Comment

James Pickands III

Professor Smith's study of the ozone data collected in Houston, Texas is quite unique. There has never been, as far as we know, a case study like this, using extreme value methods. It is extremely useful for what it shows about the pollution situation to which it applies and for the favorable light it casts upon the applicability of extreme value methods. Smith's study begins with a thorough and informative review of methods used heretofore. The data analysis and theoretical discussion which follow point the way to similar analyses of other sets of data and, parallel to this, studies with theoretical variations in the models.

The model (4.1) is an "anova-type" model which suggests a number of possible hypothesis tests, including for example "no trend". The danger is that this may find the right answer to the wrong question. The right question, which may have a (slightly) different answer, is whether the trend is sufficiently small that it is likely to be closer to 0 than to an estimator for it. Methods for making such "parameter selection" decisons are discussed throughout the books by Linhart and Zucchini (1986) and by Sakamoto, Ishiguro and Kitagawa (1986). In Section 5 of Professor Smith's paper there is a very interesting comparison of four models. For each the negative log-likelihood $-\ln L$ is given with the number of parameters M. Unfortunately, if models are nested increasing the number of parameters can only decrease $-\ln L$. What is needed is $M - \ln L$. The rationale for this is given by Sakamoto, Ishiguro and Kitagawa (1986). Calculations are easy. Apparently for both full data and split data it doesn't make much difference whether separate k or fixed $k = .25$ is used. This means that the latter should be used, in the interest of parsimony. Other possibilities could be considered. It is not completely clear to me that the scale and shape could not be fit with a couple of trigonometric terms. Another approach to parameter selection would be to estimate the parameters, for the given model, using the first half of the data and use these estimators to evaluate $-\ln L$ for the second half. In comparing models smaller is better. This method is crude but effective. It also makes it possible to evaluate and compare other factors such as the length of time needed to define cluster separation. The fact that these don't make much difference indicates that a relatively small interval will do.

It would be interesting to model the upcrossing process itself. We would have a process which would be max-stable. Max-ARMA processes would be too restrictive because of their sample path properties. As a home and basement owner, I believe that the maximum intensity of rain is not the only thing that is important. Another is the duration. For short-range forecasting, this kind of consideration is important.

Perhaps a bit further down the pike some consideration could be given to Bayes estimation (with uniform prior(s)). This would, as a practical matter, follow parameter selection. Such computation is significantly more intensive than maximum likelihood computation but as technology is advancing it is becoming less prohibitive. A bit further along will be multivariate analysis. We would consider those data points, and only those, for which at least one component exceeds the threshold for that component.

A long walk begins with a single step. Several have been taken, here. Professor Smith's study constitutes a superb and extremely fruitful and promising beginning.

James Pickands III is Professor, Department of Statistics, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6302.

ADDITIONAL REFERENCES

LINHART, H. and ZUCCHINI, W. (1986). *Model Selection.* Wiley, New York.

SAKAMOTO, Y., ISHIGURO, M. and KITAGAWA, G. (1986). *Akaike Information Statistics.* Reidel, Dordrecht.