Seasonal adjustment and forecasting in the presence of trend

By R. W. Hiorns*

A general linear statistical model for simultaneous seasonal adjustment and trend estimation is considered for one and two term deterministic trend functions. Explicit estimates of the parameters and variances are derived in a convenient computational form from which the properties of these estimates become apparent. In connection with possible uses of the model for forecasting, the Smallest Neighbourhood (SN) is introduced, within which the trend is assumed to be either linear or representable by a single term. Examples are given to show that simple techniques in certain situations may yield accurate forecasts in return for a comparatively modest amount of computational effort. A procedure "season" is given which calculates, for any general one term trend function, estimates of seasonal and trend constants together with standard errors and provides predictions, also with standard errors, for any required period.

1. Introduction

Simultaneous seasonal adjustment and trend estimation of an economic time series was suggested as long ago as 1937 by Arne Fisher and the technique would still seem to have advantages (see Lovell, 1963) over ratio-to-moving average approaches which gained currency in the pre-electronic computer age. This paper is not concerned with any controversy over possible methods, but assumes that the simultaneous adjustment and estimation has merits in certain situations and is therefore worthy of detailed study for some special cases. Situations to be considered are those where the time series may be represented by means of a model in which the trend, seasonal and random components may be arranged linearly (with respect to the parameters, or more briefly, say the model is *parameter linear*). The model may then be written

$$y = D\delta + S\sigma + e \tag{1}$$

where y is a vector of N observations $\{y_j\}$, e is the associated vector of errors (these will be assumed independent with mean zero and constant variance, w), D and S are matrices of trend variables and seasonal variables and δ and σ are the associated vectors of trend and seasonal constants. The matrix D contains in its jth row values of the trend variables corresponding to the jth observation whilst the matrix S, in its jth row, will contain only zeros except for a unity element in its ith column where i is the season of the jth observation, i.e.

$$i = j - rm \tag{2}$$

if m is the number of seasons, for some integer r in the range $0 \leqslant r \leqslant (N - m)/m$, r being the number of complete years in the data before that containing the jth observation.

As is well known, the estimates of the parameters in this model are linear, unbiased and have minimum variance because of the above error assumptions and the linearity of the model. The estimates may be expressed (see Jorgenson, 1964) in the form

$$\hat{\delta} = (D'D)^{-1}D'(I - S\{S'[I - D(D'D)^{-1}D']S\}^{-1}S' \\ [I - D(D'D)^{-1}D'])y,$$

$$\hat{\sigma} = \{S'[I - D(D'D)^{-1}D']S\}^{-1}S'[I - D(D'D)^{-1}D']y. \tag{3}$$

It would seem (see e.g. Gregg *et al.* (1964)) that linear or simple exponential trends might have some practical importance and these, or any other one-term trend model with a single deterministic component allow a convenient explicit form for the estimates to be given. Simple forms are also obtained in the following sections for models whose deterministic component consists of two terms. The fitting of certain polynomial and non-linear models becomes greatly simplified when these one- or two-term trend representations are applicable. Models of this type include the Gompertz, logistic and other modified exponential trend functions.

It is not suggested that in reality economic or other time series are always immediately representable by such simple forms, but occasionally, after suitable transformation of the data, these models may be of value.

2. General one-term trend

Suppose that the deterministic component contains one term only, the model then simplified to

$$y = f\delta + S\sigma + e \tag{4}$$

where f is a vector of values of the trend function and δ is the scalar trend constant. In the simplest case of linear trend f would contain values of the time variables and these could be the first N consecutive integers.

Corresponding to (3) the estimates of the parameters are now

$$\begin{pmatrix} \hat{\delta} \\ \hat{\sigma} \end{pmatrix} = \begin{pmatrix} f'f & f'S \\ S'f & S'S \end{pmatrix}^{-1} \begin{pmatrix} f'y \\ S'y \end{pmatrix}. \tag{5}$$

These may be written more explicitly after defining the following quantities:

$$F = f'f = \sum_{j=1}^{N} f_j^2 \text{ where } f = \{f_j\},$$

$$T = f'S \text{ so that } T = \{T_i\} = \left\{\sum_{r=0}^{n_i-1} f_{rm+i}\right\},$$

$$Y = S'y \text{ so that } Y = \{Y_i\} = \left\{\sum_{r=0}^{n_i-1} y_{rm+i}\right\}, \tag{6}$$

* *Electronic Computing Laboratory, University of Leeds.*

$$Y_F = f'y = \sum_{j=1}^{N} f_j y_j$$

and $$p = f'(I - S(S'S)^{-1}S')f = F - \sum_{i=1}^{m} \frac{T_i^2}{n_i}$$

where n_i is the number of observations in y relating to the season i.

The inverse matrix in the above equation may then be obtained in the form

$$\left(\begin{array}{c|c} A & B \\ \hline B' & C \end{array}\right) \tag{7}$$

where A is scalar, B is a row vector of m elements and C is a square matrix of order m. Furthermore,

$$A = 1/p,\ B = \{-T_i/(pn_i)\} \text{ and if } C = \{C_{ij}\}$$

then $$C_{ij} = \begin{cases} T_iT_j/(pn_in_j) & (i \neq j) \\ (1/n_i) + T_i^2/(pn_i^2) & (i = j). \end{cases}$$

Using this inverse or otherwise the estimates of δ and σ are

$$\hat{\delta} = \left(Y_F - \sum_{i=1}^{m} \frac{T_iY_i}{n_i}\right)/p \tag{8}$$

and $$\hat{\sigma}_i = \frac{1}{n_i}(Y_i - \hat{\delta}T_i) \tag{9}$$

for $\quad i = 1,2, \ldots, m.$

The inverse (7) provides the standard errors of these estimates

as $$\left.\begin{aligned} \text{s.e. }(\hat{\delta}) &= \sqrt{(\hat{w}/p)} \\ \text{s.e. }(\hat{\sigma}_i) &= \sqrt{\left\{\hat{w}\left(\frac{1}{n_i} + \frac{T_i^2}{pn_i^2}\right)\right\}} \end{aligned}\right\} \tag{10}$$

where $\hat{w}$ is the estimate of w the variance of the errors, obtained by squaring the residuals of the fitted model and dividing by $N - m - 1$. Alternatively, if

$$\hat{e} = y - f\hat{\delta} - S\hat{\sigma} \tag{11}$$

then $$\hat{w} = \frac{\hat{e}'\hat{e}}{N - m - 1} = \frac{y'y - Y_F\hat{\delta} - Y'\hat{\sigma}}{N - m - 1}. \tag{12}$$

Also important are the covariances between these estimates which are given by

$$\left.\begin{aligned} \text{cov }(\hat{\sigma}_i, \hat{\delta}) &= \frac{-\hat{w}T_i}{pn_i} \\ \text{cov }(\hat{\sigma}, \hat{\sigma}_j) &= \frac{\hat{w}T_iT_j}{pn_in_j} \end{aligned}\right\}. \tag{13}$$

Assuming normality for the distribution of e, the significance of the trend may be easily tested by noting that from (8) and (10) the quantity

$$\frac{\hat{\delta}}{\text{s.e. }(\hat{\delta})} = \frac{\left(Y_F - \sum_{i=1}^{m} \frac{T_iY_i}{n_i}\right)}{\sqrt{(pw)}} \tag{14}$$

follows the t-distribution with $(N-m-1)$ degrees of freedom.

3. Forecasting

Use of the estimates, (8) and (9), and the standard errors and covariances (10) and (13) can provide forecasts together with standard errors. Suppose we require to estimate the value of y, $\hat{y}_0$, at t_0 where $t_0 = rm + i$ for some integers r and i where $1 \leqslant i \leqslant m$. If we use the present model to obtain the expected value of y at t_0, this is

$$\hat{y}_0 = \hat{\sigma}_i + \hat{\delta}f(t_0). \tag{15}$$

The estimated variance of $\hat{y}_0$ is

$$\begin{aligned} \text{var }(\hat{y}_0) &= \text{var }(\hat{\sigma}_i + \hat{\delta}f(t_0)) \\ &= \text{var }(\hat{\sigma}_i) + f(t_0)^2 \text{ var }(\hat{\delta}) + 2f(t_0) \text{ cov }(\hat{\sigma}_i, \hat{\delta}) \\ &= \hat{w}\left[\frac{1}{n_i}\left(1 + \frac{T_i^2}{pn_i}\right) + f(t_0)^2/p - \frac{2f(t_0)T_i}{pn_i}\right] \\ &= \hat{w}\left[\frac{1}{n_i} + \frac{1}{p}\left\{\frac{T_i}{n_i} - f(t_0)\right\}^2\right]. \end{aligned} \tag{16}$$

In the special case of data representing complete years some properties of this estimate of variance may be noted. This estimate is a minimum when $f(t_0)$ is at or is as close as possible to its mean value. For simple linear trend, as would be expected, this variance is least at the midpoint of the range of observation. Other properties will be discussed later.

In applications where the deterministic component is parameter non-linear, or beyond the scope of a single term, there are two possible approaches to the problem of analysis or forecasting. Where possible it would seem preferable to extend the model by the addition of further terms, and this would probably be the first approach. This may involve a non-trivial estimation problem and may not produce suitable standard errors for estimates or forecasts. A second approach might then be adopted. This is to consider only a small neighbourhood of the present time in the series and to assume the deterministic component to be linear with time in this locality. There are two direct implications of this approach; one is the computational simplicity and the attendant suggestion that the analysis can be repeated more frequently than otherwise, and the other is the preclusion of anything but short-term forecasting.

The order of the above approaches is questionable. There is the usual objection to any generalization in the analysis or forecasting related to such series, that applications tend to be so special that each should be treated separately on its own merits. This is specially pertinent in the second approach which refers to a vague "small neighbourhood". How small this can be must be judged in practice by visual and other means unless, of course, the smallest possible neighbourhood is taken in all cases. To do this would maximize the chance of being right in applying a linear model in suspected non-linear situations but also it would maximize the standard errors of all the estimates concerned and therefore of any predictions when the deterministic component was in fact linear. These considerations are resumed in Section 6, when the smallest neighbourhood is treated in detail.

4. Linear trend

Here consider the model (4) with

$$f = \{f(t_j)\} = \{t_j\} = \{j\} \tag{17}$$

so that observations correspond to times $1, 2, \ldots, N$.

From (6)

$$F = \sum_{j=1}^{N} j^2 = \tfrac{1}{6}N(N+1)(2N+1) \tag{18}$$

and $$T_i = \sum_{r=0}^{n_i-1} (rm+i) = \tfrac{1}{2}mn_i(n_i-1) + in_i. \tag{19}$$

Where the data consists of complete years $n_i = n$ for all n and $N = mn$, these expressions reduce considerably so that after some algebra

$$p = \tfrac{1}{12}m^3n(n^2-1). \tag{20}$$

Referring back to (10), the variances of these unbiassed estimates will be seen to converge monotonically to zero as n tends to infinity, demonstrating in a simple manner the consistency of these estimates. Further, if $n = 1$, the estimates have infinite variance which is intuitively satisfactory for estimates based on one year's data.

There are some further comments which can now be made about the relation between the variances of the seasonal constants. From (19), T_i is increased linearly with i so that from (10) the relation between the variances of adjacent seasonal constants is

$$\operatorname{var}(\hat{\sigma}_i) > \operatorname{var}(\hat{\sigma}_{i-1}) \tag{21}$$

for all i in the range $1 < i \leqslant m$. More precisely, the difference between adjacent estimates is

$$\operatorname{var}(\hat{\sigma}_i) - \operatorname{var}(\hat{\sigma}_{i-1}) = \frac{w}{p_n^2}(mn - m - 1 + 2i). \tag{22}$$

Returning to the variance of a predicted value, in this case from (16)

$$\operatorname{var}(\hat{y}_0) = \hat{w}\left[\frac{1}{n} + \frac{1}{p}\left\{\frac{T_i}{n} - t_0\right\}^2\right]$$

where $t_0 = rm + i$. Using (20) this variance becomes

$$\operatorname{var}(\hat{y}_0) = w[(1/n) + 3(n-1-2r)^2/(mn(n^2-1))] \tag{23}$$

Table 1

Births in year ('000)

QUARTER	1958	1959	1960	1961	1962
1	191	192	198	204	221
2	190	198	199	208	216
3	180	187	198	205	207
4	178	173	188	187	196

The most striking feature of this quantity is its independence of i, the seasonal suffix. This implies that predictions for any year will have the same variance, whatever the season. Furthermore, the variance of a prediction for the rth year is less than that for one in the following year by an amount

$$\frac{wm^2}{p}(n-2-2r). \tag{24}$$

This is of course negative if $r < (\tfrac{1}{2})n - 1$.

In the example below, quarterly data over a five year period is used to estimate the four quarterly constants together with the slope of simple linear trend. The model is

$$y = t\delta + S\sigma + e \tag{25}$$

where $t' = (1,2,\ldots,20)$, δ is the scalar slope,

$$S' = \begin{pmatrix} 1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0 \\ 0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0 \\ 0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0 \\ 0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,1 \end{pmatrix}, \tag{26}$$

$\sigma' = (\sigma_1, \sigma_2, \sigma_3, \sigma_4)$,
where σ_i is the constant for the ith quarter y, is the vector of observations and e is the vector of errors.

The numbers of live births recorded in England and Wales for the years 1958–62 are given in **Table 1.** (Taken from the Registrar General's Quarterly Return of Births, Deaths and Marriages.)

In the calculations detailed in **Table 2,** for convenience this data is reduced by subtraction of 200 from each entry. The required intermediate quantities are given in the columns on the right.

Table 2

Calculation of estimates

OBSERVATIONS					(1)	(2)	(3)	(4)	(5)	(6)
'58	'59	'60	'61	'62	Y_i	T_i	$Y_i\,T_i$	δT_i	$Y_i - \delta T_i$	$\hat{\sigma}_i$
−9	−8	−2	4	21	6	45	270	72	−66	−13·2
−10	−2	−1	8	16	11	50	550	80	−69	−13·8
−20	−13	−2	5	7	−23	55	−1265	88	−111	−22·2
−22	−27	−12	−13	−4	−78	60	−4680	96	−174	−34·8
							−5125			

From the formulae (8) and (9),

$$p = \tfrac{1}{12} \, . \, 4^3 \, . \, 5 \, . \, 24 = 640$$

$$Y_F = -9.1 - 10.2 - 20.3 - 22.4 - 8.5 \ldots - 4.20 = -1$$

$$\hat{\delta} = \left(-1 + \frac{5125}{5}\right)/p = \frac{1024}{640} = 1{\cdot}60$$

so that, using (19), columns (4)–(6) in the table may be completed. The seasonal constants $\hat{\sigma}_i$ are in column (6) and the residuals e may be obtained from these constants and from $\hat{\delta}$, the slope estimate. The residuals are as follows:

2·6	−2·8	−3·2	−3·6	7·0
0·6	2·2	−3·2	−0·6	1·0
−2·6	−2·0	2·6	3·2	0·2
6·4	−5·0	3·6	−3·8	−1·2

From the form of the model the check that the rows sum to zero is used to verify the computation of the estimates and residuals. The variance of the residuals is estimated by their sum of squares divided by $20 - 4 - 1 = 15$ and this gives

$$\hat{w} = \frac{226{\cdot}20}{15} = 15{\cdot}08.$$

The calculation of the standard errors of the estimates is now straightforward using (8) and (9) leading to

$$\hat{\delta} = 1{\cdot}60 \pm 0{\cdot}15 \qquad \begin{aligned} \hat{\sigma}_1 &= -13{\cdot}20 \pm 2{\cdot}22 \\ \hat{\sigma}_2 &= -13{\cdot}80 \pm 2{\cdot}32 \\ \hat{\sigma}_3 &= -22{\cdot}20 \pm 2{\cdot}42 \\ \hat{\sigma}_4 &= -34{\cdot}80 \pm 2{\cdot}53. \end{aligned}$$

Predicted values for the four years immediately before and after the period covered by the above data are given in **Table 3,** together with standard errors. The latter are obtained from formula (15), and being the same for any year, are given at the foot of each column.

Table 3 also gives, where available, the errors or deviations of the observed values from the predicted values above. From the pattern and magnitude of these errors, the non-linearity over the wider range is obvious. The backward projection to 1954 well enough illustrates the danger of extending the range of prediction beyond that of observation.

5. Two-term trend

Polynomial trends have been commonly suggested and, in general, explicit estimates cannot be obtained for these. However, there are still some important special estimates which it may well be worthwhile to give directly. Consider two-term components so that in (1),

$$D\delta = f_1\delta_1 + f_2\delta_2 \tag{27}$$

where f_1 and f_2 represent any simple (non-parametric) functions of time.

Extending the previous notation naturally, the estimates of the present constants, corresponding to (5), are:

$$\begin{bmatrix} \hat{\delta}_1 \\ \hat{\delta}_2 \\ \hat{\sigma}_1 \\ \hat{\sigma}_2 \\ \vdots \\ \hat{\sigma}_m \end{bmatrix} = \begin{bmatrix} F_1 & F_{12} & T_{11} & T_{12} & \cdots & T_{1m} \\ F_{12} & F_2 & T_{21} & T_{22} & \cdots & T_{2m} \\ T_{11} & T_{21} & n_1 & 0 & \cdots & 0 \\ T_{12} & T_{22} & 0 & n_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ T_{1m} & T_{2m} & 0 & 0 & \cdots & n_m \end{bmatrix} \begin{bmatrix} Y_{F1} \\ Y_{F2} \\ Y_1 \\ \vdots \\ \vdots \\ Y_m \end{bmatrix}$$

where

$$\begin{gathered} F_1 = \sum_{j=1}^{N} f_{1j}^2, \quad F_2 = \sum_{j=1}^{N} f_{2j}^2, \\ F_{12} = \sum_{j=1}^{N} f_{1j} f_{2j} \\ T_{1i} = \sum_{r=0}^{n_i-1} f_{1(rm+i)} \quad T_{2i} = \sum_{r=0}^{n_i-1} f_{2(rm+i)} \\ Y_{F1} = \sum_{j=1}^{N} y_j f_{1j}, \quad Y_{F2} = \sum_{j=1}^{N} y_j f_{2j} \end{gathered} \tag{28}$$

and, as before, $\quad Y_i = \sum_{r=0}^{n_i-1} y_{rm+i}.$

Table 3

Predicted values and their errors

QR.	'54	'55	'56	'57	'63	'64	'65	'66	'70
1	162·8	169·2	175·6	182·0	220·4	226·8	233·2	239·6	265·2
2	163·8	170·2	176·6	183·0	215·0	221·4	227·8	234·2	259·8
3	157·0	163·4	169·8	176·2	213·6	220·0	226·4	232·8	258·4
4	146·0	152·4	158·8	165·2	203·6	210·0	216·4	222·8	248·4
standard error	4·494	3·798	3·135	2·532	2·532	3·135	3·798	4·494	7·420

Deviations: observed − predicted.

1	11·2	−0·2	3·4	2·0	1·6	−7·8	−16·6
2	11·2	0·8	5·4	2·0	5·0	6·6	−14·6
3	10·0	3·6	4·2	3·8	1·4	2·0	−13·8
4	11·0	5·6	4·2	8·8	−3·6	−6·0	−19·0

The direct estimates are readily obtained after stages of algebraic elimination and back substitution; these are conveniently set out below.

(i) *Elimination*

$$\text{Let } F_1' = F_1 - \sum_{i=1}^{m} \frac{T_{1i}^2}{n_i},$$

$$F_{12}' = F_{12} - \sum_{i=1}^{m} \frac{T_{1i}T_{2i}}{n_i}, \quad Y_{F1}' = Y_{F1} - \sum_{i=1}^{m} \frac{T_{1i}Y_i}{n_i}, \tag{29}$$

$$F_2' = F_2 - \sum_{i=1}^{m} \frac{T_{2i}^2}{n_i} - \frac{F_{12}'^2}{F_1'}$$

$$\text{and} \quad Y_{F2}' = Y_{F2} - \sum \frac{T_{2i}Y_i}{n_i} - \frac{F_{12}'Y_{F1}'}{F_1'}.$$

(ii) *Back substitution*

Directly $\qquad \delta_2 = Y_{F2}'/F_2'.$

$$\text{If} \quad Y_{F1}'' = (Y_{F1}' - \delta_2 F_{12}') \quad \text{then } \delta_1 = Y_{F1}''/F_1' \tag{30}$$

$$\text{and} \quad \hat{\sigma}_i = \frac{1}{n_i}(Y_i - \delta_1 T_{1i} - \delta_2 T_{2i}).$$

The estimates given here allow a straightforward extension to the problem of fitting models with non-linear deterministic components. In particular, the modified exponential model may be mentioned. Stevens (1951) showed that the iterative process for three-parameter asymptotic regression models with a single unknown asymptote reduces to iteration in only the non-linear parameter. This process is described in Hiorns (1965). The use of this type of model here results in an iterative process with the ratio of the δs in (30) above as the correction to the non-linear parameter at each stage. Clearly other interesting models can be dealt with in this straightforward manner.

6. The smallest neighbourhood (SN)

In order to leave at least one degree of freedom for an estimate of the variance, the total number of points N must be at least $m + 2$ for a one-term deterministic component and an m-term seasonal component.

The smallest neighbourhood (SN) is then that enclosing just $N = m + 2$ observations. Suppose that the trend function values are $f_1, f_2, \ldots, f_{m+2}$. In this case, $n_1 = n_2 = 2, n_3 = n_4 = \ldots = n_m = 1$. Also

$$T_i = \begin{cases} f_i + f_{m+i} & (i = 1,2) \\ f_i & (3 \leq i \leq m). \end{cases}$$

Define, for convenience, an operator [ ] as

$$[z_i] = [z_{m+i} - z_i].$$

Now $p = (\frac{1}{2})([f_1]^2 + [f_2]^2)$ and, using the previous relations, the estimates of the parameters become

$$\delta = \frac{[y_1][f_1] + [y_2][f_2]}{[f_1]^2 + [f_2]^2} \tag{31}$$

$$\hat{\sigma}_i = \begin{cases} (\frac{1}{2})\{(y_i + y_{m+i}) - \delta(f_i + f_{m+i}) & (i = 1,2) \\ y_3 - \delta f_3 & (3 \leq i \leq m) \end{cases}$$

$$\hat{w} = (\tfrac{1}{2})\left\{[y_1]^2 + [y_2]^2 - \frac{([y_1][f_1] + [y_2][f_2])^2}{[f_1]^2 + [f_2]^2}\right\}.$$

The estimate $\hat{y}_{m+3}$ of the month following the SN may now be obtained by direct substitution into equation (15) and the variance of this estimate now simplifies to

$$\text{var}(\hat{y}_{m+3}) = \hat{w}\left(1 + \frac{2[f_3]^2}{[f_1]^2 + [f_2]^2}\right).$$

An even simpler model may be obtained by taking the trend function f to be linear with time. Suppose $f_i = i$, the estimates then are

$$\delta = ([y_1] + [y_2])/2m$$

$$\hat{\sigma}_i = \begin{cases} (\frac{1}{2})\{(y_i + y_{m+i}) - \delta(m + 2i)\} & (i = 1,2,) \\ y_i - \delta_i & (3 \leq i \leq m) \end{cases}$$

$$\hat{w} = (\tfrac{1}{4})([y_1] - [y_2])^2. \tag{32}$$

The estimate for the month following the SN is now simplified greatly to

$$\hat{y}_{m+3} = y_3 + \tfrac{1}{2}([y_1] + [y_2]) \tag{33}$$

and its variance is var $(\hat{y}_{m+3}) = 2\hat{w}$ and its standard error takes the absolute value of $([y_1] - [y_2])/\sqrt{2}$. Confidence limits for the estimate $\hat{y}_{m+3}$ may then be seen to have the simple form

$$\hat{y}_{m+3} + k_1[y_1] + k_2[y_2], \hat{y}_{m+3} + k_2[y_1] + k_1[y_2]$$

where k_1 and k_2 are the constants $\frac{1}{2} \pm \frac{1}{2}t_\alpha$, if t_α is the appropriate t-value with one degree of freedom, so that $k_1 + k_2 = 1$. Some values for these constants with the probability α of observations occurring in the interval are given in **Table 4**.

From these values, the probability is 0·39 that the true value lies between $y_3 + [y_1]$ and $y_3 + [y_2]$. Similarly, the probability is 0·72 that the true value lies between $y_3 + 2[y_1] - [y_2]$ and $y_3 + [y_1] - 2[y_2]$.

The test for linearity given in equation (14) for this simple case reduces to comparing the ratio

$$([y_1] + [y_2])/([y_1] - [y_2])$$

with the values 6·31, 12·7 and 63·7 for t at the 10%, 5% and 1% significance levels.

Table 4

Confidence interval multipliers

k_1	k_2	α
1	0	0·39
2	−1	0·72
3	—2	0·82
4	−3	0·87
8·8	−7·8	0·95

Using only four values to estimate $\hat{y}_{m+3}$, the whole procedure is rough and obviously requires much care in application. It is probably fair to suggest that slowly changing, parameter nonlinear trend functions do occur which are exceedingly tedious to handle by least squares. Some of these would provide sufficient *local* parameter linearity to make this "smallest neighbourhood" technique worthy of serious consideration for short-term forecasting.

An example is now given showing how the SN technique, although computationally simple, might well be effective as a forecasting aid in some situations. The data used for this illustration is the airline passenger data quoted by Brown (1959) and later used by Barnard (1963). The results of applying two other forecasting methods, adaptive forecasting and the Box–Jenkins method, are compared in **Table 5** with the SN technique. In each case, forecasts were prepared each month, using only previously observed data, for the following month over the nine-year period, and the sum of absolute errors each year is given in the table for each method.

7. Conclusion

The general theory for simultaneous seasonal adjustment and trend estimation has been investigated in the special cases of models whose deterministic component consists of one or two terms. Convenient forms for the estimates of seasonal and trend constants are obtained with their variances, and properties of these quantities are exhibited. The simplicity of these forms is an attractive feature and this suggests their continuous use in a moving fashion to follow short-term trend.

A single deterministic term would not usually be adequate for more than a few years and the considerations discussed above led to a "smallest neighbourhood". This allowed just one degree of freedom for the estimation of variances, a procedure which in general will be harsh, but on occasion necessary. However, the real situation may afford some compromise if linearity is maintained for two or more years for then a "next smallest neighbourhood" (NSN) may be considered. The logical period here would be two complete years, so that two observations are used for each seasonal constant and a more satisfactory estimate of variance would then become available.

Besides these two convenient neighbourhoods, other alternatives might be considered. A possible situation is that in which linearity is maintained for part of a time series to be followed by a change to linearity according to some new regime. Overall this situation would be called non-linear, but the moving SN or NSN techniques should work well when the point of change of regime lies outside the moving neighbourhood. With this in mind, it would seem that the SN would be optimal. On closer investigation, however, the length of the neighbourhood would be better related to the distance between points of change of regime. For any regularity in the spacings of these points (e.g. the natural time between national elections) it would be necessary to make the length of the neighbourhood considered a parameter to be estimated on a given part of the time series.

Table 5

Comparison for forecasts for airline passenger data by sums of absolute errors in twelve monthly forecasts

	YEARS								
METHOD	'52	'53	'54	'55	'56	'57	'58	'59	'60
Adaptive forecasting	61	149	110	77	54	96	210	140	156
Box–Jenkins	68	142	98	109	82	85	134	138	205
SN	109	126	100	110	78	105	170	142	215

In general, in any situation it seems that a neighbourhood of optimal length might be selected to advantage. Without some measure of non-linearity, however, it would be difficult to cater for situations in advance by a general rule regarding the choice of a neighbourhood. In a given situation it is of course a trivial matter to compare the forecasting performance of some alternative techniques. This would be a natural and obvious procedure for investigating the non-linearity of a given series and the relation between this and the length of an optimal neighbourhood.

This work suggests that other neighbourhoods might be studied and clearly a "best neighbourhood" (BN) according to some criterion for a given time series, could be based upon early observations, and this adopted for all later forecasting. The choice of the criterion is not clear.

```
procedure season (y,m,f,n,N,a,e,w,P1, P2, pred, pe);
  comment This procedure calculates, simultaneously,
  estimates of seasonal and trend constants by multiple
  linear regression and provides forecasts and errors for
  a specified period. The model used has the demand
  variable represented by a trend component, seasonal
  component and random component combined additively.
  The trend term is assumed to consist of a single function
  whose values are supplied to the procedure in an array
  f[1 : N] where N is the number of observed values of
  the demand variable, also supplied, in an array y[1 : N].
  These values correspond to consecutive time periods,
  there being m seasons in a year, represented by m
  seasonal constants, but N need not be a multiple of m.
  N must satisfy N ≥ m + 2. The number of observed
  values for each season must be supplied in the array
  n[1 : m].
    Estimates are left by the procedure as follows: the
  trend constant in a[0] and the m seasonal constants in
  the remainder of the array a[0 : m]. Standard errors
  for the constants are in e[0 : m]. The residual variance
  estimate is in w.
    Forecasts (or predictions) are made for consecutive
  time periods from P1 to P2. These are left in pred
  [P1 : P2] and their standard errors in pe[P1 : P2];
```

```
array y,f,a,e,pred,pe; integer array n;
integer m,N,P1,P2; real w;

begin integer i,j,k; real F,YF,p,q;
      array T,Y[1 : m];
      k:=0; YF:=F:=p:=q:=0;
      for i:=1 step 1 until m do
      begin
            T[i]:=Y[i]:=0;
            for j:= 1 step 1 until n[i] do
            begin
                  k:=k+1; F:=F+f[k] ↑ 2;
                  T[i]:=T[i]+f[k];
                  Y[i]:=Y[i]+y[k];
                  YF:=YF+y[k]×f[k]
            end;

            p:=p−T[i] ↑ 2/n[i];
            q:=q−T[i]×Y[i]/n[i];
            e[i]:=(1+T[i] ↑ 2/n[i])/n[i]
      end;

estimates:    a[0]:=q/p;
              for i:=1 step 1 until m do
                a[i]:=(Y[i]−a[0]×T[i])/n[i];

sumsquares:   k:=0;
              for i:=1 step 1 until m do
              for j:= 1 step 1 until n[i] do
              begin k:=k+1;
                  w:=w+(y[k]−a[i]−a[0]×f[k]) ↑ 2
              end;
              w:=w/(N−m−1);
              e[0]:=sqrt(w/p);
              for i:=1 step 1 until m do
                                e[i]:=sqrt(e[i]×w);
predictions:  for k:= P1 step 1 until P2 do
              begin i:=k−m×(k−1)÷m;
                    pred[k]:=a[i]+a[0]×f[k];
                    pe[k]:=w×(1/n[i]+(T[i]/n[i]
                       −f[i]) ↑ 2/p);
                    pe[k]:=sqrt(pe[k])
              end

end of season;
```

References

BARNARD, G. A. (1963). "New Methods of Quantity Control", *Journal of Royal Statistical Society Series A*, Vol. 126, pp. 255–8.

BOX, G. E. P., and JENKINS, G. M. (1962). "Some statistical aspects of adaptive optimization and control", *Journal of Royal Statistical Society Series B*, Vol. 24, pp. 297–344.

BROWN, R. G. (1959). *Statistical Forecasting for Inventory Control*, New York: McGraw-Hill.

FISHER, ARNE (1937). "A brief note on seasonal variation", *Journal of Accountancy*, Vol. 64, pp. 174–99.

GREGG, J. V., HOSSELL, C. H., and RICHARDSON, J. T. (1964). *Mathematical Trend Curves: an aid to forecasting*, I.C.I. Monograph No. 1. London: Oliver and Boyd.

HIORNS, R. W. (1965). *The fitting of growth and allied curves of the asymptotic regression type by Stevens's method*, Cambridge University Press: Tracts for Computers Series No. XXVIII.

JORGENSON, D. W. (1964). "Minimum Variance, Linear, unbiased seasonal Adjustment of Economic Time Series", *Journal of the American Statistical Association*, Vol. 59, pp. 681–724.

LOVELL, M. C. (1963). "Seasonal adjustment of economic time series and multiple regression analysis", *Journal of the American Statistical Association*, Vol. 58, pp. 993–1010.

STEVENS, W. L. (1951). "Asymptotic regression", *Biometrics*, Vol. 7, pp. 247–67.

Book Review (*Continued from p. 134*)

and organizational, are functioning properly, and this chapter provides even newcomers to EDP with sufficient information to carry out a detailed audit.

Reference is made to the verification of balance sheet and profit and loss account items, and the suggestion is made that here may be a field for the auditor to use special computer programs, e.g. random sampling instead of a complete print-out of a file: print-out reports of items which fail to meet specified criteria. There are also two examples in some detail of the use of special computer programs, one on a large payroll application and the other in connection with the verification of the valuation of a company's hire purchase debtors. On the question as to whether an auditor has a responsibility to examine computer programs it is clear that the author has made up his mind that such a task is neither a practical nor desirable approach, and he suggests the use of test packs as a more practical method for an auditor to satisfy himself about the validity and reliability of a client's computer programs.

The preparation, use, and limitation of test packs are referred to in some detail, followed by an example of their use in relation to a sales system. Only valid data is used and the processing is carried out by a duplicate copy of the client's program. There is, however, no advice as to how to check that the duplicate used is a true copy. The "pros" and "cons" of these techniques are carefully weighed and the reader is not discouraged from seeking further development. Mr Pinkney concludes with these words, "The auditing of computer applications is still at a comparatively early stage and it has become clear, in the course of preparing this book, that a great deal of further work remains to be done. It is hoped that the suggestions contained in the preceding pages will assist these further developments."

It is thought that this book will currently be of great assistance both as a reference and a guide and will stimulate future thought and work.

F. E. HINCHCLIFF