International Journal of Probability and Statistics 2020, 9(1): 1-6
DOI: 10.5923/j.ijps.20200901.01

The Analysis of Mixed Model in Time Series Decomposition

Kelechukwu C. N. Dozie*, Hope I. Mbachu, Mirian C. Raymond

Department of Statistics Imo State University Owerri, Imo State, Nigeria

Abstract This paper examines the condition(s) under which the mixed model is the most appropriate model in descriptive time series analysis when trend-cycle component is linear. Since some existing studies have adequately characterized the additive model, the analyst is still exposed to the risk of wrongly using multiplicative model when mixed model should be used. Therefore, the aim of this study is to identify the series that admits mixed model. The method employed in this paper is the Buys-Ballot procedure developed for choice of model and choice of appropriate transformations, among other uses, based on row, column and over totals, means and variances of the Buys-Ballot table. Results show that, 1) the seasonal variance of the Buys-Ballot table, for the mixed model, a function of the slope and seasonal effect only. 2) the calculated values of the test statistic identified the mixed model correctly in 99 out the 100 stimulations.

Keywords Choice of Model, Time Series Decomposition, Mixed Model, Buys-Ballot Table

1. Introduction

The objective of time series decomposition is to separate the four time series components available in the series. That is, to de-compose an observed time series $\left(X_t, t = 1, 2, ..., n \right)$ into components, representing the trend (T_t), the seasonal (S_t), cyclical (C_t) and irregular (e_t) Kendal and Ord [9], Chatfield [2].

The models most commonly used to describe time series decomposition are the

Additive Model:

$$X_t = T_t + S_t + C_t + e_t \tag{1}$$

Multiplicative Model:

$$X_t = T_t \times S_t \times C_t \times e_t \tag{2}$$

and Mixed Model

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

for short series, the trend component is jointly estimated into the cyclical Chatfield [2] and the observed time series $\left(X_t, t = 1, 2, ..., n \right)$ can be decomposed into the trend-cycle component (M_t), seasonal component (S_t) and the irregular/residual component (e_t). Therefore, the decomposition models are

Additive Model:

$$X_t = M_t + S_t + e_t \tag{4}$$

Multiplicative Model:

$$X_t = M_t \times S_t \times e_t \tag{5}$$

and Mixed Model

$$X_t = M_t \times S_t + e_t . \tag{6}$$

It is always assumed that the seasonal effect, when it exists, has period *s*, that is, it repeats after *s* time periods.

$$S_{t+s} = S_t , \ for \ all \ t \tag{7}$$

For additive model, it is convenient to make assumption that the sum of the seasonal components over a complete period is zero, ie,

$$\sum_{j=1}^{s} S_{t+j} = 0 . \tag{8}$$

Similarly, for multiplicative and mixed models, it equally convenient to make assumption that the sum of the seasonal components over a complete period is *s*.

$$\sum_{j=1}^{s} S_{t+j} = s . \tag{9}$$

It is also assumed that the error term e_t is the Gaussian $N\left(0, \sigma_1^2 \right)$ white noise for additive and mixed models,

* Corresponding author:
kcndozie@yahoo.com (Kelechukwu C. N. Dozie)
Published online at http://journal.sapub.org/ijps

while for multiplicative model, e_t is the Gaussian $N\left(1, \sigma_2^2\right)$ white noise and that $Cov\left(e_t, e_{t+k}\right) = 0$, $\forall\ k \neq 0$.

It is assumed that (i) the appropriate model for decomposition is known; (ii) the study series satisfied the assumptions of the models. However, one of the greatest problems identified in the use of descriptive method of time series analysis is choice of appropriate model for decomposition of any study data. That is when to use any of the additive, multiplicative and mixed models for analysis is uncertain. And it is clear that; use of wrong model will definitely lead to erroneous estimates of the components. The emphasis of this paper is to identify the series that admits mixed model. This work is restricted to time series with linear using stimulated series of 100 data sets of 120 observations each stimulated from the mixed model and real life data on number of Hennessey drinks sold at Vandoz Enterprise, Owerri, Imo Stata, Nigeria for the period 2008 to 2017.

Iwueze and Nwogu [7] provided a framework for choice of model and detection of seasonal effect in time series. The framework shows that the column (seasonal) variances of the Buys-Ballot table are, for the additive model, functions of only trend parameters and for multiplicative model, functions of the trend parameters and seasonal indices. In particular, when the trend-cycle component is linear, the column variances are constant for additive model, but contain the seasonal component for the multiplicative model. Therefore, choice between additive and multiplicative models reduces to test for constant variance to identify the additive model. Thus, they suggested that any of the test for constant variance can be used to identify a series that admits the additive model. This is an improvement of what is available in the literature. However, this approach can only identify the additive model, when the column variance is constant, but does not suggest to the analyst the alternative model when the variance is not constant. The implication of this is that when the test for constant variance says the appropriate model is not the additive model; an analyst still faces the challenge of choosing between mixed model and the multiplicative model.

1.1. Buys-Ballot Procedure for Time Series Decomposition

Iwueze, *et al,* [8] observed that, for time series which contain a seasonal effect, the overall mean $(\bar{X}_{..})$ and seasonal means $\left(\bar{X}_{.j},\ j = 1, 2, ..., s\right)$ of the Buys – Ballot table are used to assess the effects either as a difference $\left(\bar{X}_{.j} - \bar{X}_{..}\right)$ or the ratio $\left(\frac{\bar{X}_{.j}}{\bar{X}_{..}}\right)$. That is, the deviation of the differences seasonal averages and the overall averages (additive model) from zero or the overall average from unity (multiplicative model) is used to assess the presence of seasonal effects. The Buys - Ballot table helps in the assessment of the trend – cycle and seasonal effect of time series data. The row means $(\bar{X}_{i.})$ estimate trend, and the differences $(\bar{X}_{.j} - \bar{X}_{..})$ or the ratio $\left(\frac{\bar{X}_{.j}}{\bar{X}_{..}}\right)$ between the column means $(\bar{X}_{.j})$ and the overall mean $(\bar{X}_{..})$ estimate the seasonal effects. Chartfield [2] stated the use of the Buys - Ballot table for inspecting time series data for the presence of trend and seasonal effects. Iwueze and Nwogu [4] developed a new estimation procedure based on row, column and overall averages of the Buys - Ballot table. This method called Buys – Ballot estimation procedure uses the periodic mean $(\bar{X}_{i.}, i = 1,2, \dots, m)$ and the overall mean $(\bar{X}_{..})$ to estimate the trend component. Seasonal means $(\bar{X}_{.j}, j = 1,2, \dots, s)$ and the overall mean $(\bar{X}_{..})$ are used to estimate the seasonal indices. Fomby [3] presented various graphs suggested by the Buys – Ballot table for inspecting time series data for the presence of seasonal effects.

2. Methodology

This section presents methodology adopted in this paper. The method adopted is the Buys-Ballot procedure for time series decomposition. This procedure has been developed for choice of transformation and choice of model, among other uses, based on the row, column and overall means and variances of the Buys-Ballot table. For detailed discussion of Buys-Ballot procedure, see Wei [11], Iwueze and Nwogu [4] and [5] Iwueze and Ohakwe [7]. The row, column and overall means and variance obtained by

Nwogu, *et al*, [10] are given in Table 1. From Table 1, it is clear that the column variances of the Buys-Ballot table, a constant multiple of square of the seasonal effect only for the mixed model. Nwogu, *et al*, [10] proposed chi-square test for choice between the mixed and the multiplicative models is based on the column variances. The proposed test is able to distinguish a series that admits mixed model from a series that admit multiplicative model. According to them, the null hypothesis to be tested is

$$H_0:\ \sigma_j^2 = \sigma_{0j}^2$$

and the appropriate model is mixed, against the alternative

$$H_1:\ \sigma_j^2 \neq \sigma_{0j}^2$$

and the appropriate model is not mixed, where

$\sigma_j^2 = \left(j = 1, 2, ..., s\right)$ is the actual variance of the jth column.

$$\sigma_{0j}^2 = \frac{b^2 n(n+s)}{12} S_j^2 + \sigma_1^2 \qquad (10)$$

and σ_1^2 is the error variance, assumed equal to 1.

They stated the statistic

$$\chi_c^2 = \frac{(m-1)\sigma_j^2}{\sigma_{0j}^2} \qquad (11)$$

follows the chi-square distribution with $m-1$ degrees of freedom, m is the number of observations in each column and s is the seasonal lag (number of columns). They also showed that under the null hypothesis, the interval $\left[\chi^2_{\frac{\alpha}{2},(m-1)}, \chi^2_{1-\frac{\alpha}{2},(m-1)}\right]$ contains the statistic (11) with 100 (1- α)% degree of confidence.

2.1. Test for Constant Variance

Although there are many tests for constant variance, Bartlett's test has been adopted in this work. Bartlett's test is more robust. Bartlett's test allows the comparison of variance of two or more samples to determine whether they are drawn from populations with equal variance. It is suitable for normally distributed data. To test the null hypothesis that the variances are equal, that is

$$H_0 : \sigma_i^2 = \sigma_j^2$$

against the alternative

$$H_1 : \sigma_i^2 \neq \sigma_j^2 \quad \text{for} \quad i \neq j$$

and at least one variance is different from others

Bartlett [1] has shown that the statistic

$$T = \frac{(N-k)\ln S_p^2 - \sum (N_i - 1)\ln S_i^2}{1 + \frac{1}{3(k-1)}\left[\sum_{i=1}^{k} \frac{1}{(N_i-1)} - \frac{1}{N-k}\right]} \tag{12}$$

follows Chi-square distribution with (k – 1) degrees of freedom. Using the parameters of the Buys-Ballot table, $N = ms,\ k = s,\ N_i = m$, the statistic in (12) is then given as

$$T_c = \frac{(ms-s)\ln \hat{\sigma}_p^2 - \sum (m-1)\ln \hat{\sigma}_j^2}{1 + \frac{1}{3(s-1)}\left[\sum_{j=1}^{s} \frac{1}{m-1} - \frac{1}{ms-s}\right]}$$

$$= \frac{(m-1)\left[s \ln \hat{\sigma}_p^2 - \sum \ln \hat{\sigma}_j^2\right]}{1 + \frac{(s+1)}{3s(m-1)}}$$

Where ms is the total number of observations, m is the number of observations in each column and s is length of the periodic interval.

Table 1. Summary of Row, Column and Overall Means and Variances of Buys-Ballot for Mixed Model

Measures	Linear trend-cycle component: $M_t = a + bt$, $t = 1, 2, ..., n = ms$
	Mixed model
$\bar{X}_{i.}$	$\left[a - bs + bsi\right] + \frac{b}{s}\sum_{j=1}^{s} jS_j + \bar{e}_{i.}$
$\bar{X}_{.j}$	$\left[a + b\left(\frac{n-s}{2}\right) + bj\right] * S_j + \bar{e}_{.j}$
$\bar{X}_{..}$	$a + b\left(\frac{n-s}{2}\right) + bC_1 + \bar{e}_{..}$
$\hat{\sigma}_{i.}^2$	$\left\{\left[(a + bs(i-1)) + bC_1\right]^2 + \text{var}\left[\left[a + bs(i-1)\right]S_j + bjS_j\right]\right\} + \sigma_1^2$
$\hat{\sigma}_{.j.}^2$	$\frac{b^2 n(n+s)}{12} S_j^2 + \sigma_1^2$
$\hat{\sigma}_x^2$	$\frac{n}{n-1}\left\{\begin{array}{l} \frac{b^2(n^2-s^2)}{12} + \left[a^2 + 2ab\left(\frac{n-s}{2}\right) + \frac{b^2(n-s)(2n-s)}{6}\right] Var(S_j) \\ + 2b\left[a + b\left(\frac{n-s}{2}\right)\right] Cov(S_j, jS_j) + b^2 Var(jS_j) \end{array}\right\} + \sigma_1^2$

Source: Nwogu, et al, (2019)

3. Empirical Examples

The purpose of this section is to present empirical examples to illustrate the application of the proposed test by Nwogu, *et al*, [10]. The empirical example consists of both stimulated series from the mixed model and real life data.

3.1. Simulations Results from Mixed Model

The stimulated series used consist of 100 data sets of 120 observations each stimulated from mixed model $X_t = (a+bt) \times S_t + e_t$, with $a=1,\ b=0.2$, $e_t \sim N(0,1)$ and S_j given in Table 2.

Each series of 120 observations has been arranged in a Buys-Ballot table as monthly data (s = 12) for 10 years (m = 10). The critical values, are for $m-1=9$ degrees of freedom, equal to 2.7 and 19.0 and at 5% level of significance. The decision rule is to reject null hypothesis, if the calculated value of the statistic lie outside the interval otherwise do not reject it. When compared with the interval 2.7 and 19.0, the calculated values of the statistic in Table 3 lie within the interval in 99 out of the 100 simulations. The test identified the mixed model correctly in 99 out of the 100 stimulations.

Table 2. Seasonal (S_j) indices used in the simulation of series

j	1	2	3	4	5	6	7	8	9	10	11	12
S_j	0.94	0.83	0.99	0.98	0.98	1.11	1.24	1.19	1.07	0.93	0.82	0.91

Table 3. Calculated Chi-Square for Mixed Model

Col	Series									
	1	2	3	4	5	6	7	8	9	10
1	8.5054	7.8937	8.8313	10.9478	10.2262	8.5270	8.1754	10.1747	8.6302	11.4643
2	10.4460	10.0819	8.4784	8.7076	7.3992	9.0979	9.3446	9.1427	10.2099	9.9454
3	9.8262	10.6055	9.3075	9.6204	8.8926	9.7582	8.8767	7.9017	9.8372	8.4075
4	9.0616	9.3167	9.2864	8.3987	8.7418	10.4852	7.9597	9.0097	9.0442	8.5360
5	9.0693	8.9255	8.1536	7.4221	8.3907	9.9461	10.2499	9.3821	8.7914	9.4324
6	9.2731	8.2256	10.8535	9.3170	10.2574	8.6887	9.7197	7.8466	9.5678	9.2602
7	9.9262	9.7359	8.3535	7.9367	8.3932	8.7275	8.4997	8.6878	9.0340	9.0346
8	7.8336	8.6354	9.4445	9.8982	8.7871	9.5260	9.5770	10.0049	8.4561	9.3923
9	7.8750	10.0158	8.2101	10.5401	8.6521	8.6745	7.6250	9.2179	8.4077	6.9103
10	8.8117	8.8229	8.8236	8.4856	9.8426	7.5166	10.5594	8.4483	9.2740	8.2414
11	8.0692	7.6919	10.6876	9.6161	9.0376	8.5530	9.0016	10.6676	7.2301	8.9335
12	9.3675	7.8286	7.9817	8.1223	9.1068	8.5058	8.2406	7.8487	9.3524	8.7364
Decision	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept

The critical values for $m-1=9$ degrees of freedom are 2.7 and 19.0

Calculated Chi-Square for Mixed Model Cont'

Col	Series									
	11	12	13	14	15	16	17	18	19	20
1	8.7755	10.5496	10.2908	9.6347	9.5487	9.46020	8.3945	10.1108	10.9183	9.41140
2	10.7630	8.3355	10.3556	7.6020	10.4923	8.23829	9.0907	8.9461	7.9310	7.79464
3	9.8913	9.0127	8.0809	9.8729	9.8590	9.56309	9.5252	8.8988	7.9123	9.72765
4	8.3711	10.3066	10.6504	9.9787	8.4176	9.34434	9.7574	9.6347	7.3661	9.55607
5	7.8818	7.2799	7.7839	8.7815	7.8585	8.91986	7.2466	8.8830	9.6954	9.99632
6	8.6641	9.4888	7.6831	9.8723	8.6037	9.70230	9.5202	8.4158	9.4467	9.03982
7	9.8038	8.5397	8.3298	7.8722	8.6224	7.96241	8.7520	8.7630	8.2131	8.57532
8	9.6343	7.8744	8.6675	8.8280	8.5870	8.13746	8.2659	9.1608	9.1081	9.10432
9	7.9875	9.6823	10.1188	9.2951	9.9179	8.67129	10.0269	8.9280	9.6243	9.53287
10	8.7172	8.7410	9.6714	9.2484	7.6658	9.48001	8.2179	7.8661	9.3762	7.52908
11	8.2105	9.2127	7.9508	10.0388	8.4098	8.89262	8.6526	8.7490	11.9767	8.96591
12	9.5470	8.6815	9.2369	7.5572	10.1038	9.91200	10.3104	9.2197	7.4892	8.53518
Decision	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept	Accept

The critical values for $m-1=9$ degrees of freedom are 2.7 and 19.0

3.2. Real Life Data

The second empirical example is based on monthly data on number of Hennessey drinks sold at Vandoz Enterprise, in Owerri Municipal of Imo State for the period 2008 to 2017. The data is given in Appendix A while the time plot is in Figure 3.1. The column variances are shown in Table 4. The first step is to check whether the data admits additive model. To test the null hypothesis that the data admits additive model, the Bartlett's test is used. The null hypothesis is rejected, if T_c is greater than the tabulated value, which for $\alpha = 0.05$ level of significance and $m-1=9$ degrees of freedom equal to 19.7 or do not reject H_0 otherwise.

From Appendix and Table 4

$$m=10,\ s=12, \hat{\sigma}_p^2 = 16.9707,\ \ln\hat{\sigma}_p^2 = 2.8315 \text{ and}$$

$$\sum_{j=1}^{s} \ln\hat{\sigma}_j^2 = 30.9062$$

Hence,

$$T_c = \frac{(9)\left[12(2.8315)-(30.9062)\right]}{1+\frac{13}{3(12)9}} = 26.5797$$

When compared with the critical value (19.7), T_c is greater, indicating that the data does not admit the additive model. Having confirmed that the data does not admit additive model, the choice now lies between mixed and multiplicative models.

In other to choose between mixed and multiplicative models, the proposed test by Nwogu, *et al*, [10] is used. The null hypothesis that the data admits mixed model is rejected, if the statistic defined in (11) lies outside the interval $\left[\chi^2_{\frac{\alpha}{2},(m-1)}, \chi^2_{1-\frac{\alpha}{2},(m-1)}\right]$ which for m-1 = 9, equals (2.7 and 19.0) or do not reject H_0 otherwise.

From Table 4,

$$\sigma_1^2 = 1,\quad b = 1.051,\quad n = 120,\quad s = 12,\quad m = 10$$

Hence, from (10)

$$\sigma_{0j}^2 = (1.051) \times 120 \left(\frac{120+12}{12}\right) s_j^2 + 1$$

and the calculated values, χ^2_{cal} given in Table 4 were obtained. When compared with the critical values (2.7 and 19.0), the calculated values of the statistic lie within the interval, indicating that the data admits mixed model.

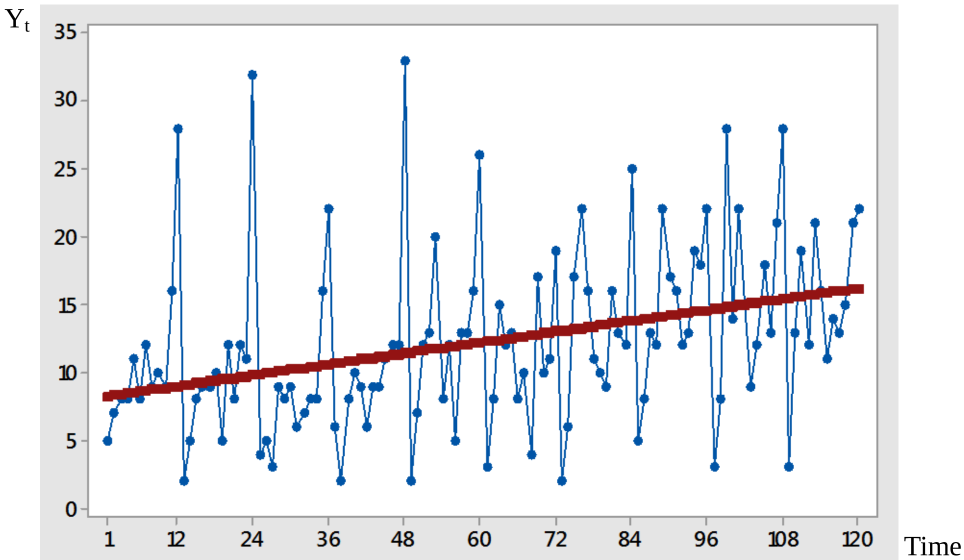

Figure 3.1. Time plot of the actual series on number of Hennessey drinks between (2008-2017)

Table 4. Seasonal effects (S_j), estimate of the column variance ($\hat{\sigma}_j^2$) and Calculated Chi-square $\left(\chi^2_{cal}\right)$

J	1	2	3	4	5	6	7	8	9	10	11	12
S_j	0.54	0.92	1.53	1.26	1.41	0.92	0.78	0.67	0.85	0.78	0.91	1.43
$\hat{\sigma}_j^2$	2.06	8.10	50.77	15.88	33.43	14.84	9.78	10.68	12.46	9.82	14.71	21.12
$\ln\hat{\sigma}_j^2$	0.72	2.09	3.93	2.77	3.51	2.70	2.28	2.37	2.52	2.28	2.69	3.05
χ^2_{cal}	4.67	7.65	18.54	8.40	14.15	13.92	12.22	17.22	13.38	12.39	14.19	8.81

4. Concluding Remarks

This paper has discussed the analysis of mixed model when trend-cycle component of time series is Linear. The emphasis is to identify the series that admit mixed model. The method adopted is Buys-Ballot procedure developed for choice model and choice of appropriate transformation, among others uses, based on row, column and overall totals, means and variances of Buys-Ballot table. Result from calculated value of the proposed test statistic shows that, the test identified mixed model correctly in 99 out of the 100 simulation In considering the mixed model, the residuals are assumed to be (i) uncorrelated; and (ii) normally distributed with mean equal to zero and constant variance. Cases, in which these assumptions are not met, are recommended for further investigation.

Appendix

Buys-Ballot table on number of Hennessey drink at Vandoz Enterprise (2008-2017)

Year	Jan.	Feb.	March	April	May	June	July	August	Sept.	Oct.	Nov.	Dec	$\bar{X}_{i.}$	$\sigma^2_{i.}$
2008	5.00	7.0	8.00	8.00	11.00	8.00	12.00	9.00	10.00	9.00	16.00	28.00	10.92	36.63
2009	2.00	5.0	8.00	9.00	9.00	10.00	5.00	12.00	8.00	12.00	11.00	32.00	10.25	56.02
2010	4.00	5.0	3.00	9.00	8.00	9.00	6.00	7.00	8.00	8.00	16.00	22.00	8.75	28.20
2011	6.00	2.0	8.00	10.00	9.00	6.00	9.00	9.00	11.00	12.00	12.00	33.00	10.58	57.90
2012	2.00	7.0	12.00	13.00	20.00	8.00	12.00	5.00	13.00	13.00	16.00	26.00	12.25	42.57
2013	3.00	8.0	15.00	12.00	13.00	8.00	10.00	4.00	17.00	10.00	11.00	19.00	10.83	23.06
2014	2.00	6.0	17.00	22.00	16.00	11.00	10.00	9.00	16.00	13.00	12.00	25.00	13.25	41.66
2015	5.00	8.0	13.00	12.00	22.00	17.00	16.00	12.00	13.00	19.00	18.00	22.00	14.75	27.48
2016	3.00	8.0	28.00	14.00	22.00	15.00	9.00	12.00	18.00	13.00	21.00	28.00	15.92	60.45
2017	3.00	13.0	19.00	12.00	21.00	16.00	11.00	14.00	13.00	15.00	21.00	22.00	15.00	28.73
$\bar{X}_{.j}$	3.50	6.9	13.10	12.10	15.10	10.80	10.00	9.30	12.70	12.40	15.40	25.70	12.25	*
$\sigma^2_{.j}$	2.06	8.1	50.77	15.88	33.43	14.84	9.77	10.68	12.46	9.82	14.71	21.12		40.27

Source: Vandoz Enterprise, Owerri, Imo State (2008-2017)

REFERENCES

[1] Bartlett, M. S. (1937). *Properties of Sufficiency and Statistical Tests.* Proceedings of the Royal Statistical Society, Series A 160, 268-282 JSTOR 96803.

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

[3] Fomby, T.B (2008). *Buys–Ballot plots: graphical methods for detecting seasonality in time series*. faculty.smu.edu/tfomby/eco5375/data/…/Buys%20Ballot%plotpdf.

[4] Iwueze, I. S. & Nwogu E.C. (2004). *Buys-Ballot estimates for time series decomposition,* Global Journal of Mathematics, 3(2), 83-98.

[5] Iwueze, I. S. & Nwogu, E.C. (2005). *Buys-Ballot estimates for exponential and s-shaped curves, for time series*, Journal of the Nigerian Association of Mathematical Physics, 9, 357-366.

[6] Iwueze, I. S. & Nwogu, E.C. (2014). *Framework for choice of models and detection of seasonal effect in time series.* Far East Journal of Theoretical Statistics 48(1), 45– 66.

[7] Iwueze, I. S. & Ohakwe, J. (2004). *Buys-Ballot estimates when stochastic trend is quadratic*. Journal of the Nigerian Association of Mathematical Physics, 8, 311-318.

[8] Iwueze, I.S., Nwogu, E.C., Ohakwe, J. & Ajaraogu. J.C (2011). *Uses of the Buys – Ballot table in time series analysis*. Applied Mathematics Journal 2, 633 –645.

[9] Kendal, M. G. & Ord, J. K. (1990). *Time Series* (3rd ed.). Charles Griffin, London.

[10] Nwogu, E.C, Iwueze, I.S. Dozie, K.C.N. & Mbachu, H.I (2019). Choice between mixed and multiplicative models in time series decomposition. International.

[11] Wei, W. W. S (1989). *Time series analysis: Univariate and multivariate methods,* Addison-Wesley publishing Company Inc, Redwood City.