STATISTICS OF EXTREMES: MODELING ECOLOGICAL DISTURBANCES

RICHARD W. KATZ,[1,3] GRACE S. BRUSH,[2] AND MARC B. PARLANGE[2,4]

[1]*Environmental and Societal Impacts Group, National Center for Atmospheric Research, Boulder, Colorado 80307 USA*
[2]*Department of Geography and Environmental Engineering, Johns Hopkins University, Baltimore, Maryland 21218 USA*

Abstract. The potential advantage of extreme value theory in modeling ecological disturbances is the central theme of this paper. The statistics of extremes have played only a very limited role in ecological modeling, despite the disproportionate influence of unusual disturbances on ecosystems. An overview of this theory is provided, with emphasis on recent developments that both make more efficient use of the available data on extremes and enable applications that are more ecologically realistic. Consistent with the emphasis on scale in ecology, scaling properties of extremes are emphasized. It is argued that the existence of distributions whose extreme upper tail is ''heavy'' (i.e., decreases at a relatively slow rate) implies that ecological disturbances are sometimes regarded as more ''surprising'' than they ought to be.

The application focuses on modeling disturbances in paleoecology. Two examples are considered: the first, a sediment yield time series for Nicolay Lake in the high Arctic, reflects only the influence of hydrologic disturbances; the second, a sediment rate time series in the Chesapeake Bay, includes both climatic and anthropogenic influences. Strong evidence supports a heavy-tailed distribution for the Nicolay Lake sediment yield, but not necessarily for the Chesapeake Bay sediment rates. For each example, it is demonstrated how the statistics of extremes can readily incorporate information about covariates, such as large-scale atmospheric–oceanic circulation patterns or land use.

Key words: generalized extreme value distribution; generalized Pareto distribution; paleoecology; peaks over threshold; return level; return period; sediment accumulation.

INTRODUCTION

It seems that the rivers know the theory. It only remains to convince the engineers of the validity of this analysis.

—Emil J. Gumbel, 1891–1966

In this quote by Emil Gumbel, ''theory'' refers to the statistical theory of extreme values. Gumbel was a pioneer in the application of this theory, particularly in fields such as climatology and hydrology (Gumbel 1958). Now, several decades after this quote appeared, the engineers are indeed convinced of the theory's utility in water resources management, building design, etc. (e.g., Katz et al. 2002). Yet perhaps this quote would remain apt if the word ''engineers'' was replaced with ''ecologists.''

Extreme events, rare but not necessarily unprecedented, play an important role in ecology (Gutschick and BassiriRad 2003). Ecological disturbances are commonly associated with the occurrence of extreme events, such as an excursion of a climate variable like temperature outside of some range (e.g., above a rel-

atively high, or below a relatively low, threshold). Compounding the problem is the specter of global climate change, with anticipated increases in the frequency of extreme events such as hot spells or intense precipitation (Folland and Karl 2001).

Given that their occurrence is by definition unusual, it has been a challenge for statisticians to devise appropriate methods for quantifying the likelihood and intensity of extreme events. Yet ecologists are largely unaware that, akin to the central limit theorem for averages, a specialized statistical theory is now available for extremes (e.g., Coles 2001). This lack of awareness exists despite at least one review article that has appeared in the ecological literature advocating the use of the statistical theory of extremes (Gaines and Denny 1993; see also Denny and Gaines 2000). This theory can model not just the frequency of rare ecological events (e.g., Dixon et al. 2005), but also their severity.

In the present paper, we build upon the review by Gaines and Denny (1993). An ecologically relevant example is provided with a compelling need for the application of extreme value theory; namely, an instance in which the distribution has a heavy tail (i.e., decreases at a relatively slow rate). Recent developments in the application of this theory, either beyond the scope of or too recent to be taken into account in Gaines and Denny (1993), are surveyed. Among other things, annual or diurnal cycles and trends, as well as more physically based variables such as El Niño events,

can be readily incorporated into the theory as covariates. The peaks over threshold (or point process approach), an alternative to block (e.g., annual) maxima, is introduced to exploit more of the information available about the upper tail of the distribution.

The paper begins with a background section covering the origin and early applications of the statistics of extremes, ecological extremes, and the previous use of the statistics of extremes in ecology. An overview of the statistical theory of extreme values with emphasis on recent developments is then provided, followed by the results from a paleoecological application of modeling disturbances using sediment accumulation. The paper concludes with a discussion of the benefits of extreme value theory in modeling ecological disturbances, as well as the unresolved issue of how to model spatial extremes.

BACKGROUND

Historical origins of statistics of extremes

In the 1920s, a number of individuals simultaneously began deriving the statistical theory of extreme values. An early theoretical breakthrough was produced by the British statisticians R. A. Fisher and L. H . C. Tippett, who derived the limiting form of the distribution of the maximum or minimum value in a random sample (Fisher and Tippett 1928). Tippett immediately applied this theory to the strength of cotton yard, a situation in which the ''weakest link'' (i.e., a minimum value) governs failure (Box 1978). This application could be viewed as a precursor to the field of engineering reliability, in which structural failure is modeled statistically. That ecological stresses reflect statistical extremes, not averages or even variances, was aptly described in Gaines and Denny (1993).

In subsequent decades, extreme value theory found application in other areas in which extreme events naturally play an important role. The first book-length treatment on the statistics of extremes covered a number of applications, many related to engineering design (Gumbel 1958). Applications directly relevant to ecology have included environmental variables such as those in climate (e.g., temperature, precipitation, wind speed), hydrology (e.g., stream flow), and oceanography (e.g., sea level, wave height), with several of these variables being included in the examples in Gaines and Denny (1993).

Ecological extremes

From a scientific perspective, the importance of extreme events in ecology is well recognized. Focusing on plants, Gutschick and BassiriRad (2003) developed the thesis that extreme events ''play a disproportionate role in shaping the physiology, ecology and evolution of organisms.'' Despite fire being an integral component of ecosystems, large fire is a graphic example of a disturbance that can disrupt ecosystem-level processes (Moritz 1997). Perhaps the ultimate extreme event results in the extinction of a population, with extinction probabilities receiving much attention in the ecological literature (Ludwig 1996). Another topic with ecological implications is longevity, particularly the variation in life spans among different species of plants and animals (Carey 2003).

Paleoecology focuses on the reconstruction of the long-term history of ecosystems through the analysis of information, such as the pollen and seeds contained in sediment cores (Brush 1989). This approach is particularly effective at uncovering the influence of disturbances on ecosystems, such as those attributable to changes in land use (e.g., deforestation) or to heavy rains. For instance, sediments have been used to reconstruct temporal and spatial histories of fire regimes in ecosystems (Lynch et al. 2003). The closely related field of paleohydrology focuses on the use of fluvial features that reflect the hydrologic cycle, generally extreme hydrologic events such as paleofloods (Gregory and Benito 2003). Extreme value theory is routinely used to estimate flood frequency and intensity on the basis of measurements of precipitation or stream flow, occasionally in combination with other historical information (Stedinger and Cohn 1986). Yet this statistical theory has rarely ever been applied to the corresponding paleohydrologic or paleoecologic data (an exception is Lamoureux [2000]).

Statistics of extremes in ecology

Despite the review article by Gaines and Denny (1993), it remains difficult to find examples of the explicit application of the statistical theory of extreme values to ecology. Instead, assumptions about tail behavior are typically made which are more restrictive and difficult to verify. The conventional approach effectively assumes that the form of distribution, chosen on the basis of a fit dominated by the majority of the observations in the center of the distribution, necessarily fits the upper and lower tails of the distribution satisfactorily as well.

All the examples of the application of the statistics of extremes presented in Gaines and Denny (1993) happen to involve variables (i.e., minimum and maximum sea surface temperature, maximum wind speed, maximum ocean wave force, and maximum human life span) whose upper (or lower) tail is either unbounded, but ''light'' (i.e., decreases at a relatively rapid rate) or bounded (the shape of the tail of a distribution will be more formally characterized in the next section). For variables such as temperature (e.g., Brown and Katz 1995) and wind speed (e.g., Palutikof et al. 1999), this characteristic of a light or bounded upper tail is known to hold more generally. Nearly all the continuous statistical distributions (e.g., normal, exponential, gamma, lognormal) commonly used to model ecological and related variables have light upper tails.

Although the statistics of extremes certainly can be useful in applications involving light or bounded tails, its need is more compelling for variables which instead possess a heavy-tailed distribution (e.g., Katz et al. 2002). One such ecological example concerns fire disturbance, in which the distribution of the largest fire within a region is heavy-tailed (Moritz 1997), work evidently inspired by Gaines and Denny (1993). Similarly, Schoenberg et al. (2003) reviewed the use of a power law (or Pareto) distribution (i.e., a form of heavy-tailed distribution [Arnold 1983]) to fit wildfire size. Rates of population spread have also been described by distributions which are heavy tailed (Clark et al. 2001). Although not necessarily synonymous with damage to ecosystems, the distribution of economic damage from disturbances such as hurricanes can be heavy-tailed as well (Katz 2002*a, b*).

The search for universal scaling laws in ecology typically involves assuming a form of power-law distribution. For example, Ferriere and Cazelles (1999) showed how a power law relationship can arise in population dynamics, and Keitt and Stanley (1998) fitted a power law relationship in analyzing the dynamics of bird populations. The connection between such scaling laws and extreme value theory will be treated in a subsequent section.

STATISTICAL THEORY OF EXTREME VALUES

Overview of theory

We provide only a relatively short overview of the essential features of the statistical theory of extreme values (for more details, see Leadbetter et al. 1983, Coles 2001, Reiss and Thomas 2001). For convenience, extremes are discussed solely in terms of maxima (or upper tails) of distributions. Nevertheless, minima (or lower tails) are effectively encompassed through the relationship

$$\min(x_1, x_2, \ldots, x_T) = -\max(-x_1, -x_2, \ldots, -x_T) \quad (1)$$

for a sample of data, $x_1, x_2, \ldots, x_T$.

A heuristic scaling argument provides some motivation for how the basic extreme value theory arises. It involves thinking about different ways to determine the maximum value in a sample of length $2T$, say data $x_1, x_2, \ldots, x_{2T}$. This highest value could be determined indirectly through separately taking the maximum of the first half and of the second half of the sample and then combining these two maxima; that is,

$$\max(x_1, x_2, \ldots, x_{2T})$$
$$= \max[\max(x_1, x_2, \ldots, x_T), \max(x_{T+1}, x_{T+2}, \ldots, x_{2T})].$$
$$(2)$$

This relationship constrains the possible form of the limiting distribution of the maximum, suitably normalized, of a sequence of random variables, say $X_1, X_2, \ldots, X_T$ (for the moment, assumed to be independent
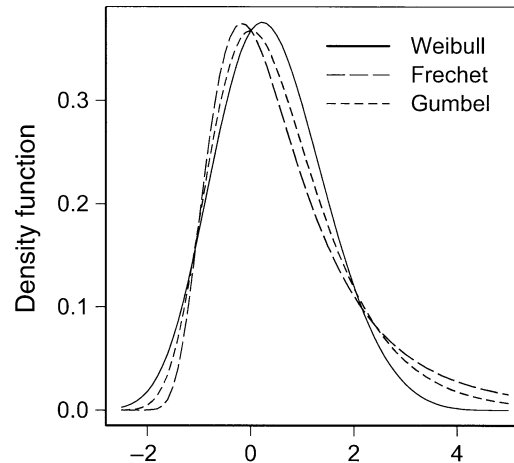

FIG. 1. Plots of the GEV (generalized extreme value) probability density function with $\mu = 0$, $\sigma = 1$, $\xi = -0.2$ (Weibull type), $\xi = 0$ (Gumbel), and $\xi = 0.2$ (Fréchet).

and identically distributed). Such a limiting distribution must satisfy the "max-stability property" (Leadbetter et al. 1983); namely, distributions for which the operation of taking the maximum of a finite sequence of independent and identically distributed random variables leads to an identical distribution, except for differing location and scale parameters (the concepts of "location" and "scale" will be explained shortly). This property is akin to the reason why the normal distribution arises in the central limit theorem; that is, the mean of normally distributed observations has exactly a normal distribution.

The max-stability property gives rise to a form of distribution known as the generalized extreme value (GEV), with cumulative distribution function

$$F(x; \mu, \sigma, \xi) = \begin{cases} \exp\{-[1 + \xi(x - \mu)/\sigma]^{-1/\xi}\}, \\ \quad 1 + \xi(x - \mu)/\sigma > 0 \quad \xi \neq 0 \\ \exp\{-\exp[-(x - \mu)/\sigma]\} \quad \xi = 0. \end{cases}$$
$$(3)$$

Here μ is termed a location, $\sigma > 0$ a scale, and ξ a shape parameter. If the random variable X has a GEV distribution (Eq. 3), then the standardized variable $(X - \mu)/\sigma$ has a distribution that does not depend on either μ or σ, only on ξ. Like the mean and standard deviation of the more familiar normal distribution, the location parameter specifies where the distribution is "centered," the scale parameter its "spread." It should be noted that Gaines and Denny (1993) used a somewhat different, but equivalent parameterization.

The shape of the GEV distribution assumes three possible types (Fig. 1), depending on the value of ξ:

(i) $\xi = 0$, a light-tailed (or Gumbel) distribution;

(ii) $\xi > 0$, a heavy-tailed (or Fréchet) distribution;

(iii) $\xi < 0$, a bounded (or Weibull) distribution.

The type (i) distribution has an unbounded upper tail

which decreases at a relatively rapid (i.e., exponential) rate. Although the type (ii) distribution also has an unbounded upper tail, it decreases at such a slow (i.e., power law) rate that its moments are infinite for all orders greater than $1/\xi$ (e.g., the variance is infinite if $\xi > 0.5$; the mean is infinite if $\xi > 1$). The type (iii) distribution has a finite upper bound at $x = \mu - (\sigma/\xi)$.

It is natural to focus on the extreme upper quantiles of the GEV distribution. Specifically, the ''return level'' associated with a ''return period'' of $1/p$ is the $(1 - p)$th quantile of the GEV distribution (e.g., when modeling annual maxima, $p = 0.01$ would correspond to a 100-yr return period), a concept used extensively in hydrology (e.g., Katz et al. 2002). This quantity can be obtained by inverting the distribution function of the GEV (Eq. 3), giving

$$F^{-1}(1 - p; \mu, \sigma, \xi)$$
$$= \begin{cases} \mu - (\sigma/\xi)\{1 - [-\ln(1 - p)]^{-\xi}\} & \xi \neq 0 \\ \mu - \sigma \ln[-\ln(1 - p)] & \xi = 0 \end{cases} \quad (4)$$

with $0 < p < 1$ (Coles 2001). The return level is linear in μ and σ, but highly nonlinear in ξ.

It is not actually necessary to assume that the observations be independent, with the limiting distribution of the maximum still being the GEV under a wide range of dependence conditions (e.g., for an autoregressive process) and the only effect being on the values of the location and scale parameters, μ and σ (Leadbetter et al. 1983). The assumption of identically distributed observations can also be relaxed, with nonstationarity being introduced through covariates. The parameters of the GEV distribution could be dependent on the time t; say,

$$\mu(t) = \mu_0 + \mu_1 t, \quad \ln[\sigma(t)] = \sigma_0 + \sigma_1 t, \quad \xi(t) = \xi,$$
$$t = 1, \ldots, T \quad (5)$$

the location parameter and the logarithm of the scale parameter (applying the logarithm to impose the constraint of $\sigma > 0$) being linear functions of time and the shape parameter being independent of time. More generally, a covariate z_t (e.g., the state of the El Niño phenomenon at time t) could be used instead of time t in Eq. 5.

As did Gaines and Denny (1993), we estimate the parameters of the GEV distribution by the method of maximum likelihood (see the Supplement). Although other types of parameter estimation techniques are sometimes used in fields such as hydrology, one advantage of the maximum likelihood approach is that covariates (as in Eq. 5) can be readily incorporated (Coles and Dixon 1999). In this way, annual or diurnal cycles and trends need not be removed before fitting the GEV distribution (as did Gaines and Denny [1993]). This approach has the advantage of allowing for cycles or trends in the upper tail not necessarily of the same form as in the center of the distribution (Smith 1989).

Recent developments

The GEV distribution (Eq. 3) is typically fitted to ''block maxima,'' such as the highest daily precipitation amount over an entire year. Although this approach is sometimes viewed as advantageous because it requires only a simplified data summary (Gaines and Denny 1993), it is disadvantageous because it does not make use of all of the information available about the upper tail of the distribution (e.g., the two highest daily precipitation amounts over the entire record might occur in the same year). An alternative approach whose idea originated in hydrology, called ''peaks over threshold'' (POT), attempts to exploit more of this information (Todorovic and Zelenhasic 1970).

We start by envisioning the process by which extremes occur, defining an extreme event in terms of exceeding a relatively high threshold, say $X > u$. Recalling that the Poisson distribution arises as an approximation to the binomial for rare events, it is reasonable to assume that the sequence of times the event occurs is governed by a Poisson process, say with rate parameter $\lambda > 0$. In this case, the number of occurrences in a time interval of length T, N_T say, has a Poisson distribution with mean λT; that is,

$$\Pr\{N_T = k\} = [(\lambda T)^k e^{-\lambda T}]/k!, \quad k = 0, 1, \ldots . \quad (6)$$

The other aspect of an extreme event is its intensity, termed the ''excess'' over the threshold, say $Y = X - u$. Consistent with the theory about the limiting distribution of the maximum being GEV, the distribution of the excess should have an approximate generalized Pareto (GP) distribution for a sufficiently high threshold (Pickands 1975). The GP has cumulative distribution function

$$F(y; \sigma^*, \xi) = \begin{cases} 1 - [1 + \xi(y/\sigma^*)]^{-1/\xi}, \\ \quad 1 + \xi(y/\sigma^*) > 0 & \xi \neq 0 \\ 1 - e^{-y/\sigma^*} & \xi = 0. \end{cases} \quad (7)$$

Here $y > 0$, with $\sigma^* > 0$ being a scale and ξ a shape parameter. If the random variable Y has a GP distribution, then the rescaled random variable Y/σ^* has a distribution which does not depend on σ^*, only on ξ. As for the GEV, the scale parameter governs the spread of the distribution.

The shape parameter of the GP distribution has precisely the same interpretation as for GEV distribution. In this case, the three possible types (Fig. 2) are:

(i) $\xi = 0$, a light-tailed (or exponential) distribution;

(ii) $\xi > 0$, a heavy-tailed (or Pareto) distribution;

(iii) $\xi < 0$, a bounded (or beta) distribution.

The type (i) or exponential distribution has a ''memoryless'' (or lack of aging) property (e.g., Ross 1970); that is,
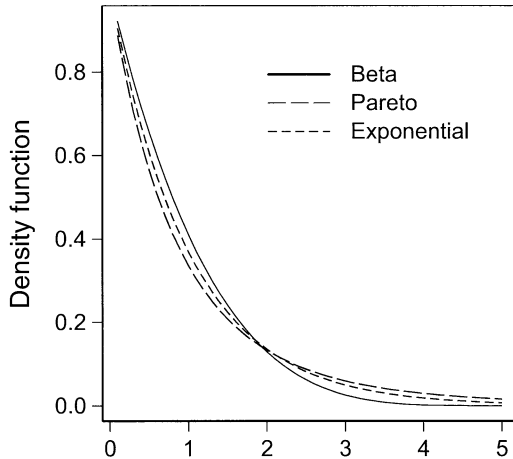
FIG. 2. Plots of the GP (generalized Pareto) probability density function with $\sigma^* = 1$, $\xi = -0.2$ (Beta type), $\xi = 0$ (exponential), and $\xi = 0.2$ (Pareto).

$$\Pr\{Y > y + y' \mid Y > y'\} = \Pr\{Y > y\}$$
$$= e^{-y/\sigma^*} \qquad y' > 0. \qquad (8)$$

In other words, if Y represents life span and has an exponential distribution, then the conditional distribution of future survival is still exponential with the same scale parameter, no matter how long the individual has already survived. This type of distribution would be reasonable for the life expectancy of many species of adult birds in which mortality is primarily the result of accidents, not aging (Ehrlich et al. 1988).

As the exponential is the only memoryless distribution (Ross 1970), the GP distribution necessarily cannot retain this property for a nonzero shape parameter. Suppose the distribution of the excess Y is exactly, instead of only approximately, GP for some threshold u. As the threshold is increased, say to some value $u' > u$, then the distribution of the excess over the higher threshold would remain the GP with the identical shape parameter ξ. Only the scale parameter needs to be adjusted; that is,

$$\sigma^*(u') = \sigma^*(u) + \xi(u' - u) \qquad (9)$$

where the dependence of the scale parameter on the threshold u is made explicit by writing $\sigma^*(u)$ (Coles 2001).

The $(1 - p)$th quantile of the GP distribution can be obtained by inverting Eq. 7, giving

$$F^{-1}(1 - p; \sigma^*, \xi) = \begin{cases} (\sigma^*/\xi)(p^{-\xi} - 1), & \xi \neq 0, \\ \sigma^*\ln(1/p), & \xi = 0 \end{cases} \qquad (10)$$

(Coles 2001). Like the GEV, a GP quantile is linear in σ^*, but highly nonlinear in ξ. In practice, to obtain a return level (analogous to that for the GEV in Eq. 4), the probability p that enters into Eq. 10 would need to be adjusted to take into account the probability of an exceedance of the threshold (Coles 2001).

The fact that the upper tail of essentially any distribution must be approximately of the GP form (Eq. 7) has implications concerning the search for ecological scaling laws. In particular, any distribution must approximately satisfy a scaling law beyond a high threshold if one permits a generalized Pareto, instead of an ordinary Pareto, form. So, at least for the upper tail of a distribution, the only issue which would remain to be addressed in practice is how large a threshold suffices for a good approximation.

A close correspondence exists between the block maxima and POT approaches to the statistical modeling of extremes, with the maximum of a sequence of observations falling below a threshold if and only if there are no exceedances of the threshold. Thus, the POT approach can be used to indirectly fit the GEV distribution, potentially making use of more information about extremes than just block maxima.

A point process representation provides a formal theoretical justification for the POT approach to fitting the GEV distribution. It consists of a two-dimensional, non-homogeneous Poisson process, combining the Poisson process (parameter λ) for the times of exceedance of the high threshold and the GP distribution (parameters σ^* and ξ) for the excesses over the threshold (Leadbetter et al. 1983, Smith 1989, Davison and Smith 1990). The parameters, λ and σ^*, of the point process are related to the parameters of the GEV distribution, μ, σ, and ξ, by

$$\ln \lambda = -(1/\xi)\ln[1 + \xi(u - \mu)/\sigma]$$
$$\sigma^* = \sigma + \xi(u - \mu) \qquad (11)$$

with the shape parameter being identical (Coles 2001).

The block maxima and POT approaches can involve a difference in time scales; for example, with time scale parameter $h \approx 1/365$ if the observations are daily and annual maxima are modeled. To convert the parameters of the GEV distribution, μ, σ, and ξ for time scale h (e.g., annual maxima of daily data), to the corresponding parameters of the GEV, μ', σ', and ξ for time scale h' (e.g., monthly maxima of daily data), it can be shown that

$$\sigma' = \sigma\delta^\xi \qquad \mu' = \mu + [\sigma'(1 - \delta^{-\xi})]/\xi \qquad (12)$$

where $\delta = h/h'$ and the shape parameter is unchanged (Coles 2001). Taking logarithms, $\ln(\sigma)$ varies linearly with $\ln(h)$, but $\ln(\mu)$ does not.

In practice, the POT approach requires the choice of a high threshold. No reliable automatic techniques for threshold selection are yet available, but some helpful diagnostics do exist (Coles 2001). The difficulty arises because of the trade-off between making the threshold high enough to ensure that the GP approximation is valid, but not so high that the number of exceedances is too small for accurate estimation of the parameters.

Another issue concerns the possible clustering of high levels, because parameter estimation techniques

such as maximum likelihood assume that the time series of excesses over the threshold be independent. For simple forms of dependence such as an autoregressive process, any apparent clustering should gradually disappear as the threshold increases. But there is some evidence that variables such as temperature do exhibit clustering that does not vanish for high thresholds (Coles 2001). To deal with this problem, the chief option is "declustering." Originally developed by hydrologists, this somewhat ad hoc and inefficient procedure requires defining clusters of high levels and using only the single highest value within a cluster (Todorovic and Zelenhasic 1970). Rather than simply being viewed as a nuisance to be eliminated, clusters might well be ecologically relevant for their own sake (e.g., reflecting a persistent hot or wet spell).

Like the block maxima approach, covariates can be incorporated into the POT approach (Eq. 5). By this means, features such as annual or diurnal cycles that are necessarily ignored in the block maxima approach can be modeled as well. Any such cycles need not be removed before analyzing extremes as did Gaines and Denny (1993). For instance, Katz et al. (2002) treated an example of daily precipitation extremes in which annual cycles in both the location and scale parameters of the GEV distribution are permitted.

But how could heavy tails arise in ecological applications? To explain such behavior, at least one chance mechanism exists that is quite plausible ecologically. Suppose that a variable Y has a distribution with an exponential (i.e., light) upper tail (Eq. 7), but that the inverse of the scale parameter (i.e., rate parameter $\nu = 1/\sigma^*$) of the exponential is itself a random variable with a gamma distribution (say with probability density function f_ν). That is, the conditional distribution of Y is given by

$$\Pr\{Y > y \mid \nu\} = e^{-\nu y}$$

$$f_\nu(\nu; \alpha, \beta) = [\beta\Gamma(\alpha)]^{-1}(\nu/\beta)^{\alpha-1}e^{-\nu/\beta} \qquad \alpha, \beta > 0.$$
(13)

Then the unconditional distribution of Y would have a heavy tail (e.g., Arnold 1983). That is,

$$\Pr\{Y > y\} = (1 + \beta\, y)^{-\alpha} \qquad (14)$$

the type (ii) (or Pareto) form under a different parameterization than that in Eq. 7 (i.e., with shape parameter $\xi = 1/\alpha > 0$). This chance mechanism, by which an unconditional heavy tail is generated from a conditional light tail, is similar to other mathematical modeling with a long tradition of use in ecology (Pielou 1977). For instance, the gamma distribution for the rate parameter could reflect some additional source of random variation in an ecological variable (e.g., associated with an unobserved covariate).

APPLICATION TO PALEOECOLOGICAL DISTURBANCES
Nicolay Lake sediment yield

The first example consists of a nearly 500-yr annual time series of sediment yield in a pristine environment
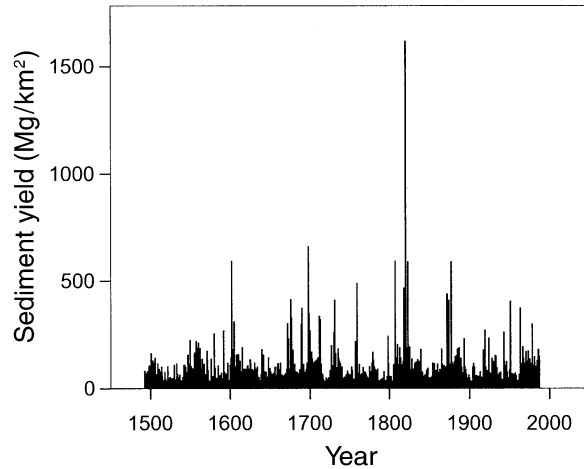

FIG. 3. Time series of annual sediment yield at Nicolay Bay, Nunavut, Canada, 1493–1987 (Source: Lamoureux 2000).

in the high Arctic, with presumably minimal anthropogenic influence (Fig. 3). These sediment yields were reconstructed from varved sediments at Nicolay Lake on Cornwall Island, Nunavut, Canada (Lamoureux 2000). Any disturbances in the time series are predominantly of a hydrologic origin, with high sediment yields being associated with summers in which intense rainfall occurs.

Lamoureux (2000) fitted a GEV distribution directly to the annual sediment yield time series (i.e., without taking block maxima). We repeat this analysis along with the alternative POT modeling approach. Table 1 summarizes the results of fitting a GEV distribution to this data by maximum likelihood. Consistent with Lamoureux (2000), the evidence of a heavy tail (i.e., $\xi > 0$) is very strong. A 95% confidence interval for the shape parameter ξ (all intervals based on the method of profile likelihood, Coles 2001) has a lower bound of about 0.3 or well above zero. As a diagnostic check, a quantile–quantile (Q–Q) plot for the fitted GEV distribution (Fig. 4a) is approximately linear, indicating that the assumed form of distribution is reasonable.

A more traditional approach in paleoclimatology or paleohydrology would be to assume a lognormal distribution; that is, the logarithmically transformed sediment yields are normally distributed (Rittenour et al. 2000). Like the GEV (or GP), a lognormal distribution is positively skewed; but unlike the GEV (or GP), it has a light upper tail (Leadbetter et al. 1983). The estimated 500-yr return level for the fitted lognormal distribution is 523.9 Mg/km^2, well below the highest and second highest observed values in the ~500-yr record of 1620.3 in 1820 and 663.7 in 1698. For the fitted GEV distribution, the estimated 500-yr return level is just under 1000 Mg/km^2 (i.e., $p = 0.002$ in Eq. 4), with the corresponding 95% confidence interval having an upper limit a bit smaller than the highest observed val-

TABLE 1. Parameter and return level estimates with standard errors (or confidence intervals) for GEV and GP distributions (as well as GEV via point process approach) fitted to time series of annual sediment yield (Mg/km²) at Nicolay Bay, Nunavut, Canada, 1493–1987.

Parameter	Estimate	1 SE (or 95% CI)
GEV distribution		
Location μ	67.54	1.85
Scale σ	35.86	1.64
Shape ξ	0.384	(0.305, 0.470)
500-yr return level	987.6	(724.9, 1435.1)
GP distribution†		
Scale σ*	47.82	5.98
Shape ξ	0.462	(0.277, 0.700)
500-yr return level	1168.9	(732.3, 2451.2)
GEV/POT†		
Location μ	62.84	6.28
Scale σ	30.64	6.33
Shape ξ	0.462	0.107

Notes: Data are from Lamoureux (2000). Abbreviations: GEV, generalized extreme value; GP, generalized Pareto; POT, peaks over threshold.

† Threshold $u = 100$ Mg/km².

ue and a lower limit somewhat greater than the second highest observation (Table 1). In other words, the GEV distribution fits the extreme upper tail of the data reasonably well, whereas the lognormal substantially underestimates the likelihood of the most extreme events.

Lamoureux (2000) made a convincing argument for why the Nicolay Lake sediment yield is a proxy for extreme hydrologic events, particularly for short spells of heavy rainfall (i.e., at most a few days in duration). Nevertheless, the fact that the GEV distribution appears to fit the annual sediment yield (without taking block maxima) well could be fortuitous. So the alternative POT approach, focusing only on the upper tail of the distribution of sediment yield, will be applied for comparison.

Table 1 includes the results of fitting the GP distribution to the excesses in annual sediment yield over a threshold of $u = 100$ Mg/km² (after a range of possible thresholds was examined). This threshold was exceeded in nearly 40% of the years (i.e., 189 out of 495). The estimated shape parameter is a bit higher than for the original GEV fit, with the 95% confidence interval being somewhat wider primarily due to the decreased sample size. Because of the heavier fitted tail, the estimated 500-yr return level (Eq. 10) is higher than for the original GEV fit, with a 95% confidence interval having a much higher upper limit well exceeding the highest observed value. The Q–Q plot for the fitted GP distribution (Fig. 4b) appears similar to that for the upper tail in the corresponding Q–Q plot for the GEV (Fig. 4a). In fitting the GP distribution, any clustering of the excesses over the threshold, a possibility suggested for the Nicolay sediment yield by Lamoureux (2002), has been ignored.

The point process approach can produce the equivalent fitted model in terms of the GEV parameterization, using the same threshold of $u = 100$ Mg/km² and a time scale parameter of $h = 1$. These parameter estimates and standard errors are also included in Table 1. They could also have been obtained indirectly from the fitted GP distribution and Poisson rate parameter λ for exceedance occurrence (estimated as 189/495 ≈ 0.382) using Eq. 11. As previously noted, the shape parameter estimates are necessarily identical. Because only the upper tail of the distribution is modeled, the GEV parameter estimates differ somewhat from those obtained by fitting the entire range of data.

It would be tempting to consider the North Atlantic Oscillation (NAO) as a covariate. This large-scale feature of the atmospheric circulation reflects a ''see-saw'' in pressure between the Azores High and the Icelandic Low (Hurrell et al. 2003). However, the NAO is well known to be ''teleconnected'' with climate in the upper latitudes of the Northern Hemisphere during the winter, but not necessarily in the summer as would be required for the Nicolay sediment yield. We use as a covariate, instead, an index of the Icelandic Low alone, the summer (i.e., June–August) mean sea level pressure at Rey-
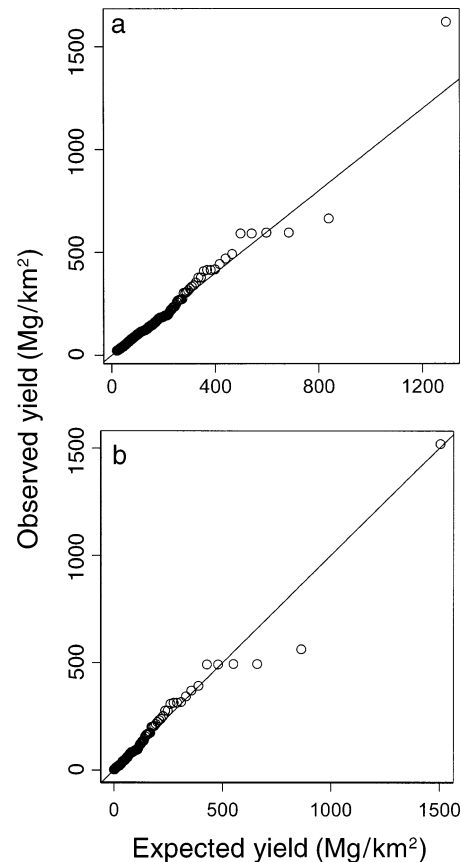

FIG. 4. Quantile–quantile plots for fit of (a) GEV and (b) GP distribution (threshold = 100 Mg/km²) to time series of sediment yield at Nicolay Bay.

TABLE 2. Parameter estimates and standard errors for GEV distribution (along with negative log likelihood), fitted to time series of sediment yield (Mg/km^2) at Nicolay Bay, 1822–1987, with and without June–August mean pressure (in hPa, minus 1000) at Reykjavik, Iceland, as covariate.

Parameter	Estimate	1 SE
GEV distribution†		
Location μ	73.44	3.36
Scale σ	37.71	2.86
Shape ξ	0.322	0.071
GEV (covariate)‡		
Intercept μ_0	84.18	10.67
Slope μ_1	−1.10	1.04
Scale σ	37.44	2.85
Shape ξ	0.329	0.071

Note: Data are from Jones et al. (1997).
† Negative log likelihood = 895.054.
‡ Negative log likelihood = 894.494.

FIG. 5. Time series of sediment rate at Furnace Bay, Maryland, USA, 1804–1980.

kjavik, Iceland. Because long time series indices of the NAO are reconstructed from proxy information such as sediment yield (Cook et al. 2002), we restrict the analysis to the period starting in 1822 for which pressure measurements are available (Jones et al. 1997).

For this considerably shorter time series of sediment yield (i.e., 166 yr), Table 2 compares the fitted GEV distribution with and without the Icelandic pressure covariate, say a random variable denoted by Z. Given a value of the pressure covariate, $Z = z$, the conditional distribution of sediment yield is assumed GEV with a location parameter depending linearly on z; that is, $\mu(z) = \mu_0 + \mu_1 z$. As anticipated, the incorporation of the pressure covariate does not significantly improve the fit. Specifically, comparing the minimized negative log likelihood function with and without the constraint of the slope parameter $\mu_1 = 0$, a P value ≈ 0.290 is obtained for likelihood ratio test. If it was important to focus on the upper tail of the distribution of sediment yield, then the pressure covariate could be introduced via the point process approach instead.

Chesapeake Bay sediment rate

The second example consists of a shorter time series of sediment rates for an estuary in which the influences of climatic and anthropogenic factors are confounded (Fig. 5). These sediment rates were derived from a core at Furnace Bay in the upper Chesapeake Bay, Maryland, USA (Brush 1989). To obtain average sediment rates, the core was divided into 2 cm-wide layers and dated on the basis of pollen profiles. Because of lower temporal resolution before 1800, we only analyze the sediment rates during 1804–1980, a total of 53 measurements with temporal resolution ranging from 2 to 8 yr. Ecologically related disturbances include heavy rainfall, sometimes associated with hurricanes, as well as land use activities such as deforestation associated with agricultural practices.

The interpretation of the Furnace Bay sediment rates is more difficult than for Nicolay Lake, because they
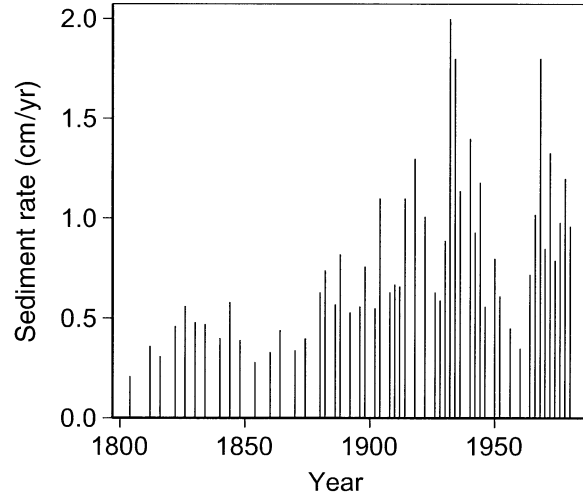
do not reflect exclusively extreme hydrologic events. In addition, even if extreme disturbances do play an important role, their influence is attenuated by the limited temporal resolution. Notwithstanding these complications, we repeat the analysis approach used in the Nicolay Lake example.

First, a GEV distribution is fitted to the sediment yield time series without taking block maxima. Table 3 includes the parameter estimates and standard errors, with the estimated shape parameter being positive but barely larger than its standard error. The Q–Q plot for the fitted GEV distribution (not shown) looks adequate. Unlike the Nicolay Lake example, a lognormal distribution would have produced reasonable extreme upper quantile estimates for the Furnace Bay data.

Table 3 also includes the results of fitting the GP distribution to the excesses in sediment rate over a threshold of $u = 0.5$ cm/yr, reducing the effective sample size from 53 to 38. Now the estimated shape pa-

TABLE 3. Parameter estimates with standard errors for GEV and GP distributions (as well as GEV via point process approach) fitted to time series of sediment rate (cm/yr) at Furnace Bay, Maryland, USA, 1804–1980.

Parameter	Estimate	1 SE
GEV distribution		
Location μ	0.567	0.042
Scale σ	0.264	0.033
Shape ξ	0.164	0.131
GP distribution†		
Scale σ^*	0.497	0.117
Shape ξ	−0.181	0.173
GEV/POT†		
Location μ	0.330	0.096
Scale σ	0.528	0.151
Shape ξ	−0.181	0.173

† Threshold $u = 0.5$ cm/yr.

TABLE 4. Parameter estimates and standard errors for GEV distribution (along with negative log likelihood and BIC statistic), with possible shifts in parameters after 1880, fitted to time series of sediment rate (cm/yr) at Furnace Bay, 1804–1980.

Parameter shift	$\mu(1)$	$\mu(2)$	$\sigma(1)$	$\sigma(2)$	$\xi(1)$	$\xi(2)$	$-\ln(L)$	BIC
None	0.567	0.567	0.264	0.264	0.164	0.164	18.165	48.241
μ	0.406	0.693	0.215	0.215	0.145	0.145	6.725	29.332
μ and σ	0.360	0.745	0.100	0.276	0.022	0.022	0.011	19.873†
μ, σ, and ξ	0.374	0.733	0.105	0.267	−0.239	0.110	−0.894	22.034

† Denotes minimum.

rameter is reversed in sign, but still about the same magnitude as its standard error. Despite only a fairly small proportion of the measurements being excluded, even weak evidence for a heavy upper tail has vanished. Again a Q–Q plot (not shown) appears satisfactory. The parameter estimates for the equivalent GEV distribution, based on the point process approach, differ quite a bit from those for the GEV directly fitted to the entire range of data.

An apparent shift in the level of sediment rates occurred around 1880 (Fig. 5), roughly the time of a transition from developing to commercial agriculture (Brush 1989). We incorporate this possible shift by allowing the parameters of the GEV distribution to change after 1880; that is, dealing with two sets of parameters, $\mu(i)$, $\sigma(i)$, $\xi(i)$, for the ith time period ($i = 1, 2$). By dividing the data set into two subsets, such a model could be fitted a single GEV distribution at a time. However, a more flexible approach involves fitting a GEV distribution with covariates. It is straightforward to allow only some of the parameters to vary, and other forms of trend in the parameters can be modeled as well.

Table 4 shows the results of fitting GEV distributions with various constraints on the parameter shifts, ranging from no change in any parameter (i.e., a single GEV distribution) to shifts in all three parameters (i.e., two completely different GEV distributions). Because several candidate models are considered, the Bayesian information criterion (BIC) is used to identify the best model (Schwarz 1978). We define the quantity

$$\mathrm{BIC}(k) = -2 \ln L(k) + k \ln(T) \qquad (15)$$

where $L(k)$ denotes the maximized likelihood function for a candidate model, requiring the estimation of k parameters, and T the sample size. This quantity is evaluated for each candidate model, the one with the smallest BIC value being selected.

For this example, the preferred model is the one in which both the location and scale parameters, but not the shape parameter, shift between the two time periods (Table 4). If tests of significance were employed instead, the same conclusions would have been reached (e.g., a likelihood ratio test, comparing the model in which no parameters are varied with the one in which only μ is varied gives a P value $< 1 \times 10^{-7}$; comparing the one in which only μ is varied with the one varying

both μ and σ, a P value ≈ 0.00025). The optimal model has an estimated shape parameter of virtually zero, suggesting the apparent positive shape parameter with no parameter shifts might be an artifact of neglecting the shift in land-use practice. Whether the lack of evidence in favor of a heavy tail reflects a real difference between Furnace Bay and Nicolay Bay in the processes which influence sedimentation or is just an artifact of the measurement process at Furnace Bay cannot be readily ascertained.

In this example, an extreme value model has been adjusted by using time as a surrogate for land use practice. Covariates other than time could be introduced into the model as well. Some of the heaviest rainfall events in the Chesapeake Bay region are associated with remnants of tropical storms (Brush 1989). The frequency of hurricanes in the North Atlantic is well known to be modulated by the El Niño phenomenon (Gray 1984), with the preference for landfall along the Atlantic Coast, as opposed to the Gulf of Mexico, being influenced by the NAO (Elsner et al. 2000). But the very low frequency of hurricane-induced precipitation events within a single water basin, along with the coarse time resolution of the sediment rate time series, would mitigate against such covariates actually improving the fit in practice.

DISCUSSION

Some potential advantages of extreme value theory in modeling ecological disturbances have been pointed out. The reliance on more conventional statistical methods can result in labeling ecological disturbances as more "surprising" than they ought to be. As one example, strong evidence has been presented in support of a heavy-tailed distribution (i.e., decreasing at a power law rate) for the Nicolay Lake sediment yield. Although not identifiable, the origin of this heavy tail is consistent with the heavy tail that precipitation amount tends to possess (Katz et al. 2002), as well as with sediment yield being an integrator over a catchment. By not exaggerating how unusual they are, the proper treatment of extreme events (e.g., large fire) as an integral part of ecosystems would be facilitated.

It has been demonstrated how the statistics of extremes can readily incorporate information about covariates, such as patterns in large-scale atmospheric–oceanic circulation or shifts in land use associated with

agricultural practices. For instance, the Furnace Bay sediment rate was shown to be better fitted by an extremal model whose parameters shift corresponding to the transition from developing to commercial agriculture. More generally, through adjustment to reflect the dynamic nature of ecosystems, the degree of ecological realism of extreme value models can be enhanced.

The issue of scale, both temporal and spatial, can be paramount in ecology. We have only focused on the temporal component in our treatment of extremes. In particular, an apparently unappreciated connection between the existence of power laws in ecology and statistical extreme value theory has been identified. Regarding spatial patterns in extremes, Gaines and Denny (1993) observed consistency in the parameter estimates of extreme value distributions for certain variables at different spatial locations. They postulated the ''possible existence of underlying principles governing these phenomena.'' While it may be difficult in practice to determine whether such consistency is attributable to a bona fide ecological invariance principle or just to the statistical theory of extreme values, it would certainly be useful to examine more systematically how ecological extremes vary spatially. In hydrology, ''regional analysis'' has long been used to obtain more accurate estimates of flood probabilities, exploiting the fact that some extremal parameters only gradually vary within a region (e.g., Hosking et al. 1985).

In principle, fully spatial-temporal modeling of extremes can be attempted through the extension of the concept of max-stability to max-stable processes, but applications so far are rather limited and unrealistic (Coles 1993). Some empirical work on how extremes scale spatially, particularly in the hydrologic context of flood estimation, has detected apparent universal behavior but connections to extreme value theory have not yet been made (Smith 1992).

Literature Cited

Arnold, B. C. 1983. Pareto distributions. International Cooperative Publishing House, Fairland, Maryland, USA.

Box, J. F. 1978. R. A. Fisher: the life of a scientist. John Wiley and Sons, New York, New York.

Brown, B. G., and R. W. Katz. 1995. Regional analysis of temperature extremes: spatial analog for climate change? Journal of Climate 8:108–119.

Brush, G. S. 1989. Rates and patterns of estuarine sediment accumulation. Limnology and Oceanography 34:1235–1246.

Carey, J. R. 2003. Longevity: the biology and demography of life span. Princeton University Press, Princeton, New Jersey, USA.

Clark, J. S., M. Lewis, and L. Horvath. 2001. Invasion by extremes: population spread with variation in dispersal and reproduction. American Naturalist 157:537–554.

Coles, S. G. 1993. Regional modelling of extreme storms via max-stable processes. Journal of the Royal Statistical Society, Series B 55:797–816.

Coles, S. 2001. An introduction to statistical modeling of extreme values. Springer-Verlag, London, UK.

Coles, S. G., and M. J. Dixon. 1999. Likelihood-based inference for extreme value models. Extremes 2:5–23.

Cook, E. R., R. D. D'Arrigo, and M. E. Mann. 2002. A well-verified, multiproxy reconstruction of the winter North Atlantic Oscillation index since A.D. 1400. Journal of Climate 15:1754–1764.

Davison, A. C., and R. L. Smith. 1990. Models for exceedances over high thresholds. Journal of the Royal Statistical Society, Series B 52:393–442.

Denny, M., and S. Gaines. 2000. Chance in biology: using probability to explore nature. Princeton University Press, Princeton, New Jersey, USA.

Dixon, P. M., A. M. Ellison, and N. J. Gotelli. 2005. Improving the precision of estimates of the frequency of rare events. Ecology 85:1114–1123.

Ehrlich, P. R., D. S. Dobkin, and D. Wheye. 1988. The birder's handbook. Simon and Schuster, New York, New York, USA.

Elsner, J. B., K.-B. Liu, and B. Kocher. 2000. Spatial variations in major U.S. hurricane activity: statistics and a physical mechanism. Journal of Climate 13:2293–2305.

Ferriere, R., and B. Cazelles. 1999. Universal power laws govern intermittent rarity in communities of interacting species. Ecology 80:1505–1521.

Fisher, R. A., and L. H. C. Tippett. 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceedings of the Cambridge Philosophical Society 24:180–190.

Folland, C. K., and T. R. Karl. 2001. Observed climate variability and change. Pages 99–181 in J. T. Houghton et al., editors. Climate change 2001: the scientific basis. Cambridge University Press, Cambridge, UK.

Gaines, S. D., and M. W. Denny. 1993. The largest, smallest, highest, lowest, longest, and shortest: extremes in ecology. Ecology 74:1677–1692.

Gray, W. M. 1984. Atlantic seasonal hurricane frequency. Part I: El Niño and 30 mb quasi-biennial oscillation influences. Monthly Weather Review 112:1649–1668.

Gregory, K. J., and G. Benito, editors. 2003. Paleohydrology: understanding global change. John Wiley and Sons, New York, New York, USA.

Gumbel, E. J. 1958. Statistics of extremes. Columbia University Press, New York, New York, USA.

Gutschick, V. P., and H. BassiriRad. 2003. Extreme events as shaping physiology, ecology, and evolution of plants: toward a unified definition and evaluation of their consequences. New Phytologist 160:21–42.

Hosking, J. R. M., J. R. Wallis, and E. F. Wood. 1985. An appraisal of the regional flood frequency procedure in the UK Flood Studies Report. Hydrological Sciences Journal 30:85–109.

Hurrell, J. W., Y. Kushnir, G. Ottersen, and M. Visbeck, editors. 2003. The North Atlantic Oscillation: climatic significance and environmental impact. Geophysical Monograph 134, American Geophysical Union, Washington, D.C., USA.

Jones, P. D., T. Jónsson, and D. Wheeler. 1997. Extension to the North Atlantic Oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. International Journal of Climatology 17:1433–1450.

Katz, R. W. 2002a. Techniques for estimating uncertainty in climate change scenarios and impact studies. Climate Research 20:167–185.

Katz, R. W. 2002*b*. Stochastic modeling of hurricane damage. Journal of Applied Meteorology **41**:754–762.

Katz, R. W., M. B. Parlange, and P. Naveau. 2002. Statistics of extremes in hydrology. Advances in Water Resources **25**:1287–1304.

Keitt, T. H., and H. E. Stanley. 1998. Dynamics of North American breeding bird populations. Nature **393**:257–260.

Lamoureux, S. 2000. Five centuries of interannual sediment yield and rainfall-induced erosion in the Canadian High Arctic recorded in lacustrine varves. Water Resources Research **36**:309–318.

Lamoureux, S. 2002. Temporal patterns of suspended sediment yield following moderate to extreme hydrological events recorded in varved lacustrine sediments. Earth Surface Processes and Landforms **27**:1107–1124.

Leadbetter, M. R., G. Lindgren, and H. Rootzén. 1983. Extremes and related properties of random sequences and processes. Springer-Verlag, New York, New York, USA.

Ludwig, D. 1996. Uncertainty and the assessment of extinction probabilities. Ecological Applications **6**:1067–1076.

Lynch, J. A., J. S. Clark, N. H. Bigelow, M. E. Edwards, and B. P. Finney. 2003. Geographical and temporal variations in fire history in boreal ecosystems of Alaska. Journal of Geophysical Research **108**:A8152, doi: 10.1029/2001JD000332.

Moritz, M. A. 1997. Analyzing extreme disturbance events: fire in Los Padres National Forest. Ecological Applications **7**:1252–1262.

Palutikof, J. P., B. B. Brabson, D. H. Lister, and S. T. Adcock. 1999. A review of methods to calculate extreme wind speeds. Meteorological Applications **6**:119–132.

Pickands, J. 1975. Statistical inference using extreme order statistics. Annals of Statistics **3**:119–131.

Pielou, E. C. 1977. Mathematical ecology. Second edition. John Wiley and Sons, New York, New York, USA.

Reiss, R.-D., and M. Thomas. 2001. Statistical analysis of extreme values: with applications to insurance, finance, hydrology, and other fields. Extended second edition. Birkhäuser, Basel, Switzerland.

Rittenour, T. M., J. Brigham-Grette, and M. E. Mann. 2000. El Niño-like climate teleconnections in New England during the late Pleistocene. Science **288**:1039–1042.

Ross, S. M. 1970. Applied probability models with optimization applications. Holden-Day, San Francisco, California, USA.

Schoenberg, F. P., R. Peng, and J. Woods. 2003. On the distribution of wildfire sizes. Environmetrics **14**:583–592.

Schwarz, G. 1978. Estimating the dimension of a model. Annals of Statistics **6**:461–464.

Smith, J. A. 1992. Representation of basin scale in flood peak distributions. Water Resources Research **28**:2993–2999.

Smith, R. L. 1989. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. Statistical Science **4**:367–393.

Stedinger, J. R., and T. A. Cohn. 1986. Flood frequency analysis with historical and paleoflood information. Water Resources **22**:785–793.

Todorovic, P., and E. Zelenhasic. 1970. A stochastic model for flood analysis. Water Resources Research **6**:1641–1648.

SUPPLEMENT

The R source code for fitting extreme value distributions is available in ESA's Electronic Data Archive: *Ecological Archives* E086-060-S1.

Special Feature