**EFFECT OF OUTLIERS ON BUYS-BALLOT ESTIMATES OF TREND PARAMETERS AND SEASONAL INDICES IN DESCRIPTIVE TIME SERIES ANALYSIS**

**Abstract**

*Buys-Ballot methodology was developed to, among other things, obtain estimates of trend parameters and seasonal indices in descriptive time series analysis (Iwueze et al, 2011). In developing this method, Iwueze and Nwogu (2004, 2005), Iwueze and Ohakwe (2005), proposed these estimates (of trend parameters and seasonal indices) were derived from row, column and overall averages and standard deviations of the Buys-Ballot table. These averages and standard deviations were known to be affected by outliers. In their proposed methods, Iwueze and Nwogu (2004, 2005), Iwueze and Ohakwe (2005), made no provision for the effects of these outliers on the proposed estimates. Hence, the purpose of this study was to assess the effects of outliers on Buys-Ballot estimates of trend parameters and seasonal indices. Simulated series were used in this study in the absence and presence of outliers. To capture the presence of outliers, 5, 10, and 15 outliers were introduced in the simulations. Comparison of the estimates (in the absence and presence of outliers) was done using the Accuracy Measures (Mean Square Error (MSE), the Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE)). The accuracy measures were based on the deviations of parameters estimated from the corresponding parameters. The results of the analysis showed MSE is 0.024 in the absence of outliers. The values of the MSE increased as the number of outliers increased to 15.37, 21.23, and 31.23 for 5,10 and 15 outliers respectively. Thus in the presence of outlier, the Buys-Ballot Estimates (BBE) is not effective in estimating the trend parameters and seasonal indices when trend-cycle component is linear for additive model. Hence, it has been recommended that in estimating the trend parameters and seasonal indices using a Buys-Ballot procedure of the Buys-Ballot table, the effect of outlier should be considered and handled before further analysis.*

**Keywords:** Buys-Ballot Estimates, Linear Trend Parameters, Seasonal Indices, Outlier, Descriptive Time Series.

**INTRODUCTION**

Buys-Ballot procedure is an alternative to the traditional descriptive time series analysis introduced by Iwueze and Nwogu (2004, 2005). In descriptive time series analysis, the traditional practice is to decompose an observed series , at time, , into its component parts; Trend component , Seasonal component ,Cyclical component , and Irregular component . The models often used for time series decomposition are; Additive, Multiplicative, and Mixed models respectively, given by,

 

 

 The additive model is based on the assumption that the deterministic components, , , and  are mutually independent and that each follows a fixed pattern for the period under review and that the four components add up to . The multiplicative model is applicable when the  values increase rapidly and the amplitude of seasonal movement tends to increase as the series progresses in time, rather than being fixed. When the original time series contains very small or zero values, the mixed model is used. (Chatfield, 2004; Kendal, 1990).

For short series when trend and cyclical components are assumed to be jointly estimated, the cyclical component is superimposed into the trend (Chatfield, 2004); and the observed time series,  can be decomposed into the trend-cycle component denoted by, Seasonal component,  and the residuals . In this case, (1.1), (1.2) and (1.3) may be written, respectively, as

  

  

  

Any of models (1.4), (1.5) or (1.6) can be used to estimate the three components and hence, affect the decomposition of a time series.

According to Iwueze and Nwogu (2004, 2005), the first step in the traditional method of time series decomposition is to estimate and eliminate for each time period from the actual data, either by subtraction (for (1.4)) or division (for (1.5)), giving a de-trended series that expresses the effect of the seasons and the irregular component. Next, the average of the de-trended series at each season is estimated to derive the seasonal effect. Calculating the average for each season and comparing it to the overall average value, either as a difference for (1.4) or as a ratio for (1.5), is typically sufficient for series with little trend. For a series which do contain a substantial trend, a more sophisticated approach may be required. After the calculation of the trend, the seasonal effect can be estimated by averaging  or  at each season, depending on whether the seasonal effect is thought to be additive or multiplicative. This gives the residual or irregular component. It is always assumed that the seasonal effect, when it exists, has periods. That is, it repeats after s time periods. That is, . For (1.4), it is convenient to make the further assumption that the sum of the seasonal components over a complete period is zero. That is, . Similarly, for (1.5), the convenient variant assumption is that the sum of the seasonal components over a complete period is . That is,  (Iwueze and Nwogu, 2014).

The traditional method of time series decomposition which is based on fitting a trend curve by some method and de-trending the series, using the de-trended series to estimate the seasonal indices, has some limitations. There are many cases where there is ‘zero’ trend and the average at each season is ‘compared’ with the overall average to obtain the seasonal indices. These limitations and difficulties in the traditional method of time series decomposition is what the Buys-Ballot estimation procedure as proposed by Iwueze and Nwogu (2004) appears to address.

The Buys-Ballot method of decomposition calculates trend with ease, averts the issue of de-trending a series before estimating the seasonal impacts, and estimates the error variance without necessarily breaking the series into its component parts from the row, column and overall averages and standard deviations of the Buys-Ballot Table. For details of Buys-Ballot procedure, see for example, Iwueze and Nwogu (2004, 2005) and Iwueze and Ohakwe (2004).

**Table 1. The Layout of a Buys-Ballot Table**

|  |  |
| --- | --- |
| Rows/Period  | Columns /(Season) |
| 1 | 2 | … | j | … |  |  |  |  |
| 1 | wps33 | wps34 | … | wps35 | … | wps36 |  |  |  |
| 2 | wps40 | wps41 | … | wps42 | … | wps43 |  |  |   |
| 3 | wps47 | wps48 | … | wps49 | … | wps50 |  |  |   |
| wps54 | wps55 | wps56 |  | wps58 |  | wps60 | wps61 | wps62 | wps63 |
|  | wps64 | wps65 | … | wps66 | … | wps67 |  |  |  |
| wps71 | wps72 | wps73 |  | wps75 |  | wps77 | wps78 | wps79 | wps80 |
|  | wps81 | wps82 | … | wps83 | … | wps84 |  |  |  |
|  |  |  | … |  | … |  |  |  |  |
|  |  |  | … |  | … |  |  |  |  |
|  |  |  | … |  | … |  |  |  | wps105 |

 Where, for  and , , , 

, , ,  

is the number of periods/years,  is the periodicity; and is the total number of observations/sample size.

Using the Buys-Ballot procedure, Iwueze and Nwogu (2004) derived the estimates of trend parameters and seasonal indices for additive and multiplicative models when trend line is linear. That is,

 (1.7)

where,

 is the intercept; is the slope; and  is the time point.

Iwueze and Nwogu (2004, 2005) used the row, column, overall averages and variances to estimate the parameters of the trend line and the seasonal indices of the Buys-Ballot Table. The row, averages are the same for both the additive and multiplicative models and are functions ofand for fixed periodic interval, .

Occasionally, interruptive events like strikes, war outbreaks, abrupt political or economic crises, unforeseen heat or cold waves, or even unanticipated typing and recording errors can have an impact on descriptive time series observations. These interruptions have the effect of producing erroneous observations that are at odds with the rest of the series. Outliers are the term typically used to describe such supposedly erroneous observations ( Aggarwal C.C., 2017)

**2. Statement of Problem**

In development of Buys-Ballot procedures proposed by Iwueze and Nwogu (2004, 2005), the row, column, overall averages and variances were used to obtain estimates of trend parameters and seasonal indices. However, in using these rows, columns, overall averages and variances to obtain trend parameters and seasonal indices, they did not acknowledge the effect of outliers. It is known that averages and variances on the bases of which these estimates were derived are affected by outliers. The question is, in a series with outliers, can the Buys-Ballot procedures as stated by Iwueze and Nwogu (2004, 2005), be used or is there any need to handle the outliers before obtaining estimates of trend parameters and seasonal indices?

Therefore, the aim of this study is to determine the effect(s) of outliers on Buys-Ballot estimates of trend parameters and seasonal indices which may be relevant to improve the estimates in such unusual circumstances. The objectives are to simulate series for linear trending curve in the absence of outliers. Introduce outliers in the series. Obtain estimates of trend parameters and seasonal indices in the absence and presence of outliers and, compare the estimates of trend parameters and seasonal indices in the absence and presence of outliers.

**Methodology**

The study is based on simulated series. In the absence of outliers. the simulated series consists of 100 series of 120 observations from the model, , with .  given in Table 2.  and  are parameters to be estimated at time ;  is the seasonal component; and  is the residual error term which is assumed to follow a normal distribution with mean,0, and variance 1.

In the presence of outliers, some percentages of outliers were introduced into the series at randomly selected positions.

Estimates of trend parameters were obtained using the Buys-Ballot procedure. According to Iwueze and Nwogu (2004) for the estimated additive model with a linear trend.

 (3.1)

Estimates of trend parameters and seasonal indices are summarized in Table 3.2

In order to establish the presence of outliers in the series, the method used is the;

**Z-score statistic** given by (Songwon, 2002) as  where,

 





Under the null hypothesis, the  follows the normal distribution with mean 0 and variance 1. The null hypothesis is rejected and declared an outlier at *a* level significance if  or not rejected otherwise, where the constant 3 is the cutoff mark.

To determine the effects of the outliers on the estimates, summary statistics such as MSE, MAE and MAPE etc. were used. These summary statistics are based on the deviations of the estimates  from the parameters used in the simulation.

, 

where,

 is the number of parameters used. Hence,

  (3.3)

  (3.4)

  (3.5)

where,





**Table 2.** Estimates of Trend Parameters and Seasonal Indices for The Additive Model

|  |  |
| --- | --- |
| PARAMETER | ADDITIVE MODEL ESTIMATES |
|  |  |
|  |  |
|  |  |

**Empirical Examples**

The empirical examples under the normal distribution consisted of 100 simulation of 120 observations, each simulated from , with , given in Table 3.. The simulations are made for four different cases, namely, No outlier, presence of 5, 10 and 15 outliers, respectively.

**Table 3.** Seasonal () indices used in the simulation of the series.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|  | -1.5 | 2.5 | 3.5 | -4.5 | -1.5 | 2.5 | 3.5 | -4.5 | -1.5 | 2.5 | 3.5 | -4.5 |

**Table 4.** Presents the estimated trend parameters and seasonal indices of the simulated series in the absence and presence of 5, 10 and 15 outliers and the estimates of their error means and standard deviations computed from the residuals.

|  |  |
| --- | --- |
| **Actual** | **ESTI MATES** |
| **Parameters** | **No Outlier** | **5 Outliers** | **10 Outliers** | **15 Outliers** |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps1.png | C:\Users\user\AppData\Local\Temp\ksohtml8784\wps2.png | C:\Users\user\AppData\Local\Temp\ksohtml8784\wps3.png | C:\Users\user\AppData\Local\Temp\ksohtml8784\wps4.png | C:\Users\user\AppData\Local\Temp\ksohtml8784\wps5.png | C:\Users\user\AppData\Local\Temp\ksohtml8784\wps6.png | C:\Users\user\AppData\Local\Temp\ksohtml8784\wps7.png | C:\Users\user\AppData\Local\Temp\ksohtml8784\wps8.png | C:\Users\user\AppData\Local\Temp\ksohtml8784\wps9.png |
| **C:\Users\user\AppData\Local\Temp\ksohtml8784\wps10.jpg** | **4.79** | **0.21** | **9.83** | **-4.83** | **10.91** | **-5.91** | **12.87** | **-7.87** |
| **C:\Users\user\AppData\Local\Temp\ksohtml8784\wps11.jpg** | **0.20** | **0** | **0.18** | **0.02** | **0.17** | **0.03** | **0.22** | **-0.02** |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps12.jpg-1.5 | -1.61 | 0.11 | -0.89 | -0.61 | -0.53 | -0.97 | -0.85 | -0.65 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps13.jpg2.5 | 2.65 | -0.15 | -0.03 | 2.53 | -0.60 | 3.10 | 0.47 | 2.03 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps14.jpg3.5 | 3.47 | 0.03 | 0.15 | 3.35 | -0.31 | 3.81 | 1.62 | 1.88 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps15.jpg-4.5 | -4.72 | 0.22 | 1.45 | -5.95 | 0.72 | -5.22 | 0.45 | -4.95 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps16.jpg-1.5 | -1.28 | -0.22 | -0.77 | -0.73 | 0.40 | -1.90 | -1.24 | -0.26 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps17.jpg2.5 | 2.48 | 0.02 | 0.03 | 2.47 | -0.51 | 3.01 | 0.42 | 2.08 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps18.jpg3.5 | 3.49 | 0.01 | 0.23 | 3.27 | -0.18 | 3.68 | -0.86 | 4.36 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps19.jpg-4.5 | -4.34 | -0.16 | -1.61 | -2.89 | -1.36 | -3.14 | -1.16 | -3.35 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps20.jpg-1.5 | -1.45 | -0.05 | -0.90 | -0.60 | -0.48 | -1.02 | 0.20 | -1.70 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps21.jpg2.5 | 2.34 | 0.16 | 0.02 | 2.48 | 0.20 | 2.30 | -0.62 | 3.12 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps22.jpg3.5 | 3.34 | 0.16 | 0.33 | 3.17 | -0.03 | 3.53 | 1.73 | 1.77 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps23.jpg-4.5 | -4.37 | -0.13 | 1.98 | -6.48 | 2.68 | -7.18 | -0.16 | -4.34 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps24.png= 0 | 0.26 | -0.26 | -0.07 | 0.07 | -0.23 | 0.23 | -0.01 | 0.01 |
| C:\Users\user\AppData\Local\Temp\ksohtml8784\wps25.png = 1 | 1.05 | -0.05 | 10.22 | -9.22 | 13.08 | -12.08 | 18.74 | -17.74 |

In Table 4, it could be seen that the actual values of the parameters  and seasonal indices used in the simulations were recovered from the corresponding estimated values, with  and . The values of the deviations of the parameter estimated from the corresponding actual values used in the simulation , are the errors made in estimating  as presented in Table 4.

For the simulated series in the presence of 5 outliers, Table 4 showed that the actual values of the parameters and seasonal indices used in the simulations were not recovered from the corresponding estimated values. The estimates of the error mean and standard deviations were also not recovered. Thus, the values of the estimated parameters were  with and  with an  recorded as against the actual values of , . The table also presented the estimated values of the error mean  and standard error  as against the actual values of  and  used in the simulation.

Table 4 also presented the estimated values of trend parameters and seasonal indices in the presence of 10 and 15 outliers respectively. The values of the estimated parameters are , and ,  respectively. This is to say, the results of the deviations of the parameters estimated from the corresponding actual values used in the simulations were not recovered.

**Assessing Performance of Trend Parameters and Seasonal Indices in the Absence and Presence of Outliers**

To access and compare the Buys-Ballot estimates of trend parameters and seasonal indices in the absence and presence of outliers, the simulated series were subjected to Accuracy Measures for effective comparison. The results of the accuracy measures are presented in **Appendix A, B, C and D** respectively while the summary statistic for 10 simulations are given in Table 5.

**Table 5** Summary statistics for the simulated series in the absence and presence of outliers

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S/No | No Outlier | 5 Outliers | 10 Outliers | 15 Outliers |
|  | MSE | MAE | MAPE | MSE | MAE | MAPE | MSE | MAE | MAPE | MSE | MAE | MAPE |
| 1 | 0.024 | 0.111 | 4.962 | 9.60 | 2.43 | 68.76 | 15.44 | 2.90 | 78.25 | 23.81 | 2.95 | 68.25 |
| 2 | 0.023 | 0.133 | 4.776 | 15.37 | 3.04 | 83.24 | 21.23 | 3.57 | 99.22 | 31.28 | 3.54 | 89.22 |
| 3 | 0.021 | 0.122 | 4.305 | 10.55 | 2.53 | 73.32 | 14.56 | 2.87 | 79.72 | 26.03 | 3.03 | 99.72 |
| 4 | 0.029 | 0.145 | 5.034 | 11.31 | 2.62 | 71.91 | 16.38 | 2.99 | 78.29 | 31.28 | 3.63 | 88.29 |
| 5 | 0.031 | 0.154 | 6.009 | 9.60 | 2.41 | 69.27 | 15.96 | 3.00 | 85.67 | 37.20 | 4.12 | 95.67 |
| 6 | 0.038 | 0.170 | 5.697 | 8.17 | 2.23 | 62.70 | 15.26 | 2.94 | 80.01 | 27.82 | 3.27 | 90.01 |
| 7 | 0.011 | 0.088 | 3.540 | 10.27 | 2.44 | 64.54 | 17.73 | 3.03 | 76.41 | 25.92 | 3.48 | 86.41 |
| 8 | 0.019 | 0.114 | 4.634 | 10.16 | 2.36 | 65.22 | 15.11 | 2.56 | 69.94 | 19.53 | 2.90 | 79.94 |
| 9 | 0.010 | 0.086 | 3.097 | 10.79 | 2.57 | 75.70 | 16.31 | 3.04 | 86.81 | 25.10 | 3.61 | 96.81 |
| 10 | 0.019 | 0.107 | 3.715 | 9.06 | 2.31 | 62.65 | 14.78 | 2.84 | 80.42 | 22.89 | 3.35 | 90.42 |

For the simulated series in the absence of outlier, **Table 5** shows that for , and , the values of the summary statistics (MSE, MAE and MAPE) are equal in almost all the simulations. However, for the values in the presence of 5, 10 and 15 outliers, the **Table 5** show that the values of the summary statistics (MSE MAE and MAPE) are unequal in most all the simulations and the difference increased as the values of the outliers’ increases. In other words, when a series is dominated by outliers, the estimates computed from the series become less precise.

This lack of agreement between estimated values in the presence of outliers and the actual values used in the simulations may be attributed to the violation of the condition for successful normality in the series by outliers.

It was observed that the values of the results of the accuracy measures for the simulated series with outliers, indicated that the more the numbers of outliers, the higher the values of the MSE, MAE and MAPE.

**Summary, Conclusion and Recommendation**

This study compared the effects of outliers on Buys-Ballot Estimates (BBE) of trend parameters and seasonal indices in descriptive time series analysis when trend-cycle component is linear for additive model. The rationale for this study was to assess and compare the performance of the Buys-Ballot Estimation in the absence and presence of outliers which may be relevant in obtaining the estimates in other such unusual circumstances.

The comparison of the estimates of trend parameters and seasonal indices in the absence and presence of outliers in this study was done using Accuracy Measures (the Mean Square Error (MSE), the Mean Absolute Error (MAE), and the Mean Absolute Percentage Error (MAPE)). These Accuracy measures were defined, for each estimation procedure, in terms of the deviations of the parameters estimated (using simulated series) from the corresponding actual values used in the simulations.

The results of the analyses showed that, for the simulated series in the absence of outlier, the actual values of the parameters and seasonal indices used in the simulations were recovered from the corresponding estimated values and the summary statistic (MSE, MAE and MAPE) were equal both in magnitude and direction in all the simulations without outlier. This indicates that, in the absence of outliers, the Buys-Ballot Estimation method is effective in estimating the trend parameters and seasonal indices when trend-cycle component is linear for additive model.

For the simulated series in the presence of outliers, the results of the analyses showed that the actual values of the parameters and seasonal indices used in the simulations were not recovered from the corresponding estimated values and the summary statistics (MSE, MAE and MAPE) were not equal both in magnitude and direction in all the simulations with outliers.

It was also recorded from the obtained results that in the presence of outliers, the estimated error means and standard deviations increases as the numbers of outliers increases. This is an indication that the more numbers of outliers in a series, the higher the error in the estimation and ineffective the results.

It is therefore, concluded that the presence of outlier affects the Buys-Ballot Estimates of trend parameters and seasonal indices when trend-cycle component is linear for additive model. Hence, it is recommended that in estimating the trend parameters and seasonal indices using a Buys-Ballot procedure of the Buys-Ballot table, the effect of outlier should be considered and handled before further analysis.

**References**

Aggarwal Charu C. (2017) *Outlier Analysis (2nd ed.).* Springer

Buys-Ballot, C.H.D., (1847). Leo Clement Periodiques de Temperatures Kemint et Fils Utrecht.

Chatfield, C. (2004). The analysis of time series: An introduction. Chapman and Hall, / CRC press,

Boca Raton.

Iwueze, I. S. and Nwogu, E.C. (2004). Buys –Ballot Estimates for Time Series Decomposition,

Global Journal of mathematics, 3(2), 83-98.

Iwueze, I. S. and 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

Iwueze, I. S., and Ohakwe, J. (2004). Uses of the Buys-Ballot Table in Time Series Analysis.

 Applied Mathematics, 2, 633-645.

Iwueze, I. S., and Nwogu, E.C., and Ajaraogu, J.C. (2011). Best linear unbiased estimate using

Buys-Ballot procedure when trend-cycle component is linear. *Pakistan Journal*

 *Statistics and Operations Research 7(2), 183-198*

Iwueze, I.S and Nwogu, E.C. (2014) Framework for Choice of Models and Detection of Seasonal Effect in Time Series. Far-East Journal of Theoretical Statistics, 48, 45-66.

Kendal, M.G. and Ord, J.K. (1990). Time series (3rd ed) Charles Griffin, London.