of crude oil CBN, (2007).

The World Bank, (2014) noted that national economy that depended on oil, would run at risk because of crude oil price instability. Nigeria's capital account had been subjected to crude oil price changes due to her reliance on oil for external revenues. This, combined with the country's large import bills, contributed to oscillations in the level of foreign reserves over time, and, as a result, the manner the reserves are maintained. Crude oil price fall had negative impact on the nation external reserves and naira exchange rate.

1.2 Motivation for the study

In the mean or variance of a stochastic process, autocorrelation and nonstationarity in time series are common. The traditional time series paradigm demands series stationarity; to account for non-stationarity, a data transformation method is used to solve time series problems using the classical model. As a result, an alternate model based on asymmetric distribution is required to handle the series' independent differencing.

Nonstationarity in the mean frequently necessitates lagging; use of unit root tests was based on the Augmented Dickey Fuller Test (ADF), Dickey Fuller (DF), Kwiatkowski Schwarch (KPSS), and Phillip Perron (PP) at order 1 or 2 or higher. In addition, non-stationarity in variance is frequently assessed before lagging to correct the process for stationarity unit root test procedures (correlogram); Partial Autocorrelation Function (PACF) or Autocorrelation Function (ACF). This extra work has aided in the problem of overcoming arimax modeling in the context of nonstationary economic time series.

In bridging knowledge gap in the application of arimax model with normal error term, this study is motivated to develop arimax model with lognormal error term to address the nonstationarity in economic time series. The choice of lognormal distribution error term is its characteristics; skewedness, long tailed and it positive values are important in determining investment decision like stock price options (Black-Scholes model).

It has been proven that the traditional method to statistical modeling and forecasting economic data of economic time series, which is based on globalization and automated market transactions, necessitates a robust approach to meet and accommodate the current scenarios. As a result, these issues are driving this academic research.

1.3 Justification for the study

Since most economic data were non-stationary by natural pattern and presence of autocorrelation, multiple linear regression modeling would be erroneous and deceptive in forecasting. The conventional arima modeling would also been inadequate in handling such data. Combining the benefits of both statistical modeling approach with their drawbacks becomes a viable option. As a result, the research was justified in developing an arimax with exogenous variables model that used the asymmetric property of the lognormal error term to correct the limitations of traditional non-stationary econometric time series differencing for better parameter measures and formulation of more accurate forecasting models.

1.4 Statement of the problem

Univariate statistical models could not fully represent statistical determination of economic impetus because of the presence of exogenous factors. Furthermore, error that was normally distributed at one point in time t may no longer be normally distributed at another point in time t (t ± 1) due to time-dependent nature of economic time series data that exhibited non-stationarity and non normality (changes in central locations, deviations from mean and covariance overtime).

The Gross Domestic Product (GDP), among other macroeconomic determinants, determines a country's economic growth; the index GDP is influenced by foreign reserves, currency rate, price index, and crude oil export. As a result, it's critical to build a better and more appropriate modeling technique that can overcome the limitations of traditional time series analysis models. Nonstationary series regression modeling using traditional techniques produces erroneous regression equations and misleading forecast outcomes. Arma – autoregressive moving average, arima – autoregressive integrated moving average, and sarma – seasonal autoregressive moving average were mixed TS - time series models not capable to accommodate exogenous variables in the model on their own.

The introduction of X exogenous variable(s) into the model armax; addressed some problems of additional variables(s). The implementation of conventional arimax model necessitates the use of strict assumptions of normality and serial stationarity in mean and variance. Independent series transformation (differencing) (in some situations it took more than order one differencing to achieve stationarity of the time series data). The usage of additional series transformations to achieve stationarity resulted to loss of original information and cost implication.

Volatility data (stock prices, interest rates, inflation rate, and other financial driven data) typically exhibited non stationarity. To bridge the gap, an alternative model was required to address the challenge of independent transformation of non stationary economic time series data.

1.5 Aim and objectives of the study

Aim: To create an arimax model assuming a lognormal error term that could solve asymmetry and nonstationarity concerns in economic time series modeling and improve forecasting.

Objectives:

1. Proposed an arimax assuming lognormal error term model,
2. Estimate the suggested model's parameters, and
3. Compare the generated model's performance to that of conventional arimax with normal error term and multiple linear regression model.

1.6 Research questions

1. In other way can non-normality problem in economic time series data be solved?
2. How can the non-stationarity problem in economic time series data be solved?
3. Can the lognormal asymmetry property help with the error term's non-normality?
4. Can the non-stationarity of a series be mitigated by lognormal long tail and positive properties?

1.7 Organization of the dissertation

This dissertation is divided into five chapters, each of which is ordered as follows: The first chapter includes a broad overview as well as some background information. It emphasized the problems of interest, the study's justification, the goals and objectives, as well as research questions. The second chapter discussed some concepts and provides a review of literature. The methodological approach of the study was presented in the third chapter. The fourth chapter contained the results and discussions. The summary and conclusion were presented in chapter five.

CHAPTER TWO

LITERATURE REVIEW

2.1 Literature review

Chapter two discussed related literature with respect to concepts of time series modeling, arimax modeling and applications. The bases of analyzing time series data were highlighted; focused on the Box-Jenkins procedures.

2.2 Conceptual clarification

Many real-world situations do not fit the assumptions of linearity or stationarity, the era of nonlinear modeling had come to supplement linear modeling in econometric time series Shittu and Yaya, (2011). The authors also pointed out that some economic time series may defy the conventional theories of stationarity and linearity. Those series were not taken into account at the raw level I(0), first order I(1), or higher order integrated levels Box and Jenkins, (1976).

2.3 Theoretical framework

Data ordered in sequence/time were in series; the term "ordered" refers to over time *t*. Statistical analyses used to found direction of change in data series over-time frame. When extrapolate; the pattern came up with future forecast. Extrapolation was a term used to describe how statistical forecasting algorithms project previous patterns/relationships into the future. Time series data in most cases would be non-stationary. The theory of time series models were based on stationary series overtime.

The mean and variance of Y_t time series data do not vary in a systematic manner and all periodic changes have been removed, the data attained stationary series. The statistical quality of a stationary process does not change with time, therefore such series were regarded to have been in statistical equilibrium. Shittu, (2011). Osabuohien, (2013) corroborated that stationary series must attain constant mean and variance. He further expressed mathematically that when Y_t satisfied the difference (∇^d) equation, it followed an arima (p, d, q) order:

$$Y_t = \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \phi_3 Y_{t-3} - ... - \phi_p Y_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + ... + \theta_q \varepsilon_{t-q} \quad (2.1)$$

$$\Rightarrow Y_t - \sum_{i=1}^{p} \phi_i Y_{t-i} = \varepsilon_t + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} \quad (2.2)$$

$$\phi(B) Y_t = \theta(B) \varepsilon_t \quad (2.3)$$

Where B denoted Backward Shift Operator.

Y_t denoted an arima (p, d, q) process if $\nabla^d Y_t$ was arima (p, d, q)

$$\phi(B) \nabla^d Y_t = \theta(B) \varepsilon_t \quad (2.4)$$

$$\phi(B)(1-B)^d Y_t = \theta(B) \varepsilon_t \quad (2.5)$$

Stationarity and invertibility, conditions were constants such that the equation's zeros were all outside the unit circle. The correlogram plot showed a spike at the seasonal lag, indicated that the data for a seasonal series had a seasonal character. Box and Jenkins, (1976), Madsen, (2008) and Meese and Geweke, (1982) all made contributions to the theories and practical application of arima model. Theoretical properties of TSA models gave a foundation for recognizing and estimating them as asserted by Osabuohien, (2013).

2.3.1 Autoregressive process

Autoregressive processes were regressions on themselves; an autoregression of order p process: denoted as AR(p). The geneal expression by Yule (1926):

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \Phi_3 Y_{t-3} + ... + \Phi_p Y_{t-p} + e_t \quad (2.6)$$

2.3.2 Moving average process

Mohammed, (2014) applied moving average (ma) model on temperature and rainfall series. The model MA was defined as:

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} - ... - \theta_q e_{t-q} \quad (2.7)$$

The equation 2.7 was mathematical expression ma(q) order q; where, Y_t stand for series observations at time t and e_t denoted the error term. There exist an independent of Y_{t-1}, Y_{t-2}, Y_{t-3}, ..., Y_{t-q}.

2.3.3 Autoregressive integrated moving average (arima) model

The strategy developed by Box and Jenkins (1970) was regarded as a watershed moment in the modern approach to time series analysis. The Box and Jenkins JB technique aimed at generating an arima model from an observed time series. The technique focused on Stationary processes, passing through appropriate preliminary data modifications. The model was a generalized version of the non-stationary arma model, which was denoted by arma (p, q):

$$Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \Phi_3 Y_{t-3} + ... + \Phi_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} - ... - \theta_q e_{t-q} \quad (2.8)$$

Where, Y_t represented the original series, for every time t, (assumed) independent of Y_{t-1}, Y_{t-2}, Y_{t-3}, …, Y_{t-p} .

A time series $\{Y_t\}$ was said to follow an integrated autoregressive moving average (arima) model if the **d^th^** difference $W_t = \nabla^d Y_t$ was a stationary arma process. If $\{W_t\}$ followed an arma (p, q) model, then the set $\{Y_t\}$ was an arima (p, d, q) process. Fortunately, for practical purposes, we can usually take d = 1 or at most 2.

Consider then an arima (p, 1, q) process; $W_t = Y_t - Y_{t-1}$ the equation become:

$$W_t = \Phi_1 W_{t-1} + \Phi_2 W_{t-2} + \Phi_3 W_{t-3} + ... + \Phi_p W_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \theta_3 e_{t-3} - ... - \theta_q e_{t-q} \quad (2.9)$$

2.3.4 Autoregressive integrated moving average with exogenous variables (arimax) model

Including exogenous variable(s) improves the arima model's ability to capture explanatory variables that may be impacting the time series' behavior. To account for external influences, the arimax model included regression model features.

The generalized arimax (p, d, q) defined as:

$$y_t = \sum_{i=1}^{p} \phi y_{t-i} + \sum_{i=1}^{q} \theta \varepsilon_{t-i} + \varepsilon_t + \sum_{k=0}^{n} \beta_k X_k \qquad (2.10)$$

where y_t is actual series

$\sum_{i=1}^{p} \phi y_{t-i}$ is the AR with p number of ϕ parameters

$\sum_{i=1}^{q} \theta \varepsilon_{t-i}$ is the MA with q number of θ parameters

ε_t is associated error term

$\sum_{k=0}^{n} \beta_k X_k$ is the Regression part with n number of β parameters

X_k is k number of associated exogenous random variable

The arimax model assumed that the error term was normal as cited by Mohammed, (2014).

2.4 Mathematical framework

2.4.1 Concept

The following ideas were discussed in the section: normal probability distribution, lognormal probability distribution, maximum likelihood, partial derivatives, and least square estimation.

2.4.2 Normal probability distribution

A family of continuous probability distribution models known as the normal distribution.

Given that $X \sim N(\mu, \sigma^2)$ then, the Probability Density function (PDF) of the population was defined as:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp^{-\frac{(x-\mu)^2}{2\sigma}} \qquad (2.11)$$

where; π and exp. were constants, -∞ < x < ∞, μ the mean of the distribution and standard deviation σ > 0.

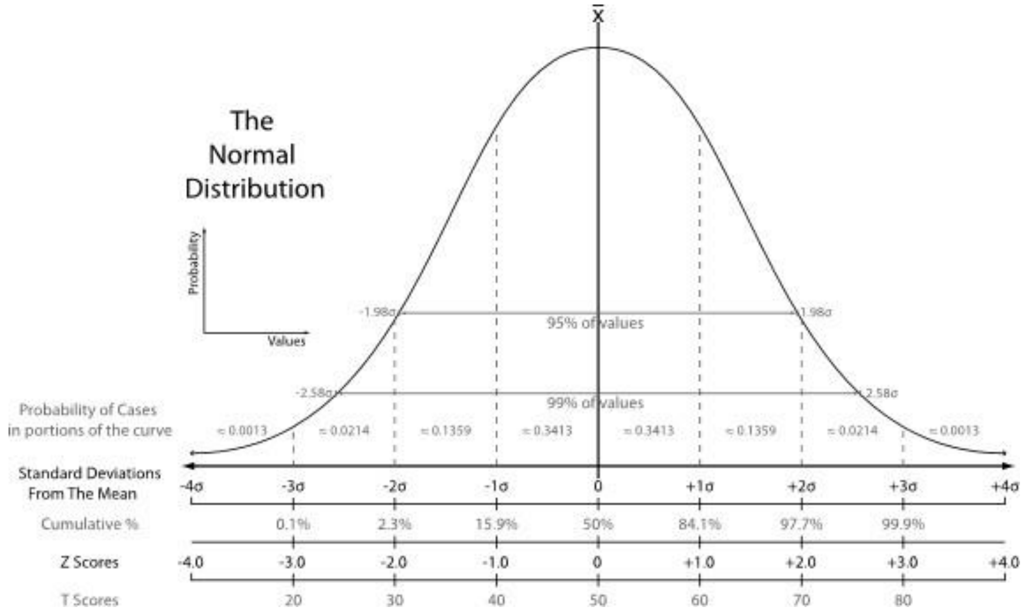


Fig. 2.1. Standard normal distribution curve (Zucchi Kristina 2018)

Fig. 2.1 The standard normal curve had symmetric, bell-shaped, the curved area size is one, zero coefficient of skweness and three coefficient of kurtosis properties.

2.4.3 Lognormal probability distribution

A continuous r. v. Y follow log-normal distribution when the distribution of $\log_e(y)$ was normal. The pdf of $\log_e(y)$ with parameter μ and σ^2 denoted $f(y;\mu,\sigma^2)$ was defined by the function:

$$f(y;\mu,\sigma^2) = \frac{1}{y_t\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2}\left|\frac{In(y_t)-\mu}{\sigma}\right|^2\right. \qquad (2.12)$$
$$for \quad 0<\mu<\infty,\; 0<y<\infty,\; \sigma>0$$

The variance σ^2 was the scale parameter and mean μ was the location parameter where y_t was random variable associated with time factor. The distribution was asymmetry and skewed to the right. The asymmetry of the lognormal distribution and the positive values result in a right-skewed curve.

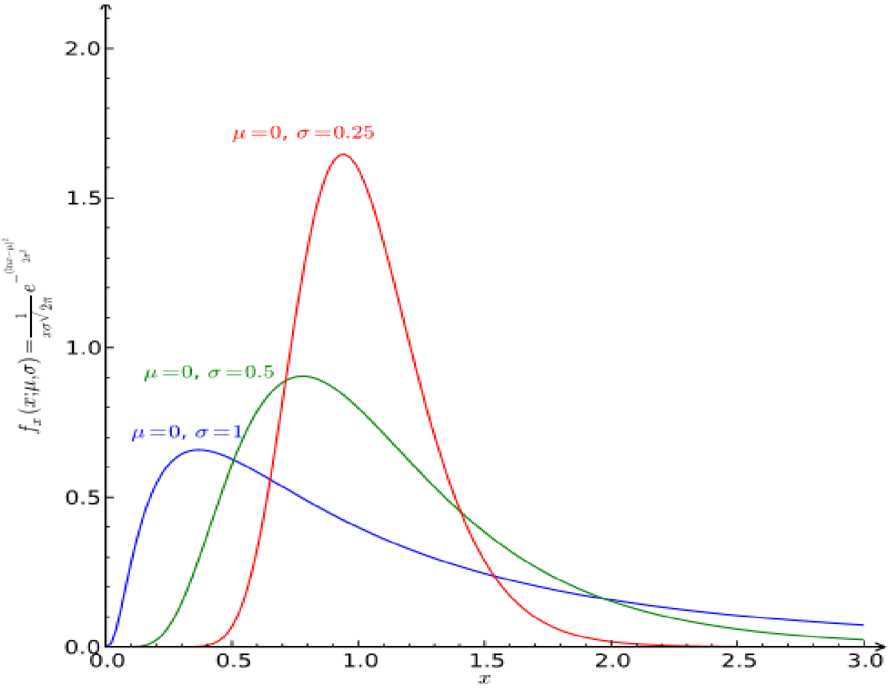


Fig. 2.2. Lognormal curve (Zucchi Kristina 2018)

As shown in Fig. 2.2 the values were positive and the lognormal distribution had asymmetric property, resulted in a right-skewed curve.

2.4.4 Mathematical principles of maximum likelihood estimator

Having found specified (adequate) distribution; the study adopted mle to estimate associated parameters of the arimax models under consideration.

Maximizing the score function in $L(\theta / y)$, the logarithm of the likelihood, would be easier. Maximum likelihood estimation (MLE) required the maximization of likelihood function $L(\theta)$ with respect to the unknown parameter θ. $L(\theta)$ was defined as an n-term product, which was difficult to maximize. Maximizing $L(\theta)$ was equivalent to maximizing $\log L(\theta)$ because log was a monotonic increasing function.

Supposed that the random variables y_1, y_2, …, y_n form a random sample from a distribution f(y/θ); if Y was continuous random variable, f(y/θ) was Probability Density Function (PDF), if Y was discrete random variable f(y/θ) was Probability Mass Function (PMF). To represent that the distribution also depends on a parameter θ, where θ could be a real valued unknown parameter or a vector of parameters. For every observed random sample y_1, y_2, …, y_n, were defined as:

$$f(y_1, y_2, ..., y_n / \theta) = f(y_1/\theta) f(y_2/\theta) ... f(y_n/\theta)$$

$$= \prod_{i=1}^{n} \mathrm{f}(\mathrm{x_i}, \theta) \qquad (2.13)$$

if $f(Y/\theta)$ was pdf, $f(y_1, y_2 ... y_n / \theta)$ was the joint density function; if $f(Y/\theta)$ was pmf, $f(y_1, y_2, ..., y_n / \theta)$ was the joint probability.

Now $f(y_1, y_2, ..., y_n / \theta)$ was the likelihood function. The likelihood function depends on the unknown parameter θ denoted as L(θ).We commonly get the mle by maximizing the natural logarithm of the likelihood, because any positively valued function reaches its maximum at the same point as its logarithm function.

Properties of mle: If $\hat{\theta}(y)$ was a maximum likelihood estimate for θ, then $g(\hat{\theta}(y))$ was a maximum likelihood estimate for $g(\theta)$. If θ was a parameter for the variance and $\hat{\theta}$ was the maximum likelihood estimator, the $\sqrt{\hat{\theta}}$ was the maximum likelihood estimator for the standard deviation.

2.5 Regression model

When the method of analysis was used to forecast or estimate the value of one variable corresponding to a given value of another variable, regression analysis was helpful in determining the most likely form of the relationship between variables.

2.5.1 Multiple linear regression

2.5.2 The model equation

Olubusoye, (2002) lecture note stated: the model of multiple linear regression explain the dependent variable y_i the i^{th} observation by a linear function of k explanatory variables; $x_{i1}, x_{i2}, ..., x_{ik} + U_i$.

A sample of n observations denoted as:

$$\begin{aligned} Y_1 &= X_{11}\beta_1 + X_{12}\beta_1 + . \ . \ . + X_{1k}\beta_k + U_1 \\ Y_2 &= X_{21}\beta_2 + X_{22}\beta_2 + . \ . \ . + X_{2k}\beta_k + U_2 \\ \vdots & \quad \vdots \quad \vdots \quad \vdots \\ Y_n &= X_{n1}\beta_n + X_{n2}\beta_n + . \ . \ . + X_{nk}\beta_k + U_n \end{aligned} \qquad (2.14)$$

The matrix format defined as:

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & . & . & x_{1k} \\ x_{21} & x_{22} & . & . & x_{2k} \\ . & . & . & . & . \\ . & . & . & . & . \\ x_{n1} & x_{n2} & . & . & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ \beta_k \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ . \\ . \\ U_k \end{bmatrix} \qquad (2.15)$$

$$Y_{nx1} = X_{nxk} \ \beta_{1xk} + U_{nx1}$$

where; Y was an n x l column vector containing the n sample, of Y values,

X was (n x k) matrix containing first, a column of ones and then all the sample values of the k-1 variables

β was (k x 1) column vector of parameters

U was n x l column vector containing the disturbance values.

2.5.3 Linear model assumptions

There were some assumptions made about the linear model as:

1. X matrix was *nonstochastic;* i.e. not random in nature.

2. E(U) = 0 (2.16)

3. $E(UU^1) = \sigma^2 I_n$ (the twin assumption)

$$UU' = \begin{bmatrix} U_1 \\ U_2 \\ . \\ . \\ U_n \end{bmatrix} \begin{bmatrix} U_1 & U_2 & . & . & U_n \end{bmatrix} = \begin{bmatrix} U_1^2 & U_1U_2 & . & . & U_1U_n \\ U_2U_1 & U_2^2 & . & . & U_1U_n \\ . & .... & . & ..... \\ . & .... & . & ..... \\ U_nU_1 & U_nU_2 & . & . & U_n^2 \end{bmatrix}_{nxn} \quad (2.17)$$

The expectation was defined as:

$$E[UU'] = \begin{bmatrix} E(U_1^2) & E(U_1U_2) & . & . & E(U_1U_n) \\ E(U_2U_1) & E(U_2^2) & . & . & E(U_1U_n) \\ . & ..... & . & ..... \\ . & ..... & . & ..... \\ E(U_nU_1) & E(U_nU_2) & . & . & E(U_n^2) \end{bmatrix}$$

$$\begin{bmatrix} \sigma_1^2 & 0 & . & . & 0 \\ 0 & \sigma_2^2 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & \sigma_n^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & 0 & . & . & 0 \\ 0 & 1 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & . & . & 1 \end{bmatrix} = \sigma^2 I_n \quad (2.18)$$

The twin folds of that assumption were:

i. $E(U_i^2) = \sigma^2 \quad \forall i$ (2.19)

Each U distribution had the same variance. The property was referred to as *homoscedasticity* (homogenous) variances.

ii. $E(U_iU_j) = 0 \quad \nabla\, i \neq j$ (2.20)

The disturbances were *pairwise uncorrelated.* When the condition failed the disturbances were said to be *autocorrelated correlated.*

4. The U vector had a multivariate normal distribution. $U \sim N(0, \sigma^2 I_n)$ Olubusoye, (2002).

2.5.3 Model estimation

Maximum likelihood estimation (mle) and least square estimation (lse) were employed in estimated models' parameters. Though distinct in principles, but produced the same estimator under certain commonly used assumptions, as described in the previous section.

2.5.4 Application of mle method on regression modeling

Least square estimator would not consider chance property of model and the terms represented by error (**e**) had a known distribution; mle was used to make some assumptions about the distribution and then maximized the probability of the sampled observations represented by the data. The assumption that the **e** were normally distributed with zero mean and variance-covariance matrix V. i.e. **e** ~ N(0, V), the likelihood defined as:

$$L = (2\pi)^{-\frac{1}{2}N} |V|^{-1/2} \exp\left\{-\frac{1}{2}(y - xb)' V^{-1}(y - xb)\right\} \qquad (2.21)$$

Maximizing this with respect to **b** was equivalent to solving

$$\frac{\partial(\log_e L)}{\partial b} = 0$$

The solution was the maximum likelihood estimator of **b** and turns out to be

$$\bar{b} = (x'V^{-1}x)^{-1} x'V^{-1} y \qquad (2.22)$$

The same as the generalized least squares estimator.

When V = σ^2I, $\bar{b}$ simplifies to $\hat{b}$, took $\hat{b}$ as the maximum likelihood estimator, because of the assumption: **e** ~ N(0, σ^2I). Under normality assumption maximum likelihood estimation leads to the same estimator, $\bar{b}$ as generalized least squares; and this reduced to the ordinary least square estimator $\hat{b}$ when V = σ^2I Chingnun, (2006).

2.5.5 Application of ordinary least square estimator (olse)

Olubusoye (2002): the olse method of parameter estimation:

Model: $Y = X\beta + U$ (2.23)

Let b = any arbitrary k - element vector.

The vector of residuals was defined as:

$\mathbf{e} = y - xb$ (2.24)

The least square principle was to choose ***b*** to minimize the residual sum of squares (RSS) **e'e** (**e**^{2} matrix):

$$RSS = e'e$$

$$= (y - xb)'(y - xb)$$

$$= y'y - b'x'y - y'xb + b'x'xb$$

$$= y'y - 2b'x'y + b'x'xb$$

$$\frac{\partial(RSS)}{\partial b} = -2x'y + 2x'xb = 0$$

$$= (x'x)b\ x'y$$

$$b = (x'x)^{-1}x'y \quad (2.25)$$

2.5.6 Mean and Variance estimation of b

Mean:

$$b = (x'x)^{-1}x'y$$

when substitute $y = X\beta + U$ gave:

$$b = (x'x)^{-1}x'(x\beta + U)$$

$$= (x'x)^{-1}x'x\beta + (x'x)^{-1}x'U = \beta + (x'x)'x'U$$

$$E(b) = \beta + (x'x)^{-1}x'E(U) = \beta \text{ (since } E(U) = 0 \text{ by assumption)} \quad (2.26)$$

Variance and covariance matrix estimator:

$$Var(b) = E\{(b - E(b))(b - E(b))'\}$$

$$b - E(b) = b - \beta = (x'x)^{-1}x'U$$

$$Var(b) = E\{((x'x)1x'U)(x'x)x'U)'\}$$

$$= E\{(x'x)^{-1\prime}x'UU'x(x'x)^{-1}\}$$

$$= (x'x)^{-1\prime}E(UU')x(x'x)^{-1}$$

$$= \sigma^2 (x'x)^{-1}x'x(x'x)^{-1}$$

$$= \sigma^2(x'x)^{-1} \quad (2.27)$$

The elements on the main diagonal gave the sampled variance of the corresponding element of **b** and the off diagonal term gave the sampled covariance.

2.5.7 Gauss Markov theorem

The most essential consequence of the least square theorem, according to the gauss markov theorem, was that no other linear unbiased estimator can have less sample variances than the ols estimator. The best linear unbiased estimators (BLUE) were those that had the least variation within the class of linear unbiased estimators Ruey (2002).

2.5.8 Properties of ols linear model

$$b = (x'x)^{-1}x'y$$

the matrix $(x'x)^{-1}x'$ were fixed numbers, *b* is a linear function of y. by definition, *b* was a linear estimation.

i. **Unbiased:**

$$Y = x\beta + U$$

$$b = (x''x)^{-1} x''(x\beta + U)$$

$$= \beta + (x''x)^{-1}x''U$$

$$E(b) = \beta \qquad (2.28)$$

ii. **Minimum variance**

Let $\hat{\beta}$ be any other linear estimator of β

Let $\hat{\beta} = [(x'x)^{-1}{}'x' + c]y$

$$\hat{\beta} = [(x'x)^{-1}{}' + c][x\beta + U]$$

$$\hat{\beta} = \beta + cx\beta + (x'x)^{-1}{}'x'U + cU$$

if $\hat{\beta}$ is unbiased, then

$$E(\hat{\beta}) = \beta \text{ iff } cx = 0$$

Thus; $Var(b) = \sigma^2 (x'x)^{-1}$

$$Var\left(\hat{\beta}\right) = E\left[\hat{\beta} - E(\hat{\beta})\right]\left[\left(\hat{\beta} - E(\hat{\beta})\right)'\right]$$

$$Var\left(\hat{\beta}\right) = E\left[\left(\hat{\beta} - \beta)\right)\left(\hat{\beta} - \beta\right)'\right]$$

$$\hat{\beta} - \beta = (x'x)^{-1} x'U + cU \qquad (2.29)$$

$$Var(\hat{\beta}) = E\left[\left[(x'x)^{-1} x'U + cU\right]\left[(x'x)^{-1} x'U + cU\right]'\right] \qquad \text{since cx = 0}$$

$$Var(\hat{\beta}) = E\left[(x'x)^{-1} x'x(x'x)^{-1} UU' + (x'x)^{-1} cxUU' + (x'x)^{-1} xcxUU' + cc'UU'\right]$$

$$Var(\hat{\beta}) = E\left[(x'x)^{-1} UU' + cc'UU'\right]$$

$$Var(\hat{\beta}) = E\left[(x'x)^{-1} UU' + cc'UU'\right]$$

$$Var(\hat{\beta}) = (x'x)^{-1} \sigma^2 + cc'\sigma^2$$

$$Var(\hat{\beta}) = \sigma^2\left[(x'x)^{-1} + cc'\right] \qquad (2.30)$$

The variances of **a** given element of $\hat{\beta}$ must necessarily be equal to or greater than corresponding element of **b**, which shows that **b**, was *BLUE*. If c = 0 then var($\hat{\beta}$) = **b.**

2.5.9 Variance of disturbance term (σ_u^2)

The estimate of σ_u^2 is based on the residual sum of squares e'e.

$$\begin{aligned} e &= y - xb \\ &= y - x(x'x)^{-1}x'y \\ &= [1 - x(x'x)^{-1}x']y \\ &= My \qquad (2.29) \end{aligned}$$

where $M = I_n - x(x'x)^{-1}x'$

$Mx = 0$

$$\begin{aligned} e &= M(x\beta + U) \\ &= Mx\beta + MU \\ &= MU \qquad \text{The matrix M was symmetric idempotent} \end{aligned}$$

$e'e = U'M'MU = U'MU$

The expectation was defined as:

$$\begin{aligned} E(e'e) &= E(U'MU) \\ &= E\{tr(U'MU)\} \\ &= E\{tr(MUU')\} \\ &= tr(M)E(UU') = tr(M)\,\sigma^2 \end{aligned}$$

$$\begin{aligned} tr(M) &= tr[I_n - x(x'x)^{-1}x'] \\ &= tr(I_n) - tr[(x'x)^{-1}x'x] = tr(I_n) - tr(I_k) \\ &= n - k \qquad (2.30) \end{aligned}$$

When $E(e'e) = \sigma^2(n - k)$

$$\sigma^2 = \frac{E(e'e)}{n-k}$$

$$S^2 = \sigma^2 = \frac{e'e}{n-k}$$

$$S^2 = \frac{(y'y - \beta'x'y)}{n-k} \qquad (2.31)$$

The **e** was defined as:

$$e = y - \hat{y} \qquad (2.32)$$

where $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k X_k$

S^2 gave an unbiased estimator of the disturbance variance.

2.5.10 Error term and the residuals

Olubusoye, (2002) emphasized that it would be a good statistical practice to examine the residuals from a fitted model for evidence of departure from the models assumptions. However, the construction of significance tests to test the discrepancies observed was not always a straight forward matter. The true errors of the model were denoted by {**e**} and the observed residuals by {$\mathbf{z_i}$}. As usual, the parameters of the model were consistently estimated, it would seemed tempting to assumed that a test statistic based on the $\mathbf{z_i}$‘‘s would have the same asymptotic distribution on the null hypothesis as the corresponding quantity computed from the $\mathbf{e_i}$‘‘s.

In a sample of **n** bivariate observations ($\mathbf{y_i}$ $\mathbf{x_i}$) fitted regression model by least squares estimation the plotted scatter diagram of residuals **z** against a third variable $\mathbf{x_2}$, to test for association between **z** and $\mathbf{x_2}$. Using an obvious notation assumptions, as $a = \sum x_2 z / \sum x_2^2$ naive test procedure would be to compute the regression coefficient of **z** on $\mathbf{x_2}$, namely $t = a\left[\sum x_2^2\right]^{\frac{1}{2}} / s$ together with the residual mean square $\mathbf{S^2}$, and to treat the statistic as asymptotically N(0, 1) standard normal. The ‘justification’ for that procedure was that the least squares estimator **b** of **β** is consistent. Hence, for each 1, $\mathbf{z_i}$ converges in probability to $\mathbf{e_i}$, while the quantity analogous to **t** in which each $\mathbf{z_i}$ is replaced by $\mathbf{e_i}$ was known to be asymptotically N(0, 1).

It would simply not be possible to put up and fit an adequate "fill model" for each test in many cases, particularly where the statistician intended to plot and analyze the residual in a number of ways from a data-analysis standpoint. In order to analyze the performance of the obvious naïve test and create adjustments where the naive test was invalid, it was necessary to give some general thought to tests based on observed residuals.

2.5.11Randomness of error term

The error term **U** was assumed to be a random variable *iff* **U** assumed various values by chance. Each element of U should be individually unimportant and should assume positive, negative and zero values.

2.5.12 Zero mean of error term

The error **U** may take values which had a zero mean. The population of all conceivable values of **U** for each period contained positive, negative or zero values all of which added up to zero. The essence of the zero mean assumption was considered axiomatically true that positive and negative values of the error term **U** have a sum equal to zero.

2.5.13 Homoscedasticity of error term

The error term **U** had a constant variance. That assumption was also called the *twin assumption*.

2.5.14 Normality of error term

The error term followed a multivariate normal distribution with zero mean and constant variance. That assumption enabled inferences to be made about the model parameters.

2.5.15 Properties of the residuals ($\hat{u}$)

For a linear model:

$$Y = XB + U \qquad (2.33)$$

The residual being the difference between the observed value of Y and the predicted value βX can be written as:

$$\hat{U} = Y - X\hat{b} \qquad (2.34)$$

where $\hat{b}$ = $(X'X)^{-1}X'Y$ is the ordinary least square estimator of β; $\hat{u}$ represented the *vector of residuals.* In a variety of ways, its elements can be plotted and otherwise investigated to see if they suggested that assumptions inherent in the assumed model were not being upheld.

Several elementary but important properties of residuals were noted as:

$$\hat{u} = Y - X(X'X)^{-1}{}'X'Y$$

$$= [I - X(X'X)^{-1}{}'X']Y$$

$$= PY \qquad (2.35)$$

where P defined as: $P = I - X(X'X)^{-1}X'$ (P satisfy idempotent property) Also,

$$PY = P(Xb + U)$$

$$= (I - X(X'X)^{-1}X')(Xb+U)$$

$$= Xb + U - Xb + X(X'X)^{-1}X'U$$

$$= U - X(X'X)^{-1}X'U$$

$$= [I - X'X)^{-1}X']U = PU \qquad (2.36)$$

The expected value of the residuals was zero.

That implied;

$$E(\hat{u}) = E(PY)$$

$$= E(PU)$$

$$= PE(U)$$

$$= 0 \qquad (2.37)$$

The variance of the residuals $P\sigma^2I$

$$Var(u) = Var(PU)$$

$$= E(\hat{u}\,\hat{u}')$$

$$= E(U'P'PU)$$

$$= E(U'PU) \text{ since P is idempotent}$$

$$= PE(U'U) \qquad (2.38)$$

$$= P\sigma^2I.$$

The covariance of the residuals and the observed value y is $P\sigma^2$

$$Cov(\hat{u}, \hat{y}) = E(\hat{u}' \hat{y})$$

$$Cov(\hat{u}, \hat{y}) = E(P'U'X'\beta)$$

$$= P'XE(U'\beta) = 0 \text{ since } PX = 0 \text{ (symmetric and idempotent)}$$

There were cases of 'outliers' among the residuals $\hat{u}$. An outlier was considered to be an observation which was far in absolute value from the range of other values. It would be three or four times the standard error in absolute value.

2.5.16 Generalized least square under non-spherical disturbance

There were situations when the assumption about the error term defined as: $E(UU') = \sigma^2 I$ fails. This implies that; the disturbance variance was not constant at each observation point (heteroscedasticity condition). The disturbance covariance at all possible pairs of observation point were not zero (auto correlation condition).
The general non-spherical disturbance matrix was specified as:

$$E(UU') = \sigma^2\Omega = V \quad (2.39)$$

where Ω and V matrices are assumed to be positive definite.

2.5.16 Generalized Least Square Estimator

We now define the model as;

$$Y = X\beta + U$$

When $E(UU') = \sigma^2\Omega$

When pre-multiply the assumed model by some (n x n) nonsingular transformation matrix T to obtain

$$TY = (TX)\,\beta + TU \quad (2.40)$$

Each element in the vector TY was linear combination of the elements in Y.

$$E(TUU'T') = \sigma^2 T\Omega T' \quad (2.41)$$

Since E(TU) = 0. If it were possible to specify T such that

$$T\Omega T' = 1$$

then we could apply OLS to the transformed variables TY and TX. Since Ω is a symmetric positive definite matrix, a nonsingular matrix P can be found such that

$$\Omega = PP$$

since P is non-singular,

$P^{-1}\Omega P^{-1'} = I$ and

$(P^{-1})'P^{-1} = \Omega^{-1} = T'T$ (2.42)

Pre-multiply the model

$Y = X\beta + U$ by P^{-1}

$P^{-1}Y = P^{-1'}X\beta + P^{-1}U$

$y^* = x^*\beta + u^*$ (2.43)

where;

$y^* = P^{-1}Y$

$x^* = P^{-1}X$

$u^* = P^{-1}U$

$E(u^*u^{*'}) = E[(P^{-1}U)(P^{-1}U)']$

$=E[P^{-1}UU'P^{-1}]$

$= P^{-1}E(UU')^{-1}$

$= \sigma^2 P^{-1}\,\Omega\, P^{-1'}$

$= \sigma^2 I$ (2.44)

Equation (2.30) satisfy all the assumptions required for the OLS model.

Applying OLS to Eq. (2.27) gives

$b^* = (X'T'TX)^{-l}\,X'T'TY$

$= (X'\Omega^{-1}X)^{-1}X'\Omega^{-1}Y$ (2.45)

with the variance-covariance matrix given by

$var(b^*) = \sigma^2(X'\Omega^{-1}X)^{-1}$ (2.46)

The estimator b* is defined to be the generalized least squares GLS estimator, b* is a best linear unbiased estimator of the model:

$Y = X\beta + U$ with

$E(UU') = \sigma^2\Omega$

An unbiased estimator of variance may be derived from the application of OLS:

$$S^2 = \frac{(TY - TXb^*)'(TY - TXb^*)}{n-k}$$

$$S^2 = \frac{(Y - Xb^*)'\ T'\ \ T\,(Y - Xb^*)}{n-k}$$

$$S^2 = \frac{(Y - TXb^*)'\Omega^{-1}(Y - Xb^*)}{n-k}$$

$$S^2 = \frac{Y'\Omega^{-1}Y - b^*X'\Omega^{-1}Y}{n-k} \quad (2.47)$$

on assumption of normality for the disturbance term

H_0: Rβ = r was based on Fisher ratio F defined as:

$$F = \frac{(r - Rb*)'\left[R\left(X'\Omega^{-1}X\right)^{-1}R'\right]^{-1}(r - Rb*)/q}{S^2} \tag{2.48}$$

had the $F_{(q,\ n\ -\ k)}$ distribution under the null hypothesis, where b* was the GLS estimator and S^2 the variance estimator. The above formulae were only operational if the elements of **Ω** were known. In most practical cases, they were not known.

2.5.17 General Least Squares Estimator

We denoted the variance-covariance matrix of the residuals by V i.e. Var($\hat{U}$) = V

the method of estimation involves minimizing

S = (Y - Xβ)V^{-1}(Y - Xβ) with respect to β

Thus:

$$\frac{\partial s}{\partial \beta} = 2x'V^{-1}x\beta - 2x'V^{-1}y \tag{2.49}$$

$$\frac{\partial s}{\partial \beta} = Q$$

S is minimized by setting Q to zero, i.e.

$$\frac{\partial s}{\partial \beta} = 0$$

$$\Rightarrow \quad x'V^{-1}x\beta = x'V^{-1}y \tag{2.50}$$

$$\beta = \left(x'V^{-1}x\right)^{-1} x'V^{-1}y$$

The generalized least square estimator becomes the ordinary least square estimator if the variance-covariance matrix V = σ^2I

2.5.18 Autoregressive Moving Average (arma) Process

Dependence was common feature in time series observations; given a series y_t, we can model that the level of its current observations depends on the level of its lagged observations. That can be represented by an Autoregressive (ar) model. AR of order one ar(1) can be define as:

$$y_t = \phi_1 y_{t-1} + \varepsilon_t \tag{2.51}$$

Where $\varepsilon_t \sim WN(0, \sigma_t^2)$ which is a strict assumption

Also, Autoregressive of order p; ar(p) can be define as:

$$y_t = \phi_1 y_{t-1} + \phi_{2t-2} + \ldots + \phi_{p\,t-p} + \varepsilon_t \qquad (2.52)$$

It can also be model that the observations of a random variable at time t were not only affected by the shock at time t, but also the shocks that had taken place before time t. if negative shock was observed in an economy (natural disaster), then it would be expected that the effect affected the economy not only for the time it took place, but also for the near future. That concept was represented by Moving Average (MA) model. The ma(1) can be define as:

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} \qquad (2.53)$$

And ma(q) defined as:

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \qquad (2.54)$$

If that two models were combined to obtain Autoregressive Moving Average of order p, q (arma (p, q)) model define as:

$$y_t = \phi_1 y_{t-1} + \phi_{2t-2} + \ldots + \phi_{p\,t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \qquad (2.55)$$

arma model provided one of the basic tools in time series modeling.

2.5.19 Lag operators

To re-write arma model in a short form, Lag (Backward Shift) operators denoted as L or B was applied. The concept was; when moved the index back one time unit and applying it k times, we move the index back k units as define:

$$By_t = y_{t-1}$$
$$B^2 y_t = y_{t-2}$$
$$\vdots$$
$$B^k y_t = y_{t-k} \qquad (2.56)$$

The lag operator is distributed over the addition operator as:

$$L(x_t + y_t) = x_{t-1} + y_{t-1} \qquad (2.57)$$

When applied the lag operator on arma model defined as:

$$AR(1):(1-\phi L)y_t = \varepsilon_t$$
$$AR(p):\left(1-\phi_1 L-\phi_2 L^2-...-\phi_p L^p\right)y_t = \varepsilon_t$$
$$MA(1): y_t = (1+\theta L)\varepsilon_t$$
$$MA(q): y_t = \left(1+\theta_1 L+\theta_2 L^2+...+\theta_q L^q\right)\varepsilon_t \quad (2.58)$$

Let $\phi_0 = 1$ also $\theta_0 = 1$ and define lag polynomials as:

$$\phi(L) = 1-\phi_1 L-\phi_2 L^2-...-\phi_p L^p$$
$$\theta(L) = 1+\theta_1 L+\theta_2 L^2+...+\theta_q L^q$$

(2.59)

With lag polynomials arma process would be rewritten in a more compacted form as:

$$AR:\phi(L)y_t = \varepsilon_t$$
$$MA: y_t = \theta(L)\varepsilon_t$$
$$ARMA:\phi(L)y_t = \theta(L)\varepsilon_t \quad (2.60)$$

2.5.20 Non-Stationary Time Series

Shittu and Yaya (2016) expressed that when the mean μ and variance σ^2 of a series X_t changes systematically with time *t* then the series X_t was referred to as been nonstationary series.

Mathematically expressed:

$$E(X_t) \neq \mu \quad \text{but} \quad \mu_t$$
$$Var(X_t) \neq \sigma^2 \quad \text{but} \quad \sigma_t^2$$

The dynamic time series techniques would no longer be appropriate in modeling the series. Hence, nonstationary series can be differenced a number of time *d* to become stationary of order *d* or *I(d)*.

An autoregressive time series ar(1) of the form:

$$y_t = \alpha y_{t-1} + \varepsilon_t$$

Was considered stationary if $|\alpha| < 1$and $\varepsilon_t \approx N(0,\sigma^2)$. The series y_t tends to return to its mean value and fluctuate around it within a more or less constant range and the variance of y_t was finite. However, if $\alpha = 1$ then y_t was nonstationary the roots of the characteristic equation lied on the unit circle. It was possible to rearrange and accumulate y_t for different periods y_{t-n} to obtained:

$$y_t = \alpha^n y_{t-n} + \sum_{j=0}^{n-1} \alpha^j \varepsilon_{t-j} \qquad (2.61)$$

As n $n \to \infty$ the equation has a constant means μ and variance $\left(\frac{\sigma^2}{1-\alpha^2}\right)$. The question of whether a variable or series is stationary depends on whether or not it had a unit root. Equation $y_t = \alpha y_{t-1} + \varepsilon_t$ rewritten as:

$$(1-\alpha L)y_t = \varepsilon_t \qquad (2.62)$$

Where L denoted the lag operator ($Ly_t = y_{t-1}$) and $L^2 y_t = y_{t-2}$….. $L^k y_t = y_{t-k}$

The characteristic root was $(1-\alpha L) = 0$ it could be observed that the roots of the equation were all greater than unity in absolute terms. In the equation there was only one root $(L = \frac{1}{1-\alpha})$ series. Stationarity requires that $|\alpha| < 1$, if $|\alpha| > 0$ the y_t would be nonstationary and explosive.

Consider an autoregressive process of order p; ar(p) model given as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \ldots + \phi_P Y_{t-P} + \varepsilon_t \qquad (2.63)$$

Which can also be written using backshift method as:

$$\psi(L) = \varepsilon_t$$

Where $\psi(L) = 1 - \phi_1 L^1 - \phi_2 L^2 - \ldots - \phi_P L^P$ was a polynomial in lag L. if the roots of the characteristic equation $\psi(L) = 0$ were all greater than unity in absolute term, was said to be stationary, otherwise y_t was nonstationary.

Using non stationary series produces unreliable and spurious results and leads to poor understanding of the process it represent. Also the forecast performance may be very poor.

2.5.21 Spurious regression

According to Shittu and Yaya, (2016) spurious or nonsense regression occurs when one or more of the ordinary least square assumptions were violated. That also occurs in non-stationary series where dynamic models arima (p, q) were no longer appropriate.

A regression of two non-stationary and uncorrelated series y_t and x_t as:

$$y_t = \beta_0 + \beta_1 x_t + \varepsilon \quad \text{where;}$$
$$\mathrm{y_t} = y_{t-1} + \mu_t,\ \mu_t \approx N(0,1) \qquad (2.54)$$
$$x_t = x_t + \mu_t, \quad \mu_t \approx N(0,1) \text{ and } \varepsilon_t \text{ is the disturbance term.}$$

It should be expected that $\beta_1 = 0$ and the coefficient of determination R should tend towards zero. However, because of the non-stationary nature of the data, implying that ε_t was also non-stationary, any tendency for both series to be growing due to correlation which was picked up by the regression model even though each series was growing for very different reason and at rates that were uncorrelated. It y_t was differenced once we had:

$$\Delta y_t = \phi_0 + \phi_1 \Delta X_t + \varepsilon_t \qquad (2.55)$$

It converged to zero in probability. Thus, correlation between non-stationary series do not imply the kind of causal relationship that might be inferred from stationary series. That gave important reason which regression models – deterministic or dynamic need to be diagnosed for normality and stationarity to ensure validity. For dynamic regression, each of the involving series had to be subjected to stationarity test before processing for causal relationship between them.

2.5.22 Assumptions of error term and properties of residuals

The expected value of the error term was the same as that of the residual which was zero. The variance was assumed to be constant throughout in the case of the error term, but the variance of the residual was not constant. Instead a variance-covariance matrix existed for the residuals.

Var(U) = σ^2I (constant) but Var(u) = P σ^2I (2.56)

where P = (I - X(X''X)$^{-1}$X')

Also the error term was assumed to follow a normal distribution with mean zero and constant variance σ^2. The residuals on the other hand may not necessarily follow a normal distribution it may follow distributions like exponential, Poisson etc or even normal distribution with non-zero mean and having a variance-covariance matrix i.e. non-central normal distribution.

2.5.23 Estimation of parameters of the model such that violations were removed

In most cases as discussed earlier the variances of the residuals may not be constant throughout. Thus the residuals had a variance-covariance matrix denoted by $P\sigma^2$ instead of $\sigma^2 I$ as assumed for the error term. If the ordinary least squares estimator was applied in that case of estimating the parameters of the model the variances of the estimates would no longer be minimum although they will still remain unbiased. The generalized least square estimator was therefore applied.

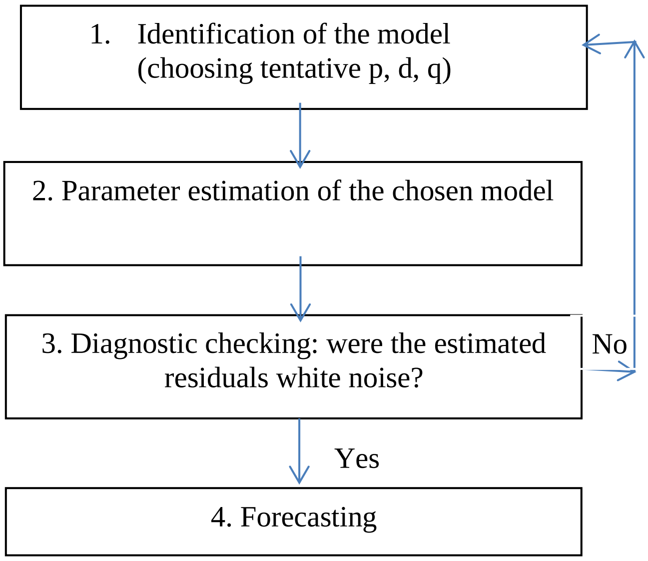


Fig. 2.3 Flowchart of Box-Jenkins framework (Shaib and Umar, 2019)

2.6 Review of related studies

According to Shangodoyin, (2002), various economic and financial time series data encountered in practice: plots of nominal Gross Domestic Product (GDP) for each given country had a tendency to trend upward over time, and that upward trend should be considered into any series forecast. The introduction of a deterministic temporal trend was used to explain trends that grow upward over time.

With stochastic processes, the theoretical breakthroughs in time series analysis began early (probability process). Yule and Walker's work in the 1920s and 1930s was credited with been the first to apply autoregressive models to data. During that period, the moving average was introduced to eliminate periodic fluctuations in time series, such as seasonal swings. Herman Wold proposed the arma (Autoregressive Moving Average) model for stationary series, but he was unable to build a likelihood function that would allow maximum likelihood (ML) parameter estimation. That was completed in 1970, when Box and Jenkins published their classic book "Time Series Analysis," which detailed the entire modeling technique for individual series, including specification, estimate, diagnostics, and forecasting.
Many forecasting and seasonal adjustment strategies can be traced back to the Box-Jenkins (1970) models, which were possibly the most widely utilized.

In order to analyze time series data, the Autoregressive Integrated Moving-Average (arima) model was used. The model has a lot of flexibility when it comes to assessing different time series and making reliable projections. Box and Jenkins (1976) introduced the Arima model technique. The method studied univariate stochastic time series (error term of the time series); nevertheless, the analyzed time series must be stationary for this to be possible. That is, the series' Mean, Variance, and Covariance remained consistent across time. Most economic and financial time series, on the other hand, indicate trends across time.

Stationarity was critical because if the series were non-stationary, all of the traditional regression analysis conclusions would be invalidated. Non-stationary series regressions are referred to as "spurious" since they may have no meaning. A stationary series' long-term forecasts would converge to the series' unconditional mean Shittu and

Yaya (2016).

Only stationary time series might be used with such techniques. Economic time series, on the other hand, frequently showed a rising trend, indicating non-stationarity and the presence of a unit root. Unit root tests were established mostly in the 1980s.

Dynamic Linear Models: The popularity of Gaussian Distribution can be attributed to a number of factors. To begin with, many physical processes were roughly Gaussian; this was linked to the central limit theorem. Second, it was both analytically elegant and user-friendly. Third, it was simple to modify; if a Gaussian variable was propagated through a linear function with a Gaussian output variable Lee and Roberts, (2008).

Dynamic Linear Models: The Gaussian Distribution is quite popular for a variety of reasons. To begin with, many physical processes were Gaussian in nature; this was linked to the central limit theorem. Second, it was easy to use and analytically elegant. Finally, if a Gaussian variable was propagated through a linear function with a Gaussian output variable, it could be easily adjusted Lee and Roberts, (2008).

Shittu and Yaya (2016) emphasized that a behavioral process over a period of time had relied on time series models; time series models have a wide range of applications, including sales forecasting, weather forecasts, inventory investigations, and so on. Time series models have been found to be one of the most successful approaches of forecasting decisions that contain future uncertainty. Most of the time, the future course of action and decisions for such processes are determined by the expected outcome. The necessity for those expected results had prompted businesses to develop forecasting tools in order to be better prepared for an uncertain future. Those models can also be integrated with other data mining approaches to aid in understanding data behavior and the prediction of future trends and patterns in data activity.

There are two approaches to modeling non-Gaussian time series: keep the general autoregressive moving average (arima) framework and accept non-Gaussian white noise, or forgo the linearity assumption. In the first case, the challenge was determining the suitable white noise distribution so that the time series model exhibited a non-Gaussian feature. In the latter scenario, one must choose from an infinite number of nonlinear forms that typically express the time series as a nonlinear function of its lagged values to find an appropriate explicit model.

The exchange rate is the value of one country's currency in terms of another country's currency. It is also known as the worth or value of a country's money in terms of the currency of another country. Changes in exchange rate policies to correct both internal and external sectors cause exchange rates to vary from time to time. Because exchange rates fluctuate, the rate can either appreciate or depreciate.

Jimoh, (2017) observed that if the amount of home country currency required to purchase a foreign currency decreases, the exchange rate appreciates; if the amount of home country currency required to purchase a foreign currency increases, the exchange rate depreciates. It has been simple to quantify or determine the extent of external sector activity in Nigeria's economy since the country's exchange rate was established in 1986. In a similar vein, determining the extent to which the Nigerian economy affects foreign trade had been easier. The author also cited that if the amount of home country currency required to purchase a foreign currency decreases, the exchange rate appreciates; conversely, if the amount of home country currency required to purchase a foreign currency increases, the exchange rate depreciates. It has been simple to quantify or determine the extent of external sector participation in Nigeria's economy since 1986, when the country's exchange rate was first established.
The inventory was divided into four categories: high turnover products, medium turnover items, and low turnover items, according to the article. Because of its limited number of parameters and high % best fit, the article concluded that armax is the best linear model. Furthermore, using a mathematical inventory model, we will be able to monitor the system and implement suitable control techniques. However, the article did not take into account the multiple outlet system or the seasonality impact, which are both common in time series data.

Bruce, (2013) modeled deformity at Southern Maine using arma and arimax time series models. Two statistical approaches to forecasting long-term disability benefit claims were considered in that study. Although both models were capable of producing accurate four-quarter projections, the later model was able to reflect the impact of external factors such as the health of the economy and management controllable policies. Both models, on the other hand, outperform the widely used seasonally adjusted four-quarter moving average (SAMA) model. However, data from monthly and yearly time series were not taken into account.

Wei, (2002) investigated the armax model's least square identification. He found that the standard least square method's success was often attributed to its simple concept and ease of implementation. The main disadvantage was that least square parameter estimates were only unbiased in the rare case where the underlying system model's equation error was white noise. The study introduced an auxiliary linear regression model that was equivalent to the armax model, which was an innovation. In an innovative usage of extra delayed outputs, an estimated noise covariance vector was then obtained from that auxiliary model, which identifies the source of the noise induced bias in the least square estimate. To accomplish unbiased identification of the model, the paper developed Bias Error Least Square (BELSX) with extra delayed output. The influence of the bias on the results, on the other hand, stemmed from Monte-Carlo simulations of real-world data, which were not checked.

Yaya, (2016) looked at the persistence of volatility and asymmetry in naira exchange rates before and after the crisis. The study used two time series modeling approaches: fractional and generalized autoregressive conditional heteroscedasticity (GARCH). The paper used six daily official naira exchange rates, including those with central and West African Francs (naira-CFA) and US Dollars (naira-USD). The analysis found that during the post-crisis period, the US Dollar exchange rate rose astronomically in all series, before stabilizing after a few months. The use of the US dollar as a common foreign currency in Nigeria was credited with the stability.

Ranjeeta Bisoi, (2014) referenced work on hybrid decision tree: revealed that forecasting and categorization were two significant characteristics of time series analyses. Because of its flexibility in modeling numerous stationary processes, statistical-based approaches such as linear autoregressive (AR) models have been used extensively in time-series forecasting. As also observed by Fan and Yao, (2003) and Weron and Misiorek, (2008).

The model takes a linear relationship between the delayed variables as well as provides rough estimate of real-world problems, failing to effectively anticipate the evolution of nonlinear and non-stationary processes in most cases. When changes occurred with time seasons the elements would rapidly fluctuate the data. arma model performance suffers significantly. At times, methods based on the evolution of the increments were employed in minimizing initial lack of stationarity.

However, because differencing enhances high-frequency noise in time series, determining the order of an arima model requires a significant amount of effort. Most traditional arima models were limited to solving first-order non-stationary problems in economic data. Engle, (1982) initiated arch model to solve 2^{nd} order non stationary conditional variance.

Bollerslev, (1986) originated the generalized autoregressive conditional heterodasticity (GARCH) model which depicts the error (variance) as an expression of autoregressive process, allowing for a more compact description of the time series. Tong (1990) also presented threshold nonlinear arma models (TAR), which have been successfully employed for modeling time domains economic series as observed by Yadav, (1994).

Lineesh, (2010) applied wavelet to partition sequence into orthogonal pattern series; then used the model to forecast each decomposed series using arma and tar models.

Krishnamurthy and Yin, (2002) worked on nonlinear time series forecasting using hidden Markov and AR models regime, the study switched AR parameters in time in accordance to the outcome finite-state Markov chain. However, the property of linearity in autoregressive structure limits the method for nonlinear and stable time series forecasting.

According to Park and Sandberg, (1991) observed that neural network approaches had several distinct advantages among which was being nonlinear also useful in complicated modeling.

Zhang, (2012) published a review on NN models on time series forecasting. Forecast had also been done using FNNs with recurrent feedback connections as noted by De Groot and Wuertz, (1991).

Menezes and Barreto, (2008) designed an RNS nonlinear AR model with external recursion capable of capturing trends of various kinds Chen, (1991, 1992).

Barreto, (2007) studied time series forecasting: employed self-organizing map neural network models were reviewed as global models of FNN and used the properties of the distribution in achieving non-linear forecast of series.

As noted by Smola and Scholkop, (2004) SVM-based methods could be used as a class of generalized regression models in which the parameters were derived using convex quadratic support vector regression technique.

Cao, (2003) used LR model to reduce structural risk at upper bound of generalization error; resulted to superior prediction performance. Fu-Yuan Huang, (2008) observed that ANN was an effective forecast tool in a wide range of applications. Some scholars used fuzzy logic theory because it is an effective tool for dealing with uncertainties. That was applied in Taiwan stock exchange forecast as noted by Cheng, (2007).

Yu Lixin, (2005) forecasted financial time series with FNN model with genetic and gradient descent learning algorithms utilized alternately by substitution method updated the parameters until the error was minimal. Slim Chokri, (2006) used Mackey glass time series data to fit HNF based on Kalman filter in predicting monetary series.

Fu-yuan Huang, (2008) used robust PSO steps to analyze stock market data.

CHAPTER THREE

METHODOLOGY

3.1 Methodology of the study

We discussed the research study design/framework, approach, and methodologies. Exploratory methodological framework was used to project the past behavior of non-stationary economic time series data into the future.

3.2 Data source and data collection

The real-life data used were collected from publication of Central Bank of Nigeria (2017) and Nigerian National Petroleum Corporation (2017). The data on economic time series variables: Monthly External Reserves (Million USD) denoted as Y_t Monthly Official Exchange Rate (Naira to 1 USD) denoted as X_{t1}, Monthly Crude Oil Export (Million Barrel per Day mbd) denoted as X_{t2} Monthly Crude Oil Price (USD/Barrel) denoted as X_{t3}. The variables X_{t1}, X_{t2} and X_{t3} were the three exogenous variables considered.

Data collection sheet was designed in accordance to the data structure and the variables of interest. Those consist of the Year, Month, Period, External Reserve, Exchange Rate, Crude Oil Export and Crude Oil Price. In-line with the structure the required data were recorded from the source accordingly. The data structure was automated in computer spreadsheet package compactable to statistical packages; MS Excel, Eview, SPSS and R program were used in aiding data storage, data analyses.

3.3 Sample size criteria

Inclusion criteria: Monthly data recorded from January, 2008 to December, 2017. This covered 10 years period of 120 sample size was used as in-sample data to model the process while 12 observations (Jan, - De, 2018) data were used as out-sample data to generate forecast.

3.4 Methods of data analysis

The statistical data analyses tools adopted were: Time plot (Graphical Method), Summary Statistic, Correlogram, Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), Augmented Dickey Fuller (ADF), Q-statistic, Jarqua-Berran Statistic, Unit Root Test, Residual Test, Forecast plot, Akaike Information Criteria (AIC), Log-likelihood, Mean Square Error Shittu and Yaya (2016). Data analyses processes were aided by computer statistical software packages; R-Program, Eview, Microsoft Excel and SPSS.

3.5 Derivation of arimax with lognormal error term

3.5.1 Conventional arimax with normal error term

Time series mixed model with exogenous variable defined as:

$$y_t = \sum_{i=1}^{p} \phi y_{t-i} + \sum_{i=1}^{q} \theta \varepsilon_{t-i} + \varepsilon_t + \sum_{k=0}^{n} \beta_k X_k \qquad (3.1)$$

where y_t - series (output)

$\sum_{i=1}^{p} \phi y_{t-i}$ - AR with p number of ϕ parameters (Autoregressive Coefficients)

$\sum_{i=1}^{q} \theta \varepsilon_{t-i}$ - MA with q number of θ parameters (Moving Average Coefficients)

ε_t - associated error term

$\sum_{k=0}^{n} \beta_k X_k$ - Regression part with n number of β parameters (Regression Coefficients)

X_k - associated exogenous random variable (input variable).

when expressed (3.1) in backward shift operator, we had :

$$\phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t + \beta_0 + \beta_k X_k \qquad (3.2)$$

Formation of polynomial expression of equation (3.2) defined as:

The Polynomial;

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$$

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q \qquad (3.3)$$

The associated parameters were ϕ, θ, β and σ^2

3.5.2 Derivation of arimax (1, 0, 1) with lognormal error term

$$y_t = \phi_1 y_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t + \beta_0 + \beta_1 x_1 \quad (3.4)$$

expressed in backward shift operator

$$\phi_1(B)(1-B)y_t = \theta_1(B)\varepsilon_t + \beta_0 + \beta_1 \mathrm{x}_1$$

$$\Rightarrow (1\text{-}\phi_1 B)(1-B)y_t = (1+\theta_1)\varepsilon_t + \beta_0 + \beta_1 x_1 \quad (3.5)$$

expresing equation (3.2) in terms of the error ε_t we had :

$$(1\text{-}\phi_1 B)(1-B)y_t - \beta_0 - \beta_1 x_1 = (1+\theta_1)\varepsilon_t \quad (3.6)$$

$$\varepsilon_\mathrm{t} = \frac{(1\text{-}\phi_1 B)(1-B)y_t - \beta_0 - \beta_1 x_1}{(1+\theta_1)} \quad (3.7)$$

Take the natural log of equation (3.7)

$$In\varepsilon_\mathrm{t} = In\left[\frac{(1\text{-}\phi_1 B)(1-B)y_t - \beta_0 - \beta_1 x_1}{(1+\theta_1)}\right] \quad (3.8)$$

$$= \mathrm{In}\left[(1\text{-}\phi_1 B)(1-B)y_t - \beta_0 - \beta_1 x_1\right] \text{-} \mathrm{In}(1+\theta_1) \quad (3.9)$$

The equation of lognormal function defined as:

$$f(y;\mu,\sigma^2) = \frac{1}{y_t\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2}\left|\frac{In(y_t)-\mu}{\sigma}\right|^2\right. \quad (3.10)$$

$$for\ 0 < \mu < \infty,\ 0 < y < \infty,\ \sigma > 0$$

$In\varepsilon_t$ = $In(y_t) - \mu$ representing the residual

Substitution of Error term of ARIMAX (1, 1, 1) Model into lognormal function;

$$\Rightarrow f(\varepsilon_t) = \frac{1}{y_t\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2}\left|\frac{In(\varepsilon_t)}{\sigma}\right|^2\right. \quad (3.11)$$

When we substitute equation (3.9)

$$\Rightarrow f(\varepsilon_t) = \frac{1}{y_t\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{1}{2}\left|\frac{\mathrm{In}\left[(1\text{-}\phi_1 B)(1-B)y_t - \beta_0 - \beta_1 x_1\right] \text{-} \mathrm{In}(1+\theta_1)}{\sigma}\right|^2\right. \quad (3.12)$$

Expressing equation (3.12) in likelihood function we had:

$$L(\varepsilon_t) = \frac{(2\pi\sigma^2)^{-\frac{n}{2}}}{\prod_{t-1}^{n} y_t}\exp\left\{-\frac{1}{2\sigma^2}\left|\mathrm{In}\left[(1\text{-}\phi_1 B)(1-B)y_t - \beta_0 - \beta_1 x_1\right] \text{-} \mathrm{In}(1+\theta_1)\right|^2\right. \quad (3.13)$$

When we take the log of likelihood function $L(\varepsilon_t)$ we had:

$$InL(\varepsilon_t) = \frac{n}{2}In(2\pi\sigma^2) - \sum(Iny_t) - \frac{1}{2\sigma^2}\sum\left|\mathrm{In}\left[(1\text{-}\phi_1 B)(1-B)y_t - \beta_0 - \beta_1 x_1\right] \text{-} \mathrm{In}(1+\theta_1)\right|^2 \quad (3.14)$$

Also by expansion

$$(1-\phi_1 B)(1-B)y_t = y_t - By_t - By_t - \phi_1 By_t + \phi_1 B^2 y_t = y_t - \mathrm{y}_{t-1} - \phi_1 \mathrm{y}_{t-1} + \phi_1 \mathrm{y}_{t-2} \quad (3.15)$$

When substitute into (3.14) we obtained

$$InL(\varepsilon_t) =$$

$$\frac{n}{2} In(2\pi\sigma^2) - \sum (Iny_t) - \frac{1}{2\sigma^2} \sum \left| \mathrm{In}\left[y_t - \mathrm{y}_{t-1} - \phi_1 \mathrm{y}_{t-1} + \phi_1 \mathrm{y}_{t-2} - \beta_0 - \beta_1 x_1\right] - \mathrm{In}(1+\theta_1) \right|^2 \quad (3.16)$$

To estimate the associated parameters we took partial derivative of equation (3.16) with respect to ϕ, θ, β and σ^2 respectively and equated to zero.

3.5.3 Estimation of ϕ_1 parameter

Partial derivative of equation 3.16 with respect to ϕ_1 when equated to zero we obtained :

$$\frac{\partial InL(\varepsilon_t)}{\partial \phi_1}$$

$$= -2\sum \left| \frac{In(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)}{2\sigma^2} \right| \sum (y_{t-2} - y_{t-1}) = 0$$

$$\Rightarrow \left(\sum y_{t-2} - \sum y_{t-1}\right) In \sum \left|\left(y_t - y_{t-1} - \phi_1 y_{t-1+} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1 - In(1+\theta_1)\right)\right| = 0$$

$$\Rightarrow In \sum \left|\left(y_t - y_{t-1} - \phi_1 y_{t-1+} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1\right)\right| = In(1+\theta_1) \quad (3.17)$$

When open brackets and collect like terms:

$$\Rightarrow \sum y_t - \sum y_{t-1} + \phi_1 \sum (y_{t-2} - y_{t-1}) - n\beta_0 - \beta_1 \sum x_1 = 1 + \theta_1 \quad (3.18)$$

solve for ϕ_1

$$\phi_1 \sum (y_{t-2} y_{t-1}) = 1 + \theta_1 + \sum y_{t-1} - \sum y_t + n\beta_0 + \beta_1 \sum x_1 \quad (3.19)$$

$$\Rightarrow \phi_1 \sum (\mathrm{y}_{t-2} - y_{t-1}) = (1+\theta_1) + \sum (y_{t-1} - y_t) + n\beta_0 + \beta_1 \sum x_1 \quad (3.20)$$

Therefore;

$$\hat{\phi}_1 = \frac{(1+\hat{\theta}_1) + \sum (y_{t-1} - y_t) + n\hat{\beta}_0 + \hat{\beta}_1 \sum x_1}{\sum (y_{t-2} - y_{t-1})} \quad (3.21)$$

3.5.4 Estimation of θ_1 parameter

Partial derivative of equation (3.16) with respect to θ_1 and equate to zero:

$$\frac{\partial InL(\varepsilon_t)}{\partial\theta_1}$$

$$= -\frac{2}{2\sigma^2}\cdot\frac{1}{1+\theta_1}.In\sum\left|(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)\right| = 0 \qquad (3.22)$$

$$\Rightarrow \frac{In\sum\left|(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)\right|}{\sigma^2(1+\theta_1)} = 0 \qquad (3.23)$$

$$\Rightarrow In\sum\left|(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)\right| = 0 \qquad (3.24)$$

$$\Rightarrow In\sum\left|(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1)\right| = In(1+\theta_1) \qquad (3.25)$$

$$\therefore \hat{\theta}_1 = \left|\sum(y_t - y_{t-1}) - \hat{\phi}_1\sum(y_{t-1} - y_{t-2}) - n\hat{\beta}_0 - \hat{\beta}_1\sum x_1 - 1\right| \qquad (3.26)$$

3.5.5 Estimation of β_0 Parameter

Partial derivative of equation (3.16) with respect to β_0 and equate to zero:

$$\frac{\partial InL(\varepsilon_t)}{\partial\beta_0}$$

$$= -\frac{2}{2\sigma^2} \cdot \frac{In\sum\left|(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)\right|}{\sum(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1)} = 0 \qquad (3.27)$$

$$\Rightarrow In\sum\left|(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)\right| = 0$$

$$\Rightarrow In\sum\left|(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1)\right| = In(1+\theta_1)$$

$$= \sum(y_t - y_{t-1}) - \phi_1\sum(y_{t-1} - y_{t-2}) - n\beta_0 - \beta_1\sum x_1 = 1+\theta_1$$

$$n\beta_0 = \sum(y_t - y_{t-1}) - \phi_1\sum(y_{t-1} - y_{t-2}) - \beta_1\sum x_1 - 1 - \theta_1$$

$$\therefore \hat{\beta}_0 = \frac{\sum(y_t - y_{t-1}) - \hat{\phi}_1\sum(y_{t-1} - y_{t-2}) - \hat{\beta}_1\sum x_1 - 1 - \hat{\theta}_1}{n} \qquad (3.28)$$

3.5.6 Estimation of β_1 parameter

Partial derivative of equation (3.16) with respect to β_1 and equate to zero:

$$\frac{\partial InL(\varepsilon_t)}{\partial\beta_1}$$

$$= -\frac{2}{2\sigma^2}.\frac{-\sum x_1\left(In\sum\left|(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)\right|\right)}{\sum(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1)} = 0 \qquad (3.29)$$

when cross multiplied

$$\Rightarrow \sum x_1\left(In\sum\left|(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)\right|\right)$$

Distribute $\sum x_1$

$$\Rightarrow In\sum x_1\left(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2}\right) - \beta_0\sum x_1 - \beta_1\sum x_1^2 = In\sum x_1(1+\theta_1)$$

$$= \sum x_1\left(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2}\right) - \beta_0\sum x_1 - \sum x_1 - \theta_1\sum x_1 = \beta_1\sum x_1^2$$

Divide through by $\sum x_1^2$

$$\beta_1 = \frac{\sum x_1\left(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2}\right) - \beta_0\sum x_1 - \sum x_1 - \theta_1\sum x_1}{\sum x_1^2}$$

$$= \frac{\sum x_1\left(y_t - y_{t-1} - \phi_1(y_{t-1} + y_{t-2}) - \beta_0 - 1 - \theta_1\right)}{\sum x_1^2}$$

$$\therefore \hat{\beta}_1 = \frac{y_t - y_{t-1} - \hat{\phi}_1\sum(y_{t-1} - y_{t-2}) - \hat{\beta}_0 - 1 - \hat{\theta}_1}{\sum x_1} \qquad (3.30)$$

3.5.7 Estimation of σ^2 parameter

Partial derivative of equation (3.16) with respect to σ^2 and equate to zero:

$$\frac{\partial InL(\varepsilon_t)}{\partial \beta_1}$$

$$= -\frac{n}{2}\left(\frac{2\pi}{2\pi\sigma^2}\right) + \frac{1}{2(\sigma^2)^2}\sum\left[In(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)\right]^2 = 0$$

$$\Rightarrow \frac{n}{2\sigma^2} = \frac{1}{2(\sigma^2)^2}\sum\left[In(y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1) - In(1+\theta_1)\right]^2 \qquad (3.31)$$

Multiplied through by $\frac{2\sigma^4}{\mathrm{n}}$ obtained:

$$\hat{\sigma}^2 = \frac{\sum\left[In(y_t - y_{t-1} - \hat{\phi}_1 y_{t-1} + \hat{\phi}_1 y_{t-2} - \hat{\beta}_0 - \hat{\beta}_1 x_1) - In\left(1+\hat{\theta}_1\right)\right]^2}{n} \qquad (3.32)$$

The derivation had unique solutions for the estimation of arimax (1, 0, 1) with lognormal error model parameters as we applied the derived equations (3.21), (3.26), (3.28), (3.30) and (3.32).

Derived equation of arimax (1, 0, 1) with lognormal error fitted as:

$$\hat{Y}_t = \hat{\phi}_1 y_{t-1} + \hat{\theta}_1 \varepsilon_{t-1} + \hat{\beta}_0 + \hat{\beta}_1 x_1 \qquad (3.33)$$

3.6 Model identification techniques

The statistical techniques applied to examine model identification were Autocorrelation Function (acf) and Partial Autocorrelation Function (pacf).

3.6.1 Graphical analysis

The time plot of series exhibited pattern (likely nature) of time series. The time plot suggested the mean changed overtime; that may imply the series was not stationary. That gave initial clue for more formal tests of stationarity.

3.6.2 Sample autocorrelation function (sacf)

Autocorrelation measured the correlation between successive observations in the series under study. Times series X_t; the covariance between X_t and X_{t-k} defined as:

$$\begin{aligned} \gamma_k &= Cov(X_t, X_{t-k}) \\ &= E[(X_t - \mu)(X_{t-k} - \mu)] \\ &where\ \ E(X_t) = \mu \end{aligned} \tag{3.34}$$

The autocorrelation function between X_t and X_{t-k} denoted by $\hat{\rho}_k$ (rho k) defined as:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\text{covariance at lag k}}{\text{variance}} = \frac{Cov(X_t, X_{t-k})}{\sqrt{Var(X_t)}\sqrt{Var(X_{t-k})}} \tag{3.35}$$

Where;

$$\hat{\gamma}_k = \frac{\sum (X_t - \overline{X}_k)(X_{t+k} - \overline{X})}{n} \tag{3.36}$$

$$\hat{\gamma}_0 = \frac{\sum (X_t - \overline{X})^2}{n} \tag{3.37}$$

n denoted the sample size while $\overline{X}$ denoted the sample mean Shittu and Yaya (2016).

3.6.3 Partial autocorrelation function (pacf)

The Autocorrelation function between X_t and X_{t-k} allowed for the effect of the intervened values of X_{t-1}, X_{t-2}. …, X_{t-k}. The Partial Autocorrelation Function (PACF) expressed in matrix as:

$$\rho_k = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \vdots & \vdots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{bmatrix} \tag{3.38}$$

The partial autocorrelation ϕ_{kk} defined as:

$$\phi_{kk} = \frac{\rho_k^*}{\rho_k} \tag{3.39}$$

Where ρ_k^* is the matrix ρ_k with the last column replaced by vector $(\rho_1, \rho_2, ..., \rho_k)^1$. Durbin had presented an iterative way of computing the partial autocorrelation from the relations:

$$\hat{\phi}_{k+1,k+1} = \frac{\gamma_{k+1} - \sum_{i=1}^{k} \phi_{kj} \gamma_{k+1-j}}{1 - \sum_{i=1}^{k} \phi_{kj} \gamma_j} \tag{3.40}$$

In stationary series condition; determination of order of the model used traditional Box-Jenkins approach which applied the combination of acf and pacf functions for possible decline of the curves (cut off) on either of the two curves. A series generated by an ar (p) decays if the (acf) ρ_k decays exponentially to zero and ϕ_{kk} (pacf) cuts off at lag p. The observed cut-off point of ϕ_{kk} determined the appropriate order of the ar model Shittu and Yaya (2016).

Moving average (ma) process of order q was determined by the (acf) ρ_k cut off at a particular lag and the ϕ_{kk} pacf decays exponentially to zero. The point at which ρ_k cuts-off (q) determined the lag of the ma model fit. If neither acf nor the pacf cuts-off, it suggested an arma (p, q) model. The order (p, q) would sometimes difficult to determine by mere inspection of acf and pacf.

Correlogram: the plot of autocorrelation function on non negative lags k, k = 1, 2, 3, …, N/4 gave visual inspection. Correlogram helped in determined whether a series was random, stationary or alternated series and whether it contained a trend or seasonal fluctuations.

3.6.4 Unit root test

Test for stationarity of time series variables were considered to avoid spurious model results. The stationarity consideration of series; when the mean and variance remained constant overtime (do not change overtime or drift). Unit root test was applied as a diagnostic procedure for testing the variables under consideration with respect to stationarity behavior of the variables based on their associated data pattern. Unit root helped in measuring the unchanging behavior of data mean and variance overtime and at level (order) of stationarity Shaib (2019).

The test statistic applied: Dickey Fuller (df) mathematically expressed as:

$$Y_t = \rho Y_{t-1} + \varepsilon_t$$
$$Y_t - Y_{t-1} = \rho Y_{t-1} - Y_{t-1} + \varepsilon_t$$
$$\Delta Y_t = (\rho - 1) y_{t-1} + \varepsilon_t \qquad 3.41$$

$$AR(1) = YY_{t-1} + \varepsilon_t$$

Equation (3.41) was used to test for the statistical significance of series coefficients with respect to stationarity condition.

Table 3.1 Model specification criteria

MODEL	ACF PATTERN	PACF PATTERN
ar(p)	Exponential decay or damped sine wave pattern or both.	Significant spikes through first lag.
ma(q)	Significant spikes through first lag.	Decline Exponentially.
arma(p, q)	Exponential decay.	Exponential decay.

3.7 Model validation/selection criteria

Approaches used as criteria for selecting the best model in terms of the order of a model were: Best Fit Criterion, Final Prediction Error Criteria, Akaike's Information Criterion, Portmanteau Lack-of-fit test.

3.7.1 Best fit criterion

The best model structure minimized the prediction error. Best fit criterion was used for model validation by the highest fit. The best fit was measured by the Coefficient of Determination denoted as R^2, expressed as:

$$R^2 = \left(1 - \frac{\sum_{i=1}^{N} \varepsilon^2}{\sum_{i=1}^{N} (y - \hat{y})^2}\right) x100\% \qquad (3.41)$$

R^2 lied between 0 and 1; the closer R^2 to 1, the better the fit.

3.7.2 Final prediction error criterion (FPE)

The FPE evaluated the model quality, where the model was tested on a new set of data. The most accurate model had the smallest FPE.

The FPE equation defined as:

$$FPE = \frac{1 + n/N}{1 - n/N} * V \qquad (3.42)$$

Variance V, time length N and number of parameters n.

3.7.3 Akaike's information criterion (AIC)

AIC was used for goodness-of-fit test defined as:

$$AIC = \log\left[V\left(1 + \frac{2n}{N}\right)\right] \qquad (3.43)$$

The model with the lowest AIC most preferred.

3.7.4 Portmanteau lack-of-fit test

The test statistic defined as:

$$Q = N \sum_{k=1}^{K} r_k^2 \qquad (3.44)$$

The Q statistic computed from the lowest K autocorrelations, at k = 1, 2, ….., K follows a Chi-square distribution with (K – p – q) degrees of freedom, where p and q were the AR and MA orders of the model and N was the length of the time series Shaib (2019).

3.7.5 Model consideration criteria

The best fitted model considered; met the following criteria:

1. Most significant coefficients
2. Lowest volatility (Variance)
3. Highest Adjusted R^2
4. Lowest AIC
5. Highest log likelihood
6. Smallest absolute means square error

3.8 Residuals diagnostic check

3.8.1 Graphical method test was used to explore the residuals of the fitted models. Beside graphical test other quantitative tests were used to supplement the purely qualitative approach were Jarque-Bera test and Ljung-Box test.

3.8.2 *Jarque-Bera Test*

According to Mohammed (2014) normality assumption can be checked by using Jarque-Bera test which determined goodness-of-fit measured departure from normality, based on the sample kurtosis (k) and skewedness(s). The test statistic Jarque-Bera (JB) defined as:

$$JB = \frac{n}{6}\left(s^2 + \frac{(k-3)^2}{4}\right) \approx \chi^2_{(2)} \tag{3.45}$$

The notations in the equation: n was observations, k parameters. JB assumed $\chi^2_{,2}$ *df.*

3.8.3 *Ljung-Box Test*

Ljung-Box test was used to check autocorrelation among the residuals. The criterion; if a model fit well, the residuals should not be correlated and the correlation should be small. The null hypothesis was stated as:

$H_0 : \rho_1(e) = \rho_2(e) = \ldots\ldots = \rho_k(e) = 0$

$$Q = N(N+1)\sum_{i=1}^{k}(N-k)\rho_k^2(e) \qquad (3.46)$$

Where, N denoted the number of observations used to estimate the model. The statistic Q followed the chi-square distribution with (k-q) degree of freedom, where q was the number of parameters estimated in the model. The criterion: if Q value was large (significantly large from zero), it was considered that the residuals autocorrelation set were significantly different from zero and random shocks of estimated model were probably auto-correlated.

3.9 Model forecast

The model forecast criteria used on the fitted model were:

1. The essence of fitted arimax model was to forecast future values of the series
2. Used past values of the series itself
3. The series spoke for themselves
4. Forecast was based on the final selected model
5. Plotted the forecast graph
6. Verified how successful the forecast had been in predicted future values of the series
7. Conclusion was drawn

3.10 Scope of the Study

The scope of the study was; on mixed time series model applicable in modeling and forecasting time series. Autoregressive Integrated Moving Average with Exogenous variables (arimax) as well as consideration of Multiple Linear Regression model (mlr). The study made use of real-life data obtained from secondary source to verify the applicability of the models and established adequacy (robustness) of the fitted models meeting the expectation of time series analyses.

The time series economic variables under consideration were: Nigeria External Reserves, Exchange Rates, Crude Oil Export and Crude Oil Prices. Real-life monthly data of ten (10) years period were gathered from Central Bank of Nigeria (CBN) Statistical Bulletin (2017) and Nigerian National Petroleum Corporation (NNPC).

3.11 Definition of terms

Y_t – Monthly External Reserves (Million US$ (Dep. Var.))

X_{t1} – Monthly Official Exchange Rate (Naira to 1 US$)

X_{t2} – Monthly Crude Oil Export (Million Barrel per Day mbd)

X_{t3} – Monthly Crude Oil Price (US$/Barrel)

TSD – Time Series Data

ARMA – Autoregressive Moving Average

ARIMA – Autoregressive Integrated Moving Average

ARIMAX – Autoregressive Integrated Moving Average with Exogenous Variable

MLR – Multiple Linear Regression model under normality assumptions

SARIMA – Seasonal Autoregressive Integrated Moving Average

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1 Results and discussion

Chapter four presents the results of data analyses and corresponding discussions. Data exploratory analyses results were presented in section 4.2, section 4.3 contained descriptive analyses results, model identification results/plots shown in section 4.4, the diagnostic test results in section 4.5. Section 4.6 presented the fitted model equations, section 4.7 contained model selection results, models' in-sample forecast plots were presented in section 4.8. The residual plots were shown in section 4.9 while section 4.10 presented out-sample forecast plots.

4.2 Exploratory analysis of data

The graphical exploratory analyses results of the time series data under consideration were shown in the figures (fig. 4.1 to fig. 4.8) followed by corresponding discussions. The time plots showed the behavoural pattern (trend) of the economic time series data during the period under consideration. The corresponding histogram plot described the series distribution shape; to ascertain the normality status of the time series data of interest. The normal curve property of having bell-shaped distribution inferred the status decision of the series at raw level.

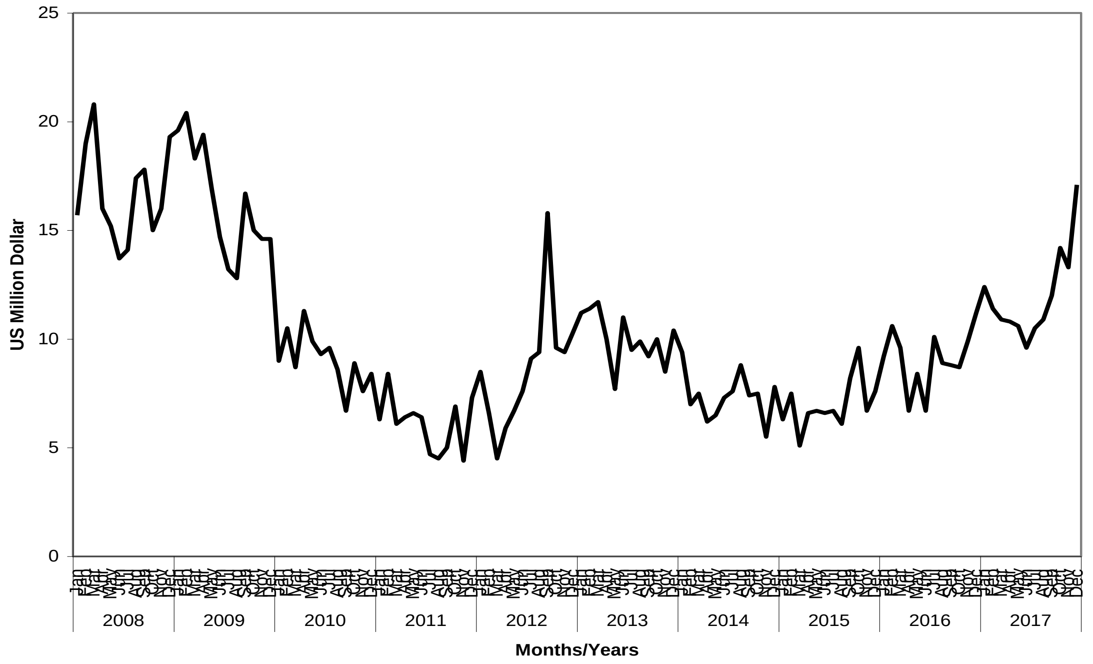


Fig. 4.1. Time plot of external reserve (Jan. 2008 - Dec. 2017 (US million dollar))

Fig. 4.1 showed the trend pattern of Monthly External Reserves (US $ Million) time series data set between the periods of Jan. 2008 to Dec. 2017. The series tends to fluctuate; there was an observed rapid decline between Jan, 2010 through Aug, 2012 Sept, 2012 experienced sharp increase and a drop in Jan, 2013. The trend pattern exhibited stochastic process and non-stationarity of the series. The behavior of the series might be traced to the fall of crude oil price at the global market and increase in foreign exchange rate during that period.

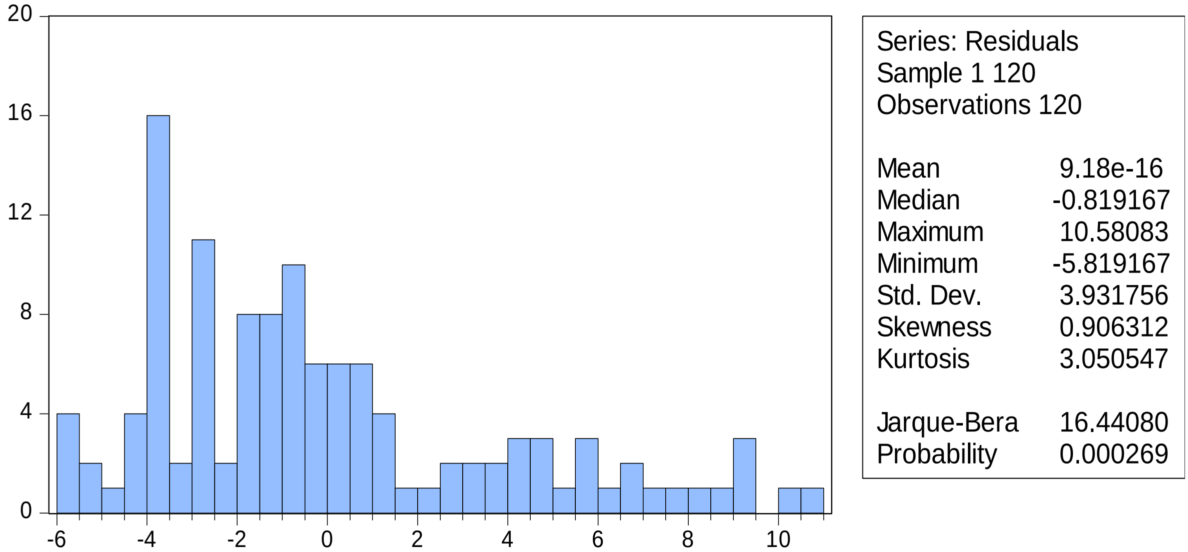


Fig. 4.2. External reserve histogram plot and statistic

The histogram (Fig. 4.2) plot showed deviation from normal curve, the curve exhibited long tail and gradual decline to the right (right skewed distribution). That indicated the decline of external reserve over time period. The skewness value was 0.91 and kurtosis value was about 3.10 those values showed that the series was not normally distributed. The Jarque-Bera (JB) value was 16.44 with P-value of $0.2x10^{-4}$ which revealed that the series was non-stationary.

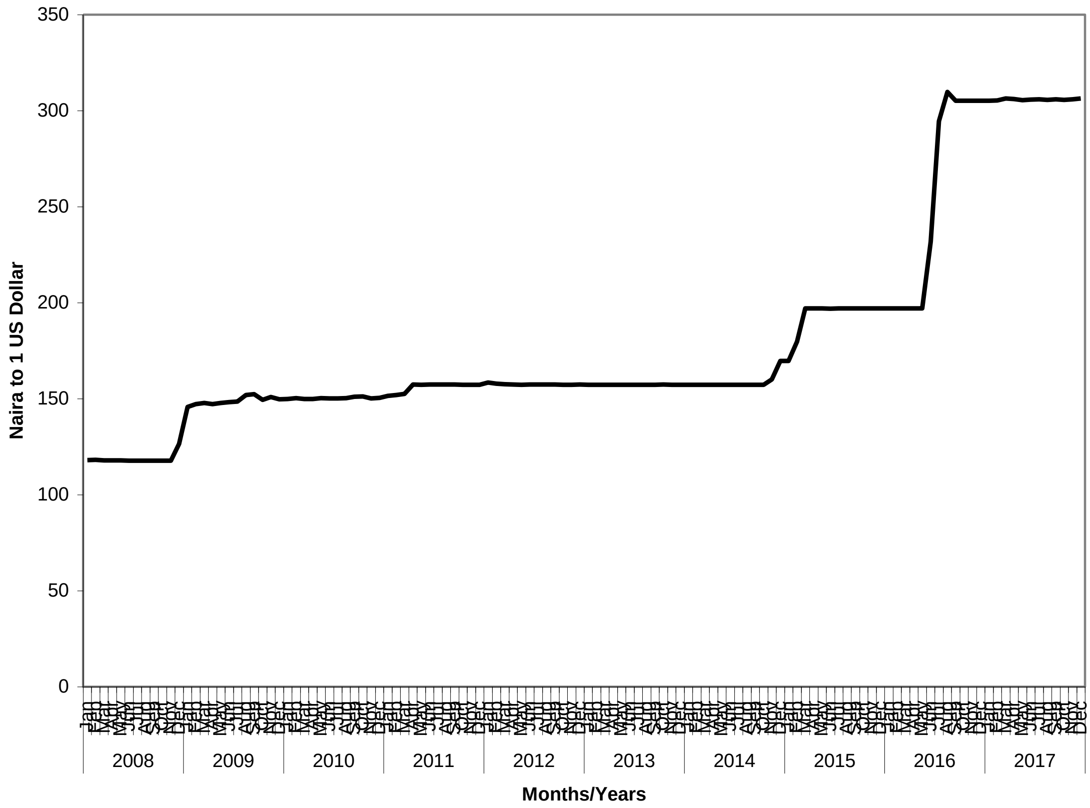


Fig. 4.3. Time plot of exchange rate (Jan. 2008 - Dec. 2017 (Naira to US $1))

The time plot of Exchange Rate in Naira per 1 US$ as shown in Fig. 4.3 indicated steady exchange rate of N120/US$ between Jan, 2008 to Dec, 2008; then, an increase of about N150/US$ in Jan, 2009 and gradual increase through Dec, 2014. In Jan, 2015 an increase of N200/US$ was observed, in the month of Jun, 2016 a sharp increase of above N300/US$ was observed. That could be inflation and high cost of imported goods and services resulted to reduction of external reserve under study. The series exhibited a stochastic process, and non-stationarity. That resulted to high cost of importation of goods and services which affected external reserves negatively.

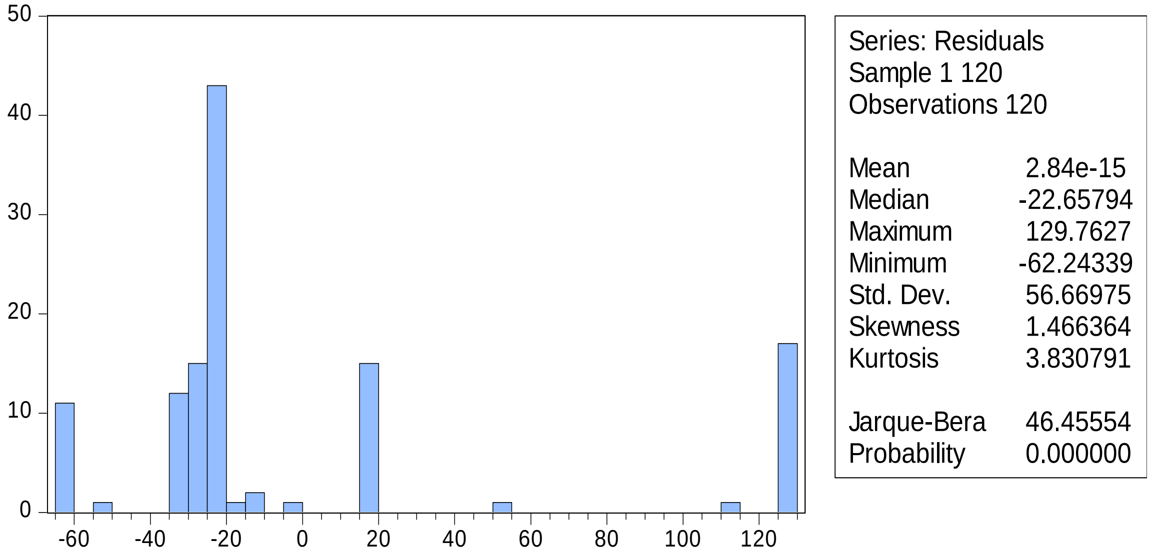


Fig. 4.4. Exchange rate histogram plot and statistic

The histogram (right skewed distribution) chart (Fig. 4.4) showed some periods of no significant changes in exchange rate and a sharp deviation from normal curve. The skewness value was 1.47 and kurtosis value was 3.38, the Jarque-Bera value was 46.46. The exchange rate time series data deviated from normal distribution, and non-stationary time series.

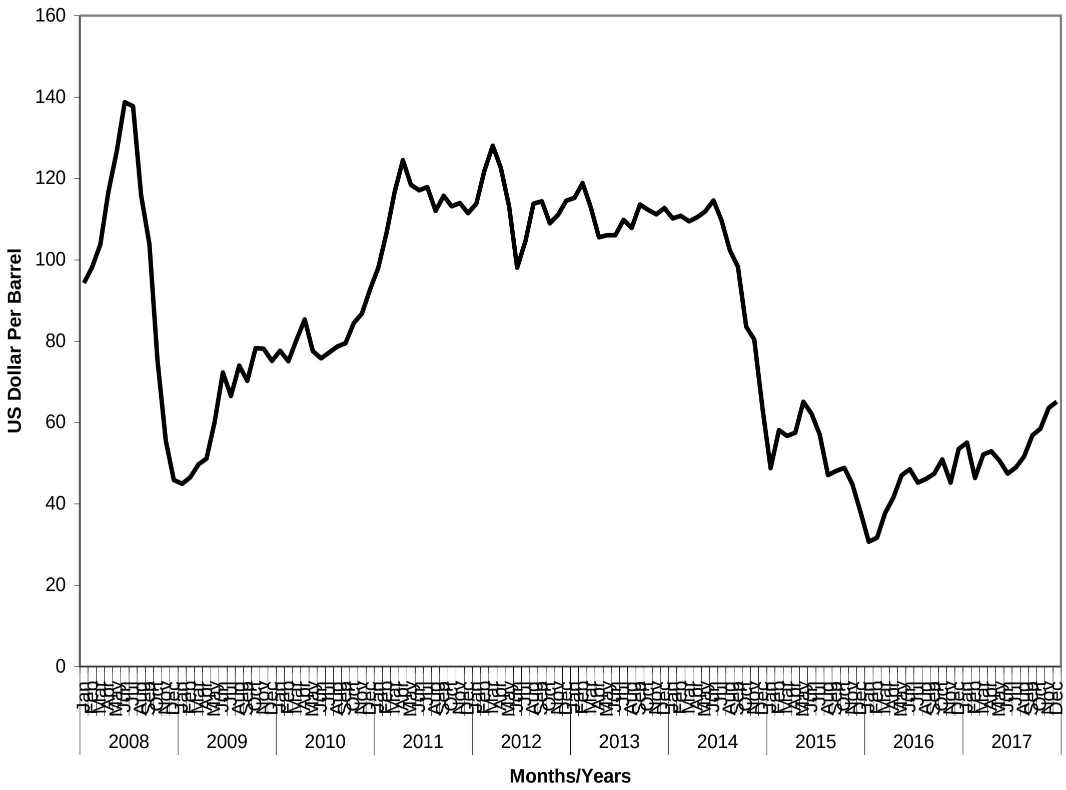


Fig. 4.5. Time plot of crude oil price in US dollar/barrel (2008 – 2017)

Fig. 4.5 showed the time plot of monthly crude oil price (US$/Barrel) exhibited fluctuation over the period under study. The series indicated highest crude oil price in June, 2008. The price of crude oil dropped to $44.95/barrel in January 2009, there was indication of fluctuation in crude oil price $110 – $130 per barrel between Jan 2011 to Dec 2014. In January 2015 experienced another sharp dropped of crude oil price of $40 per barrel. The minimum crude oil price was $31.70 per barrel in February 2016. The state of crude oil price affected external reserve directly being the main sources of Nigerian foreign income.

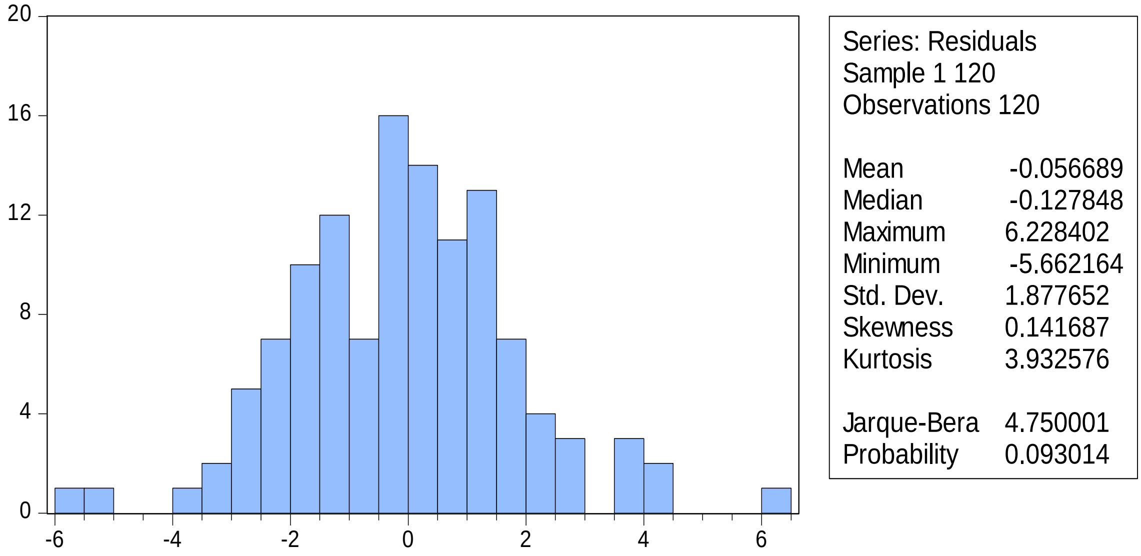


Fig. 4.6. Crude oil price histogram plot and statistic

The histogram (Fig. 4.6) exhibited approximately normal distribution; plot showed approximately normal curve. The skewness value was 0.14 and kurtosis value was about 3.93. The Jarque-Bera (JB) value was 4.75 with P-value of 0.9×10^{-1} which revealed that the series was non-stationary. Crude oil price experienced sharp drop and increased price during the period under study. That was typical shorter period of business cycle.

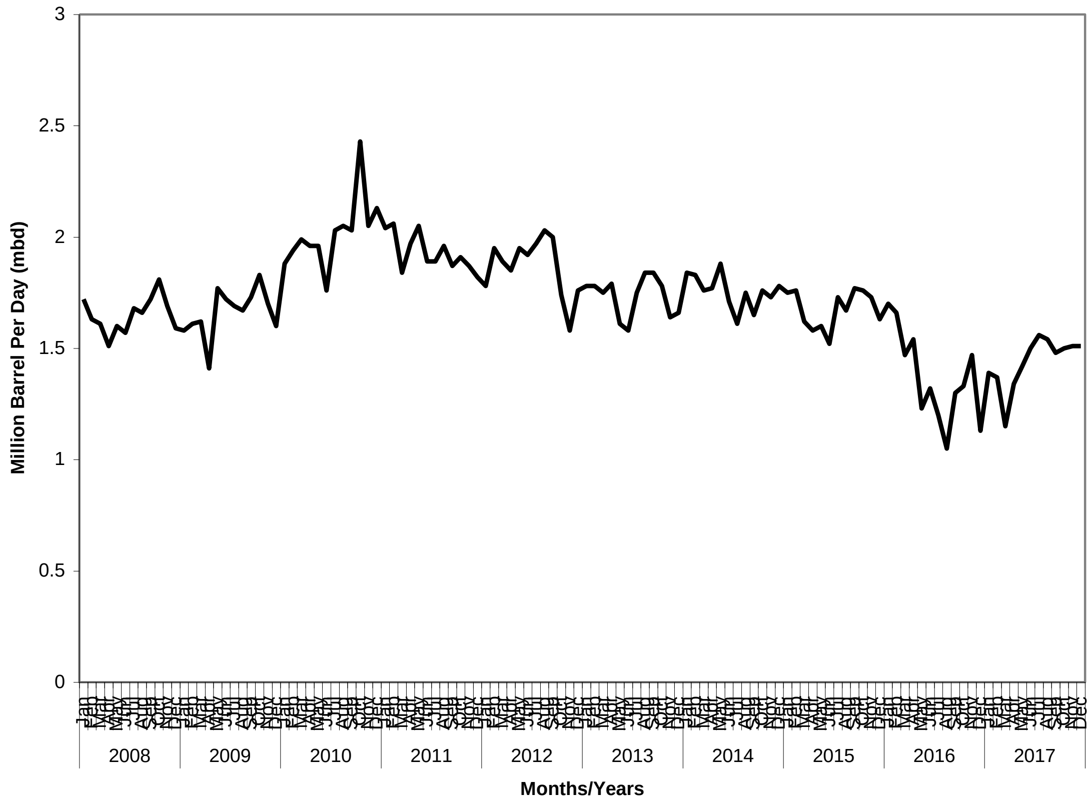


Fig. 4.7. Time plot of crude oil export in million barrels per day (2008 – 2017)

Fig. 4.7 showed the time plot of monthly crude oil export: the series trend fluctuated over the period under study. The crude oil export market operated within the boundary (limit) of 1.05mbd and 2.43mbd (million barrel per day). The highest crude oil export of 2.43mbd was observed in Nov. 2010 while the lowest crude oil export of 1.05mbd was noticed in Jun. 2016, the range value was 1.35mbd.

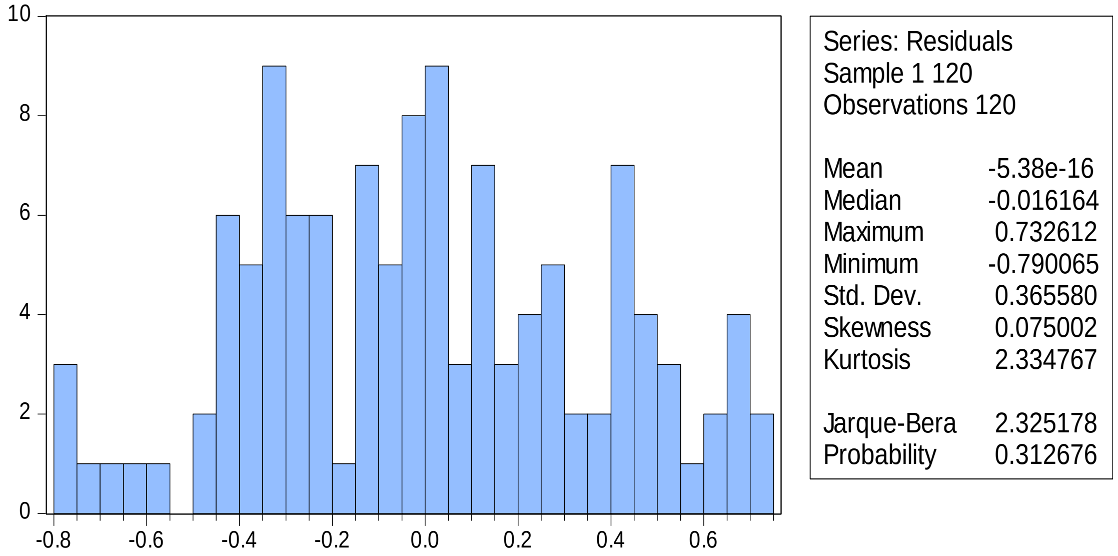


Fig. 4.8. Crude oil export histogram plot and statistic

The histogram monthly crude oil export (Fig. 4.8) plot exhibited a bimodal distribution. There was relatively regular exportation operation of crude oil market but affected by crude oil price market forces. The plot showed deviation from normal curve. The skewness value was 0.08 and kurtosis value was about 2.33. The Jarque-Bera (JB) value was 2.33 with P-value of 0.31 which revealed that the series was non-stationary at raw level. That confirmed the nature of time series data. The crude oil export had not been stable over time; which might lead to negative impact on the economic activities of the nation Nigeria among other factors.

4.3: Descriptive analyses results

Section 4.3 presented descriptive analyses results; table 4.1 showed the summary statistic of the four economic time series variables understudy.

Notations:

Y_t – External reserve (in million \$);

X_1 – Exchange rate (in naira to \$1);

X_2 – Crude oil export (in million barrel per day) and

X_3 – Crude oil price (\$ per barrel).

Table 4.1. Summary statistic of time series variables under study

	Y_t	X_1	X_2	X_3
Mean	**10.21917**	**179.9677**	**1.709833**	**83.52367**
Median	**9.400000**	**157.3098**	**1.730000**	**80.34500**
Maximum	**20.80000**	**309.7304**	**2.430000**	**138.7400**
Minimum	**4.400000**	**117.7243**	**1.050000**	**30.66000**
Std. Dev.	**3.931756**	**56.66975**	**0.224885**	**29.17091**
Skewness	**0.906312**	**1.466364**	**-0.295559**	**-0.061853**
Kurtosis	**3.050547**	**3.830791**	**3.744180**	**1.565548**
Jarque-Bera	**16.44080**	**46.45554**	**4.516120**	**10.36478**
Probability	**0-000269**	**0.000000**	**0.104553**	**0.005615**

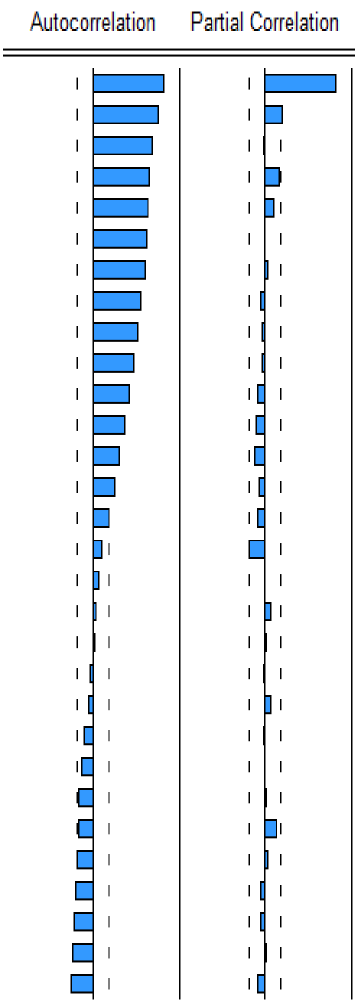
Table 4.1 showed some indices/test results of the economic variables under study; the series – external reserve, exchange rate and crude oil price follows normal distribution as indicated by JB-Jarque-Bera statistic ($p < 5\%$ critical value) while crude oil export series distribution deviated from normal as $p > 5\%$ critical value.

There is indication of slight skewness; the series external reserve and exchange rate were positively skewed while crude oil export and crude oil price were negatively skewed. The average external reserves within the period of Jan. 2008 – Dec. 2017 was about $10.23 million while that of exchange rate was about N180 per $1. The crude oil export was in average of 2 million barrel per day (mbd) and the average crude oil price was about $84 per barrel.

The maximum external reserve observed during the period was about $21 million in March 2008 while the minimum was about $4.4 million observed in Dec. 2011 and Mar 2012. The maximum exchange rate was in Jan 2017 of about N310 per $1 while the minimum exchange rate was about N118 per $1. The maximum crude oil export was about 2.4 mbd while minimum observed crude oil export was about 1.1 mbd. The crude oil maximum price was about $139 million per barrel while the minimum price was about $30.67 million per barrel.

4.4 Arimax with normal error model identification

Section 4.4 presented the results and discussion of time series model order identification. The plotted correlogram, autocorrelation function (acf), partial autocorrelation function (pacf) and Q-statistic (Q) at raw level I(0) and at first difference I(1) with respect to the economic time series variables of interest. The plots were shown from fig. 4.9 to fig. 4.16 with their respective discussions.



	AC	PAC	Q-Stat	Prob
1	0.860	0.860	90.899	0.000
2	0.793	0.209	169.00	0.000
3	0.713	-0.024	232.64	0.000
4	0.688	0.165	292.36	0.000
5	0.666	0.101	348.78	0.000
6	0.639	0.007	401.30	0.000
7	0.616	0.040	450.47	0.000
8	0.575	-0.041	493.64	0.000
9	0.535	-0.040	531.36	0.000
10	0.492	-0.032	563.54	0.000
11	0.442	-0.083	589.73	0.000
12	0.385	-0.100	609.77	0.000
13	0.316	-0.124	623.47	0.000
14	0.257	-0.079	632.58	0.000
15	0.194	-0.096	637.81	0.000
16	0.113	-0.185	639.61	0.000
17	0.066	-0.002	640.22	0.000
18	0.038	0.074	640.42	0.000
19	0.014	0.013	640.45	0.000
20	-0.026	-0.018	640.55	0.000
21	-0.057	0.064	641.02	0.000
22	-0.102	-0.012	642.57	0.000
23	-0.140	0.006	645.51	0.000
24	-0.174	0.017	650.14	0.000
25	-0.172	0.135	654.72	0.000
26	-0.182	0.030	659.91	0.000
27	-0.203	-0.047	666.39	0.000
28	-0.233	-0.053	675.01	0.000
29	-0.241	0.017	684.36	0.000
30	-0.262	-0.096	695.51	0.000

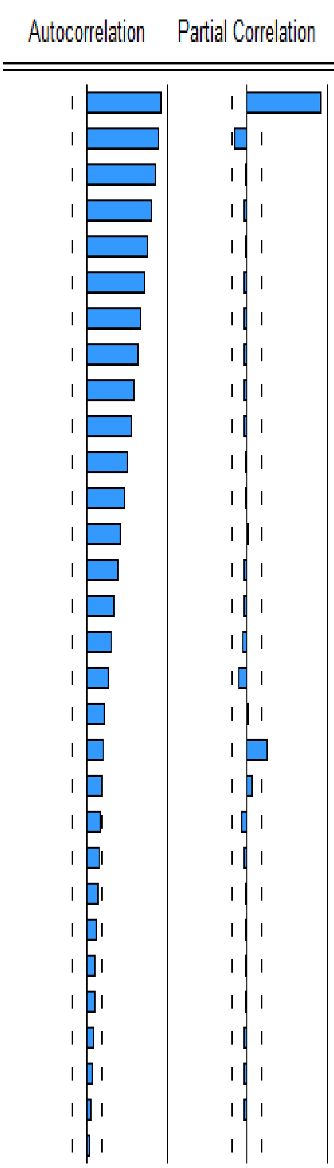
Fig. 4.9: Acf, pacf plots and Q-statistic of external reserve at raw revel I(0)

Fig. 4.9 showed the correlogram of external reserve series at raw level I(0): Autocorrelation Function (acf), Partial Autocorrelation Function (pacf). The plot exhibited exponential decay from lag 1. The acf at lag 1 has value 0.860 to -0.262 at lag 30 while pacf at lag 1 has value 0.860 to -0.096 at lag 30. This indicated non stationary series. Hence, the series required transformation to achieve stationarity by differencing (data stationary).

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.293	-0.293	10.448	0.001
		2	0.023	-0.069	10.513	0.005
		3	-0.136	-0.164	12.814	0.005
		4	0.014	-0.085	12.840	0.012
		5	0.043	0.011	13.069	0.023
		6	-0.003	-0.008	13.070	0.042
		7	0.018	0.013	13.112	0.069
		8	-0.013	0.008	13.135	0.107
		9	0.046	0.055	13.410	0.145
		10	0.010	0.051	13.423	0.201
		11	-0.028	-0.003	13.529	0.260
		12	0.062	0.077	14.051	0.297
		13	-0.026	0.028	14.144	0.364
		14	0.061	0.069	14.654	0.402
		15	0.058	0.131	15.116	0.443
		16	-0.108	-0.047	16.760	0.401
		17	-0.073	-0.119	17.514	0.420
		18	0.055	0.007	17.946	0.459
		19	0.042	0.019	18.199	0.509
		20	-0.055	-0.086	18.635	0.546
		21	0.034	-0.005	18.804	0.598
		22	-0.041	-0.032	19.057	0.642
		23	0.010	-0.041	19.071	0.697
		24	-0.058	-0.098	19.589	0.720
		25	0.045	-0.001	19.905	0.752
		26	0.021	0.034	19.970	0.793
		27	0.029	0.029	20.105	0.826
		28	-0.041	-0.004	20.366	0.851
		29	0.051	0.087	20.780	0.867
		30	-0.185	-0.160	26.305	0.659

Fig. 4.10. Acf, pacf plots and Q-statistic of external reserve at first difference I(1)

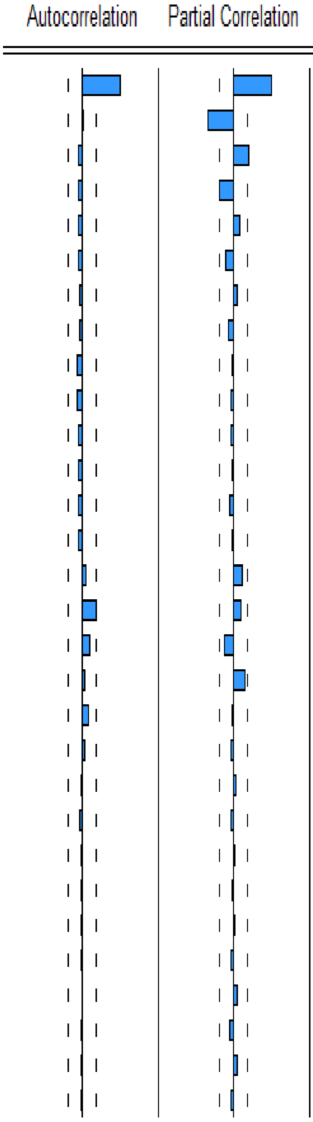
Fig 4.10 showed that external reserve series achieved stationarity at first difference I(1). The correlogram at first difference I(1) level, the plot indicated significant spike at lag 1 of acf and lag 1 of pacf which suggested arima(1, 1, 1) model order.



Lag	AC	PAC	Q-Stat	Prob
1	0.966	0.966	114.73	0.000
2	0.923	-0.149	220.32	0.000
3	0.878	-0.019	316.87	0.000
4	0.834	-0.025	404.69	0.000
5	0.790	-0.024	484.10	0.000
6	0.745	-0.026	555.45	0.000
7	0.701	-0.025	619.10	0.000
8	0.656	-0.029	675.41	0.000
9	0.611	-0.033	724.69	0.000
10	0.566	-0.026	767.35	0.000
11	0.522	-0.022	803.90	0.000
12	0.478	-0.009	834.93	0.000
13	0.438	0.013	861.24	0.000
14	0.399	-0.036	883.20	0.000
15	0.359	-0.027	901.17	0.000
16	0.318	-0.054	915.38	0.000
17	0.271	-0.105	925.84	0.000
18	0.227	0.022	933.26	0.000
19	0.203	0.259	939.24	0.000
20	0.189	0.071	944.49	0.000
21	0.175	-0.070	949.01	0.000
22	0.160	-0.031	952.81	0.000
23	0.145	-0.015	955.96	0.000
24	0.130	-0.022	958.52	0.000
25	0.114	-0.015	960.54	0.000
26	0.099	-0.018	962.07	0.000
27	0.084	-0.028	963.18	0.000
28	0.068	-0.033	963.93	0.000
29	0.053	-0.025	964.38	0.000
30	0.037	-0.007	964.61	0.000

Fig. 4.11. Acf, pacf plots and Q-statistic of exchange rate at raw level I(0)

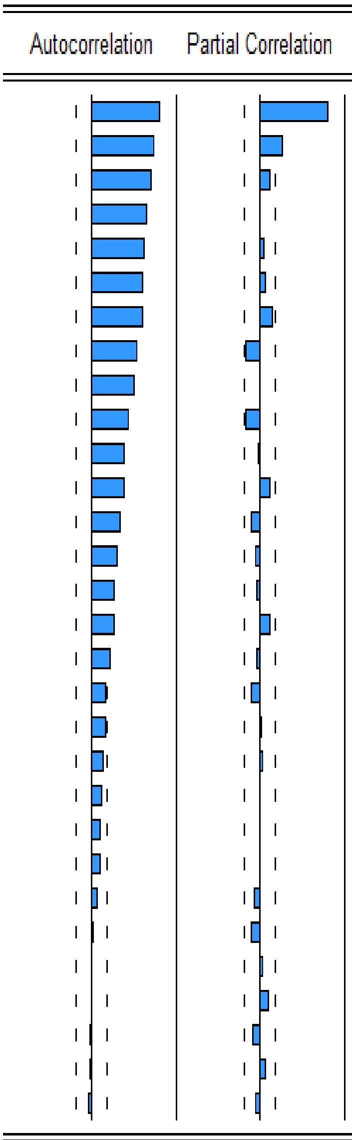
Fig. 4.11 showed the correlogram (acf and pacf) of exchange rate (1US$ per Naira), the plot exhibited exponential decay from lag 1 with acf value of 0.966 to 0.037 at lag 30 while the pacf from lag 1 value 0.966 to -0.007, that indicated non-stationary time series. The series required differencing to achieve stationary condition.



	AC	PAC	Q-Stat	Prob
1	0.517	0.517	32.592	0.000
2	0.009	-0.352	32.603	0.000
3	-0.043	0.206	32.830	0.000
4	-0.043	-0.181	33.059	0.000
5	-0.049	0.086	33.361	0.000
6	-0.048	-0.101	33.659	0.000
7	-0.037	0.047	33.836	0.000
8	-0.035	-0.072	33.994	0.000
9	-0.059	-0.019	34.456	0.000
10	-0.060	-0.027	34.936	0.000
11	-0.051	-0.034	35.289	0.000
12	-0.046	-0.015	35.571	0.000
13	-0.046	-0.043	35.855	0.001
14	-0.047	-0.017	36.156	0.001
15	0.048	0.123	36.480	0.002
16	0.189	0.113	41.500	0.000
17	0.109	-0.124	43.181	0.000
18	0.038	0.162	43.389	0.001
19	0.089	-0.017	44.539	0.001
20	0.037	-0.026	44.737	0.001
21	-0.019	0.027	44.791	0.002
22	-0.025	-0.037	44.887	0.003
23	-0.021	0.023	44.955	0.004
24	-0.015	-0.021	44.990	0.006
25	-0.019	0.024	45.042	0.008
26	-0.021	-0.029	45.112	0.011
27	-0.007	0.049	45.119	0.016
28	-0.016	-0.052	45.159	0.021
29	-0.024	0.042	45.252	0.028
30	-0.024	-0.032	45.342	0.036

Fig. 4.12. Acf, pacf plots and Q-statistic of exchange rate at first difference I(1)

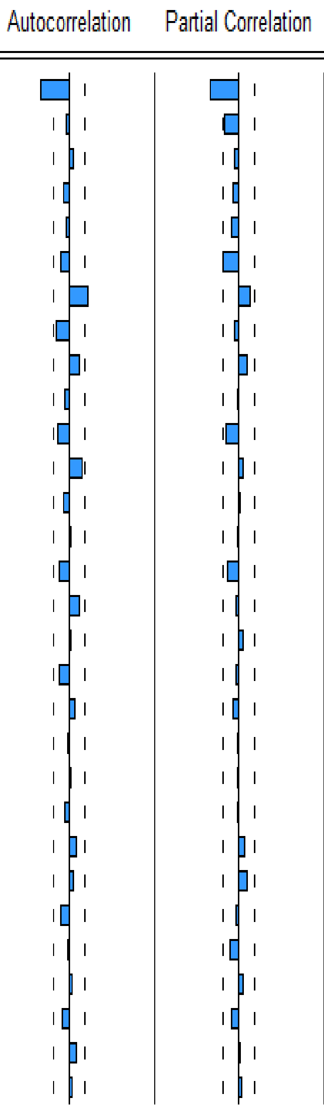
Fig. 4.12 showed the exchange rate first difference plot: correlogram of exchange rate series achieved stationarity at first difference I(1) level. There was indication of statistical significance spike at lag 1 of acf and pacf as well as lag 2 of pacf suggesting two model order of arima(1, 1, 1) and arima(2, 1, 1).



	AC	PAC	Q-Stat	Prob
1	0.831	0.831	84.899	0.000
2	0.777	0.280	159.79	0.000
3	0.735	0.131	227.35	0.000
4	0.678	0.001	285.43	0.000
5	0.644	0.047	338.25	0.000
6	0.619	0.060	387.47	0.000
7	0.626	0.156	438.26	0.000
8	0.553	-0.172	478.22	0.000
9	0.526	-0.002	514.72	0.000
10	0.451	-0.175	541.84	0.000
11	0.402	-0.018	563.58	0.000
12	0.401	0.129	585.41	0.000
13	0.341	-0.108	601.31	0.000
14	0.313	-0.046	614.82	0.000
15	0.273	-0.026	625.23	0.000
16	0.274	0.120	635.82	0.000
17	0.225	-0.026	642.99	0.000
18	0.177	-0.102	647.46	0.000
19	0.172	0.016	651.75	0.000
20	0.141	0.035	654.65	0.000
21	0.124	0.000	656.93	0.000
22	0.099	0.007	658.39	0.000
23	0.099	0.004	659.86	0.000
24	0.068	-0.065	660.56	0.000
25	0.018	-0.113	660.61	0.000
26	0.006	0.029	660.62	0.000
27	-0.002	0.103	660.62	0.000
28	-0.019	-0.089	660.68	0.000
29	-0.011	0.074	660.70	0.000
30	-0.029	-0.054	660.83	0.000

Fig. 4.13. Acf, pacf plots and Q-statistic of crude oil export at raw level I(0)

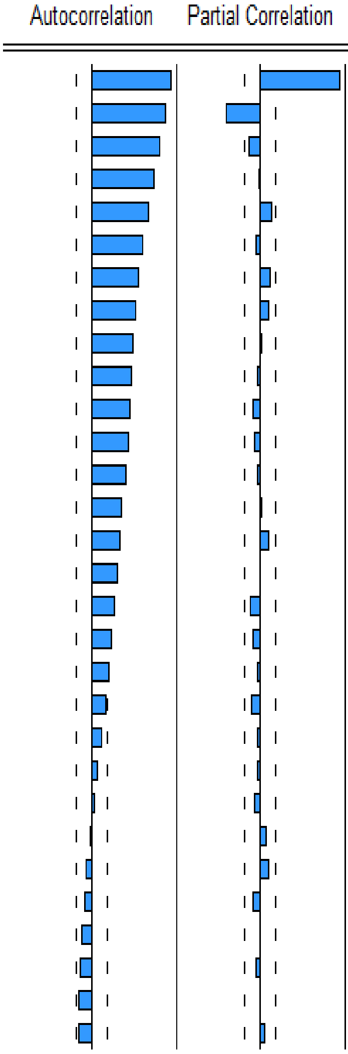
The Autocorrelation Function (acf) of crude oil export (in mbd) shown in Fig 4.13 exhibited exponential decay slowly in lag 1 a value 0.831 to lag 30 value -0.029 revealed non-stationarity series; the pacf had significant spikes at lag 1 and lag 2. The acf and pacf exhibited exponential decay (rapid decline) which suggested an arima (p, q) model. Differencing was suggested to achieve stationarity of the series.



	AC	PAC	Q-Stat	Prob
1	-0.348	-0.348	14.757	0.000
2	-0.036	-0.178	14.916	0.001
3	0.041	-0.042	15.127	0.002
4	-0.061	-0.077	15.595	0.004
5	-0.026	-0.086	15.679	0.008
6	-0.103	-0.188	17.037	0.009
7	0.234	0.136	24.050	0.001
8	-0.148	-0.045	26.899	0.001
9	0.124	0.110	28.911	0.001
10	-0.055	-0.016	29.310	0.001
11	-0.142	-0.160	32.002	0.001
12	0.152	0.045	35.103	0.000
13	-0.063	0.024	35.638	0.001
14	0.020	-0.014	35.692	0.001
15	-0.125	-0.147	37.853	0.001
16	0.126	-0.033	40.084	0.001
17	0.009	0.052	40.095	0.001
18	-0.119	-0.040	42.110	0.001
19	0.072	-0.061	42.853	0.001
20	-0.015	-0.012	42.886	0.002
21	0.020	-0.018	42.944	0.003
22	-0.057	-0.015	43.420	0.004
23	0.096	0.075	44.807	0.004
24	0.049	0.110	45.175	0.006
25	-0.103	-0.035	46.807	0.005
26	-0.011	-0.113	46.828	0.007
27	0.029	0.048	46.961	0.010
28	-0.088	-0.091	48.180	0.010
29	0.087	0.024	49.394	0.010
30	0.032	0.026	49.560	0.014

Fig. 4.14. Acf, pacf plots and Q-statistic of crude oil export at first difference I(1)

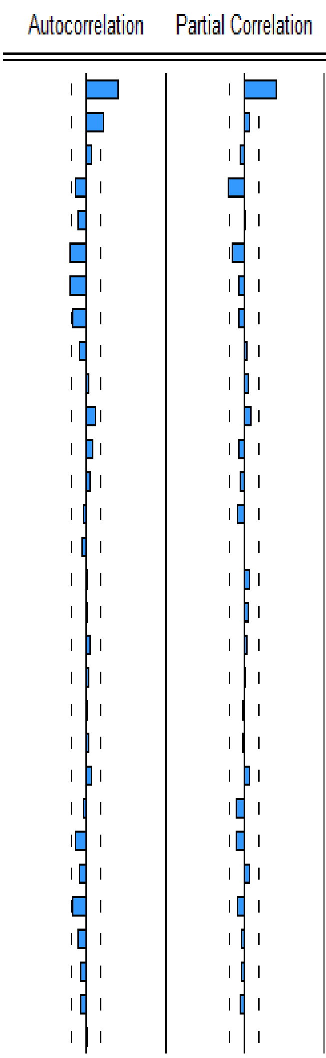
Fig. 4.14 showed the first difference I(1) correlogram of crude oil export. The series achieved stationarity at I(1) level, the acf and pacf (cut-off) were outside standard error bound (or outside 95% confidence interval) had significant spikes at lag 1 which suggested model of order 1 arima(1, 1, 1).



	AC	PAC	Q-Stat	Prob
1	0.969	0.969	115.55	0.000
2	0.914	-0.421	219.10	0.000
3	0.843	-0.133	308.12	0.000
4	0.767	-0.017	382.42	0.000
5	0.697	0.134	444.26	0.000
6	0.631	-0.052	495.40	0.000
7	0.577	0.117	538.50	0.000
8	0.537	0.100	576.21	0.000
9	0.509	0.020	610.36	0.000
10	0.489	-0.033	642.14	0.000
11	0.468	-0.098	671.56	0.000
12	0.443	-0.064	698.15	0.000
13	0.412	-0.035	721.38	0.000
14	0.377	0.015	740.99	0.000
15	0.344	0.106	757.46	0.000
16	0.313	0.001	771.21	0.000
17	0.279	-0.119	782.31	0.000
18	0.244	-0.096	790.83	0.000
19	0.206	-0.027	797.01	0.000
20	0.166	-0.105	801.04	0.000
21	0.124	-0.031	803.30	0.000
22	0.078	-0.040	804.20	0.000
23	0.026	-0.077	804.30	0.000
24	-0.023	0.058	804.38	0.000
25	-0.063	0.104	805.00	0.000
26	-0.098	-0.093	806.49	0.000
27	-0.122	0.007	808.83	0.000
28	-0.141	-0.047	811.98	0.000
29	-0.155	0.001	815.85	0.000
30	-0.163	0.055	820.19	0.000

Fig. 4.15. Acf, pacf plots and Q-statistic of crude oil price at raw level I(0)

Fig. 4.15 showed crude oil price series of acf and pacf at level I(0) which exhibited non-stationarity as indicated by the correlogram which decay exponentially from lag 1 with a value of 0.969 to -0.163 at lag 30 while the pacf value 0.969 at lag 1 to 0.055 at lag 30; hence, the need for differencing.



	AC	PAC	Q-Stat	Prob
1	0.418	0.418	21.278	0.000
2	0.227	0.064	27.634	0.000
3	0.074	-0.051	28.310	0.000
4	-0.142	-0.205	30.817	0.000
5	-0.110	0.020	32.352	0.000
6	-0.212	-0.148	38.057	0.000
7	-0.207	-0.069	43.567	0.000
8	-0.173	-0.070	47.437	0.000
9	-0.092	0.037	48.533	0.000
10	0.026	0.044	48.624	0.000
11	0.124	0.087	50.669	0.000
12	0.096	-0.058	51.916	0.000
13	0.049	-0.054	52.243	0.000
14	-0.027	-0.088	52.341	0.000
15	-0.047	0.001	52.649	0.000
16	0.013	0.072	52.673	0.000
17	0.018	0.046	52.719	0.000
18	0.046	0.041	53.017	0.000
19	0.039	0.011	53.233	0.000
20	0.021	-0.014	53.297	0.000
21	0.032	-0.011	53.450	0.000
22	0.075	0.065	54.276	0.000
23	-0.037	-0.110	54.485	0.000
24	-0.138	-0.111	57.369	0.000
25	-0.085	0.074	58.474	0.000
26	-0.170	-0.088	62.930	0.000
27	-0.110	-0.034	64.820	0.000
28	-0.067	-0.031	65.523	0.000
29	-0.078	-0.055	66.487	0.000
30	0.009	-0.002	66.499	0.000

Fig. 4.16. Acf, pacf plots and Q-statistic of crude oil price at first difference I(1)

Fig. 4.16 showed the first difference of the crude oil price series: correlogram, the acf and pacf indicated significant spike at lag 1. The acf and pacf at lag 1 has value of 0.418, the model of arima (1,1,1) was suggested for crude oil price series.

4.5 Series Stationarity test

Test of stationary of the economic time series variables of interest were considered to avoid spurious model results. The time series data were considered stationary when the mean and variance remain constant overtime. Unit root diagnostic procedure was applied; Dickey Fuller test results of the variables were presented in table 4.2.

Table 4.2. Dickey fuller (df) unit root test

Critical Levels	External Reserve (Y_t)	Exchange Rate (X_1)	Crude Oil Export (X_2)	Crude Oil Price (X_3)
1%	2.5836	2.5836	2.5836	2.5836
5%	**1.9428**	**1.9428**	**1.9428**	**1.9428**
10%	1.6172	1.6172	1.6192	1.6172
Absolute DF Statistic				
Level I(0)	0.6149 (0.5399**)	1.7430 (0.0841**)	0.2978 (0.7664**)	1.2492 (0.2143**)
First Diff I(1)	5.6363 (0.0000*)	4.0083 (0.0000*)	6.1591 (0.0000*)	5.1259 (0.0000*)

P-values in parenthesis ( )

* Statistically significant at 5% critical level

** Not statistically significant at 5% critical level

The table 4.2 showed the unit root test results of dickey fuller (df) to collaborate stationarity check of the series of acf and pacf plots. The four economic time series indicated presence of unit root at raw level. The absolute values of dickey fuller t-statistic value of the series were less than 5% (0.05) critical level. The stationarity of the series were achieved at first difference I(1). The results of df values of I(1) the p-value presented in parenthesis were statistically significant at 5%.

The dickey fuller statistic values of the economic time series variables: external reserve, exchange rate, crude oil export and crude oil price at raw level were; 0.6149, 1.7430, 0.2978 and 1.2492 respectively were less than 1.9428 critical value at 5%. That revealed non stationarity of the series at raw level I(0). At first difference of the series, dickey fuller values were; 5.6363, 4.0083, 6.1591 and 5.1259 respectively greater than 1.9428 critical value at 5%. That implied the series had achieved stationarity at first difference I(1).

Table 4.3. Parameters' coefficients estimated

Model	Parameters	Coefficients
Multiple linear regression	Constant	35.5787
	X_1	-0.0401
	X_2	-9.0747
	X_3	-0.0315
Arimax(1, 1, 1) assuming normal error of three exogenous variables	ar(1)	-0.3127
	ma(1)	-0.9941
	X_1	0.0238
	X_2	0.2725
	X_3	-0.0296
Arimax(1, 0, 1) assuming lognormal error of three exogenous variables	ar(1)	0.4252
	ma(1)	2.4068
	X_1	0.0362
	X_2	6.0537
	X_3	0.0189

Table 4.3 showed the estimated parameter values which were the coefficient values of the variables; X_1 – exchange rate (naira to 1 US$), X_2 – crude oil export (mbd), X_3 – crude oil price (US$/barrel) under consideration. Fitted multiple linear regression parameters coefficients: 35.5787 (intercept), -0.0401, -9.0747 and -0.0315 respectively; that of arimax(1, 1, 1) assuming normal error: -0.3127 (ar(1)), -0.9941 (ma(1)), 0.0238, 0.2725 and -0.0296 respectively. That of arimax(1, 0, 1) assuming lognormal error: 0.4252 (ar(1)), 2.4068 (ma(1)), 0.0362, 0.0537 and 0.0189 respectively.

4.6 Fitted Equations

The fitted equations models:

The Dependent Variable: External Reserve (US$ in Million) - Y_t

Three Exogenous Variables: X_1 – Exchange Rate (Naira to 1 US$), X_2 – Crude Oil Export (mbd) and X_3 – Crude Oil Price (US$/Barrel)

4.6.1: Fitted Multiple Regression Equation

$$\hat{Y}_t = 35.5787 - 0.0401X_1 - 9.0747X_2 - 0.03127X_3 \quad (4.1)$$

4.6.2: Fitted arimax (1,1,1) with Normal Error Equation

$$\hat{Y}_t = -0.3127y_{t-1} - 0.9941\varepsilon_{t-1} + 0.0238X_1 + 0.2725X_2 - 0.0296X_3 \quad (4.2)$$

4.6.3 Fitted arimax (1,0,1) with Lognormal Error Equation

$$\hat{Y}_t = 0.4252y_{t-1} + 2.4068\varepsilon_{t-1} + 0.0362X_1 + 6.0537X_2 + 0.0189X_3 \quad (4.3)$$

4.7 Model selection

The three fitted models 4.1, 4.2 and 4.3 were compared to ascertain best (most suitable for forecast) based on selection criteria: smallest aic, largest loglikelihood (logic) and smallest mean square forecast error (msfe). Table 4.4 present the results.

Table 4.4. Comparison fitted models (diagnostic check analyses results)

Model	Log Likelihood (Loglik)	AIC	MSFE
Multiple Linear Regression	-317.41	5.36	12.409
Arimax (1, 1, 1) with Normal Error at first difference	-240.23	490.45	12.484
Arimax (1, 0, 1) with Lognormal Error at raw series (non-stationary series)	1344.47**	-0.41*	1.766***

* Lowest akiake information criteria (aic) ** Highest loglikelihood (Loglik)

*** Smallest mean square forecast error (msfe)

Based on results table 4.4 and 4.5, the study considered arimax(1, 0, 1) assuming lognormal error was considered more appropriate and best fitted model based on the criteria of highest Loglikelihood (Loglik) value, the smallest Akaike Information Criteria (aic) value and the smallest Mean Square Forecast Error (msfe).

Table 4.5. Out-of-sample forecast summary (Jan – Dec, 2018) of the three fitted models

A	B	C	D	E	F	G	H	I	J	K	L
121	Jan	12.4	10.01	11.00	13.11	2.39	1.40	-0.71	5.73	1.97	0.50
122	Feb	11.4	10.45	13.06	9.92	0.95	-1.66	1.48	0.90	2.76	2.19
123	Mar	10.9	12.21	12.67	9.73	-1.33	-1.77	1.17	1.76	3.14	1.37
124	Apr	10.8	10.49	12.45	11.01	0.31	-1.65	-0.21	0.10	2.73	0.04
125	May	10.6	9.86	12.66	11.13	0.74	-2.06	-0.53	0.55	4.25	0.28
126	June	9.6	9.22	13.64	10.24	0.38	-4.04	-0.64	0.14	16.31	0.41
127	July	10.5	8.62	12.04	13.06	1.88	-1.54	-2.56	3.52	2.36	6.55
128	Aug	10.9	8.73	12.16	12.44	2.17	-1.26	-1.54	4.70	1.60	2.36
129	Sep	12	9.11	11.18	13.33	2.89	0.82	-1.33	8.37	0.67	1.77
130	Oct	14.2	8.88	9.69	15.12	5.32	4.51	-0.92	28.27	20.33	0.86
131	Nov	13.3	8.62	11.94	11.18	4.68	1.36	2.12	21.89	1.84	4.51
132	Dec	17.1	8.56	7.52	17.71	8.54	9.58	-0.61	72.99	91.85	0.37
									148.91	149.819	21.19
							MSFE		**12.41**	**12.48**	**1.77***

* Smallest mean square forecast error (msfe)

Notations in Table 4.5:

A – Time t

B – Year 2018 (months)

C – Actual data (Y_t)

D – Multiple linear regression forecast

E – Arimax with normal error forecast

F – Arimax with lognormal forecast

G – Multiple linear regression (Residual)

H – Arimax with normal error (Residual)

I – Arimax with lognormal (Residual)

J – Multiple linear regression (Square residual)

K – Arimax with normal error (Square residual)

L – Arimax with lognormal error (Square residual)

Table 4.5 contained the mean square forecast error (msfe) value of the proposed arimax with lognormal error was 1.766; the multiple linear regression msfe was 12.409 and that of arimax with normal error msfe was 12.484. The proposed model had the smallest msfe compared to the other two models. Hence, the proposed model was considered better and improved good fitting of non-normal error and non-stationary time series of economic data.

4.8 In-sample forecast analysis of three fitted models: Fig. 4.17 – Fig. 4.19

Fig. 4.17 showed the in-sample forecast trend of fitted multiple linear regression model exhibited fluctuation over the period of time. The trend indicated non-stationary series at raw level; which was common to economic time series data.

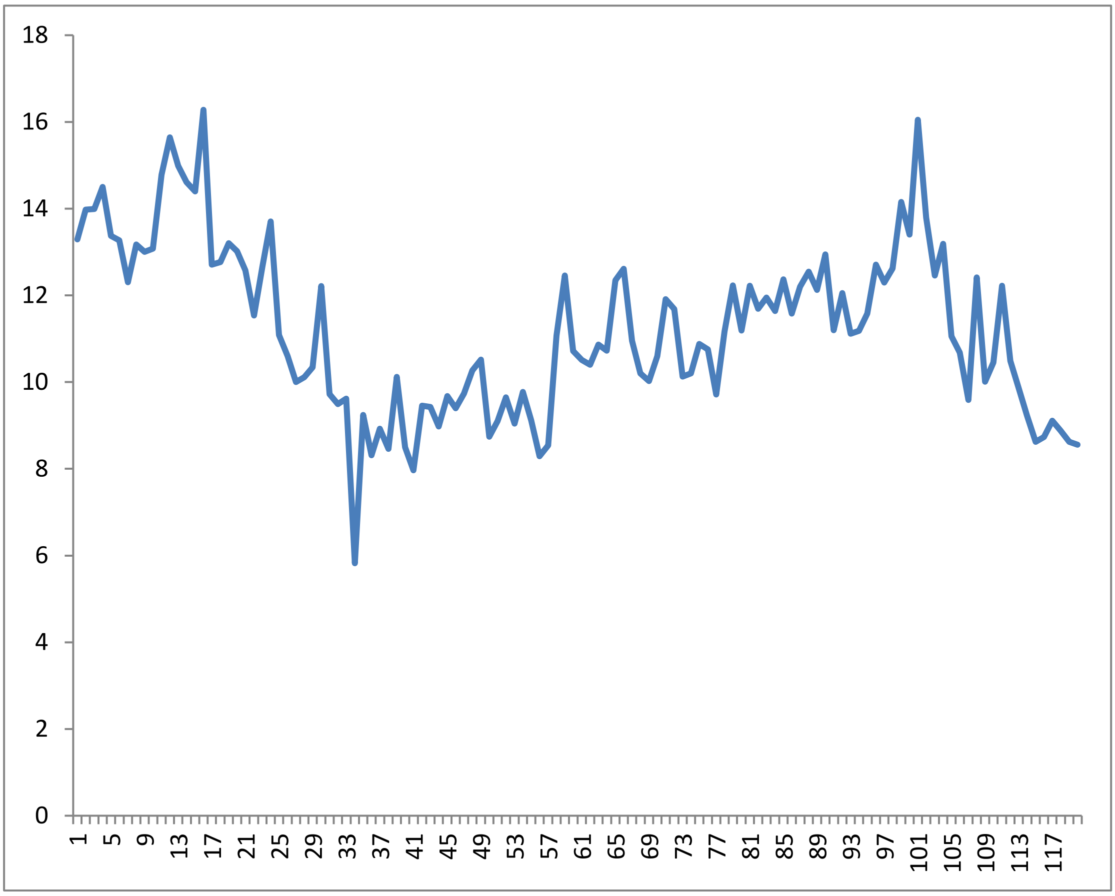


Fig. 4.17. In-sample forecast plot of multiple linear regression model

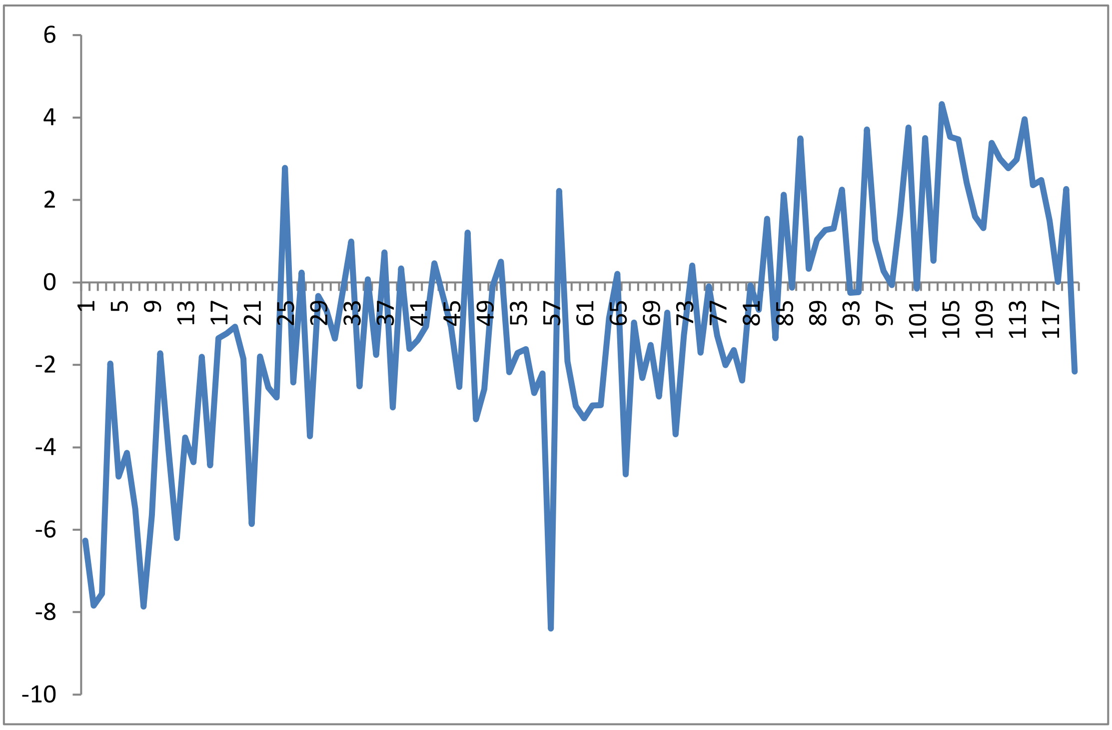


Fig. 4.18. In-sample forecast plot of arimax (1, 1, 1) model assuming normal error

Fig. 4.18 showed in-sample forecast: the trend fluctuated from negative during the early months of the out-of-sample year. There was indication of non-stationary of series at raw level.

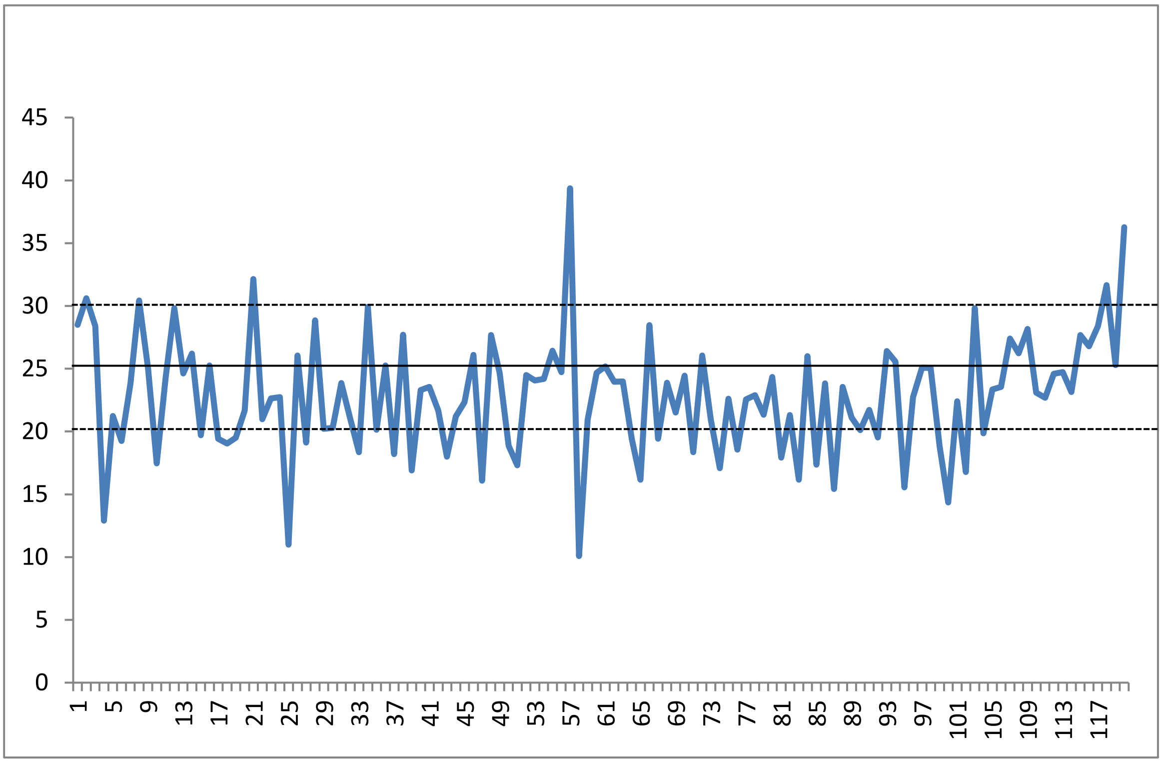


Fig. 4.19. In-sample forecast plot of arimax(1, 0, 1) assuming lognormal error

Fig. 4.19 showed the in-sample forecast plot of the proposed arimax (1, 0, 1) model with lognormal error exhibited some level of stationary (consistency of mean and variance) series compared to in-sample forecast plot of arimax (1, 1, 1) with normal error and multiple linear regression. That showed the superiority of the proposed model of arimax with lognormal error.

4.9 Residual analyses of in-sample forecast

The residual analyses of the forecast from the three fitted models under consideration were presented in Fig. 4.20 to Fig. 4.22.

Fig. 2.20 residual plot of multiple linear regression indicated fluctuation and non stationrity of the series residual which deviated from white noise error. Fig. 4.21 shown non stationarity in level and in slope of the residual plot of fitted arimax (1, 1, 1) assumed normal error and fig. 4.22 revealed constant mean at three stages of time as indicated on the plot. Tha implied stability in the residual of fitted arimax (1, 0,1) model assumed lognormal error.

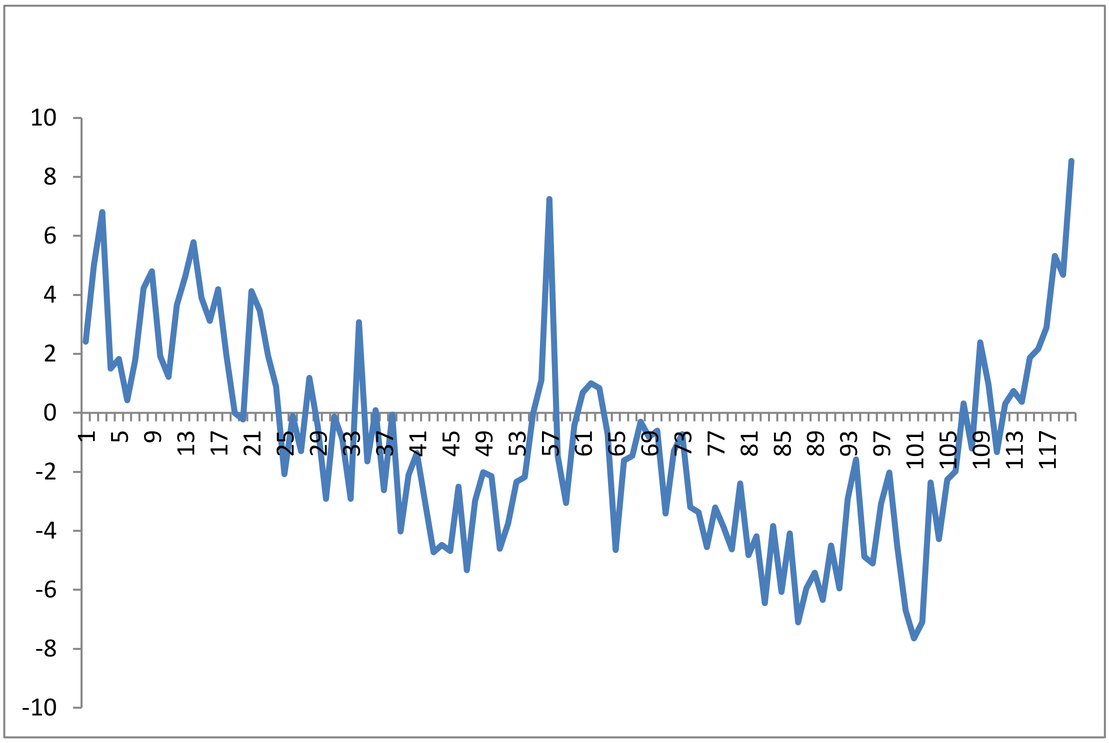


Fig. 4.20. Residual plot of fitted multiple linear regression (mlr)

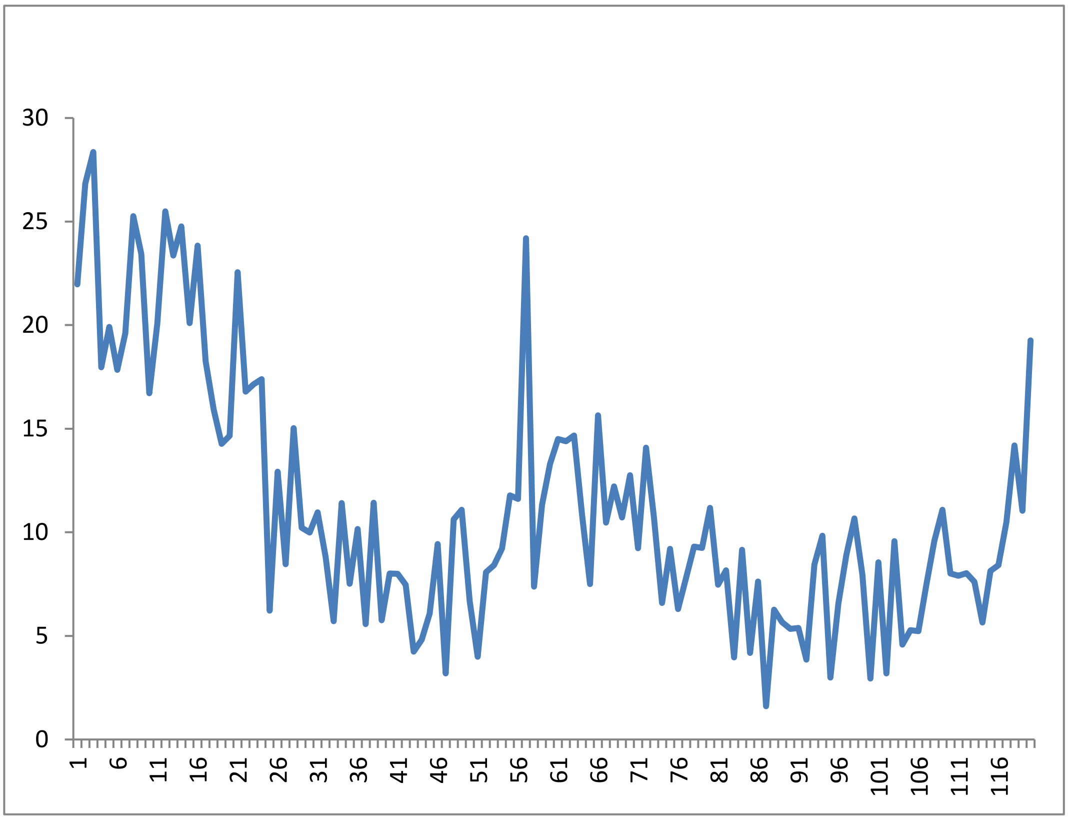


Fig. 4.21. Residual plot of arimax (1,1,1) assuming normal error

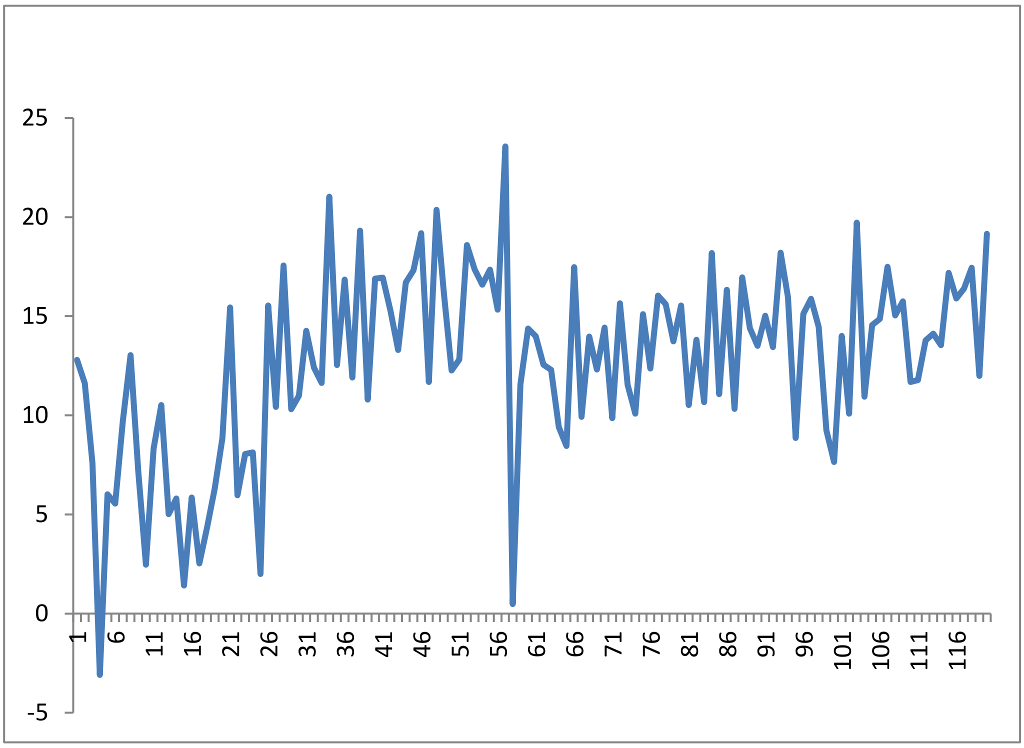


Fig. 4.22: Residual plot of arimax (1,0,1) assuming lognormal error

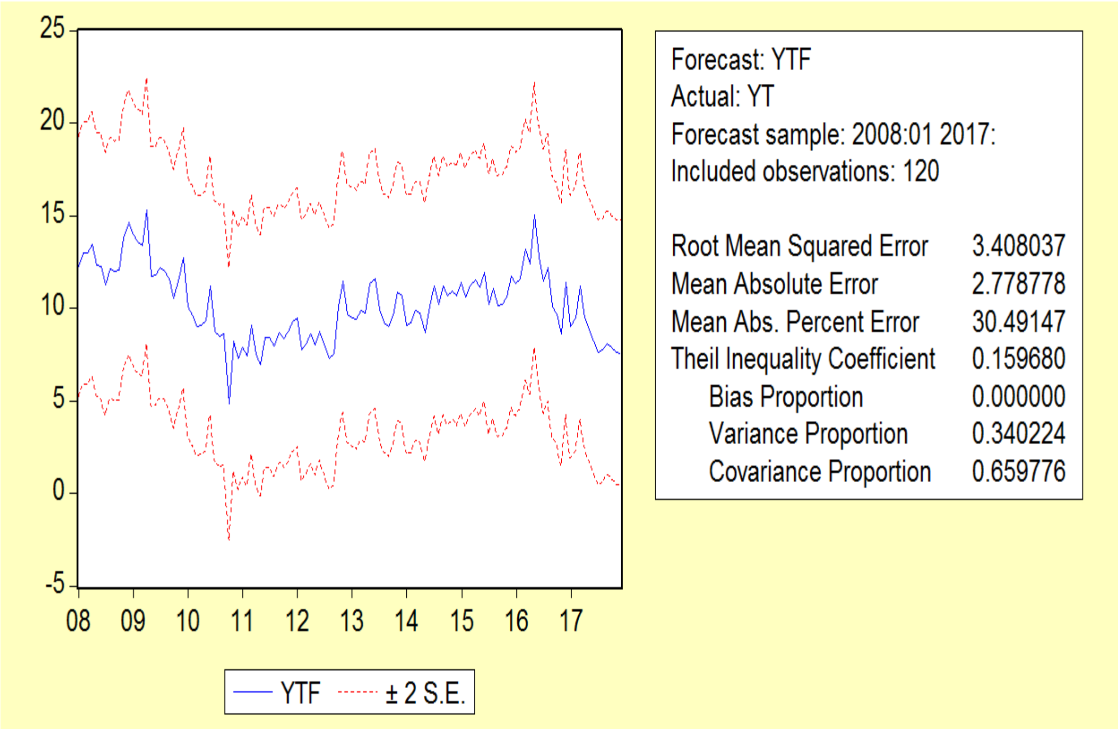


Fig. 4.23. Forecast plot of external reserve

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.346	0.346	14.733	0.000
		2	0.206	0.098	19.995	0.000
		3	0.037	-0.070	20.169	0.000
		4	0.065	0.060	20.695	0.000
		5	0.116	0.102	22.412	0.000
		6	0.079	-0.004	23.205	0.001
		7	-0.006	-0.070	23.210	0.002
		8	0.000	0.021	23.210	0.003
		9	0.017	0.028	23.249	0.006
		10	0.010	-0.025	23.262	0.010
		11	0.022	0.013	23.325	0.016
		12	0.028	0.035	23.433	0.024
		13	0.027	0.008	23.530	0.036
		14	0.044	0.022	23.801	0.048
		15	-0.079	-0.123	24.682	0.054
		16	-0.067	-0.020	25.309	0.065
		17	-0.027	0.037	25.411	0.086
		18	0.112	0.131	27.197	0.075
		19	0.106	0.029	28.827	0.069
		20	0.001	-0.078	28.827	0.091
		21	-0.087	-0.069	29.939	0.093
		22	-0.095	-0.043	31.276	0.091
		23	-0.098	-0.081	32.726	0.086
		24	-0.078	-0.042	33.662	0.091
		25	-0.070	0.005	34.407	0.099
		26	-0.040	0.035	34.655	0.119
		27	-0.104	-0.092	36.347	0.108
		28	0.001	0.080	36.348	0.134
		29	-0.048	-0.024	36.714	0.154
		30	0.102	0.123	38.403	0.140

Fig. 4.24. Acf, pacf and Q-sta plots of external reserve residuals

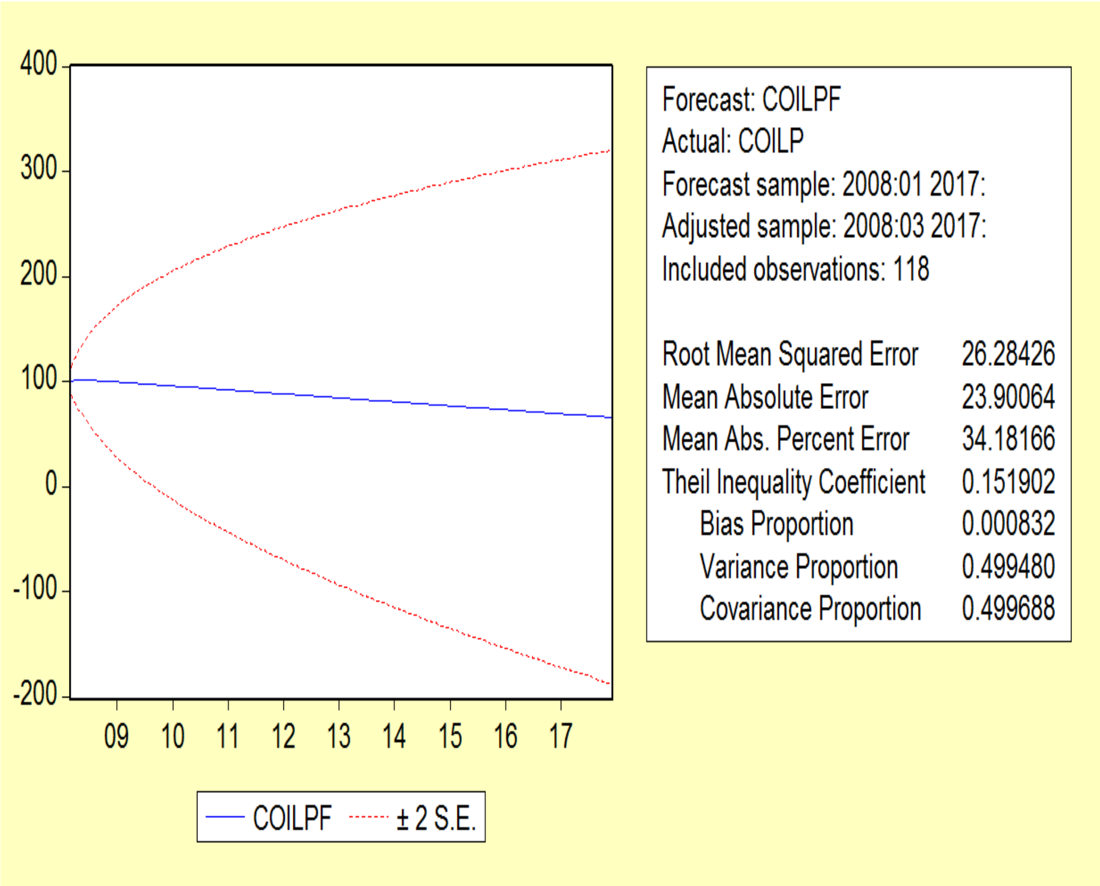


Fig. 4.25: Crude oil price forecast plot

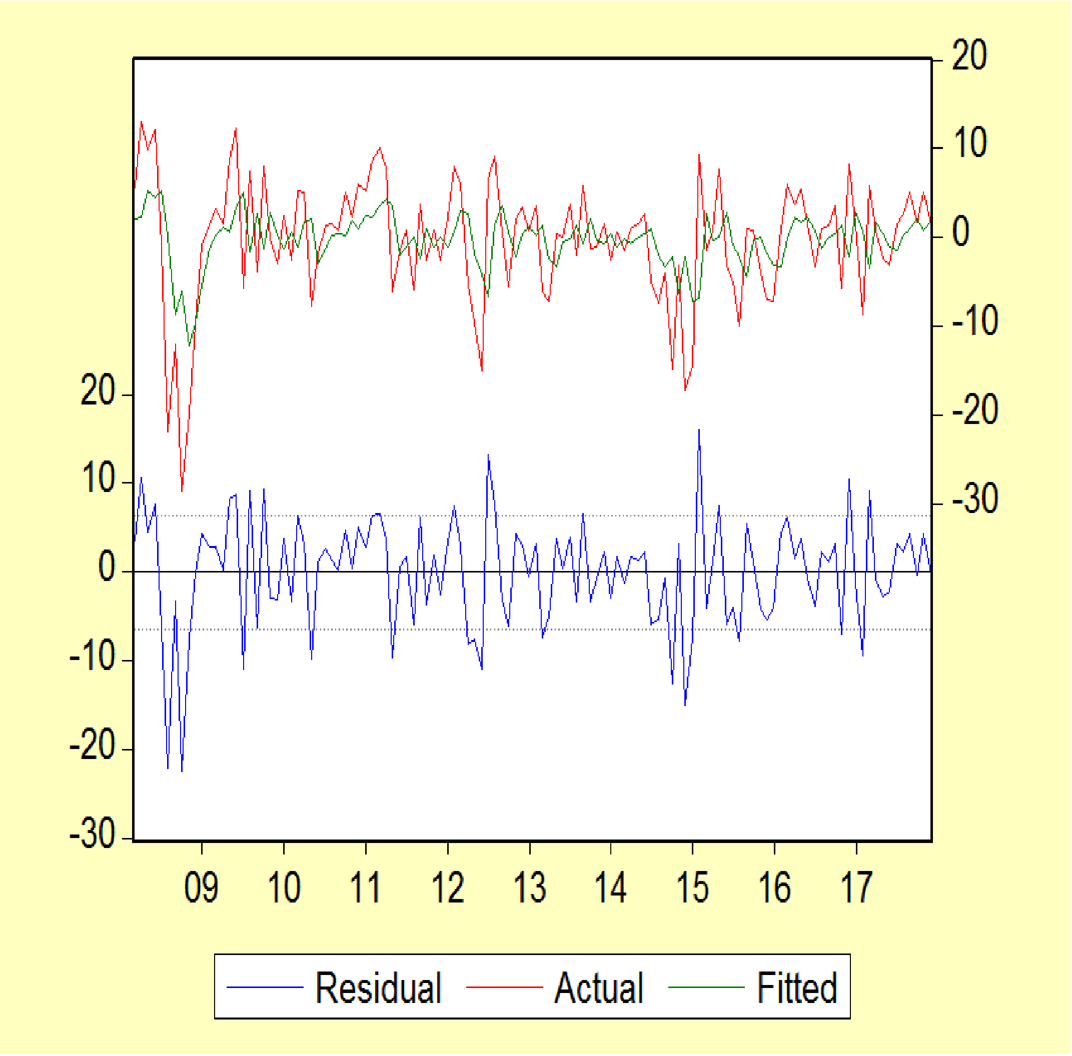


Fig. 4.26. Residual plot of crude oil price

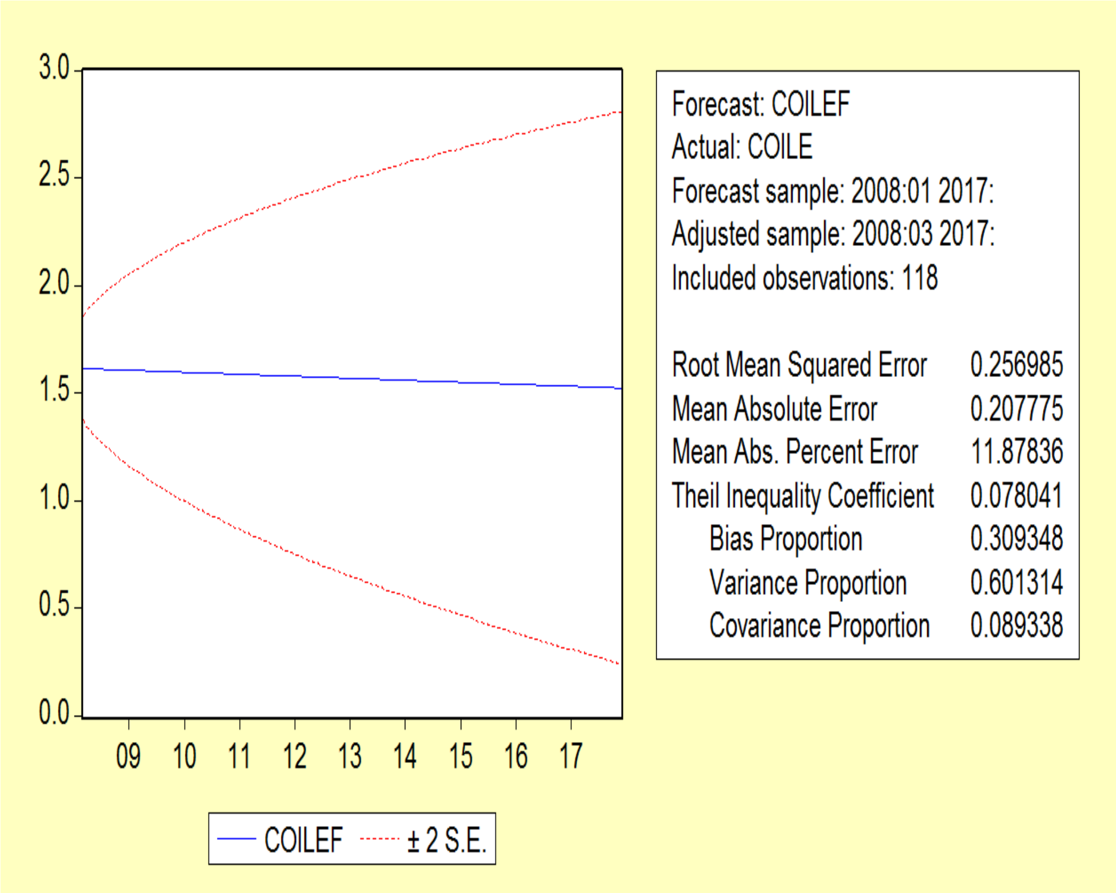


Fig. 4.27. Crude oil export plot

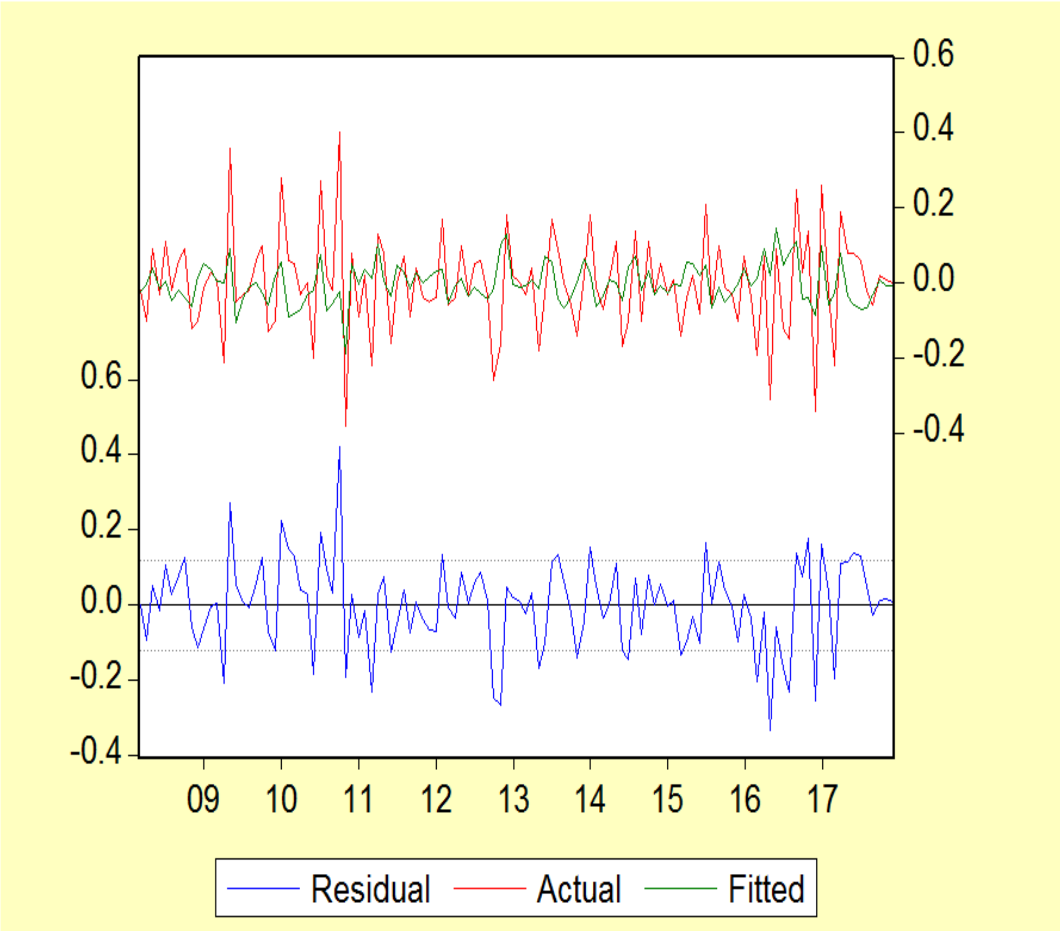


Fig. 4.28. Residual plot of crude oil export

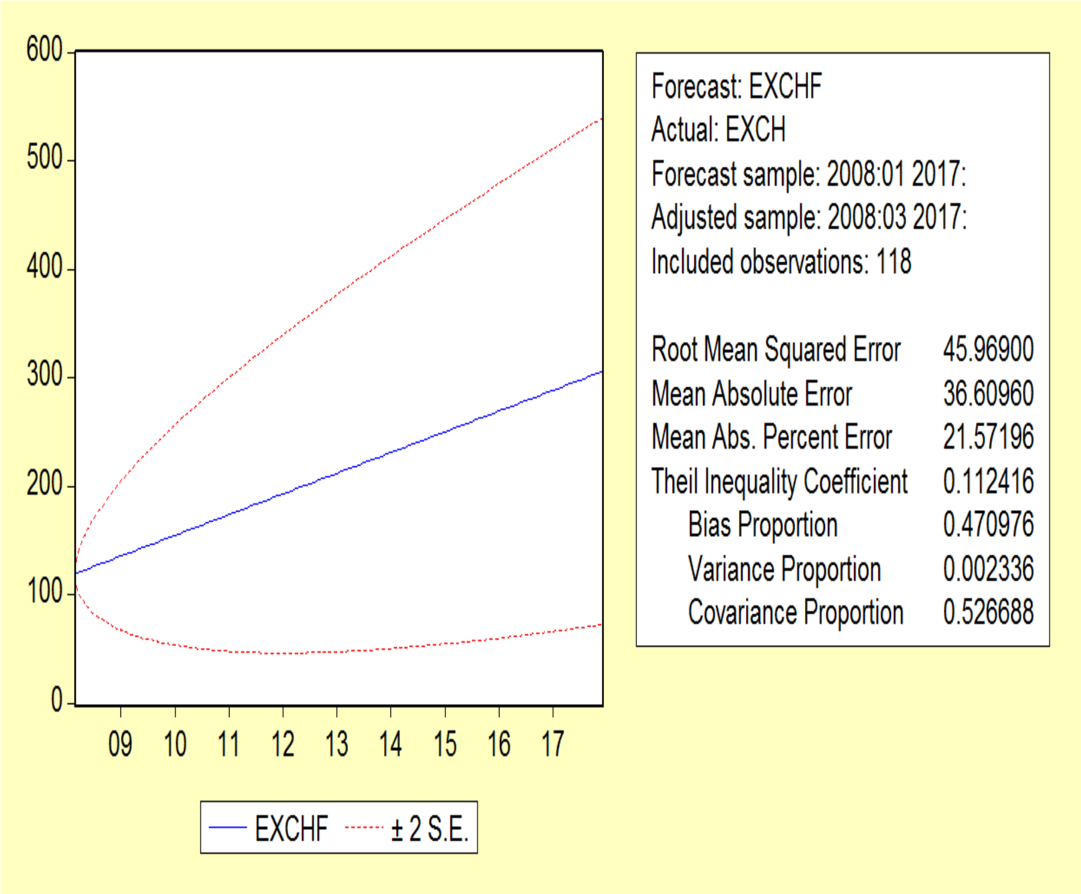


Fig. 4.29. Exchange rate forecast plot

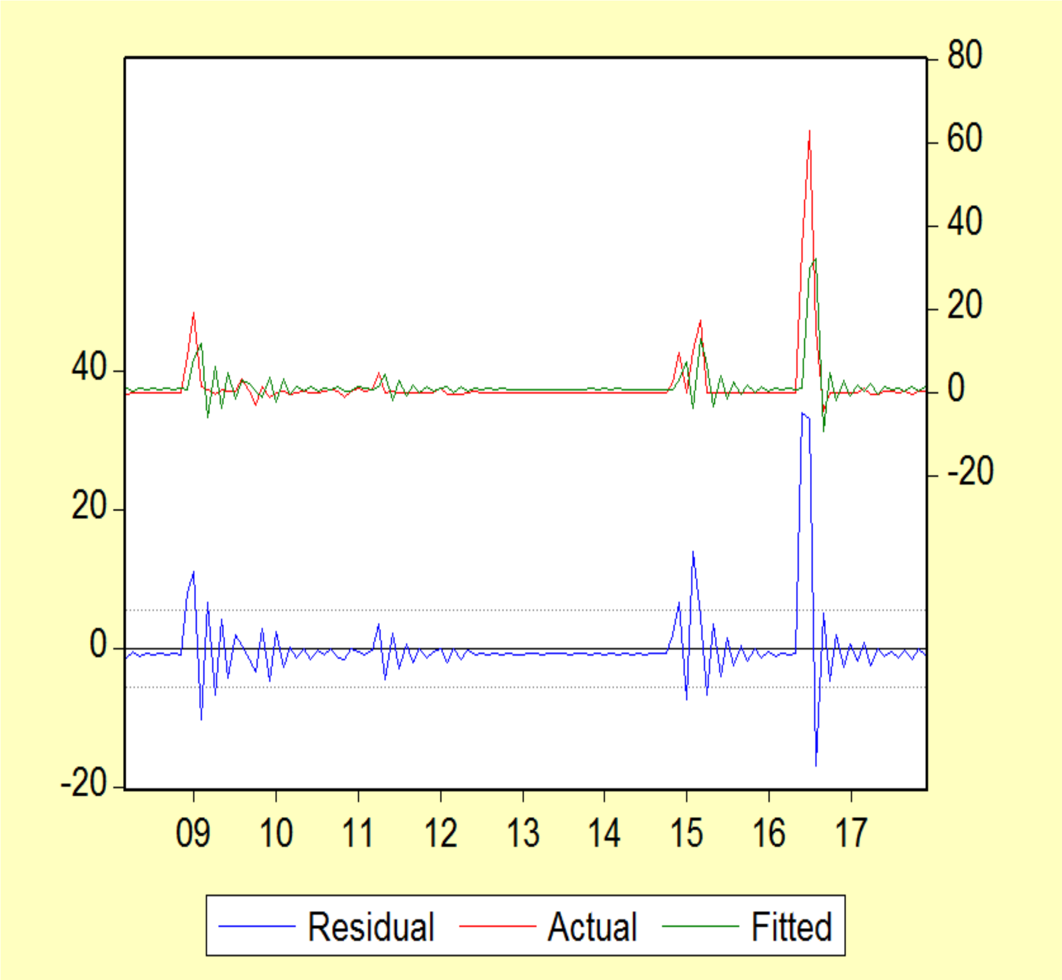


Fig. 4.30. Residual of exchange rate plot

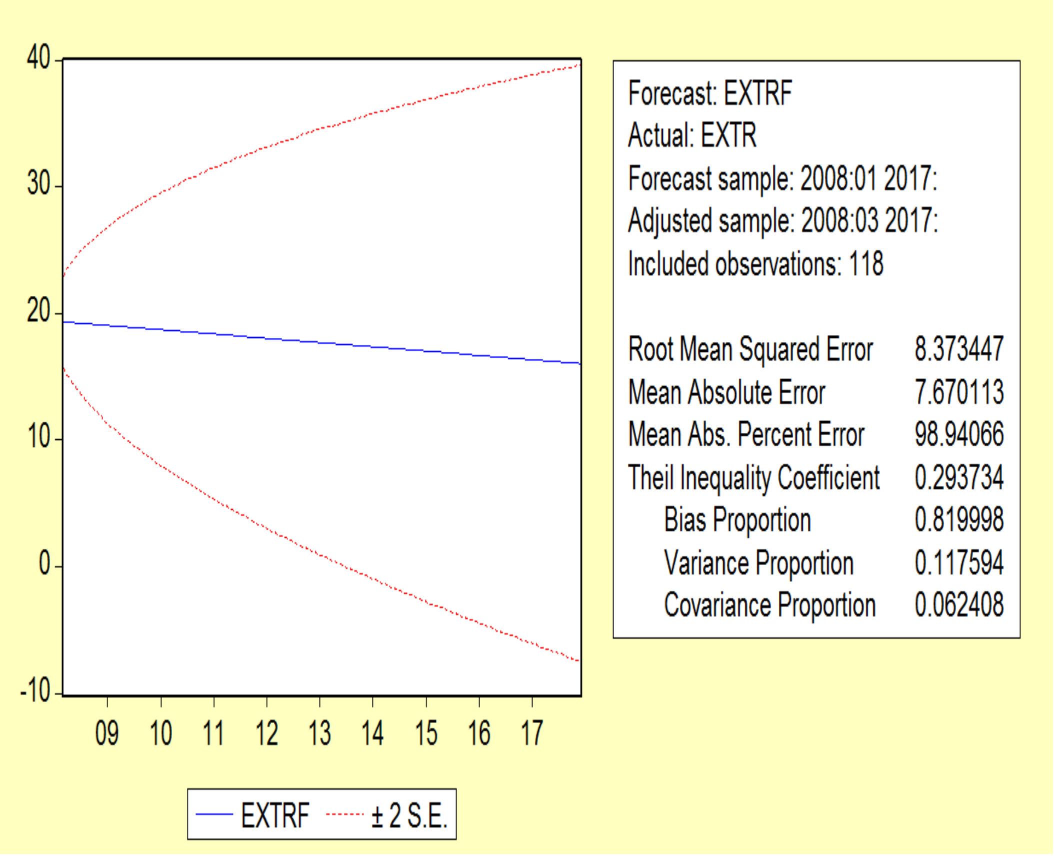


Fig. 4.31. External reserve forecast plot

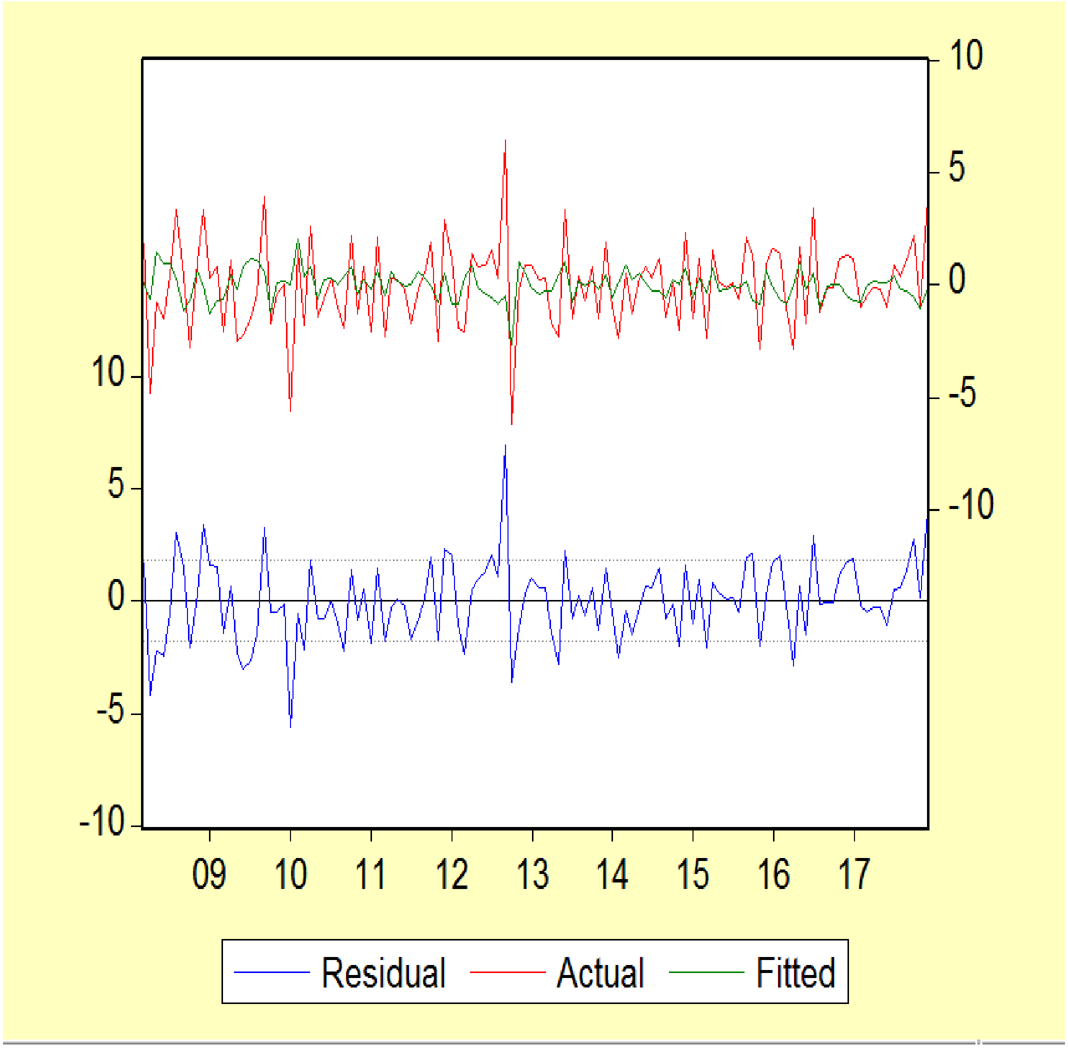


Fig. 4.32. Residual plot of external reserve

Table 4.6. Multiple linear regression fitted model results

Included observations: 120

EXTR= C(1)*EXCH + C(2)*COILE + C(3)*COILP + C(4)

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	-0.040067	0.007906	-5.067829	0.0000
C(2)	-9.074738	1.891591	-4.797409	0.0000
C(3)	-0.031519	0.013839	-2.277590	0.0246
C(4)	35.57873	4.227570	8.415883	0.0000

R-squared	0.242348	Mean dependent var	10.21917
Adjusted R-squared	0.222753	S.D. dependent var	3.931756
S.E. of regression	3.466299	Akaike info criterion	5.356817
Sum squared resid	1393.766	Schwarz criterion	5.449733
Log likelihood	-317.4090	Durbin-Watson stat	0.451389

As shown in Table 4.6 the fitted equation of multiple linear regression model obtained as:

$$\hat{Y}_t = 35.58 - 0.04X_{t1} - 9.07X_{t2} - 0.03X_{t3}$$

(4.23) (0.01) (1.90) (0.01)

Where;

$\hat{Y}_t$ - External reserve, X_{t1} - Exchange rate, X_{t2} - Crude oil export, X_{t3} - Crude oil price and *t* – time (in months).

The coefficient values of the variables estimated were so stated while the values in brackets were the associated standard error of the estimated coefficients. The t-statistic indicated significance of the coefficients at 0.05 critical levels while the model's coefficient of determination R-square value 0.24. That showed the fitted model could account for about 24% of the total variation which implied about 76% unaccounted (error). The model may not be adequate for forecast external reserve with respect to time factor.

The Durbin-Watson statistic was about 0.45 less than value of 2.0 criteria. Hence there was cleared indication of autocorrelation among the series. That considered true since the data under study were time series data.

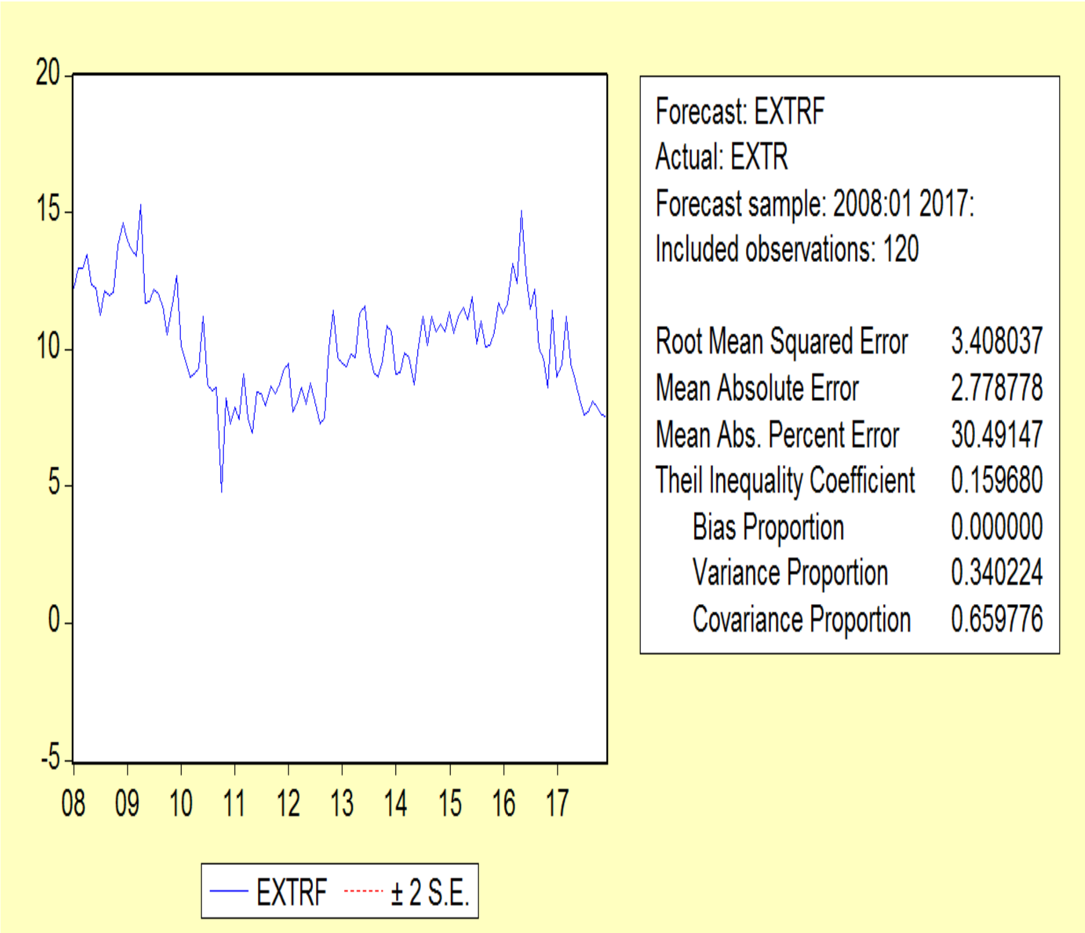


Fig. 4.33. Multiple linear regression model forecast plot

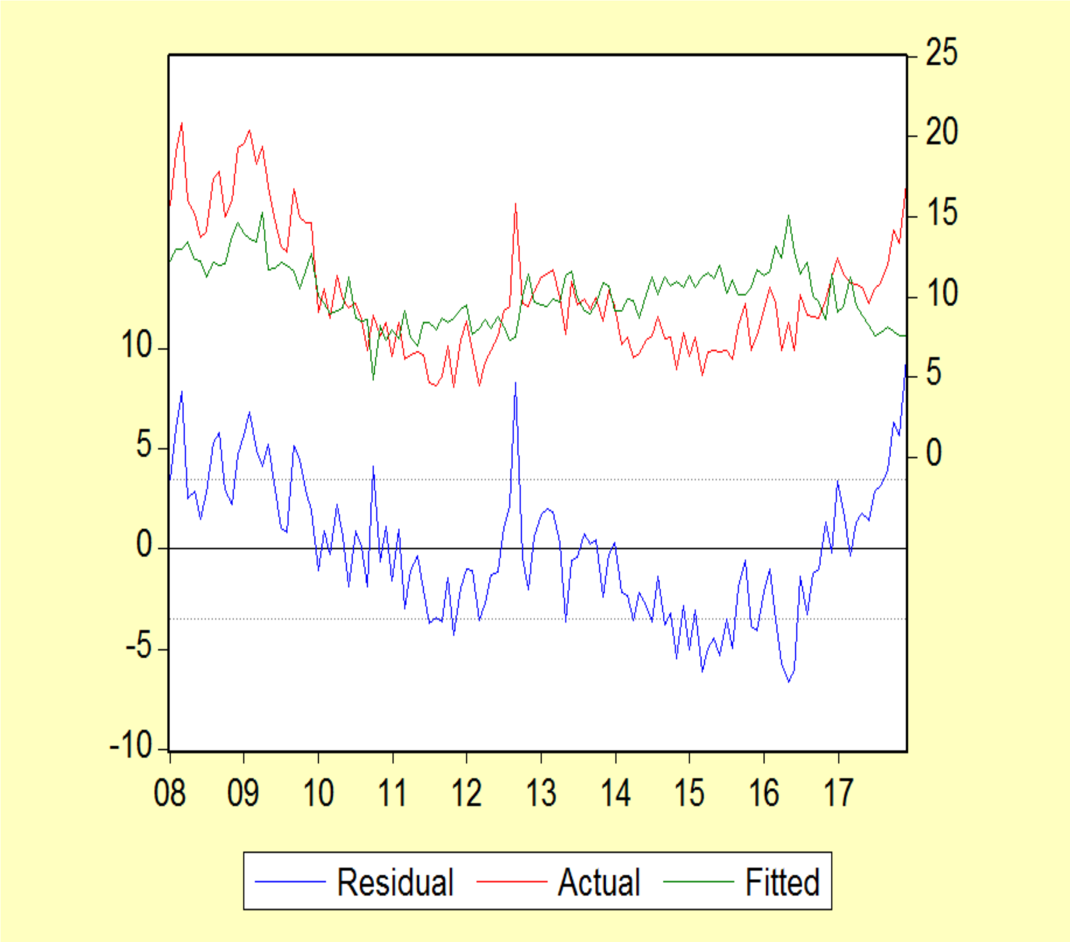


Fig. 4.34. Residual plot of fitted multiple linear regression model

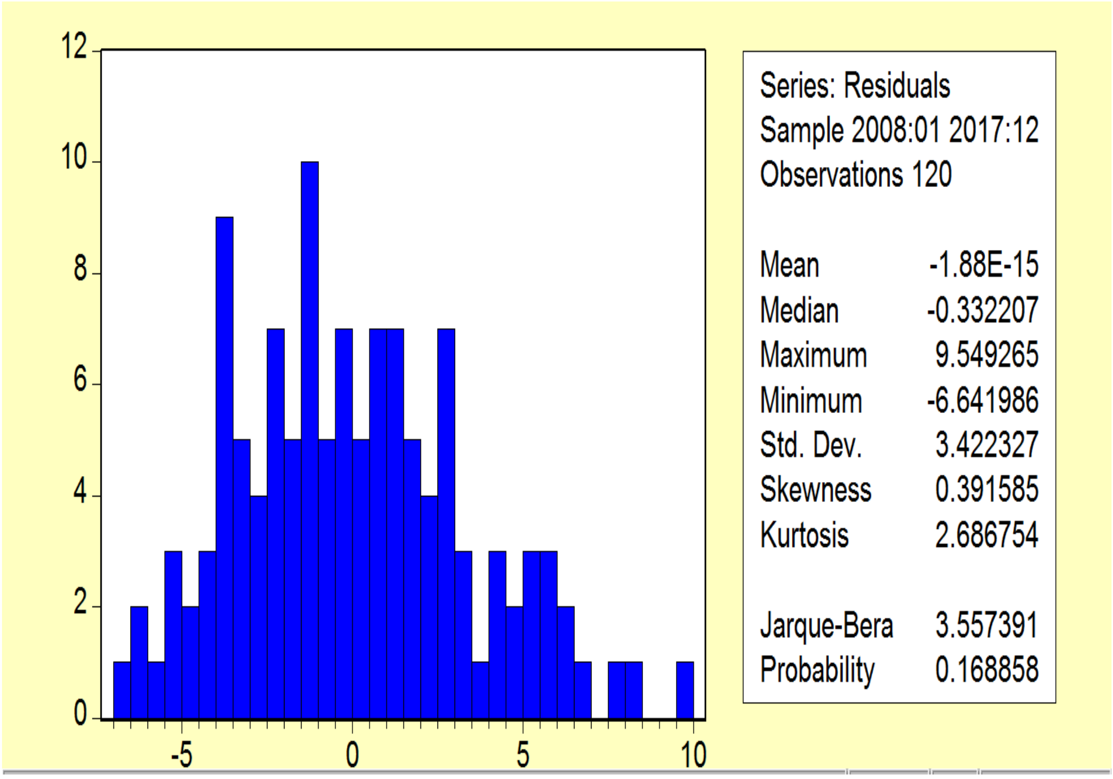


Fig. 4.35. Normality test of fitted multiple linear regression residual

The plot of normality test of residual of the fitted multiple linear regression model showed that the residual were not normally distributed. The histogram plot (Fig. 4.35) of the residual series deviated from normal curve. The pattern did not conform with bell shape property of normal curve.

The skew value (0.39) was positive and greater than zero property of normal curve. The kurtosis value (2.69) deviated from normality peak value of 3.0 of normal curve. The Jarque-Bera value (3.56) with p-value of 0.17 indicated the residual deviation from normality assumption. Hence, the residual series were not *white noise*. The multiple linear regression was not adequate in fitting series that were non-normally distributed and non-stationary.

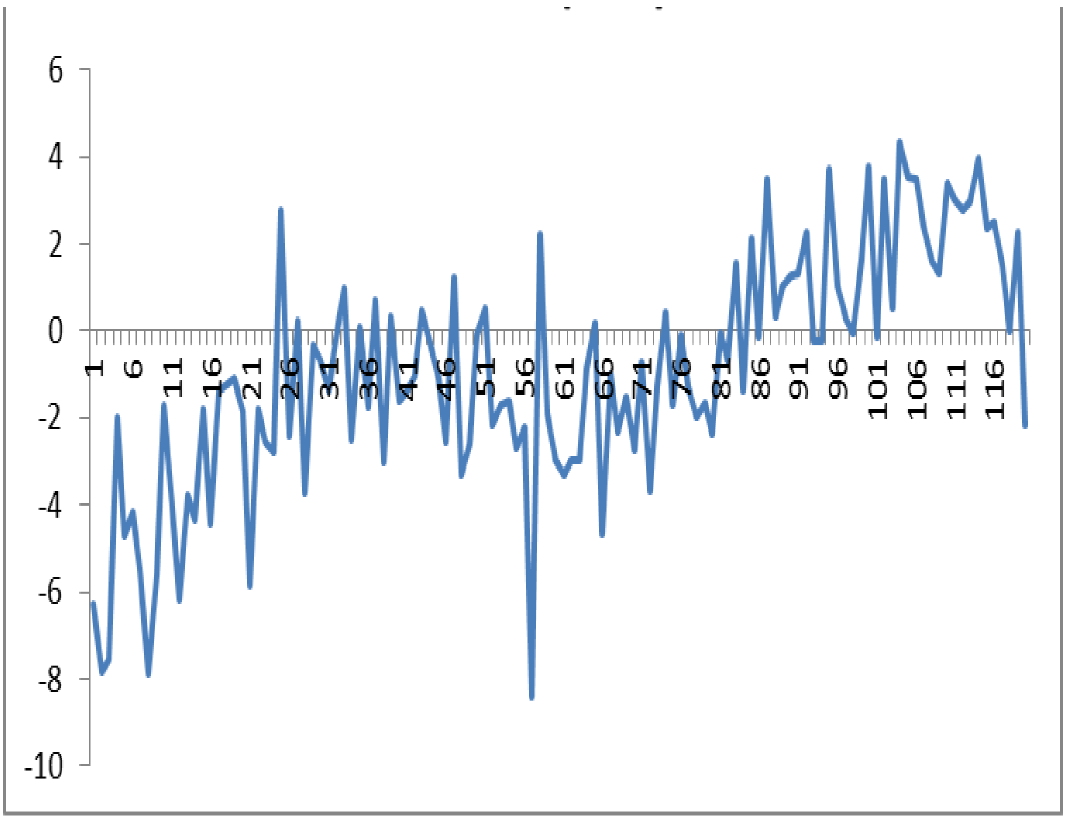


Fig. 4.36. Arimax (1, 1, 1) with normal error in-sample forecast plot

4.10 Fitted Models Forecast Plots

The three fitted models were subjected to In-sample forecast to enable comparison of their forecast capability. Figures 4.36 – Fig. 4.38 showed the plots.

The pattern of the series in Fig. 4.36 showed non-stationarity in level and in slope but exhibited homogeneity in some defined period of time.

The proposed model arimax assuming lognormal error in-sample forecast plot as shown in Fig. 4.37 exhibited stationarity in level.

The pattern of the series in Fig. 4.38 showed non-stationarity in level and in slope but exhibited homogeneity in some defined period of time.

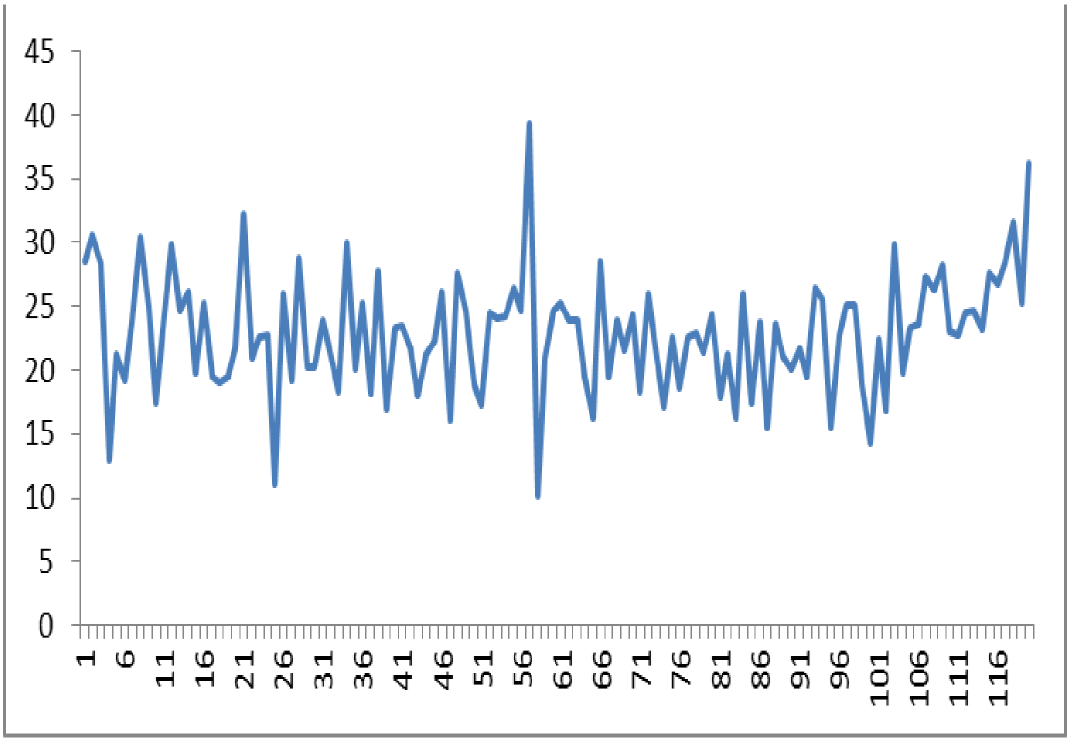


Fig. 4.37. Arimax (1, 0, 1) with lognormal error in-sample forecast plot

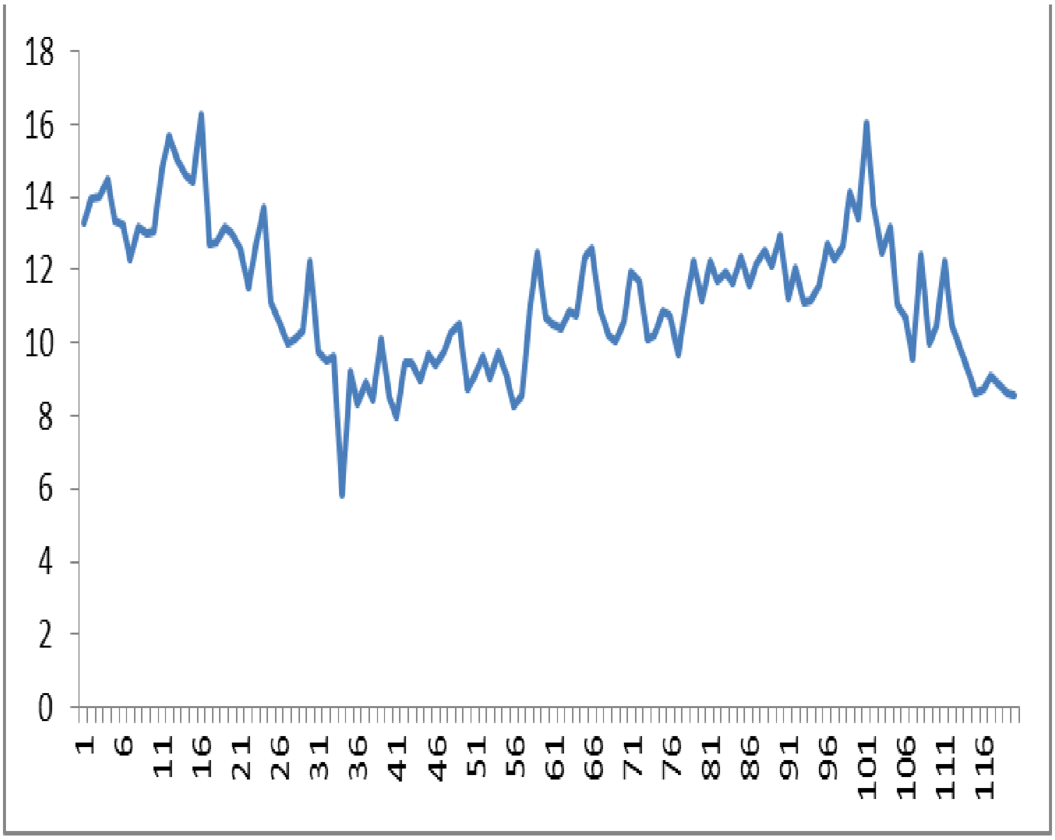


Fig. 4.38. Multiple linear regression in-sample forecast plot

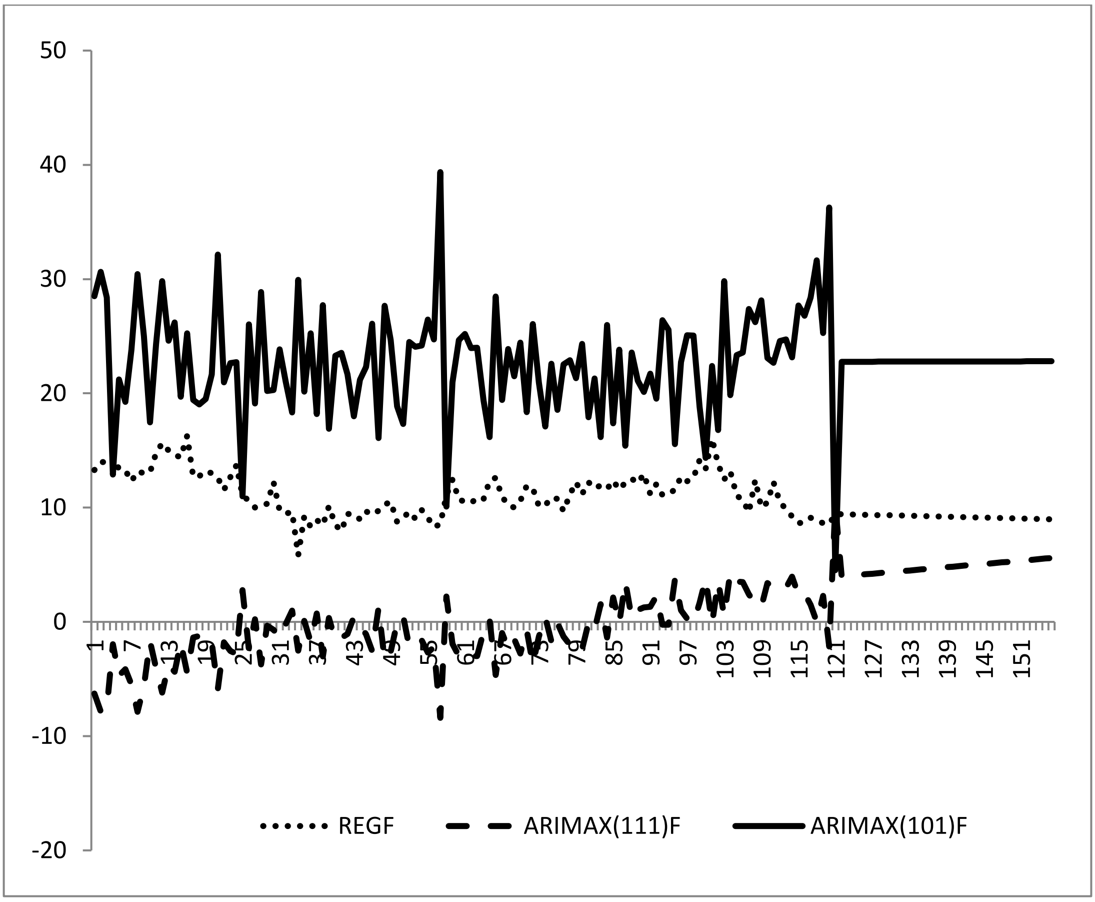


Fig. 4.39. Combined forecast plot of the three fitted models

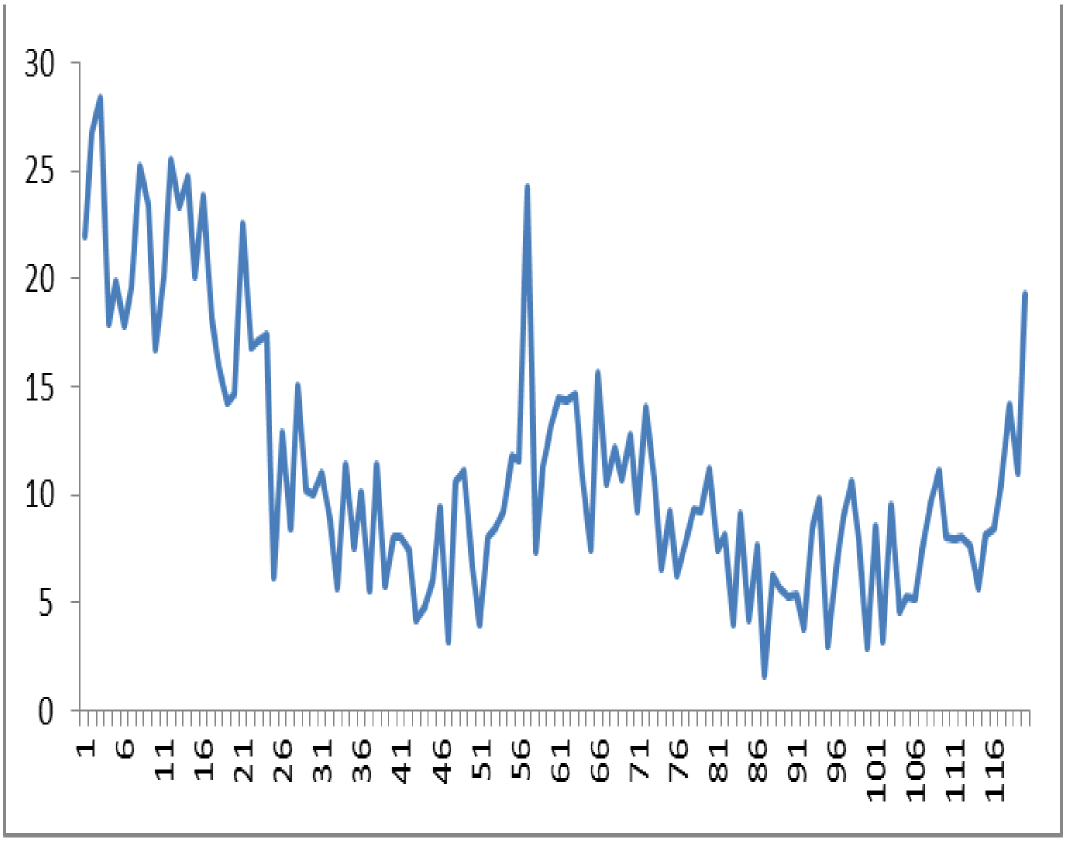


Fig. 4.40. Arimax (1, 1, 1) assuming normal error residual plot

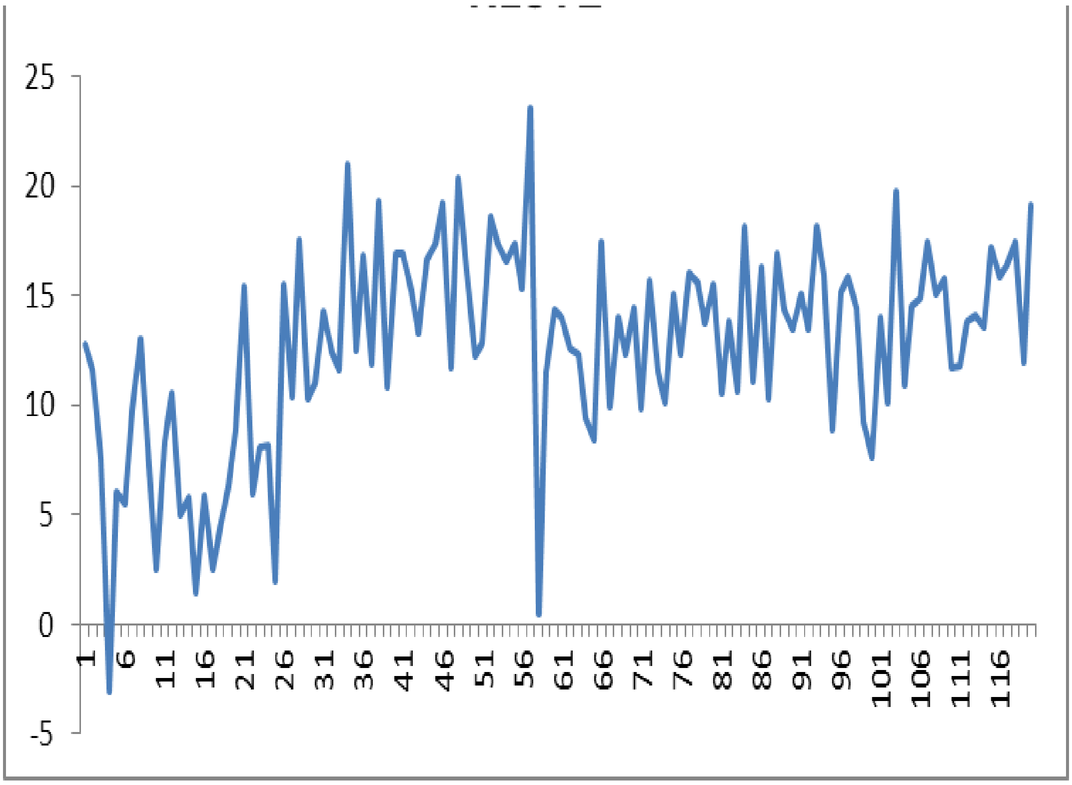


Fig. 4.41. Arimax (1, 0, 1) with lognormal error residual plot

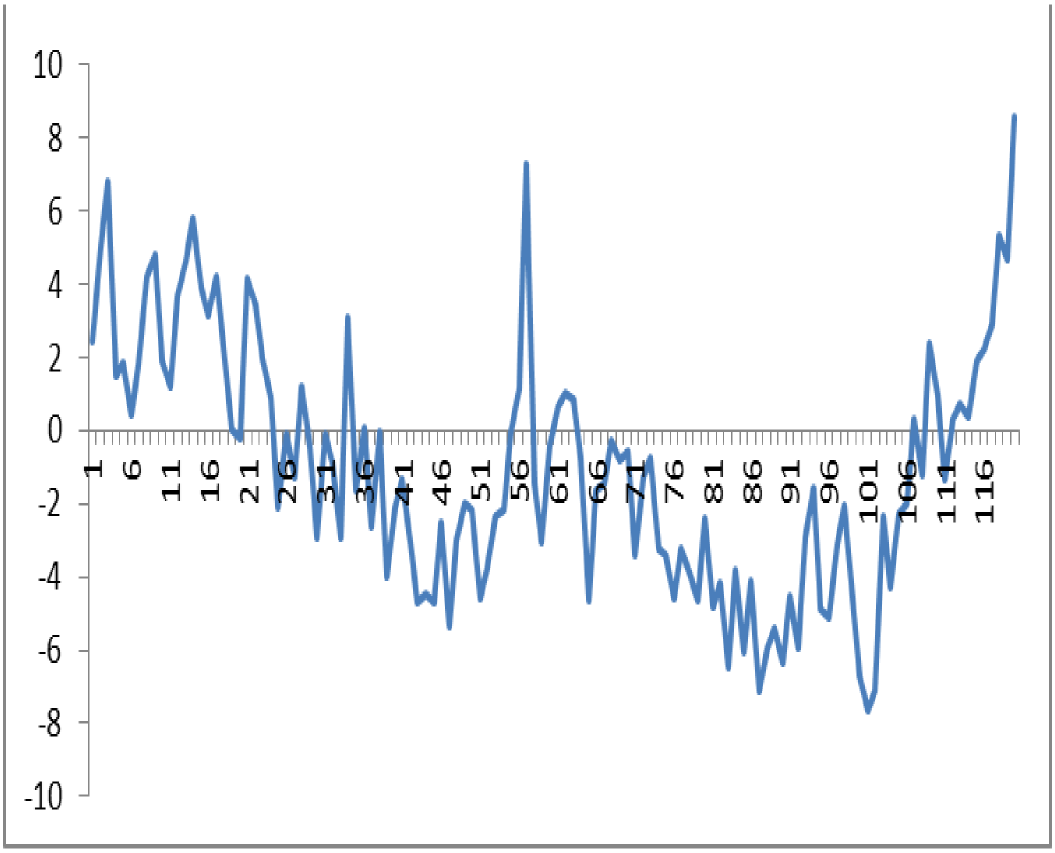


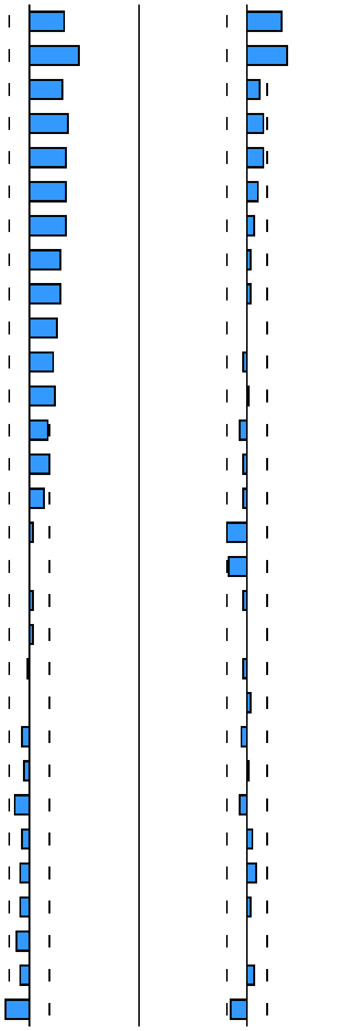
Fig. 4.42. Multiple linear regression residual plot

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.541	0.541	36.061	0.000
		2	0.610	0.448	82.187	0.000
		3	0.518	0.165	115.78	0.000
		4	0.567	0.215	156.39	0.000
		5	0.549	0.175	194.74	0.000
		6	0.514	0.053	228.63	0.000
		7	0.493	0.029	260.11	0.000
		8	0.437	-0.051	285.08	0.000
		9	0.412	-0.066	307.52	0.000
		10	0.387	-0.042	327.45	0.000
		11	0.345	-0.074	343.41	0.000
		12	0.337	-0.018	358.77	0.000
		13	0.277	-0.059	369.31	0.000
		14	0.278	-0.003	380.00	0.000
		15	0.267	0.055	389.92	0.000
		16	0.170	-0.124	393.99	0.000
		17	0.165	-0.061	397.87	0.000
		18	0.192	0.110	403.17	0.000
		19	0.174	0.049	407.55	0.000
		20	0.136	-0.015	410.24	0.000
		21	0.120	0.020	412.36	0.000
		22	0.077	-0.050	413.24	0.000
		23	0.068	-0.040	413.95	0.000
		24	0.013	-0.112	413.97	0.000
		25	0.056	0.022	414.45	0.000
		26	0.024	0.026	414.54	0.000
		27	0.034	0.030	414.72	0.000
		28	-0.021	-0.021	414.79	0.000
		29	0.010	0.050	414.81	0.000
		30	-0.091	-0.151	416.15	0.000

Fig. 4.43. Multiple linear regression model acf, pacf chart

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.335	0.335	13.821	0.000
		2	0.459	0.391	39.967	0.000
		3	0.314	0.119	52.319	0.000
		4	0.369	0.150	69.512	0.000
		5	0.357	0.157	85.769	0.000
		6	0.347	0.098	101.19	0.000
		7	0.339	0.080	116.09	0.000
		8	0.299	0.027	127.76	0.000
		9	0.298	0.028	139.44	0.000
		10	0.260	-0.005	148.40	0.000
		11	0.232	-0.038	155.65	0.000
		12	0.249	0.021	164.03	0.000
		13	0.179	-0.058	168.41	0.000
		14	0.184	-0.041	173.10	0.000
		15	0.143	-0.038	175.93	0.000
		16	0.026	-0.194	176.03	0.000
		17	-0.004	-0.179	176.04	0.000
		18	0.027	-0.026	176.14	0.000
		19	0.025	0.000	176.23	0.000
		20	-0.009	-0.038	176.24	0.000
		21	0.004	0.031	176.24	0.000
		22	-0.081	-0.046	177.22	0.000
		23	-0.042	0.025	177.49	0.000
		24	-0.144	-0.078	180.67	0.000
		25	-0.065	0.048	181.32	0.000
		26	-0.085	0.083	182.45	0.000
		27	-0.084	0.033	183.57	0.000
		28	-0.127	0.003	186.14	0.000
		29	-0.088	0.081	187.38	0.000
		30	-0.234	-0.161	196.29	0.000

Fig. 4.44. Arimax (1, 1, 1) model assuming normal error acf, pacf of external reserve (1998 – 2007).



	AC	PAC	Q-Stat	Prob
1	0.335	0.335	13.821	0.000
2	0.459	0.391	39.967	0.000
3	0.314	0.119	52.319	0.000
4	0.369	0.150	69.512	0.000
5	0.357	0.157	85.769	0.000
6	0.347	0.098	101.19	0.000
7	0.339	0.080	116.09	0.000
8	0.299	0.027	127.76	0.000
9	0.298	0.028	139.44	0.000
10	0.260	-0.005	148.40	0.000
11	0.232	-0.038	155.65	0.000
12	0.249	0.021	164.03	0.000
13	0.179	-0.058	168.41	0.000
14	0.184	-0.041	173.10	0.000
15	0.143	-0.038	175.93	0.000
16	0.026	-0.194	176.03	0.000
17	-0.004	-0.179	176.04	0.000
18	0.027	-0.026	176.14	0.000
19	0.025	0.000	176.23	0.000
20	-0.009	-0.038	176.24	0.000
21	0.004	0.031	176.24	0.000
22	-0.081	-0.046	177.22	0.000
23	-0.042	0.025	177.49	0.000
24	-0.144	-0.078	180.67	0.000
25	-0.065	0.048	181.32	0.000
26	-0.085	0.083	182.45	0.000
27	-0.084	0.033	183.57	0.000
28	-0.127	0.003	186.14	0.000
29	-0.088	0.081	187.38	0.000
30	-0.234	-0.161	196.29	0.000

Fig. 4.45: Arimax (1, 0, 1) assuming lognormal error acf, pacf of external reserve (1998 – 2007).

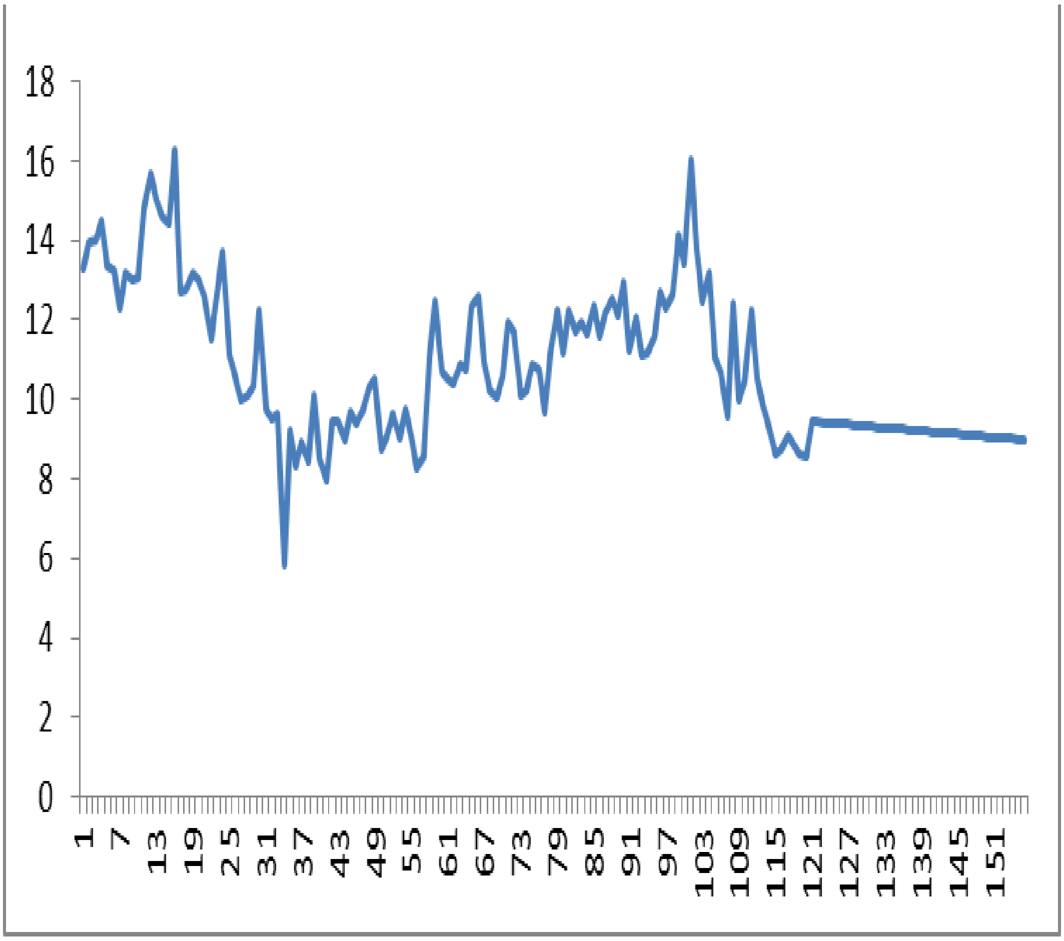


Fig. 4.46. Multiple linear regression out of sample forecast plot

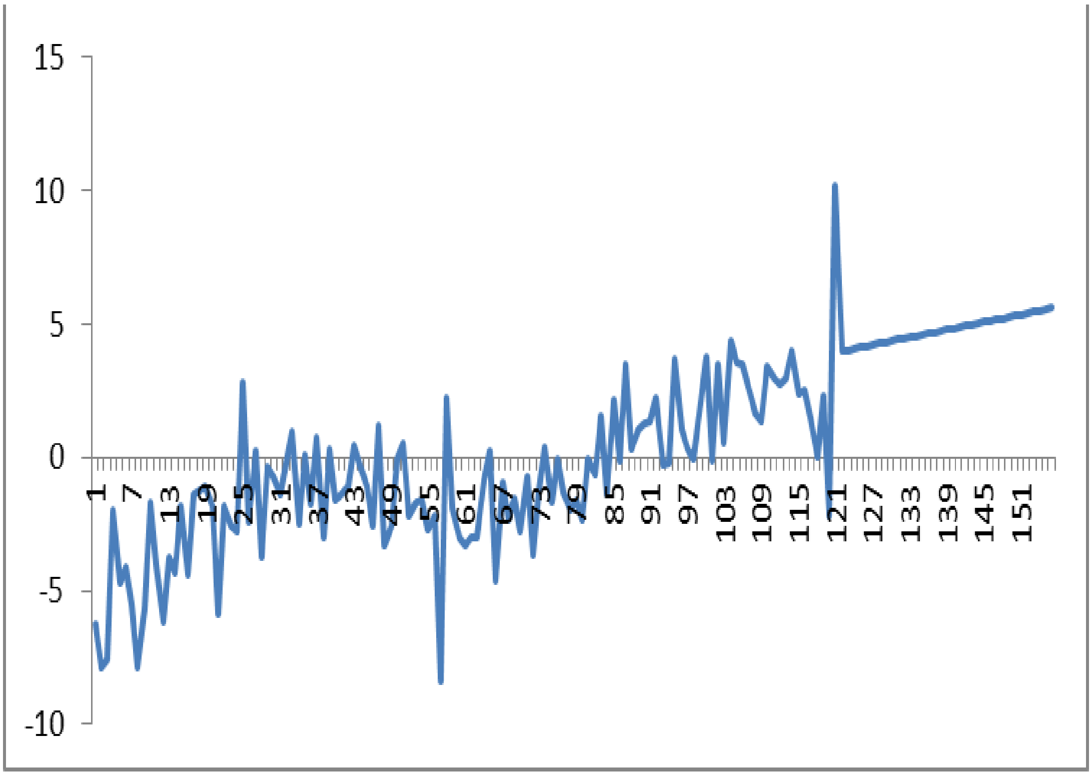


Fig. 4.47. Arimax (1, 1, 1) model assuming normal error out of sample plot

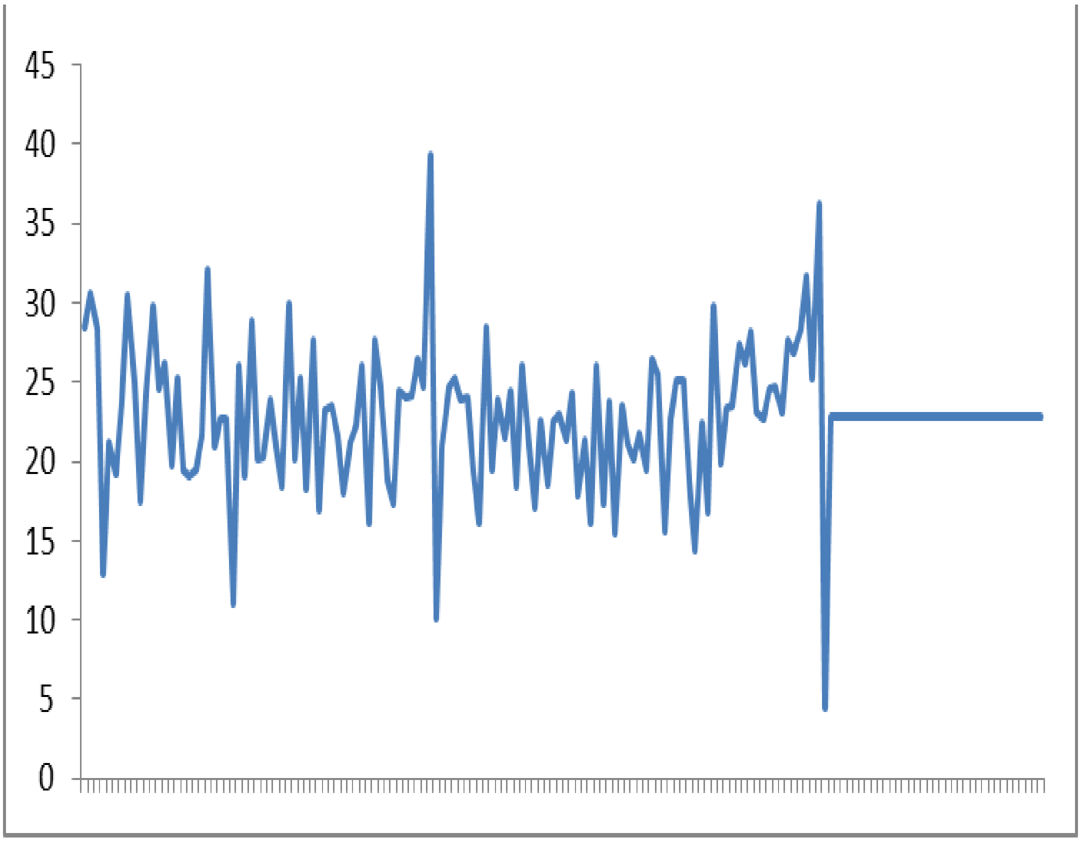


Fig. 4.48. Arimax (1, 0, 1) model assuming lognormal error out of sample plot

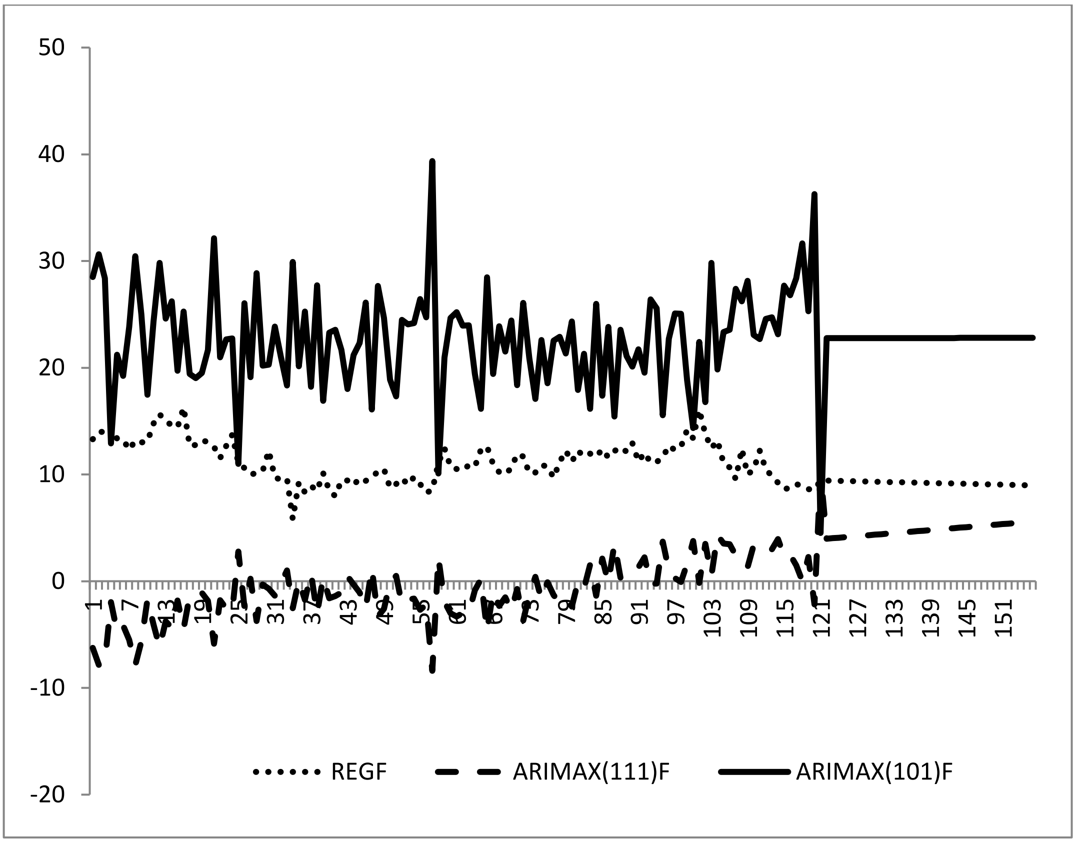


Fig. 4.49. Combined out of sample plot of arimax (1, 0, 1) model assuming lognormal error and multiple linear regression and arimax (1, 1, 1) with normal error term

CHAPTER FIVE

SUMMARY AND CONCLUSION

5.1 Summary of findings

The developed arimax assuming lognormal error term was capable of providing better model and improved forecasting ability of non-normal error and non-stationary economic time series without going through the task of traditional differencing of the series to achieve stationarity of time series before used in traditional models.

The performance of the proposed model was compared with that of arimax (1, 1, 1) assuming normal error term and multiple linear regression (mlr) model.

The error term of arimax (1, 0, 1) assuming normal error defined as:

$$\varepsilon_t = \frac{(1-\phi_1 B)y_t - \beta_0 - \beta_1 x_1}{(1+\theta_1)}$$

where the lag operator $B = y_{t-1}$; the parameter ϕ_1 represented the coefficient of the autoregressive model (ar), θ_1 was the coefficient of moving average (ma), β_0 was the intercept and β_1 was the slope of the regression part of the model.

The proposed non-normal error term was obtained as:

$$f(\varepsilon_t) = \frac{1}{y_t\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left|\frac{\text{In}\left[(1-\phi_1 B)y_t - \beta_0 - \beta_1 x_1\right] - \text{In}(1+\theta_1)}{\sigma}\right|^2\right.$$

while the loglikelihood function was derived as;

$$InL(\varepsilon_t) =$$
$$\frac{n}{2}In(2\pi\sigma^2) - \sum(Iny_t) - \frac{1}{2\sigma^2}\sum\left|\text{In}\left[y_t - y_{t-1} - \phi_1 y_{t-1} + \phi_1 y_{t-2} - \beta_0 - \beta_1 x_1\right] - \text{In}(1+\theta_1)\right|^2$$

The partial derivative of loglikelihood function equation with respect to each parameter and equate to zero, we obtained the parameter estimator.

The four economic time series were found to be non-stationary and non-normally distributed. The Loglik values of mlr, conventional arimax(1, 1, 1) with normal error and proposed arimax(1, 0, 1) with lognormal error term were -317.41, -240.23 and 1344.47; AIC values were 5.36, 490.45 and -0.41; while MSFE values were 12.41, 12.48 and 1.77.

The proposed model had the highest Loglik value, smallest AIC and smallest msfe values when compared with conventional arimax(1, 1, 1) with normal error and mlr model. Hence, the proposed model was considered best.

5.2 Conclusion

The proposed arimax with lognormal error term has ameliorated the task of differencing to achieve stationarity, stabilization of economic time series variables to attained normality before modeling. If considered, the proposed arimax (1, 0, 1) with lognormal error term was superior compared to traditional arimax (1, 1, 1) model with normal error and multiple linear regression in terms modeling process, validity and accurate forecast ability of economic time series data.

The study recommended the proposed arimax with lognormal error to be conveniently applied in modeling economic time series data that violated normality and stationarity assumptions without undergoing independent stabilization of the series. That saves time and cost implication. The model proved to be superior to the time series models of arimax with normal error and multiple linear regression model.

The study had contributed to knowledge: Conventional arimax with normal error indicated been not appropriate in modeling non-normal error and non-stationary economic time series data due to stringent assumptions of normal error and stationary series. The process of differencing at first, second or higher order to achieve stationary series lead to loss of vital information of time series data originality.

The study had proposed lognormal error innovation of conventional arimax (p, d, q) model and it estimation properties. The proposed lognormal error was incorporated into arimax (1, 0, 1) model. Whereby bridging the gap of differencing economic time series data that violated normality of error term and stationarity of series. By that proposed model, information loss by differencing the original time series data were captured whereby improving forecast capability.

REFERENCES

Arulampalam, M.S., Maskell, S., Gordon, N. and Clapp, T. 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2), 174–188.

Balabin, R. M. and Lomakina, E. I. 2011. Support vector machine regression (SVR/LSSVM) – an alternative to neural networks (ANN) for analytical chemistry, Comparison of nonlinear methods on near infrared (NIR) spectroscopy data. *Analyst*, 136(8), 1703–1712.

Barreto, G., Hammer B. and Hitzler, P. 2007. *Time series prediction with the self organizing map: a review perspectives of neural-symbolic integration, in perspectives of neural-symbolic integration*. Hammer B. and Hitzler, P. Eds. Springer: Berlin, pp. 135–158.

Bollerslev, T. 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.

Box, G. E. P, Jekins G. M. & Reinsel G. C. 1996. *Time series analysis forecasting and control.* Prentice-Hall International, Inc. Upper Saddle River, New Jersey 07458.

Box, G. E. P., and Jenkins, G. M. 1976. *Time series analysis, forecasting and control*, San Francisco, Holden-Day, California, USA.

Bruce, H. Andrews, Matthew D. Dean, Robert Swain and Caroline Cole 2013. Building arima and arimax models for predicting long-term disability benefit application rates in the public/private sectors. © 2013 *Society of Actuaries*.

Cao, L. 2003. Support vector machines experts for time series forecasting. *Neurocomputing*, 51, 321–339.

Central Bank of Nigeria, 2014. The 93rd monetary policy committee communiqué. *CBN Bulletin* January 20 – 21 2014.

Central Bank of Nigeria 2017. Crude oil export. *Statistical Bulletin*, Vol. 28 December 2017

Chang Pei-Chann, Fan Chin-Yuan, Chen Shih-Hsin 2007. Financial time series data forecasting by Wavelet and TSK fuzzy rule based system. *Fourth International Conference on Fuzzy Systems and knowledge Discovery,* Institute of Electrical and Electronics Engineers.

Chen, S., Billings, S.A. and Grant, P.M. 1992. Recursive hybrid algorithm for nonlinear system identification using radial basis function networks. *International Journal of Control*, 55(5), 1051–1070.

Chen, S., Cowan, C. F. N. and Grant, P. M. 1991. Orthogonal least squares learning algorithm for radial basis function networks. *IEEE Transactions on Neural Networks*, 2(2), 302–309.

Cheng Ching-hsue, Chen Tai-Liang, Teoh Hia-jong 2007. Multiple-period modified fuzzy time series for forecasting. *Institute of Electrical and Electronics Engineers, Fourth International Conference on Fuzzy systems and Knowledge Discovery*, 0-7695-2874.

Chingnun Lee 2006. *Maximum likelihood estimation*. Chp 17 mle of a Gaussian ar(1) process. www.scribd.com July, 2019.

Damodar N. Gujarati, Dawn C. Porter 2009. *Basic econometrics.* International edition published by McGraw-Hill/Irwin, a business unit of The McGraw-Hill Companies, Inc. 1221 Avenue of the Americas, New York, NY 10020.

De Groot, C. and Wuertz, D. 1991. Analysis of univariate time series with connectionist nets: a case study of two classical examples. *Neurocomputing,* 3(4), 177– 192.

Engle, R. F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50(4), 987–1007.

Fan, J. and Yao, Q. 2003. *Nonlinear Time series: nonparametric and parametric methods.* Springer-Verlag, New York, NY.

Fu-yuan Huang 2008. Forecasting stock price using a genetic fuzzy neural network. *Institute of Electrical and Electronics Engineers, International Conference in Computer Science and Information Technology*. 978-0-7695-3308.

Fu-Yuan Huang 2008. Integration of an improved particle swarm optimization algorithm and fuzzy neural network for shanghai stock market prediction. *IEEE workshop on power electronics and intelligent transportation system*, 978-07695-3342.

Grudnitski, G. and Osburn, L. 1993. Forecasting S and P and gold futures prices: an application of neural networks. *Journal of futures markets.* 13(6), 631–643.

Hubbard R. G. & O'Brien A. P. 2009. *Macroeconomics* 2nd Edition, Pearson Education, Inc, New Jersey, 07458.

Jarque, Carlos M. Bera, Anil K. 1980. Efficient test for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3) 255 – 259.

Jimoh Nimot Temitope 2017. On the performance of arima and arimax modelling technique. A project submitted to the Department of Statistics, Faculty of Science, University of Ibadan. xiv+ 180p.

Kodogiannis, V. and Lolis, A. 2002. Forecasting financial time series using neural network and fuzzy system-based techniques. *Neural Computing and Applications*, 11(2), 90–102.

Krishnamurthy, V. and Yin, G.G. 2002. Recursive algorithms for estimation of hidden Markov models and autoregressive models with Markov regime. *IEEE Transactions on Information Theory*, 48(2), 458–476.

Kuan, C.M. and Liu, T. (1995) Forecasting exchange rates using feed forward and recurrent neural networks. *Journal of Applied Econometrics*, 10(4), 347–364.

Kuo I-Hong, Horng Shi-Jinn, Chen Yuan-Hsin, Run Ray-Shine, Kao Tzong-Wann, Chen Rong-Jian, Lai Jui-Lin, Lin Tsung-Lieu 2009. Forecasting TAIFEX based on fuzzy time series and particle swarm optimization. *Expert System with Applications*.

Lapedes, A. and Farber, R. 1987. Nonlinear signal processing using neural networks: prediction and system modeling. Report, *Los Alamos Nationary Laboratory, Los Alamos, NM*.

Lee S. M. and Roberts S. J. 2008. Multivariate Time Series Forecasting in Incomplete Environments. *Technical Report* PARG 08 – 03, Robotics Research Group, Dept. of Engineering Science University of Oxford.

Lineesh, M.C. and John, C. J. 2010. Analysis of non-stationary time series using wavelet decomposition. *Nature and Science*, 8(1), 53–59.

Nigerian Stock Exchange Market 2017. *Official exchange rate*. Retrieved August 14, 2018, from https://en.wikipedia.org/wiki/Nigerian_Stock_Exchange/

Madsen H. 2008. *Time series analysis.* Chapman and Hall/CRC, London.

Meese R. and J. Geweke 1982. A comparison of autoregressive univaraite forecasting procedures for macroeconomic time series. University of California, Berkeley, CA, USA.

Menezes, J. M. P. and Barreto, G. A. 2008. Long-term time series prediction with the NARX network: an empirical evaluation. *Neurocomputing*, 71(16-18), 3335–3343.

Michael J. Alenyi *2016.* Exchange rate and external reserves in nigeria: a threshold cointegration analysis. *CBN Journal of Applied Statistics Vol. 7 No. 1(b) (June, 2016)* 233

Mohammed Amir Hamjah 2014. Temperature and rainfall effects on spice crops production and forecasting the production in Bangladesh: an application of Box-Jenkins arimax model mathematical theory and modeling. Retrieved from www.iiste.org ISSN 2224-5804 (Paper) ISSN 2225-0522 (Online) Vol.4, No.10, 2014.

Ngozi E. Nwachukwu, Abdulkadir I. Ali, Ismaila S. Abdullahi, Mohammed A. Shettima, Solomon S. Zirra, Bola S. Falade, and Osabuohien-Irabor

Osarumwense 2013. Applicability of Box Jenkins sarima model in rainfall forecasting: A Case Study of Port-Harcourt South South Nigeria, Canadian *Journal on Computing in Mathematics, Natural Sciences, Engineering and Medicine* Vol. 4 No. 1, Feb. 2013.

Olubusoye O. E. 2002. Advanced econometric methods. STA 767 lecture note. Department of Statistics, University of Ibadan, Nigeria.

Park, J. and Sandberg, I. W. 1991. Universal approximation using radial-basis-function networks. *Neural Computation*, 3(2), 246–257.

Rabiner, L. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2), 257–286.

Ranjeeta Bisoi, P. K. Dash, P. K. Nayak 2014. A hybrid decision tree and new frequency filtering S-Transform for simultaneous power signal disturbance pattern recognition. *International Journal of Knowledge-based and Intelligent Engineering Systems (KES) Journal, IOS Press*, Vol.18, pp.229-245.

Ruey S. Tsay (2002). *Analysis of financial time series*. John Wiley & Sons, Inc.

Shangodoyin, D. K., Ojo, J. F. 2002. *Elements of time series analysis*. Rasmed Publications Ltd, Ibadan.

Sahib, I. O. and Umar, S. S. 2019. *Understanding time series analysis: concepts, theories & applications*. Chapter 11 pg 78 – 82. Ahmedo Print, Auchi, Edo State, Nigeria.

Shittu, O. I. and Yaya, O. S. 2011. *Introduction to time series analysis*. Fasco Printing Works, Mokola, Ibadan.

Shittu, O. I. and Yaya, O. S. 2016. *Introduction to time series analysis*. Baba Tunde Intercontinental Print, No. 35, Gbadebo Street, Mokola, Ibadan, Nigeria. Chapter 3 pg. 42 – 67, Chapter 9 pg. 181 - 193.

Slim, Chokri 2006. Neuro-fuzzy network based on extended kalman filtering for financial time series. *Proceeding of Whorl Academy of Science, Engineering and Technology*, Vol 15, ISSN 1307-6884.

Smola, A. and Scholkop, B. 2004. A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199–222.

Sofia Rachad, Benayard Nsiri and Bahloul Bensassi 2015. System identification of inventory system using arx and armax models. *International Journal of Control and Automation* Vol. 8 No. 12 (2015) pp.283 – 294.

Tong, H. 1990 *Non-Linear Time Series: A Dynamical System Approach*, Clarendon Press, Oxford, UK.

Wan, E. A. and Vander Merwe, R. (2000) The unscented kalman filter for nonlinear estimation. *The IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium*, IEEE, Lake Louise, Alberta, Canada, pp. 153–158.

Wang, Z., Liu, X., Liu, Y., Liang, J. and Vinciotti, V. 2009. An extended Kalman filtering approach to modeling nonlinear dynamic gene regulatory networks via short gene expression time series. *IEEE Transactions on Computational Biology Bioinformatics*, 6(3), 410–419.

Wei Xing Zheng 2002. On least-squares identification of armax models. School of Quantitative Methods and Mathematical Sciences University of Western Sydney Ponrith South DC NSW 1797, Australia IFAC 15th Triennial World Congress, Barcelona, Spain

Weron, R. and Misiorek, A. 2008. Forecasting spot electricity prices: a comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24(4), 744–763.

Yadav, P. K., Pope, P. F. and Paudyal, K. 1994. Threshold autoregressive modeling in finance: the price differences of equivalent assets. *Mathematical Finance*, 4(2), 205–221.

Yaya O. S. 2016. Volatility persistence in naira exchange rates: A Pre-and Post-Global Financial Crisis Analysis. *Journal of the Nigerian Statistical Association* Vol 28, 2016 pg. 44 - 58

Yee, P. and Haykin, S. 1999. A dynamic regularized radial basis function network for nonlinear, nonstationary time series prediction. *IEEE Transactions on Signal Processing*, 47(9), 2503–2521.

Yu Lixin and Zhang Yan Qing 2005. Evolutionary Fuzzy Neural Networks for Hybrid financial Prediction. *Institute of Electrical and Electronics Engineers Transaction on Systems Man and Cybernetics*, Vol 35 No 2.

Yule, G. U., 1926. Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, 89

Zhang, G. P. 2012. *Neural networks for time-series forecasting*, in Handbook of Natural Computing, Rozenberg, G., Back, T. and Kok, J. (eds), Springer, Berlin, Germany, pp. 461–477.

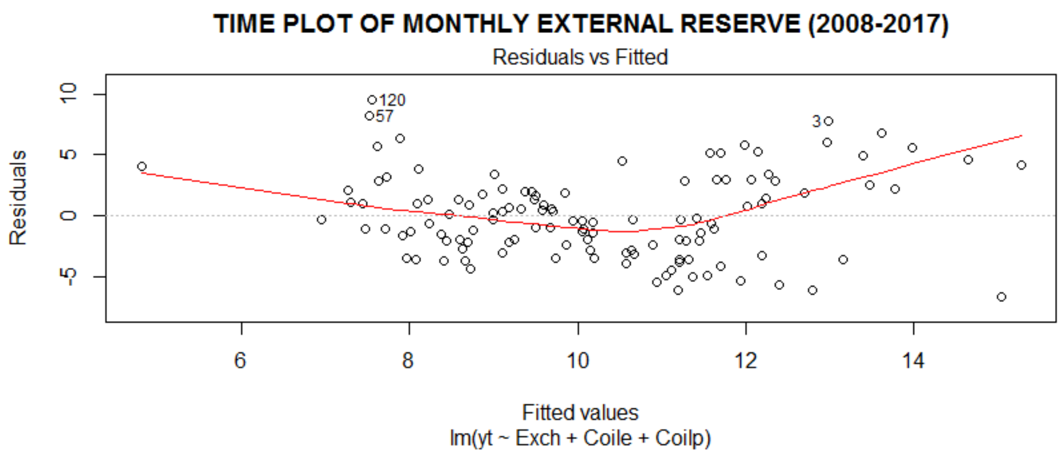
Zhang, G. P. and Berardi, V. L. 2001. Time series forecasting with neural network ensembles: an application for exchange rate prediction. *Journal of the Operational Research Society*, 52, 652–664.

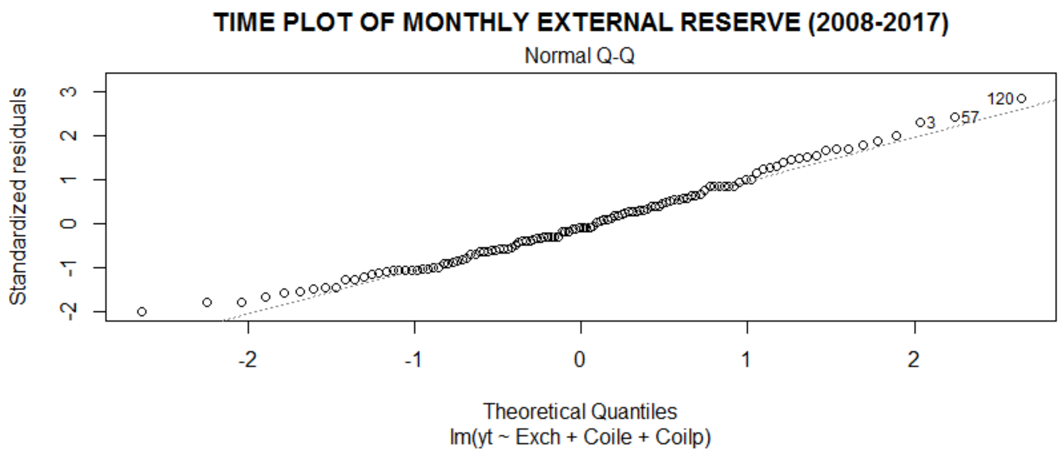
APPENDIX I

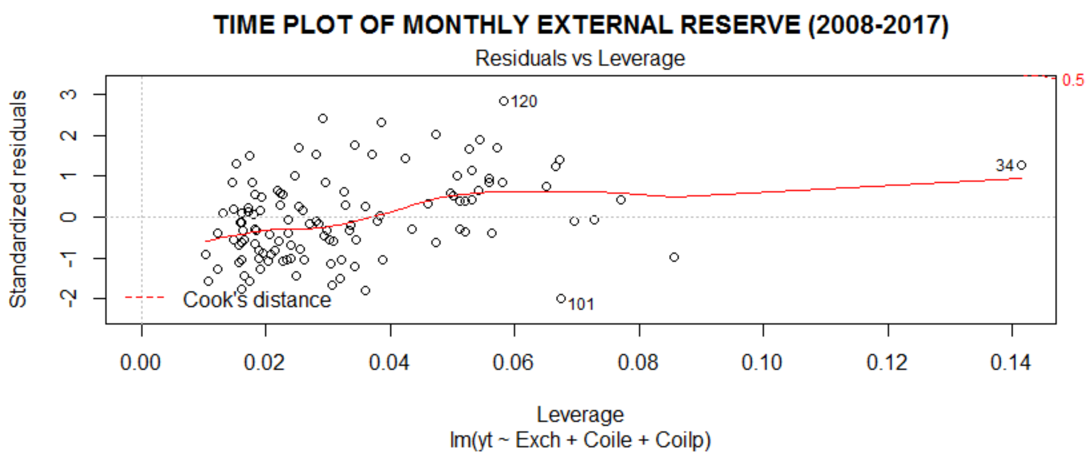
```
# R Code and Data Analyses Results of ARIMAX with Lognormal Error Term by Bello A. O. M. Phil P
roject Work#
> External=read.csv("Exteroil2.csv")
> attach(External)
> View(External)
> Month
  [1]   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15
 [16]  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
 [31]  31  32  33  34  35  36  37  38  39  40  41  42  43  44  45
 [46]  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
 [61]  61  62  63  64  65  66  67  68  69  70  71  72  73  74  75
 [76]  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90
 [91]  91  92  93  94  95  96  97  98  99 100 101 102 103 104 105
[106] 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
> Exch
  [1] 117.9768 118.2100 117.9218 117.8737 117.8342 117.8086
  [7] 117.7671 117.7420 117.7256 117.7243 117.7433 126.4756
 [13] 145.7803 147.1444 147.7226 147.2272 147.8427 148.2018
 [19] 148.5890 151.8580 152.3015 149.3550 150.8460 149.6928
 [25] 149.7792 150.2224 149.8285 149.8927 150.3125 150.1915
 [31] 150.0986 150.2667 151.0332 151.2500 150.2211 150.4799
 [37] 151.5455 151.9391 152.5071 157.3314 157.2762 157.4388
 [43] 157.4342 157.3796 157.3429 157.3156 157.3000 157.2742
 [49] 158.3868 157.8681 157.5875 157.3314 157.2762 157.4388
 [55] 157.4342 157.3796 157.3429 157.3156 157.3080 157.3240
 [61] 157.3012 157.2994 157.3115 157.3051 157.3008 157.3065
 [67] 157.3167 157.3136 157.3157 157.4166 157.2734 157.2742
 [73] 157.2916 157.3075 157.3008 157.2918 157.2873 157.2873
 [79] 157.2873 157.2873 157.3006 157.3141 159.9961 169.6800
 [85] 169.6800 179.7400 197.0700 197.0000 197.0000 196.9200
 [91] 196.9700 197.0000 197.0000 196.9900 196.9900 196.9900
 [97] 197.0000 197.0000 197.0000 197.0000 197.0000 231.7614
[103] 294.5722 309.7304 305.2250 305.2125 305.1818 305.2237
[109] 305.2024 305.3125 306.4022 306.0528 305.5381 305.7150
[115] 305.8619 305.6674 305.8868 305.6238 305.9045 306.3139
> Coile
  [1] 1.72 1.63 1.61 1.51 1.60 1.57 1.68 1.66 1.72 1.81 1.69 1.59
 [13] 1.58 1.61 1.62 1.41 1.77 1.72 1.69 1.67 1.73 1.83 1.70 1.60
 [25] 1.88 1.94 1.99 1.96 1.96 1.76 2.03 2.05 2.03 2.43 2.05 2.13
 [37] 2.04 2.06 1.84 1.97 2.05 1.89 1.89 1.96 1.87 1.91 1.87 1.82
 [49] 1.78 1.95 1.89 1.85 1.95 1.92 1.97 2.03 2.00 1.74 1.58 1.76
 [61] 1.78 1.78 1.75 1.79 1.61 1.58 1.75 1.84 1.84 1.78 1.64 1.66
 [73] 1.84 1.83 1.76 1.77 1.88 1.71 1.61 1.75 1.65 1.76 1.73 1.78
 [85] 1.75 1.76 1.62 1.58 1.60 1.52 1.73 1.67 1.77 1.76 1.73 1.63
 [97] 1.70 1.66 1.47 1.54 1.23 1.32 1.20 1.05 1.30 1.33 1.47 1.13
[109] 1.39 1.37 1.15 1.34 1.42 1.50 1.56 1.54 1.48 1.50 1.51 1.51
> Coilp
  [1]  94.26  98.15 103.73 116.73 126.57 138.74 137.74 115.84
  [9] 103.82  75.31  55.51  45.87  44.95  46.52  49.70  51.16
 [17]  60.02  72.24  66.52  74.00  70.22  78.25  78.11  75.11
 [25]  77.62  75.06  80.27  85.29  77.54  75.79  77.18  78.67
 [33]  79.45  84.42  86.71  92.79  97.96 106.57 116.56 124.49
 [41] 118.43 117.03 117.86 111.99 115.73 113.12 113.92 111.46
 [49] 113.81 121.87 128.00 122.62 113.08  98.06 104.62 113.76
 [57] 114.36 108.92 111.05 114.49 115.24 118.81 112.79 105.55
 [65] 106.00 106.06 109.78 107.84 113.59 112.29 111.14 112.75
 [73] 110.19 110.83 109.47 110.41 111.90 114.60 109.63 102.33
 [81]  98.27  83.50  80.42  63.28  48.81  58.09  56.69  57.45
```

```
 [89]  65.08  62.06  57.01  47.09  48.08  48.86  44.82  37.80
 [97]  30.66  31.70  37.76  41.59  47.01  48.46  45.25  46.15
[105]  47.43  50.94  45.25  53.48  55.01  46.39  52.13  52.94
[113]  50.57  47.42  49.01  51.64  56.79  58.46  63.56  65.11
> Extr
  [1] 15.7 19.0 20.8 16.0 15.2 13.7 14.1 17.4 17.8 15.0 16.0 19.3
 [13] 19.6 20.4 18.3 19.4 16.9 14.7 13.2 12.8 16.7 15.0 14.6 14.6
 [25]  9.0 10.5  8.7 11.3  9.9  9.3  9.6  8.6  6.7  8.9  7.6  8.4
 [37]  6.3  8.4  6.1  6.4  6.6  6.4  4.7  4.5  5.0  6.9  4.4  7.3
 [49]  8.5  6.6  4.5  5.9  6.7  7.6  9.1  9.4 15.8  9.6  9.4 10.3
 [61] 11.2 11.4 11.7 10.0  7.7 11.0  9.5  9.9  9.2 10.0  8.5 10.4
 [73]  9.4  7.0  7.5  6.2  6.5  7.3  7.6  8.8  7.4  7.5  5.5  7.8
 [85]  6.3  7.5  5.1  6.6  6.7  6.6  6.7  6.1  8.2  9.6  6.7  7.6
 [97]  9.2 10.6  9.6  6.7  8.4  6.7 10.1  8.9  8.8  8.7  9.9 11.2
[109] 12.4 11.4 10.9 10.8 10.6  9.6 10.5 10.9 12.0 14.2 13.3 17.1
> y=ts(External$X.1, frequency = 12,start = 2008)
Error in ts(External$X.1, frequency = 12, start = 2008) :
  'ts' object must have one or more observations
> summary(y)

Call:
lm(formula = yt ~ Exch + Coile + Coilp)

Residuals:
    Min      1Q  Median      3Q     Max
-6.6420 -2.4743 -0.3322  2.1511  9.5493

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 35.578734   4.227570   8.416 1.17e-13 ***
Exch        -0.040067   0.007906  -5.068 1.54e-06 ***
Coile       -9.074738   1.891591  -4.797 4.82e-06 ***
Coilp       -0.031519   0.013839  -2.278   0.0246 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.466 on 116 degrees of freedom
Multiple R-squared:  0.2423,	Adjusted R-squared:  0.2228
F-statistic: 12.37 on 3 and 116 DF,  p-value: 4.462e-07

> # plot series#
> plot(y, main="TIME PLOT OF MONTHLY EXTERNAL RESERVE (2008-2017)")
Hit <Return> to see next plot: adf.test(y)
```

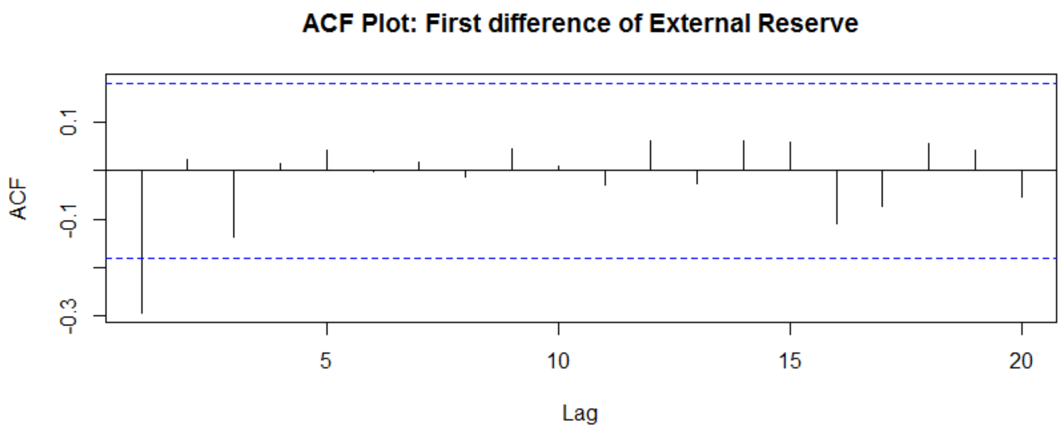
Hit <Return> to see next plot:

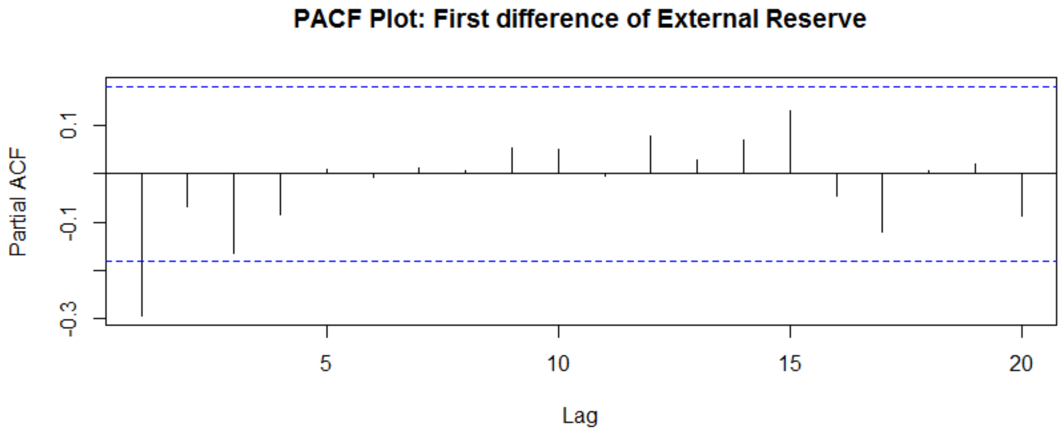

TIME PLOT OF MONTHLY EXTERNAL RESERVE (2008-2017)
Residuals vs Fitted
Residuals
Fitted values
lm(yt ~ Exch + Coile + Coilp)
120
57
3


TIME PLOT OF MONTHLY EXTERNAL RESERVE (2008-2017)
Normal Q-Q
Standardized residuals
Theoretical Quantiles
lm(yt ~ Exch + Coile + Coilp)
120
3
57

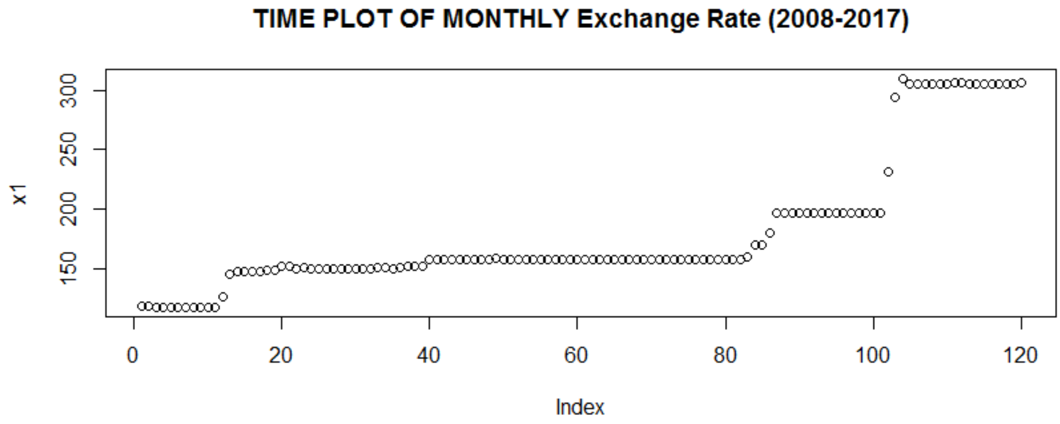

TIME PLOT OF MONTHLY EXTERNAL RESERVE (2008-2017)
Residuals vs Leverage
Standardized residuals
Leverage
lm(yt ~ Exch + Coile + Coilp)
Cook's distance
0.5
120
34
101

```
 > plot(y, main="TIME PLOT OF MONTHLY EXTERNAL RESERVE (2008-2017)")
Hit <Return> to see next plot: adf.test(y)
Hit <Return> to see next plot: adf.test(y)
Hit <Return> to see next plot: pacf(y)
Hit <Return> to see next plot: d=diff(y)
> adf.test(d, main = "First difference of External Reserve")
Error in adf.test(d, main = "First difference of External Reserve") :
  unused argument (main = "First difference of External Reserve")
> acf(d, main = "ACF Plot: First difference of External Reserve")
```






```
 > pacf(d, main = "PACF Plot: First difference of External Reserve")
> summary(d, main = "First difference of External Reserve")
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-6.20000 -1.40000  0.20000  0.01176  1.20000  6.40000
> Exch..xt.
Error: object 'Exch..xt.' not found
> x1=ts(External$X.2,frequency = 12,start = 2008)
Error in ts(External$X.2, frequency = 12, start = 2008) :
  'ts' object must have one or more observations
> plot(x1, main="TIME PLOT OF MONTHLY Exchange Rate (2008-2017)")
```

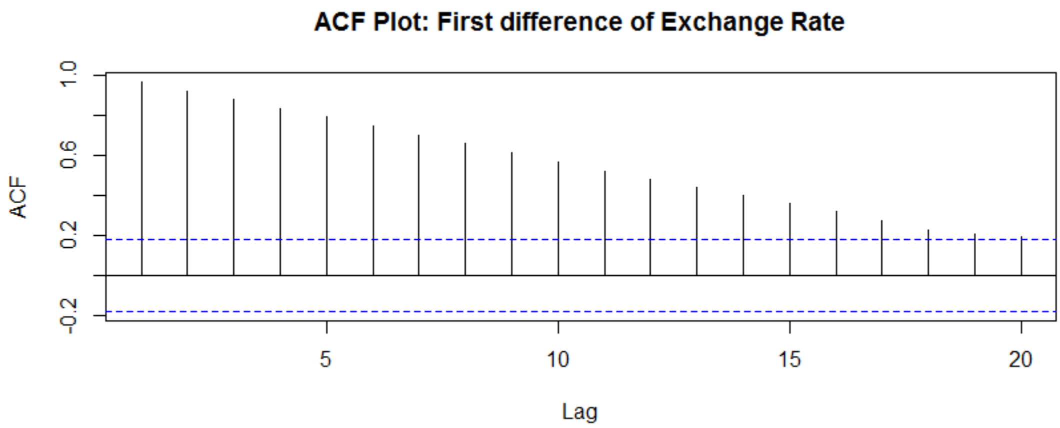


adf.test(x1)

Augmented Dickey-Fuller Test

data: x1
Dickey-Fuller = -1.1698, Lag order = 4, p-value = 0.9081
alternative hypothesis: stationary

> acf(x1, main = "ACF Plot: First difference of Exchange Rate")



```
> adf.test(x1)

        Augmented Dickey-Fuller Test

data:  x1
Dickey-Fuller = -1.1698, Lag order = 4, p-value = 0.9081
alternative hypothesis: stationary

> acf(x1, main = "ACF Plot: First difference of Exchange Rate")
> summary(x1)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  117.7   150.4   157.3   180.0   197.0   309.7
> xdifference=diff(x1)
> model2<-arimax(d,order=c(1,1,1),xreg=xdifference)
> model2

Call:
arimax(x = d, order = c(1, 1, 1), xreg = xdifference)

Coefficients:
          ar1      ma1    xreg
      -0.3081  -0.9936  0.0244
s.e.   0.0908   0.0675  0.0201

sigma^2 estimated as 3.359:  log likelihood = -241.29,  aic = 488.58
> model2<-arima(d,order=c(1,1,1))
> L=(238.81)
> log*(L)
Error in log * (L) : non-numeric argument to binary operator
> n=length(External.Reserve..Yt.)
Error: object 'External.Reserve..Yt.' not found
> k=log(n)
> p=4
> BIC=print(-2*log(L) + k*p)
[1] 8.19863
> ###########log normal model
> stock2 <- read.csv("arimax3 data.csv")
> attach(stock2)
The following object is masked from External:

    year

> yt=Extr
> x1=Exch
> x2=Coile
> x3=Coilp
> x1<-c(Exchange.Rate..xt.)
> x2<-c()
> yt<-c(External.Reserve..Yt.)
> yt1<-c(yt.1)
> yt2<-c(yt.2)
> llik<-function(sigma,theta1,phi1,beta0,beta1){
+   n=120
+   ll=(n/2)*log (2*pi*sigma^2)-sum(log(yt))-
+     (1/(2*sigma^2))*sum((abs(log(abs(yt-yt1-phi1*yt1+phi1*yt2-beta0-beta1*x1)
+                              - log(1+theta1)))^2))
+   return(-ll)
+ }
> y1<-mle2(minuslogl=llik,start=list(sigma=2.5,theta1=1.5,
+                               phi1=0.5,beta0=1.4,beta1=5.2))
```

```
> print(y1)

Call:
mle2(minuslogl = llik, start = list(sigma = 2.5, theta1 = 1.5,
    phi1 = 0.5, beta0 = 1.4, beta1 = 5.2))

Coefficients:
      sigma       theta1         phi1        beta0        beta1
36162.590999     1.642915     0.498075     1.464711    16.063444

Log-likelihood: 1099.07
> #### AIC
> L=1099.07
> log(L)
[1] 7.00222
> k=2
> p=4
> AIC=print(-2*log(L) + k*p)
[1] -6.004439
> #####BIC
> n=length(External.Reserve..Yt.)
> k=log(n)
> p=4
> BIC=print(-2*log(L) + k*p)
[1] 5.145528
> fit <- arimax(d,order=c(1,1,1),xreg=xdifference)
> #####   arimax with lognomal error term##
> External=read.csv("Exteroil2.csv")
> Extern=read.csv("arimax3 data.csv")
> attach(External)

> yt.1
  [1] 13.0 15.7 19.0 20.8 16.0 15.2 13.7 14.1 17.4 17.8 15.0 16.0
 [13] 19.3 19.6 20.4 18.3 19.4 16.9 14.7 13.2 12.8 16.7 15.0 14.6
 [25] 14.6  9.0 10.5  8.7 11.3  9.9  9.3  9.6  8.6  6.7  8.9  7.6
 [37]  8.4  6.3  8.4  6.1  6.4  6.6  6.4  4.7  4.5  5.0  6.9  4.4
 [49]  7.3  8.5  6.6  4.5  5.9  6.7  7.6  9.1  9.4 15.8  9.6  9.4
 [61] 10.3 11.2 11.4 11.7 10.0  7.7 11.0  9.5  9.9  9.2 10.0  8.5
 [73] 10.4  9.4  7.0  7.5  6.2  6.5  7.3  7.6  8.8  7.4  7.5  5.5
 [85]  7.8  6.3  7.5  5.1  6.6  6.7  6.6  6.7  6.1  8.2  9.6  6.7
 [97]  7.6  9.2 10.6  9.6  6.7  8.4  6.7 10.1  8.9  8.8  8.7  9.9
[109] 11.2 12.4 11.4 10.9 10.8 10.6  9.6 10.5 10.9 12.0 14.2 13.3
> yt=Extr
> length(yt)
[1] 120
> x1=Exch
> x2=Coile
> x3=Coilp
> yt1<-c(yt.1)
> yt2<-c(yt.2)
> llik<-function(sigma,theta1,phi1,beta0,beta1,beta2,beta3){
+   n=120
+   ll=(n/2)*log (2*pi*sigma^2)-sum(log(yt))-
+     (1/(2*sigma^2))*sum((abs(log(abs(yt-yt1-phi1*yt1+phi1*yt2-beta0-beta1*x1-beta2
*x2-beta3*x3)
+                          - log(1+theta1)))^2))
+   return(-ll)
+ }
> y1<-mle2(minuslogl=llik,start=list(sigma=2.5,theta1=1.5,
+                            phi1=0.5,beta0=2.4,beta1=5.2,beta2=2.4,beta3=6.4))
```

```
> print(y1)

Call:
mle2(minuslogl = llik, start = list(sigma = 2.5, theta1 = 1.5,
    phi1 = 0.5, beta0 = 2.4, beta1 = 5.2, beta2 = 2.4, beta3 = 6.4))

Coefficients:
       sigma       theta1         phi1        beta0        beta1
2.794984e+05 2.406881e+00 4.251954e-01 4.507027e+00 3.618276e+02
       beta2        beta3
6.053757e+00 1.888530e+02

Log-likelihood: 1344.47
> ## Multiple Regression##
> y=lm(yt~Exch+Coile+Coilp)
> y

Call:
lm(formula = yt ~ Exch + Coile + Coilp)

Coefficients:
(Intercept)         Exch        Coile        Coilp
   35.57873     -0.04007     -9.07474     -0.03152

> summary(y)

Call:
lm(formula = yt ~ Exch + Coile + Coilp)



Residuals:
    Min      1Q  Median      3Q     Max
-6.6420 -2.4743 -0.3322  2.1511  9.5493

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 35.578734   4.227570   8.416 1.17e-13 ***
Exch        -0.040067   0.007906  -5.068 1.54e-06 ***
Coile       -9.074738   1.891591  -4.797 4.82e-06 ***
Coilp       -0.031519   0.013839  -2.278   0.0246 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.466 on 116 degrees of freedom
Multiple R-squared:  0.2423,        Adjusted R-squared:  0.2228
F-statistic: 12.37 on 3 and 116 DF,  p-value: 4.462e-07

> #### AIC
> L=1344.47
> log(L)
[1] 7.203755
> k=2
> p=7
> AIC=print(-2*log(L) + k*p)
[1] -0.4075103
> #####BIC
> n=length(yt)
> k=log(n)
> p=7
```

```
> BIC=print(-2*log(L) + k*p)
[1] 19.10493
> fit<- arima(yt,xreg=External[,(x1,x2,x3)],order=c(1,1,1),)
Error: unexpected ',' in "fit<- arima(yt,xreg=External[,(x1,"
> length(yt)
[1] 120
> View(yt)
```

APPENDIX II

R code

R Code and Data Analyses Results of ARIMAX with Lognormal Error Term by Bello A. O. M. Phil Project Work#

```
External=read.csv("Exteroil2.csv")
attach(External)
View(External)
Month
Exch
Coile
Coilp
Extr
y=ts(External$X.1, frequency = 12,start = 2008)
summary(y)
# plot series#
plot(y, main="TIME PLOT OF MONTHLY EXTERNAL RESERVE (2008-2017)")
adf.test(y)
acf(y)
pacf(y)
d=diff(y)
adf.test(d, main = "First difference of External Reserve")
acf(d, main = "ACF Plot: First difference of External Reserve")
pacf(d, main = "PACF Plot: First difference of External Reserve")
summary(d, main = "First difference of External Reserve")
Exch..xt.
x1=ts(External$X.2,frequency = 12,start = 2008)
plot(x1, main="TIME PLOT OF MONTHLY Exchange Rate (2008-2017)")
adf.test(x1)
acf(x1, main = "ACF Plot: First difference of Exchange Rate")
pacf(x1, main = "PACF Plot: First difference of Exchange Rate")
summary(x1)
xdifference=diff(x1)


model2<-arimax(d,order=c(1,1,1),xreg=xdifference)
model2
model2<-arima(d,order=c(1,1,1))

L=(238.81)
log*(L)
n=length(External.Reserve..Yt.)
k=log(n)
p=4
BIC=print(-2*log(L) + k*p)
```

```
##log normal model
stock2 <- read.csv("arimax3 data.csv")
attach(stock2)
yt=Extr
x1=Exch
x2=Coile
x3=Coilp
x1<-c(Exchange.Rate..xt.)
x2<-c()
yt<-c(External.Reserve..Yt.)
yt1<-c(yt.1)
yt2<-c(yt.2)
llik<-function(sigma,theta1,phi1,beta0,beta1){
  n=120
  ll=(n/2)*log (2*pi*sigma^2)-sum(log(yt))-
    (1/(2*sigma^2))*sum((abs(log(abs(yt-yt1-phi1*yt1+phi1*yt2-beta0-beta1*x1)
                              - log(1+theta1)))^2))
  return(-ll)
}
y1<-mle2(minuslogl=llik,start=list(sigma=2.5,theta1=1.5,
                               phi1=0.5,beta0=1.4,beta1=5.2))
print(y1)
#### AIC
L=1099.07
log(L)
k=2
p=4
AIC=print(-2*log(L) + k*p)
#####BIC
n=length(External.Reserve..Yt.)
k=log(n)
p=4
BIC=print(-2*log(L) + k*p)

fit <- arimax(d,order=c(1,1,1),xreg=xdifference)
#####   arimax with lognomal error term of 3x variables##
External=read.csv("Exteroil2.csv")
Extern=read.csv("arimax3 data.csv")
attach(External)
attach(Extern)
yt.1

yt=Extr
length(yt)
x1=Exch
x2=Coile
x3=Coilp
yt1<-c(yt.1)
yt2<-c(yt.2)
llik<-function(sigma,theta1,phi1,beta0,beta1,beta2,beta3){
  n=120
  ll=(n/2)*log (2*pi*sigma^2)-sum(log(yt))-
    (1/(2*sigma^2))*sum((abs(log(abs(yt-yt1-phi1*yt1+phi1*yt2-beta0-beta1*x1-beta2*x2-
beta3*x3)
                              - log(1+theta1)))^2))
```

```
  return(-ll)
}
y1<-mle2(minuslogl=llik,start=list(sigma=2.5,theta1=1.5,
                        phi1=0.5,beta0=2.4,beta1=5.2,beta2=2.4,beta3=6.4))
print(y1)

## Multiple Regression##
y=lm(yt~Exch+Coile+Coilp)
y
summary(y)

#### AIC
L=1344.47
log(L)
k=2
p=7
AIC=print(-2*log(L) + k*p)
#####BIC
n=length(yt)
k=log(n)
p=7
BIC=print(-2*log(L) + k*p)

fit<- arima(yt,xreg=External[,(x1,x2,x3)],order=c(1,1,1),)


length(yt)
View(yt)
```

```
# R Code of Data Analyses on#
# Topic: Development of ARIMAX with Lognormal Error Term#
# by Bello A. O. #
# Mat. No. 113238#
# Program: M. Phil (Time Series) Dissertation #
# Initialize variables#
External=read.csv("Exteroil2.csv")
attach(External)
View(External)
Month
Exch
Coile
Coilp
Extr
y=ts(External$X.1, frequency = 12, start = 2008)
summary(y)
# plot series#
plot(y, main="TIME PLOT OF MONTHLY EXTERNAL RESERVE (2008-2017)")
adf.test(y)
acf(y)
pacf(y)
d=diff(y)
adf.test(d, main = "First difference of External Reserve")
```

```
acf(d, main = "ACF Plot: First difference of External Reserve")
pacf(d, main = "PACF Plot: First difference of External Reserve")
summary(d, main = "First difference of External Reserve")
Exch..xt.
x1=ts(External$X.2,frequency = 12,start = 2008)
plot(x1, main="TIME PLOT OF MONTHLY Exchange Rate (2008-2017)")
adf.test(x1)
acf(x1, main = "ACF Plot: First difference of Exchange Rate")
pacf(x1, main = "PACF Plot: First difference of Exchange Rate")
summary(x1)
xdifference=diff(x1)


model2<-arimax(d,order=c(1,1,1),xreg=xdifference)
model2
length(d)
model2<-arima(d,order=c(1,1,1))
a=forecast(model2,30)a
plot(a)
L=(238.81)
log*(L)
n=length(External.Reserve..Yt.)
k=log(n)
p=4
BIC=print(-2*log(L) + k*p)


#Development of ARIMAX model with lognormal error#
stock2 <- read.csv("arimax3 data.csv")
attach(stock2)
yt=Extr
x1=Exch
x2=Coile
x3=Coilp
x1<-c(Exchange.Rate..xt.)
x2<-c()
yt<-c(External.Reserve..Yt.)
yt1<-c(yt.1)
yt2<-c(yt.2)
llik<-function(sigma,theta1,phi1,beta0,beta1){
  n=120
  ll=(n/2)*log (2*pi*sigma^2)-sum(log(yt))-
    (1/(2*sigma^2))*sum((abs(log(abs(yt-yt1-phi1*yt1+phi1*yt2-beta0-beta1*x1)
                              - log(1+theta1)))^2))
  return(-ll)
}
y1<-mle2(minuslogl=llik,start=list(sigma=2.5,theta1=1.5,
                               phi1=0.5,beta0=1.4,beta1=5.2))
print(y1)

#AIC#
L=1099.07
log(L)
k=2
p=4
```

```
AIC=print(-2*log(L) + k*p)

#BIC#
n=length(External.Reserve..Yt.)
k=log(n)
p=4
BIC=print(-2*log(L) + k*p)

# arimax with normal error term of 3x variables#
fit <- arimax(d,order=c(1,1,1),xreg=xdifference)

# arimax with lognomal error term of 3x variables#
External=read.csv("Exteroil2.csv")
Extern=read.csv("arimax3 data.csv")
attach(External)
attach(Extern)
yt.1

yt=Extr
length(yt)
x1=Exch
x2=Coile
x3=Coilp
yt1<-c(yt.1)
yt2<-c(yt.2)

# Likelihood Function of arimax with lognomal error term of 3x variables#
llik<-function(sigma,theta1,phi1,beta0,beta1,beta2,beta3){
  n=120
  ll=(n/2)*log (2*pi*sigma^2)-sum(log(yt))-
    (1/(2*sigma^2))*sum((abs(log(abs(yt-yt1-phi1*yt1+phi1*yt2-beta0-beta1*x1-beta2*x2-
beta3*x3)
                    - log(1+theta1)))^2))
  return(-ll)
}
y1<-mle2(minuslogl=llik,start=list(sigma=2.5,theta1=1.5,
                    phi1=0.5,beta0=2.4,beta1=5.2,beta2=2.4,beta3=6.4))
print(y1)

#Fitted  ARIMAX(1,1,1) MODEL WITH X1, X2, X3#
yt=-0.3127*yt1-0.9941*yt2+0.0238*x1+0.2725*x2-0.0296*x3

#Fitted of Developed ARIMAX(1,0,1) model with lognormal error term of  X1, X2 and X3#
yt2=0.4252*yt1+2.4068*yt2+0.0362*x1+6.0537*x2+0.0189*x3

#Forecast of ARIMAX(1,1,1) MODEL WITH X1, X2, X3#
forecast<-forecast(yt, h = 30)## predict 30 months
plot(forecast)

#Forecast of Developed ARIMAX(1,0,1) model with lognormal error term of  X1, X2 and X3#
forecast<-forecast(yt2, h = 30)## predict 30 months
plot(forecast)

# Multiple Linear Regression with 3 independent variables X1, X2 and X3#
y=lm(yt~Exch+Coile+Coilp)
```

```
y
summary(y)
# forecast #
forecast<-forecast(yt, h = 30)## predict 30 yrs
plot(forecast)

# AIC #
L=1344.47
log(L)
k=2
p=7
AIC=print(-2*log(L) + k*p)

d=cbind(x1,x2,x3)
fit<- arima(yt,xreg=d,order=c(1,1,1),)
fit
## forecast
forecast<-forecast(fit, h = 30)## predict 30 yrs
plot(forecast)
```