MIKE WEST

University of Warwick

The authors are to be congratulated on a succinct mathematical development of time series influence functionals that generalises and usefully extends existing concepts and techniques of classical robustness. The definition directly caters to the types of outliers specific to time series and reflects the authors' experience with practical modelling and analysis using robust techniques. Central to the paper is the general replacement model for linear (gaussian) series which, to my mind, is the most interesting contribution. There are many ways of representing the various outlier types associated with time series, all closely related to this general model. My own preference is for the simple, state space type of model, which allows for the various outliers hierarchically. Such a model is

$$\begin{aligned} Y_t &= Z_t + \nu_t, \\ Z_t &= X_t + \delta_t, \qquad (1) \\ X_t &= G_t(\mathbf{X}_{t-1}) + \omega_t, \end{aligned}$$

where Y_t is the observed series; ν_t is a zero mean observational noise process, typically comprising independent errors; X_t is the core process defined as a function of $\mathbf{X}_{t-1} = \{X_{t-1}, X_{t-2}, \dots\}$ and the process evolution noise ω_t; and δ_t represents a superimposed contamination process. Additive, or purely observational, outliers are modelled by large ν_t and changes in the core process X_t are modelled by large ω_t. The δ_t process introduces patchy outliers that may be viewed as purely stochastic or related to independent variables via regression or transfer function effects.

Models such as (1), and more complex versions of them, have been used extensively by Bayesian forecasters in applications where protection against additive outliers and adaptation to changes in the Z_t process, via δ_t or ω_t, are of importance. The authors are, of course, familiar with this approach but, in their current paper, part company with Bayesian forecasters in important ways. In my opinion the techniques proposed are most appropriate for fast processing of long series of observations with short sampling intervals when a sustained, stationary core process is evident. Such applications may arise commonly in the engineering fields with which the authors are familiar. In such areas, where interest centres on the estimation of the stable core process and large amounts of data are available, the dominating *data smoothing* feature of robust techniques is very relevant. Otherwise, the *objectives* of the time series modelling activity must be more carefully considered. Models such as (1) are directly geared to the specific operational requirements of sequential *forecasting* that is a primary goal for Bayesian modellers. Here the ability to detect and distinguish the various outlier types and adapt forecasts appropriately is paramount. Omnibus robust methods would tend to oversmooth and hence obscure the *local* behavior that is so relevant in short-term forecasting. Simple sequential techniques for outlier detection, intervention and adaption to change are described in West (1986) and applied in dynamic Bayesian forecasting by West and Harrison (1986). The occurrence of patchy outliers, and possible explanation using independent variables hitherto omitted from the model, is of great interest in improving forecasts, with the emphasis on identifying and estimating the δ_t process and its development into the future.

A related point of contention is the assumption of a stable core process defined by constant parameters whose estimation is the primary goal. A key underlying principle in dynamic Bayesian modelling is the rejection of stationarity in general and the associated allowance for parametric changes over time. Unlike the above mentioned engineering application areas, business and economic series, typically rather short in length, exhibit only *local* stability, global nonstationary, with both sustained, steady, small changes and more marked, abrupt changes in defining parameter values. Thus the primary goal of the author's robust estimation techniques would appear to be limited in scope for application. Can it be adapted to allow for dynamic parameters changing over time? This would be particularly important if independent variables were to be incorporated. Change over time of regression coefficients is not only expected as a general, steady dynamic, but also vital in allowing for the unpredictable effects of further related variables not recognised as being of importance.

The authors may be interested in considering extensions of their techniques, and their outlier models, to nonstandard problems such as those arising with non-gaussian processes. Outlier models apart, there are important questions raised as soon as the non-gaussian nature of time series is admitted. Suppose for example, that the series is discrete, or simply binary. Binary series arise both naturally and by construction in many areas. Particular examples, quite common in applications where data rates are high, and data reduction necessary, concern series derived from underlying, continuous processes via *clipping* operations. Specifically, if Y_t is such a basic process, a binary series S_t is derived by clipping Y_t at level A if S_t has the representation

$$S_t = \begin{cases} 1, & \text{if } Y_t \geq A, \\ 0, & \text{if } Y_t < A. \end{cases}$$

Clearly the theoretical characteristics of the S_t series may be derived from any suitable continuous time series model for Y_t. In the outlier modelling framework, models such as (1) should produce interesting contaminated binary series.

My own approach to practical modelling for non-gaussian series has, however, been somewhat different, being based on the development of dependence models for the S_t series directly. In the binary case, the family of dynamic generalised linear (and nonlinear) models introduced in West, Harrison and Migon (1985) provides a rich class of process structures currently under study. As a simple example, a first order autoregressive type of model for S_t, that parallels the standard linear, gaussian state space model, is given by taking

$$P(S_t = 1|\pi_t) = \pi_t \qquad (0 < \pi_t < 1),$$

where, setting $Z_t = \log(\pi_t/(1 - \pi_t))$, then

$$Z_t = \phi Z_{t-1} + \omega_t \qquad (2)$$

for some (gaussian?) noise sequence ω_t. Thus the unobservable AR process Z_t, provides the "success" probability for S_t after transformation. Generalisations to higher order processes, transforms other than the log-odds, and time-varying parameters (ϕ_t rather than simply ϕ), are evident. The outlier model (1) can now be extended to this binary series by a minor extension of (2) to

$$Z_t = X_t + \delta_t,$$
$$X_t = \phi X_{t-1} + \omega_t,$$

incorporating changes via ω_t series, patchy outliers, and, now, observational outliers through appropriate models for the δ_t series. The only point of significant difference between this model and (1) is that the sampling model is now Bernoulli, rather than gaussian, which leads to a slightly different view of the way in which observational outliers are generated. A closely related, but structurally quite different, class of models for binary series provides for dynamic evolution of transition probabilities in Markov chains. The first order case, for example, has a basic model for $P(S_t = 1|\pi_t)$ as above, but, instead of the continuous process model for the log-odds probability Z_t in (2), a discrete version

$$Z_t = \theta_t + \phi_t S_{t-1} + \omega_t, \tag{3}$$

where θ_t and ϕ_t are time-varying process parameters and ω_t, as usual, process evolution noise. Concerning outlier models, a basic problem arises with (3) in that the observations S_t are fed back into the process model, so a little more thought is required in modelling pure observational outliers. Perhaps the authors have some comments on such problems.

REFERENCES

WEST, M. (1986). Bayesian model monitoring. *J. Roy. Statist. Soc. Ser. B* **48** 70–78.

WEST, M. and HARRISON, P. J. (1986). Monitoring and adaptation in Bayesian forecasting models. To appear in *J. Amer. Statist. Assoc.*

WEST, M., HARRISON, P. J. and MIGON, H. S. (1985). Dynamic generalized linear models and Bayesian forecasting (with discussion). *J. Amer. Statist. Assoc.* **80** 73–97.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WARWICK
COVENTRY CV4 7AL
ENGLAND

REJOINDER

R. DOUGLAS MARTIN AND VICTOR J. YOHAI

University of Washington and University of Buenos Aires and CEMA, Buenos Aires

The discussants have provided us with more than ample food for thought concerning a myriad of issues related to our work on influence functionals for time series. Leading issues include the following: (1) Relationships and dif-