Comment

Harry Joe

Professor Smith's article is a timely paper since environmental issues are very much in the news these days. Extreme value inference is important for environmental time series because regulations are generally based on the allowable number of exceedances above high thresholds within certain time periods. Smith has demonstrated very well the application of theoretical results from point processes and extreme value theory for statistical inference. In my discussion, I will elaborate on a few things in the paper.

First I comment on the exploratory data analysis and data reduction. Thanks to Dr. David Fairley at the Bay Area Air Quality Management District, I have some hourly ozone data for many stations in the San Francisco area. There is a strong diurnal pattern in the hourly average concentrations, with the large values in the afternoons (this is reported in Cleveland, Kleiner and Warner, 1976; and Davison and Hemphill, 1987). In this case of a strong diurnal pattern, one might reduce the data to daily maxima of hourly averages. There would in general be serial dependence for the daily maxima, so that there could be runs of several consecutive days where the daily maxima exceed a high threshold. This reasoning suggests that a cluster interval of 72 hours (or more) is better than a cluster interval of 24 hours. Also from the diurnal pattern, one can argue that some of the missing values would not exceed the threshold used for deciding peaks of clusters so that the p_{ij} used in Section 4 could be bigger than the actual observation period. Davison and Hemphill (1987) mention that it is rare to have an exceedance of over 8 parts per hundred million between 9 pm and 9 am.

As mentioned in Section 3 , the approach of this paper avoids the difficult modeling of the times series. For the ozone data, the modeling of the daily maxima of hourly averages may be easier than the modeling of the hourly average time series. In fact, Hirtzel and Quon (1981) perform autocorrelation analyses on both series over summer months and discover that correlation persists at large time lags. If one wanted to make inferences about the average cluster size above a threshold as well as the frequency of exceedance above the threshold, then the modeling of the time series may be more necessary; I will be interested in these inferences and others for some time series for personal exposures to a pollutant in a microenvironment (cf. Duan, 1982). As Smith mentions, a simple model would be a decomposition into a seasonal component and a stationary series. I am thinking of modeling the stationary series as a (first order) Markov chain; some probabilistic theory for these stationary series is included in O'Brien (1987) and Rootzén (1988). This simple class of models is enough to allow for an arbitrary marginal distribution and various degrees of clustering above high thresholds. Starting with a bivariate distribution function G with survival function $\bar{G}$, density g, and identical marginal distributions F and marginal densities f, a Markov sequence with transition probability kernel $h(x_t \mid x_{t-1}) = g(x_{t-1}, x_t)/f(x_{t-1})$ exists. A parametric family of G leads to a parametric family of kernels h. Let $\bar{F} = 1 - F$. In some simple situations, $\bar{G}(x, x)/\bar{F}(x) \sim c(\bar{F}(x))^{\alpha}$ as x approaches the upper support point of F, where $0 \le c \le 1$ and $\alpha \ge 0$. Clustering above high thresholds will depend on how close α is to zero and how close c is to 1. Using some results in O'Brien (1987), if $\alpha = 0$, then the extremal index in (3.8) is at most $1 - c$.

A special case of the Markov (order one) sequences is with G bivariate normal for which an AR(1) sequence is obtained. However, for making inferences for extremes, an assumption of normal tails for the marginal distribution F is too strong and clustering above high thresholds does not occur for ARMA models. From extreme value theory, a weaker assumption is that the tail of F is approximately generalized Pareto (this requires that F is in the domain of attraction of an extreme value distribution). Hence classical time series methods are not always usable for extreme value inferences. This is an important point of the paper.

Next I comment on the likelihood in Section 4. Note that the likelihood with $k_j = 0$ and $\mu_{ij} = \alpha_j$ for all j has a closed form maximum likelihood estimate. In this case, the log-likelihood (log of (4.2)) becomes

$$\sum_{i,j} \{-p_{ij}\exp[\alpha_j/\sigma_j] - N_{ij}\log \sigma_j - N_{ij}(\bar{y}_{ij+} - \alpha_j)/\sigma_j\}$$

$$= \sum_j \{-p_{+j}\exp[\alpha_j/\sigma_j] - N_{+j}\log \sigma_j - N_{+j}(\bar{y}_{+j+} - \alpha_j)/\sigma_j\},$$

where a subscript of + means that a subscript has been added over and the bar over y denotes a mean.

Harry Joe is Associate Professor, Department of Statistics, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5.

It is straightforward to show that

$$(1) \qquad \hat{\sigma}_j = \bar{y}_{+j+}, \quad \hat{\alpha}_j = \bar{y}_{+j+}\log[N_{+j}/p_{+j}].$$

This gives some idea about the magnitudes of σ and α and provides an alternative initial estimate for the likelihood; the method mentioned by Smith based on the Gumbel distribution also comes from taking $k_j = 0$. From (1), $\hat{\alpha}$ decreases as p_{+j} increases. This may happen even if k_j is not fixed at zero, so that one must be careful in how to deal with missing values at hours when the ozone levels are typically low.

Finally I have some comments referring to Sections 4 and 5. The mean exceedance rates are based on a unit of a cluster and so one must take into account the average number of days in a cluster in order to compare with the quantities specified by air quality standards. Chock (1982) raises the question of whether it is reasonable to count an adverse multiday meteorological event several times as having exceeded a threshold.

Smith points out that the analysis needs to be repeated at other sites to get a firm indication of a downward trend in crossing rates at high levels. Walker (1985) reports on ozone trends in California and Texas over a period of 10 years and concludes that there is little evidence that annual average ozone or average peak ozone has been reduced. Walker's analysis is not an extreme value analysis, but he does mention two confounding factors for the ozone trend that are relevant here. These are trends in analytical methodology (for measurements) and data quality assurance. The EPA made ultraviolet photometry the basic calibration procedure for all official ambient ozone monitors in 1979, and data prior to this year are generally adjusted (calibrated) in order to study trends from 1973 on. Was this true of the ozone data in this analysis? Concerning the data quality trend, Walker states that more recently many high values are invalidated as outliers where earlier they were accepted. The methodology in this paper has wide applicability but one must be careful with potential confounding factors in making conclusions.

In conclusion, Professor Smith is to be commended for an excellent paper that develops statistical methodology for an important application and mentions important areas of developing research.

ADDITIONAL REFERENCES

CHOCK, D. P. (1982). On the non-randomness of high-pollution days. *Atmospheric Environment* **16** 2855–2862.

CLEVELAND, W. S., KLEINER, B. and WARNER, J. L. (1976). Robust statistical methods and photochemical air pollution data. *J. Air Pollution Control Assoc.* **26** 36–38.

DUAN, N. (1982). Models for human exposure to air pollution. *Environment International* **8** 305–309.

HIRTZEL, C. S. and QUON, J. E. (1981). Statistical analysis of continuous ozone measurements. *Atmospheric Environment* **15** 1025–1034.

O'BRIEN, G. L. (1987). Extreme values for stationary and Markov sequences. *Ann. Probab.* **15** 281–291.

ROOTZÉN, H. (1988). Maxima and exceedances of stationary Markov chains. *Adv. in Appl. Probab.* **20** 371–390.

WALKER, H. M. (1985). Ten-year ozone trends in California and Texas. *J. Air Pollution Control Assoc.* **35** 903–912.

Comment

Ishay Weissman

I enjoyed Richard Smith's study of ground-level ozone data using extreme value theory. Smith should be commended for undertaking this project, and congratulated for his lucid analysis and exposition.

After describing the data, Smith gives some theoretical background, just enough for the reader who is not an expert in extreme value theory to understand the analysis that follows. The paper as a whole was written in a free-flowing format that makes it interesting and enjoyable to read. The author applied simple descriptive methods (tables, histograms, boxplots, etc.) as well as sophisticated ones (generalized extreme value models, generalized Pareto models with and without trend, etc.). The latter have been developed to a large extent by Smith himself in earlier works.

Due to time pressure, I will only make a few short comments.

1. EXTREME VALUE ASPECTS

I totally agree with Smith's decision to concentrate on high exceedances. Ozone as well as other pollutants become serious health-hazards when they exceed certain levels (thresholds). The current ozone standard, as Smith puts it, permits no more than three exceedances above 12 parts per 100 million in any 3-year period. Hence, looking at high exceedances is only

Ishay Weissman is Associate Professor of Statistics, Faculty of Industrial Engineering and Management, The Technion, Haifa 32000, Israel.