GLOBAL JOURNAL OF MATHEMATICAL SCIENCES VOL. 3, NO.2, 2004: 83 - 98
 ISSN 1596 - 6208

BUYS-BALLOT ESTIMATES FOR TIME SERIES DECOMPOSITION

I. S. IWUEZE and E. C. NWOGU

(Received 6 August 2003; Revision accepted 17 May, 2004).

ABSTRACT

An estimation procedure based on the Buys – Ballot (1847) table for time series decomposition is given in this paper. We give two alternative methods called the Chain Base Estimation and Fixed Base Estimation methods. Simulated examples are used to illustrate the methods, while comparing them with the least squares approach. U.S. quarterly beer production is re-analysed and the descriptive model obtained is shown to outperform the ARIMA model of Wei (1989) in terms of forecasts.

KEY WORDS: Trend, Seasonality, Cycles, Decomposition, Periodicity, Buys-Ballot Estimate.

1. INTRODUCTION

One of the aims of time series analysis is description of a series. Description includes the examination of trend, seasonality, cycles, turning points, changes in level, trend and scale and so on that may influence the series. This is also an important preliminary to modelling, when it has to be decided whether and how to seasonally adjust, to transform, to deal with outliers, and whether to fit a model to the entire history or only part of it.

In the examination of trend, seasonality and cycles, a time series is often described as having trends, seasonal effects, cyclic patterns and the irregular or random component. Since emphasis in time series analysis is on model building, the following additive and multiplicative models are always considered:

Additive: $X_t = T_t + S_t + C_t + I_t,\ t = 1,2,..,n,$ (1.1)

Multiplicative: $X_t = T_t \times S_t \times C_t \times I_t,\ t = 1,2,..,n$ (1.2)

where, for the time t, X_t denotes the observed value of the series, T_t is the trend, S_t the seasonal term, C_t the cyclic term, and I_t is the irregular component of the series.

Other analysts (Chatfield (1980), Kendal (1973)) may go further to consider 'mixed' models.

Cyclical variation refers to the long term oscillations or swings about the trend and only long period sets of data will show cyclical fluctuation of any appreciable magnitude. If short period of time are involved (which is true of all examples of this paper), the cyclical component is superimposed into the trend (Chatfield (1980), p. 13) and we obtain a

trend–cycle component denoted by M_t. In this case, equations (1.1) and (1.2) may be written respectively, as

$$X_t = M_t + S_t + I_t,\ t = 1,2,..,n, \tag{1.3}$$

and

$$X_t = M_t \times S_t \times I_t,\ t = 1,2,..,n \tag{1.4}$$

Using (1.3) or (1.4) we can estimate the three components of our model and hence 'decompose the series into its component parts. A summary of the traditional methods of decomposition of time series will be given in Section 2.

I. S. IWUEZE, Department of Statistics, Abia State Univesity, P. M B. 2000, Uturu, Nigeria.
E. C. NWOGU, Dept. of Mathematical and Physical Sciences, Michael Okpara University of Agriculture, Umudike, Nigeria

If a time series contains seasonal effects with period s (length of the periodic interval), we expect observations separated by multiples of s to be similar: X_t should be similar to $X_{t \pm is}$, $i = 1,2,3,...$ To analyse the data, it is helpful to arrange the series in a two-dimensional table (Table 1), according to the period and season, including the totals and / or averages. Such two-way tables that display within-period pattern, that are similar from period to period are known as Buys- Ballot tables. Wold (1938) credits these arrangements of the table to Buys-Ballot (1847)

Buys – Ballot table helps in the assessment of the trend – cycle (simply referred to as trend) and seasonal effect of time series data. The row averages ($\bar{X}_{i.}$) estimate trend, and the differences ($\bar{X}_{.j}$ - $\bar{X}_{..}$) or the ratio ($\bar{X}_{.j}$ / $\bar{X}_{..}$) between the column averages ($\bar{X}_{.j}$) and the overall average ($\bar{X}_{..}$) estimate the seasonal effects. Outside this crude procedure of assessing trend and seasonal effects, can these row, column and the overall averages be used for the efficient estimation of the trend and seasonal effects? We give in Section 3, a new estimation procedure called Buys – Ballot estimates, that is based on these averages.

Section 4 will be devoted to the application of the two different model building procedures to a number of simulated and real time series data

Table 1: Buys – Ballot table for a seasonal time series

PERIOD	1	2	...	J	...	s	TOTAL	AVERAGE
1	X_1	X_2	...	X_j	...	X_s	$T_{1.}$	$\bar{X}_{1.}$
2	X_{s+1}	X_{s+2}	...	X_{s+j}	...	X_{2s}	$T_{2.}$	$\bar{X}_{2.}$
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
I	$X_{(i-1)s+1}$	$X_{(i-1)s+2}$	...	$X_{(i-1)s+j}$	...	X_{is}	$T_{i.}$	$\bar{X}_{i.}$
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
M	$X_{(m-1)s+1}$	$X_{(m-1)s+2}$	...	$X_{(m-1)s+j}$	...	X_{ms}	$T_{m.}$	$\bar{X}_{m.}$
TOTAL	$T_{.1}$	$T_{.2}$	...	$T_{.j}$	...	$T_{.s}$	T	
AVERAGE	$\bar{X}_{.1}$	$\bar{X}_{.2}$	...	$\bar{X}_{.j}$	...	$\bar{X}_{.s}$		$\bar{X}_{..}$

where, m = number of periods
s = length of the periodic interval/length of periodicity
n = ms

2. TRADITIONAL METHOD OF DECOMPOSITION

The task of the analyst dealing with a time series for descriptive purposes is to segregate each factor or component in so far as this is possible. By isolating or removing individual components the impact of each may be assessed (Chatfield (1980)). Either of the models (1.3) or (1.4) may be used to effect the decomposition.

The first step will usually be to estimate M_t and then to eliminate M_t for each time period from the actual data either by subtraction (for equation (1.3)) or division (for equation (1.4)), giving a detrended series which expresses the effect of the seasons and the irregular component. Of all the methods of trend analysis, the fitting of a mathematical trend curve to time series data are now more usually adopted, and we concentrate on these here. The mathematical trend curve is more often taken to be a polynomial of order p = 1 or 2. The parameters of the trend curve are obtained by least squares estimation procedure (hereafter LSE) which has been implemented in the main statistical computer packages, especially MINITAB.

To make things a little more precise, we shall define a constant level to be a 'zero' trend, and we shall assume that the seasonal effect when it exists has period s, that is, it repeats after s time periods.

$$S_{t+s} = S_t, \text{ for all } t \quad (2.1)$$

For equation (1.3) it is convenient to make the further assumption that the sum of the seasonal components over a complete period is zero.

$$\sum_{j=1}^{s} S_{t+j} = 0 \qquad (2.2)$$

Similarly, for equation (1.4) the convenient variant assumption is that the sum of the seasonal components over a complete period is s.

$$\sum_{j=1}^{s} S_{t+j} = s \qquad (2.3)$$

For series showing little trend, it is usually adequate to simply calculate the average at each season and compare it with the overall average figure, either as a difference for equation (1.3) or as a ratio for equation (1.4). 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 $X_t - M_t$ or X_t/M_t at each season, depending on whether the seasonal effect is thought to be additive or multiplicative.

We can obtain the detrended, deseasonalized series by eliminating the trend - cycle M_t and seasonal component S_t for each time period from the actual data by subtraction (for equation (1.3)) or division (for equation (1.4)). This gives the residual or irregular component.

Having fitted a model to a time series, one often wants to see if the residuals are purely random. Testing residuals for randomness is a somewhat different problem. Our own preference in testing residuals for randomness is just to look at the first few values of the autocorrelation function (ACF), particularly at lag one and the first seasonal lag (if any) and see if any are significantly different from zero. For detailed discussion of residual analysis see Box and Jenkins (1976), Ljung and Box (1978).

3. BUYS – BALLOT ESTIMATES

The work described in Section 2 is based on (i) fitting a trend curve by some method and detrending the series (ii) using the detrended series to estimate the seasonal indices. 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. We now look at a new proposal that (i) computes the trend easily and (ii) gets over this problem of detrending a series before the seasonal effects are computed. We will restrict our discussion to the case where the trend is a straight line. That is,

$$M_t = a + bt, \text{ for all } t \qquad (3.1)$$

3.1. Additive Model

For the additive model, we consider a linear trend–cycle component (3.1) and seasonal component of period s. Ignoring the irregular component we obtain the following row, column and overall totals and means.

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

$$= [a + b((i-1)s + S_1] + [a + b((i-1)s+2) + S_2] + ...+ [a + b((i-1)s + s) + S_s]$$

$$= as + b(i-1)s^2 + b(1 + 2 + ... + s) + \sum_{j=1}^{s} S_j$$

$$= as + b(i-1)s^2 + \frac{bs}{2}(s+1)$$

$$= as + \frac{bs}{2}[(2i-1)s+1] \tag{3.2}$$

since the seasonal components over a complete period is zero.

$$\overline{X}_{i.} = \frac{T_i}{s} = a + \frac{b}{2}[(2i-1)s+1] \tag{3.3}$$

$$T_{.j} = \sum_{i=1}^{m} X_{(i-1)s+j,} \qquad , j = 1,2,\ldots,s$$

$$= [a + bj + S_j] + [a + b(j+s) + S_{j+s}] + [a + b(j+2s) + S_{j+2s}]$$

$$+ \ldots + [a + b(j + (m-1)s) + S_{j+(m-1)s}]$$

$$= ma + mbj + bs(1 + 2 + 3 + \ldots + (m-1)) + mS_j$$

$$T_{.j} = ma + mbj + \frac{bsm}{2}(m-1) + mS_j$$

$$= ma + \frac{mb}{2}[2j + n - s] + mS_j \tag{3.4}$$

$$\overline{X}_{.j} = \frac{T_{.j}}{m} = a + \frac{b}{2}(2j+n-s) + S_j \tag{3.5}$$

$$T = \sum_{i=1}^{m} T_{i.} = \sum_{j=1}^{s} T_{.j}$$

$$= na + \frac{bn}{2}(n+1) \tag{3.6}$$

$$\overline{X}_{..} = \frac{T}{n} = a + \frac{b}{2}(n+1) \tag{3.7}$$

3.2. Multiplicative Model

For the multiplicative model, we again consider a linear trend-cycle component (3.1) and a seasonal component of period s such that $\sum S_t = s$. Again, we ignore the irregular components to obtain the following results

$$T_{i.} = [a + ((i-1)s + 1)b]S_1 + [a + ((i-1)s + 2)b]S_2 + \ldots [a + ((i-1)s + s)b]S_s$$

$$= a(S_1 + S_2 + \ldots + S_s) + bs(i-1)(S_1 + S_2 + \ldots + S_s) + b(S_1 + 2S_2 + 3S_3 + \ldots + sS_s)$$

$$= as + bs^2(i-1) + bC \tag{3.8}$$

where,

$$C = S_1 + 2S_2 + \ldots + sS_s \tag{3.9}$$

Now from (1.4), we obtain

$$S_t = X_t / M_t I_t \tag{3.10}$$

If there were no seasonal components and no residual variation, the original data would be the same as the trend. That is, we expect

$$S_t = 1.0. \qquad (3.11)$$

Thus,

$$\begin{aligned} C &= S_1 + 2S_2 + 3S_3 + \ldots + sS_s \\ &= S_1 + S_2 + S_3 + \ldots + S_{s-1} + S_s \quad = s \\ &\quad + S_2 + S_3 + \ldots + S_{s-1} + S_s \quad = s-1 \\ &\quad + S_3 + \ldots + S_{s-1} + S_s \quad = s-2 \\ &\quad \vdots \\ &\quad + S_{s-1} + S_s \quad = s-(s-2) \\ &\quad + S_s \quad = s-(s-1) \end{aligned}$$

$$= s^2 - (1 + 2 + 3 + \ldots + (s-1)) = s^2 - \frac{s}{2}(s-1)$$

$$= \frac{s(s+1)}{2} \qquad (3.12a)$$

Thus,

$$T_{i.} = as + bs^2(i-1) + \frac{bs}{2}(s+1)$$

$$= as + \frac{bs}{2}[(2i-1)s+1] \qquad (3.12b)$$

$$\bar{X}_{i.} = a + \frac{b}{2}[(2i-1)s+1] \qquad (3.13)$$

results, totally in agreement with the additive case.

$$T_{.j} = [a + jb]S_j + [a + (s+j)b]S_j + [a + (2s+j)b]S_j + \ldots + [a + ((m-1)s+j)b]S_j$$

$$= [ma + mbj + bs(1 + 2 + 3 \ldots + (m-1))S_j$$

$$= [ma + mbj + \frac{mbs}{2}(m-1)]S_j$$

$$= [ma + \frac{mb}{2}(2j + n-s)]S_j \qquad (3.14)$$

$$\bar{X}_{.j} = [a + \frac{b}{2}(2j + n - s)]S_j \qquad (3.15)$$

Finally,

$$T = \sum_{i=1}^{m} T_{i.} = \sum_{j=1}^{s} T_{.j}$$

$$= na + \frac{bn}{2}(n+1) \qquad (3.16)$$

$$\overline{X}_{..} = a + \frac{b}{2}(n+1) \qquad (3.17)$$

Again, we obtain results in agreement with the additive case.

3.3 Estimates.

We now use the row, column and overall averages to estimate the parameters of the trend line and the seasonal indices.

(1). Estimation of a and b

The row averages are the same for both the additive and multiplicative models and are functions of a and b for fixed periodic interval s. Using (3.3) or (3.13), we obtain

$$\nabla \overline{X}_{i.} = \overline{X}_{i.} - \overline{X}_{(i-1).}, \quad i = 2,3,...m$$

$$= bs \qquad (3.18)$$

and

$$\overline{X}_{i.} - \overline{X}_{1.} = (i-1)bs, \quad i = 2,3,...,m \qquad (3.19)$$

From (3.18), we obtain an estimate of b as

$$\hat{b}_i^{(1)} = \frac{1}{s}\nabla \overline{X}_{i.} = \frac{1}{s}[\overline{X}_{i.} - \overline{X}_{(i-1)\cdot}] \qquad (3.20)$$

and from (3.19), we obtain

$$\hat{b}_i^{(2)} = \frac{1}{(i-1)s}(\overline{X}_{i} - \overline{X}_{1.}) \qquad (3.21)$$

The computation of 'b' is done by expressing the changes in the periodic averages as differences with reference to the average value at some earlier period. Two alternatives are possible.

(i) Equation (3.20) computes 'b' from the relative periodic average changes (Chain Base Estimation method (CBE))

(ii) Equation (3.21) computes 'b' using the first period as the earlier period (Fixed Base Estimation Method (FBE)).

The two possibilities will each give rise to (m – 1) different estimates of 'b'. The average of these (m – 1) different estimates will be taken as the Buys – Ballot estimate of 'b'.

Having estimated 'b', we use (3.3) or (3.13) to find m different estimates of 'a'. Again the average of these m different values of 'a' will be taken as the Buys-Ballot estimate of 'a'.

That is,

$$\hat{a} = \frac{1}{m}\sum_{i=1}^{m} a_i = \frac{1}{m}\sum_{i=1}^{m}\left\langle \overline{X}_{i.} - \frac{b}{2}[(2i-1)S+1]\right\rangle$$

$$= \frac{1}{m}\left[m\overline{X}_{..} - \frac{\hat{b}m}{2}(ms+1)\right]$$

$$= \overline{X}_{..} - \frac{\hat{b}}{2}(n+1) \qquad (3.22)$$

(2.) Estimate of S_j, j = 1,2,..,s

The seasonal indices are thereafter obtained from (3.5) for the additive model and from (3.15) for th multiplicative model. For the additive model

$$\hat{S}_j = \overline{X}_{.j} - \hat{a} - \frac{\hat{b}}{2}[2j+n-s]$$

$$= \overline{X}_{.j} - \overline{X}_{..} - \frac{\hat{b}}{2}(2j-s-1) \qquad (3.23)$$

and for the multiplicative model

$$\hat{S}_j = \overline{X}_{.j}\Big/\left[\hat{a} + \frac{\hat{b}}{2}(2j+n-s)\right]$$

$$= \overline{X}_{.j}\Big/\left[\overline{X}_{..} + \frac{\hat{b}}{2}(2j-s-1)\right] \qquad (3.24)$$

When there is no trend (b = 0), it is clear from (3.23) and (3.24) that

$$\hat{S}_j = \overline{X}_{.j} - \overline{X} \qquad (3.25)$$

for the additive model, and

$$\hat{S}_j = \overline{X}_{.j}/\overline{X}_{..} \qquad (3.26)$$

for the multiplicative; results already indicated in Section 2.

4. EMPIRICAL EXAMPLES

4.1 Simulation of Additive Model

This example shows a simulation of 100 values from an additive model

$$X_t = a + bt + S_t + e_t \qquad (4.1)$$

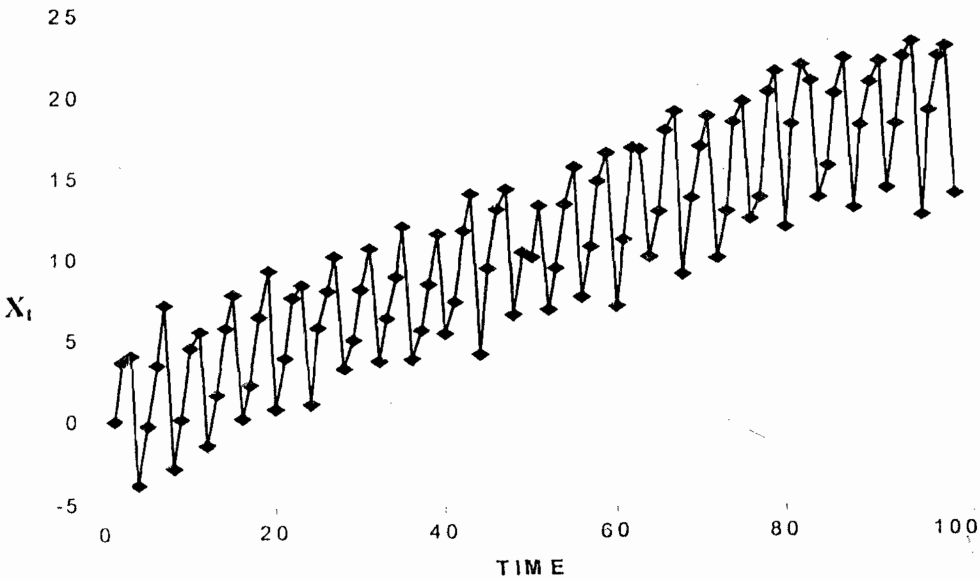

Fig.1: A simulated additive series; $X_t = 1.0 + 0.2t + S_t + e_t$ with $S_1 = -1.5$, $S_2 = 2.5$, $S_3 = 3.5$, $S_4 = -4.5$, $e_t \sim N(0, 1)$

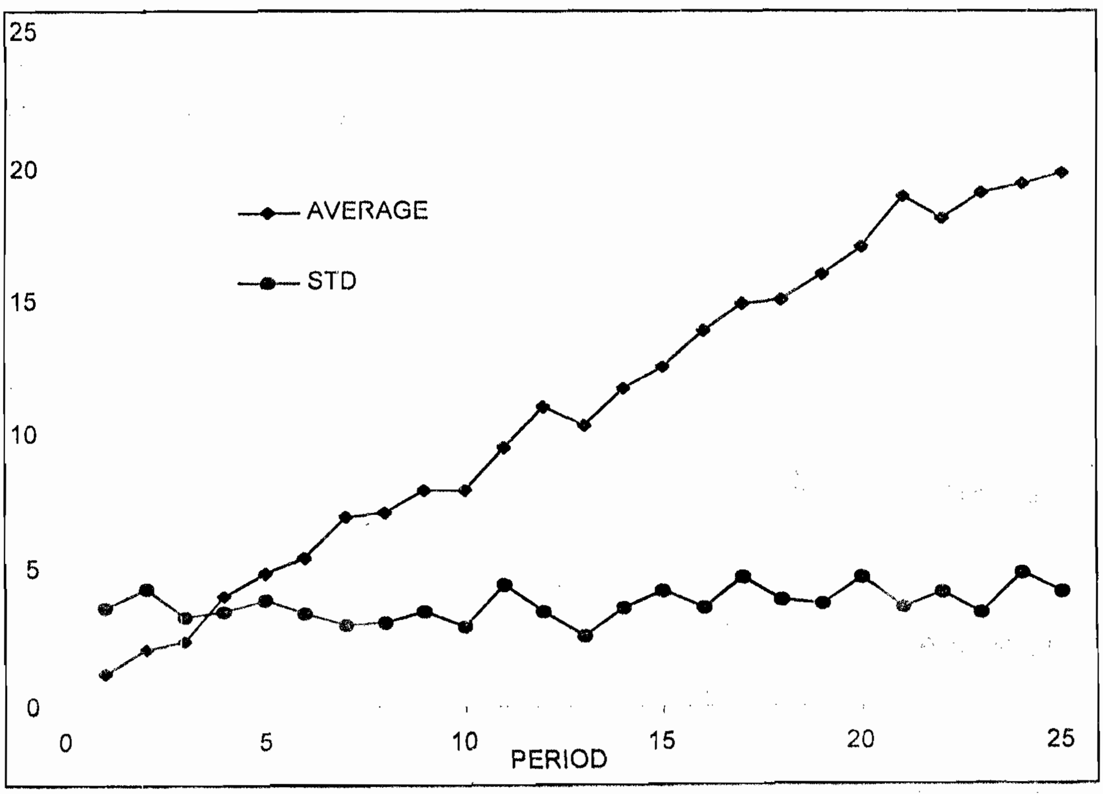

Fig.2: Mean and Standard deviation of $X_t = 1.0 + 0.2t + S_t + e_t$ with $S_1 = -1.5$, $S_2 = 2.5$, $S_3 = 3.5$, $S_4 = -4.5$, $e_t \sim N(0, 1)$

with a = 1.0, b = 0.2, $S_1 = -1.5$, $S_2 = 2.5$, $S_3 = 3.5$, $S_4 = -4.5$ and e_t being Gaussian N(0 , 1) white noise. The series is listed in Table 2 with its row and column averages and standard deviation. As shown in Figures 1 and 2, it is clearly seasonal with an upward trend and stable variance. There is an upsurge of the series in the second and third quarters and a sharp drop in the first and fourth quarters. There is a reasonably stable seasonal pattern over the periods suggesting the additive model.

In order to assess the forecasting performance of our models in examples 4.1 and 4.2, we use only the first 96 observations of the series for model construction. When an additive model is adequate, we can use it to forecast future values. For a forecast origin, say $t = n_o$, forecasts can be calculated directly from (1.3) to give the ℓ - step ahead forecast as

Table 2: Simulated data from $X_t = a + bt + S_t + e_t$ with $s = 4$, $a = 1.0$, $b = 0.2$, $S_1 = -1.5$, $S_2 = 2.5$, $S_3 = 3.5$, $S_4 = -4.5$, $e_t \sim N(0,1)$.

	SEASON						
PERIOD	I	II	III	IV	TOTAL	AVERAGE	STD
1	0.2819	3.9074	4.2887	-3.5180	4.9600	1.240000	3.649937
2	-0.0577	3.7369	7.3987	-2.5590	8.5189	2.129725	4.363259
3	0.3625	4.7249	5.7068	-1.1943	9.5999	2.399975	3.337189
4	1.8663	5.9175	7.9937	0.3982	16.1757	4.043925	3.518729
5	2.3951	6.6492	9.4859	0.9470	19.4772	4.869300	3.915233
6	4.0932	7.7541	8.5962	1.2083	21.6518	5.412950	3.417352
7	5.9109	8.2203	10.3576	3.3829	27.8717	6.967925	3.001558
8	5.1878	8.3220	10.8463	3.9272	28.2833	7.070825	3.122382
9	6.5015	9.0910	12.1945	3.9848	31.7718	7.942950	3.518448
10	5.7133	8.6010	11.7572	5.5106	31.5821	7.895525	2.935995
11	7.4686	11.8789	14.2801	4.2953	37.9229	9.480725	4.461783
12	9.5368	13.2334	14.5086	6.7287	44.0075	11.001875	3.544164
13	10.5704	10.2714	13.4543	6.9778	41.2739	10.318475	2.649477
14	9.5822	13.5691	15.8684	7.8203	46.8400	11.710000	3.669961
15	10.9075	15.0572	16.7887	7.2621	50.0155	12.503875	4.277998
16	11.2988	17.0678	17.0045	10.2369	55.6080	13.902000	3.644969
17	13.1137	18.2031	19.3734	9.1720	59.8622	14.965550	4.722480
18	13.9217	17.1874	19.0695	10.1657	60.3443	15.086075	3.909352
19	13.0921	18.6454	19.9549	12.5806	64.2730	16.068250	3.775753
20	13.9338	20.5122	21.7753	12.0939	68.3152	17.078800	4.781409
21	18.5196	22.1662	21.1888	13.9998	75.8744	18.968600	3.653515
22	15.9577	20.4450	22.6041	13.3204	72.3272	18.081800	4.211834
23	18.4417	21.0441	22.3748	14.5645	76.4251	19.106275	3.440339
24	18.5176	22.6546	23.5647	12.8141	77.5510	19.387750	4.902015
25	19.3495	22.6847	23.2599	14.1771	79.4712	19.867800	4.167092
TOTAL	236.4665	331.5448	373.6956	168.2969	1110.0038		
AVERAGE	9.458660	13.261792	14.947824	6.731878		11.100038	
STD	6.156215	6.411410	5.936666	5.541034			

STD = Standard Deviation

$$\hat{X}_{n_o}(\ell) = \hat{M}_{n_o+\ell} + \hat{S}_{n_o} \qquad (4.2)$$

where,

$$\hat{M}_{n_{o+\ell}} = \hat{a} + \hat{b}(n_o + \ell) \qquad (4.3)$$

Let the ℓ-step ahea

$$\hat{e}_\ell \quad n_o + \quad n_o(\ell) \qquad 4)$$

The comparison is usually based on the following summary statistics

1. Mean percentage error, which is also referred to as bias since it measures forecast bias

$$\text{MPE} = \left(\frac{1}{m_o} \sum_{\ell=1}^{m_o} \frac{e_\ell}{X_{n_o+\ell}} \right) 100\% \qquad (4.5)$$

2. **Mean square error**

$$MSE = \frac{1}{m_o}\sum_{\ell=1}^{m_o} e_\ell^2 \tag{4.6}$$

3. **Mean absolute error**

$$MAE = \frac{1}{m_o}\sum_{\ell=1}^{m_o} |e_\ell| \tag{4.7}$$

4. **Mean absolute percentage error**

$$MAPE = \left(\frac{1}{m_o}\sum_{\ell=1}^{m_o}\left|\frac{e_\ell}{X_{n_o+\ell}}\right|\right)100\% \tag{4.8}$$

LSE Estimates

Trend analysis by LSE procedures (all trend analysis in this paper are performed using MINITAB) gave the following estimates

$$\hat{M}_t = \underset{(\pm 0.7053)}{0.8971} + \underset{(\pm 0.0126)}{0.2028t} \tag{4.9}$$

where values in parentheses below the parameter estimates are the associated standard errors. The associated seasonal analysis procedure estimates the seasonal indices as

$$\hat{S}_1 = -1.3822,\ \hat{S}_2 = 2.2377,\ \hat{S}_3 = 3.7672,\ \hat{S}_4 = -4.6154$$

which, as can be noticed, sum to 0.0073. Since we require the seasonals to sum to zero we add a correction factor of

$$\frac{-0.0073}{4} = -0.0018$$

10

to the values above to give

$$\hat{S}_1 = -1.3840,\ \hat{S}_2 = 2.2359,\ \hat{S}_3 = 3.7654,\ \hat{S}_4 = -4.6173$$

We can then examine the irregular/residual component of our series which we estimate by subtracting both the trend line and seasonal effects. The residual ACF indicate no model inadequacy with residual mean = 0.0002 and residual variance = 0.9107

Buys-Ballot Estimates

The computational procedure for the Buys-Ballot estimates for the trend line is laid out in Table 3. The seasonal indices are estimated using Equation (3.23). The two separate Buys-Ballot estimates can each be used to obtain component analysis tables. The irregular components obtained can then be checked for randomness. For both methods, the residual ACF's indicate no model inadequacy.

Table 3: Buys-Ballot estimates for the trend line of the additive model

S/No	$\overline{X}_{i.}$	CBE			FBE		
		$\nabla\overline{X}_{i.}$	$\hat{b}^{(1)}$	$\hat{a}^{(1)}$	$\overline{X}_{i.} - \overline{X}_{1.}$	$\hat{b}^{(2)}$	$\hat{a}^{(2)}$
1	1.240000	-	-	0.746855	-	-	0.722263
2	2.129725	0.889725	0.222431	0.847548	0.889725	0.222431	0.783608
3	2.399975	0.270250	0.067563	0.328766	1.159975	0.144997	0.225478
4	4.043925	1.643950	0.410988	1.183684	2.803925	0.233660	1.041048
5	4.869300	0.825375	0.206344	1.220027	3.629300	0.226831	1.038043
6	5.412950	0.543650	0.135913	0.974645	4.172950	0.208648	0.753313
7	6.967925	1.554975	0.388744	1.740588	5.727925	0.238664	1.479908
8	7.070825	0.102900	0.025725	1.054456	5.830825	0.208244	0.754428
9	7.942950	0.872125	0.218031	1.137549	6.702950	0.209467	0.798173
10	7.895525	-0.047425	-0.011856	0.301092	6.655525	0.184876	-0.077633
11	9.480725	1.585200	0.396300	1.097260	8.240725	0.206018	0.679188
12	11.001875	1.521150	0.380288	1.829378	9.761875	0.221861	1.371958
13	10.318475	-0.683400	-0.170850	0.356946	9.078475	0.189135	-0.139823
14	11.710000	1.391525	0.347881	0.959439	10.470000	0.201346	0.423323
15	12.503875	0.793875	0.198469	0.964282	11.263875	0.201141	0.388818
16	13.902000	1.398125	0.349531	1.573375	12.662000	0.211033	0.958563
17	14.96550	1.063550	0.265888	1.847893	13.725550	0.214462	1.193733
18	15.086075	0.120525	0.030131	1.179386	13.846075	0.203619	0.485878
19	16.068250	0.982175	0.245544	1.372529	14.828250	0.205948	0.639673
20	17.078800	1.010550	0.252638	1.594047	15.838800	0.208405	0.821843
21	18.968600	1.889800	0.472450	2.694815	17.728600	0.221608	1.883263
22	18.081800	-0.886800	-0.221700	1.018983	16.841800	0.200498	0.168083
23	19.106275	1.024475	0.256119	1.254426	17.866275	0.203026	0.364178
24	19.387750	0.281475	0.070369	0.746869	18.147750	0.197258	-0.182728
TOTAL			4.536941	28.024838		4.763176	16.574579
AVERAGE			0.197258	1.167702		0.207095	0.690607
STD.			0.183475	0.539379		0.018768	0.509483

The estimates and the corresponding error means and variances are shown in Table 4. The results indicate that CBE method better estimates the error variance. Finally, we calculate the ℓ - step – ahead forecasts $\hat{X}_{96}(\ell)$, for ℓ = 1, 2, 3, 4 from the forecast origin 96 for the three competing methods. The forecast errors and the corresponding summary statistics are shown in Table 5. The results indicate that CBE outperforms LSE and FBE in terms of forecasts.

Table 4: Summary of estimates (Additive Model)

PARAMETER	ESTIMATION METHOD			
	ACTUAL	LSE	CBE	FBE
b	0.2000	0.2028	0.1973	0.2071
		(± 0.0126)	(± 0.1835)	(± 0.0188)
a	1.0000	0.8971	1.1677	0.6906
		(± 0.7053)	(± 0.5394)	(± 0.5094)
S_1	-1.5000	-1.3840	-1.3923	-1.3776
S_2	2.5000	2.2359	2.2331	2.2380
S_3	3.5000	3.7654	3.7682	3.7633
S_4	-4.5000	-4.6173	-4.6089	-4.6236
Error Mean	0.0000	0.0002	-0.0021	-0.0002
Error Variance	1.0000	0.9107	0.9497	0.9197

Table 5: Comparison of forecasts between estimation methods (Additive Model)

Lead Time	Actual Value	LSE		CBE		FBE	
		Forecast Value	Error	Forecast Value	Error	Forecast Value	Error
1	19.3495	19.1847	0.1648	18.9135	0.4360	19.4018	-0.0523
2	22.6847	23.0074	-0.3227	22.7362	-0.0515	23.2244	-0.5397
3	23.2599	24.7397	-1.4798	24.4685	-1.2086	24.9567	-1.6968
4	14.1771	16.5598	-2.3827	16.2888	-2.1117	16.7769	-2.5998
MPE		-5.93%		-4.52%		-7.07%	
MSE		2.00		1.53		2.48	
MAE		1.09		0.95		1.22	
MAPE		6.36%		5.64%		7.07%	

4.2. Simulation of Multiplicative Model

The second example shows a simulation of 100 values from a multiplicative model

$$X_t = (a + bt)\, S_t\, e_t \tag{4.10}$$

with $a = 1.0$, $b = 0.2$, $S_1 = 0.6$, $S_2 = 1.1$, $S_3 = 0.9$ $S_4 = 1.4$ and e_t being Gaussian N (1.0 , 0.10) white noise. The Buys-Ballot table of the series is listed in Table 6. As shown in Figures 3 and 4, it is clearly seasonal with an upward trend and the variance appears to increase with the mean; suggesting the multiplicative model. The standard deviation is directly proportional to the mean showing that a logarithmic transformation is necessary to stabilize the variance.

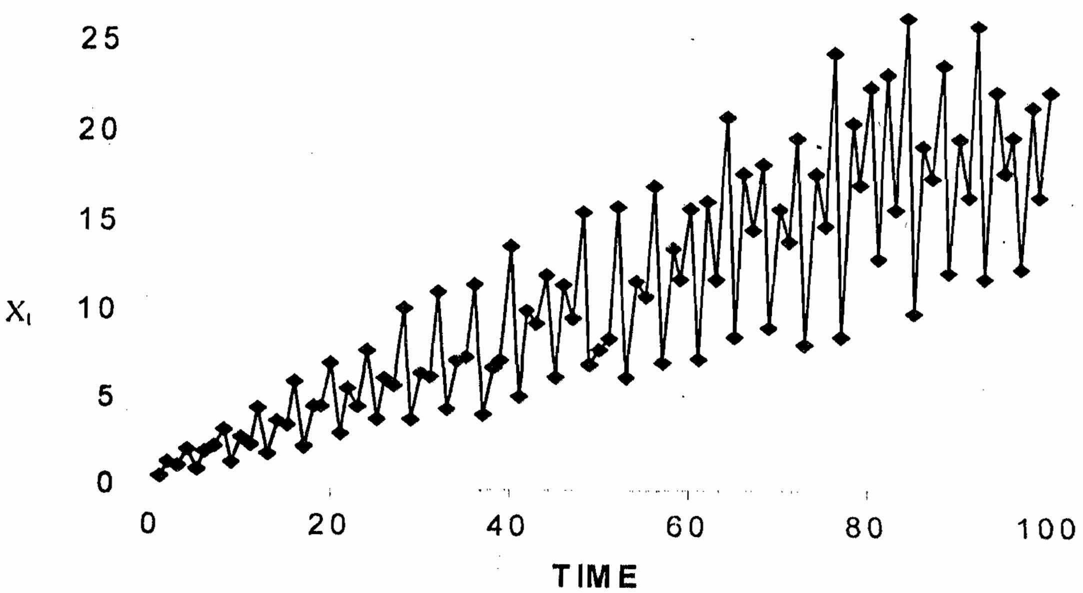

Fig.3: A simulated multiplicative series; $X_t = (1.0 + 0.2t)\, S_t\, e_t$ with $S_1 = 0.6$, $S_2 = 1.1$, $S_3 = 0.9$, $S_4 = 1.4$, $e_t \sim N(1.0 , 0.01)$

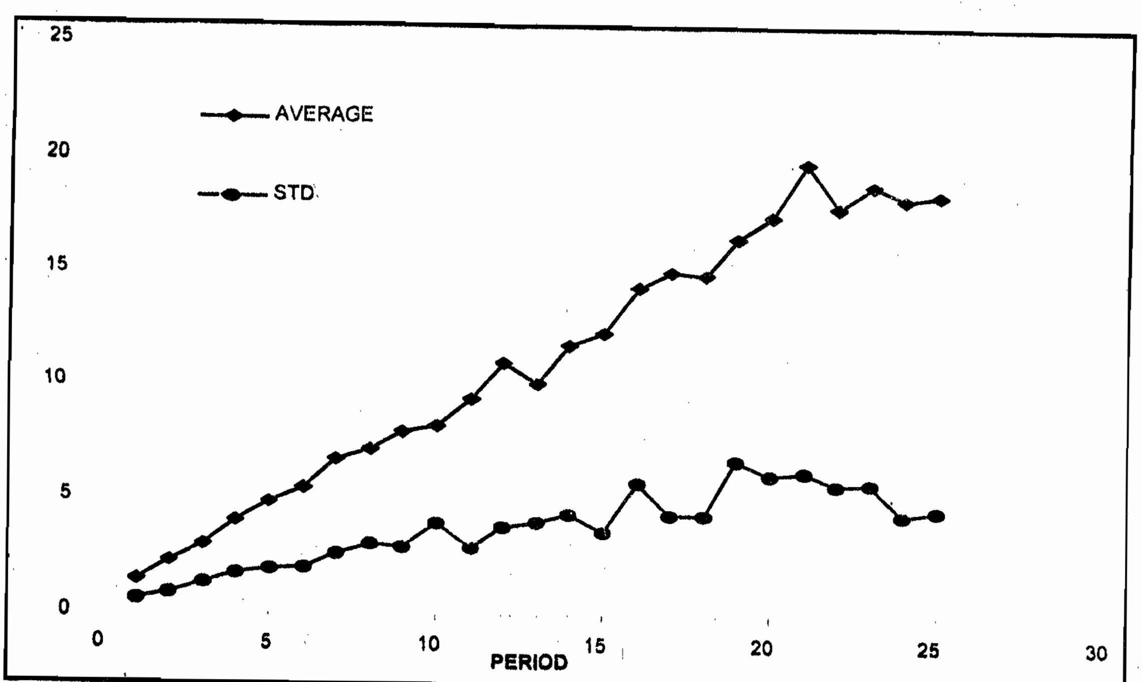

Fig.4: Mean and standard deviation of $X_t = (1.0 + 0.2t)\, S_t\, e_t$ with $S_1 = 0.6$, $S_2 = 1.1$, $S_3 = 0.9$, $S_4 = 1.4$. $e_t \sim N(1.0 , 0.01)$

For each of the estimation methods under consideration, estimation of the trend line is the same for both the additive and multiplicative models. Computational procedures described in Section 4.1 are used to obtain the trend estimates listed in Table 7. However, the seasonal analysis methods are different. For the LSE method, we average the ratios X_t/M_t at each season, while for the CBE and FBE methods we use equation (3.24). Results obtained are also listed in Table 7

Table 6. Simulated data from $X_t = (a + bt) S_t e_t$ with $s = 4$, $a = 1.0$, $b = 0.2$, $S_1 = 0.6$, $S_2 = 1.1$ $S_3 = 0.9$, $S_4 = 1.4$, $e_t \sim N(1.0, 0.10)$

	SEASON						
PERIOD	I	II	III	IV	TOTAL	AVERAGE	STD
1	0.7619	1.5411	1.3232	2.3139	5.9401	1.485025	0.642722
2	1.1331	2.1869	2.4837	3.4001	9.2038	2.300950	0.934239
3	1.5225	3.0442	2.5940	4.7151	11.8758	2.968950	1.327593
4	2.1095	4.0201	3.7777	6.2905	16.1978	4.049450	1.718567
5	2.5067	4.8319	4.8323	7.3129	19.4838	4.870950	1.962636
6	3.2427	5.8533	4.7861	8.0455	21.9276	5.481900	2.017270
7	4.1079	6.4928	6.0236	10.4254	27.0497	6.762425	2.650940
8	4.0342	6.7929	6.5748	11.4242	28.8261	7.206525	3.078018
9	4.7431	7.5427	7.7000	11.8070	31.7928	7.948200	2.909126
10	4.4419	7.0959	7.4901	13.8734	32.9013	8.225325	4.001300
11	5.3923	10.3182	9.6596	12.3416	37.7117	9.427925	2.922444
12	6.6221	11.8185	9.9296	15.7730	44.1432	11.035800	3.819167
13	7.3032	8.1934	8.8243	16.0842	40.4051	10.101275	4.037131
14	6.5996	12.0313	11.1979	17.2855	47.1143	11.778575	4.379842
15	7.4456	13.8007	12.0830	15.9470	49.2763	12.319075	3.613103
16	7.6022	16.4613	12.1231	21.1301	57.3167	14.329175	5.799928
17	8.9155	17.9678	14.8695	18.5432	60.2960	15.074000	4.411287
18	9.4321	15.9842	14.1855	19.9769	59.5786	14.894650	4.372555
19	8.4166	17.9803	15.0551	24.6772	66.1202	16.532300	6.744944
20	8.8893	20.8387	17.3507	22.8335	69.9122	17.478050	6.157858
21	13.2298	23.4775	15.9807	26.6639	79.3519	19.837975	6.281877
22	10.2143	19.5095	17.7260	24.0089	71.4597	17.864925	5.745047
23	12.5678	19.9472	16.7181	26.2488	75.4819	18.870475	5.772216
24	12.2511	22.5523	18.1165	20.1187	73.0386	18.259650	4.397199
25	12.7902	21.7189	16.7729	22.5707	73.8527	18.463175	4.564763
TOTAL	166.2752	302.0016	258.1780	383.8121	1110.2669		
AVERAGE	6.651008	12.080064	10.327120	15.352404		11.102669	
STD	3.803829	7.079607	5.403085	7.358425			

With the estimates of Table 7, component analysis tables are obtained for each estimation method and the irregular components obtained are checked for randomness. The residual ACF's of the LSE and FBE methods indicate no model inadequacies, while the residual ACF of the CBE has significant spikes at Lags 1 – 5, and hence the method lead to an inadequate model.

Finally, the ℓ - step- ahead forecasts for the two adequate methods are computed and listed in Table 8. The results indicate that FBE out performs LSE in terms of forecasts.

Table 7: Summary of estimates (Multiplicative Model)

			ESTIMATION METHOD	
PARAMETER	ACTUAL	LSE	CBE	FBE
b	0.2000	0.[illegible]	0.1823	0.2027
		(± 0.0129)	(± 0.2420)	(± 0.0126)
a	1.0000	0.8696	1.9528	0.9653
		(± 0.7225)	(± 0.9125)	(± 0.7451)
S_1	0.6000	0.6106	0.6078	0.6095
S_2	1.1000	1.0756	1.0910	1.0920
S_3	0.9000	0.9267	0.9239	0.9236
S_4	1.4000	1.3871	1.3598	1.3560
Error Mean	1.0000	0.9996	0.9538	0.9[illegible]67
Error Variance	0.1000	0.0963	0.1495	0.0963

Table 8: Comparison of forecasts between estimation methods (Multiplicative Model)

Lead Time	Actual Value	LSE		FBE	
		Forecast Value	Error	Forecast Value	Error
1	12.7902	12.6550	0.1352	12.5723	0.2179
2	21.7189	22.5125	-0.7936	22.7463	-1.0274
3	16.7729	19.5857	-2.8128	19.4131	-2.6402
4	22.5707	29.6002	-7.0295	28.7951	-6.2244
MPE		-12.63%		-11.59%	
MSE		14.49		11.70	
MAE		2.69		2.53	
MAPE		13.16%		12.44%	

4.3 U.S. Beer Production

Table 9 shows the Buys-Ballot table of 32 consecutive quarters of U.S. beer production, in millions of barrels, from the first quarter of 1975 to the fourth quarter of 1982. As shown in Table 9 and figure 5, it is clearly seasonal with a slight upward trend. There is an upsurge of the series almost of equal magnitude in the second and third quarters and a sharp drop (again of almost equal magnitude) in the first and fourth quarters. The yearly standard deviations are stable while the seasonal standard deviations differ, indicating that the series needs some transformation to make the seasonal effect additive.

Wei (1989), ignoring the stochastic trend in the series, used 30 observations of the series for ARIMA model construction. Based on the forecasting performance of his models, he settled on the model.

$$\underset{}{(1-B^4)X_t} = \underset{(\pm 0.09)}{1.49} + \underset{(\pm 0.16)}{(1-0.87B^4)e_t} \qquad (4.11)$$

with $\hat{\sigma}_e^2 = 2.39$.

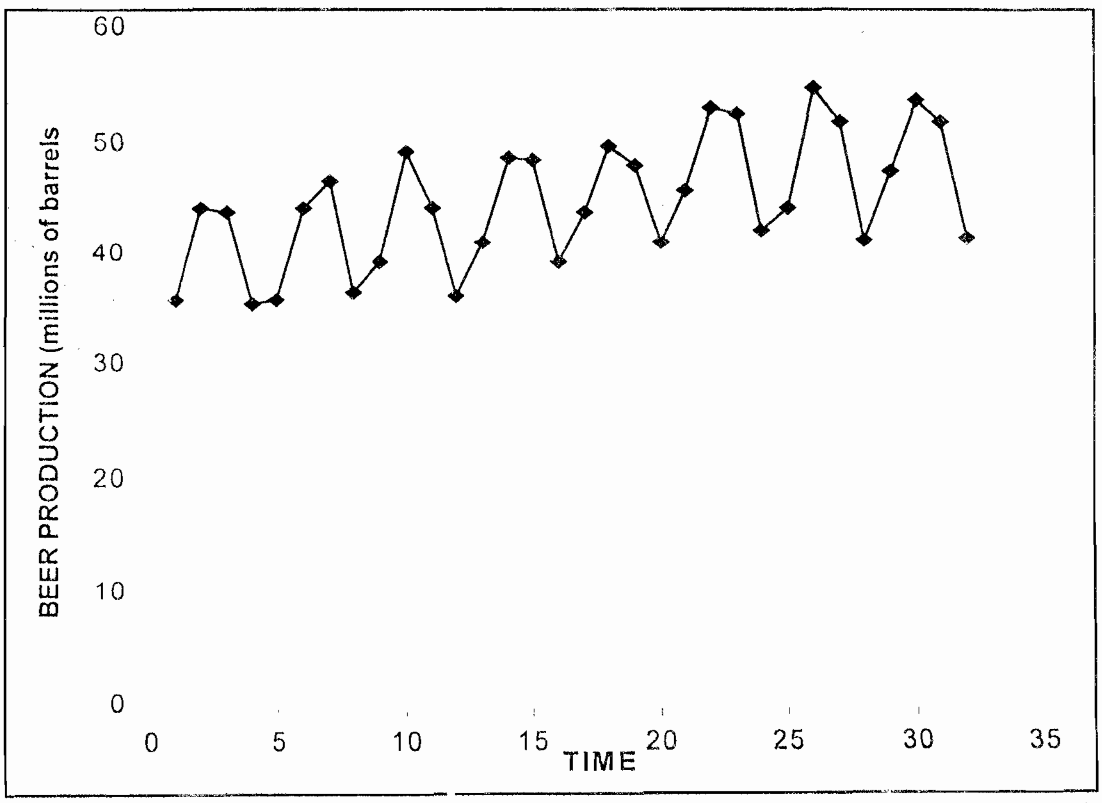

Fig.5: U.S. quarterly beer production, in millions of barrels, between 1975 and 1982

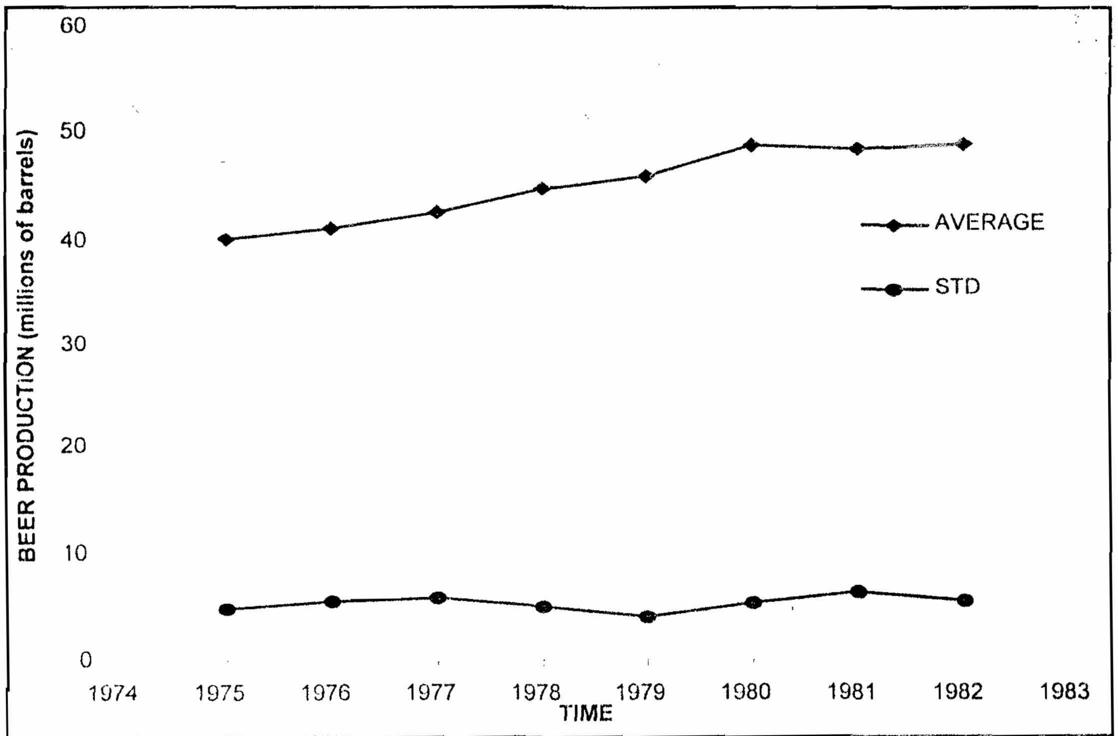

Fig.6: Yearly means and standard deviations of U.S. beer production

Table 9: U.S. Quarterly beer production in millions of barrels, between 1975 and 1982.

YEAR	QUARTER				TOTAL	AVERAGE	STD.
	I	II	III	IV			
1975	36.14	44.60	44.15	35.72	160.61	40.1525	4.8822
1976	36.19	44.63	46.95	36.90	164.67	41.1675	5.4287
1977	39.66	49.72	44.49	36.54	170.41	42.6025	5.7629
1978	41.44	49.07	48.98	39.59	179.08	44.7700	4.9711
1979	44.29	50.09	48.42	41.39	184.19	46.0475	3.9476
1980	46.11	53.44	53.00	42.52	195.07	48.7675	5.3491
1981	44.61	55.18	52.24	41.66	193.69	48.4225	6.3378
1982	47.84	54.27	52.31	41.83	196.25	49.0625	5.5217
TOTAL	336.28	401.00	390.54	316.15	1443.97		
AVERAGE	42.0350	50.1250	48.8175	39.5188		45.1241	
STD.	4.4228	4.0659	3.4967	2.7413			

The purpose of the present analysis is to demonstrate, using the Buys – Ballot modeling procedures, that descriptive models may sometimes outperform the complicated ARIMA models. Table 10 shows a summary of adequate (in terms of adequacy of residual ACF's) estimates of the additive and multiplicative models using the LSE, CBE and FBE methods. The one step ahead and two step ahead forecasts $\hat{X}_{30}(\ell)$ for $\ell = 1$ and 2 from the forecast origin 30 are calculated for each method. The forecast errors and the corresponding summary statistics are shown in Table 11.

Table 10: Summary of estimates for U.S. beer production.

PARAMETER	ESTIMATION METHOD					
	ADDITIVE MODEL			MULTIPLICATIVE MODEL		
	LSE	CBE	FBE	LSE	CBE	FBE
B	0.3804	0.3894	0.3540	0.3804	0.3894	0.3540
	(± 0.1008)	(± 0.2678)	(± 0.0588)	(± 0.1008)	(± 0.2678)	(± 0.0588)
A	39.0986	38.9484	39.5323	39.0986	38.9484	39.5323
	(± 1.7900)	(± 0.5690)	(± 0.6892)	(± 1.7900)	(± 0.5690)	(± 0.6892)
S_1	-2.6916	-2.2977	-2.3508	0.9385	0.9478	0.9467
S_2	5.0180	5.4029	5.3852	1.1116	1.1204	1.1200
S_3	3.5919	3.2071	3.2248	1.0809	1.0708	1.0712
S_4	-5.9184	-6.3123	-6.2592	0.8691	0.8610	0.8620
Error mean	-0.0780	-0.0930	-0.1260	0.9983	0.9983	0.9971
Error Variance	1.5475	1.7082	1.7345	0.0008	0.0009	0.0009

The results of Table 11 indicate that in terms of forecasts

(i) The multiplicative model outperforms the additive model for all estimation methods.

(ii) The FBE method outperforms the LSE and CBE methods

(iii) The FBE method of the multiplicative model outperforms the ARIMA model as listed in Table 8.13 of Wei (1989).

5. CONCLUSION

We have here outlined a new technique for the estimation of trend-cycle and seasonal components in descriptive time series analysis. No attempt has been made to discuss this technique when the trend-cycle component is not linear. Application when trend-cycle component is quadratic is already in preparation.

This technique is computationally simple when compared with other descriptive methods. The estimation of the slope of the line (b) is easily computed from periodic averages while the computation reduces to

$$\hat{b} = \frac{1}{m-1}\sum_{i=2}^{m}\hat{b}_i^{(1)} = \frac{1}{s(m-1)}\sum_{i=2}^{m}\left(\overline{X}_{i.} - \overline{X}_{(i-1).}\right)$$

$$= \frac{1}{(m-1)s}\left(\overline{X}_{m.} - \overline{X}_{1.}\right) \qquad (5.1)$$

for the CBE method, and

$$\hat{b} = \frac{1}{m-1}\sum_{i=2}^{m}\hat{b}_i^{(2)} = \frac{1}{(m-1)s}\sum_{i=2}^{m}\frac{(\overline{X}_{i.} - \overline{X}_{1.})}{(i-1)}$$

$$= \frac{1}{(m-1)s}\left\langle\sum_{i=2}^{m}\frac{\overline{X}_{i.}}{(i-1)} - \overline{X}_{1.}\sum_{i=2}^{m}1/(i-1)\right\rangle \qquad (5.2)$$

for the FBE method.

Equations (5.1) and (5.2) show clearly that the CBE method takes into consideration only the first and the last periodic averages, while the FBE method takes into consideration all the periodic averages. We, therefore, recommend the FBE method when it leads to adequate fit in terms of the randomness of the residuals.

REFERENCES

Box, G.E.P and Jenkins, G.M., 1976 Time Series Analysis, Forecasting and Control, 2nd ed., Holden Day, San Francisco

Buys-Ballot, C.H.D., 1847, Leo Claemert Periodiques de Temperature, Kemint et Fils, Utrecht.

Chatfield, C., 1980,. The Analysis of Time Series: An introduction, Chapman and Hall, London.

Kendal, M.G., 1976, Time Series, 2nd ed, Griffin, London.

Ljung, G.M. and Box G.E.P., 1978, On a measure of lack of fit in time series models, Biometrika 65, 297-303.

Wei, W.W., 1989, Time Series Analysis: Univariate and Multivariate Methods, Addison -- Wesley publishing Company, Inc. Redwood City.

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