FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 9 No. 1, January, 2025, pp 41 - 45
DOI: https://doi.org/10.33003/fjs-2025-0901-2785

SELECTING AN ADEQUATE MODEL FOR TIME SERIES DECOMPOSITION WHEN THE TREND CURVE IS QUADRATIC

***Enegesele Dennis, Eriyeva Godspower and Ejemah Thomas**

Department of Statistics, Delta State University of Science and Technology, Ozoro, Delta State

*Corresponding authors' email: den4_js@yahoo.com

ABSTRACT

The Buys-Ballot (B-B) approach for the decomposition of additive and multiplicative models in descriptive time series (TS) was examined in this paper. The selection of an adequate model is very important as it shows the underlying structure of the series because the fitted model will be used for future forecasting. Mis-specifying the model characteristics of the data is consequential and can result in biased tests and false predictions The Buys-Ballot method was demonstrated for the selection of an appropriate model and a statistical test that will aid in the selection between additive and multiplicative models was proposed when the trending curve is quadratic. The model identified was used for the forecast. Using the B-B technique, the column variances for the additive model do not contain the seasonal effect while that of the multiplicative model contains the seasonal effect. This distinction was then applied to select between additive and multiplicative models. The chi-square test was proposed for the selection between additive and multiplicative models for the decomposition of TS data. The results when applied to a quadratic trend curve reveal that the appropriate model for the decomposition of the data is the additive model as all calculated chi-square values are within the chi-square acceptance region based on a two-tail test at a 1% level of significance. The additive model identified was then used to decompose the series and the trend analysis model was used for the forecast of the series. The chi-square test was proposed to justify the Buys-Ballot method for distinguishing when a model should be decomposed with an additive or multiplicative model in time series decomposition.

Keywords: Buys-Ballot, Seasonal effect, Column Variances, Additive Model, Decomposition, Forecasting, Chi-Square

INTRODUCTION

Identifying the pattern as portrayed by the sequence of TS plot and forecasting are the two specific goals of TS analysis. These goals demand that time series data patterns be identified and completely described (Box and Jenkins, 1976; Brockwell and Davis 2002; Iwueze and Nwogu, 2014). In addition to studying the nature of the data, the goals are better implemented if the proper model is used for the analysis.

The decomposition of a time series involves splitting the observed TS into components denoted by trend (P_t), seasonal (S_t), cyclical (C_t), and irregular(e_t)components. Chatfield (2004) noted that for short series, the cyclical component can be included in the trend. The TS ($Y_t, t = 1,2,...,n$)decomposition can the be referred to as trend-cycle (P_t)seasonal (S_t) and residual (e_t) components.

Time series models can either be decomposed using additive, multiplicative, or mixed (combining both models). In this paper, the emphasis will be on additive and multiplicative models respectively. These are;

Additive: $Y_t = P_t + S_t + e_t$ (1)

Multiplicative : $Y_t = P_t \times S_t \times e_t$ (2)

The time point is t, P_t is the trend-cycle; S_t the seasonal, and $e_t$the random components. It is assumed that for model (1), e_t the error component is Gaussian $N(0, \sigma_1^2)$ white noise, and the summation of the seasonal component over a complete cycle is zero,($\sum_{j=0}^{s} S_j = 0$). It is also assumed that for the multiplicative model (2), $e_t$is the Gaussian $N(1, \sigma_2^2)$ white noise and for a complete cycle, the summation of the seasonal component is equal to the period i.e ($\sum_{j=0}^{s} S_j = s$)

A major concern in TS decomposition is the identification of the right model. The time series plot of the entire series is used to differentiate between the additive and multiplicative models. In some TS plots, the additive model is applicable if the amplitude of both the seasonal and irregular variations is constant as the level of the trend rises or falls. However, when there is a direct increase in the level of trend between the seasonal and irregular variations, the multiplicative model should be adopted,(Chatfield, 2004). There are occasions when such a plot (graph) is difficult to interpret and therefore may not be easy to ascertain whether a series is a multiplicative or additive model. This difficulty in the interpretation of the graph is what this paper intends to address.

Iwueze et al (2011) noted that the selection of a desired model can be achieved through the interrelationship of the seasonal means and the seasonal standard deviation. They noted that the graph for both means and standard deviation can aid the selection of an appropriate model for decomposition. However, no statistical test was provided for the selection of an adequate model. The method of seasonal quotient and difference as proposed by Justo and Rivera (2010) noted that when the seasonal quotient for the coefficient of variation is higher than the seasonal difference, the additive model should be used for decomposition otherwise the model is multiplicative is applicable. This method also failed to provide any statistical test to justify the choice of model for decomposition.

In the literature, more researchers have started to use machine learning (ML) or a more hybrid model to complete time series prediction tasks. Generally, the Hybrid prediction models consist of two parts: signal decomposition and signal prediction. The most commonly used signal decomposition methods include EMD, EEMD, VMD, etc. Hao *et al*. (2024) applied a decomposition-guided mechanism to inherit the advantages of the decomposition method and ML methods for nonstationary time series prediction without introducing the end-effect problem in the hybrid model in time series forecasting. Jan and Anna (2019) studied the applicability and usefulness of time series decomposition in analyzing the

changeability in timber prices and supply in Poland. They employed the time series multiplicative model. The elements of the time series were determined utilizing the Census X11 method, while cyclicality was separated from the trend employing the Hodrick–Prescott filter.
Atajeromavwo *et al.* (2024) studied the estimation of oil spillage and salvage revenue in Kokori oil field using numerical method and python algorithm. They developed a linear model and compared it with the trapezoidal method and salvage revenue. They emphasis on the need of adopting accurate estimation model for environmental and economic purpose.
Iwueze and Nwogu (2014) presented a guide for the selection of an adequate model that depends on the row, column, and overall averages and variances of the B-B table. They noted that for the additive model, the column variances are the trending curve of the TS while for the multiplicative model, it is the square of the seasonal effect and the product of the trending curve. Hence, the problem of selecting an appropriate model reduces to testing the hypothesis that the column variance of the B-B table is equal to the trending curve. In particular, when there is no upward or downward trend, the problem reduces to test for constant variance.
The objective of this paper is to provide a statistical test that will aid in differentiating between additive and multiplicative models when the trending curve is quadratic. The significance of this paper is to fill the research gap experienced in the literature by proposing a statistical test that will aid in the model identification for decomposition. This paper will mainly focus on the additive and multiplicative model for decomposition in TS analysis.

MATERIALS AND METHODS

Buys-Ballot (B-B) Table

The B-B table can be used to ascertain the effect of seasonal variation in time series data. To analyse the data, Iwueze and Nwogu (2014) posited that it is important to involve the periodic and seasonal totals$\left(T_{i.} and T_{.j}\right)$, period and seasonal averages$\left(\bar{Y}_{i.} and \bar{Y}_{.j}\right)$, and the overall total and mean$\left(T_{..} and \bar{Y}_{..}\right)$
According to Wei (1989), Wold (1938) credits these arrangements of time series data into a two-dimensional table to Buys-Ballot (1847). Thus, in the literature, the table is known as the Buys-Ballot table (as shown in Table 1).
This paper will adopt the method proposed by Iwueze and Nwogu (2014).

Table 1: Buys-Ballot (B-B) Table

Months(i)	Seasons								
	1	**2**	**…**	**j**	**…**	**s**	T_i	$\bar{Y}_{i.}$	$\hat{\sigma}_{i.}$
1	Y_1	Y_2	…	Y_j	…	Y_s	T_1	$\bar{Y}_1$	$\breve{\sigma}_1$
2	Y_{s+1}	Y_{s+2}	…	Y_{s+j}	…	Y_{2s}	T_2	$\bar{Y}_2$	$\breve{\sigma}_2$
3	Y_{2s+1}	Y_{2s+2}	…	Y_{2s+j}	…	Y_{3s}	T_3	$\bar{Y}_3$	$\breve{\sigma}_3$
…	…	…	…	…	…	…	…	…	…
k	$Y_{(k-1)s+1}$	$Y_{(k-1)s+2}$	…	$Y_{(k-1)s+j}$	…	$Y_{(k-1)s+s}$	T_k	$\bar{Y}_k$	$\breve{\sigma}_k$
…	…	…	…	…	…	…	…	…	…
$n.$	$Y_{(n-1)s+1}$	$Y_{(n-1)s+2}$	…	$Y_{(n-1)s+j}$	…	Y_{ns}	T_n	$\bar{Y}_{n.}$	$\breve{\sigma}_n$
$\bar{Y}_{.j}$	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$	…	$\bar{Y}_{.j}$	…	$\bar{Y}_{.s}$	-	$\bar{Y}_{..}$	-
$\breve{\sigma}_{.j}$	$\breve{\sigma}_{.1}$	$\breve{\sigma}_{.2}$	…	$\breve{\sigma}_{.j}$	…	$\breve{\sigma}_{.s}$	-	-	$\breve{\sigma}_{..}$

Iwueze and Nwogu (2014)

$where$
$\bar{Y}_{.j} = \frac{1}{n}\sum_{k=1}^{n} Y_{(k-1)s+j}, k = 1,2,\ldots,$ is the average
$\breve{\sigma}_{.j} = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}\left(Y_{(k-1)s+j} - \bar{Y}_{.j}\right)^2},$ is the standard deviation
$where\ Y_t, n, s and m = ks$ is the series, periods, periodicity, and total sample size respectively
Definition: let $\bar{Y}_{.j}$, be the column mean and $\sigma_{.j}^2$ be column variances for additive model

The TS observation at time t is given by
$Y_t = a + bt + ct^2 + S_t + e_t.$
The expression $t = (k-1)s + j$ can be written in the form of the row (k) and column (j) of the B-B table. The results of the B- B techniques are:

$$\bar{Y}_{.j} = a + \frac{b}{2}(m-s) + \frac{c}{6}(2m-s)(m-s) + [b + c(m-s)]j + cj^2 + S_j + \bar{e}_{.j}$$

$$\breve{\sigma}_{.j}^2 = \frac{m(m+s)}{180}[(2m-s)(8m-11s)c^2 + 30(m-s)bc + 15b^2] + \frac{m(m+s)}{3}[(m-s)c^2 + bc]j + \left[\frac{m(m+s)c^2}{3}\right]j^2 + \sigma_1^2$$

For the multiplicative model:
Let $\bar{Y}_{.j}$ be the column mean and $\sigma_{.j}^2$ be column variances for multiplicative model
The TS observation at time t is given by
$Y_t = (a + bt + ct^2) * S_t * e_t.$
The expression $t = (k-1)s + j$ can be in the form of the row (k) and column (j) of the B-B table. The results of the B- B techniques are:

$$\bar{Y}_{.j} = \begin{Bmatrix}[a + bj + cj^2]e_{ij} + s[b + 2cj](k-1)e_{ij} \\ cs^2(k-1)^2 e_{ij}\end{Bmatrix} * S_j$$

$$\sigma_{.j}^2 = \begin{Bmatrix} a^2 + \frac{c^2}{30}[6m^4 - 15m^3 s + 10m^2 s^2 - s^4] + 2abj + [b^2 + 2ac]j^2 + 2bcj^3 + c^2 j^4 \\ + \frac{(m-s)(b+2cj)}{6}[(b+2cj)(2m-s) + 3cm(m-s)] \\ + \frac{(m-s)(a+bj+cj^2)}{3}[3(b+2cj) + c(2m-s)] \end{Bmatrix} \sigma_2^2 S_{.j}^2$$

Table 2: Summary of Row and Column Variances for Additive and Multiplicative Models

	Linear trend-cycle component: $N_t = a + bt + ct^2, t = 1,2,\ldots nm = ks$	
	Additive	**Multiplicative**
$\bar{Y}_{.j}$	$a + \frac{b}{2}(m-s) + \frac{c}{6}(2m-s)(m-s)$ $[b + c(m-s)]j + cj^2 + S_j + \bar{e}_{.j}$	$\begin{Bmatrix}[a + bj + cj^2]\bar{e}_{.j} + s[b + 2cj](k-1)\bar{e}_{.j'} \\ cs^2(k-1)^2\bar{e}_{.j''}\end{Bmatrix} * S_j$
$\sigma^2_{.j.}$	$\frac{m(m+s)}{180}[(2m-s)(8m-11s)c^2 + 30(m-s)bc + 15b^2]$ $+\frac{m(m+s)}{3}[(m-s)c^2 + bc]j + \left[\frac{m(m+s)c^2}{3}\right]j^2 + \sigma_1^2$	$\begin{Bmatrix} a^2 + \frac{c^2}{30}[6m^4 - 15m^3s + 10n^{m2}s^2 - s^4] + 2abj + [b^2 + 2ac]j^2 \\ +2bcj^3 + c^2j^4 + \frac{(m-s)(b+2cj)}{6}[(b+2cj)(2n-s) + 3cm(m-s)] \\ +\frac{(m-s)(a+bj+cj^2)}{3}[3(b+2cj) + c(2m-s)] \end{Bmatrix}\sigma_2^2 S_j^2$

$Footnote: \bar{e}_{.j'} = \frac{1}{n}\sum_{i=1}^{n}(k-1)e_{ij}\ , \bar{e}_{.j''} = \frac{1}{n}\sum_{i=1}^{n}(k-1)^2\, e_{ij}$
$\sigma_1^2 = ErrorVariance(Additive\ mod\ e\ l), \sigma_2^2 = ErrorVariance(Multiplicative\ mod\ e\ l)$

The difficulty in selecting an appropriate model for decomposition in a quadratic trend component reduces to testing the null hypothesis:
$H_0: \sigma_j^2 = \xi_j(\theta), j = 1,2,\ldots s$
and the jth column variance is simply the trend value. Against the alternative
$H_1: \sigma_j^2 \neq \xi_j(\theta)$,for at least one j
and the jth column variance is different from the trend value. For the quadratic trend curve, the column variance is given as
$\hat{\sigma}_{.j}^2 = (\theta_0 + \theta_1 j + \theta_2 j^2) + \sigma_1^2, for additive\ mod\ e\ l$
$where$
$\theta_0 = \frac{m(m+s)}{180}[(2m-s)(8m-11s)c^2 + 30(m-s)bc + 15b^2]$
$\theta_1 = \frac{m(m+s)}{3}[(m-s)c^2 + bc]\ ,\ \theta_2 = \left[\frac{m(m+s)c^2}{3}\right]$
The test statistic is set to be the hypothesized variance of each column divided by the nominal value of the variance (ie the value to be tested). The statistic has a Chi-Square distribution with n-1 degree of freedom. Hence, the Chi-Square test statistic was proposed in this paper and the test for the null hypothesis is given by

$$\chi^2 = \frac{(n-1)\hat{\sigma}_j^2}{\sigma_0^2} \quad (3)$$

The observations in each column are represented by n,$\hat{\sigma}_j^2$ which is the jth column computed variances and $\sigma_0^2 = \xi_j(\theta)$is the hypothesized variance for each column. Under the null hypothesis, the statistic in (3) complies with the Chi-Square distribution with n-1 degrees of freedom (Spiegel, 1975; Mood et al., 1974). An assumption of the Chi-Square test used in this paper is based on its asymptotic null distribution. Based on this two-tailed test at α level of significance, the null hypothesis would not be rejected if
$\chi^2_{\alpha/2} < \frac{(n-1)\hat{\sigma}_j^2}{\sigma_0^2} < \chi^2_{1-\alpha/2}$
or rejected otherwise.
Thus, the appropriate model is additive when the null hypothesis is not rejected or multiplicative when it is rejected.

Empirical Examples

The data used for the illustrative example is Nigeria's Spot component price of oil (US Dollar per Barrel) from 1991-2021. The characteristics of the data from figure 1 shows that the original data has a quadratic trend curve. Thus, the application of B-B table technique was used to ascertain if the additive or multiplicative models should be adopted for TS decomposition. The column variances and the trend values using the B-B table method are shown below.

$$\sigma_{.j}^2 = \begin{pmatrix}921.27\\1038.09\\1157.35\\1221.97\\1168.87\\1199.73\\1279.41\\1177.37\\1092.78\\1017.79\\1032.34\\1042.91\end{pmatrix}, \sigma_0^2 = \begin{pmatrix}969.5232\\1049.123\\1112.642\\1160.078\\1191.432\\1206.705\\1205.896\\1189.005\\1156.032\\1106.978\\1041.841\\960.6232\end{pmatrix}, \frac{(n-1)(\sigma_{.j}^2)}{\sigma_0^2} = \begin{pmatrix}28.51\\29.68\\31.21\\31.60\\29.43\\29.83\\31.83\\29.71\\28.36\\27.58\\29.73\\32.57\end{pmatrix}$$

At 0.01, under the null hypothesis, the chi-square is stated as

$$\chi^2: 13.79 \leq (n-1)\frac{\sigma_{.j}^2}{\sigma_0^2} \leq 53.67, j = 1,2,\ldots,12$$

The numerical example shown above shows that all calculated chi-square values are within the acceptance region $13.79 \leq (n-1)\frac{\sigma_{.j}^2}{\sigma_0^2} \leq 53.67$ at 1% level of significance indicating that the additive model is the necessary model for the decomposition of the data

Decomposition and Forecast

The additive model identified was used to decompose the series. The P--P plot in Figure 3 shows that the residual seems to be normally distributed indicating that an adequate model was used for decomposition. Also, most of the data points lie along the normal line and those outside the line deviate in a similar pattern below and above the normal line. The quadratic trend model was used for the forecast. The original TS plot and trend analysis are given in Figures 1 and 2 respectively

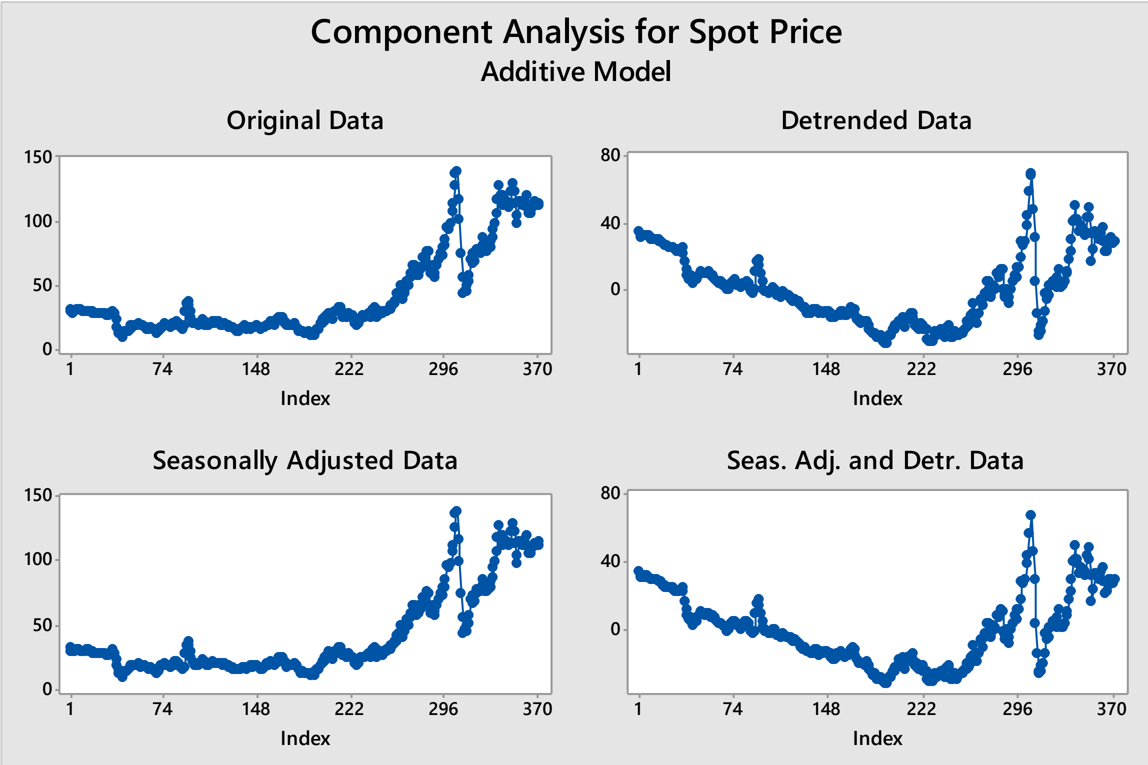

Figure 1: Original Series of Nigeria Spot Component Price

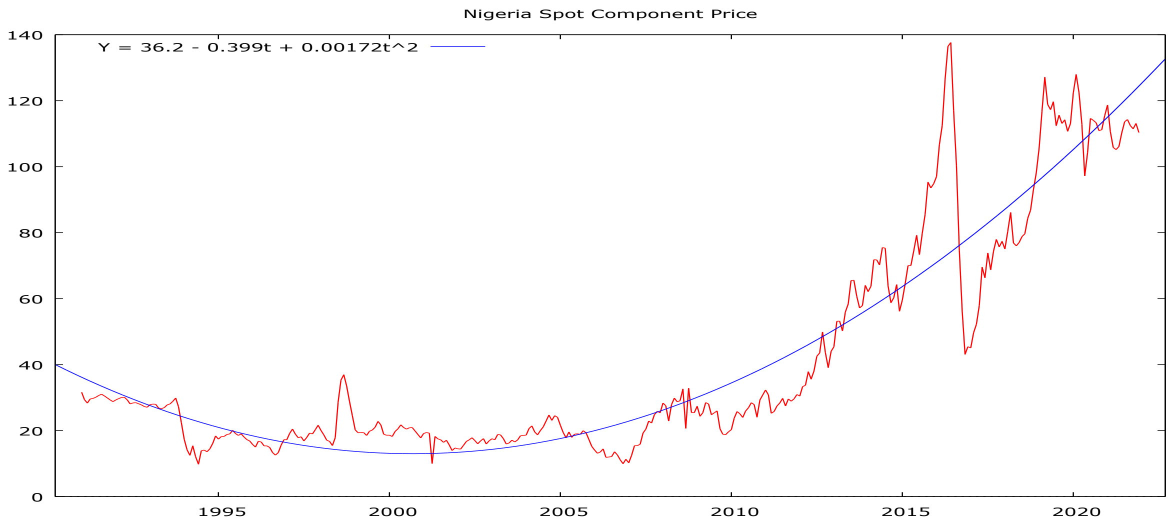

Figure 2: Fitted Quadratic Trend Curve

Fitted Trend Equation

Yt = 36.88 - 0.4047×t + 0.001716×t^2

Months	Forecast for 2022
January	124.640
February	125.517
March	126.397
April	127.281
May	128.168
June	129.059
July	129.953
August	130.851
September	131.752
October	132.656
November	133.564
December	134.476

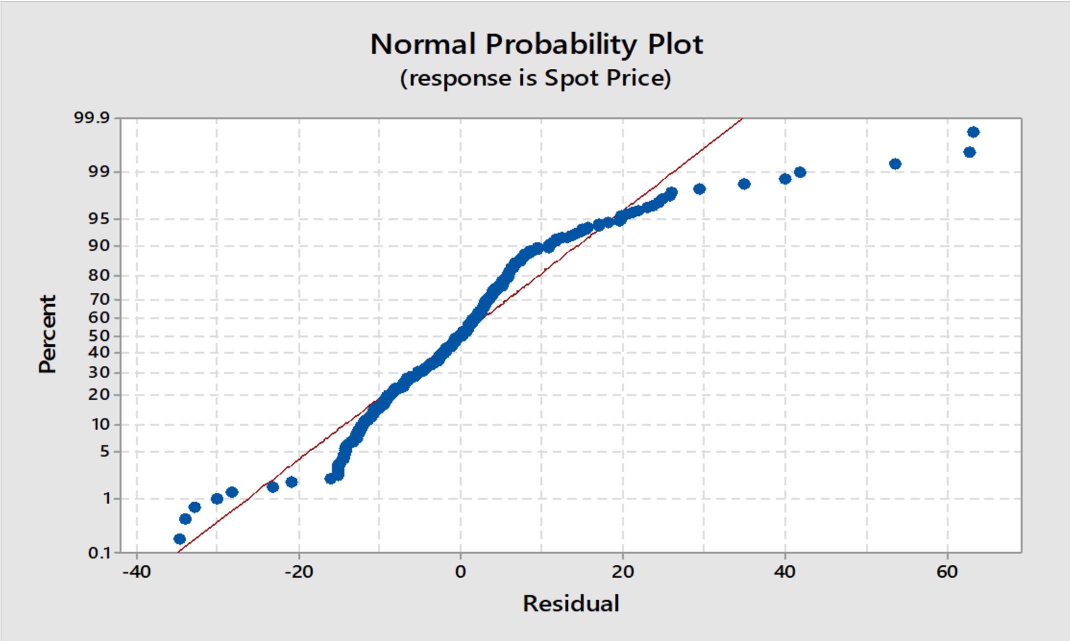

Figure 3: Normal Probability Plot of Decomposed Residual

CONCLUSION

A main concern in the application of descriptive TS analysis is that of the selection of a model for the decomposition of TS data. From the available literature, there is no definite test for the selection of a model. The B-B procedure reveals that (a) the additive and multiplicative models have different column mean and variances, and (b) for the additive and multiplicative models, the column variances imitate the shape of the trending series when the trend curve is quadratic. However, while the column variances of the additive model exclude the seasonal component, the column variances of the multiplicative model contain the seasonal component. This characteristic is what has been used to develop the test for the choice of an adequate model for decomposition. The problem of choice between additive and multiplicative models reduces to a Chi-Square test for variances based on a two-tail test at α level of significance. The calculated chi-square values are within the acceptance region $13.79 \leq (m-1)\frac{\sigma_j^2}{\sigma_0^2} \leq 53.67$ at 1% level of significance. Thus, the additive model was recommended. The normal probability plot shows that an appropriate decomposition model was used to decompose the series. The regression trend equation was then used for the forecast. Decomposition often plays a vital role in making time series better as well as improving the forecast. The chi-square test proposed in this paper justify the Buys-Ballot method for distinguishing when a model should be decomposed with an additive or multiplicative model when the trend curve is quadratic. The Buys-Ballot method can be extended to other series whose trend curve is exponential.

REFERENCES

Atajeromavwo Edafe John, Okiemute Dickson Ofuyekpone, and Rume Elizabeth Yoro (2024). Estimation of Oil Spillage and Salvage Revenue in Kokori Oil Field using Numerical methods and Python Algorithm. FUDMA Journal of Sciences (fjs) vol 8(5).

Box G. E. P and Jenkins M, (1976) Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco.

Brockwell P.J and Davis R.A (2002). "Introduction to Time Series and Forecasting 2nd Edition, Springer New York.

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

Iwueze I. S, Nwogu E.C, Ohakwe J. and Ajaraogu J.C (2011), "Uses of the Buys-Ballot Table in Time Series Analysis", Applied Mathematics, pg 633-645.

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, Vol. 48 No. 1, pp 45-66.

Hao Wang, Lubna AI Tarawneh, Changqing Cheng, and Yu Jin (2024). A decomposition-guided mechanism for non-stationary time series forecasting. AIP Advances 14, 015254. https://doi.org/10.1063/5.0153647 .

Jan Banas and Anna Kozuch (2019). The Application of Time Series Decomposition for the Identification and Analysis of fluctuation in Timber supply and price. Forest 10(11) 990; https://doi.org/10.3390/f10110990 .

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

Mood, A., Graybill, F. and Boes, D. (1974) Introduction to the Theory of Statistics. 3rd Edition, McGraw-Hill, New York, 122-123.

Spiegel, M.R. (1975) Probability and Statistics. McGraw-Hill, New York, 372.

Wei W.W.W, (1989): Time Series Analysis, Univariate and Multivariate Methods, Addison- Wesley, Redwood City.