BIOMETRICS 39, 867–878
December 1983

Monitoring Renal Transplants: An Application of the Multiprocess Kalman Filter

A. F. M. Smith

Department of Mathematics, University of Nottingham, University Park, Nottingham NG7 2RD, England

and

M. West

Department of Statistics, University of Warwick, Coventry, Warwickshire, CV4 7AL, England

SUMMARY

The multiprocess Kalman filter offers a powerful general framework for the modelling and analysis of noisy time series which are subject to abrupt changes in pattern. It has considerable potential application to many forms of biological series used in clinical monitoring. In particular, the approach can be used to provide on-line probabilities of whether changes have occurred, as well as to identify the type of change that is involved. In this paper, we extend and illustrate the methodology within the context of a particular case study. The general features of the problem, and the approach adopted, will be seen to have wide application.

1. Introduction

In many situations where clinical monitoring is based upon series of quantitative measurements, the detection and interpretation of abrupt changes in the pattern of the time series is of paramount importance. Often, however, such series of data are difficult to interpret, even when facilities for visual inspection of graphical plots are available together with simple statistical summaries. In particular, many series are extremely noisy as a result of considerable biological variation and errors arising in the collection, measurement and processing of the data. In addition, the series may be subject to several different types of abrupt change. Some of these changes will correspond to biological events of direct interest and importance, possibly calling for immediate clinical intervention, whereas others will be of less direct clinical interest. It is therefore important not only to be able to detect changes in pattern, but also to distinguish between different forms of change; the need to do this often precludes the use of simple monitoring techniques such as the CUSUM procedure.

The multiprocess Kalman filter, introduced to statisticians by Harrison and Stevens (1976), provides a flexible general framework within which to model and analyse noisy time series subject to abrupt changes of pattern. In this paper, we illustrate both the modelling and analysis aspects of the approach by presenting a detailed case study carried out in collaboration with the Renal Unit at the City Hospital, Nottingham. Although the reported study is a particular one, it will be seen that it exhibits features that are present in many other contexts. The detailed methodology thus carries over, with minor modifications, to a number of other applications.

The particular problem studied was that of developing an on-line statistical procedure for

Key words: Multiprocess Kalman filter; Bayesian statistics; Recursive estimation; Serum creatinine; Renal transplants.

monitoring the progress of kidney function in individual patients who had recently received transplants. The basis of the procedure was to be a series of regular observations of the chemical indicator serum creatinine.

In §2, we discuss the relationship of serum creatinine to overall kidney function, and develop a model for the observed series of clinical readings. This model incorporates a variety of noise inputs resulting from data collection and measurement procedures, together with specific forms of abrupt changes that result from biological or clinical events. The model is seen to fit into the framework of the linear growth model introduced by Harrison and Stevens (1976, §5.2).

The latter model is itself a special case of the multiprocess Kalman filter (Harrison and Stevens, 1976, §5). In §3, we develop the necessary mathematical framework for calculating on-line probabilities of the occurrence of changes. In particular, we introduce an extension to the standard Kalman-filter recursive estimation procedure, which enables us to learn efficiently about unknown aspects of the noise characteristics of the series. This leads to increased sensitivity in detecting changes. Results of the case study are summarized in §4 and some concluding remarks follow in §5.

2. A Time-Series Model for Serum Creatinine

2.1 *Features of the Serum Creatinine Series*

Clinicians monitoring the progress of renal transplant patients are concerned to detect changes in the level of functioning of the transplanted kidney, particularly those changes which indicate the possible onset of a rejection.

The level of renal functioning is indicated by an unobservable factor, the 'glomerular filtration rule' (GFR), which measures the rate of clearance of various substances through the kidney. An indirect approach to learning about GFR is to measure blood and/or urine concentrations of selected chemicals and then to infer GFR by using basic relationships suggested by kidney physiology. In fact, most of the available chemical series appear to be too noisy for this strategy to be practicable. One of the few to show promise as an indicator in this context is the serum creatinine series, which forms the basis of the current study.

If a kidney is functioning normally, GFR is constant and creatinine is excreted at a constant rate. If GFR is increasing—perhaps as the result of improving function in a recently transplanted kidney—the blood concentration of creatinine will decrease. Conversely, if GFR decreases there will be an increase in observed concentration of creatinine. Sudden changes in the form of evolution of serum creatinine levels thus draw attention to the occurrence of significant biological events relating to changes in the functioning of the transplanted organ.

To obtain a more precise mathematical description of the evolution of the creatinine series, we note that: (i) in theory, GFR is proportional to the reciprocal of creatinine; (ii) since creatinine is being measured in terms of concentration, an adjustment should be made to compensate for measurement distortion induced by changes in body fluid. Using an adjustment formula based on body weight, Knapp *et al.* (1977) demonstrated that a plot of reciprocal body-weight-adjusted serum creatinine against time has the form of a succession of approximately linear trends, the directions of which switch as the patient moves from a period of improving kidney function to one of deteriorating function, or vice versa.

This switching straight-line representation of reciprocal body-weight-adjusted serum creatinine was successfully used by Smith and Cook (1980) as the basis for retrospective identification of points of change from improving to deteriorating kidney function (i.e. identification of the onset of rejection of the transplanted organ). In the present study, however, we are interested in developing an on-line procedure. Also, there are a number of

complicating features of the creatinine series which were not taken into account in the simpler retrospective analysis:

(i) Serum creatinine measurements were made at intervals of eight hours over a period of several weeks following transplant, and must therefore be modelled as a time series. The typical pattern of evolution of the series exhibits an initial period of poor renal function, followed by a gradual improvement and then by a rather erratic period. Rejection episodes alternate with periods of improvement, until—in cases that are eventually successful—the accepted organ approaches a level of stable functioning. It is clearly much easier to pick out such general features and trends retrospectively than it is to judge the state of renal function day by day using each new creatinine value as it becomes available.

(ii) Dialysis treatment is often provided in the early stages of postoperative care to support the transplanted kidney. The effect of dialysis treatment is to produce a sudden sharp drop in level of creatinine. This change in level contrasts with the changes in slope that correspond to the onsets, or reversals, of rejection episodes. From the point of view of an on-line statistical monitoring procedure, it would be useful to be able to distinguish these forms of change without recourse to complete patient histories.

(iii) Blood samples are not always collected precisely on schedule (in this case, at intervals of eight hours), and this introduces timing errors into any form of analysis that assumes equally spaced observations.

(iv) There are multiplicative measurement errors entering into the observed creatinine levels.

(v) Measurements of creatinine concentration are in units of μmol/l, but are only quoted to the nearest 10. There is, therefore, additional noise in the series due to reporting errors.

(vi) There are definite possibilities of gross errors, or outliers, in series values, either as a result of equipment malfunction, blood-sample contamination, or mistakes in data transcription.

Clearly, some of these features are more important than others and some might ultimately be disregarded as having negligible effect. However, in this paper we wish to illustrate the fact that all such features can be modelled and their influence can be incorporated into the analysis within the Kalman-filter framework.

2.2 *A Model for Steady Evolution of the Series*

If we denote by ϕ_{0_t} the actual serum creatinine level at Time t (measured in integer multiples of eight hours) and denote by ϕ_t the body-fluid-adjusted value, then

$$\phi_t = \omega_t \phi_{0_t}, \tag{2.1}$$

where ω_t, the adjustment factor, is a known function of observed body weight (see Knapp *et al.*, 1977). Assuming approximate linear trends during periods of steady evolution of renal function (where improvement or deterioration proceeds uninterruptedly), we may write

$$\left.\begin{aligned} \phi_t^{-1} &= \mu_t \\ &= \mu + \beta t, \end{aligned}\right\} \tag{2.2}$$

denoting the reciprocal body-weight-adjusted actual serum creatinine at Time t by μ_t.

If we take into account measurement errors and reporting errors (see §2.1), but temporarily ignore timing errors, then x_{0_t}, the measured value of ϕ_{0_t}, satisfies

$$x_{0_t} = e_t \phi_{0_t} + u_t, \tag{2.3}$$

where u_t denotes the reporting error and e_t denotes the measurement error. If we assume that $e_t = 1 + \varepsilon_t$, where ε_t is symmetrically distributed with mean 0 and constant variance, then, writing $a_t = \varepsilon_t \phi_{0_t}$, we can rewrite (2.3) in the form

$$x_{0_t} = \phi_{0_t} + a_t + u_t. \tag{2.4}$$

In what follows, we shall assume that the distribution of a_t is approximately N(0, $c^2\phi_0^2$), for some constant $c > 0$, and that the distribution of u_t is uniform over [−5, +5].

Letting x_t denote the body-weight-corrected version of x_{0_t}, we have

$$\left.\begin{aligned} x_t &= \omega_t x_{0_t} \\ &= \phi_t + \omega_t(a_t + u_t). \end{aligned}\right\} \tag{2.5}$$

Writing $y_t = x_t^{-1}$, from (2.2) and (2.5) we obtain

$$y_t = \mu_t(1 + s_t)^{-1}, \tag{2.6}$$

where

$$s_t = \mu_t\omega_t(a_t + u_t). \tag{2.7}$$

In Appendix 1, we show that $|s_t| \ll 1$, so that (2.6) may be approximated by

$$y_t = \mu_t(1 - s_t). \tag{2.8}$$

To complete our model for the reciprocal body-weight-adjusted observed serum creatinine series, we now introduce a further factor to take account of the timing error (see §2.1). If we assume that there is a symmetrically distributed perturbation, r_t, say, around t, the scheduled time of data collection, then the right-hand side of (2.2) should really be changed to $\mu + \beta(t + r_t)$ and (2.8) should be redefined. If, instead, we retain the notation $\mu_t = \mu + \beta t$, then (2.8) must be rewritten in the form

$$y_t = \mu_t + v_t, \tag{2.9}$$

where

$$v_t = -s_t\mu_t + (1 - s_t)\beta r_t. \tag{2.10}$$

In our application, it is reasonable to assume that $|r_t| \leqslant 1/16$; in other words, that blood samples are collected within 30 minutes of the scheduled times. With this assumption, it is shown in Appendix 2 that the distribution of v_t in (2.9) is approximately N(0, $c^2\mu_t^2$).

2.3 *A Model for Sudden Changes in the Series*

Combining (2.2) and (2.9), we can express our model for steadily evolving parts of the series in the form

$$y_t = \mu_t + v_t, \tag{2.11}$$

$$\mu_t = \mu_{t-1} + \beta_t + \gamma_t, \tag{2.12}$$

$$\beta_t = \beta_{t-1} + \delta_t, \tag{2.13}$$

where $v_t \sim$ N(0, $c^2\mu_t^2$), and γ_t and δ_t each have N(0, 0) distributions. By permitting the variances of γ_t and δ_t to be small but nonzero, we could allow for approximate nondeterministic linear trends in μ_t. In any case, Equations (2.11) to (2.13)—the linear growth model of Harrison and Stevens (1976, §5.2)—provide a convenient starting point for the extension of the steadily evolving model to incorporate possible sudden changes in the series.

Returning to the discussion of §2.1, we see that there are three types of abrupt change in the steadily evolving pattern that need to be considered.

Outliers. A gross error corresponds to a sudden perturbation to (2.11). Instead of y_t being an unbiased measurement of μ_t, with variance $c^2\mu_t^2$, we obtain a measurement whose accuracy is in considerable doubt. A convenient way of modelling this is to regard (2.11) as still applying, but with a variance considerably larger than $c^2\mu_t^2$.

Changes in level. Dialysis just prior to Time t produces a sudden perturbation to (2.12). Again, a convenient way to model this is to utilize the form of (2.12), but this time take γ_t as normally distributed about 0 with a large variance.

Changes in slope. Actual changes in trend, from improvement to deterioration, or vice versa, produce sudden perturbations in (2.13). These will be modelled by retaining the form of (2.13), but taking δ_t as normally distributed about 0 with a large variance.

Throughout, we shall assume that v_t, γ_t and δ_t are independent random quantities. If we express their distributions in the form N(0, $c^2\mu_t^2 K_v$), N(0, $c^2 K_\gamma$) and N(0, $c^2 K_\delta$), respectively, then by choosing various combinations of K_v, K_γ and K_δ to be zero or to be large positive numbers as appropriate, we can model, via (2.11) to (2.13), the four possible states of the series at any instant: steady state; change of level; change of slope; outlier.

In the next section, we recall the representation of this model within the multiprocess-Kalman-filter framework (Harrison and Stevens, 1976, §5), and present a recursive procedure for calculating on-line probabilities of which of the four states pertains at any given time.

3. The Multiprocess Kalman Filter

3.1 *Representation of the Linear Growth Model*

The system described by (2.11) to (2.13) can be rewritten in the form

$$\left.\begin{aligned} y_t &= \mathbf{F}\boldsymbol{\theta}_t + v_t, \\ \boldsymbol{\theta}_t &= \mathbf{G}\boldsymbol{\theta}_{t-1} + \mathbf{w}_t, \end{aligned}\right\} \tag{3.1}$$

where

$$\mathbf{F} = [1 \quad 0],$$

$$\mathbf{G} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

$$\boldsymbol{\theta}_t^T = [\mu_t, \beta_t],$$

$$\mathbf{w}_t^T = [\gamma_t + \delta_t \quad \delta_t],$$

and v_t and $\mathbf{w}_t$ are normally distributed with zero means.

To define V_t, the variance of v_t, and $\mathbf{W}_t$, the variance–covariance matrix of $\mathbf{w}_t$, we first introduce some further notation. We write $M_t^{(j)}$ to denote the assumption that the observation at Time t belongs to State j, where the four possible states are numbered as follows:

$j = 1$, steady state;

$j = 2$, change in level;

$j = 3$, change in slope;

$j = 4$, outlier.

We then write $V_t^{(j)}$ and $\mathbf{W}_t^{(j)}$ to denote the values of V_t and $\mathbf{W}_t$, respectively, when $M_t^{(j)}$ is assumed. Similarly, we write $K_v^{(j)}$, $K_\gamma^{(j)}$ and $K_\delta^{(j)}$ to denote the corresponding values of K_v, K_γ and K_δ, as discussed at the end of §2.3.

It then follows that

$$\begin{aligned} \operatorname{var}(v_t \mid \mu_t, M_t^{(j)}) &= V_t^{(j)} \\ &= c^2\mu_t^2 K_v^{(j)}, \\ \operatorname{var}(\mathbf{w}_t \mid M_t^{(j)}) &= \mathbf{W}_t^{(j)} \\ &= c^2 \mathbf{K}_w^{(j)}, \end{aligned} \tag{3.2}$$

where

$$\mathbf{K}_w^{(j)} = \begin{bmatrix} K_\gamma^{(j)} + K_\delta^{(j)} & K_\delta^{(j)} \\ K_\delta^{(j)} & K_\delta^{(j)} \end{bmatrix}.$$

A discussion of the choice of suitable values of $K_v^{(j)}$ etc. is given in §4.1.

3.2 *Calculation of Probabilities: c^2 Known.*

In this section, we shall outline the recursive formulae required for calculating on-line probabilities for the states of the system (3.1) assuming a known value of c^2. We shall adopt the notation $D_t = (y_1, \ldots, y_t)^T$ and assume that, for any event E defined in terms of the history of the system prior to Time t,

$$\begin{aligned} p(M_t^{(j)} \mid E) &= p(M_t^{(j)}) \\ &= p_0^{(j)}, \end{aligned} \tag{3.3}$$

the *a priori* probability of being in State j. Further discussion of the choice of these values is given in §4.1.

We now make the assumption (to which we shall return shortly) that the conditional distribution of $\boldsymbol{\theta}_{t-1}$, given $M_{t-1}^{(i)}$ and D_{t-1}, is given by

$$(\boldsymbol{\theta}_{t-1} \mid M_{t-1}^{(i)}, D_{t-1}) \sim \mathrm{N}(\mathbf{m}_{t-1}^{(i)}, c^2\mathbf{C}_{t-1}^{(i)}), \tag{3.4}$$

and we recall from (3.1) that

$$\left.\begin{aligned} (y_t \mid M_t^{(j)}, \boldsymbol{\theta}_t) &\sim \mathrm{N}(\mathbf{F}\boldsymbol{\theta}_t, V_t^{(j)}), \\ (\boldsymbol{\theta}_t \mid M_t^{(j)}, \boldsymbol{\theta}_{t-1}) &\sim \mathrm{N}(\mathbf{G}\boldsymbol{\theta}_{t-1}, \mathbf{W}_t^{(j)}). \end{aligned}\right\} \tag{3.5}$$

To proceed with the normal analysis, we require the variances $V_t^{(j)}$ to be known. However, from (3.2), they depend on the unknown μ_t. At this stage, therefore, we make an approximation and replace μ_t by $\mathrm{E}(\mu_t \mid D_{t-1})$, which we denote by $\bar{\mu}_t$ [a more refined approximation could be based on $\mathrm{E}(\mu_t^2 \mid D_{t-1})$, if required].

It follows from standard results (see, for example, Lindley and Smith, 1972), that

$$(y_t \mid M_t^{(j)}, M_{t-1}^{(i)}, D_{t-1}) \sim \mathrm{N}[\mathbf{FGm}_{t-1}^{(i)}, c^2\{K_v^{(j)}\bar{\mu}_t^2 + \mathbf{F}(\mathbf{GC}_{t-1}^{(i)}\mathbf{G}^T + \mathbf{K}_w^{(j)})\mathbf{F}^T\}], \tag{3.6}$$

$$(\boldsymbol{\theta}_t \mid M_t^{(j)}, M_{t-1}^{(i)}, D_t) \sim \mathrm{N}(\mathbf{m}_t^{(i,j)}, c^2\mathbf{C}_t^{(i,j)}), \tag{3.7}$$

where

$$(\mathbf{C}_t^{(i,j)})^{-1} = \mathbf{F}^T(K_v^{(j)}\bar{\mu}_t^2)^{-1}\mathbf{F} + (\mathbf{GC}_{t-1}^{(i)}\mathbf{G}^T + \mathbf{K}_w^{(j)})^{-1} \tag{3.8}$$

and

$$\mathbf{m}_t^{(i,j)} = \mathbf{C}_t^{(i,j)}\mathbf{F}^T(K_v^{(j)}\bar{\mu}_t^2)^{-1}(\mathbf{y}_t - \mathbf{FGm}_{t-1}^{(i)}) + \mathbf{Gm}_{t-1}^{(i)} \tag{3.9}$$

These forms provide the basis for a recursive updating procedure, starting from an initial assumption $\boldsymbol{\theta}_0 \sim \mathrm{N}(\mathbf{m}_0, c^2\mathbf{C}_0)$. We shall discuss the choice of $\mathbf{m}_0$, $\mathbf{C}_0$ in §4.1.

To complete the development of the recursive procedure, we note that by straightforward application of Bayes' theorem, together with (3.3), we obtain, for $1 \leqslant i, j \leqslant 4$,

$$\begin{aligned} p_t^{(i,j)} &= p(M_t^{(j)}, M_{t-1}^{(i)} \mid D_t) \\ &\propto p(y_t \mid M_t^{(j)}, M_{t-1}^{(i)}, D_{t-1})p(M_{t-1}^{(i)} \mid D_{t-1})p_0^{(j)} \\ &\propto p(y_t \mid M_t^{(j)}, M_{t-1}^{(i)}, D_{t-1})p_{t-1}^{(i)}p_0^{(j)}, \end{aligned} \tag{3.10}$$

where, in general, $p_t^{(j)}$ denotes the probability, given D_t, that y_t is an observation from the jth state of the system. Since

$$p_t^{(j)} = \sum_{i=1}^{4} p_t^{(i,j)}, \tag{3.11}$$

(3.10), together with (3.6), enables us to calculate the $p_t^{(j)}$ values recursively. Given our initial assumption about the distribution of $\boldsymbol{\theta}_0$, $p(y_1 \mid M_1^{(j)})$ is obtained immediately from (3.6) by replacing $\mathbf{m}_{t-1}^{(i)}$ and $\mathbf{C}_{t-1}^{(i)}$ by $\mathbf{m}_0$ and $\mathbf{C}_0$, respectively, and approximating $\mu_t = \mu_1$ by $\bar{\mu}_1$, the first

component of $\mathbf{m}_0$. The recursion for $p_t^{(j)}$ then has the starting value

$$p_1^{(j)} \propto p(y_1 \mid M_1^{(j)}) p_0^{(j)}. \tag{3.12}$$

We return now to the distributional assumption summarized by (3.4). Whilst this form is essential if we are to retain the simple recursive forms of (3.6) to (3.12), we note that it must be regarded as an approximation. To see this, we note from (3.7) that

$$(\boldsymbol{\theta}_t \mid M_t^{(j)}, D_t) \sim \sum_{i=1}^{4} (p_t^{(i,j)}/p_t^{(j)}) \mathrm{N}(\mathbf{m}_t^{(i,j)}, c^2 \mathbf{C}_t^{(i,j)}), \tag{3.13}$$

which only has the general form of (3.4) if we replace the right-hand side by an $\mathrm{N}(\mathbf{m}_t^{(j)}, c^2 \mathbf{C}_t^{(j)})$ approximation, where

$$\mathbf{m}_t^{(j)} = \sum_{i=1}^{4} (p_t^{(i,j)}/p_t^{(j)}) \mathbf{m}_t^{(i,j)} \tag{3.14}$$

and

$$\mathbf{C}_t^{(j)} = \sum_{i=1}^{4} (p_t^{(i,j)}/p_t^{(j)}) \{\mathbf{C}_t^{(i,j)} + (\mathbf{m}_t^{(i,j)} - \mathbf{m}_t^{(j)})(\mathbf{m}_t^{(i,j)} - \mathbf{m}_t^{(j)})^T\}. \tag{3.15}$$

Using this form of 'mixture-collapsing' procedure (introduced by Harrison and Stevens, 1976), and replacing μ_t in (3.2) by $\bar{\mu}_t$, the first component of

$$\mathbf{m}_{t-1} = \sum_{j=1}^{4} \mathbf{p}_{t-1}^{(j)} \mathbf{m}_{t-1}^{(j)}, \tag{3.16}$$

we obtain a closed-form, easily calculated, recursive procedure which, in particular, returns on-line probabilities, $p_t^{(j)}$, $j = 1, \ldots, 4$, of the current state of the system.

In addition to assessing the current state of the system, it will be useful to have available quantities such as $p(M_t^{(j)} \mid D_{t+1})$ and $p(M_t^{(j)} \mid D_{t+2})$, which provide one-step-back and two-step-back assessments of the previous states of the system in the light of current information.

To calculate these quantities, we observe that

$$\begin{aligned} p(M_t^{(j)} \mid D_{t+1}) &= p(M_t^{(j)} \mid D_t, y_{t+1}) \\ &\propto p(y_{t+1} \mid M_t^{(j)}, D_t) p(M_t^{(j)} \mid D_t) \\ &\propto \left\{ \sum_{k=1}^{4} p(y_{t+1} \mid M_{t+1}^{(k)}, M_t^{(j)}, D_t) p_0^{(k)} \right\} p_t^{(j)}, \end{aligned} \tag{3.17}$$

and that (3.6), (3.10) and (3.11) provide explicit forms for the right-hand side of (3.17). Similarly,

$$\begin{aligned} p(M_t^{(j)} \mid D_{t+2}) &= p(M_t^{(j)} \mid D_{t+1}, y_{t+2}) \\ &\propto p(y_{t+2} \mid M_t^{(j)}, D_{t+1}) p(M_t^{(j)} \mid D_{t+1}). \end{aligned} \tag{3.18}$$

The right-hand side of (3.18) can be rewritten in the form

$$\sum_{i=1}^{4} \sum_{k=1}^{4} p(y_{t+2} \mid M_{t+2}^{(i)}, M_{t+1}^{(k)}, M_t^{(j)}, D_{t+1}) p(M_{t+2}^{(i)}, M_{t+1}^{(k)}, M_t^{(j)} \mid D_{t+1}), \tag{3.19}$$

which, given (3.3) and the structure of (3.5), reduces to

$$\sum_{i=1}^{4} \sum_{k=1}^{4} p(y_{t+2} \mid M_{t+2}^{(i)}, M_{t+1}^{(k)}, D_{t+1}) p_0^{(i)} p_{t+1}^{(j,k)}. \tag{3.20}$$

This can be calculated straightforwardly by use of (3.6) and (3.10). In §4.2 there is a discussion of the uses of (3.11), (3.17) and (3.20) in the renal transplant study.

3.3 *Calculation of Probabilities: c^2 Unknown*

In the previous section, we assumed c^2 to be known. We now extend the analysis to cover the case of $\lambda = c^{-2}$ unknown, using ideas from West (1981).

We retain the assumption (3.3) and rewrite (3.4) in the form

$$(\boldsymbol{\theta}_{t-1} \mid M_{t-1}^{(i)}, D_{t-1}, \lambda) \sim \mathrm{N}(\mathbf{m}_{t-1}^{(i)}, \lambda^{-1}\mathbf{C}_{t-1}^{(i)}). \tag{3.21}$$

To deal with the unknown scale parameter λ, we shall assume that, given $M_{t-1}^{(i)}$, D_{t-1}, the distribution of λ can be approximated by

$$(\lambda \mid M_{t-1}^{(i)}, D_{t-1}) \sim G(\tfrac{1}{2}n_{t-1}, \tfrac{1}{2}b_{t-1}^{(i)}), \tag{3.22}$$

where $U \sim G(\alpha, \beta)$ signifies that U has a gamma distribution with density

$$p(u) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} u^{\alpha-1}\exp(-\beta u), \quad u > 0. \tag{3.23}$$

We assume that, initially, λ has a $G(\frac{1}{2}n_0, \frac{1}{2}b_0)$ distribution. Together, (3.21) and (3.22) constitute the assumption of the normal-gamma joint-conjugate form for $(\boldsymbol{\theta}_{t-1}, \lambda)$; see, for example, DeGroot (1970). Standard Bayesian analysis (De Groot, 1970, §9.6) shows immediately that

$$(\boldsymbol{\theta}_t \mid M_t^{(j)}, M_{t-1}^{(i)}, D_t, \lambda) \sim \mathrm{N}(\mathbf{m}_t^{(i,j)}, \lambda^{-1}\mathbf{C}_t^{(i,j)}) \tag{3.24}$$

and

$$(\lambda \mid M_t^{(j)}, M_{t-1}^{(i)}, D_t) \sim G(\tfrac{1}{2}n_t, \tfrac{1}{2}b_t^{(i,j)}), \tag{3.25}$$

where $\mathbf{m}_t^{(i,j)}$ and $\mathbf{C}_t^{(i,j)}$ are given by (3.8) and (3.9),

$$n_t = n_{t-1} + 1 \tag{3.26}$$

and

$$b_t^{(i,j)} = b_{t-1}^{(i)} + \{K_v^{(j)}\bar{\mu}_t^2 + \mathbf{F}(\mathbf{G}\mathbf{C}_{t-1}^{(i)}\mathbf{G}^T + \mathbf{K}_w^{(j)})\mathbf{F}^T\}^{-1}(y_t - \mathbf{F}\mathbf{G}\boldsymbol{m}_{t-1}^{(i)})^2. \tag{3.27}$$

To calculate $p_t^{(i,j)}$, given by (3.10), we note that

$$p(y_t \mid M_t^{(j)}, M_{t-1}^{(i)}, D_{t-1}) = \int_0^{\infty} p(y_t \mid M_t^{(j)}, M_{t-1}^{(i)}, D_{t-1}, \lambda)p(\lambda \mid M_{t-1}^{(i)}, D_{t-1})\, d\lambda. \tag{3.28}$$

The first term in the integral is given by (3.6), writing $c^2 = \lambda^{-1}$, and the second term is defined by (3.22). It follows (see, for example, Aitchison and Dunsmore, 1975) that (3.28) has the form of a t-density, proportional to

$$\{K_v^{(j)}\bar{\mu}_t^2 + \mathbf{F}(\mathbf{G}\mathbf{C}_{t-1}^{(i)}\mathbf{G}^T + \mathbf{K}_w^{(j)})\mathbf{F}^T\}^{-\frac{1}{2}}(b_t^{(i,j)})^{-\frac{1}{2}n_t}. \tag{3.29}$$

Calculation of $p_t^{(i,j)}$ and $p_t^{(j)}$ now follows from (3.10) to (3.12), where the latter has the factor $p(y_1 \mid M_1^{(j)})$ proportional to (3.29) with $\mathbf{m}_{t-1}^{(i)}$ and $\mathbf{C}_{t-1}^{(i)}$ replaced by $\mathbf{m}_0$ and $\mathbf{C}_0$, respectively, n_{t-1} and $b_{t-1}^{(i)}$, replaced by n_0 and b_0, respectively, and μ_t approximated by the first component of $\mathbf{m}_0$. Collapsing of the mixture distribution arising for $\boldsymbol{\theta}_t$ follows from (3.13) to (3.16), but with (3.13) now representing the distribution of $(\boldsymbol{\theta}_t \mid M_t^{(j)}, D_t, \lambda)$. Similarly, we need a collapsing procedure for the distribution of λ, since, from (3.25),

$$(\lambda \mid M_t^{(j)}, D_t) \sim \sum_{i=1}^{4} (p_t^{(i,j)}/p_t^{(j)})G(\tfrac{1}{2}n_t, \tfrac{1}{2}b_t^{(i,j)}). \tag{3.30}$$

Recalling that the mean of a $G(\alpha, \beta)$ distribution is α/β, if we seek to replace (3.30) by a single $G(\frac{1}{2}n_t, \frac{1}{2}b_t^{(j)})$ distribution it is natural to define $b_t^{(j)}$ by

$$(b_t^{(j)})^{-1} = \sum_{i=1}^{4} (p_t^{(i,j)}/p_t^{(j)})(b_t^{(i,j)})^{-1}. \tag{3.31}$$

Alternatively, all the collapsing procedures used in this section can be justified as minimizing the Kullback–Liebler divergence between the mixture and the simple density. As in §3.2, we approximate μ_t, wherever it occurs, by the first component of $\mathbf{m}_{t-1}$ defined by (3.16).

Current, one-step-back and two-step-back probability assessments of the state of the system again follow from (3.10), (3.11) and (3.17) to (3.20), where the factor previously defined by (3.6) is now given by (3.29).

4. Application to Renal Monitoring

4.1 *Inputs*

In the case of known coefficient of variation c, we need to supply initial values for $p_0^{(j)}$, $K_v^{(j)}$, $K_\gamma^{(j)}$ and $K_\delta^{(j)}$, $j = 1, \ldots, 4$, as well as for $\mathbf{m}_0$ and $\mathbf{C}_0$.

On the basis of careful retrospective study of past creatinine series, the following parameters were used:

$$(p_0^{(1)}, p_0^{(2)}, p_0^{(3)}, p_0^{(4)}) = (.85, .06, .07, .02).$$

These parameters reflect the following assumed rates: outliers, about 2%; steady state observations, 85%; observations affected by dialysis, about 6%, with the remaining 7% corresponding to changes in slope.

The variance multiples that were found to be most suitable (following empirical trials with the system) were as follows:

j	$K_v^{(j)}$	$K_\gamma^{(j)}$	$K_\delta^{(j)}$
1	1	0	0
2	1	90	0
3	1	0	60
4	100	0	0

.

As far as the parameters $\mathbf{m}_0$ and $\mathbf{C}_0$ are concerned, clinicians are usually fairly confident that initial creatinine levels are about 1000 μmol/l, so that $\mu_0 \simeq 0.001$, and that, initially, changes are fairly slow so that $\beta_0 \simeq 0$. We have therefore taken

$$\mathbf{m}_0 = \begin{pmatrix} 0.001 \\ 0 \end{pmatrix},$$

$$\mathbf{C}_0 = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix},$$

the various values in $\mathbf{C}_0$ reflecting relatively great uncertainty about μ_0 and β_0 (given the typical value of μ_0).

In the case of unknown c^2, we need, in addition, to specify n_0 and b_0. In general, the hospital quality-control laboratory can provide rather good estimates of the coefficient of variation c. However, we have found that the learning procedure described in §3.3 is extremely good, even when we set $n_0 = 0$ and $b_0 = 0$ to reflect very vague prior knowledge of $\lambda = c^{-2}$. The adequacy of the approximation in other applications would, of course, require separate validation in each case.

4.2 *Outputs*

Figure 1 shows the output from an on-line computer system, based on §3.1 and §3.2, in use at the Renal Unit of the City Hospital, Nottingham. It should be borne in mind that, in practice, output is displayed gradually, evolving from left to right as successive observations become available.

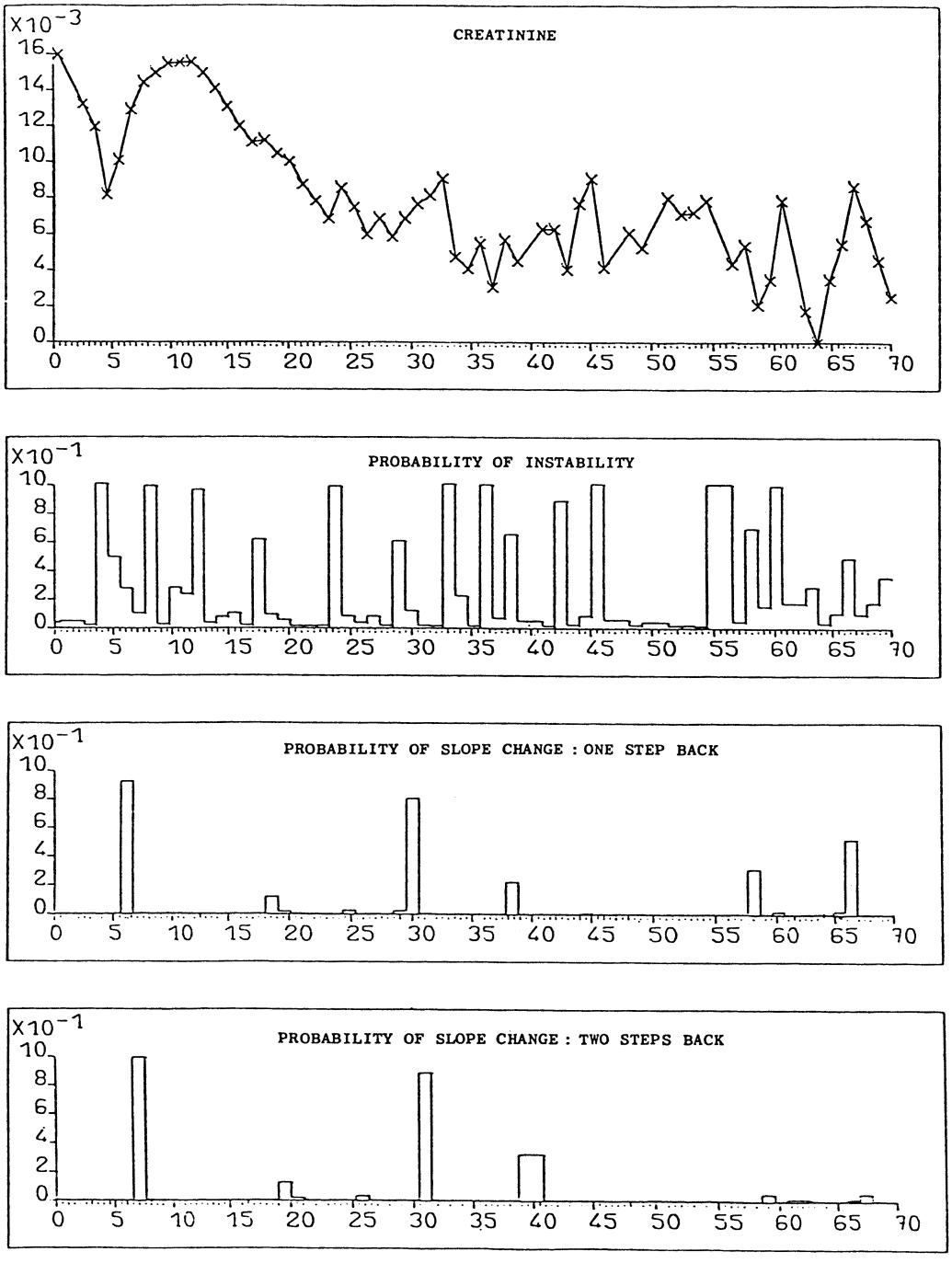

Figure 1. On-line monitoring of serum creatinine.

The upper graph displays the weight-corrected creatinine values plotted on an inverted reciprocal scale. A change of slope from negative to positive thus indicates a deterioration of renal function. (In fact, only every third value is plotted; the analysis uses *all* the creatinine values.)

The second graph displays values of $1 - p_t^{(1)}$, the posterior probability that at Time t there is some form of instability in the evolution of the series.

The third graph displays values of $p(M_t^{(3)} \mid D_{t+1})$ in cases where $\mathrm{E}(\beta_t \mid D_t)$, the second component of $\mathbf{m}_t = \sum_j p_t^{(j)} \mathbf{m}_t^{(j)}$, is negative. High values occurring on this plot indicate points at which, looking one-step-back, there is high probability of a shift from decreasing to increasing creatinine.

The fourth graph displays values of $p(M_t^{(3)} \mid D_{t+2})$ corresponding to the time points selected for the third graph. This acts as a 'final' confirmatory check that 'interesting' values of the third graph represent genuine changes in slope (rather than a wobble in a change of level, or two successive outliers, or whatever).

In general, these probabilities can be combined with specific utility assessments in order to arrive at optimal decisions concerning clinical intervention or, alternatively, rule-of-thumb 'cut-off' values can be chosen such that action is taken if, say, probabilities on the fourth graph exceed the cut-off value. This latter kind of strategy has been adopted at Nottingham, and has resulted in a system which, essentially, flags precisely those changes that experienced clinicians subsequently agree to be onsets of rejection episodes and, moreover, identifies them, on average, at least a day before the clinicians. Details of how the cut-off values were chosen and of the validation of the system were reported by Trimble *et al.* (1983).

5. Discussion

We have presented a methodology for modelling and monitoring biological time series, which has the flexibility to deal with a wide range of problems that involve a variety of noise inputs and possible abrupt changes in pattern. The features present in the renal monitoring study will be seen to occur widely in many applications, and we conjecture that the approach illustrated here will also prove successful in a large number of other situations.

The detailed development presented in §3 is related to the work of Harrison and Stevens (1976), but departs from their analysis in two respects. First, our emphasis is much more on using the system-state probabilities as ends in themselves rather than as aids to forecasting the future evolution of the system. Secondly, the procedure for learning about the unknown coefficient of variation given in §3.3 has been shown in detailed numerical studies to be a considerable improvement over the method proposed by Harrison and Stevens, and overcomes the criticisms of the multiprocess-Kalman-filter approach made by Stoodley and Mirnia (1979).

On-line monitoring systems developed with this approach have two possible functions. First, after suitable validation they may come to be accepted by a group of clinicians as providing genuine objective guidance in the context of noisy unstable series where unaided eyeball approaches lead to much controversy and uncertainty. Secondly, in the case of less noisy series, where eyeball approaches would not be too unreasonable, computer implementation of such procedures can free clinicians from what would be a routine but time-consuming task. We have experienced both kinds of reaction in the renal context, where we are developing extensions of these procedures to deal with series of indicators involving cyclical features and to model several indicators simultaneously by using a multivariate form of (3.1). More detailed discussions of the clinical aspects of the collaboration and the validation of the procedure were provided by I. M. G. Trimble (in a M.Phil. thesis at the University of Nottingham, 1980) and Trimble *et al.* (1983). A non-technical summary of the approach was given by Smith *et al.* (1983).

ACKNOWLEDGEMENTS

We are indebted to Dr M. S. Knapp and his colleagues at the Renal Unit, City Hospital, Nottingham for continuing close collaboration on this project. In particular, we should like to thank Ian Trimble for his major contribution to our understanding of many of the points discussed in §2.1.

Résumé

La méthode du filtre de Kalman constitue un outil général et puissant pour modéliser et analyser les séries temporelles avec bruit qui présentent des variations brusques de comportement. Un important domaine d'application est celui des nombreuses séries temporelles biologiques recueillies au cours du monitorage clinique. En particulier, la méthode peut être utilisée pour fournir —on line—des probabilités de survenue de variations et identifier le type de variation. Cet article développe la méthodologie existante et l'illustre par un exemple particulier. On verra qu'elle peut servir à de nombreuses applications.

References

Aitchison, J. and Dunsmore, I. R. (1975). *Statistical Prediction Analysis.* Cambridge: Cambridge University Press.

De Groot, M. H. (1970). *Optimal Statistical Decisions.* New York: McGraw-Hill.

Harrison, P. J. and Stevens, C. F. (1976). Bayesian forecasting (with Discussion). *Journal of the Royal Statistical Society, Series B* **38,** 205–247.

Knapp, M. S., Blamey, R. W., Cove-Smith, J. R. and Heath, M. (1977). Monitoring the function of renal transplants. *Lancet* **2,** 1183.

Lindley, D. V. and Smith, A. F. M. (1972). Bayes estimates for the linear model (with Discussion). *Journal of the Royal Statistical Society, Series B* **34,** 1–41.

Smith, A. F. M. and Cook, D. G. (1980). Straight lines with a change-point: a Bayesian analysis of some renal transplant data. *Applied Statistics* **29,** 180–189.

Smith, A. F. M., West, M., Gordon, K., Knapp, M. S. and Trimble, I. M. G. (1983). Monitoring kidney transplant patients. *Statistician* **32,** 46–54.

Stoodley, K. D. C. and Mirnia, M. (1979). The automatic detection of transients, step changes and slope changes in the monitoring of medical series. *Statistician* **28,** 163–170.

Trimble, I. M. G., West, M., Knapp, M. S., Pownall, R. and Smith, A. F. M. (1983). Detection of renal allograft rejection by computer. *British Medical Journal* **286,** 1695–1699.

West, M. (1981). Robust sequential approximate Bayesian estimation. *Journal of the Royal Statistical Society, Series B* **43,** 157–166.

Received February 1982; *revised August* 1982

Appendix 1

From (2.1), (2.2) and (2.7), $s_t = \phi_{0_t}^{-1}(a_t + u_t)$, where a_t has an approximate $N(0, c^2\phi_{0_t}^2)$ distribution with $c \doteq 0.1$, and $|u_t| < 5$. Actual serum creatinine values are such that almost all values lie in the range $100 < \phi_{0_t} < 1000$, so that $\phi_{0_t}^{-1} < k$, where k is certainly less than 0.1 and, typically, is closer to 0.01. It follows that $|s_t| < |\phi_{0_t}^{-1} a_t| + 5k$, where $\phi_{0_t}^{-1} a_t \sim N(0, c^2)$, so that, with high probability, $|s_t| < 5(c + k)$. For nearly all values of t, we therefore have $|s_t| \ll 1$.

Appendix 2

Given the form $v_t = -s_t\mu_t + (1 - s_t)\beta r_t$, it follows that v_t is symmetrically distributed about 0, with variance, for a given value of μ_t, equal to

$$\mu_t^4\omega_t^2 E(u_t^2) + \mu_t^2 c^2 + \beta^2 E(r_t^2)\{1 + \omega_t^2\mu_t^2 E(u_t^2) + c^2\}.$$

We have assumed that $|r_t| \leqslant 1/16$ and, if we further assume that r_t is approximately normally distributed, it follows that $E(r_t^2) \doteq 4 \times 10^{-4}$. We almost always have $.001 < \mu_t < .01$, so that $\mu_t^2 < 10^{-4}$, and, typically, β^2 is similarly bounded. It can therefore be seen, since $E(u_t^2) \doteq 10$, that the first two terms in the variance dominate and that, if $\mu_t^2 < 10^{-4}$, the variance is well approximated by $c^2\mu_t^2$.