Comment

David R. Brillinger

Professor Rao formulates the problem of the prediction of future values Y_t, $t = t_{p+1}, t_{p+2}, \cdots$ of an individual's measurements, given values at "times" $t = t_1, t_2, \cdots, t_p$. He proceeds, in a way he pioneered many years earlier, by assuming that there exist several $\{Y_t\}$, each a realization of some stochastic process. These processes have some parameters in common and some particular to the individual. In his examples, the parameters are linear, the measurements are at common times and estimation is by the method of moments, generally.

I would like to describe a problem of some practical importance and to show, by presenting empirical results, that through the availability of modern computing and numerical tools one can handle nonlinear forms and irregular time points in a direct likelihood-based manner. The results obtained will be viewed by some as nonstandard, but they are intuitively plausible.

One of the important problems in seismology and earthquake engineering is to obtain an expression for the maximum earth motion occurring at a specified location in the course of a large nearby earthquake. This information is important for, among other things, the choosing of sites for critical facilities and for the understanding of damage that occurred in the course of a particular earthquake. In Joyner and Boore (1981) one can find a list of the maximum accelerations recorded at available seismometer locations for some 23 large earthquakes that occurred mainly in the western United States over a time period dating back to 1940. The principal data may be denoted A_{ij}, M_i and d_{ij} with i indexing event, with j indexing measurement within event, with A_{ij} maximum acceleration, with M_i earthquake magnitude and with d_{ij} horizontal distance of the jth seismometer recording the ith event from the epicenter of that event.

Joyner and Boore (1981) proposed an attenuation law of the form

$$\log A = \alpha + \beta M - \log(\sqrt{d^2 + \delta^2}) + \gamma\sqrt{d^2 + \delta^2}$$

with α, β, γ, δ unknowns. Here d, distance from the epicenter, plays the role of t, the time parameter of growth curves. This law was set down employing physical reasoning to an extent. The parameter δ represents depth of the event in an average sense.

For some events there is only one observation, so it is not possible to reasonably estimate their individual parameters from their individual data. Further one event has 38 observations and so one has to be concerned that its "peculiarities" do not dominate the coefficients determined.

One approach to the problem is to seek to borrow strength in estimating the coefficients of one earthquake from the data available for others. "Borrowing strength" is a term introduced in Tukey (1961) for the class of statistical procedures that seek to improve on naive estimates by incorporating data from parallel but formally distinct circumstances. One way to borrow strength formally is to introduce a random effects model and to proceed in what some call an empirical Bayes fashion. In the attenuation law situation, the earthquakes at hand can be viewed as representatives of a population of earthquakes. One can thus set down the model

$$\log A_{ij} = \alpha_i + \beta_i M_i - \log(\sqrt{d_{ij}^2 + \delta_i^2}) - \gamma_i\sqrt{d_{ij}^2 + \delta_i^2} + \varepsilon_{ij} \tag{1}$$

with $\alpha_i, \beta_i, \gamma_i, \delta_i, i = 1, \cdots, I$, independent realizations of random variables with means μ_α, μ_β, μ_γ, μ_δ and variances σ_α^2, σ_β^2, σ_γ^2, σ_δ^2, respectively. The ε_{ij} are independent mean 0 variance σ^2 noises. The connection of the model (1) with Professor Rao's should be apparent. Some more details of the problem may be found in Brillinger (1987).

If one further assumes that the variates appearing are independent and normal, then one can set down the likelihood function. The exact likelihood involves integrals over the variates common to events. In some cases such integrals can be evaluated exactly. Professor Rao's cases are examples and so are those of Dempster, Rubin and Tsutakawa (1981). In the results to be presented, because of the nonlinearities in the parameters, the integrations were carried out numerically (employing 9-point Gauss-Hermite integration and a Sun workstation), and approximate maximum likelihood estimates evaluated. Before presenting the results obtained we remark that the ideas of borrowing strength and "improving estimates" have been around since the early part of this century (see Berger, 1985, page 168, for example). We note also that the idea of numerically integrating out variables appearing in

David R. Brillinger is Professor, Department of Statistics, University of California, Berkeley, California 94720.

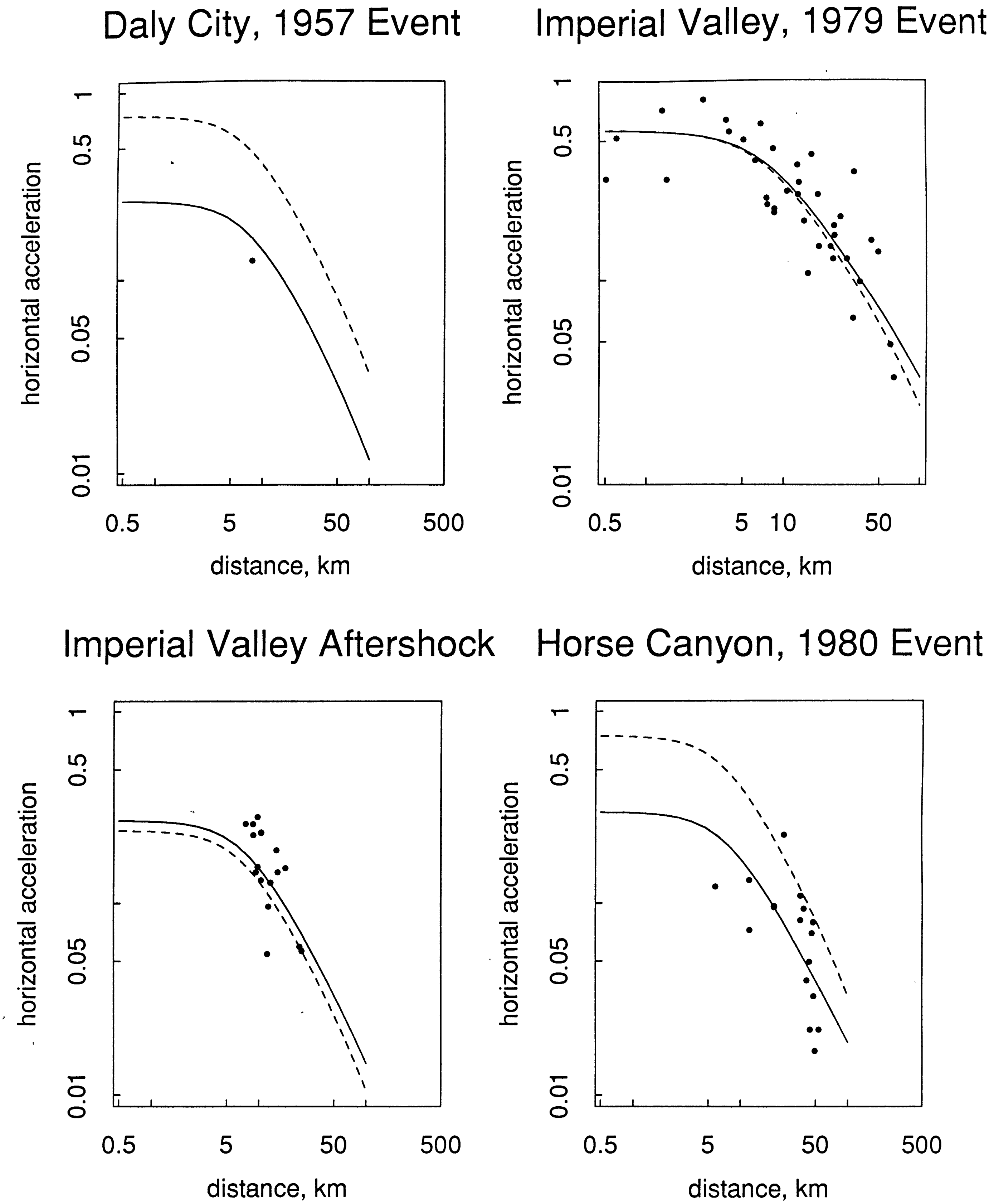

FIG. 1. *For the four indicated earthquakes, observed maximum accelerations are plotted at corresponding differences from the event's epicenter. The solid curve is the "borrowed strength" estimate. The dashed one is that of Joyner and Boore* (1981).

likelihood functions was implemented as early as Bock and Lieberman (1970).

As an estimate of the attenuation curve for the ith event one now has

$$E\{\alpha_i + \beta_i M_i - \log(\sqrt{d^2 + \delta_i^2}) - \gamma_i \sqrt{d^2 + \delta_i^2} + \varepsilon_i \mid \text{all the data}\} \tag{2}$$

substituting the maximum likelihood estimates for the unknown parameters. Figure 1 presents the results of evaluating this for the data of Joyner and Boore (1981) and four representative events. The solid curves provide the evaluations of expression (2) for the events indicated. The dashed curves provide the fit obtained in Joyner and Boore (1981). (These researchers employed a two-stage nonlinear regression procedure: at the first stage, magnitudes were not included, just dummy variables for events; at the second stage, the dummy variables values were regressed on magnitude.) In each case, the random effects fit has moved nearer to (been shrunk toward) the central mass of the data. The first event had but one observation, so borrowing strength has been crucial. The second event is the one with the largest amount of data. The curves are not too different in this case. The third event had a fair amount of data, but it is all at about the same distance from the epicenter and the importance of employing a functional form is clear. If one uses the Joyner-Boore formula then all events with the same magnitude have the same curve. The final graph shows that even with a fair number of data points, the curve can be shifted a good distance by the borrowing operation. Of course, in the application of these techniques one has to be concerned about their robustness to nonnormality and their resistance to outliers. Quek (1987) has done some work on these problems.

Turning specifically to Professor Rao's paper, the formulas presented provided yet another example of the great skills he has at determining exact, pretty results in complicated situations. My one concern is with the "poor" showing of the empirical Bayes procedure in his examples. I wonder if it doesn't result, at least in part, from the use of unbiasedness to obtain a divisor in expression (4.4.6). (In the example I present, maximum likelihood estimates were employed.) I wonder about perhaps improving the performance of the empirical Bayes procedure by determining values to minimize an expression like

$$\sum_i (\| U_i - X\beta_i \|^2 + g(\beta_i - \gamma)' \Gamma^{-1} (\beta_i - \gamma))$$

with g chosen by cross-validation in the manner of Wahba (1977), for example.

Throughout my career I have learned from and enjoyed Professor Rao's papers. This present paper is no exception.

ADDITIONAL REFERENCES

BERGER, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis*, 2nd ed. Springer, New York.

BOCK, R. D. and LIEBERMAN, M. (1970). Fitting a response model for n dichotomously scored items. *Psychometrika* **35** 179–197.

BRILLINGER, D. R. (1987). Some examples of the statistical analysis of seismological data. *Proc. Centennial Anniversary Symp. Seismographic Stations Univ. California.* Univ. California Press, Berkeley.

DEMPSTER, A. P., RUBIN, D. B. and TSUTAKAWA, R. K. (1981). Estimation in covariance component models. *J. Amer. Statist. Assoc.* **76** 341–353.

JOYNER, W. B. and BOORE, D. M. (1981). Peak horizontal acceleration and velocity from strong-motion records including records from the 1979 Imperial Valley, California, earthquake. *Bull. Seismol. Soc. Amer.* **71** 2011–2038.

QUEK, S. A. (1987). The mixed effects model: Fitting and validation. Ph.D. thesis, Univ. California, Berkeley.

TUKEY, J. W. (1961). Data analysis and behavioral science or learning to bear the quantitative man's burden by shunning badmandments. Paper 8 in *The Collected Works of John W. Tukey* (L. V. Jones, ed.) **3** (1986) 187–389. Wadsworth, Monterey, Calif.

WAHBA, G. (1977). Practical approximate solutions to linear operator equations when the data are noisy. *SIAM J. Numer. Anal.* **14** 651–667.