Open Access Library Journal
2020, Volume 7, e6435
ISSN Online: 2333-9721
ISSN Print: 2333-9705

Additive Decomposition with Arima Model Forecasts When the Trend Component Is Quadratic

Biu O. Emmanuel[1], Dennis Enegesele[2], C. O. Arimie[3]

[1]Department of Mathematics & Statistics, University of Port Harcourt, Port Harcourt, Nigeria
[2]Department of Mathematics and Computing Sciences, KolaDaisi University, Ibadan, Nigeria
[3]Department of Radiology, University of Port Harcourt Teaching Hospital, Port Harcourt, Nigeria
Email: emmanuelbiu@yahoo.com, dennis.enegesele@koladaisiuniversity.edu.ng, codarimie@yahoo.com

How to cite this paper: Emmanuel, B.O., Enegesele, D. and Arimie, C.O. (2020) Additive Decomposition with Arima Model Forecasts When the Trend Component Is Quadratic. *Open Access Library Journal*, **7**: e6435.
https://doi.org/10.4236/oalib.1106435

Received: May 16, 2020
Accepted: July 10, 2020
Published: July 13, 2020

Abstract

This paper demonstrates the use of Buys-Ballot table for identification of decomposition model using graphical method, when the trend cycle component is quadratic. A suitable ARIMA model was fitted, and was used for forecasting. Using the Buys-Ballot techniques, the column means, variances and standard deviation were estimated for the model identification. The additive model had no seasonal effect but, the multiplicative model had seasonal effect. The result of the illustrative example using the data of Nigeria Spot component price of oil (US Dollar per Barrel) showed the additive model to be the appropriate model for decomposition of this series. AR(2) model was identified as a suitable ARIMA model for the de-trended Nigeria Spot component price of oil. This was used to make forecast for the next twelve months. The obtained expected oil prices were compared with the observed prices. The comparison of expected and observed prices showed no significance difference between them, using Mean Absolute Percentage Error (MAPE).

Subject Areas

Applied Statistical Mathematics

Keywords

Buys-Ballot Table, Seasonal Averages, Column Variances, Trend Analysis, Identification of Model, Decomposition

1. Introduction

There are two main reasons for time series analysis: 1) pattern identification of a series and 2) forecasting. These goals require that the observed time series data pattern be identified and described [1]. Besides identifying the pattern, the two main reasons will be better evaluated only when the right model is used for the analysis.

Descriptive time series, which is also known as decomposition of time series is the separation of an observed time series into its components represented by the trend (T_t), the seasonal (S_t), cyclical (C_t) and irregular (e_t) components. When the period of time is small, the cyclical component is embedded into the trend and the observed time series ($X_t, t = 1, 2, \cdots, n$) can be decomposed into the trend-cycle component (N_t) seasonal component (S_t) and the irregular/residual component (e_t) [2].

There are three main decomposition models in descriptive time series:

Additive Model:

$$X_t = N_t + S_t + e_t \tag{1}$$

Multiplicative Model:

$$X_t = N_t \times S_t \times e_t \tag{2}$$

Pseudo-Additive/Mixed Model:

$$X_t = N_t \times S_t + e_t \tag{3}$$

where, for time point *t*, N_t is the trend-cycle component; S_t is the seasonal component and e_t is the irregular or random component.

The assumption for the additive model (1) is that the irregular/error component e_t is Gaussian $N\left(0, \sigma_1^2\right)$ white noise and the sum of the seasonal component over a complete period is zero, ($\sum_{j=0}^{s} S_j = 0$) while for the multiplicative model (2), e_t is the Gaussian $N\left(1, \sigma_2^2\right)$ white noise and the sum of the seasonal component over a complete period is ($\sum_{j=0}^{s} S_j = s$).

The major problem in the use of descriptive time series is the choice of adequate model for time series decomposition. The methods used in the literature to make choice between additive, multiplicative and pseudo-additive models are the graphical and non-graphical methods. In time series where the amplitude of both the seasonal and irregular variations does not change as the level of the trend rises or falls, the additive model is adopted. However, when the amplitude of both the seasonal and irregular variations increases as the level of the trend rises, the multiplicative model is adopted [2]. Iwueze *et al.* [3] states that using Buys-Ballot table, the relationship between the seasonal means and the seasonal standard deviation could provide an insight of a desired model. In that study, the time plot for means and standard deviation was used for the choice of model.

The method of coefficient of variation of the seasonal quotients and differences was proposed by Justo and Rivera [4]. They opined that if the coefficient of variation for the seasonal quotients is greater than the coefficient of variation for the seasonal differences, the model for decomposition is additive otherwise, and the model is multiplicative. This method did not provide the choice and the use of seasonal quotients and a difference was not stated. Iwueze and Nwogu [1] provided a framework for choice of model in descriptive time series based on the Buys-Ballot table. According to them, the column (seasonal) variances of the Buys-Ballot table are simply the trending curve of the time series for the additive model and the product of the trending curve and square of the seasonal effect for the multiplicative model.

The ultimate objective of this paper is to identify and remove trend curve (quadratic) of a time series, using the identified decomposition model (additive or multiplicative model) then, fit an ARIMA model to the de-trended series and use the fitted model for forecasts. The significance of this paper is that it will improve the certainty of the analyst in choosing a suitable model for decomposition of a time series when trend is quadratic.

2. Method

2.1. Buys-Ballot Table

A Buys-Ballot table gives the summary of time series data arranged in m rows and s column for possible seasonal variation. In other to analyze the data, it is necessary to include the period and seasonal totals ($T_{i.}$ and $T_{.j}$), period and seasonal averages ($\bar{X}_{i.}$ and $\bar{X}_{.j}$), the grand total and mean ($T_{..}$ and $\bar{X}_{..}$). Wold [5] credits these arrangements of time series data to Buys-Ballot [6] hence, the table is referred to as Buys-Ballot table in the literature and is as shown in Table 1. In Table 1 below, the rows represent the periods/years while the column are the seasons.

For better understanding of Table 1, we have defined the column (*j*) totals, averages and standard deviation as follows:

Table 1. Buys-ballot table.

Period(*i*)	Seasons								
	1	2	…	j	…	s	$T_{i.}$	$\bar{X}_{i.}$	$\hat{\sigma}_{i.}$
1	X_1	X_2	…	X_j	…	X_s	$T_{1.}$	$\bar{X}_{1.}$	$\hat{\sigma}_1$
2	X_{s+1}	X_{s+2}	…	X_{s+j}	…	X_{2s}	$T_{2.}$	$\bar{X}_{2.}$	$\hat{\sigma}_2$
3	X_{2s+1}	X_{2s+2}	…	X_{2s+j}	…	X_{3s}	$T_{3.}$	$\bar{X}_{3.}$	$\hat{\sigma}_3$
…	…	…	…	…	…	…	…	…	…
i	$X_{(i-1)s+1}$	$X_{(i-1)s+2}$	…	$X_{(i-1)s+j}$	…	$X_{(i-1)s+s}$	$T_{i.}$	$\bar{X}_{i.}$	$\hat{\sigma}_{i.}$
…	…	…	…	…	…	…	…	…	…
m	$X_{(m-1)s+1}$	$X_{(m-1)s+2}$	…	$X_{(m-1)s+j}$	…	X_{ms}	$T_{m.}$	$\bar{X}_{m.}$	$\hat{\sigma}_{m.}$
$T_{.j}$	$T_{.1}$	$T_{.2}$	…	$T_{.j}$	…	$T_{.s}$	$T_{..}$	-	-
$\bar{X}_{.j}$	$\bar{X}_{.1}$	$\bar{X}_{.2}$	…	$\bar{X}_{.j}$	…	$\bar{X}_{.s}$	-	$\bar{X}_{..}$	-
$\hat{\sigma}_{.j}$	$\hat{\sigma}_{.1}$	$\hat{\sigma}_{.2}$	…	$\hat{\sigma}_{.j}$	…	$\hat{\sigma}_{.s}$	-	-	$\hat{\sigma}_{..}$

Source: Iwueze and Nwogu [1].

$$T_{.j} = \sum_{i=1}^{m} X_{(i-1)s+j}, \quad i = 1,2,\cdots,s \tag{4}$$

$$\bar{X}_{.j} = \frac{T_{.j}}{m} = \frac{1}{m}\sum_{i=1}^{m} X_{(i-1)s+j}, \quad j = 1,2,\cdots,s \tag{5}$$

$$\hat{\sigma}_{.j} = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}\left(X_{(i-1)s+j} - \bar{X}_{.j}\right)^2}, \quad j = 1,2,\cdots,s \tag{6}$$

where $X_t, t = 1,2,\cdots,n$ is the series, m is the number of periods/years, s is the periodicity, and $n = ms$ is the overall number of observation/sample size.

$T_{.j}$ = Total for jth season, $\bar{X}_{.j}$ = Average of *j*th season

$\hat{\sigma}_{.j}$ = Standard deviation for *j*th season.

Let, $X_t = a + bt + ct^2 + S_t + e_t$ be the observation of the time series at time, *t*.

Define $\bar{X}_{.j}$ as column mean and $\sigma^2_{.j}$ as column variances for additive model.

We can write $t = (i-1)s + j$ in terms of the row (*i*) and column (*j*) of the Buys-Ballot table.

For the multiplicative model:

Let, $X_t = (a + bt + ct^2) * S_t * e_t$ be the observation of time series at time, *t*.

Define $\bar{X}_{.j}$ as column mean and $\sigma^2_{.j}$ as column variances for multiplicative model.

We can write $t = (i-1)s + j$ in terms of the row (*i*) and column (*j*) of the Buys-Ballot table [3].

Several papers in the statistics literature have discussed the use of time plot of the entire series to make the appropriate choice between additive and multiplicative models. This makes a review of some of the works done as regards the choice of models in a descriptive time series analysis necessary in order to highlight the import of this study.

The time plot of a series can be used to choose between additive and multiplicative models. If the seasonal variation stays roughly the same size regardless of the mean level, then it is additive, but if it increases in size in direct proportion to the mean level, then it is said to be multiplicative [2].

Iwueze *et al.* [3] on the uses Buys-Ballot table, states that the relationship between the seasonal means and the seasonal standard deviation could give an indication of a desired model. Additive model should be employed when the seasonal standard deviation shows no applicable increase/decrease relative to any increase or decrease in the seasonal means. The multiplicative model should be employed when the seasonal standard deviation shows applicable increase/decrease relative to any increase or decrease in the seasonal means. This comparison was done using time plot for means and standard deviation.

Figures 1-4 respectively illustrate two (2) time series with their trend components, on which the choice of appropriate model can be easily decided. In the first case, the additive model was the appropriate choice as the differences between the trend and observed data (the seasonal differences) for the same periods in different years are almost the same, while in the second case the multiplicative

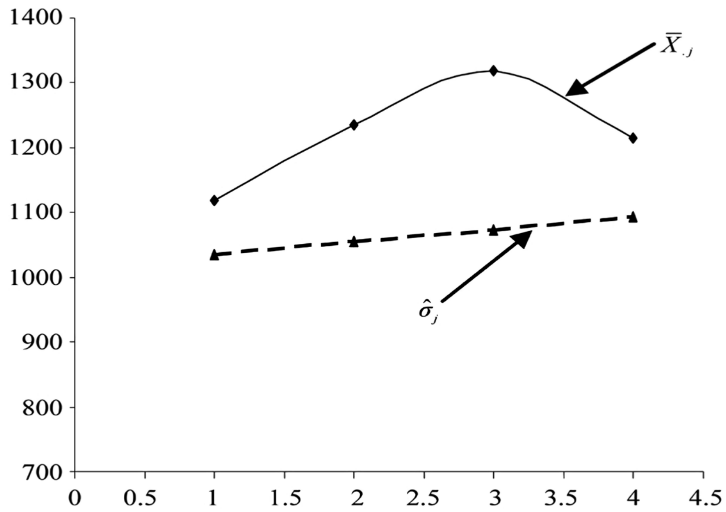


Figure 1. Additive model with non-seasonal effect.

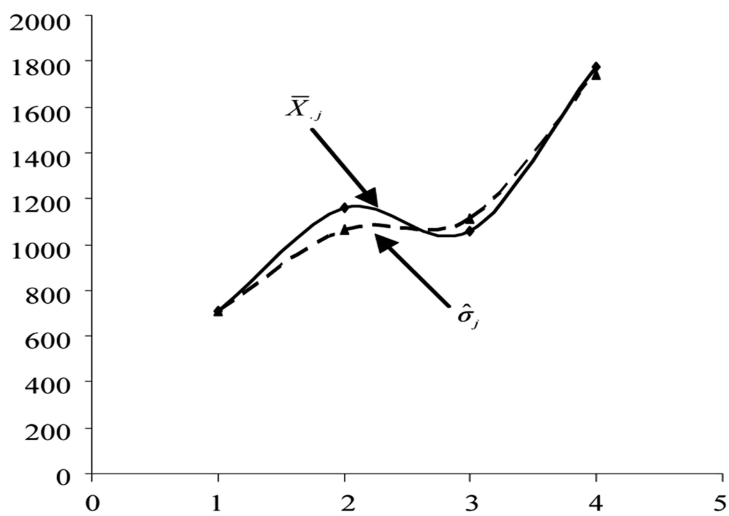


Figure 2. Multiplicative model with non-seasonal effect.

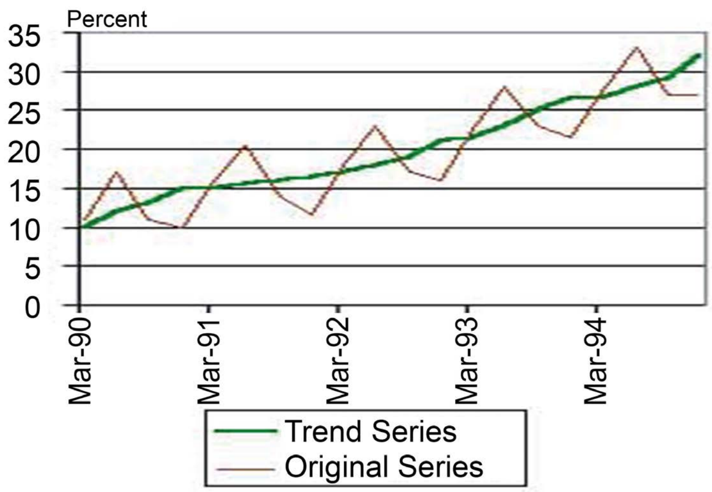


Figure 3. Time series appropriate for additive decomposition with seasonal effect.

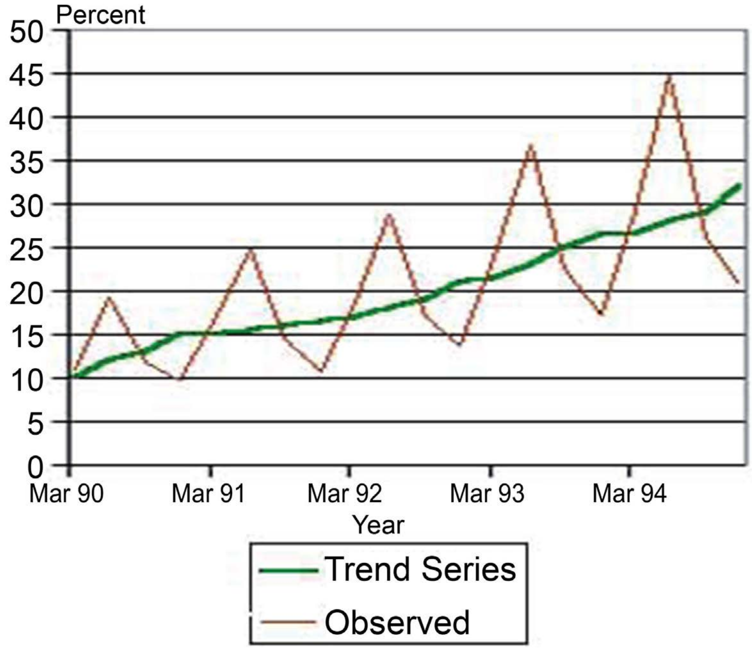


Figure 4. Time series appropriate for multiplicative decomposition with seasonal effect. **Source:** http://www.stats.govt.nz/surveys_and_methods/methods/data-analysis/seasonal-adjustment/theunderlying-model.aspx.

model was chosen, because the ratios of the trend and observed data (the seasonal indices) for the same periods in different years are almost the same. Thus, the appropriate model is either additive or multiplicative.

2.2. Autoregressive Moving Average (ARMA) Model

Frequently, after achieving stationarity, a time series contains AR(p) and MA(q) components of certain orders which can be combined and used for forecasting. It can also be called mixed process. Thus, a mixture of autoregressive process of order p, AR(p), and moving average of order q, MA(q), denoted as ARMA(p, q), is of the form;

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q} \quad (7)$$

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q}$$

$$\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right) X_t = \left(1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q\right) e_t$$

$$\phi(B) X_t = \theta(B) e_t \quad (8)$$

where, $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ and $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ and e_t is a sequence of independently and identically distributed (*iid*) random variables with $E(e_t) = 0$ and $Var(e_t) = \sigma_e^2$.

For stationarity, the roots of $\phi(B) = 0$ lie outside the unit circle and for invertibility condition, the roots of $\theta(B) = 0$ all lie outside the unit circle.

For the ARMA(p, q) process, there are q autocorrelations, $\rho_1, \rho_2, \cdots, \rho_q$ whose values depend directly on the choice of the q moving average parameters

$\theta_1, \theta_2, \cdots, \theta_q$, as well as on the p autoregressive parameters $\phi_1, \phi_2, \cdots, \phi_q$.

2.2.1. Model Selection Criteria

When more than one model is selected from the process enumerated in Equation (8), the Akaike's Information Criterion (AIC) is then used to select the most suitable model amongst them. The Akaike's Information Criterion is most commonly given as:

$$\text{AIC} = N \log \hat{\sigma}^2 + 2r \tag{9}$$

where, r is the number of model parameters, N = Effective number of data point used in the estimation procedure and $\hat{\sigma}^2$ is the estimated residual variance (Mean sum of squared error (MSE)) [7] [8]. The model that minimizes the AIC criterion is the best model.

2.2.2. Accuracy Measures of the Estimated Values

1) Estimated Errors

To gauge the accuracy of our estimates, the estimated errors will be used to compare the expected estimated forecast values and observed values for 2013. This is done by subtracting the estimated forecast values (EFV) from the original values or [actual values (AV)] to obtain the estimate errors [9]. The estimate error (e_i) is denoted by

$$e_i = \text{AV}_i - \text{EFV}_i,\ \ i = 1, 2, \cdots, v \tag{10}$$

However, the accurate measures used in this paper are Mean Absolute Percentage Error (MAPE).

2) Mean Absolute Percentage Error (MAPE)

This accounts for the percentage of deviation between the actual values and estimates [10]. This can be obtained as:

$$\text{MAPE} = 100 \times \left[\frac{1}{v} \sum_{i=1}^{v} \frac{e_i}{\text{AV}_i} \right] \ \ \left(\text{AV}_i \neq 0 \right) \tag{11}$$

where, v is the number of forecast values.

This method is now illustrated in Section 3.0 with the use of real-life time series data.

3. Empirical Illustration

The section is divided into two parts: 1) Identifying between additive or multiplicative model for time series decomposition. 2) ARIMA model and forecasting.

3.1. Identification of Additive or Multiplicative Model

Nigeria Spot component price of oil (US Dollar per Barrel) from 1983-2013 data (**Appendix A**) was applied to the Buys-Ballot table to ascertain if the additive or multiplicative models should be used for Time Series decomposition. The column means, variances and standard deviation were obtained. Then, the trend behaviour using means and standard deviation is shown below in **Figure 5**.

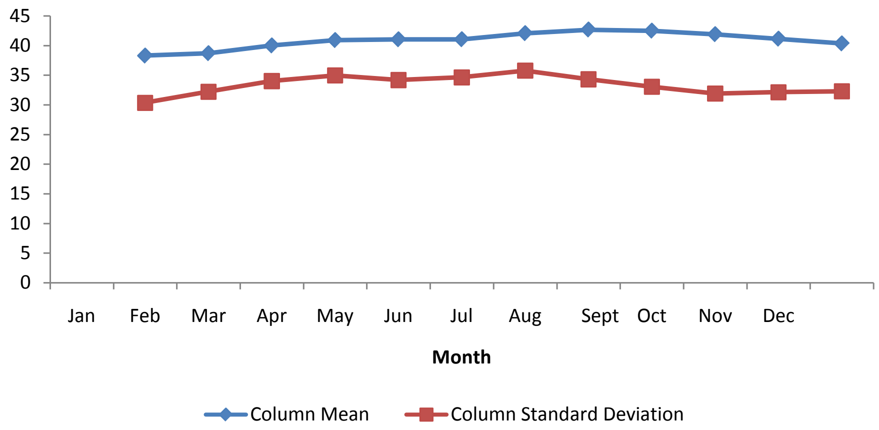


Figure 5. decomposition model identification.

The time plot for means and standard deviation in **Figure 5** was used for the choice of model. **Figure 5** show that the standard deviation variation stays roughly the same size regardless of the mean level, which implies additive [2]. Thus, the appropriate model for decomposition Nigeria Spot component price of oil (US Dollar per Barrel) is the additive model.

3.2. ARMA Model and Forecasting

This section is divided into three parts: de-trend the series using the appropriate model for decomposition (*i.e.* additive model), ARMA modeling of the series and Forecasting.

Data: Nigeria Spot component price of oil (US Dollar per Barrel) **(1983-2013)**

The Nigeria Spot component price of oil (US Dollar per Barrel) is a monthly data comprising 372 data points given in 31 years (1983-2013). **Appendix A** shows a complete presentation of the data in Buys-Ballot table form and its series plot is shown as **Figure 6**. Note that the first 30 years (1983-2012) was used for model building (X_t) and the last year (2013) was used for forecasts comparison.

Examining **Figure 6**, we noticed that NSCPO series appreciate from January 1983 to December 2012, which indicate a trend component which is either any of the polynomials of suitable order. Hence, we fitted several polynomials trend curve of order m_0 in other to identify which curve represents the trend component in the NSCPO series, using R-square (R^2) ([11] [12] [13]). The polynomial trend that R^2 is close to 1, with its entire coefficient being significant is the best trend curve.

3.2.1. Trend Identification and the De-Trended Series

The general polynomial trend is expressed as:

$$\hat{X}_t = \hat{C}_0 + \hat{C}_1 t + \hat{C}_2 t^2 + \cdots + \hat{C}_m t^m \qquad (12)$$

where, m is the order of the fitted polynomial at which R^2 is close to 1.

We used Equation (12) *to estimate the coefficients of the polynomials,* $C_k, k = 1, 2, \cdots, m$ *in* **Figure 7** *using Microsoft Excel. Then, we also used regression*

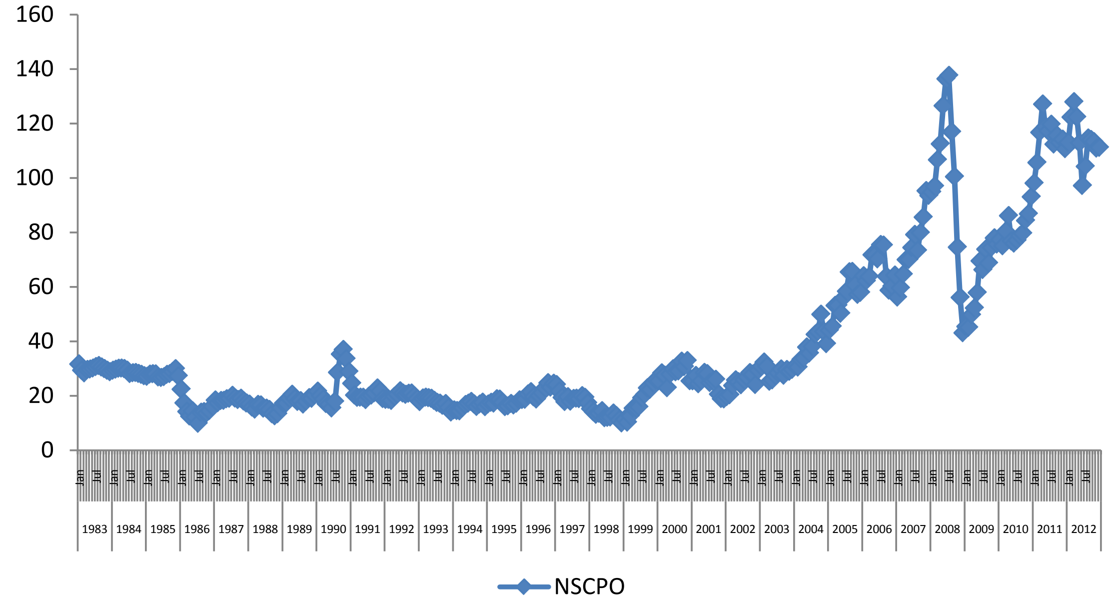


Figure 6. Nigeria spot component price of oil (NSCPO) (1983-2012) "X_t".

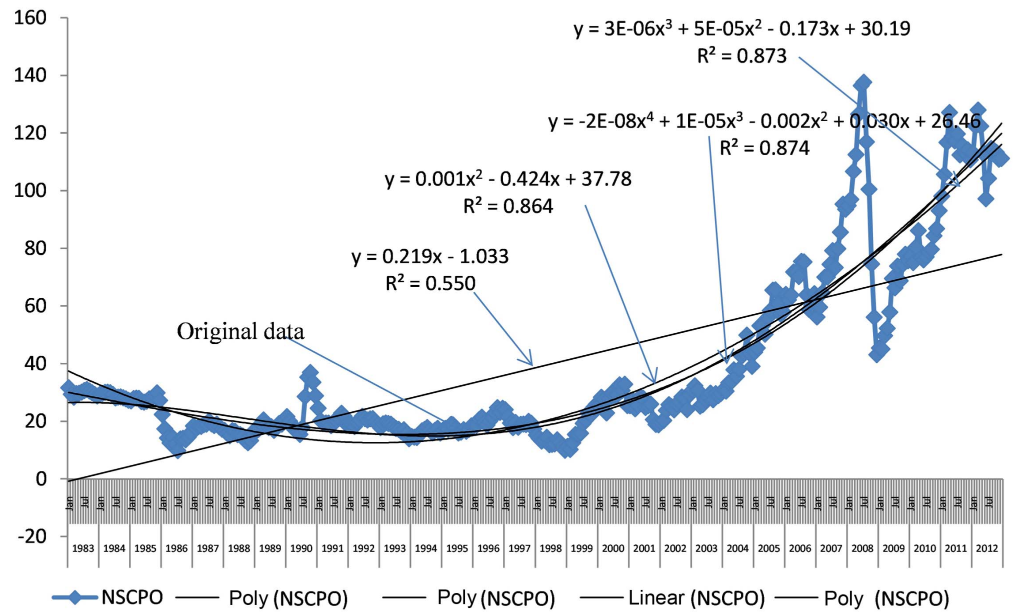


Figure 7. Trend curves fitted to NSCPO series (X_t).

method to fit the coefficients of the polynomials of order m with the help of Minitab 17 in Appendix B, summarized in Table 2.

From Table 2, one of the estimated coefficients of the polynomial of order 3 is not significant and two estimated coefficients of the polynomial of order 4 are also not significant. However, when these coefficients are remove from the fitted polynomial trend curve, the resultant trend curve is quadratic as shown in Table 2 above (Hint: over fitting method). Therefore, the best fitted polynomial trend curve is quadratic trend with R^2 = 0.86%. Another reason why the polynomials of order three and four cannot be used for this analysis is that they fit negative

Table 2. Fitted polynomials of the NSCPO series.

order m	Estimated Coefficients (p-values)					R^2	REMARK
	$\hat{C}_0$	$\hat{C}_1$	$\hat{C}_2$	$\hat{C}_3$	$\hat{C}_4$		
1	−1.033 **(0.636)**	0.219 (0.000)				55.0%	Linear Trend
2	**37.781 (0.000)**	**−0.4242 (0.000**	**1.782 × 10^{-3} (0.000)**			**86.4%**	**Quadratic trend**
3	30.198 (0.000)	−0.1739 (0.000)	5.07 × 10^{-5} (0.888)**	3.20 × 10^{-6} (0.000)		87.30%	Cubic Trend
4	26.462 (0.000)	0.0305 (0.786)**	−0.0025 (0.05)**	1.413 × 10^{-5} (0.000)	−2.0 × 10^{-8} (0.036)	87.4%	Quartic trend

Footnote: **p-values are greater than the appropriate critical value (0.05) and the bold trend is the optimal order (m = 2) identified.

values to values that are positive. Thus, we conclude that the trend component is represented by the quadratic trend curve.

The fitted polynomial (quadratic trend; m = 2) is given as

$$\hat{X}_t = C_0 + C_1 t + C_2 t^2 \tag{13}$$

By substitution of the coefficients in Table 2 into Equation (13), we have

$$\hat{X}_t = 37.781 - 0.4242t + \left(1.782 \times 10^{-3}\right)t^2; \quad t = 1, 2, \cdots, 360 \tag{14}$$

Next, we used the additive model identified in removing the trend of the NSCPO Series since the trend curve is the quadratic. The identified model decomposition process is the additive model in Section 3.1 and Equation (1) is the representation of the additive model.

- **Additive Model Decomposition:**

$$Y_t = X_t - \hat{X}_t, \; t = 1, 2, \cdots, 360 \tag{15}$$

Then,

$$Y_t = X_t - \left(37.781 - 0.4242t + \left(1.782 \times 10^{-3}\right)t^2\right), \; t = 1, 2, \cdots, 360 \tag{16}$$

Figure 8 shows that the series is stationary and the variance is constant which indicated that the quadratic trend have been removed completely by the use of Equation (16) [or by the Additive Method.]. We now fit the best ARMA(p, q) model to the de-trended series [represented by (16)].

3.2.2. ARMA Modelling of the Series (16)

From Figure 8 above, the autocorrelation function (ACF) and partial autocorrelation function (PACF) are shown in Figure 9 and Figure 10, respectively.

Figure 9 show significant spikes at lags 1, 2, 3, 4 and 5 while Figure 10 showed significant spikes at lags 1 and 2 only. Figure 9 suggests MA(5) while Figure 10 suggest AR(2) process. However, suitable ARMA(p, q) models ($p + q \le 5$) may also be appropriate.

On the other hand, a test is appropriate, to test if the constant mean "μ" be included in the models. In this case, the hypothesis of interest is given as

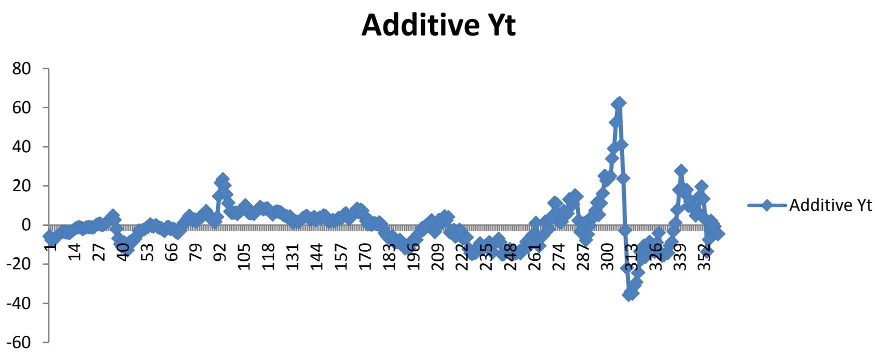


Figure 8. De-trended series of the NSCPO series (Y_t).

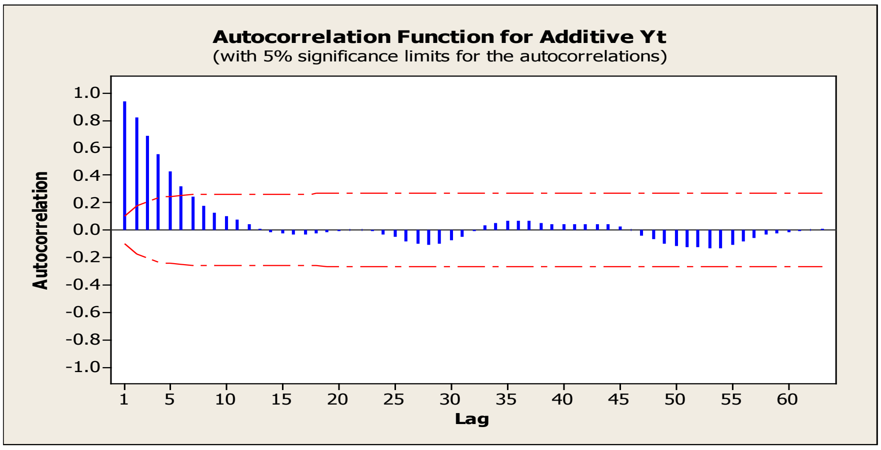


Figure 9. ACF correlogram of NSCPO "Y_t", Equation (16).

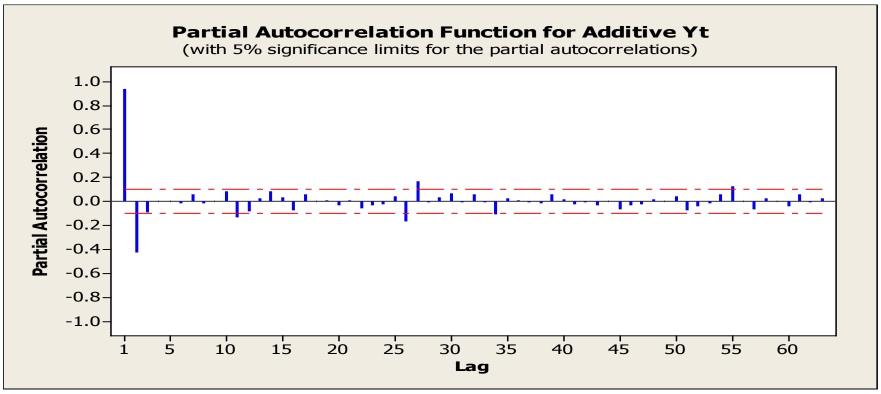


Figure 10. PACF correlogram of NSCPO "Y_t", Equation (16).

$$H_0 : \mu = 0 \text{ against } H_1 : \mu \neq 0 \tag{17}$$

The test statistics for testing $H_0 : \mu = 0$ against $H_1 : \mu \neq 0$ is

$$t = \frac{\bar{Y}_t}{std(Y_t)} \tag{18}$$

The computed t-value is $t = -9.32 \times 10^{-7}$ with p-value = 1.000 (see **Appendix C**). This p-value is greater than the critical p-value = 0.05 hence, we accepted $H_0 : \mu = 0$. This implies that μ should not be included in the model.

Various ARMA(p, q) models were fitted to Equation (16) with respective residuals being white noise and the summary is shown in Table 3. The model selection criteria used to select the best model amongst models is Akaike's Information Criterion (AIC) [Section 2.2.1, Equation (17)] is also detailed out in Table 3.

The identified model using Akaike's Information Criterion in Table 3 is AR(2) model. AR(2) can be expressed as

$$Y_t = \phi_1 T_{t-1} + \phi_1 Y_{t-2} + e_t \tag{19}$$

Estimation of the Parameter of the Identified AR(2) for (16)

Estimates were obtained by use of Minitab 17 software and the results are tabulated in Table 4.

The residuals ACF and PACF in Figure 11 and Figure 12 reveal that the model is adequate for Equation (16). The adequacies of the model were also checked by the use of Ljung-Box Chi-square statistics [14], and the results are summarized in Table 5.

Table 3. AIC values for ARMA(p, q) models ($p+q \le 5$).

Model	r	σ^2	N	AIC
AR(1)	1	16.03	360	1000.81
AR(2)	**2**	**13.09**	**360**	**929.87**
MA(5)	5	14.12	360	963.13
ARMA(1, 1)	2	13.81	360	949.14
ARMA(1, 2)	3	13.24	360	935.97
ARMA(1, 3)	4	13.15	360	935.51
ARMA(1, 4)	5	13.14	360	937.24
ARMA(1, 5)	6	13.05	360	936.76

Table 4. Parameter estimates of AR(2) Model for (16).

AR(2) Model
$\varphi_1 = 0.3389 \pm 0.0477$
$\varphi_2 = -0.4315 \pm 0.0477$
$\sigma^2 = 13.09$

Footnote: values after (±) are their standard errors.

Table 5. (Ljung-box) chi-square statistic for adequacy of (19).

k	df	AR(2)	Chi-square Table $\chi^2_{(h)}$
		$Q(k)$	
12	10	14.8	18.3
24	22	32.8	33.9
36	34	55.3	58.8
48	46	67.0	67.5

Footnote: k is the lags.

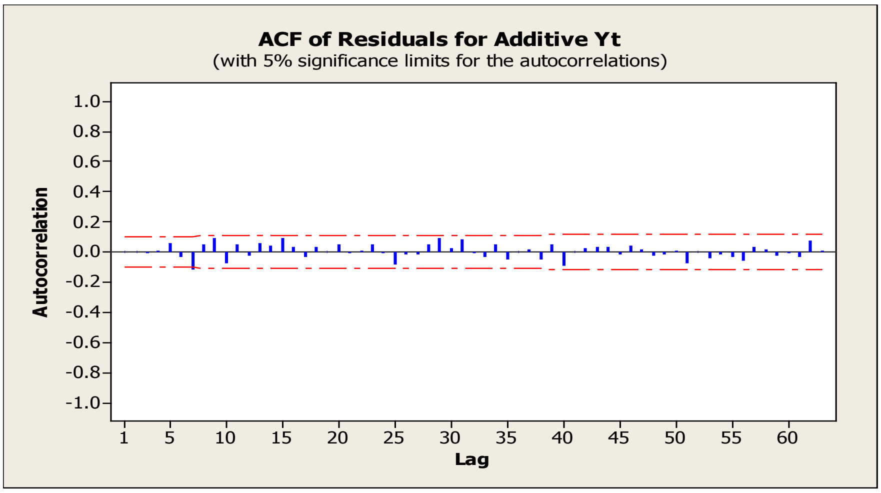


Figure 11. Residual ACF correlogram of AR(2).

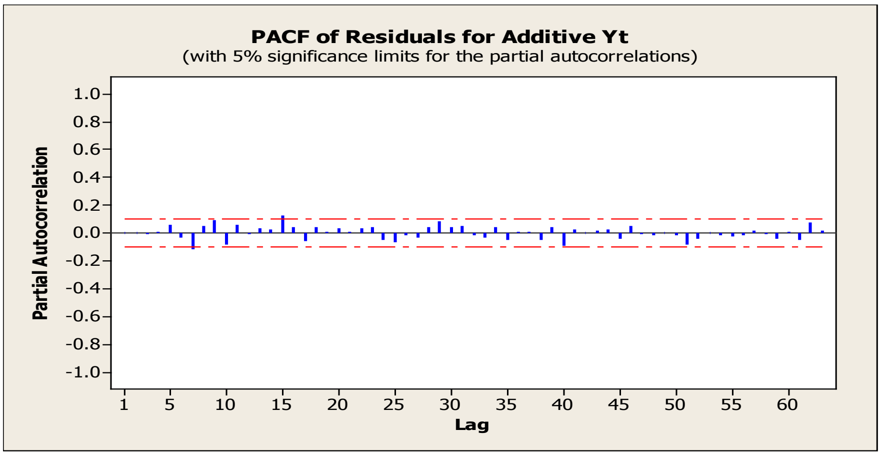


Figure 12. Residual PACF correlogram of AR(2).

In Table 5, comparing $Q(k)$ with $X^2_{(df)}$, [*i.e.* $Q(k) < X^2_{(df)}, k = 12, 24, 36, 48$], it is obvious that the model is adequate and they can be used for forecasting.

3.2.3. Forecasting

We obtained forecasts for AR(2) model in Equation (19) starting at the origin point 360 for 12 months. The ℓ-step forecasts denoted by $\hat{Y}_t(\ell_i)$ for $i = 1, 2, \cdots, 12$ with the 95% confidence intervals of the forecasts are shown in **Appendix D**. Similarity the forecasts values using Equation (19) was obtained. The expected estimated forecast values for NSCPO Series (X_t); *i.e.* ($X_t = Y_t + \hat{X}_t$) and the estimated error (e_i) between actual values for the year 2013 and the expected estimated forecast values, using Equation (14) are shown in Table 6.

In Table 6 below, the accuracy measures of the estimated forecasts confirmed that additive model is the suitable method, because it is closer to the original values for 2013. [Hint: using the MAPE accuracy measure, the additive decomposition method shows 8.64% (or 9%) less than the original values for the year 2013).

Table 6. Comparison of forecast values obtained by Equation (14) and Equation (19) with the original values for 2013.

Period (t)	Months	Actual (2013)	Expected estimated Forecast Values			Estimated Errors "e_i"
			AR(2) Y_t	**Quadratic Trend** $\hat{X}_t$	$\left(X_t = Y_t + \hat{X}_t\right)$	
361	January	115.41	−4.6296	116.88	112.25	3.16
362	February	118.69	−4.1159	117.74	113.62	5.07
363	March	110.57	−3.5135	118.61	115.10	−4.53
364	April	105.17	−2.9284	119.48	116.55	−11.38
365	May	105.83	−2.4050	120.36	117.96	−12.13
366	June	106.12	−1.9566	121.24	119.28	−13.16
367	July	110.21	−1.5821	122.12	120.54	−10.33
368	August	113.62	−1.2741	123.00	121.73	−8.11
369	September	114.3	−1.0233	123.89	122.87	−8.57
370	October	112.44	−0.8204	124.78	123.96	−11.52
371	November	111.47	−0.6570	125.68	125.02	−13.55
372	December	113.11	−0.5257	126.58	126.05	−12.94
MAPE						**8.64**

4. Conclusion

The Buys-Ballot procedure has shown that 1) the column mean and variances are not the same for the two models (Additive and Multiplicative), 2) when the trend is quadratic, the column variances mimic the shape of the trending series for both the additive and multiplicative models for the illustrative example. The result of the illustrative example using the data of Nigeria Spot component price of oil (US Dollar per Barrel) showed the additive model to be the appropriate model for decomposition of the series, based on the relationship between means and standard deviation. Finally, AR(2) was found to be adequate for the series under consideration; hence it was used to forecast. The comparison of the expected and observed prices showed no significant difference between them, using Mean Absolute Percentage Error (MAPE). Hence, we concluded that the additive model should be adopted when the trend component is quadratic (*i.e.*, when the variation does not change as the level of the trend rises or falls).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Iwueze, I.S. and Nwogu, E.C. (2014) Framework for Choice of Model and Detection of Seasonal Effect in Time Series. *Far East Journal of Theoretical Statistics*, **48**,

45-66.

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

[3] Iwueze, I.S., Nwogu, E.C., Ohakwe, J. and Ajaraogu, J.C. (2011) Uses of the Buys-Ballot Table in Time Series Analysis. *Journal of Applied Mathematics*, **2**, 633-645. https://doi.org/10.4236/am.2011.25084

[4] Justo, P. and Rivera, M.A. (2010) Descriptive Analysis of Time Series Applied to Housing Prices in Spain. Management Mathematics for European Schools, 94342-CP-1-2001-DE-COMENIUS-C21.

[5] Wold, H. (1938) A Study in the Analysis of Stationary Time Series. 2nd Edition, Almqrist and Witsett, Stockholm.

[6] Buys-Ballot, C.H.D. (1847) Leo Claemert Periodiques de Temperature. Kemint et Fills, Utrecht.

[7] Akaike, H. (1974) A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, **19**, 716-723. https://doi.org/10.1109/TAC.1974.1100705

[8] Biu, O.E. and Iwueze, I.S. (2011) Application of Statistical Interventions Analysis to Oil and Gas Series in Niger Delta, Nigeria. *Nigeria Annual International Conference and Exhibition*, Abuja, 3-4. https://doi.org/10.2118/150744-MS

[9] Bradley, H., Christophe, T. and Denis, C. (2014) Standard Errors: A Review and Evaluation of Standard Errors Estimators Using Monte Carlo Simulations. *The Quantitative Methods for Psychology*, **10**, 107. https://doi.org/10.20982/tqmp.10.2.p107

[10] Smith, S. (1987) Tests of Accuracy and Bias for County Population Projections. *Journal of the American Statistical Association*, **82**, 991-1003. https://doi.org/10.1080/01621459.1987.10478528

[11] Henry, G.A. (2010) Comparison of Akaike Information Criterion (AIC) and Bayesian Criterion (BIC) in Selection of an Asymmetric Price Relationship. *Journal of Development and Agricultural Economics*, **2**, 001-006.

[12] Richard, A.S. (1994) Model Comparisons and R^2. *The American Statistician*, **48**, 113-117. https://doi.org/10.1080/00031305.1994.10476036

[13] Draper, N.R. and Smith, H. (1998) Applied Regression Analysis. John Wiley and Sons Inc., New York. https://doi.org/10.1002/9781118625590

[14] Ljung, G.M. and Box, G.E.P. (1978) On a Measure of Lack of Fit in Time Series Models. *Biometrika*, **65**, 67-77. https://doi.org/10.1093/biomet/65.2.297

Appendix A

Buys-Ballot Table for Nigeria Spot component price of oil (US Dollar per Barrel)

Year	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
1983	31.65	29.38	28.35	29.6	29.76	30.13	30.66	31.04	30.51	29.93	29.31	28.78
1984	29.32	29.7	30.05	30.11	29.93	29.33	28.1	28.4	28.48	28.15	27.8	27.31
1985	27.09	27.89	28.04	27.98	26.85	26.66	26.97	27.79	28.09	28.87	29.89	27.25
1986	22.38	17.4	14.21	12.5	14.41	12.09	9.82	13.91	14.11	13.6	14.49	16.12
1987	18.4	17.43	18.2	18.25	18.84	19.01	20.1	19.1	18.55	19.03	18.05	17.28
1988	16.85	15.8	15.02	16.76	16.57	15.38	15.39	14.9	13.34	12.59	13.33	15.59
1989	17.27	17.23	19.18	20.49	18.96	17.85	18.14	16.87	17.93	19.26	19	20.28
1990	21.64	20.04	18.73	17.1	16.75	15.46	17.86	28.6	35.31	36.95	33.55	28.85
1991	24.55	20.25	19.36	19.41	19.39	18.53	19.81	20.18	21.03	22.81	21.74	18.91
1992	18.6	18.63	18.22	19.75	20.55	21.76	20.94	20.46	20.86	20.95	19.91	18.83
1993	17.8	19.13	19.42	19.24	19.01	18.25	17.51	17.22	16.44	17.08	15.66	13.96
1994	14.74	14.5	14.4	15.55	16.72	17.21	17.85	16.98	16.01	16.89	17.58	15.94
1995	16.92	17.54	17.24	18.84	18.71	17.58	15.95	16.25	17.11	16.56	17.19	18.44
1996	18.55	18.64	20.64	21.43	19.58	18.73	20.04	21.15	22.95	24.74	23.11	24.53
1997	24.04	21.65	19.39	17.82	19.6	17.95	18.95	19.04	18.89	19.98	19.36	17.34
1998	15.25	14.11	13.14	13.51	14.46	11.89	12.01	12.14	13.59	12.66	11.15	9.96
1999	11.33	10.24	12.56	15.44	15.45	15.86	19.28	20.44	22.9	22.3	24.8	25.86
2000	25.41	28.36	27.54	22.91	27.87	29.86	28.75	29.06	32.65	30.67	32.86	25.47
2001	25.43	27.4	24.35	25.43	28.51	28.06	24.81	25.41	25.98	20.6	18.92	18.78
2002	19.65	20.3	23.76	25.79	25.1	23.98	25.93	26.94	28.46	27.9	24.07	29.27
2003	30.78	32.33	30.83	25.27	25.78	27.46	28.39	29.79	27.47	29.59	28.93	29.64
2004	30.94	30.47	33.34	33.74	37.87	35.6	38.08	42.55	43.56	49.91	43.6	39.08
2005	44.01	45.43	53.15	53.18	50.23	55.93	58.4	65.49	65.6	60.74	57.18	57.91
2006	64.04	62.12	63.8	71.8	71.75	70.22	75.49	75.29	63.87	58.75	60.32	64.28
2007	56.18	59.58	64.6	70.01	70.03	74.45	79.21	73.34	79.87	85.6	95.32	93.55
2008	94.85	96.98	106.68	112.52	126.55	136.44	137.64	116.93	100.48	74.57	56.11	43.1
2009	45.44	45.07	49.7	52.24	57.87	69.55	66.31	73.84	68.74	74.41	77.96	75.68
2010	77.39	75.04	80.4	86.14	76.87	76	77.04	78.82	79.65	84.35	86.83	93.08
2011	98.1	105.66	116.75	127.12	118.88	117.27	119.69	112.41	115.63	113.09	114.21	110.71
2012	113.08	122.36	127.98	122.36	112.87	97.19	104.24	114.63	114.06	113.31	110.91	111.19
2013	115.41	118.69	110.57	105.17	105.83	106.12	110.21	113.62	114.3	112.44	111.47	113.11

Source: Central Bank of Nigeria Statistical Bulletin 2014.

Months	Column Mean	Column Variances	Column Standard Deviation
Jan	38.29	921.26	30.35
Feb	38.69	1038.06	32.22
Mar	39.99	1157.33	34.02
Apr	40.89	1221.96	34.96
May	41.02	1168.89	34.19
Jun	41.03	1199.70	34.64
Jul	42.05	1279.42	35.77
Aug	42.66	1177.37	34.31
Sept	42.47	1092.76	33.06
Oct	41.88	1017.76	31.90
Nov	41.12	1032.33	32.13
Dec	40.33	1042.91	32.29

Appendix B. Minitab 17 Output

Regression Analysis: NSCPO versus *t*

The regression equation is

$$NSCPO = -1.03 + 0.219t$$

Predictor	Coef	SE Coef	T	P
Constant	−1.033	2.181	−0.47	0.636
t	0.21910	0.01047	20.93	0.000

S = 20.6454 R-Sq = 55.0% R-Sq(adj) = 54.9%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	1	186,648	186,648	437.90	0.000
Residual Error	358	152,592	426		
Total	359	339,240			

Regression Analysis: NSCPO versus t, t^2

The regression equation is

$$NSCPO = 37.8 - 0.424\,t + 0.00178\,t^2$$

Predictor	Coef	SE Coef	T	P
Constant	37.781	1.803	20.95	0.000
t	−0.42423	0.02307	−18.39	0.000
t^2	0.00178209	0.00006188	28.80	0.000

S = 11.3404 R-Sq = 86.5% R-Sq(adj) = 86.4%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	293,328	146,664	1140.43	0.000
Residual Error	357	44,912	129		
Total	359	339,240			

Regression Analysis: NSCPO versus t, t^2, t^3

The regression equation is

$$NSCPO = 30.2 - 0.174t + 0.000051t^2 + 0.000003t^3$$

Predictor	Coef	SE Coef	T	P
Constant	30.198	2.343	12.89	0.000
t	−0.17388	0.05612	−3.10	0.002
t^2	0.0000507	0.0003610	0.14	0.888
t^3	0.00000320	0.00000066	4.86	0.000

S = 10.9968 R-Sq = 87.3% R-Sq(adj) = 87.2%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	3	296,189	98,730	816.42	0.000
Residual Error	356	43,051	121		
Total	359	339,240			

Regression Analysis: NSCPO versus t, t^2, t^3, t^4

The regression equation is

$$NSCPO = 26.5 + 0.031t - 0.00249\ t^2 + 0.000014t^3 - 0.000000t^4$$

Predictor	Coef	SE Coef	T	P
Constant	26.462	2.933	9.02	0.000
T	0.0305	0.1122	0.27	0.786
t^2	−0.002489	0.001262	−1.97	0.049
t^3	0.00001413	0.00000525	2.69	0.007
t^4	−0.00000002	0.00000001	−2.10	0.036

S = 10.9445 R-Sq = 87.5% R-Sq(adj) = 87.3%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	296,717	74179	619.28	0.000
Residual Error	355	42,523	120		
Total	359	339,240			

Appendix C

One-Sample T: Additive Y_t

Test of mu = 0 vs not = 0

Variable	N	Mean	StDev	SE Mean	95% CI	T	P
99Additive Yt	360	−0.000001	11.308743	0.596023	(−1.172136, 1.172135)	-9.32×10^{-7}	1.000

Appendix D

Trend Analysis for NSCPO (Quadratic Trend $\hat{X}_t$)

Data	NSCPO
Length	360
NMissing	0

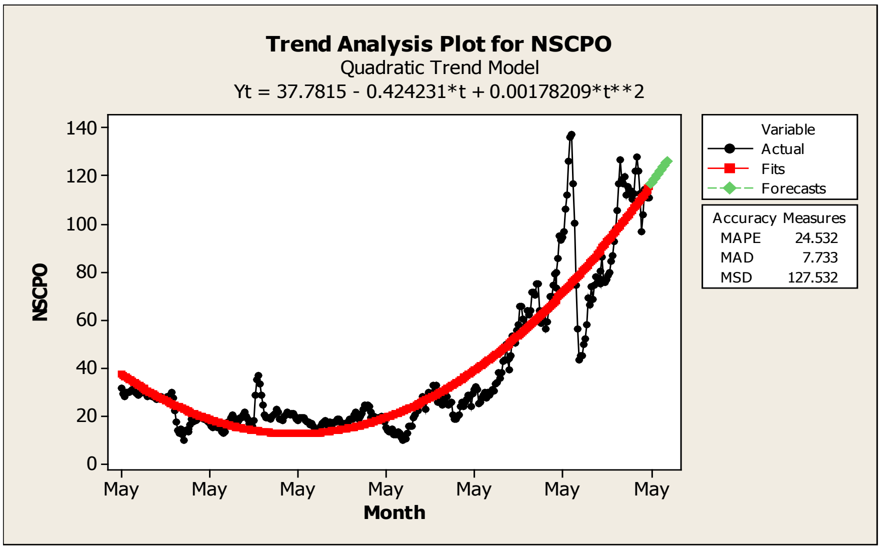
Fitted Trend Equation

$$Y_t = 37.7815 - 0.424231t + 0.00178209t^2$$

Accuracy	Measures
MAPE	24.532
MAD	7.733
MSD	127.532

Forecasts for 2013

Period	Forecast
May	116.878
Jun	117.742
Jul	118.610
Aug	119.481
Sep	120.356
Oct	121.235
Nov	122.117
Dec	123.002
Jan	123.892
Feb	124.784
Mar	125.681
Apr	126.580



AR(2) Y_t Forecasts for 2013

ARIMA Model: Additive Y_t

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	1.3389	0.0477	28.08	0.000
AR	2	−0.4315	0.0477	−9.05	0.000

Number of observations: 360

Residuals: SS = 4685.13 (backforecasts excluded)

MS = 13.09 DF = 358

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

Lag	12	24	36	48
Chi-Square	14.8	32.8	65.3	78.0
DF	10	22	34	46
P-Value	0.141	0.064	0.001	0.002

Forecasts from period 360

		95 Percent Limits		
Period	Forecast	Lower	Upper	Actual
361	−4.6296	−11.7215	2.4623	
362	−4.1159	−15.9676	7.7357	
363	−3.5134	−18.7995	11.7726	
364	−2.9284	−20.5811	14.7243	
365	−2.4050	−21.6465	16.8365	
366	−1.9566	−22.2478	18.3346	
367	−1.5821	−22.5595	19.3953	
368	−1.2741	−22.6970	20.1487	
369	−1.0233	−22.7339	20.6872	
370	−0.8204	−22.7162	21.0753	
371	−0.6570	−22.6716	21.3577	
372	−0.5257	−22.6165	21.5652	