Buys-Ballot Modeling of Nigerian Domestic Crude Oil Production (2006 – 2012)

Okororie Chigozie[1], Egwim Kenneth C[2], Eke Charles N.[2], Onuoha Desmond O.[2]

[1]*Department of Statistics, Imo State University, Owerri, Nigeria*
[2]*Department of Statistics, Federal Polytechnic Nekede, Owerri, Nigeria*
E-mail: goziestat21@yahoo.com and desonuoha@yahoo.com

Abstract

This paper has examined time series analysis of Nigeria domestic crude oil production, the descriptive approach of time series analysis. Buys-Ballot table procedure was used in assessing variance stability (transformation), choice of model and seasonal effect. Inverse square root transformation was carried to stabilize the variance, quadratic trend was fitted and the error component was found to be random and normally distributed with mean zero and some constant variance i.e. $e_t \sim N(0, 0.022270)$.

Keywords: *Buys-Ballot Table, Data Transformation, Periodic Averages, Choice of Model, Trend Assessment, Seasonal Indices*

1.0 Introduction

The general time series is a mixture of four components [7]. The components are;

(i) The trend or long term movement of the series.
(ii) The seasonal component
(iii) The cyclical component or fluctuations about the trend
(iv) The irregular, random or error component.

The trend and cyclical components could be combined to give the trend-cycle component especially in short time series [7]

A time series is a sequence of observations made usually (but not necessarily) at equal interval of time. Examples of time series are records of daily temperature, records of monthly births, average monthly rainfall, number of barrels per month of crude oil export, etc., there are three main models which are used in descriptive time series analysis. They are the multiplicative, additive and the mixed models. The models are expressed mathematically as;

Multiplicative:

$$X_t = T_t * S_t * C_t * I_t$$

$$X_t = M_t * S_t * I_t \qquad 1$$

Additive:

$$X_t = T_t + S_t + C_t + I_t$$

$$X_t = M_t + S_t + I_t \qquad 2$$

Mixed:

$$X_t = T_t * S_t * C_t + I_t$$
$$X_t = M_t * S_t + I_t \qquad 3$$

Where T_t is the trend over time t, S_t is the seasonal variation over time t, C_t is the cyclical variation over time t, I_t is the irregular or error variation over time t and M_t is the trend-cycle (when short periods are observed) variation over time t.

In statistics, data transformations refer to the application of a deterministic mathematical function to each point in a data set – that is each data point z_i replaced with the transformed value $y_i = f(z_i)$, where f is a function. Transforms are usually applied so that the data appear to more closely meet the assumptions of a statistical inference procedure that is to be applied, or to improve the interpretability or appearance of graphs.

The criterion on how data should be transformed or whether transformation is applicable at all should emanate from the particular statistical analysis to be carried out.

Generally there are a number of reasons for data transformation like easy visualization, improvement in interpretation, variance stabilization, ensuring a normally distributed data, additivity of the seasonal effect. However, for time series data there are three major reasons namely;

(a) For easy visualization.
(b) To improve interpretability, even if no formal statistical analysis or visualization is to be performed.
(c) When there is evidence of substantial skew in the data, it is common to transform the data to a symmetric distribution before constructing a confidence interval.
(d) Another reason is to stabilize the variance because we expect the variance to be same for each possible expected value (this is known as Homoscedasticity).

Below are some of the common transformations [1].

Table 1: Bartlett's transformation for some values of β

S/No	1	2	3	4	5	6	7
Values of β	0	$\frac{1}{2}$	1	$\frac{3}{2}$	2	3	-1
Transformation	No transformation	$\sqrt{X}$	Log$_e$ X_t	$\frac{1}{\sqrt{X}}$	$\frac{1}{X_t}$	$\frac{1}{X_t^2}$	X_t^2

Source: Iwueze et al (2009).

Iwueze and Akpanta [6] has shown below that table 1 above could be regarded as a power function where

$$Y_t = \begin{cases} \log_e X_t; & \beta = 1 \\ X^{1-\beta}; & \beta \neq 1 \end{cases} \qquad (4)$$

In this paper, we shall explore the appropirate transformation that will be suitable for analysis modelling and analysis Nigeria crude oil domestic production to ensure that the met some appropriate inference procedure.

In the analysis of this time series data under the descriptive approach, we would determine;

1. The appropriate transformation to be adopted (if applicable).
2. The appropriate choice of the model

to be adopted i.e. multiplicative, additive or mixed model using Iwueze and Nworgu [6].

3. The trend pattern of domestic crude oil production.
4. The probability of predicting the production of crude oil domestically in 2013.

2. Choice of Appropriate Transformation

The first stage in analyzing a time series data is to plot a graph of the values or observations against time [3].

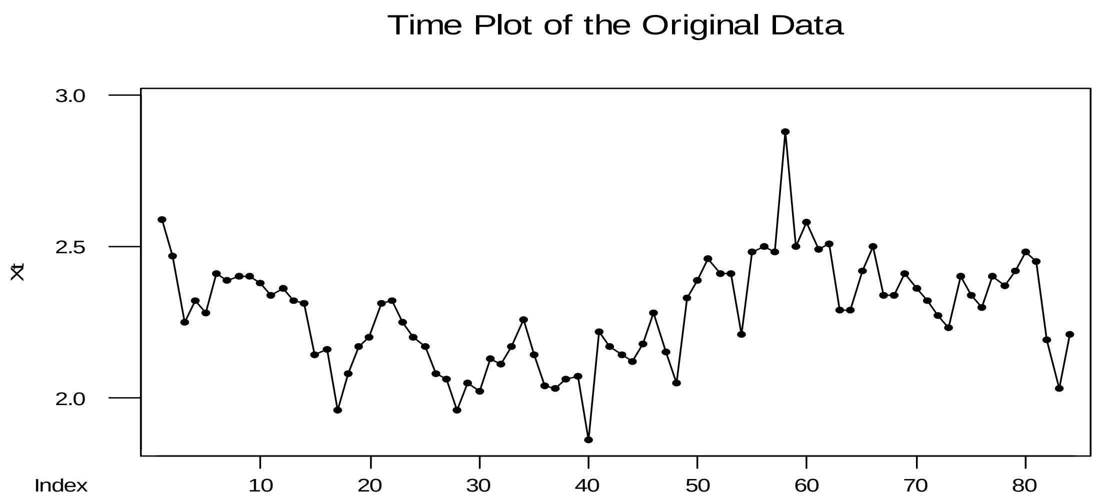

Fig. 1: Buys-Ballot table of the data is shown below

Table 2: Buys-Ballot table

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2006	2.590	2.470	2.250	2.320	2.280	2.410	2.390	2.400	2.400	2.380	2.340	2.360
2007	2.320	2.310	2.140	2.160	1.960	2.080	2.170	2.200	2.310	2.320	2.250	2.200
2008	2.170	2.080	2.060	1.960	2.050	2.020	2.130	2.110	2.170	2.260	2.140	2.040
2009	2.030	2.060	2.070	1.860	2.220	2.170	2.140	2.120	2.180	2.280	2.150	2.050
2010	2.330	2.390	2.440	2.410	2.410	2.210	2.480	2.500	2.480	2.880	2.500	2.580
2011	2.490	2.510	2.290	2.290	2.420	2.500	2.340	2.340	2.410	2.360	2.320	2.270
2012	2.230	2.400	2.340	2.300	2.400	2.370	2.420	2.480	2.450	2.190	2.030	2.210
Mean	2.309	2.317	2.227	2.186	2.249	2.251	2.296	2.307	2.343	2.381	2.247	2.244
Std Dev	0.176	0.167	0.143	0.204	0.184	0.179	0.146	0.164	0.126	0.229	0.156	0.187

Source: www.cbn.gov.ng/rates.

Year	Ave	Std dev	Log Ave	Log Std Dev
2006	2.382	0.089	0.868	-2.423
2007	2.202	0.110	0.789	-2.206
2008	2.099	0.081	0.742	-2.514
2009	2.111	0.108	0.747	-2.224
2010	2.467	0.161	0.903	-1.829
2011	2.378	0.086	0.866	-2.453
2012	2.318	0.131	0.841	-2.030

Bartlett [1] have structured transformation guide with respect to varying values of $\hat{\beta}$ otherwise known as the slope of regression line of natural logarithm of the periodic standardard deviations ($\log\left(\hat{\sigma}\right)$) against the corresponding means ($\log\left(\bar{x}\right)$)

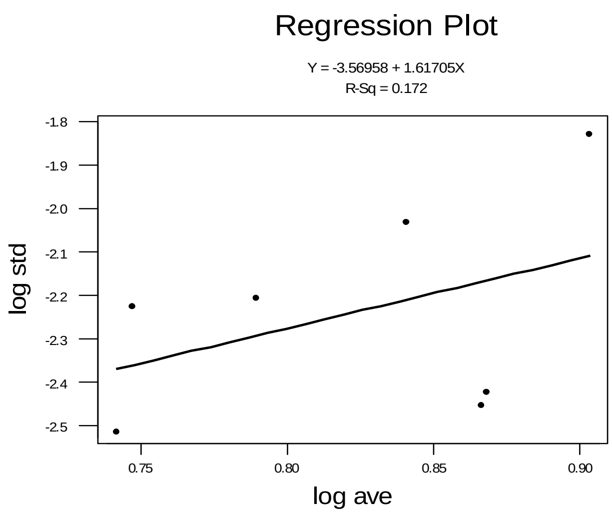

Fig. 2

The regression equation is y = - 3.57 + 1.62x

Predictor Coef StDev T P
Constant -3.570 1.310 -2.73 0.042
x 1.617 1.589 1.02 0.035

S = 0.2472 R-Sq = 17.2% R-Sq(adj) = 0.6%

In other to determine the appropriate transformation, the periodic (annual) means and standard deviations only will be used. The slope of the regression equation of natural logarithm of the periodic mean standard deviation $\left(\log\hat{\sigma}\right)$ on the logarithm of the periodic means $\left(\log_e \bar{y}\right)$ of the study data in table 4.1 was found to be

$\hat{\beta} = 1.62$ with the standard error 1.589 and coefficient of determination $R^2 = 0.172$.

Test for Significance of $\hat{\beta}$

Since the slope of the regression line is not exactly 1.5 but lies between 1.5 and 2. Since 1.62 is closer to 1.5 than to 2, we shall test for the significance of 1.5 using t - test. The test procedure is given below

$$H_0 : \hat{\beta} = 1.5 \quad vs \quad H_1 : \hat{\beta} \neq 1.5$$

$$\alpha = 0.05, \quad \alpha/2 = 0.025; \quad 1 - \alpha/2 = 0.975$$

Degrees of freedom, df = n-1 = 7-1 = 6.
Decision Rule: Reject H_0 iff
$t_{cal} \geq t_{tab}, \textit{otherwise accept}\ H_0$. Therefore the tabulated values

$$t_{tab} = t_{1-\alpha/2, 9} = t_{0.975, 9} = \pm\ 2.75$$

$$\frac{\hat{\beta} - \beta}{stderror of\ \hat{\beta}} = \frac{1.62 - 1.5}{1.589} \quad \Rightarrow t_{cal} = 0.0755$$

Decision: since t_{cal} = 0.0755 falls inside ±2.57 i.e. the acceptance region, we accept H_0 and conclude that $\hat{\beta}$ = 1.62 is significantly and statistically not different from 1.5 at 5% level of significance. Therefore $\hat{\beta}$ corresponds to inverse square root ($\frac{1}{\sqrt{X_t}}$) transformation [1]. (Refer to Table 1).

Time plot of the transformed data is shown below in figure 3.

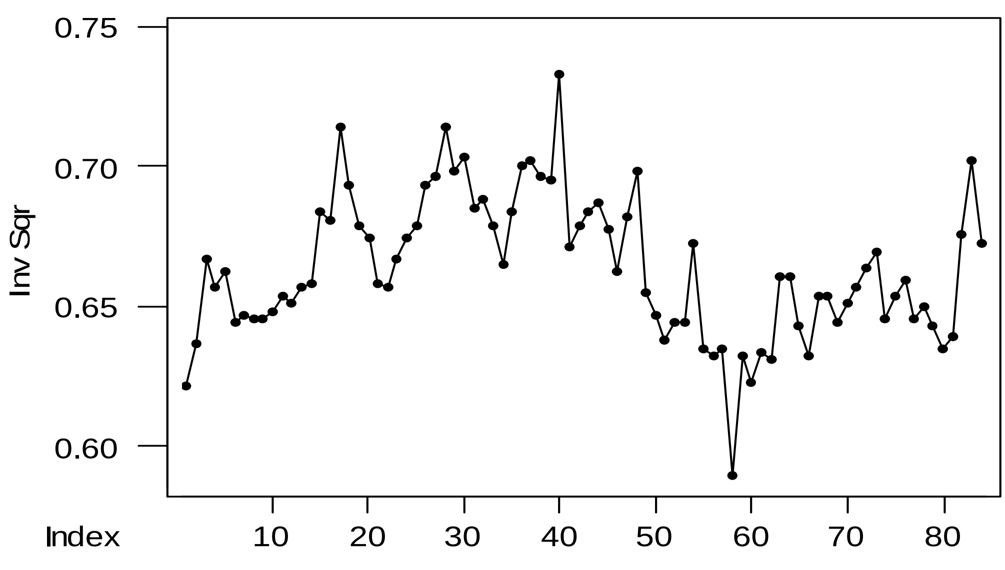

Fig. 3: Time Plot of the Inverse Square Root Transformed Data

3. Choice of Appropriate Model

Using Buys-Ballot table, the relationship between the periodic means $\left(\bar{X}_{.j}, j = 1,2,\cdots,n \right)$ and the periodic standard deviations $\left(\hat{\sigma}_{.j}, j = 1,2,\cdots,n \right)$, where n is the number of years, gives an indication of the desired model. An additive model is appropriate when the seasonal standard deviations show no appreciable ; increase/decrease relative to any increase or decrease in the seasonal means. On the other hand, a multiplicative model is usually appropriate when the seasonal standard deviations show appreciable increase/decrease relative to any increase or decrease in the seasonal means [5][6].

This is explicitly shown in figure 4 below

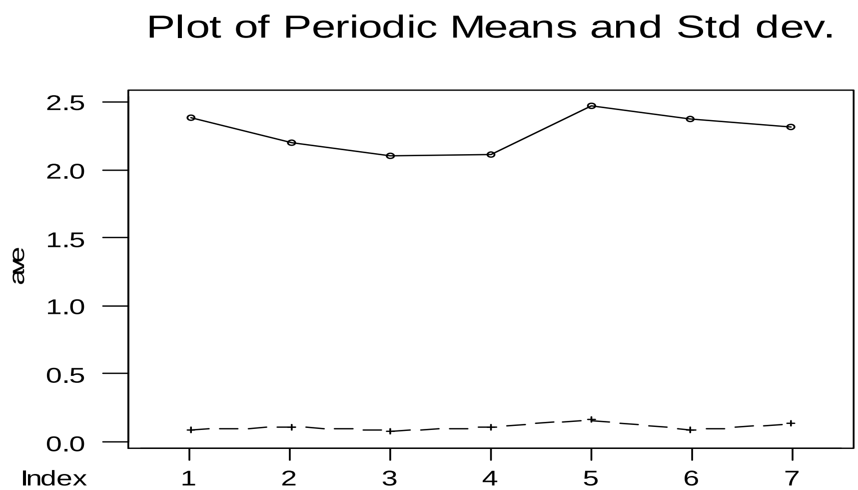

Fig. 4: Plot of Periodic Means and Standard Deviation

From the above plot, the plot of standard deviation is fairly stable on the horizontal line while the means are fluctuating. This indicates that additive model is the appropriate model.

4. Trend Analysis

Trend could be linear or curvilinear. In this work, two suspected trend (linear and quadratic) patterns based on figure 3 above, was plotted for determine the line of best fit.

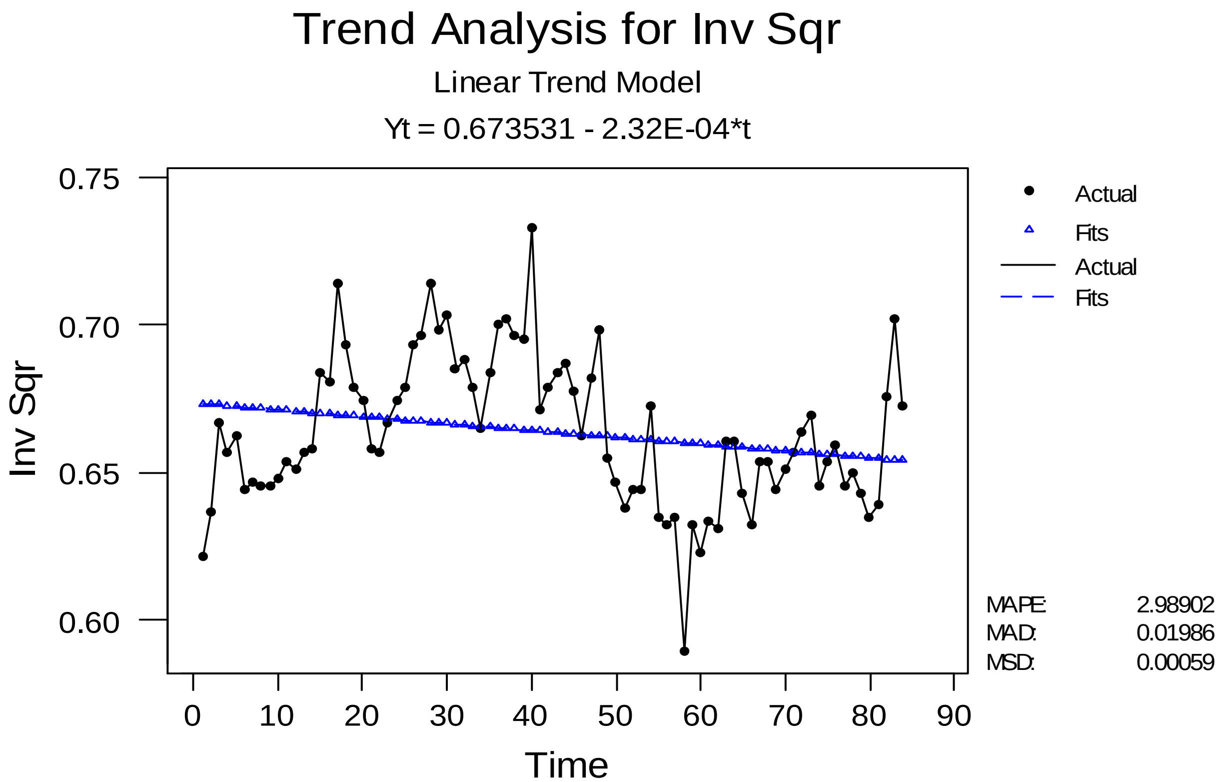

Fig. 5: Trend Analysis for Inv Sqr

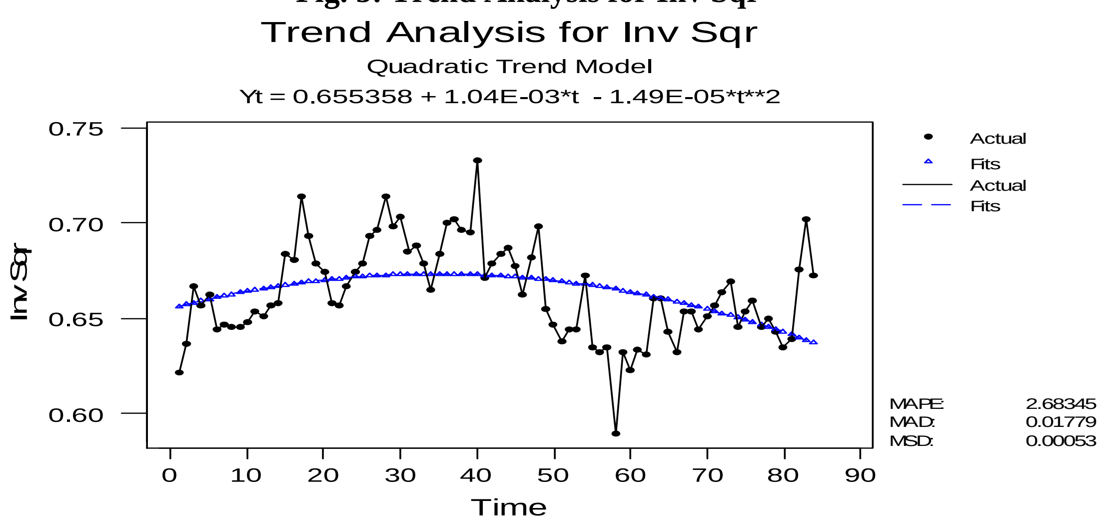

Fig. 6: Trend Analysis of Inv Sqr

Table 3: Accuracy Measure

TYPE OF TREND	ACURACY MEASURE		
	MAPE	MAD	MSD
LINEAR	2.98902	0.01986	0.00059
QUADRATIC	2.68345	0.01779	0.00053

From the table 4.2 above, the accuracy measures (Mean Absolute Percentage Error (MAPE), Mean Absolute deviation (MAD) and Mean Square Deviation (MSD)) shows that the quadratic trend has the least error of estimate when compared with the linear trend. This shows that quadratic trend is the line of best fit and was fitted accordingly.

The fitted line equation is given as
$Y_t = 0.655362 + 0.00104*t - 0.0000149*t^2$; t = 1, 2. . . 84.

Substituting successive values of t (in the range specified above) in the equation gives the values of trend-cycle (M_t) component. For an additive model, subtracting the trend-cycle component from the series gives the detrended series.

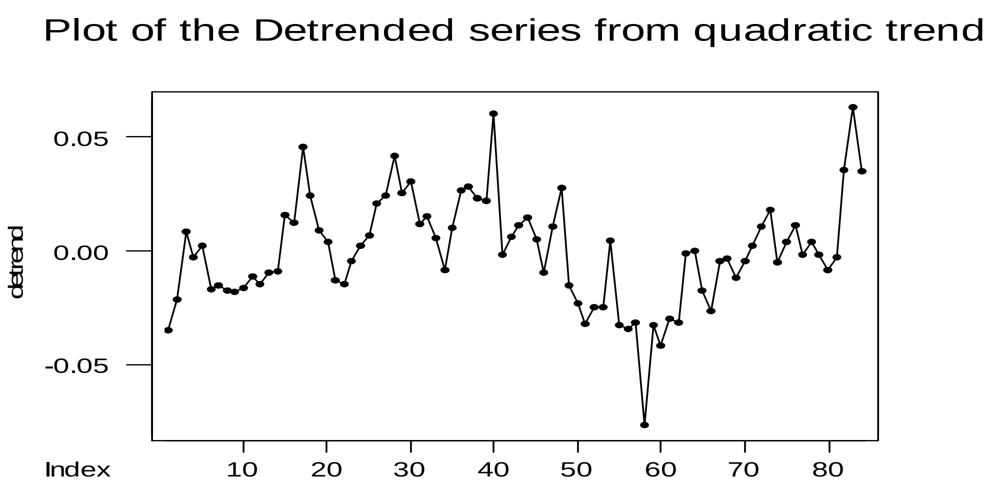

Fig. 7: Plot of the De-trended series from quadratic trend

5.0 Seasonal Analysis

Traditional method of estimating the seasonal component involves arranging the detrended series in a Buys-Ballot table. Seasonal means is expected to sum to zero for an additive model.

Also, for time series which contain a seasonal effect, the overall average $\left(\bar{X}_{..}\right)$ and the seasonal average

$\left(\bar{X}_{.j},\ j=1,2,\cdots,s\right)$ of the Buys-Ballot table are used to assess the effects either as a difference $\left(\bar{X}_{.j}-\bar{X}_{..}\right)$ or a ratio $\left(\bar{X}_{.j}\big/\bar{X}_{..}\right)$ Iwueze et al, 2011. That is the deviations of the differences seasonal averages and the overall average (multiplicative model) from unity is used. The wider the deviations; the greater the seasonal effects.

.

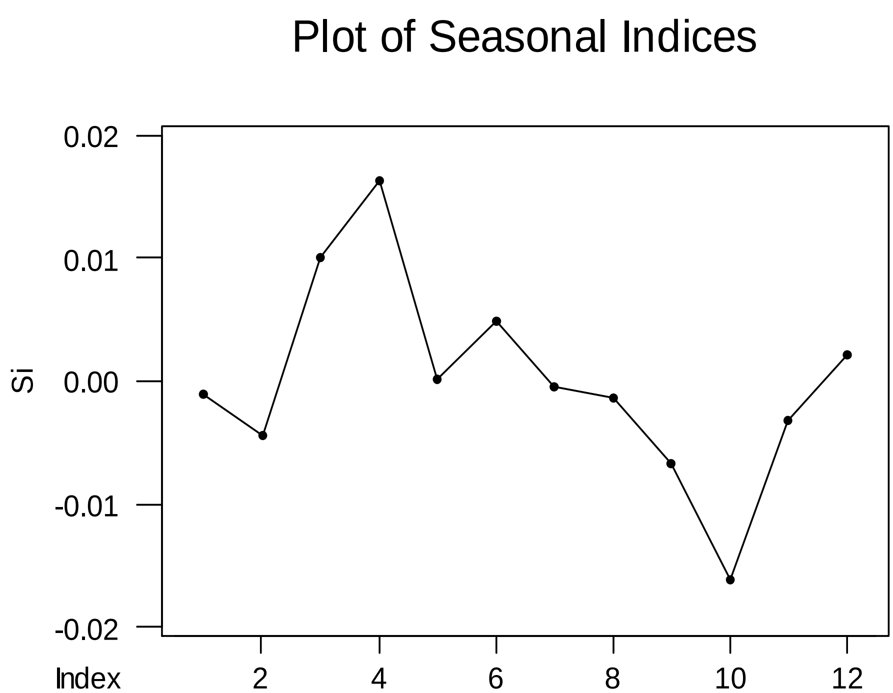

Fig. 8: Plot of Seasonal Indices

The plot above clearly shows the presence of seasonal effect on the series. Table 4 below shows the estimated seasonal component (S_t).

Table 4: Estimated seasonal Indices

S/n	Month	Index
1	January	-0.0010641
2	February	-0.0043782
3	March	0.0100355
4	April	0.0162693
5	May	0.0001183
6	June	0.0048721
7	July	-0.0004820
8	August	-0.0014307
9	September	-0.0066805
10	October	-0.0161239
11	November	-0.0032735
12	December	0.0021377
	Total	**0.00000**

6.0 Estimation of Error Component

Estimation of error or irregular involves subtracting the trend-cycle and seasonal indices from the error i.e. $E_t = Y_t - M_t - S_t$

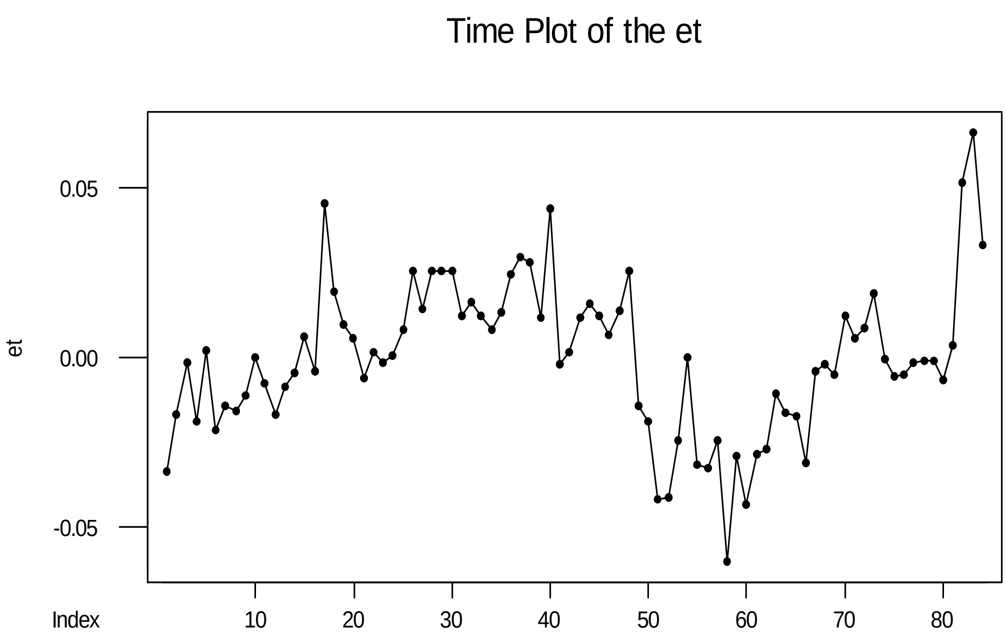

Fig. 9: Time Plot of the et

7. 0 Validation of Normality of Error Component

In time series analysis using Ordinary Least 1Squares (OLS) in estimating the trend parameters we make assumptions which include that the error component is random. In practice, this is not always the case, to verify this; we conduct a test on the randomness of the error component. In this work, we used the Kolmogorov-Smirnov (K-S) test of the non-parametric family. The plot is shown below:

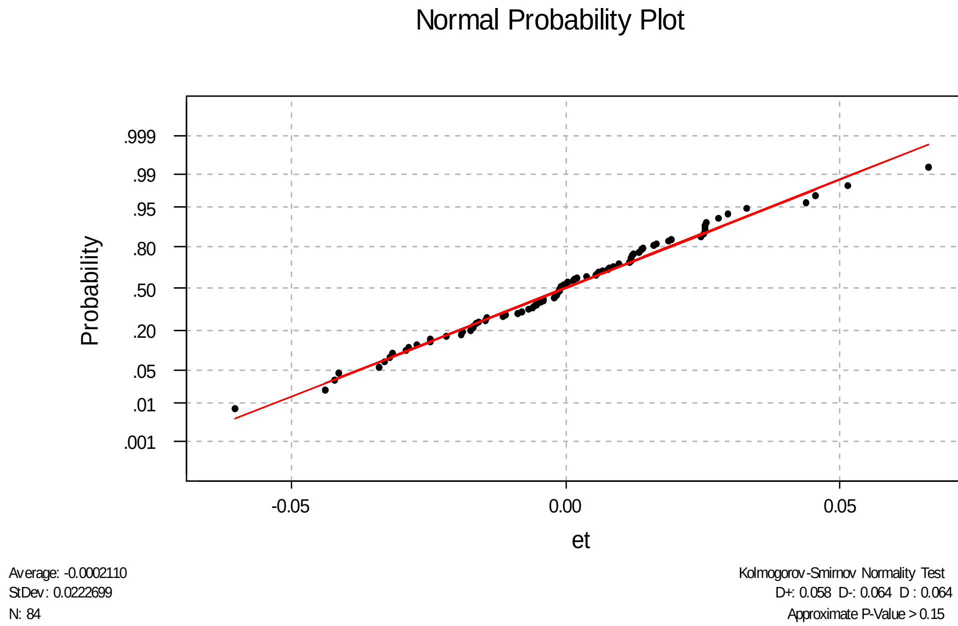

Fig 10: Normal Probability Plot

D-statistic (D- since it has the highest value when compared with D+) is estimated to be 0.064 and the approximate p-value > 0.15. This shows that the error component is random with mean -0.00021103 and standard deviation 0.022270 i.e. $e_t \sim N(0, 0.022270)$.

This shows that the fitted model is fit and appropriate for forecasting future trend pattern.

8.0 Conclusion

This paper has examined time series analysis of Nigeria domestic crude oil under the descriptive approach. . Buys-Ballot table procedure was used in assessing variance stability (transformation), choice of model and seasonal effect. As expected of roduction data, there is influence of seasonal effect where some seasons (months) are seen as the peak while some are seen as the low point in the production.

References

[1] Bartlett, M. S. (1947). "*The Uses of Transformations*": Biometrika, Vol.3, pp. 39-52.
Box, G.E.P., Jenkins, G.M. and Reinsel, G. C. (1994). *"Time Series Analysis, Forecasting and Control, 3rd Ed"*. Englewood Cliffs: Prentice-Hall.

[2] Chatfield, C. (2004). *"The Analysis of time Series: An Introduction"*. Boca Raton: Chapman and Hall/CRC Press.

[3] Esan, E. O. and Okafor, R. O. (2010). *Basic Statistical Methods*. Lagos: Toniichristo Concepts, Lagos

[4] Iwueze, I.S., Nwogu, E. C. and Ajaraogu, J. C. (2011). "*Uses of the Buys-Ballot table in Time Series Analysis*. Journal of Applied Mathematics, 2, 633-645.

[5] Iwueze, I.S. and Akpanta, A.C. (2009). *"On Applying the Bartlett Transformation Method to Time Series Data"*. Journal of Mathematical Sciences, Vol. 20, No. 3, pp. 227-243.

[6] Iwueze, I. S. and Nwogu, E. C. (2004). *"Buys-Ballot for Time Series Decomposition"*. Global Journal of Mathematical Science, 3 (2): 83-89

[7] Kendall, M. G. and Ord, J. K. (1990). *"Time Series, 3rd Ed"*. London: Charles Griffin.