Qüestiió - V. 7, n.º 4 (desembre 1983) pp. 563-575

THE FORMULATION OF STRUCTURAL TIME SERIES MODELS IN DISCRETE AND CONTINUOUS TIME

A. C. HARVEY
LONDON SCHOOL OF ECONOMICS

This paper sets out an approach to modelling univariate time series, including those in which observations are available on a daily basis. An underlying continuous time model is formulated and it is shown that this model has important implications for the way in which a discrete model is set up. It is also shown that the continuous time model allows observations subject to temporal aggregation and irregularly spaced observations to be handled relatively easily. The extension to cases where explanatory variables are to be included in the model is also discussed.

Keywords: CONTINUOUS TIME MODEL; DAILY OBSERVATIONS; FORECASTING; IRREGULAR OBSERVATIONS; KALMAN FILTER; STRUCTURAL TIME SERIES MODEL.

1. INTRODUCTION.

Structural time series models differ from traditional ARIMA models in that the various components of a time series, for example the trend and seasonal effects, are modelled explicitly; see Engle /6/ , Harvey and Todd /12/ and Kitagawa /20/ . Harvey and Todd /12/ , in particular, argue that the structural approach has a certain methodological advantages over the Box-Jenkins - approach in that it enables the user to be aware of exactly what he is doing. This is particularly important when dealing with economic time series which contain observations for a relatively small number of years. The argument in Harvey and Todd /12/ , was made with respect to monthly and quarterly observations, but when daily observations are available the case in favour of a structural approach is even stronger. This is because the Box-Jenkins model selection procedure tends to become unmaneagable.

This article first sets out a model for handling daily data. It is then shown how this discrete time model can be derived from a model formulated in continuous time. This is important for two reasons, one methodological, the other technical. The methodological - reason is that any model formulated in discrete time suffers from the disadvantage that casting the dynamics in terms of the time period of the observations is arbitrary; see Bergstom / 2 / . This point becomes apparent with ARIMA models when it is realized that the order of an ARIMA model will, in general, change if observations are obtained at more (or less) frequent intervals or if there is temporal aggregation; see, for example, - Amemiya and Wu (1972). The technical reason for considering the continuous time representa-- tions is that it enables one to derive efficient ways of handling irregularly spaced observations and observations subject to - temporal aggregation.

2. MODELLING DAILY DATA.

The models employed in Harvey and Todd /12/ and Kitagawa /20/ , are for monthly or quarterly data. They are formulated in terms of a trend, a seasonal and an irregular component, although cycles can be incorporated into the models if desired. The components are usually set up in such a way that the forecast function consists of a linear trend with a fixed seasonal pattern imposed on it. The models are stochastic in that the trend and

- A. C. Harvey - Dep. of Statistics - London School of Economics - University of London - Houghton Street London WC2A 2AE (England).

- Article rebut el Novembre del 1983.

- Aquest treball es una ponencia invitada al "First Catalan International Symposium on Statistics" a Barcelona del 23 al 24 de Setembre de 1983.

seasonal components are allowed to change slowly over time. From the practical point of view, this means that past observations are discounted in making forecasts. Hence the models are local, rather than global. Globals models do, however, arise as a special case and Franzini and Harvey /7/, show how it is possible to test for global models in small samples. The general model can be written as follows

$$y_t = \mu_t + \psi_t + \gamma_t + \theta_t + \theta_t^* + \varepsilon_t, \qquad (2.1)$$

where μ_t, ψ_t, γ_t, θ_t and ε_t are the trend, cyclical, seasonal, daily and irregular components respectively. The term θ_t^* represents calender effects, such as public holidays. In many applications y_t will be in logarithms.

The specification of the various components is as follows.

Trend - the level of the process, μ_t, and the slope, β_t, are generated by the multivariate random walk process

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \qquad (2.2a)$$

$$\beta_t = \beta_{t-1} + \zeta_t \qquad (2.2b)$$

where $(\eta_t \ \zeta_t)'$ is a multivariate white noise process with mean zero and covariance matrix, Q^η. Some restrictions must be placed on Q^η in order to ensure identifiability; cf. the discussion of identifiability of unobserved components models in Hotta /16/. One possibility is to let it be diagonal. Another possibility suggested by the continuous time formulation, is to set

$$Q^\eta = \begin{bmatrix} \bar\sigma_\eta^2 + \frac{1}{3}\bar\sigma_\zeta^2 & \frac{1}{2}\bar\sigma_\zeta^2 \\ \frac{1}{2}\bar\sigma_\zeta^2 & \bar\sigma_\zeta^2 \end{bmatrix} \qquad (2.2c)$$

The interpretation of the two parameters, $\bar\sigma_\eta^2$ and $\bar\sigma_\zeta^2$, will be clear from the discussion in section 4. The above model can be extended so that it becomes a local approximation to a higher polynomial time trend; see Harrison and Stevens /10/. For most purposes (2.2) will suffice.

Cycle - A trade cycle centred around a frequency of λ_c can be represented by ψ_t where

$$\psi_t = \psi_{t-1}\cos\lambda_c + \psi_{t-1}^*\sin\lambda_c + \kappa_t \qquad (2.3a)$$

$$\psi_t^* = -\psi_{t-1}\sin\lambda_c + \psi_{t-1}^*\cos\lambda_c + \kappa_t^* \qquad (2.3b)$$

where κ_t and κ_t^* are white noise disturbances with covariance matrix Q^κ. The continuous time formulation of Section 4 suggests that this covariance matrix be scalar, i.e. $Q^\kappa = \sigma_\kappa^2 I_2$. The restrictions which this implies are more than enough to ensure identifiability.

The cycle defined by (2.3) is nonstationary and its forecast function will have a cyclical form which persists indefinitely into the future. A stationary cyclical component can be constructed by introducing an additional parameter, ρ, into (2.3) so that it becomes

$$\begin{bmatrix} \psi_t \\ \psi_t^* \end{bmatrix} = \rho \begin{bmatrix} \cos\lambda_c & \sin\lambda_c \\ -\sin\lambda_c & \cos\lambda_c \end{bmatrix} \begin{bmatrix} \psi_{t-1} \\ \psi_{t-1}^* \end{bmatrix} + \begin{bmatrix} \kappa_t \\ \kappa_t^* \end{bmatrix} \qquad (2.4a)$$

or, in matrix terms,

$$\psi_t = C\,\psi_{t-1} + \kappa_t, \qquad (2.4b)$$

where $0 \le \rho < 1$; cf. Harrison and Akram /9/. This process can be expressed as

$$\psi_t = \frac{(1 - \rho\cos\lambda_c L)\kappa_t + (\rho\sin\lambda_c L)\kappa_t^*}{1 - 2\rho\cos\lambda_c L + \rho^2 L^2} \qquad (2.5)$$

where L is the lag operator. Thus ψ_t is a special case of an ARMA (2,1) process. The attraction of (2.4) is that its parameters are directly related to the parameters of the (pseudo) cycle.

Seasonal - For monthly or quarterly data, seasonality can be modelled by a set of dummy variables or by a set of trigonometric terms. In Harvey and Todd /12/ and Kitagawa /20/, a dummy variable formulation is employed. For daily data a trigonometric representation is preferable for several reasons, the main one being to ensure continuity from the last day of one month to the first day of the next.

A seasonal model is set up by defining cycles at frequencies $\lambda_j = 2\pi j/M$, $j=1,2,\dots,h$, where M is the number of days in the year. (If observations are available every day in the year, then M = 365 or 366). The model for each frequency is similar to the trade cycle model, i.e.

$$\begin{bmatrix} \gamma_{jt} \\ \gamma^*_{jt} \end{bmatrix} = \begin{bmatrix} \cos\lambda_j & \sin\lambda_j \\ -\sin\lambda_j & \cos\lambda_j \end{bmatrix} \begin{bmatrix} \gamma_{j,t-1} \\ \gamma^*_{j,t-1} \end{bmatrix} + \begin{bmatrix} \omega_{it} \\ \omega^*_{jt} \end{bmatrix} \quad j=1,2,\ldots \qquad (2.6a)$$

or in matrix terms.

$$\gamma_{jt} = S_j\, \gamma_{j,t-1} + \omega_{jt} \quad , \; j = 1,2,\ldots \qquad (2.7b)$$

The seasonal effect is given by

$$\gamma_t = \gamma_{1t} + \gamma_{2t} + \gamma_{3t} + \ldots \qquad (2.8)$$

Again the continuous time formulation suggests letting the covariances of each ω_{jt} depend on only a single parameter, i.e., $\text{Var}(\omega_{jt}) = \sigma^2_j I_2$.

Daily - Let w be the number of different types of day in a week and let k_j be the number of days of the j-th type for $j = 1,\ldots,w$. Thus, for example, if all week days are alike but both Saturdays and Sundays are different, $w = 3$, $k_1 = 5$, and $k_2 = k_3 = 1$. The effect associated with the j-th type of day is θ_{jt}, where

$$\theta_{jt} = \theta_{j,t-1} + \xi_{jt}, \quad j = 1,\ldots,w-1 \;, \qquad (2.9)$$

the disturbance term ξ_{jt} having zero mean and variance

$$\text{Var}(\xi_{jt}) = \sigma^2_\xi(1-k^2_j/K), \; j = 1,\ldots,w-1,$$

$$K = \sum_{j=1}^{w} k^2_j \qquad (2\text{-}10)$$

the covariances between the disturbances are

$$E(\xi_{jt}\, \xi_{ht}) = -\,\sigma^2_\xi k_j k_h/K, \qquad j \neq h, \quad j,h=1,\ldots,w-1. \qquad (2.11)$$

The effect for the w-th type of day is defined by the requirement that the sum of the daily effects over a week should be zero, i.e.,

$$\theta_{wt} = -\,k_w^{-1} \sum_{j=1}^{w-1} k_j\, \theta_{jt} \qquad (2.12)$$

It can be verified directly that θ_{wt} has exactly the same properties as the other daily effects, i.e. it follows a random walk of the form (2.9) with a variance (2.10) and covariances (2.11), see Appendix. Thus it is immaterial which effect is taken to be the w-th.

The above model can be written in matrix form by defining the $(w-1)\times 1$ vector $k = (k_1,\ldots,k_{w-1})'$. Then if ξ_t is the $(w-1)\times 1$ vector of disturbances in

$$\theta_t = \theta_{t-1} + \xi_t, \qquad (2.13)$$

we have

$$E(\xi_t\, \xi_t') = \sigma^2_\xi(I-K^{-1}kk')$$

In the special case when $k_j = 1$ for $j = 1,\ldots,7$, the model takes essentially the same form as the one used by Harrison and Stevens /10/ to model seasonal effects.

Calender Effects - Calender effects arise on certain days throughout the year due to public holidays, religious festivals and so on. These effects can be modelled in the same way as daily effects, with w replaced by w^* which denotes the number of different types of special days throughout the year. As with daily effects a constraint analogous to (2.12) ensures that the sum of the calender effects throughout the past 365 or 366 days is equal to zero.

Calender effects should be distinguished from unique events such as royal weddings and earthquakes. Within the sample period events of this kind can be treated as missing observations or modelled explicitly by the introduction of dummy variables. These two approaches are equivalent unless the effect lasts several days and constraints are put on the coefficients of the dummy variables.

Irregular Component - The specification of (2.1) is completed by the requiriment that the irregular component, ε_t, be white noise. An extension of the model could be obtained by replacing ε_t by a stationary AR or ARMA process, but this possibility will not be pursued here.

3. STATISTICAL TREATMENT OF THE DAILY MODEL.

Given that the observations are normally distributed, unknown parameters can be estimated and tests of specification carried out by working with a likelihood function computed via the prediction error decomposition. The prediction errors are computed by setting up the model in state space form and applying the Kalman filter.

State Space Formulation

The structural model defined in section 2 can be handled statistically by writting it in state space form. The transition equation consists of (2.2), (2.4), (2,7), (2.13) and an equation for the calender effects, Θ_t, -- having the same form as (2.13). Thus

$$\begin{bmatrix}\mu_t\\ \beta_t\\ \psi_t\\ \gamma_{1t}\\ \gamma_{2t}\\ \vdots\\ \Theta_t\\ \Theta_t^*\end{bmatrix} = \begin{bmatrix} \begin{matrix}1&1\\0&1\end{matrix} & & & & & & \\ & C & & & & & \\ & & S_1 & & & & \\ & & & S_2 & & & \\ & & & & \ddots & & \\ & & & & & I & \\ & & & & & & I\end{bmatrix} \begin{bmatrix}\mu_{t-1}\\ \beta_{t-1}\\ \psi_{t-1}\\ \gamma_{1,t-1}\\ \gamma_{2,t-1}\\ \vdots\\ \Theta_{t-1}\\ \Theta_{t-1}^*\end{bmatrix} + \begin{bmatrix}\eta_t\\ \zeta_t\\ \kappa_t\\ \omega_{1t}\\ \omega_{2t}\\ \vdots\\ \xi_t\\ \xi_t^*\end{bmatrix} \quad (3.1a)$$

or, more compactly,

$$\alpha_t = T_t \alpha_{t-1} + \eta_t \quad (3.1b)$$

Both the transition matrix T_t, and the covariance matrix of η_t which will be denoted by Q_t, are block diagonal. The associated measurement equation is

$$y_t = z_t' \alpha_t + \varepsilon_t = z_1' \alpha_{1t} + z_{2t}' \alpha_{2t} + \varepsilon_t \quad (3.2)$$

where the state vector has been partitioned so that $\alpha_{1t} = (\mu_t, \beta_t, \psi_t', \gamma_{1t}', \gamma_{2t}', \ldots)'$ and $\alpha_{2t} = (\Theta_t' \; \Theta_t{}')'$. The vector z_1 is time invariant, consisting of alternate ones and zeroes, i.e., $z_1 = (1\ 0\ 1\ 0\ 1\ \ldots)'$. The $(w + w^* - 2) \times 1$ vector, z_t, changes over time so that each day it picks out the appropriate daily effect from Θ_t and the appropriate calender effect from Θ_t. Thus the first w-1 elements are such that for a day of type j, there is a one in the j-th position and zeroes elsewhere. The exception is for the w-th type of day in which case (2.12) leads to the requirement that the j-th element be equal to $-k_j/k_w$ for $j = 1,\ldots,w-1$. Calender effects are treated in a similar way.

Kalman Filter

Let a_{t-1} denote the minimum mean square -- error estimator (MMSE) of α_{t-1} based on all the information up to and including time t, and let P_{t-1} denote the covariance matrix of its estimation error, $a_{t-1} - \alpha_{t-1}$. Similarly let $a_{t/t-1}$ denote the MMSE of α_t based on all the information up to time t-1 and let $P_{t/t-1}$ be the covariance matrix of its estimation error. The Kalman filter consists of the prediction equations

$$a_{t/t-1} = T_t a_{t-1} \quad (3.3a)$$

and

$$P_{t/t-1} = T_t P_{t-1} T_t' + Q_t, \quad (3.3b)$$

together with the updating equations

$$a_t = a_{t/t-1} + P_{t/t-1} z_t \nu_t / f_t \quad (3.4a)$$

and

$$P_t = P_{t/t-1} - P_{t/t-1} z_t f_t^{-1} z_t' P_{t/t-1} \quad (3.4b)$$

where $\nu_t = y_t - z_t' a_{t/t-1}$ is the one-step ahead prediction error and

$$f_t = z_t' P_{t/t-1} z_t + h_t \quad (3.4c)$$

The term h_t is the variance of ε_t. Further details can be found in, for example, Anderson and Moore /1/ and Harvey /11/.

Maximum Likelihood Estimation

The state vector, α_t, contains n elements all of which are nonstationary. If the first n observations are regarded as fixed, starting values for the Kalman filter can be constructed from these observations and the likelihood function formed from the subsequent T-n one-step ahead prediction errors, i. e.

$$\log L = -\frac{(T-n)}{2}\log 2\pi - \frac{1}{2}\sum_{t=n+1}^{T}\log f_t - \frac{1}{2}\sum_{t=n+1}^{T} \nu_t^2/f_t \tag{3.5}$$

where ν_t is the one-step ahead prediction error at time t and f_t is its variance,(3.4c).

These quantities are obtained directly from the Kalman filter. Note that instead of explicitly calculating starting values from the first n observations, the recursions can be started off at t = 0 with an arbitrary a_o and a covariance matrix of estimation errors equal to κI, where κ is a large but finite number [1]. In any case the likelihood must be maximized numerically with respect to the hyperparameters in the model, i.e., ρ, λ and all the variances of the disturbances in the vector η_t. (Note that one of the variances can always be concentrated out of the likelihood function).

A slightly different likelihood function is obtained if the initial state vector, α_o, is assumed to be fixed and the first n observations are random. The Kalman filter is again needed to compute the likelihood function,but the starting value problem is solved by a modification of an algorithm due to Rosenberg /24/ or by the method given in Wecker and Ansley /25/ ; see Harvey and Peters /15/, for further discussion.

Missing observations can be handled quite easily within a state space framework; see Jones /17/ , Harvey and Pierse /13/ and Kohn and Ansley /21/ . All that needs to be done is to omit the updating equations corresponding to a missing observation. Temporal aggregation is a related problem in which missing values are incorporated into the next subsequent observation. This situation arises often in economics with flow variables, i.e. variables which can only be measured with respect to a particular period of time. One way of handling the problem is to augment the state vector with past values of y_t; cf. Harvey and Pierse /13/ . However, in some cases, a more elegant solution can be derived from the continuous time formulation; see section 5.

Predictions of Future Observations

Once any unknown parameters in the model have been estimated, ℓ-step ahead predictions together with this conditional MSEs, can be made by repeated application of the Kalman filter prediction equations. The forecast function for (2.1) is

$$\begin{aligned}\tilde{y}_{T+\ell/T} &= \\ &= m_T + b_T\ell + \rho^{\ell}(\tilde{\psi}_T \cos\lambda_c\ell + \tilde{\psi}^*_T \sin\lambda_c\ell) \\ &\quad + \sum_j (c_{jT}\cos\lambda_j\ell + c^*_{jT}\sin\lambda_j\ell). \\ &\quad + \{\text{Daily and calender effects at } T+\ell\}, \\ &\quad \ell = 1,2,\ldots\end{aligned} \tag{3.6}$$

where m_T, b_T, $\tilde{\psi}_T$, $\tilde{\psi}^*_T$, c^*_{jT} and c_{jT} are estimates of μ_T, β_T etc. computed from the Kalman filter. These estimates are conditional on estimates of the unknown hyperparameters, ρ, λ_c, σ^2_η and so on in the model.

The forecast function in (3.6) is obtained very easily by observing that if a_T is the estimator of α_T from the Kalman filter at time T, then

$$\tilde{y}_{T+\ell/T} = z'_{T+\ell}\, \underset{\sim}{T}^{\ell}\, a_T, \quad \ell = 1,2,\ldots \tag{3.7}$$

since the transition matrix, $\underset{\sim}{T}$, is time invariant.

Estimates of the MSEs of these predictions can also be computed by repeated application of the covariance updating equation (3.3b). Of course these MSEs are conditional on the hyperparameters.

4. CONTINUOUS TIME FORMULATION

A continuous time model is, in some ways, more fundamental than a discrete model. It is not tied to a particular arbitrary time interval and it removes the constraint that the observations be at regular intervals. In order to allow for this possibility it will be assumed that observations are available at times t_τ , $\tau = 1,\ldots,T$.

Throughout this section we will adopt the convention of writing a variable which is a continuous function of time as, for example, $\mu(t)$. Writing μ_τ will denote the value of $\mu(t)$ at $t = t_\tau$. The other important convention is that a multivariate continuous white noise process, $\eta(t)$ say, with mean vector zero and covariance matrix $\bar{Q}$ will be defined as follows. Let

$$\eta^*(r,s) = \int_r^s \eta(t)dt \tag{4.1a}$$

Then

$$E[\eta^*(r,s)] = 0 \tag{4.1b}$$

and

$$E[\eta^*(r,s)\eta^*(r,s)'] = (s-r)\bar{Q} \tag{4.1c}$$

and

$$E[\eta^*(r_1,s_1)\eta^*(r_2,s_2)'] = $$

$$= 0 \quad \text{for } r_1 < s_1 < r_2 < s_2 \tag{4.1d}$$

The matrix $\bar{Q}$ will be referred to as the covariance matrix of $\eta(t)$

A first-order stochastic differential equation for an $n\times1$ vector of continuous variables, $\alpha(t)$, can be written:

$$\frac{d}{dt}[\alpha(t)] = A\,\alpha(t) + \eta(t) \tag{4.2}$$

where A is an $n\times n$ matrix of parameters and $\eta(t)$ is a multivariate continuous white noise process as defined in (4.1). The relationship between $\alpha(t_\tau)$ and $\alpha(t_{\tau-1})$, is

$$\alpha(t_\tau) = e^{A\delta}\,\alpha(t_{\tau-1}) + \int_0^\delta e^{A(\delta-s)}\,\eta(t_{\tau-1}+s)ds \tag{4.3}$$

where $\delta_\tau = t_\tau - t_{\tau-1}$, but the τ subscript has been dropped from δ for notational convenience. Expression (4.3) can be written as a discrete time transition equation of the form (3.1b) by noting that the transition matrix is

$$T_\tau = e^{A\delta}\,, \tag{4.4}$$

while the disturbance term has a mean of zero and a covariance matrix

$$Q_\tau = \int_0^\delta e^{A(\delta-s)}\,\bar{Q}\,e^{A'(\delta-s)}\,ds \tag{4.5}$$

In keeping with the convention introduced at the beginning of this section the subscript τ will replace the subscript t in the state space model and the Kalman Filter. Note that both T_τ and Q_τ depend on δ, but there is no requirement for δ to remain constant over time.

Equation (4.3) can be regarded as being equivalent to a discrete time transition equation of the form (3.1b). All that is required is a slight change of notation to set $\alpha(t_\tau)=\alpha_\tau$ and to let η_τ be the disturbance integral on the right hand side of (4.3). In order to make this discrete time transition equation operational it is necessary to evaluate (4.4) and (4.5). In Jones /19/ and Harvey and Stock /14/ , the problem of evaluating the integral in (4.5) is solved by diagonalizing the matrix A. In the continuous time analogue of (2.1), however, this is unnecessary as the derivations can be carried out directly.

Trend - A continuous time linear trend model can be expressed as

$$\frac{d}{dt}\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\begin{bmatrix}\mu(t)\\ \beta(t)\end{bmatrix} + \begin{bmatrix}\bar{\eta}(t)\\ \bar{\zeta}(t)\end{bmatrix} \; ; \tag{4.6}$$

see, for example, Jones /19/ , Kitagawa /20/ , or, in a slightly different context, Wecker and Ansley /25/ . Bearing in mind the definition of a matrix exponential,

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \ldots \tag{4.7}$$

it can be seen that the discrete time transition equation corresponding to (4.6) is

$$\begin{bmatrix}\mu_\tau \\ \beta_\tau\end{bmatrix} = \begin{bmatrix}1 & \delta \\ 0 & 1\end{bmatrix}\begin{bmatrix}\mu_{\tau-1} \\ \beta_{\tau-1}\end{bmatrix} + \begin{bmatrix}\eta_\tau \\ \zeta_\tau\end{bmatrix} \quad (4.8)$$

If the covariance matrix of $\{\bar{\eta}(t), \bar{\zeta}(t)\}$, is diag.$(\bar{\sigma}_\eta^2, \bar{\sigma}_\zeta^2)$, then the covariance matrix of (η_τ, ζ_τ), is

$$Q_\tau^\eta = \delta \left[\begin{array}{c|c}\bar{\sigma}_\eta^2 + \frac{\delta^2}{3}\bar{\sigma}_\zeta^2 & \frac{\delta}{2}\bar{\sigma}_\zeta^2 \\ \hline \frac{\delta}{2}\bar{\sigma}_\zeta^2 & \bar{\sigma}_\zeta^2\end{array}\right] \quad (4.9)$$

when $\delta = 1$, (4.8) reduces to (2.2a) and (2.2b) while (4.9) is identical to (2.2c). Thus letting the disturbances associated with the level and slope be independent of each other in continuous time does not imply that they will be independent of each other in discrete time. However if the observations are equally spaced, specifying Q to be of the form (2.2c) rather than being diagonal is of little practical importance with respect to forecasting. Nevertheless the Q^η matrix suggested by the continuous time formulation does resolve the problem of how best to specify this matrix for a discrete model; cf. the slightly different formulations adopted in Harrison and Stevens /10/ and Harvey and Todd /12/ and the discussion in Harrison and Akram /9/ , pp.32-3).

Cycle - A continuous cycle can be formulated as

$$\frac{d}{dt}\begin{bmatrix}\psi(t) \\ \psi^*(t)\end{bmatrix} = \begin{bmatrix}\bar{\rho} & \lambda_c \\ -\lambda_c & \bar{\rho}\end{bmatrix}\begin{bmatrix}\psi(t) \\ \psi^*(t)\end{bmatrix} + \begin{bmatrix}\kappa(t) \\ \kappa^*(t)\end{bmatrix} \quad (4.10)$$

where $\bar{\rho}$ and λ_c are parameters, the latter being the frequency of the cycle. The characteristic roots of the matrix containing these parameters are $\bar{\rho} + i\lambda_c$ and $\bar{\rho} - i\lambda_c$, where $i = \sqrt{-1}$. Since the condition for stationarity is that these roots have negative real parts, a stationary process must have $\bar{\rho} < 0$.

Using the definition of a matrix exponential, (4.7), it can be seen that, when $\delta = 1$, the discrete time model corresponding to (4.10) is of the form (2.4c), with $\bar{\rho} = \log\rho$. The non stationary model (2.3) is obtained when $\bar{\rho} = 0$. More generally, the transition matrix in (2.4a) is

$$C_\tau = \rho^\delta \begin{bmatrix}\cos\lambda\delta & \sin\lambda\delta \\ -\sin\lambda\delta & \cos\lambda\delta\end{bmatrix} \quad (4.11)$$

As regards the properties of the disturbances, it can be seen from (4.5) that if $\kappa(t)$ and $\kappa^*(t)$ are independent of each other and have equal variances, then in a discrete time κ_t and κ_t^* will also be independent and have equal variances for any δ. If $\bar{\sigma}_c^2$ denotes the common variance of $\kappa(t)$ and $\kappa^*(t)$. The covariance matrix of $(\kappa_t, \kappa_t^*)'$ is:

$$Q_\tau^\kappa = \rho^{2\delta}\,\bar{\sigma}_c^2 \cdot \delta I. \quad (4.12)$$

Letting $\kappa(t)$ and $\kappa^*(t)$ be independent and making their variances the same amounts to imposing one more restriction then is necessary for identifiability. However, having the specification of the discrete time model in (2.4a) independent of the period between observations is an attractive property.

A pseudo-cyclical process can also be formulated in continuous time by letting $\psi(t)$ follow a continuous AR(2) or ARMA (2,1) process; see Phadke and Wu /23/ . The attraction of (4.10) is that it is set up directly in terms of the parameters of interest.

Seasonal - The continuous time seasonal model consists of several pairs of components, $\lambda_j(t)$ and $\lambda_j^*(t)$, generated by a process of the form (4.10) with $\bar{\rho} = 0$ and λ_c replaced by the appropriate seasonal frequency λ_j. For irregularly spaced data the form of the transition matrices, S_{jt}, and the covariance matrices of the disturbances can be deduced directly from (4.11) and (4.12). Again the restrictions on the continuous time disturbances are more than enough to ensure identifiability in discrete time, but the discrete time specification holds irrespective of the length of time between observations.

Daily and Calender Effects - The continuous time daily model is

$$\frac{d}{dt}\,\Theta(t) = \xi(t) \quad (4.13a)$$

with

$$\mathrm{Var}[\xi(t)] = \bar{\sigma}_\xi^2[I - K^{-1}kk'] \quad (4.13b)$$

where κ and k are as defined in Section 2. Since $e^0 = I$, the discrete time model is of the form (2.1b), and in the general case of irregular observations

$$Var(\xi_t) = \bar{\sigma}_\xi^2 \delta [I - K^{-1} kk'] \quad (4.14)$$

Calender effects are treated in exactly the same way.

State Space Formulation and Maximum Likelihood Estimation.

For equally spaced observations, i.e. $\delta_\tau = 1$, the discrete time transition matrix derived from the continuous time formulation is as - in (3.1a). For irregularly spaced observations the transition matrix and the covariance matrix of the vector are no longer time invariant.

The state space model is completed by the specification of a measurement equation However, the possibility of irregularly spaced observations means that it is now necessary to discriminate between what economists call stocks and flows. A stock variable can be measured at one particular point in time, while a flow variable is an integral over a period of time. The money supply is an example of a stock, while income is a flow. In treating flows in this paper it is assumed that y_τ is the cumulated effect from $t_{\tau-1}$ to t_τ.

The measurement equation for a stock variable is

$$y_\tau = z_\tau' \alpha_\tau + \varepsilon_\tau \quad , \quad \tau = 1,\ldots,T \quad (4.15)$$

where ε_τ is a white noise disturbance term. It is straight-forward to apply the Kalman filter to (4.3) and (4.15); see, for example, Jones /19/ or Harvey and Stock /14/ . As in the discrete case, starting values can be constructed from the first n observations, where n is the number of nonstationary components in the stats vector. The likelihood function is of the form (3.5), but with the t subscript replaced by a τ subscript to indicate the possibility of irregularly spaced observations.

When the observations are flow variables, the measurement equation can be defined as an integral. If the state vector is partitioned as in (3.2) this takes the form

$$y_\tau = z_1' \int_0^\delta \alpha_1(t_{\tau-1} + s)ds + \int_0^\delta z_2'(s)\alpha_2(t_{\tau-1}+s) + \varepsilon_\tau$$

$$t = 1,\ldots,T, \quad (4.16)$$

where $z_2(s)$ is a continuous function, which may change each time there is a new day in the interval $t_{\tau-1}$ to t_τ. The irregular term ε_τ may or may not have a variance which is proportional to δ, depending on its interpretation. An integral measurement equation like (4.16) requires certain modifications in the Kalman filter. In the case of the model considered here, these modifications take a special form, the details of which are set out in the next section.

5. FLOW VARIABLES.

The modifications in the Kalman filter needed to handle a flow variable are set out below. It is then shown that when the observations are equally spaced, the forecast function for nonstationary components is the same as for a stock variable.

Modifications to the Kalman Filter.

Given the measurement equation in (4.16), it can be seen that the MMSE of y_τ at time $t_{\tau-1}$ is

$$\tilde{y}_{\tau/\tau-1} = z_1'\left[\int_0^\delta e^{A_1 s}\,ds\right]a_{1,\tau-1} + \left[\int_0^\delta z_2(s)ds\right]a_{2,\tau-1} \quad (5.1)$$

where A_1 denotes the block in the A matrix corresponding to components other than the daily and calender effects. The unit of time will be taken to be one day and it will be assumed, for simplicity, that δ is an integer so that the integral in the second term on the right hand side of (5.1) can be replaced by a summation. Given this change, the continuous function, $z_2(s)$, will be replaced by the discrete quantity, z_{2s}, the definition of which is analogous to the definition of z_{2t} in (3.2). Expression (5.1) can now be rewritten as

$$\tilde{y}_{\tau/\tau-1} = w_{1\tau}' a_{1,\tau-1} + w_{2\tau}' a_{2,\tau-1} = w_\tau' a_{\tau-1} \quad (5.2)$$

where

$$w'_{1\tau} = z'_1 \int_0^\delta e^{A_1 s}\, ds \qquad (5.3)$$

and

$$w_{2\tau} = \sum_{s=1}^{\delta} z_{2s} \qquad (5.4)$$

The prediction error is

$$\nu_\tau = y_\tau - \tilde{y}_{\tau/\tau-1} = w'_\tau(\alpha_{\tau-1} - a_{\tau-1}) + z'_1\int_0^\delta\int_0^s e^{A_1(s-u)}\eta_1(t_{\tau-1}+u)\,du\,ds + \sum_{s=1}^{\delta} z_{2s}\int_0^s \eta_2(t_{\tau-1}+u)\,du \qquad (5.5)$$

where the partitioning of the disturbance in (4.2) into $\eta_1(t)$ and $\eta_2(t)$ corresponds to the partitioning of the state vector into $\alpha_1(t)$ and $\alpha_2(t)$. Writing the covariance matrices of $\eta_1(t)$ and $\eta_2(t)$ as $\bar{Q}_1$ and $\bar{Q}_2$ respectively, and bearing in mind that they are independent of each other, the variance of the prediction error is

$$f_\tau = E(\nu_\tau^2) = w'_\tau P_{\tau-1} w_\tau + z'_1 Q_{1\tau}^{ff} z_1 + \sum_{s=1}^{\delta}\sum_{r=1}^{\delta} \min(r,s).z'_{2s}\bar{Q}_2 z_{2s} \qquad (5.6)$$

where

$$Q_{1\tau}^{ff} = \int_0^\delta\int_0^\delta\int_0^{\min(r,s)} e^{A_1(s-u)}\bar{Q}_1 e^{A'_1(r-u)}\,du\,dr\,ds \qquad (5.7)$$

Now consider the modifications required for the Kalman filter. The prediction equations, (3.3) are unchanged, but by following Harvey and Stock /14/ , it can be shown that the updating equations, (3.4), need to be replaced by

$$a_\tau = a_{\tau/\tau-1} + p^f_{\tau/\tau-1}\nu_\tau / f_\tau \qquad (5.8a)$$

and

$$P_\tau = P_{\tau/\tau-1} - p^f_{\tau/\tau-1} f_\tau^{-1} p^{f\prime}_{\tau/\tau-1} \qquad (5.8b)$$

where f_τ is defined by (5.6) and

$$p^f_{\tau/\tau-1} = T_\tau P_{\tau-1} w_\tau + \begin{bmatrix} q^{f\prime}_{1\tau} & q^{f\prime}_{2\tau}\end{bmatrix}' \qquad (5.9)$$

The $(n - w - w^* - 2)\times 1$ vector $q^f_{1\tau}$ is defined by $q^f_{1\tau} = Q^f_{1\tau} z_1$ where

$$Q^f_{1\tau} = \int_0^\delta\int_0^s e^{A_1(\delta-u)}\bar{Q}_1 e^{A'_1(\delta-s)}\,du\,ds \qquad (5.10)$$

while the $(w+ w^* - 2)\times 1$ vector $q^f_{2\tau}$ is

$$q^f_{2\tau} = \sum_{s=1}^{\delta} s\,\bar{Q}_2 z_{2s} \qquad (5.11)$$

The modified Kalman filter can be applied to yield a likelihood function for the y's exactly as before. However, in order to make it operational it is necessary to be able to evaluate $w_{1\tau}, z'_1 Q^{ff}_{1\tau} z_1$ and $q^f_{1\tau}$. Fortunately the block diagonality of A_1 and Q_1 makes it possible to obtain expressions for the trend, seasonal and cyclical components independently.

Trend - the trend elements in $w'_{1\tau}$ are

$$(1\quad 0)\int_0^\delta \begin{pmatrix}1 & s\\ 0 & 1\end{pmatrix} ds = \left(\delta,\ \delta^2/2\right) \qquad (5.12)$$

The trend contribution on the term $z'_1 Q^{ff}_{1\tau} z_1$ in (5.6) is

$$\int_0^\delta\int_0^\delta\int_0^{\min(r,s)} \{\bar{\sigma}_\eta^2 + \bar{\sigma}_\zeta^2(r-u)(s-u)\}\,du\,dr\,ds = \bar{\sigma}_\eta^2 \cdot \frac{\delta^3}{3} + \bar{\sigma}_\zeta^2 \cdot \frac{\delta^5}{20} \qquad (5.13)$$

Finally evaluating the appropriate block in $Q^f_{1\tau}$ and post-multiplying by the first two elements of z_1, i.e. $(1\quad 0)'$, gives the following two elements of $q^f_{1\tau}$

$$\left[\frac{\sigma^2}{2}\bar{\sigma}_\eta^2 + \frac{\delta^4}{8}\bar{\sigma}_\zeta^2\ ,\quad \frac{\delta^3}{6}\bar{\sigma}_\zeta^2\right]' \qquad (5.14)$$

Seasonal/Cycle - In considering a trigonometric component of the form (4.10), only the case $\bar{\rho} = 0$ will be considered. Formulae can be obtained for $\bar{\rho} < 0$ but they are rather more tedious to derive.

A pair of typical trigonometric components in $w'_{1\tau}$ are

$$\int_0^\delta [\cos\lambda s,\ \sin\lambda s]\,ds = \lambda^{-1}[\sin\lambda\delta\ ,\ 1 - \cos\lambda\delta]$$

The contribution to the term $z'_1 Q^{ff}_{1\tau} z_1$ is evaluated by first noting that the covariance matrix of the disturbances is scalar, i.e. $\bar{\sigma}_c^2$, and that if A_c denotes the matrix

$$A_c = \begin{bmatrix} 0 & \lambda\\ -\lambda & 0\end{bmatrix}$$

then $A_c + A'_c = 0$. Thus the appropriate block in (5.7) is

$$\bar{\sigma}_c^2 \int_0^\delta \int_0^\delta \int_0^{\min(r,s)} e^{A_c s} \cdot e^{A_c' r} \, du \, dr \, ds$$

The top left hand element of this matrix, i.e. the appropriate term in $z_1' Q z_1$, is

$$\bar{\sigma}_c^2 \int_0^\delta \int_0^\delta \min(r,s).\cos\lambda(r-s) dr\, ds$$

$$= \bar{\sigma}_c^2 \int_0^\delta \int_0^s r \cos\lambda(r-s).dr\, ds +$$

$$+ \bar{\sigma}_c^2 \int_0^\delta \int_0^r s \cos\lambda(r-s)\, ds\, dr$$

$$= \frac{2\,\bar{\sigma}_c^2\,\delta}{\lambda^2}\left[1 - \frac{\sin(\lambda\delta)}{\lambda\delta}\right] \qquad (5.16)$$

By similar reasoning, the corresponding pair of elements in q_{1T}^f is:

$$\left[\bar{\sigma}_c^2 (1 - \cos\lambda\delta)/\lambda^2 \quad , \quad 0\right]' \qquad (5.17)$$

Predictions

In making predictions of future values of flow variables ℓ steps ahead it is necessary to distinguish between the total accumulated effect from time t_T to $t_T + \ell$, and the amount of the flow in one time period at time $t_T+\ell$, i.e, from $t_T+\ell-1$ to $t_T+\ell$. The prediction of the former quantity is given by expression (5.1) with $\delta = \ell$ and $a_{\tau-1}$ replaced by a_T. Predictions for the latter are obtained by a similar formula but with the limits of integration being from $\ell - 1$ to ℓ. It is predictions of this second kind with which we are concerned below, i.e.

$$\tilde{y}^*_{T+\ell/T} = $$

$$= z_1'\left[\int_{\ell-1}^{\ell} e^{A_1 s} ds\right] a_{1T} + \left[\int_{\ell-1}^{\ell} z_2(s).ds\right] a_{2T}. \qquad (5.18)$$

For the trend component, the first term in square brackets on the right hand side of (5.18) is, when pre-multiplied by z_1',

$$\int_{\ell-1}^{\ell} (1 \quad s) ds = (1 \ , \ \ell - \tfrac{1}{2}) \qquad (5.19)$$

Thus the forecast function for the trend component is linear.

For a trigonometric term with $\bar{\rho} = 0$, the term corresponding to (5.19) is

$$\int_{\ell-1}^{\ell} (\cos\lambda s, \quad \sin\lambda s) ds$$

Post-multiplying by the appropriate two elements in a_T, and re-arranging, gives a forecast function of the form

$$c_T.\cos\lambda\ell + d_T.\sin\lambda\ell \qquad (5.20)$$

where c_T and d_T depend on the two elements in a_T.

The last term in (5.18) indicates that the appropriate forecast for a daily or calender effect is cimply the latest estimate of that effect. Taking this result together with (5.19) and (5.20) therefore shows that for a flow variable, the forecast function for nonstationary components takes the same form as the corresponding forecast function for a stock. Thus, for equally spaced observations, a flow variable can, for practical purposes, be treated in exactly the same way as a stock.

6. EXPLANATORY VARIABLES.

Explanatory variables, such as national income or rainfall can also be introduced into the model. Thus, if x_t denotes a kxl vector of explanatory variables at time t and δ_t denotes a corresponding vector of parameters, the model in (2.1) becomes

$$y_t = \mu_t + \psi_t + \gamma_t + \Theta_t + \Theta_t^* + x_t'\delta_t + \varepsilon_t, \qquad t=1,\ldots,T \qquad (6.1)$$

The parameters in δ_t change over time according to a random walk, i.e.

$$\delta_t = \delta_{t-1} + \nu_t \qquad (6.2)$$

where ν_t is a vector of random disturbances with mean zero and diagonal covariance matrix. Models of this kind have been proposed quite frequently in the econometric and engineering literature; see, inter alia, Cooley and -- Prescott /5/ , Garbade /8/ and Mehra /22/ . The appropriate statistical treatment consists of adding δ_t to the state vector in (3.1).

The continuous time analogue of (6.2) is

$$\frac{d}{dt}\delta(t) = \nu(t) \qquad (6.3)$$

and if the covariance matrix of $\upsilon(t)$ is diagonal, the implied covariance matrix of υ_τ is also diagonal, and proportional to the time between observations, δ. When y_τ is a stock variable, the appropriate values of the explanatory variables at time t_τ simply appear in the measurement equation. When y_τ is a flow variable, matters are more complicated since the term

$$z_3' \int_0^\delta \alpha_3(t_{\tau-1}+s)x(t_{\tau-1}+s)ds \qquad (6.4)$$

needs to be added to the right hand side of (4.16). The notation in (6.4) is such that $\alpha_3(t)$ denotes $\delta(t)$ while z_3 is a vector of ones. The problem arises because $x(t)$ is continuous and so a continuous record is needed to evaluate (6.4). If this is not available some kind of approximation must be used; cf. the paper by Phillips in Bergstrom /2/ , Ch.8). Note that explanatory variables which are flows need to be handled slightly differently from those which are stocks.

7. CONCLUSION.

Setting up a time series model in continuous time provides a rationale for certain types of formulations in discrete time. More specifically it leads to a class of models which do not depend on a particular, arbitrary, time interval. Furthermore it enables irregular patterns of observations to be handled. Although the emphasis in the paper has been on daily observations, all the techniques discussed are relevant for monthly, quarterly and annual observations. It is also possible to extend these techniques to situations where observations are available several times a day. Situations ot this kind arise in modelling the demand for electricity or telephone calls; see Bunn /4/ . In this context a set of sines and cosines may well provide a viable means of modelling the "profile" of demand for a particular day of the week.

8. APPENDIX - DAILY EFFECTS.

The definition of Θ_{wt} in (2.12) implies that the disturbance term, ξ_{wt}, obeys a similar restriction, i.e.

$$\xi_{wt} = -k_w^{-1}\sum_{j=1}^{w-1} k_j \ \xi_{jt}$$

This can be written in matrix terms as

$$\xi_{wt} = -k_w^{-1} k' \xi_t$$

where k and ξ_t are defined as for (2.13), and so

$$Var(\xi_{wt}) = k_w^{-2} k' \ Var(\xi)k$$

$$= k_w^{-2} k' (I - K^{-1} kk')k.\sigma_\xi^2$$

$$= k_w^{-2} [k'k - K^{-1}k'k.k'k].\sigma^2$$

$$= \sigma_\xi^2 k_w^{-2} [k'k(k'k+k_w^2)-(k'k)^2]/K$$

$$= \sigma_\xi^2 k'k/K = \sigma_\xi^2(K-k_w^2)/K = \sigma_\xi^2 [1 - k_w^2/K]$$

This is the same expression as (2.10). Furthermore

$$E(\xi_{wt}\ \xi_{jt}) = -k_w^{-1}\sum_{h=1}^{w-1} k_h\ E(\xi_{jt}\ \xi_{ht})$$

$$= \frac{-\sigma_\xi^2}{k_w}\left[k_j - \frac{k_j \sum k_h^2}{K}\right] = -\sigma_\xi^2 \frac{k_j}{k_w}\left[1 - \frac{(K-k_w^2)}{K}\right] =$$

$$= -\sigma_\xi^2 \frac{k_w k_j}{K}$$

which is (2.11).

9. FOOTNOTES.

This paper was written in August 1983 while I was a visitor at the Institute of Mathematics, Federal University of Rio de Janeiro. I am grateful to CNPq and FINEP for financial support. I am also grateful to the SSRC for support in connection with the LSE Econometrics Programme on Dynamic Econometric Models, Expectations, Innovations and Choice (DEMEIC).

1. When the model contains calendar effects it will not always be the case that starting values are formed from the first n observations due to multicollinearity. Brown, Durbin and Evans /3/ ,(p.152-3) discuss how to handle a situation of this kind in recursive least squares and the solution in this more general case is similar. If the large K

starting value technique is used, however, no changes are needed.

10. REFERENCES.

/1/ ANDERSON, B.D.O. and MOORE, J.B. "Optimal Filtering", Englewood Cliffs: Prentice-Hall (1979).

/2/ BERGSTROM, A.R. "Statistical Inference in Continuous Time Economic Models", Amsterdam: North Holland (1976).

/3/ BROWN, R.L., DURBIN, J. and EVANS, J.M., "Techniques for testing the constancy of regression relationships over time", Journal of the Royal Statistical Society, Series B, 37, 141-192 (1975).

/4/ BUNN, D.W., "Short-term forecasting: A Review of Procedures in the Electricity Supply Industry", Journal of the Operational Research Society, 33, 533-545 (1982).

/5/ COOLEY, T.F. and E.C. PRESCOTT, "Estimation in the Presence of Stochastic Parameter Variation", Econometrica, 44, 167-184 (1978).

/6/ ENGLE, R.F., "Estimating structural models of seasonality", in A. Zellner (ed), Seasonal Analysis of Economic Time Series", Washington, D.C.: Bureau of the Census, 281-308. (1978).

/7/ FRANZINI, L. and HARVEY, A.C., "Testing for Deterministic Trend and Seasonal Components in Time Series Models", Biometrika, 70, 673-682 (1983)

/8/ GARDADE, K., "Two methods for examining the stability of regression coefficients" Journal of the American Statistical Association, 72, 54-63 (1977).

/9/ HARRISON, P.J., and AKRAM, M., Generalised exponentiality weighted regression and parsimonious dynamic linear modelling O.D. Anderson (ed). Time Series Analysis: Theory and Practice 3. Amsterdam: North Holland Publishing Co. 19-42. (1983).

/10/ HARRISON, P.J. and STEVENS, C.F., "Bayesian forecasting", Journal of the Royal Statistical Society", Series B, 38, 205-247 (1976).

/11/ HARVEY, A.C., "Time Series Models", Deddington, Oxford: Philip Alland and New York: John Wiley/Halstead Press. (1981).

/12/ HARVEY, A.C. and TODD, P., "Forecasting economic time series with structural - and Box-Jenkins models: a case study", with discussion, Journal of Business and Economic Statistics, 1, 299-315 (1983).

/13/ HARVEY, A.C. and PIERSE, R.G., "Esti--mating missing observations in economic time series", Journal of American Statistical Association, (forthcoming) (1984).

/14/ HARVEY, A.C. and STOCK, J.H., "The Estimation of Higher Order Continuous Time Autoregressive Models", LSE Econometrics Programme Discussion Paper No. A. 38 (1983).

/15/ HARVEY, A.C. and PETERS, S., "Estimation Procedures for Structural Time Series Models", Unpublished paper, LSE. (1984).

/16/ HOTTA, L.K., "Identification and Testing of Hypotheses in Unobserved Components Models", Ph.D. Thesis, University of London (1983).

/17/ JONES, R.H., "Maximum Likelihood fitting of ARIMA models to time series with missing observations", Technometrics, 22, 389-395 (1980).

/18/ JONES, R.H., "Fitting a Continuous Time Autoregression to Discrete Data", Applied Time Series Analysis II, ed. by D.F.Findley. New York: Academic Press, (1981) 651-682.

/19/ JONES, R.H., "Fitting Multivariate Models to Unequally Spaced Data", in E. Parzen (ed), Time Series Analysis of Irregular Observations, New York: Springer-Verlag, (1984), to appear.

/20/ KITAGAWA, G., "A Nonstationary time series model and its fitting by a recursive filter", Journal of Time Series Analysis, 2, 103-116 (1981).

/21/ KOHN, R. and ANSLEY, C.F. "Fixed Interval Estimation in State Space Models when some of the Data are Missing or Aggregated", Biometrika, 70, 683-688 (1983).

/22/ MEHRA, R.K, "Kalman filters and their applications to forecasting", TIMS Studies in the Management Sciences, 12, 75-94 (1979).

/23/ PHADKE, M.S. and S.M. Wu, "Modelling of Continuous Stochastic Processes from Discrete Observations with Applications to Sunspots Data", Journal of the American Statistical Association, 69, 325-329 (1974).

/24/ ROSENBERG, B., "Random Coefficient Models: the Analysis of a Cross Section of Time Series by Stochastically Convergent Parameter Regression", Annals of Economic and Social Measurement, 2, 399-428 (1973).

/25/ WECKER, W.E. and ANSLEY, C.F., "The Signal Extraction Approach to Nonlinear Regression and Spline Smoothing", Journal of the American Statistical Association, 78, 81-89 (1983).