Far East Journal of Theoretical Statistics
© 2014 Pushpa Publishing House, Allahabad, India
Published Online: September 2014
Available online at http://pphmj.com/journals/fjts.htm
Volume 48, Number 1, 2014, Pages 45-66

FRAMEWORK FOR CHOICE OF MODELS AND DETECTION OF SEASONAL EFFECT IN TIME SERIES

Iheanyi S. Iwueze and Eleazer C. Nwogu

Department of Statistics
Federal University of Technology
Owerri, Nigeria

Abstract

Identification of patterns and choice of model in time series data is critical to facilitate forecasting. Two patterns that may be presented are trend and seasonality and the two competing models are the additive and multiplicative models. This paper uses the Buys Ballot table: (1) to provide an overview of recent developments in the identification and measure of trend and seasonality and choice of models in classical time series data analysis, and (2) to provide new insights into the development of new methodologies for effective identification of patterns and choice of models in classical time series data analysis when the trend is monotonous and the seasonal pattern is stable.

1. Introduction

There are two main goals of time series analysis: (1) identifying the nature of the phenomenon represented by the sequence of observations, and (2) forecasting (predicting future values of the time series variable). Both of these goals require that the pattern of observed time series data is identified and more or less formally described.

Received: January 20, 2014; Revised: April 8, 2014; Accepted: July 5, 2014

2010 Mathematics Subject Classification: 37M10.

Keywords and phrases: trend, seasonality, pattern identification, choice of models, additive and multiplicative models.

In time series analysis, it is assumed that the data consist of a systematic pattern (usually a set of identifiable components) and random noise (error). Most time series patterns can be described in terms of four basic classes of components: trend (denoted as T_t), seasonal (denoted as S_t), cyclical (denoted as C_t) and irregular (denoted as I_t or e_t) components, where t stands for the particular point in time. These four classes of time series components may or may not coexist in real-life data.

The trend represents a general systematic linear or (most often) nonlinear component that changes over time and does not repeat or at least does not repeat within the time range captured by out data. As long as the trend is monotonous (consistently increasing or decreasing), the identification of trend component is not very difficult. Trend analysis (methods for estimating the trend parameters) can be done by three important methods: (1) smoothing (Box and Jenkins [5], Velleman and Hoaglin [33], Makridakis et al. [25], Gardner [11] and Montgomery et al. [27]), (2) fitting a mathematical function (Chatfield [7] and Kendall and Ord [20]) and (3) differencing to make the series stationary in the ARIMA methodology (Box and Jenkins [5], Pankratz [30] and Wei [35]). Tests for trend are given in Kendall and Ord [20]. Correlation analysis can also be used to assess trend. If a time series contains a trend, then the values of the autocorrelations will not come to zero except for very large values of the lag (Chatfield [7]). Table 1 gives a summary of the common functions used to approximate trend.

Many time series exhibit a variation which repeats itself in systematic intervals over time and this behaviour is known as seasonal dependency (seasonality). By seasonality, we mean periodic fluctuations. It is formally defined as correlational dependency of order k between each ith element of the series and the (i-k)th element (Kendall and Ord [20]) and measured by autocorrelation (a correlation between X_t and X_{t-k}); k is usually called the *lag*. Seasonality can be visually identified in the series as a pattern that repeats every k elements. The following graphical techniques can be used to detect seasonality: (1) a run sequence plot (Chambers et al. [6]), (2) a seasonal subseries plot (Cleveland [8]), (3) multiple box plots (Chambers et al. [6]) and (4) the autocorrelation plot (Box and Jenkins [5]). Both the seasonal subseries plot and the box plot assume that the seasonal periods are known. In most cases, the seasonal periods are easy to find (4 for quarterly data, 12 for monthly data, etc). If there is significant seasonality, then the autocorrelation plot should show spikes at multiples of lags equal to the period, the seasonal lag (for quarterly data, we would expect to see significant spikes at lags 4, 8, 12, 16, and so on). Davey and Flores [10] proposed a method which adds statistical tests of seasonal indexes for the multiplicative model that helps identify seasonality with greater confidence. Tests for seasonality are also given in Kendall and Ord [20]. The seasonal component, S_t, is associated with the property that $S_{(i-1)s+j} = S_j$, $i = 1, 2, \ldots$.

Apart from seasonal effects, some time series exhibit variation at a fixed period or at periods that are not fixed but which are predictable. The difference between a cyclical component and a seasonal component is that the latter occurs at regular (seasonal) intervals, while cyclical factors have usually a longer duration that varies from cycle to cycle. For short duration of data, the trend and cyclical components are customarily combined into a trend-cycle component, denoted as M_t (Chatfield [7]).

After trend and seasonal effects have been removed from a set of data, we are left with a series of residuals, which may or may not be random. A visual examination of the run sequence plot may be enough to see that a series is random or not random. A variety of tests exist for randomness and they are described in Kendall and Ord [20]. If a time series is completely random, then its autocorrelations are zero for all lags. Tests based on the autocorrelation of the residuals are called *portmanteau lack-of-fit tests* in Box and Jenkins [5] and Ljung and Box [23].

In addition to identifying the patterns (the components), our two goals are better achieved if and only if the correct model is used for analysis. The specific functional relationship between these components can assume different forms. However, two straightforward possibilities are that they combine in an additive (additive seasonality) or a multiplicative

(multiplicative seasonality) fashion: additive model (when trend and cyclical components are combined) is

$$X_t = M_t + S_t + e_t, \tag{1.1}$$

where $\sum_{j=1}^{s} S_j = 0$ and $e_t \sim N(0, \sigma_1^2)$.

Multiplicative model (when trend and cyclical components are combined) is

$$X_t = M_t * S_t * e_t, \tag{1.2}$$

where $\sum_{j=1}^{s} S_j = s$ and $e_t \sim N(1, \sigma_2^2)$.

The run sequence plot (time plot) is 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 said to be *additive*, but if it increases in size in direct proportion to the mean level, then it is said to be *multiplicative* (Chatfield [7]). Multiplicative seasonality implies that the seasonal changes increase with the overall trend (i.e., the variance is correlated with then mean over the segments of the series). In the plot of the series, the distinguishing characteristic between these two types of seasonal components is that in the additive case, the series shows steady seasonal fluctuations, regardless of the overall level of the series; in the multiplicative case, the size of the seasonal fluctuations varies, depending on the overall level of the series.

The purpose of classical decomposition method is to isolate those components, that is, to decompose the series into trend effect, seasonal effects, and remaining variability. The traditional method of time series decomposition for the deterministic trend/seasonal (DTDS) model, based on (i) fitting a trend curve to the entire series by some method and detrending the series (ii) using the detrended series to estimate the seasonal indices is time consuming and complex (Chatfield [7], Kendall and Ord [20] and Wei [35]).

This paper proposes a methodology to do three things: (1) improve on the traditional method of time series decomposition by showing that trend estimation is different for additive and multiplicative models while seasonal estimates can be done without first detrending the series, (2) improve the degree of confidence in the characterization of seasonality in a time series by performing statistical tests on the seasonal indices and (3) improve the degree of confidence in the choice between additive or multiplicative models by performing statistical tests.

Table 1. Tests needed for choosing an appropriate curve to fit the given data

S/No.	Curve name	Mathematical form of the curve	Property to be tested for $y = X_t$
1	Straight line	$X_t = a + bt$	The first differences of given y's are nearly constant
2	Parabola	$X_t = a + bt + ct^2$	The second differences of given y's are nearly constant
3	Cubic	$X_t = a + bt + ct^2 + dt^3$	The third differences of given y's are nearly constant
4	Polynomial of order p	$X_t = a_0 + a_1 t + a_2 t^2 + \cdots + a_p t^p$	The pth differences of given y's are nearly constant
5	Exponential	$X_t = ae^{bt}$	The first differences of $[\log y]$'s are nearly constant
6	Exponential parabola	$X_t = ae^{bt+ct^2}$	The second differences of $[\log y]$'s are nearly constant
7	Modified exponential	$X_t = ab^t + c$	The first differences of y change by a constant percentage
8	Gompertz	$X_t = ab^{c^t}$	The first differences of $[\log y]$'s change by a constant percentage
9	Logistic	$X_t = \dfrac{a}{1 + bc^t}$	The first differences of $\left[\dfrac{1}{y}\right]$'s change by a constant percentage

Note. X_t stands for the observed value of the time series at time t.

2. Buys Ballot Table and its Uses

A seasonal time series data is conventionally arranged into m rows and s columns. By arranging a seasonal series of length n into m rows and s columns, as shown in Table 2, the rows represent the periods/years while the columns are the seasons. This two-dimensional arrangement of a series is referred to as the Buys Ballot table (see Wold [36] and Iwueze and Nwogu [15]).

Table 2. Buys Ballot table

Row	Columns								
	1	2	…	j	…	s	$T_{i.}$	$\overline{X}_{i.}$	$\hat{\sigma}_{i.}$
1	X_1	X_2	…	X_j	…	X_s	$T_{1.}$	$\overline{X}_{1.}$	$\hat{\sigma}_{1.}$
2	X_{s+1}	X_{s+2}	…	X_{s+j}	…	X_{2s}	$T_{2.}$	$\overline{X}_{2.}$	$\hat{\sigma}_{2.}$
3	X_{2s+1}	X_{2s+2}	…	X_{2s+j}	…	X_{3s}	$T_{3.}$	$\overline{X}_{3.}$	$\hat{\sigma}_{3.}$
…	…	…	…	…	…	…	…	…	…
i	$X_{(i-1)s+1}$	$X_{(i-1)s+2}$	…	$X_{(i-1)s+j}$	…	$X_{(i-1)s+s}$	$T_{i.}$	$\overline{X}_{i.}$	$\hat{\sigma}_{i.}$
…	…	…	…	…	…	…	…	…	…
m	$X_{(m-1)s+1}$	$X_{(m-1)s+2}$	…	$X_{(m-1)s+j}$	…	X_{ms}	$T_{m.}$	$\overline{X}_{m.}$	$\hat{\sigma}_{m.}$
$T_{.j}$	$T_{.1}$	$T_{.2}$	…	$T_{.j}$	…	$T_{.s}$	$T_{..}$		
$\overline{X}_{.j}$	$\overline{X}_{.1}$	$\overline{X}_{.2}$	…	$\overline{X}_{.j}$	…	$\overline{X}_{.s}$		$\overline{X}_{..}$	
$\hat{\sigma}_{.j}$	$\hat{\sigma}_{.1}$	$\hat{\sigma}_{.2}$	…	$\hat{\sigma}_{.j}$	…	$\hat{\sigma}_{.s}$			$\hat{\sigma}_{..}$

$T_{i.}$ = Total for ith period/year

$T_{.}$ = Sum of all observations

$\overline{X}_{.j}$ = Average for jth season

$\hat{\sigma}_{.i}$ = Standard deviation for ith period/year

$\hat{\sigma}_{..}$ = Overall observation standard deviation

s = Number of seasons per period/years

$T_{.j}$ = Total for jth season

$\overline{X}_{i.}$ = Average for ith period/year

$\overline{X}_{..}$ = Overall observation average

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

m = Number of period/years

$n = ms$ = Number of observations.

Chatfield [7] suggested the use of the Buys Ballot table for inspecting time series data for the presence of trend and seasonal effects. In addition to the inspection for the presence of trend and seasonal effects, Iwueze and Nwogu [15], Iwueze and Ohakwe [17] and Iwueze and Nwogu [16] proposed a Buys Ballot Estimation Procedure for the estimation of trend parameters and seasonal indices. Fomby [12] presented various graphs suggested by the Buys Ballot table for inspecting time series data for the presence of seasonal effects. Fomby [13] in his study of Stable Seasonal Pattern (SSP) models gave an adaptation of Friedman's two-way analysis of variance by ranks test for seasonality in time series data. Iwueze et al. [18] gave five (5) uses of the Buys Ballot table in time series analysis. Iwueze et al. [19] and Nwogu et al. [28] provided the best linear unbiased estimates of the Buys Ballot estimates.

All the above methods wrongly assume that the periodic/yearly averages $\overline{X}_{i.},\ i = 1, 2, \ldots, m$ for both additive and multiplicative models are functions of the periods/years that reveal the basic trend in the data but will have nothing to say about seasonality. This wrong assumption on the periodic/yearly averages has lead to the use of the same formulae for the estimation of trend parameters in additive and multiplicative models. To overcome this problem, we have introduced the computation of periodic/yearly sample variances and seasonal sample variances into the Buys Ballot table. In contrast, the seasonal averages $\overline{X}_{.j},\ \ j = 1, 2, \ldots, s$ reveal an interesting seasonal pattern because they are functions of the trend parameters and the seasonal indices.

3. Row, Column and Overall Averages and Variances of the Buys Ballot Table for Selected Trend Curves

The summaries of the row, column and overall averages and variances are shown in Tables 3.1 through 3.3 for the selected trending curves under the additive and multiplicative models. As the tables show, the row, column and overall averages and variances are not the same for the two models.

(a) (i) For the additive model, the row averages mimic the shape of the trending series and do not contain the seasonal component and (ii) for the multiplicative model, the row averages also mimic the shape of the trending series but contain the seasonal component.

(b) The column averages mimic the shape of the trending series and contain the seasonal component for both additive and multiplicative models.

(c) The row variances contain both the trending parameters and the seasonal component for both additive and multiplicative models.

(d) Just like the row averages, (i) for the additive model, the column variances mimic the shape of the trending series and do not contain the seasonal component and (ii) for the multiplicative model, the column variances also mimic the shape of the trending series but contain the seasonal component.

These characteristics are what could be used for (i) choice of the appropriate model for decomposition, (ii) assessment and estimation of trend and (iii) assessment and estimation of seasonality.

Table 3.1a. Summary of totals and averages of a series when trend-cycle components are linear: $(M_t = a + bt)$

Totals and averages	Linear trend-cycle component: $M_t = a + bt$, $t = 1, 2, ..., n = ms$	
	Additive model	Multiplicative model
$T_{i.}$	$as - \frac{bs}{2}(s-1) + (bs^2)i,$	$(a - bs)s + b\sum_{j=1}^{s} jS_j + (bs^2)i$
$T_{.j}$	$ma + \frac{mb}{2}(n-s) + mbj + mS_j$	$m\left\{a + \frac{b}{2}(n-s) + bj\right\}S_j$
$T_{..}$	$na + nb\left(\frac{(n+1)}{2}\right)$	$n\left\{a + \frac{b}{2}(n-s) + \frac{b}{s}\sum_{j=1}^{s} jS_j\right\}$
$\overline{X}_{i.}$	$a - b\left(\frac{s-1}{2}\right) + (bs)i$	$a - \frac{b}{s}\left(s^2 - \sum_{j=1}^{s} jS_j\right) + (bs)i$
$\overline{X}_{.j}$	$a + \frac{b}{2}(n-s) + bj + S_j,$	$\left\{a + \frac{b}{2}(n-s) + bj\right\}S_j$
$\overline{X}_{..}$	$a + \frac{b(n+1)}{2}$	$a + \frac{b}{2}(n-s) + \frac{b}{s}\sum_{j=1}^{s} jS_j$

Table 3.1b. Summary of sample variances of a series when trend-cycle components are linear: $(M_t = a + bt)$

Sample variances	Linear trend-cycle component: $M_t = a + bt,\ t = 1, 2, \ldots, n = ms$	
	Additive model	Multiplicative model
$\hat{\sigma}_i^2$	$b^2\left(\frac{s(s+1)}{12}\right) + \left(\frac{2b}{s-1}\right)\sum_{j=1}^{s} jS_j + \frac{1}{s-1}\sum_{j=1}^{s} S_j^2$	$\frac{1}{s-1}\left\{\begin{array}{l} \{a + bs[(i-1)]\}^2 \sum_{j=1}^{s}(S_j - 1)^2 + b^2 \sum_{j=1}^{s}\left(jS_j - \frac{C_1}{s}\right) \\ + 2b\{a + b[(i-1)s]\}\sum_{j=1}^{s}(S_j - 1)\left(jS_j - \frac{C_1}{s}\right) \end{array}\right\}$
$\hat{\sigma}_j^2$	$b^2\left(\frac{n(n+s)}{12}\right)$	$b^2\left(\frac{n(n+s)}{12}\right)S_j^2$
$\hat{\sigma}^2$	$b^2\left(\frac{n(n+1)}{12}\right) + \frac{1}{n-1}\left\{2bm\sum_{j=1}^{s} jS_j + m\sum_{j=1}^{s} S_j^2\right\}$	$\frac{1}{n-1}\left\{\begin{array}{l} \frac{b^2 n(n-s)(n+s)}{12} \\ + m\left[a^s + ab(n-s) + \frac{b^2(n-s)(2n-s)}{6}\right]\sum_{j=1}^{s}(S_j - 1)^2 \\ + mb^2\sum_{j=1}^{s}\left(jS_j - \frac{C_1}{s}\right) \\ + 2b\left(ma + \frac{nb(m-1)}{2}\right)\sum(S_j - 1)\left(jS_j - \frac{C_1}{s}\right) \end{array}\right\}$

Note. $C_1 = \sum_{j=1}^{s} jS_j$.

Table 3.2a. Summary of totals and averages of a series when trend-cycle components are quadratic: $(M_t = a + bt + ct^2)$

Totals and averages	Quadratic trend-cycle component: $M_t = a + bt + ct^2,\ t = 1, 2, \ldots, n = ms$	
	Additive model	Multiplicative model
$T_{i.}$	$as - \frac{bs}{2}(s-1) + \frac{cs}{6}(s-1)(2s-1) + s^2[b - c(s-1)]i + (cs^3)i^2$	$as - b(s^2 - C_1) + cs\left(s^2 - 2C_1 + \frac{C_2}{s}\right) + s^2\left[b - \frac{2c}{s}(s^2 - C_1)\right]i + (cs^3)i^2$
$T_{.j}$	$am + \frac{bm}{2}(n-s) + \frac{cm}{6}(n-s)(2n-s) + m[b + c(n-s)]j + mcj^2 + mS_j$	$\left\{\begin{array}{l} am + \frac{bm}{2}(n-s) + \frac{cm}{6}(n-s)(2n-s) \\ + m[b + c(n-s)]j + mcj^2 \end{array}\right\} S_j$

$T_{..}$	$an + \frac{bn}{2}(n+1) + \frac{cn}{6}(n+1)(2n+1)$	$an + bn\left[\frac{n-s}{2} + \frac{C_1}{s}\right]$ $+ cn\left[\frac{(n-s)(2n-s)}{6} + (n-s)\frac{C_1}{s} + \frac{C_2}{s}\right]$
$\overline{X}_{i.}$	$a - \frac{b}{2}(s-1) + \frac{c}{6}(s-1)(2s-1)$ $+ s[b - c(s-1)]i + (cs^2)i^2$	$a - \frac{b}{s}(s^2 - C_1) + c\left(s^2 - 2C_1 + \frac{C_2}{s}\right)$ $+ s\left[b - \frac{2c}{s}(s^2 - C_1)\right]i + (cs^2)i^2$
$\overline{X}_{.j}$	$a + \frac{b}{2}(n-s) + \frac{c}{6}(n-s)(2n-s)$ $+ [b + c(n-s)]j + cj^2 + S_j$	$\left\{\begin{array}{l} a + \frac{b}{2}(n-s) + \frac{c}{6}(n-s)(2n-s) \\ + [b + c(n-s)]j + cj^2 \end{array}\right\} S_j$
$\overline{X}_{..}$	$a + \frac{b}{2}(n+1) + \frac{c}{6}(n+1)(2n+1)$	$a + b\left[\frac{n-s}{2} + \frac{C_1}{s}\right]$ $+ c\left[\frac{(n-s)(2n-s)}{6} + (n-s)\frac{C_1}{s} + \frac{C_2}{s}\right]$

Table 3.2b. Summary of sample variances of a series when trend-cycle components are quadratic: $(M_t = a + bt + ct^2)$

Sample variances	Quadratic trend-cycle component: $M_t = a + bt + ct^2$, $t = 1, 2, \ldots, n = ms$	
	Additive model	Multiplicative model
$\hat{\sigma}_i^2$	$\frac{s(s+1)}{180}\{(2s-1)(8s-11)c^2 - 30(s-1)bc + 15b^2\}$ $+ \frac{1}{s-1}\left\{\sum_{j=1}^{s} S_j^2 + 2[b - 2cs]C_1 + 2cC_2\right\}$ $+ \left\{\frac{s^2(s+1)}{3}\left[bc - c^2(s-1) + \frac{4csC_1}{s-1}\right]\right\}i$ $+ \left[\frac{s^3(s+1)c^2}{3}\right]i^2$	$\frac{1}{s-1}\left\{\begin{array}{l} \{a + bs[(i-1)]\}^2 \sum_{j=1}^{s}(S_j - 1)^2 + b^2 \sum_{j=1}^{s}\left(jS_j - \frac{C_1}{s}\right) \\ + 2b\{a + b[(i-1)s]\} \sum_{j=1}^{s}(S_j - 1)\left(jS_j - \frac{C_1}{s}\right) \end{array}\right\}$
$\hat{\sigma}_j^2$	$\frac{n(n+s)}{180}[(2n-s)(8n-11s)c^2 + 30(n-s)bc + 15b^2]$ $+ \frac{n(n+s)}{3}[(n-s)c^2 + bc]j + \left[\frac{n(n+s)c^2}{3}\right]j^2$	$\left\{\begin{array}{l} \frac{n(n+s)}{180}[(2n-s)(8n-11s)c^2 + 30(n-s)bc + 15b^2] \\ + \frac{n(n+s)}{3}[(n-s)c^2 + bc]j + \left[\frac{n(n+s)c^2}{3}\right]j^2 \end{array}\right\} S_j^2$
$\hat{\sigma}^2$	See Appendix	See Appendix

Note. $C_1 = \sum_{j=1}^{s} jS_j$, $C_2 = \sum_{j=1}^{s} j^2 S_j$.

Table 3.3a. Summary of totals and averages of a series when trend-cycle components are exponential: $(M_t = be^{ct})$

Totals and averages	Exponential trend-cycle component: $M_t = be^{ct}$, $t = 1, 2, ..., n = ms$	
	Additive model	Multiplicative model
$T_{i.}$	$b\left(\frac{e^{(1-s)} - e^{c}}{1 - e^{c}}\right)e^{csi}$	$\left[\frac{b}{e^{cs}}\sum_{j=1}^{s} e^{cj}S_j\right]e^{csi}$
$T_{.j}$	$b\left(\frac{1 - e^{cn}}{1 - e^{cs}}\right)e^{cj} + mS_j$	$b\left(\frac{1 - e^{cn}}{1 - e^{cs}}\right)e^{cj}S_j$
$T_{..}$	$be^{c}\left(\frac{1 - e^{cn}}{1 - e^{c}}\right)$	$b\left(\frac{1 - e^{cn}}{1 - e^{cs}}\right)\sum_{j=1}^{s} e^{cj}S_j$
$\overline{X}_{i.}$	$\frac{b}{s}\left(\frac{e^{(1-s)} - e^{c}}{1 - e^{c}}\right)e^{csi}$	$\left[\frac{b}{se^{cs}}\sum_{j=1}^{s} e^{cj}S_j\right]e^{csi}$
$\overline{X}_{.j}$	$\frac{b}{m}\left(\frac{1 - e^{cn}}{1 - e^{cs}}\right)e^{cj} + S_j$	$\frac{b}{m}\left(\frac{1 - e^{cn}}{1 - e^{cs}}\right)e^{cj}S_j$
$\overline{X}_{..}$	$\frac{be^{c}}{n}\left(\frac{1 - e^{cn}}{1 - e^{c}}\right)$	$\frac{b}{n}\left(\frac{1 - e^{cn}}{1 - e^{cs}}\right)\sum_{j=1}^{s} e^{cj}S_j$

Table 3.3b. Summary of sample variances of a series when trend-cycle components are exponential: $(M_t = be^{ct})$

Sample variances	Exponential trend-cycle component: $M_t = be^{ct}$, $t = 1, 2, ..., n = ms$	
	Additive model	Multiplicative model
$\hat{\sigma}_i^2$	$b^2 e^{2c[(i-1)s+1]}\left[\left(\frac{1-e^{2cs}}{1-e^{2c}}\right) - \frac{1}{s}\left(\frac{1-e^{cs}}{1-e^{c}}\right)\right] + \sum_{j=1}^{s} S_j^2 + 2be^{c(i-1)s}\sum_{j=1}^{s} e^{cj}S_j$	$\frac{b^2}{(s-1)e^{2cs}}\left[\sum_{j=1}^{s} e^{2cj}S_j^2 - \frac{1}{s}\left(\sum_{j=1}^{s} e^{cj}S_j\right)^2\right]e^{2csi}$
$\hat{\sigma}_j^2$	$\frac{b^2}{m-1}\left[\left(\frac{1-e^{2cn}}{1-e^{2cs}}\right) - \frac{1}{m}\left(\frac{1-e^{cn}}{1-e^{cs}}\right)^2\right]e^{2cj}$	$\frac{b^2}{m-1}\left[\left(\frac{1-e^{2cn}}{1-e^{2cs}}\right) - \frac{1}{m}\left(\frac{1-e^{cn}}{1-e^{cs}}\right)^2\right]e^{2cs}S_j^s$
$\hat{\sigma}^2$	$\frac{b^2e^{2c}}{n-1}\left[\left(\frac{1-e^{2cn}}{1-e^{2c}}\right) - \frac{1}{n}\left(\frac{1-e^{cn}}{1-e^{c}}\right)^2\right] + \frac{m}{m-1}\sum_{j=1}^{s} S_j^2 + \frac{2b}{n-1}\sum_{j=1}^{s} e^{cj}S_j$	$\frac{b^2}{(n-1)}\left[\left(\frac{1-e^{2cn}}{1-e^{2cs}}\right)\sum_{j=1}^{s} e^{2cj}S_j^2 - \frac{1}{n}\left(\frac{1-e^{cn}}{1-e^{cs}}\right)^2\left(\sum_{j=1}^{s} e^{cj}S_j\right)^2\right]$

4. Choice between Additive and Multiplicative Models

Decomposition methods require that the model structure (additive or multiplicative) is known and more or less formally described. From Tables 3.1 through 3.3, it is clear that the column variances, which depend only on the trend parameters for the additive model, will aid the choice of model.

(a) Linear trend component

For the purposes of selection of appropriate model for decomposition, an analyst only needs to compare the seasonal/column variances. If seasonal/column variances are the same (see Tables 3.1b), then the appropriate model is additive. However, if seasonal/column variances vary, then: (i) the appropriate model is multiplicative and (ii) the series contains seasonal effects.

Therefore, the problem of selection of appropriate model for decomposition when trend component is linear reduces to that of testing the null hypothesis:

$$H_0 : \sigma_{.1}^2 = \sigma_{.2}^2 = \cdots = \sigma_{.s}^2 \tag{4.1}$$

against the alternative hypothesis

$$H_1 : \sigma_{.j}^2 \neq \sigma_{.j'}^2, \text{ for at least one } j \neq j'. \tag{4.2}$$

Available tests that can be used to assess equality of variances of two or more samples are the F-test (Snedecor and Cochran [32]), the Bartlett's (Bartlett [2]) test, Levene's (Levene [22]) test, Box-Anderson (Box and Anderson [4]) test and Jacknife (Layard [21], O'Brien [29] and Sharma [31]) tests. The F-test and Bartlett's test are sensitive to departures from normality while Levene's test, as an alternative to the Bartlett test, is less sensitive to departures from normality. Box-Andersen test is a variation of Bartlett's test (Miller [26]). There are nonparametric tests that do not rely on the assumption that the variables have a normal distribution. The squared ranks (nonparametric) test can be used to assess equality of variances across two or more independent random samples which have been measured using a scale that is at least interval (Conover [9]). Other nonparametric tests for constant variance include those by Wang and Zhou [34], Allingham and Rayner [1] and Beersma and Buishand [3]. Outside the normality assumptions of the parametric tests, all parametric and nonparametric tests assume the following: (i) random samples, (ii) independence within samples, (iii) mutual independence between samples and (iv) measurement scale is at least interval. The assumptions of random samples, independence within samples and mutual independence between samples are typically violated for a time series. Hence, parametric and nonparametric tests of equality of the column variances ought to be compared in terms of robustness (methods with good performance when there are small departures from assumptions) and power (the probability that the test will reject the null hypothesis when the null hypothesis, H_0, is false (i.e., the probability of not committing a Type II error)) in simulation experiments.

(b) Quadratic trend component

Here, the row average is quadratic in i (equation (4.3)) while the column variance is quadratic in j (equation (4.5)). The values of a, b and c can be estimated from equation (4.3) and equation (4.5):

$$\overline{X}_i = a - \frac{b}{2}(s-1) + \frac{c}{6}(s-1)(2s-1) + s[b - c(s-1)]i + (cs^2)i^2 \quad (4.3)$$

$$= \alpha_0 + \alpha_1 i + \alpha_2 i^2, \quad (4.4)$$

$$\hat{\sigma}^2_{.j} = \frac{n(n+s)}{180}[(2n-s)(8n-11s)c^2 + 30(n-s)bc + 15b^2]$$

$$+ \frac{n(n+s)}{3}[(n-s)c^2 + bc]j + \left[\frac{n(n+s)c^2}{3}\right]j^2 \quad (4.5)$$

$$= \beta_0 + \beta_1 j + \beta_2 j^2. \quad (4.6)$$

We have to devise a suitable test for equality of the values of a, b and c obtained from equations (4.3) and (4.5).

(c) Exponential trend component

Here, the row average is exponential in i (equation (4.7)) while the column variance is exponential in j (equation (4.8)). The values of b and c can be estimated from equation (4.7) and equation (4.8):

$$\overline{X}_{i.} = \frac{b}{s}\left(\frac{e^{(1-s)} - e^c}{1 - e^c}\right)e^{csi} = \theta_0 e^{\theta_1 i}, \quad (4.7)$$

$$\hat{\sigma}^2_{.j} = \frac{b^2}{m-1}\left[\left(\frac{1-e^{2cn}}{1-e^{2cs}}\right) - \frac{1}{m}\left(\frac{1-e^{cn}}{1-e^{cs}}\right)^2\right]e^{2cj} = \phi_0 e^{\phi_1 j}. \quad (4.8)$$

We have to devise a suitable test for equality of the values of b and c obtained from equations (4.7) and (4.8).

5. Test for Seasonality in Additive Model

For all trending curves shown in Table 4, we note the following:

(1) $\overline{X}_{.j} - \overline{X}_{..}$ contains the parameters of the trending curve and the seasonal component at season j.

(2) $\sum_{j=1}^{s} (\overline{X}_{.j} - \overline{X}_{..}) = 0.$ Hence,

(3) $\hat{\sigma}^2_{(\overline{X}_{.j} - \overline{X}_{.})}$ depends on the trending parameters.

(4) Because of (1) and (2), a time series data should be detrended before test for seasonality.

Table 4. Properties of $(\overline{X}_{.j} - \overline{X}_{..})$

Trend	Linear	Quadratic	Exponential
$\overline{X}_{.j} - \overline{X}_{..}$	$b\left(j - \frac{s+1}{2}\right) + S_j$	$[b + c(n-s)]\left(j - \frac{(s+1)}{2}\right)$ $+ c\left(j^2 - \frac{(s+1)(2s+1)}{6}\right) + S_j$	$\frac{b(1-e^{cn})}{m}\left\{\frac{e^{cj}}{1-e^{cs}} - \frac{e^{c}}{s(1-e^{c})}\right\} + S_j$
$\sum(\overline{X}_{.j} - \overline{X})^2_{..}$	$\frac{b^2 s(s+1)(s-1)}{12}$ $+ 2bC_1 + \sum S_j^2$	$\frac{(s-1)s(s+1)}{180}\left\{\begin{array}{l} c^2(2s+1)(8s+11) \\ + 15[b + c(n-s)][b + c(n+s+2)] \end{array}\right\}$ $+ \left\{\sum_{j=1}^{s} S_j^2 + 2[b + c(n-s)]\sum_{j=1}^{s} jS_j + 2c\sum_{j=1}^{s} j^2 S_j\right\}$	$\frac{[be^{cn}(1-e^{cn})]^2}{m}\left\{\frac{1-e^{2cs}}{(1-e^{2c})(1-e^{cs})^2} - \frac{1}{s(1-e^{c})^2}\right\}$ $\cdot \frac{2be^{cn}(1-e^{cn})}{m^2(1-e^{cs})} C_3 + \sum S_j$
$\frac{\sum(\overline{X}_{.j} - \overline{X})^2}{s-1}$	$\frac{b^2 s(s+1)}{12}$ $+ \frac{1}{s-1}\left\{2bC_1 + \sum S_j^2\right\}$	$\frac{s(s+1)}{180}\left\{\begin{array}{l} c^2(2s+1)(8s+11) \\ + 15[b + c(n-s)][b + c(n+s+2)] \end{array}\right\}$ $+ \frac{1}{s-1}\left\{\sum_{j=1}^{s} S_j^2 + 2[b + c(n-s)]\sum_{j=1}^{s} jS_j + 2c\sum_{j=1}^{s} j^2 S_j\right\}$	$\frac{[be^{cn}(1-e^{cn})]^2}{m^2(s-1)}\left\{\frac{1-e^{2cs}}{(1-e^{2c})(1-e^{cs})^2} - \frac{1}{s(1-e^{c})^2}\right\}$ $\cdot \frac{2be^{cn}(1-e^{cn})}{m(s-1)(1-e^{cs})} C_3 + \frac{1}{(s-1)}\sum S_j$
under $H_0: S_j = 0$, for all j	$\frac{b^2 s(s+1)}{12}$	$\frac{s(s+1)}{180}\left\{\begin{array}{l} c^2(2s+1)(8s+11) \\ + 15[b + c(n-s)][b + c(n+s+2)] \end{array}\right\}$	$\frac{[be^{cn}(1-e^{cn})]^2}{m^2(s-1)}\left\{\frac{1-e^{2cs}}{(1-e^{2c})(1-e^{cs})^2} - \frac{1}{s(1-e^{c})^2}\right\}$
χ^2_c	$\frac{\hat{\sigma}^2_{\overline{X}_{.j}}}{b^2 s(s+1)/12}$	$\hat{\sigma}^2_{\overline{X}_{.j}} \Big/ \frac{s(s+1)}{180}\left\{\begin{array}{l} c^2(2s+1)(8s+11) \\ + 15[b + c(n-s)][b + c(n+s+2)] \end{array}\right\}$	$\hat{\sigma}^2_{\overline{X}_{.j}} \Big/ \frac{[be^{cn}(1-e^{cn})]^2}{m^2(s-1)}\left\{\frac{1-e^{2cs}}{(1-e^{2c})(1-e^{cs})^2} - \frac{1}{s(1-e^{c})^2}\right\}$

6. Estimates of Trend Parameters and Seasonal Indices from Row, Column and Overall Averages and Variances

Using the expressions in Tables 3.1 through 3.3, the estimates of trend parameters and seasonal indices given in Tables 5.1 through 5.3 are derived, respectively, for the three selected trending curves.

Table 5.1. Estimates of parameters of linear trend-cycle components and seasonal indices

Parameter	Model	
	Additive model	Multiplicative model
a	$a' + \hat{b}\left(\frac{s-1}{2}\right)$	$a' + \hat{b}\left(s - \frac{C_1}{s}\right)$
b	$\frac{b'}{s}$	$\frac{b'}{s}$
S_j	$\overline{X}_{.j} - \left(\hat{a} + \frac{\hat{b}}{2}(n - s + 2j)\right)$	$\hat{S}_j = \frac{\hat{\sigma}_j}{\hat{b}\sqrt{\frac{n(n+s)}{12}}}$

Note. (1) a', b' and c' are estimates derived from the regression equations of row averages on row number.

(2) Additive and multiplicative models give different estimates.

Table 5.2. Estimates of parameters of quadratic trend-cycle component and seasonal indices

Parameter	Model	
	Additive model	Multiplicative model
a	$\hat{a}' + \left(\frac{s-1}{2}\right)\hat{b} - \left(\frac{(s-1)(2s-1)}{6}\right)\hat{c}$	$\hat{a}' + \hat{b}\left(s - \frac{\hat{C}_1}{s}\right) - \hat{c}\left(s^2 - 2\hat{C}_1 + \frac{\hat{C}_2}{s}\right)$
b	$\hat{b} = \frac{\hat{b}'}{s} + \hat{c}(s-1)$	$\frac{1}{s}[b' + 2c(s^2 - C_1)]$
c	$\hat{c} = \frac{c'}{s^2}$	$\hat{c} = \frac{c'}{s^2}$

S_j	$\hat{S}_j = \overline{X}_{.j} - d_j$	$\hat{S}_j = \overline{X}_{.j} / d_j$
d_j	$\hat{a} + \frac{\hat{b}}{2}(n-s) + \frac{\hat{c}(n-s)(2n-s)}{6} + (\hat{b} + \hat{c}(n-s))j + \hat{c}j^2$	$\hat{a} + \frac{\hat{b}}{2}(n-s) + \frac{\hat{c}(n-s)(2n-s)}{6} + (\hat{b} + \hat{c}(n-s))j + \hat{c}j^2$

Note. (1) $\hat{C}_1 = \sum_{j=1}^{s} j\hat{S}_j,\ \hat{C}_2 = \sum_{j=1}^{s} j^2\hat{S}_j.$

(2) a', b' and c' are estimates derived from the regression equations of row averages on row number.

(3) Additive and multiplicative models give different estimates.

Table 5.3. Estimates of parameters of exponential trend-cycle component and seasonal indices

Parameter	Model	
	Additive model	Multiplicative model
b	$\hat{b}'s\left(\frac{1-e^{-\hat{c}}}{1-e^{-\hat{c}s}}\right)$	$\hat{b} = \frac{\hat{b}'se^{\hat{c}s}}{\hat{C}_3},\ \hat{C}_3 = \sum_{j=1}^{s} e^{\hat{c}j}\hat{S}_j$
c	$\frac{\hat{c}'}{s}$	$\hat{c} = \frac{\hat{c}'}{s}$
S_j	$\overline{X}_{.j} - \frac{\hat{b}}{m}\left(\frac{1-e^{\hat{c}s}}{1-e^{\hat{c}}}\right)e^{\hat{c}j}$	$\begin{cases} \frac{s\overline{X}_{.1}}{\sum_{j=1}^{s}\overline{X}_{.j}e^{-(j-1)\hat{c}}}, & j = 1 \\ \left(\frac{\overline{X}_{.j}}{\overline{X}_{.1}}\right)\hat{S}_1 e^{-(j-1)\hat{c}}, & j = 1, 2, ..., s \end{cases}$

Note. (1) a', b' and c' are estimates derived from the regression equations of row averages on row number.

(2) Additive and multiplicative models give different estimates.

7. Concluding Remarks

The Buys Ballot table is very useful for diagnosing the presence or absence of trend and seasonal effects in time series. It is also useful for the estimation of trend parameters and seasonal indices in time series. It helps in determining the model structure: additive or multiplicative model. The Buys Ballot procedures discussed in this paper are easy to understand and easy to apply.

It does have its drawbacks, however. First, the Buys Ballot procedure does not have an explicit way of taking into account cycles, missing values and outliers that might be presented in the data. The Buys Ballot procedure as developed is for data that has stable seasonal pattern and will perform poorly in the presence of seasonal patterns that are not stable over time.

Appendix

Overall sample variance for quadratic trend.

(a) Additive model

$$\hat{\sigma}^2_{..} = \frac{nc^2}{n-1}\left\{\begin{array}{l}\frac{(n^2-s^2)(2n-s)(8n-11s)}{180} + \frac{(s^2-1)(2s+1)(8s+1)}{180} \\ + \frac{(n-s)(s+1)(6n^2+7ns-n+s^2+5s+6)}{36}\end{array}\right\}$$

$$+ \frac{bcn(n+1)^2}{6} + \frac{b^2 n(n+1)}{12}$$

$$+ \frac{n}{s(n-1)}\left\{\sum_{j=1}^{s} S_j^2 + 2[b + c(n-s)]C_1 + 2cC_2\right\}.$$

(b) Multiplicative model

$$\hat{\sigma}^2 = \frac{n(n^2-s^2)}{n-1}\left\{\frac{c^2(2n-s)(8n-11s)}{180}\right.$$

$$+\frac{1}{12}\left[\left(b+\frac{C_1}{s^2}\right)^2+2(n-s)\left(b+\frac{C_1}{s^2}\right)\right]\Bigg\}$$

$$+\frac{n}{n-1}\left\{\left[a+b\left(\frac{n-s}{2}\right)+\frac{c(n-s)(2n-s)}{6}\right]^2\right.$$

$$\left.+\frac{n^2-s^2}{12}[b+c(n-s)]^2+\frac{c^2(n^2-4s^2)}{12}\right\}\operatorname{var}(S_j)$$

$$+\frac{n}{n-1}\left\{\left[a+\left(b+\frac{c(n-s)}{2}\right)^2+\frac{c^2(n-s)(13n-5s)}{12}\right]\operatorname{var}(jS_j)\right.$$

$$\left.+c^2\operatorname{var}(j^2S_j)\right\}$$

$$+\frac{2n}{n-1}\left[a+b\left(\frac{n-s}{2}\right)+\frac{cn(n-s)}{2}\right][b+c(n-s)]\operatorname{cov}(S_j,\ jS_j)$$

$$+\frac{2nc}{n-1}\left\{\left[a+bc\left(\frac{n-s}{2}\right)+\frac{c(n-s)(2n-s)}{6}\right]\operatorname{cov}(S_j,\ j^2S_j)\right.$$

$$\left.+[b+c(n-s)]\operatorname{cov}(jS_j,\ j^2S_j)\right\}.$$

References

[1] D. Allingham and J. C. W. Rayner, A nonparametric two-sample Wald test of equality of variances, Hindawi Publishing Corporation Advances in Decision Sciences, 2011.

[2] M. S. Bartlett, Properties of sufficiency and statistical tests, Proc. Roy Soc. Ser. A 160 (1937), 268-282.

[3] J. J. Beersma and T. A. Buishand, A simple test for equality of variances in monthly climate data, J. Climate 12 (1999), 1770-1779.

[4] G. E. P. Box and S. L. Anderson, Permutation theory in the derivation of robust criteria and the study of departures from assumption, J. Roy Statist. Soc. Ser. B 17 (1955), 1-26.

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

[6] J. M. Chambers, W. S. Cleveland, B. Kleiner and P. A. Tukey, Graphical Methods for Data Analysis, Wadsworth, Belmont, CA, 1983.

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

[8] W. S. Cleveland, Visualizing Data, AT&T, Murray Hill, NJ, 1993.

[9] W. J. Conover, Practical Nonparametric Statistics, 3rd ed., John Wiley and Sons, New York, 1999.

[10] A. M. Davey and B. E. Flores, Identification of seasonality in time series: a note, Math. Comput. Modelling 18(6) (1993), 73-81.

[11] E. S. Gardner, Jr., Exponential smoothing: the state of the art, J. Forecast. 4 (1985), 1-28.

[12] T. B. Fomby, Buys Ballot plots: graphical methods for detecting seasonality in time series, 2008.
faculty.smu.edu/tfomby/eco5375/data/…/Buys%20Ballot%20Plots.pdf.

[13] T. B. Fomby, Stable seasonal pattern (SSP) model, 2010.
faculty.smu.edu/tfomby/eco5375/data/Notes/ Stable%20seasonal%20Pattern%20Model_v2.pdf.

[14] I. S. Iwueze, Some implications of truncating the $N(1, \sigma^2)$ on the left at zero, Journal of Applied Sciences 7 (2007), 189-195.

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

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

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

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

[19] I. S. Iwueze, E. C. Nwogu and J. C. Ajaraogu, Best linear unbiased estimate using Buys-Ballot procedure when trend-cycle component is linear, Pakistan Journal of Statistics and Operations Research 7(2) (2011b), 183-198.

[20] M. Kendall and J. K. Ord, Time Series, 3rd ed., Griffin, London, 1990.

[21] M. W. J. Layard, Robust large-sample tests for homogeneity of variances, J. Amer. Statist. Assoc. 68 (1973), 195-198.

[22] H. Levene, Robust test for equality of variances, Contributions to Probability and Statistics, I. Olkin, S. G. Ghury, W. Hoeffding, W. G. Madow and H. B. Mann, eds., Stanford Univ. Press, Stanford, 1960, pp. 275-292.

[23] G. M. Ljung and G. E. P. Box, On a measure of lack of fit in time series models, Biometrika 65 (1978), 297-304.

[24] S. G. Makridakis and S. C. Wheelwright, Forecasting Methods for Management, 5th ed., Wiley, New York, 1989.

[25] S. G. Makridakis, S. C. Wheelwright and V. E. McGee, Forecasting: Methods and Applications, 2nd ed., Wiley, New York, 1983.

[26] R. G. Miller, Jr., Beyond ANOVA, Basics of Applied Statistics, John Wiley & Sons, New York, 1986.

[27] D. C. Montgomery, L. A. Johnson and J. S. Gardiner, Forecasting and Time Series Analysis, 2nd ed., McGraw-Hill, New York, 1990.

[28] E. C. Nwogu, I. S. Iwueze and H. C. Iwu, Best linear unbiased estimate of linear trend-cycle component based on FBE-derived variables, J. Math. Comput. Sci. 2(2) (2012), 189-205.

[29] R. G. O'Brien, Robust techniques for testing heterogeneity of variance effects in factorial designs, Psychometrika 43 (1978), 327-342.

[30] A. Pankratz, Forecasting with Univariate Box-Jenkins Models: Concepts and Cases, Wiley, New York, 1983.

[31] S. C. Sharma, A new Jacknife test for homogeneity of variance, Comm. Statist. Simulation Comput. 20 (1991), 479-495.

[32] G. W. Snedecor and W. G. Cochran, Statistical Methods, 8th ed., Iowa State University Press, 1989.

[33] P. F. Velleman and D. C. Hoaglin, Applications, Basics, and Computing of Exploratory Data Analysis, Duxbury Press, Belmont, CA, 1981.

[34] L. Wang and Xiao-Hua Zhou, A fully nonparametric diagnostic test for homogeneity of variances, Canad. J. Statist. 33 (2005), 545-558.

[35] W. W. Wei, Time Series Analysis: Univariate and Multivariate Methods, Addison-Wesley, New York, 1989.

[36] H. Wold, A Study in the Analysis of Stationary Time Series, 2nd ed., Almqrist and Witsett, Stockholm, 1938.