JONES, R. H. (1984). Fitting multivariate models to unequally spaced data. In *Time Series Analysis of Irregularly Observed Data. Lecture Notes in Statist.* (E. Parzen, ed.) **25** 158–188. Springer, New York.

SHUMWAY, R. H. (1984). Some applications of the EM algorithm to analyzing incomplete time series data. In *Time Series Analysis of Irregularly Observed Data. Lecture Notes in Statist.* (E. Parzen, ed.) **25** 290–324. Springer, New York.

TONG, H. (1977). Some comments on the Canadian lynx data. *J. Roy. Statist. Soc. Ser. A* **140** 432–436.

TONG, H. (1983). *Threshold Models in Non-linear Time Series Analysis. Lecture Notes in Statist.* **21**. Springer, New York.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

J. FRANKE AND E. J. HANNAN

University of Frankfurt and Australian National University

This paper by Martin and Yohai will stimulate much future research. The authors are to be congratulated for that and for the presentation of the paper, which stresses statistical intuition and avoids technical detail, where possible.

The central point of their approach to the generalisation of the concept of influence function to a time series setting is the explicit dependence of that function on the arc along which the measure of the observed process approaches that of the nominal process. This emphasis on a specific model for the contamination is necessary because of the great range of possibilities for contamination in a time series setting. However, one can, consequently, ask how strongly the conclusions with respect to robustness and relative performance drawn from this influence function depend on the contamination model. The model (2.2) is very general but the major part of the paper and the examples in Section 5, in particular, deal only with z_i^γ given by (2.4). Consider, for example

$$(1) \qquad y_i = x_i + \eta_i, \qquad \eta_i = \sum_0^{k-1} \beta_j \varepsilon_{i-j},$$

where the ε_i are i.i.d. with distribution $(1-p)\delta_0 + pH$. Here the contamination is generated by impulses which excite a linear system whose effect is imposed on the nominal process. The model (1) is included in the general model (2.2) and both (1) and (2.4) could generate similar patterns of outlier patches so that it would be difficult to distinguish between them from the data. Of course it can be hoped that conclusions from the influence function, based on (2.4), will not differ substantially from those that would have been derived via (1), for example, since essential aspects of the influence function, such as gross error sensitivities, are essentially qualitative in nature and small numerical differences will be of no consequence. However, basically different types of outlier, e.g., isolated outliers compared to those occurring in patches, appear to lead to large differences, as

Figures 1 and 2 of Martin and Yohai show, and this is an agreeable feature of their approach.

The main example of the paper, for practice, is the robust estimation of ARMA parameters. In considering various estimation procedures via the influence function it is necessary to consider the purpose of the analysis. Three, not entirely disjoint, purposes appear to have been as follows. (1) The use of an ARMA model to approximate to the structure of a series so as to obtain understanding. (2) The use of an ARMA model to construct a spectral estimate. (3) The use of an ARMA model for prediction and control. It is not obvious that estimates of ARMA parameters that are robust lead to spectral estimates with good properties. A spectral density influence function is needed to judge this. The same kind of thing can be said of control and prediction and again it would be interesting to transfer the concept of influence function to the control setting. It seems that the main purpose of influence function analysis must be the first of the three listed above.

If one fits an ARMA model with this first objective in mind one has to be aware of problems of misspecification, which have been only recently examined. (See, for example, Hannan (1982) and Shibata (1980).) There is a fair understanding of what happens when the model order is chosen too small or too large or, more realistically, the truth does not lie in any of the parametric sets examined, even when the order is left to be determined from the data. For robust ARMA procedures one has to face the question whether the influence function, derived on the basis of a given parametric model being specified, can be usefully interpreted in the realistic setting. Again the extension of the idea of the influence function to a more general setting, where the ARMA model is no more than an approximation to the truth, is desirable.

Finally we discuss Section 8.2. Examples of the effects of a few isolated outliers on spectral estimation have been given in Bloomfield (1976, Section 5.3), Kleiner, Martin, and Thomson (1975), and Tukey (1984, Chapter 29). Having in mind the intuitive understanding these give it seems unsatisfactory that the spectral density influence curve $\mathrm{IC}_s(\xi)$, derived on the basis of an additive outlier model, does not depend on frequency. Small aberrations at high frequencies with low power may be as important as larger effects at frequencies with large power. The performance of the spectral estimator might better be measured in terms of relative mean square error. See for example Priestley (1981, Section 7.2). One might better examine the influence function of $\log S_n(f)$ as an estimate of $\log S(f)$.

REFERENCES

BLOOMFIELD, P. (1976). *Fourier Analysis of Time Series: An Introduction*. Wiley, New York.

HANNAN, E. J. (1982). Testing for autocorrelation and Akaike's criterion. In *Essays of Statistical Science* (J. Gani and E. J. Hannan, eds.) 403–412. Applied Probability Trust, Sheffield.

KLEINER, B., MARTIN, R. D. and THOMSON, D. J. (1979). Robust estimation of power spectra. *J. Roy. Statist. Soc. Ser. B* **41** 313–351.

PRIESTLEY, M. B. (1981). *Spectral Analysis and Time Series*. Academic, New York.

SHIBATA, R. (1980). Asymptotically efficient selection of the order of the model for estimating parameters of a linear process. *Ann. Statist.* **8** 147–164.

TUKEY, J. W. (1984). *The Collected Notes of John Tukey* **2** (D. R. Brillinger, ed.). Wadsworth, Monterey, Calif.

UNIVERSITY OF FRANKFURT
DEPARTMENT OF MATHEMATICS
JOHANN WOLFGANG GOETHE UNIVERSITY
P.O. BOX 111319, 6000 FRANKFURT
WEST GERMANY

AUSTRALIAN NATIONAL UNIVERSITY
DEPARTMENT OF STATISTICS
MATHEMATICS BUILDING, I.A.S.
AUSTRALIAN NATIONAL UNIVERSITY
G.P.O. BOX 4, CANBERRA, 2601
AUSTRALIA

HANS R. KÜNSCH

ETH, Zurich

Martin and Yohai provide an interesting study on the effect of atypical observations on the behavior of estimators in time series. The influence functional given by Definition 4.2 is the infinitesimal asymptotic bias in a one-parameter family of contaminations of a given model. The bias was also the starting point of my own paper (1984, cf. Section 1.2), but I treated only a smaller class of estimators and I focused on different aspects. So let me explain the differences between the two approaches and discuss their advantages and disadvantages.

Heuristically the connection between ICH and IF is as follows. ICH is the derivative in all directions, i.e., the gradient of $\mathbf{T}$. Hence by the chain rule of differential calculus one gets, formally,

$$\mathrm{IF} = \frac{d}{d\gamma}\mathbf{T}(\mu_y^\gamma) = < \text{grad } \mathbf{T}, \frac{d}{d\gamma}\mu_y^\gamma > = \int \mathrm{ICH}(\mathbf{y}_1)\frac{d}{d\gamma}\mu_y(d\mathbf{y}_1).$$

If $\mathbf{T}$ depends only on the m-dimensional marginal, we can find $(d/d\gamma)\mu_y^\gamma$ in the model (2.4) by the following argument. Ignoring terms of order $o(\gamma)$, there is at most one block of outliers intersecting with $(1, 0, \dots, 2 - m)$, and the initial point of this block is distributed uniformly over $(1, 0, \dots, 3 - m - k)$. To me, the most important theoretical contribution of Martin and Yohai is Theorem 4.2 where they show that the same result also holds for $m = \infty$, at least if $\tilde{\psi}$ depends only weakly on values far away. Since the uniform distribution on all integers is not finite, a bounded $\tilde{\psi}$ is not sufficient for the boundedness of $(d/d\gamma)\mathbf{T}(\mu_y^\gamma)$.

Some of the arguments in the proof of Theorem 4.2 involve the specific contamination model while others are valid more generally. Since the latter may be useful in other situations, I propose to split it in the following way.

THEOREM 4.1′. *Let* $\mathbf{T}$ *be a* $\tilde{\psi}$ *estimate with* $\mathbf{t}_0 = \mathbf{T}(\mu_x)$ *and put* $\mathbf{m}(\gamma, \mathbf{t}) = E[\tilde{\psi}(\mathbf{y}_1^\gamma, \mathbf{t})]$. *If*

(a′) $\mathbf{T}(\mu_y^\gamma) - \mathbf{t}_0 = O(\gamma)$,

(b′) $\mathbf{m}(0, \mathbf{t})$ *is differentiable at* $\mathbf{t} = \mathbf{t}_0$ *and the derivative* C *is nonsingular*,

(c′) $\mathbf{b}(\mathbf{t}) = \lim(\mathbf{m}(\gamma, \mathbf{t}) - \mathbf{m}(0, \mathbf{t}))/\gamma$ *exists and the convergence is uniform for* $|\mathbf{t} - \mathbf{t}_0| \le \varepsilon_0$,