Katz, R. W.; T. Faragó, 1989: Applications of
the extreme values in climatology.
Proc. of 4th Intern. Meeting
on Statist. Climatology, (Rotorua, March 27-31, 1989

Applications of the theory of extreme values in climatology

R. W. Katz
ESIG/NCAR, Boulder, Colorado

T. Farago
Central Meteor. Institute, Budapest, Hungary

Introduction and Summary

The importance of considering extreme climate events in assessing the impacts of climate on society is now widely recognized (e.g., Mearns et al., 1984; Wigley, 1985). Among other things, scenarios of future climate must adequately reflect any possible changes in the likelihood of extreme events. Little attention, however, has been devoted to developing methods that are appropriate for making assessments of the impact of extreme climate events.

Nevertheless, a well-developed statistical theory of extremes does exist (e.g., Leadbetter et al., 1983). Although there have been many attempts to apply it in climatology, this theory has not yet been successfully integrated with realistic applications. Most of the research has either: (i) dwelled on the theory of extremes with only a superficial application to climate (e.g., Tiago de Oliveira, 1986); or (ii) treated realistic climate problems without the proper application of theory (e.g., Tabony, 1983).

We review some of the specific issues that arise in attempting to apply the theory of extreme values to climate, issues for which clear-cut solutions have yet to be identified. First, the classical theory of extreme values is briefly outlined, both for the maximum of a sequence as well as for the number of exceedances of a threshold. Then the issue of the effect of serial correlation of climate time series on the theory of extreme values is considered. In particular, a blatant example of the misapplication of a theory based on means to extremes is identified in the climate literature. Next, the issue of how to take into account the seasonal cycles of climate time series is considered. The various approaches that have been proposed for dealing with this problem each possess certain advantages and disadvantages, making it unclear which one is preferable. Finally, one way in which the application of the theory of extreme values might aid in the resolution of certain questions related to climate impacts is suggested. Specifically, the question of how the likelihood of extreme climate events would change with changes in the average and variability of climate is addressed.

Classical Theory of Extreme Values

In this section, the so-called classical theory of extreme values is briefly reviewed. This theory is concerned with the extreme values of a sequence $X_1, X_2, \cdots X_n$ of independent and identically distributed (with common distribution function F, say) random variables. The distribution of the maximum

$$M_n = \max(X_1, X_2, \cdots, X_n) \tag{1}$$

as the sample size n tends toward infinity is of interest. In addition, the distribution of the number of exceedances of a threshold

$$N_n = \sum_{i=1}^{n} \chi_{\{X_i > c_n\}} \tag{2}$$

where χ is the indicator function, is considered as the threshold c_n and sample size n both become large.

At least from a probabilistic point of view, the classical theory of extreme values has been derived in nearly complete generality (e.g., Leadbetter et al., 1983). This theory is developed in a parallel, if not completely analogous, fashion to that for averages. In place of the Central Limit Theorem, it can be shown that the distribution of the maximum M_n tends toward one of three possible distributions (termed Type I, Type II, and Type III extreme value distributions). Specifically, if

$$\lim_{n\to\infty} \Pr\{a_n(M_n - b_n) \leq x\} = G(x) \tag{3}$$

where $a_n > 0$ and b_n are normalizing constants, then the distribution function G must be one of three possible forms. For instance, the Type I extreme value distribution, most commonly applied in climatology, is given by

$$G(x) = \exp(-e^{-x}) \tag{4}$$

where, $-\infty < x < \infty$.
except for a change in location and scale.

Necessary and sufficient conditions for the three possible forms of distribution function G in (3) to arise depend upon the shape of the right-hand tail of the distribution function F of the individual random variables [i.e., the behavior of $1 - F(x)$ for large x]. The Type I extreme value distribution appears in (3) when F has an infinite right-hand tail [i.e., $F(x) < 1$ for all x] with $1 - F(x)$ decreasing to zero at a sufficiently fast enough rate as x tends to infinity. In particular, nearly all of the distributions (e.g., exponential, gamma, lognormal, normal, Weibull) commonly fit to individual observations of climate variables give rise to the Type I extreme value distribution as the limiting distribution for the maximum.

It can also be shown that the limiting distribution of the number of exceedances N_n is Poisson, if the threshold c_n is increased at such a rate that, as the sample size n increases, the mean number of exceedances of the threshold remains roughly constant. Specifically,

$$\lim_{n\to\infty} \Pr\{N_n = k\} = \frac{e^{-\tau}\tau^k}{k!} \tag{5}$$

where

$$\tau = \lim_{n\to\infty} n[1 - F(c_n)] \tag{6}$$

This Poisson approximation has been applied, for instance, to the exceedance of thresholds by time series of wind speeds (Ross, 1987) and to the occurrence of below freezing temperatures (Waylen, 1988). Moreover, Revfeim and Hessell (1984) represent the sequence of occurrences of extreme events (e.g., wind gusts) as a Poisson process.

Persistence of Climate

Although the classical theory of extremes makes the asssumption of independence, it is well known that climate time series typically possess positive autocorrelation. This question of the effects of serial correlation on extremes has created much confusion among climatologists. In particular, it is apparently commonly believed that the concept of the 'effective number of independent samples' is relevant in accounting for the effects of dependence on extremes (e.g., Tabony, 1983). This claim is a blatant example of the misapplication of a theory based on means to extremes.

In an attempt to clarify this issue, the nature of the extension of the classical theory of extreme values to the case of weak dependence is briefly reviewed. Recall that the Central Limit Theorem for averages still holds for stationary processes with weak dependence, but the normalizing constant representing the standard deviation of a time average must be adjusted. In particular, the persistence of climate time series inflates the variance of a time average relative to the independent case. This adjustment for variance inflation has led to the concept of the 'effective number of independent samples' (e.g., Madden, 1979).

The effects of dependence on extreme values are, however, quite different. For weak dependence, the normalizing constants a_n and b_n in (3) for the limiting distribution of the maximum are exactly the same as in the independent case. Even the rate of convergence to the limiting extreme value distribution has been shown in some cases (e.g., stationary normal process) to be essentially unaffected by dependence. Similarly, the Poisson limit (5) for the number of exceedances of a threshold still holds under weak dependence. In this regard, it has been suggested (e.g., by Buishand, 1986) that the Poisson approximation works better in practice for climate time series if 'clusters' of individual exceedances are combined into a single exceedance event.

The type of stochastic processes usually employed to model climate time series (e.g., autoregressive-moving average process) do possess weak enough dependence that the classical theory of extreme values is still in force. Specifically, when dealing with a stationary normal process, the dependence is weak enough if the autocorrelations $\rho_n = \mathrm{Corr}(X_i, X_{i+n})$ tend to zero at a fast enough rate that

$$\lim_{n\to\infty} \rho_n \log n = 0 \tag{7}$$

An issue that does have significant implications for applications of the theory of extremes concerns whether climate time series actually possess stong dependence (or 'long memory'). Under strong dependence [when, in particular, (7) does not hold in the case of a normal process], other forms of limiting distributions for the maximum than the Type I, II, and III extreme value distributions are possible (Leadbetter et al., 1983). In this regard, the so-called Hurst Phenomenon, which concerns the behavior of the 'rescaled adjusted range' (a rather complex form of extreme value statistic), is exhibited by many time series considered in water resources research (e.g., precipitation, stream flow). One of the possible explanations for this phenomenon is strong dependence (Mandelbrot and Wallis, 1968). Given this combination of theoretical developments and empirical observations, it is surprising that this issue has received so little attention in the climatological literature.

Seasonal Cycles

Although the classical theory of extremes makes the assumption of identical distributions, climate time series naturally contain seasonal cycles. Several approaches have been proposed for dealing with this problem. One technique involves simply standardizing each random variable to remove any seasonal cycles that might be present in the mean and variance; that is, constructing a new sequence of random variables

$$X_i' = \frac{X_i - E(X_i)}{[\mathrm{Var}\ (X_i)]^{\frac{1}{2}}} \quad . \tag{8}$$

The classical theory of extremes can then be applied to the standardized sequence X_1', $X_2', \cdots, X_n'$ (as has been done, for example, in the case of extreme wind speeds by Zwiers, 1987). Although this procedure is theoretically sound, it may leave much to be desired from a practical point of view. The method measures extreme values at different times within the year relative to their respective mean and variance, rather than in absolute terms. Such a measure would not be very meaningful for many climate impact studies in which thresholds for extreme values are taken as fixed quantities.

A second approach consists of applying extreme value theory separately to individual time periods (e.g., months or seasons) and then combining these results to obtain a single distribution for the annual maximum (Carter and Challenor, 1984). If it is assumed that the maxima for the individual time periods are independent, then the combined distribution is simply the product of the individual extreme value distributions. A debate has raged in the climatological literature over whether this method is superior to simply fitting a single extreme value distribution directly to the annual maxima (Tabony, 1983).

The classical theory of extreme values has also been extended to this particular case of nonstationary time series. Expressions showing how the normalizing constants a_n and b_n in (3) become more complex for periodically varying means have been derived (Horowitz, 1980; Leadbetter et al., 1983).

However, it is not clear how realistic some of these theoretical developments are for climatological applications. Ballerini and McCormick (1988), for instance, show that the form of limiting distribution for the maximum in the case of a stationary process with weak dependence and periodically varying mean and variance is determined by the period of maximal standard deviation. They use an example of time series of daily temperature to motivate the problem. However, temperature is generally most variable in the winter, not the season during which the annual maximum temperature ordinarily occurs!

Climate Impacts

As mentioned in Section 1, the need to consider extreme events in climate impact studies is now recognized. Mearns et al. (1984) addressed the question of how the probability of occurrence of an extreme event would change as the overall climate changes (e.g., in terms of means and variances). Extreme high temperature events, such as the maximum temperature on a single day or on a run of consecutive days exceeding a threshold, were considered. A simulation approach was employed to obtain numerical results concerning the relationship between the likelihood of these extreme events and the mean, variance, and autocorrelation of time series of daily maximum temperature. More insight into the nature of these nonlinear relationships could be provided by direct reliance on extreme value theory. Specifically, this approach would enable analytical results to be derived. If progress is to be made on factoring consideration of extreme events into climate impact assessment, then the theory of extreme values needs to be exploited.

Acknowledgements

Research supported in part by U.S. National Science Foundation (U.S.-Eastern Europe Cooperative Science Program) and Hungarian Academy of Sciences.

References

Ballerini, R., and W.P. McCormick, 1988: Extreme value theory for processes with periodic variances (abstract). *Advances in Applied Probability*, **20**, 9.

Buishand, T.A., 1986: Extreme value analysis of climatological data. *Third International Conference on Statistical Climatology*. Austrian Society of Meteorology, Vienna, Austria, pp. 145–158.

Carter, D.J.T., and P.G. Challenor, 1984: Comments on 'Extreme value analysis in meteorology' by R.C. Tabony (with reply by R.C. Tabony). *Meteorological Magazine*, **113**, 43–52.

Horowitz, J., 1980: Extreme values from a nonstationary stochastic process: An application to air quality analysis. *Technometrics*, **22**, 469–478.

Leadbetter, M.R., G. Lindgren, and H. Rootzen, 1983: *Extremes and Related Properties of Random Sequences and Processes*. Springer-Verlag, New York.

Madden, R.A., 1979: A simple approximation for the variance of meteorological time averages. *Journal of Applied Meteorology*, **18**, 703–706.

Mandelbrot, B.B., and J. Wallis, 1968: Noah, Joseph, and operational hydrology. *Water Resources Research*, **4**, 909–918.

Mearns, L.O., R.W. Katz, and S.H. Schneider, 1984: Extreme high-temperature events: Changes in their probabilities with changes in mean temperature. *Journal of Climate and Applied Meteorology*, **23**, 1601–1613.

Revfeim, K.J.A., and J.W.D. Hessell, 1984: More realistic distributions for extreme wind gusts. *Quarterly Journal of the Royal Meteorological Society*, **110**, 505–514.

Ross, W.H., 1987: A peaks-over-threshold analysis of extreme wind speeds. *Canadian Journal of Statistics*, **15**, 328–335.

Tabony, R.C., 1983: Extreme value analysis in meteorology. *Meteorological Magazine*, **112**, 77–98.

Tiago de Oliveira, J., 1986: Extreme values and meteorology. *Theoretical and Applied Climatology*, **37**, 184–193.

Waylen, P.R., 1988: Statistical analysis of freezing temperatures in central and southern Florida. *Journal of Climatology*, **8**, 607–628.

Wigley, T.M.L., 1985: Impact of extreme events. *Nature*, **316**, 106–107.

Zwiers, F.W., 1987: An extreme-value analysis of wind speeds at five Canadian locations. *Canadian Journal of Statistics*, **15**, 317–327.