

Statistics of extremes in hydrology

Richard W. Katz ^{a,*}, Marc B. Parlange ^b, Philippe Naveau ^c

^a Environmental and Societal Impacts Group, National Center for Atmospheric Research, Boulder, CO 80307, USA

^b Department of Geography and Environmental Engineering, Johns Hopkins University, Baltimore, MD 21218, USA

^c Institut Pierre Simon Laplace, Laboratoire de Météorologie Dynamique, École Polytechnique 91 128 Palaiseau, France

Received 7 December 2001; received in revised form 4 February 2002; accepted 27 May 2002

Abstract

The statistics of extremes have played an important role in engineering practice for water resources design and management. How recent developments in the statistical theory of extreme values can be applied to improve the rigor of hydrologic applications and to make such analyses more physically meaningful is the central theme of this paper. Such methodological developments primarily relate to maximum likelihood estimation in the presence of covariates, in combination with either the block maxima or peaks over threshold approaches. Topics that are treated include trends in hydrologic extremes, with the anticipated intensification of the hydrologic cycle as part of global climate change. In an attempt to link downscaling (i.e., relating large-scale atmosphere–ocean circulation to smaller-scale hydrologic variables) with the statistics of extremes, statistical downscaling of hydrologic extremes is considered. Future challenges are reviewed, such as the development of more rigorous statistical methodology for regional analysis of extremes, as well as the extension of Bayesian methods to more fully quantify uncertainty in extremal estimation. Examples include precipitation and streamflow extremes, as well as economic damage associated with such extreme events, with consideration of trends and dependence on patterns in atmosphere–ocean circulation (e.g., El Niño phenomenon).

© 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Climate change; Covariates; Maximum likelihood; Statistical downscaling

1. Introduction

The mathematician, carried along on his flood of symbols, dealing apparently with purely formal truths, may still reach results of endless importance for our description of the physical universe. (Karl Pearson, 1857–1936)

This quote by Karl Pearson, a pioneer in the field of statistics, makes use of the term “flood” in a sense that has no connection to hydrology. In fact, some would argue that mathematical statistics have made little, if any, contribution to the physical understanding of hydrologic phenomenon such as floods. On the other hand, the statistics of extremes have played an important role in engineering practice for water resources design and management. How recent developments in the statistical theory of extreme values can be applied

both to improve the rigor of hydrologic applications and to make such analyses more physically meaningful is the central theme of this paper.

Although the fundamental probabilistic theory of extreme values has been well developed for a long time (e.g., [59]), the statistical modeling of extremes remains a subject of active research (for a recent text, see [15]). In the present paper, a statistical modeling approach based on maximum likelihood (ML) estimation, in the possible presence of covariates, will be advocated [15,87]. Such covariates could incorporate trends, cycles, or actual physical variables (e.g., measures of large-scale atmosphere–ocean circulation patterns such as the El Niño phenomenon). Probability weighted moments (PWM) (or *L*-moments) are more popular than ML in applications to hydrologic extremes, both because of their computational simplicity and because of their good performance for small samples [40,43]. But the PWM technique has the disadvantage of not being able to readily incorporate covariates. On the other hand, it is still straightforward to apply the ML technique in the presence of covariates, in conjunction with either the

* Corresponding author.

E-mail addresses: rwk@ucar.edu (R.W. Katz), mbparlange@jhu.edu (M.B. Parlange), naveau@lmd.polytechnique.fr (P. Naveau).

block (e.g., annual) maxima or the peaks over threshold (POT) (commonly termed “partial duration series” in the hydrologic literature) approaches.

In the present paper, the applications to hydrologic extremes are limited to those about which we are most familiar. In particular, our focus is on climate-related issues, both variability and change. As part of global climate change, an accelerated hydrologic cycle (including an increase in heavy precipitation) is anticipated on a theoretical basis [95,96], is predicted by numerical models of the climate system [21], and has been detected in observed precipitation [32,48] and, possibly, streamflow [33,62] and evaporation [8,93]. Improved understanding of large-scale atmosphere–ocean sources of low frequency variation (such as the El Niño phenomenon) has led to efforts in “statistical downscaling” for smaller-scale meteorological and hydrologic variables [60].

In Section 2 the use of the statistics of extremes in hydrology and the characteristics of hydrologic extremes are reviewed. Section 3 contains a description of recent methodological developments in the statistics of extremes, including a point process model that combines the block maxima and POT techniques and a comparison of estimation techniques (some technical details about ML estimation are relegated to an appendix). The issue of trends in hydrologic extremes in conjunction with global climate change is the subject of Section 4, whereas an attempt to link downscaling with the statistics of extremes is made in Section 5. Finally, problems in hydrologic extremes for which future developments in the statistics of extremes are needed (e.g., regional analysis) are outlined in Section 6.

2. Hydrologic extremes

2.1. Historical use of statistics of extremes

Some of the earliest applications of the statistical theory of extreme values were to hydrology and to closely related problems in climatology. In a paper published in 1941, the statistician Emil Gumbel, a pioneer in the application of the statistics of extremes, considered the problem of the statistical estimation of the “return period” of floods. He cautioned that:

In order to apply any theory we have to suppose that the data are homogeneous, i.e. that no systematical change of climate and no important change in the basin have occurred within the observation period and that no such changes will take place in the period for which extrapolations are made. (p. 187 in [34])

Thus early on, the spectre of possible climate change was raised, an issue that will receive attention in the

present paper. Gumbel also grappled with opponents to the statistics of extremes on a physical basis; in the context of floods, the question of how could the distribution of largest values be unbounded. He observed that “Some authors have tried to introduce upper and lower limits to the discharges, even though it is doubtful that such limits exist” (p. 163 in [34]). In a somewhat different guise, this same conundrum will be confronted in the present paper.

Quite a bit later, Gumbel published the first book on the statistics of extremes. Many of the applications dealt with hydrology or climatology (pp. 236–245 in [35]), and he noted that: “The oldest problems connected with extreme values arise from floods” (p. 4 in [35]). Evidently, he was still confronted with resistance on a physical basis to the statistical theory of extreme values, being compelled to state:

The exploration of how unlimited distributions behave at infinity is just part of the common general effort of mathematics and science to transgress the finite, as calculus has done since Newton’s time for the infinite, and nuclear physics is doing for the infinitesimal. (p. 2 in [35])

Not surprisingly, several important original contributions to extreme value theory have been made by hydrologists in collaboration with statisticians. A few of these contributions are cited here: namely, the POT approach to extremal modeling, PWM estimation of parameters of extremal distributions, and regional analysis of extremes. The POT approach originated in hydrology quite a while ago [81,94]. Its rationale is that if additional information about the extreme upper tail were used besides the annual maxima (i.e., other relatively high values in the sample), then more accurate estimates of the parameters and quantiles of extreme value distributions would be obtained [64,65]. Early work in hydrology usually assumed an exponential distribution for the excess over a high threshold, being equivalent to a Gumbel distribution for the maximum [94]. The theoretical basis for the POT approach will be addressed in Section 3.1.

First proposed by Landwehr et al. [58] for extreme value distributions, the parameter estimation technique of PWM (i.e., linear combinations of “*L*-moments”, the term that commonly appears in the hydrology literature) also has a well-established tradition in hydrology [40]. This method has been advocated not just because of its good performance for small-sample sizes [43], but as a good choice of starting values for the iterative numerical procedure required to obtain ML estimates [39] and as an alternative to ML that is faster to compute and that produces “feasible” values (i.e., finite mean). The comparative performance of the ML and PWM techniques will be considered in Section 3.2.

Regional analysis is another device for making use of more available information that originated with estimation of hydrologic extremes in mind [22,44]. The basic idea is that if a region is relatively homogenous, then the estimation of extreme quantiles at a given site can be improved by using extreme observations at other sites as well (i.e., a trade-off between space and time). In practice, it is assumed that the region is homogenous, after adjustment for spatial variations in parameters (e.g., only the mean might be allowed to vary across the region). It is also common in hydrology to make use of PWM in performing a regional analysis [42]. Future statistical developments that could aid in improving the statistical rigor of regional analysis will be addressed in Section 6.1.

2.2. Characteristics of hydrologic extremes

Before discussing the statistical characteristics of hydrologic extremes, the generalized extreme value (GEV) distribution is defined. It arises as the limiting distribution of the maximum of a series of independent and identically distributed observations [59] (how these conditions can be relaxed will be discussed later, including Sections 2.2.3 and 3.1.1). It also arises in conjunction with physical systems that possess scale invariance [4]. The cumulative distribution function of the GEV is given by:

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

Here $\mu, \sigma > 0$, and γ are the location, scale, and shape parameters, respectively. The parameterization for the shape parameter γ in Eq. (1) follows the notational convention prevalent today in the statistics literature; in the hydrologic literature, it is still common to parameterize in terms of $\gamma^* = -\gamma$ instead.

If the shape parameter $\gamma > 0$, then the GEV distribution is said to be *heavy tailed*. Because its probability density function decreases at so slow a rate in the upper tail, the moments of the GEV are infinite for orders greater than $1/\gamma$ (e.g., the variance is infinite if $\gamma > 1/2$; the mean is infinite if $\gamma > 1$). If $\gamma < 0$, then the distribution has a bounded upper tail. The case of $\gamma = 0$ in Eq. (1), obtained by taking the limit of the general expression as $\gamma \rightarrow 0$, is termed the Gumbel distribution (i.e., an unbounded, thin tail). Used to estimate design values (or return levels), the quantile function $F^{-1}(1 - p; \mu, \sigma, \gamma)$, $0 < p < 1$, of the GEV distribution can be expressed as:

$$F^{-1}(1 - p; \mu, \sigma, \gamma) = \begin{cases} \mu + (\sigma/\gamma)\{[-\ln(1 - p)]^{-\gamma} - 1\}, & \gamma \neq 0, \\ \mu + \sigma\{-\ln[-\ln(1 - p)]\}, & \gamma = 0. \end{cases} \quad (2)$$

2.2.1. Heavy tails

There is much evidence that the distributions of hydrologic variables are heavy tailed. Such evidence is generally based on fitting the GEV distribution, with a likelihood ratio test being one way to test whether the shape parameter $\gamma = 0$ (i.e., thin tail) when ML estimation is used. One issue concerns the chance mechanism by which a heavy-tailed distribution arises. A mixture of Gumbel distributions, as might arise when the extremal distributions depends on the time of day or year or on some covariate such as the El Niño phenomenon, can resemble the GEV [68].

Although it can be difficult to determine from only a single site unless the record is relatively long, the distribution of maximum precipitation amount (e.g., hourly or daily) appears consistently to have a heavy tail [29,86,87]. Regional analyses of precipitation extremes, in which the shape parameter γ is constrained to be constant within the area, provide clear evidence of heavy tails [9,10]. An example of extreme daily precipitation amount will be treated in Section 2.3.1.

Streamflow (e.g., daily or monthly) is routinely found to possess a heavy tail, especially in arid and semi-arid regions, with this tail tending to heavier than that of precipitation [3,31,63,68]. An open question concerns the extent to which the heavy upper tail of streamflow is directly attributable to that of precipitation, as opposed to the integrative nature of runoff. Statistical theory, in combination with the physical relationship between rainfall and runoff, suggests that any heavy-tailed characteristics of precipitation ought to be inherited by streamflow. This issue could be more systematically studied through the use of rainfall–runoff models [92]. An example of peak streamflow will be treated in Section 2.3.2.

The extreme tails of other hydrologic variables have not received as much attention, but some likewise appear to be heavy tailed. For example, Lamoureux [57], in an analysis of a long record (nearly 500 yr) of sedimentation rates for a high latitude lake (apparently quite sensitive to extreme high precipitation), found a heavy tail. Further, the highest sedimentation rates for estuaries appear to be related to major storms [7]. The long records produced in paleohydrology and related fields would surely benefit from more use of the statistics of extremes.

Finally, there is some evidence that the distribution of the economic damage associated with extreme hydrologic events can be heavy tailed. For instance, Katz [52] obtained results indicating that the distribution of economic damage from hurricanes making landfall in the US (of which flooding is a major source of damage) has a heavy tail. It remains to determine the extent to which damage being heavy tailed is attributable to the underlying climate or hydrologic variables, as opposed to an inherent tendency of income or wealth. An example of

economic damage from floods will be treated in Section 4.2.1.

2.2.2. Annual and diurnal cycles

Because hydrologic variables can exhibit marked periodic behavior on both diurnal and annual time scales, naturally their extremes do as well. However, such cycles in extremes have not received much attention, as the block maxima technique does not require their explicit modeling. Still annual cycles in hydrologic extremes have been modeled in connection with devising a more efficient, indirect approach to estimating the distribution of annual maxima [11,78]. Like the POT and regional analysis approaches, the basic motivation is that monthly or seasonal maxima constitute additional information about the upper tail of the distribution. Although for some regions and certain seasons, diurnal cycles in precipitation amounts can be quite large [55], such features have generally not been introduced directly into the statistical modeling of extremes.

Another impetus for modeling annual cycles is related to the increased adoption of the POT approach instead of block maxima, necessitating the inclusion of any annual cycles. In fact, the early paper by Todorovic and Zelenhasic [94] included an application of the POT method to streamflow, allowing the rate of occurrence of exceedances of a high threshold to have an annual cycle and considering the possibility that the distribution of excesses over the threshold depends on the season as well (also see [50]). Using the POT approach to estimate the parameters, Smith [86,87] found annual cycles in both the location and scale parameters, μ and σ , of the GEV distribution in modeling extreme daily precipitation amount at a large number of sites across the US. An example of modeling the annual cycle in extreme daily precipitation amount will be given in Section 2.3.1.

2.2.3. Temporal and spatial dependence

Because hydrologic variables exhibit substantial dependence over a wide range of temporal and spatial scales, it should be anticipated that their extremes do as well. Temporal dependence enters into the POT method for estimating the distribution of annual maxima. Although there is some evidence that exceedances of high thresholds occur in clusters for hourly or daily precipitation, it is still somewhat inconclusive [2,20]. “Decustering” procedures (i.e., making use only of the single highest exceedance within a cluster) are routinely employed in applications of the POT approach to hydrologic time series to avoid the effects of dependence. In particular, Todorovic and Zelenhasic (p. 1642 in [94]) recommended that “When a flood hydrograph is a multiple peaked hydrograph . . . , only the largest peak is taken into consideration”.

Although it is clear that there must be some spatial dependence among hydrologic extremes at least over

small scales, it is rare to take such dependence explicitly into account. In Section 6.1, this issue will be discussed in conjunction with the topic of regional analysis of extremes.

2.3. Examples

To illustrate some of the characteristics typical of hydrologic extremes, two examples are treated: one concerned with estimating the distribution of the annual maximum of daily precipitation amount, another with estimating the distribution of annual peak streamflow.

2.3.1. Fort Collins maximum precipitation

Daily precipitation amounts at a single location (Fort Collins, CO, USA) are analyzed for the time period 1900–1999 (source: Colorado Climate Center, Colorado State Univ. <http://ulysses.atmos.colostate.edu>). Because of a flood that occurred on 28 July 1997, this data set is of special interest (although this particular rain gauge is not actually situated within the area that received the most intense precipitation) [72]. Fig. 1 shows the time series of annual maximum daily precipitation amount derived from the original daily data. When the GEV distribution is fitted by ML directly to the sample of 100 annual maxima (i.e., block maxima approach), some evidence of a heavy tail is obtained ($\hat{\gamma} = 0.174$, P -value ≈ 0.038 for likelihood ratio test of $\gamma = 0$).

As an alternative to the block maxima approach, the same GEV distribution for the annual maximum is fitted by ML, indirectly using the POT method, to the time series of daily precipitation amount at Fort Collins (with a relatively low threshold of 0.40 in. to avoid the complication of varying the threshold itself depending on the time of year; no declustering was applied). In this ap-

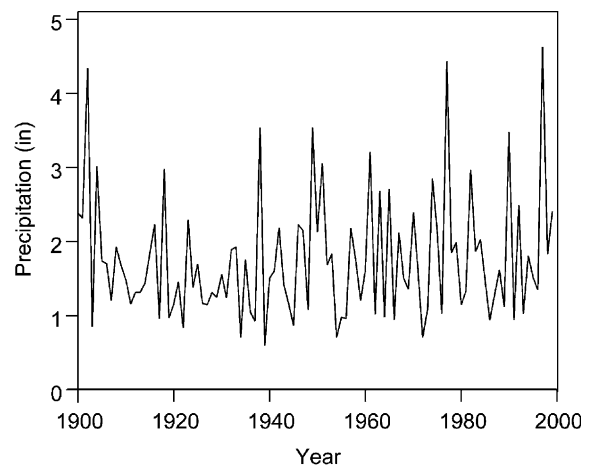


Fig. 1. Annual maximum of daily precipitation amount at Fort Collins, CO, USA, 1900–1999.

proach, annual cycles for the location and scale parameters, μ and σ , are explicitly modeled. Formally, it is assumed that the annual maximum of daily precipitation amount has a GEV distribution with parameters $\mu(t)$ and $\ln \sigma(t)$ (taking logarithm constrains the scale parameter to be positive), $t = 1, 2, \dots, 365$, possibly depending on particular day t within a given year through sine waves:

$$\begin{aligned}\mu(t) &= \mu_0 + \mu_1 \sin(2\pi t/T) + \mu_2 \cos(2\pi t/T), \\ \ln \sigma(t) &= \sigma_0 + \sigma_1 \sin(2\pi t/T) + \sigma_2 \cos(2\pi t/T).\end{aligned}\quad (3)$$

Here $T = 365\frac{1}{4}$ days and the shape parameter γ is taken constant (although an annual cycle in γ would be permissible as well).

Both the incorporation of an annual cycle in the location parameter $\mu(t)$ and the subsequent addition of an annual cycle in the log-transformed scale parameter $\sigma(t)$ produce large reductions in the minimized negative log-likelihood function (i.e., P -values near zero for likelihood ratio tests of $\mu_1 = \mu_2 = 0$ and $\sigma_1 = \sigma_2 = 0$ in Eq. (3); additional terms in Fourier series were not considered). The estimated shape parameter is nearly the same as for the block maxima method ($\hat{\gamma} = 0.182$), but the evidence for a heavy tail is much stronger because so much more data enters into the estimation procedure (P -value near zero for likelihood ratio test of $\gamma = 0$). Further, our method of analysis eliminates the possibility that this heavy tail is simply an artifact of annual cycles.

One way to portray the nature of the annual cycle in extremes is through an “effective” design value. This quantity has an interpretation similar to that for an ordinary design value (i.e., the quantile corresponding to a specified return period), except that it varies depending on the time of year. These estimated design values are obtained by substituting the parameter estimates into Eq. (2) for the quantile function of the GEV distribution. Fig. 2 shows the effective design value for a 100-yr return period (i.e., $p = 0.01$ in Eq. (2)), with the parameters of the GEV distribution being rescaled to reflect the maximum of daily precipitation amount over a month (average length = $(365.25)/12$ day) instead of a year. The design values range from a low in mid-January of about 1.1 in. to a high in mid-July of 4.3 in. (observed daily maximum over 100-yr period is 4.63 in. reflecting the flood in 1997). To give a rough feeling for the actual annual cycle in extreme precipitation, the observed monthly maximum of daily precipitation for each year in the 100-yr record is also included in Fig. 2. Consistent with the effective design values for the fitted GEV distribution, a marked tendency is evident toward higher precipitation extremes in summer than in winter.

It is also of interest to estimate the return period for the high precipitation event in July 1997 (i.e., 4.63 in.). With annual cycles in the parameters of the GEV dis-

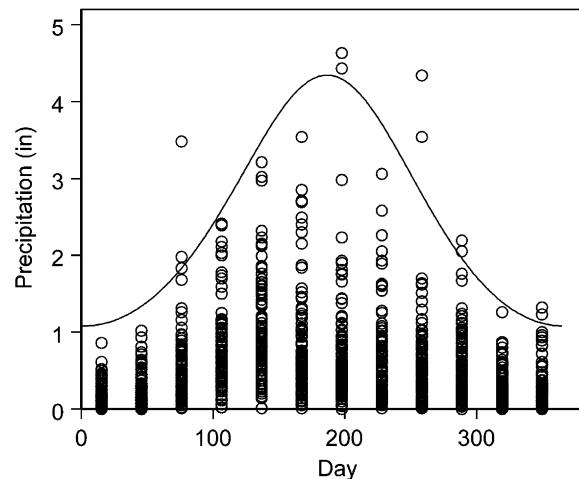


Fig. 2. Annual cycle in effective design value (100-yr return period) for fitted GEV distribution for monthly maximum of Fort Collins daily precipitation. Observed values of monthly maximum of daily precipitation indicated by circles.

tribution (Eq. (3)), the determination of a return period involves combining probabilities that differ depending on the day within a year (see Chapter 7 in [15]). Refitting only the data for the time period 1900–1996 (i.e., leaving out the 1997 flood event; the parameter estimates are not very sensitive to whether this event is included), the estimated return period for a daily amount of 4.63 in. is roughly 50.8 yr for the GEV distribution, as compared to 562.3 yr for the corresponding Gumbel distribution. Such a large difference in estimated return periods amply illustrates the effect of neglecting a heavy tail.

Details on the implementation of ML with covariates (such as the annual cycles in the parameters as specified in Eq. (3)) for the GEV distribution will be described in Section 3 and Appendix A.

2.3.2. Potomac River peak flow

Fig. 3 shows a relatively long time series of the annual peak instantaneous flow (1 cfs = $0.028317 \text{ m}^3/\text{s}$; water year October–September) of the Potomac River at Point of Rocks, MD, USA, for the time period 1895–2000 (source: US Geological Survey <http://water.usgs.gov/nwis/peak>). Smith [82] also analyzed the same time series, except that the available record then ended at 1986. He fitted the GEV distribution to the data by ML obtaining an estimated shape parameter of $\hat{\gamma} \approx 0.42$, but expressed concern that this distribution does not necessarily fit the upper tail of annual peak flow well.

When the GEV distribution is again fitted by ML to the longer time series of 106 annual maxima, fairly strong evidence of a heavy tail is obtained ($\hat{\gamma} = 0.191$, P -value ≈ 0.002 for likelihood ratio test of $\gamma = 0$). With the increased sample size and smaller estimated shape parameter compared to those obtained in [82], a quantile–quantile (Q – Q) plot for the GEV distribution

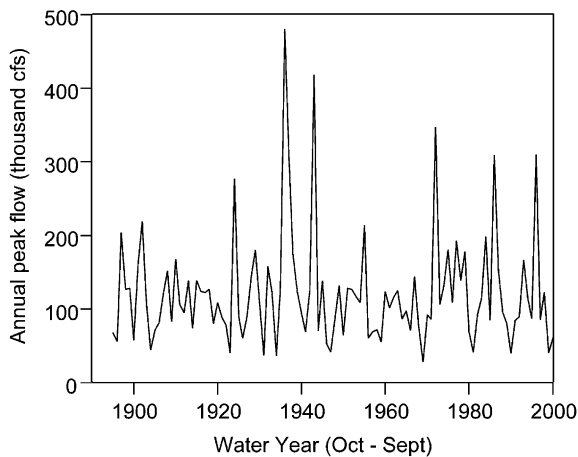


Fig. 3. Annual peak flow of Potomac River at Point of Rocks, MD, USA, 1895–2000.

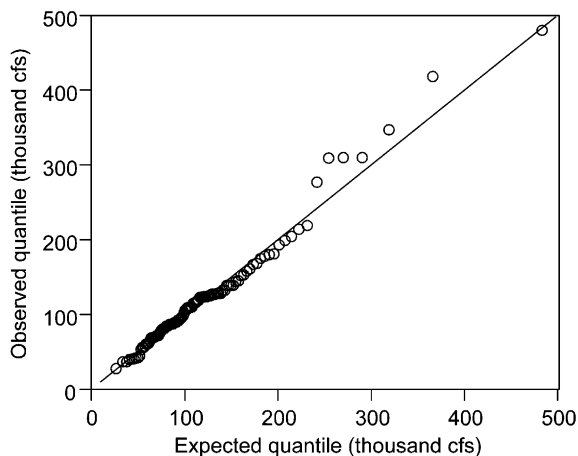


Fig. 4. $Q-Q$ plot for fit of GEV distribution to annual peak flow of Potomac River (line of equality indicates perfect fit).

(Fig. 4) indicates that the fit is reasonably adequate, even in the upper tail. In Section 5.2.2, another annual peak flow time series will be analyzed for which the fit of the GEV distribution does not appear to be acceptable.

3. Methodological developments

3.1. Theoretical framework

Underlying the POT method is a formal statistical model, consisting of a Poisson process for the occurrence of an exceedance of a high threshold and a generalized Pareto (GP) distribution for the excess over the threshold (termed “Poisson–GP model”). A basic reference for the point process representation of extremes is Leadbetter et al. (Chapter 5 in [59], also see Chapter 7 in [15]).

3.1.1. Poisson–GP model

Arising as an approximation for the distribution of excesses above a high threshold, the cumulative distribution and quantile functions for the GP are given by:

$$\begin{aligned} F(x; \sigma^*, \gamma) &= 1 - [1 + \gamma(x/\sigma^*)]^{-1/\gamma}, \\ \sigma^* &> 0, \quad 1 + \gamma(x/\sigma^*) > 0, \\ F^{-1}(1 - p; \sigma^*, \gamma) &= (\sigma^*/\gamma)(p^{-\gamma} - 1), \quad 0 < p < 1. \end{aligned} \quad (4)$$

Here σ^* and γ are the scale and shape parameters, respectively. The interpretation of the shape parameter γ is equivalent to that for the GEV distribution (e.g., if $\gamma > 0$, then the GP distribution is heavy tailed). By convention, $\gamma = 0$ refers to the limiting case obtained as $\gamma \rightarrow 0$ in Eq. (4) of the exponential distribution (i.e., an unbounded, thin tail).

Let $X_1, X_2, \dots, X_n$, denote a time series (assumed, for now, to be independent and identically distributed) whose high extreme values are of interest. The Poisson–GP model consists of two components (Chapter 4 in [15,25], Chapter 5 in [77]): (i) the occurrences of exceedances of some high threshold u (i.e., $X_i > u$, for some i) are generated by a Poisson process (with rate parameter λ); and (ii) the excesses over threshold u (i.e., $X_i - u$, for some i) have a GP distribution (with scale and shape parameters, σ^* and γ). The scale parameter σ^* of the GP distribution differs from that for the GEV by an amount depending on the threshold u (see Eq. (A.3) in Appendix A). As previously mentioned, the assumption of independence can be relaxed by dealing with cluster maxima instead of all exceedances, and one way to relax the assumption of identical distribution is by letting the parameters of the Poisson–GP model depend on covariates (e.g., annual or diurnal cycles).

3.1.2. Point process approach

Among others, Smith [85] developed the statistical theory needed to apply the point process approach to the statistics of extremes. In essence, this approach involves representing the two components of the Poisson–GP model (i.e., the occurrence of exceedances and the excesses over a high threshold) jointly as a two-dimensional nonhomogenous Poisson process (one dimension is time, the other the excess values). In this way, features of the GEV distribution for block maxima and the POT approach can be combined. In particular, the GEV distribution can be indirectly fitted via the POT method, but still in terms of the GEV parameterization. In this way, the scale parameter σ is invariant with respect to the choice of threshold u , the extension to time-dependent parameters (e.g., covariates) is immediate, and even thresholds that vary with time (e.g., because of annual cycles or trends) are permissible (Chapter 7 in [15]).

Suppose that it is desired to fit the GEV distribution, with parameters μ , σ , and γ , for the maximum over some time period denoted by $1/h$. In other words, the time

scaling constant h constitutes the length of the sampling frequency relative to the period over which the maximum is being taken (e.g., in the Fort Collins precipitation example of Section 2.3.1, $h = 1/365.25$ because the time series is daily and the annual maximum is being modeled; although to construct Fig. 2, $h = 12/(365.25)$).

It is convenient to view the data values as a two-dimensional vector (x, z) , where x denotes a value of the variable and z indicates whether or not this value exceeds the threshold u (i.e., $z = 1$ if $x > u$, $z = 0$ otherwise). Let $f(x, z; \mu, \sigma, \gamma)$ denote the joint density corresponding to the two-dimensional nonhomogenous Poisson process that generates these values (x, z) (but note that the distribution of x , when $x < u$, is not actually specified). Then this density can be expressed as (see Appendix A and [15,87]):

$$\begin{aligned} -\ln f(x, z; \mu, \sigma, \gamma) &= z\{\ln[\sigma + \gamma(u - \mu)] + (1/\gamma) \ln[1 + \gamma(u - \mu)/\sigma]\} \\ &\quad + z(1 + 1/\gamma) \ln\{1 + [\gamma(x - u)]/[\sigma + \gamma(u - \mu)]\} \\ &\quad + h[1 + \gamma(u - \mu)/\sigma]^{-1/\gamma}, \\ &\quad 1 + [\gamma(x - u)]/[\sigma + \gamma(u - \mu)] > 0. \end{aligned} \quad (5)$$

Let the bivariate random sample $\{(X_i, Z_i) : i = 1, 2, \dots, n\}$ correspond to the representation (x, z) above (again, we do not actually need to keep track of any value $X_i < u$). The ML technique involves evaluating Eq. (5) at each member of the sample (i.e., $-\ln f(X_i, Z_i; \mu, \sigma, \gamma)$, $i = 1, 2, \dots, n$), summing these n terms, and determining by numerical methods the values of the parameters, μ , σ , and γ , that minimize this sum. The parameters of the GEV distribution in Eq. (5) could actually depend on time or other covariates (e.g., as in the annual cycles specified by Eq. (3) for the Fort Collins example), expressed in full generality as $\mu(t)$, $\sigma(t)$, and $\gamma(t)$.

3.2. Parameter estimation

In the present paper, we make use exclusively of ML because of the need to fit structural models. Nevertheless, it has been known for a long time that the performance of ML can be extremely erratic for small samples (say $n \leq 25$), especially in the estimation of extreme quantiles of the GEV distribution (Eq. (2)). For instance, Martins and Stedinger (p. 739 in [66]) pointed out that “absurd values of the GEV shape parameter ... can be generated”. They provided an example in which a random sample of size $n = 15$ is generated from a GEV distribution with shape parameter $\hat{\gamma} = 0.2$, yet the ML technique yields an estimate of $\hat{\gamma} = 2.48$. Such small-sample behavior could not have been ruled out on a theoretical basis, because the primary justification for ML arises from its asymptotic properties (i.e., for large samples).

Given the need in applied hydrology for extreme quantile estimates even when the historical record is quite short, the poor performance of ML provided the impetus for the development of alternative estimation techniques. When the shape parameter γ of the GEV distribution is positive (i.e., heavy tail), Hosking et al. [43] showed that PWM estimators are superior to ML for small samples (e.g., $n = 15$ or 25) in estimating upper quantiles of the GEV (e.g., $1 - p = 0.9$ or 0.99). They attributed the poor performance of ML to a relatively few cases in which much larger estimates of the shape parameter than the true positive value are produced (i.e., consistent with the example in [66]). Hosking [40] discussed more generally the advantages of PWM over ML.

As mentioned earlier, PWM estimators are only defined for a GEV distribution with finite mean (i.e., shape parameter $\gamma < 1$). Although it has been argued [43] that this constraint is not restrictive in practice (e.g., for peak streamflow), it still appears that it is imposed primarily for pragmatic, not physical reasons. Because PWM estimators have not received much attention in the mathematical statistics literature (for one exception, see [70]), a theoretical explanation of their good performance has been lacking.

Recent work by Coles and Dixon [16] indicates that one source of the apparent superiority of PWM for small samples is related to its constraint of finite mean. If the same constraint (i.e., $\gamma < 1$) is imposed on ML (by the technique of penalized likelihood estimation in [16]), then the performance of ML is comparable, perhaps even superior to PWM for small samples. Similarly, Martins and Stedinger [66] constrained the shape parameter through a Bayesian prior distribution whose support has an upper bound of $\gamma = 1/2$ (i.e., constraint of finite variance) (Bayesian methods will be discussed in Section 6.2). In an attempt to avoid the limitations of both ML and PWM, Morrison and Smith [68] combined these two estimation methods. The question remains of whether it is really justified to impose a constraint such as finite mean (recall the quote by Gumbel in Section 2 about the issue of a distribution being unbounded). In particular, are the users of the GEV distribution in hydrologic applications aware that moments above some finite order would be infinite, no matter how small a positive value the shape parameter assumes?

3.3. Standard errors, resampling, and residuals

One advantage of the ML method is that approximate standard errors for estimated parameters and design values can be automatically produced, either via the information matrix (e.g., “Extremes” software [30]) or through profile likelihood (Chapter 3 in [15]). But like the parameter estimates themselves, such standard errors can be quite unreliable for small-sample sizes.

“Resampling”, an alternative approach for determining standard errors, is becoming increasingly popular [23,28]. A technique, such as the “bootstrap”, is used to manufacture “new” samples from the original one. The bootstrap generates new samples by drawing at random with replacement from the original sample. By refitting the estimator to a large number of bootstrap samples and calculating the standard deviation of these estimates across the samples, more realistic standard errors can be obtained. Some software for statistics of extremes can produce bootstrap standard errors and confidence intervals for parameters and quantiles (e.g., “Xtremes” software [77]). We note that the bootstrap can be applied to other estimation procedures (e.g., PWM), not just ML.

One difficulty with fitting more complex models, such as the GEV distribution with covariates, is that resampling procedures are based on the premise that the original sample consists of independent and identically distributed data. Generally, this issue can be circumvented by making use of the fitted model to convert the original data to generalized “residuals”. Then the resampling exercise is performed in terms of these residuals.

To be more specific, suppose that the random variable X_t has a GEV distribution (Eq. (1)) with time-dependent parameters, $\mu(t)$, $\sigma(t)$, $\gamma(t)$. By design, in this model the observations would not be identically distributed (rather, nonstationary). But the residual random variable ϵ_t , obtained by the transformation

$$\epsilon_t = \{1 + \gamma(t)[X_t - \mu(t)]/\sigma(t)\}^{-1/\gamma(t)}, \quad (6)$$

would be identically distributed (i.e., exponential with unit scale parameter). New samples could be generated through a multi-stage procedure: (i) convert the original data to residuals using Eq. (6) (with estimates for the model fit to the original data being substituted in place of the unknown parameters); (ii) produce new samples of residuals by the conventional bootstrap; and (iii) form new samples of data by converting the bootstrap residuals through the inverse of Eq. (6) [24]:

$$X_t = \mu(t) + \sigma(t)[\epsilon_t^{-\gamma(t)} - 1]/\gamma(t). \quad (7)$$

Essentially the same strategy is effective for producing diagnostic displays. For instance, $Q-Q$ plots (recall Fig. 4) are likewise designed for independent and identically distributed data. One could use Eq. (6) for the GEV distribution with time-dependent parameters, and then construct a $Q-Q$ plot of the empirical quantiles for the residuals as compared to those for the theoretical exponential distribution. Similarly, if the POT approach were being employed, one might be interested in checking how well the GP distribution with time-dependent parameters, $\sigma(t)$ and $\gamma(t)$, fit the excesses over a threshold (e.g., as an aid in determining whether the threshold is high enough). In this case, the excesses can

be transformed to an exponential distribution (with unit scale parameter) through

$$\epsilon_t = [1/\gamma(t)] \ln\{1 + \gamma(t)[X_t/\sigma(t)]\} \quad (8)$$

(for more detailed discussion of graphical procedures to examine the fit of extreme value models, see [15,87,88]).

4. Change in hydrologic extremes

The methodology for fitting extremal distributions by ML with time varying parameters has been described in Section 3. In addition, an example has already been treated in Section 2.3.1 (involving annual cycles in parameters, as opposed to a long-term trend). So everything is in place to fit trends in hydrologic extremes. For instance, a simple candidate model for trends in the GEV distribution (Eq. (1)) could involve linear trends in the location and log-transformed scale parameters (although a trend in the shape parameter would be permissible as well); that is,

$$\mu(t) = \mu_0 + \mu_1 t, \quad \ln \sigma(t) = \sigma_0 + \sigma_1 t, \quad \gamma(t) = \gamma, \quad (9)$$

where t denotes time (e.g., in units of days or years). It should be noted that some have questioned whether it is reasonable to assume linear trends as in Eq. (9), with nonparametric alternatives being feasible [38,75].

Notwithstanding the fact that the statistics of extremes are widely applied in hydrology, only rarely has this methodology been applied to detect trends in hydrologic extremes. Thus the evidence of trends to be cited is based primarily on methodology, such as least squares regression, that does not necessarily reflect properties of extremes like heavy tails. Moreover, because the statistics of extremes have not been relied on, any detected trends generally are not expressed in terms of design values or return periods (as desirable for water resources design and management).

4.1. Evidence of trends

As mentioned in the Introduction, it is anticipated that the hydrologic cycle would intensify as part of the enhanced greenhouse effect on global climate. In this subsection, the evidence of trends in precipitation and streamflow is reviewed.

4.1.1. Precipitation

A tendency toward increases in the frequency of extreme high precipitation has been detected in many areas of the world, including the US [49]. This effect has been generally measured in terms of trends in high quantiles of daily precipitation amount, but also shows up for the maximum daily precipitation amount over a month or season [48]. To more easily detect an effect, these analyses involved aggregation over a number of sites within

quite large regions (e.g., the US is divided into nine subregions). So it is unclear what such trends in precipitation extremes necessarily imply for design values or returns periods within smaller regions; in particular, water basins. Moreover, their relevance for high streamflow and floods has been questioned [14].

Information about predicted changes in extremes with global warming is also available from deterministic numerical models of the global climate system, known as general circulation models (GCMs). Zwiers and Kharin [102] compared GCM control run output for the present climate with GCM experiment run output for a doubling in the atmospheric concentration of carbon dioxide (CO_2). The GEV distribution was fitted by the method of PWM to the annual maxima of the synthetic daily precipitation amounts for both the present and doubled CO_2 climate at each of several thousand grid points. The estimated design values increase nearly everywhere across the world, but the differences are not necessarily statistically significant. In a follow-up study [56], the transient response of precipitation extremes, in which the CO_2 concentration is gradually increased instead of doubled, was examined. Because GCMs do not necessarily produce realistic precipitation extremes for the present climate, these results should be viewed with caution.

4.1.2. Streamflow

Peak streamflow, a variable that is closely related to precipitation extremes as well as other hydrometeorological factors (e.g., snowmelt), does not necessarily exhibit an increasing trend, notwithstanding other evidence in support of an intensified hydrologic cycle. For example, Robson et al. [79] were unable to detect any trend in annual peak flow for data pooled across the U.K. Similarly, Zhang et al. [101] found virtually no sites in Canada with increasing trends (but did detect some decreasing trends) in annual peak flow and high quantiles of daily mean flow. Although Lins and Slack [62] uncovered increasing trends in streamflow across much of the US, most of these trends were confined to low and moderate quantiles, not relevant for peak flow or flooding. Nevertheless, Groisman et al. [33] concluded that, given the well-documented increasing trends in extreme precipitation and the relationship between extreme precipitation and high flow, it is likely that increasing trends in high flow have occurred in the eastern US. On the other hand, they attribute the lack of trends in peak flow in the western US to the compensating effect of decreases in snow cover extent.

Other hydrometeorological variables should reflect any acceleration in the hydrologic cycle as well. For instance, Brutsaert and Parlange [8] pointed out that the limited data available (in some parts of the world) for evaporation are consistent with this hypothesis, with Szilagyi et al. [93] finding intensified actual evapotran-

spiration over the US. Ultimately, trends in hydrologic extremes should be seen in societal impacts as well. Although a marked increasing trend in flood damage in the US has occurred, Pielke and Downton [74] showed that it is primarily attributable to shifts in societal vulnerability, not necessarily any trends in precipitation extremes or streamflow. The trend in US flood damage will be illustrated in the example in Section 4.2.1. Regional analysis (Section 6.1) would constitute a more powerful approach to the detection of trends.

4.2. Examples

4.2.1. US flood damage

Fig. 5 shows the time series of annual total economic damage due to floods in the US for the time period 1932–1997 (for source of data, see [74]). The damages were totaled over the calendar year through 1982, thereafter over the water year (October–September), and have been adjusted for inflation (constant 1995 \$). An increasing trend, especially in the highest values, is clear from the plot of the raw time series. Much discussion has centered around whether this trend reflects a change in climate [74].

When the GP distribution is fitted by ML to this damage data set, with a linear trend in the log-transformed scale parameter (i.e., $\ln \sigma(t) = \sigma_0 + \sigma_1 t$), an estimated slope of $\hat{\sigma}_1 = 0.0275$ (corresponding to an increase in damage of about 2.79% per yr) is obtained (P -value $< 10^{-4}$ for likelihood ratio test of $\sigma_1 = 0$). There is no evidence of a heavy tail ($\hat{\gamma} = 0.029$), and a residual Q - Q plot (Fig. 6) using Eq. (8) to adjust for the trend in the scale parameter appears acceptable. Also included in Fig. 5 is the median of the GP distribution (i.e., quantile function given by Eq. (4) with $p = 0.5$) for the fitted trend in the scale parameter, increasing from about 0.64 to 3.82 billion \$ over the period of record.

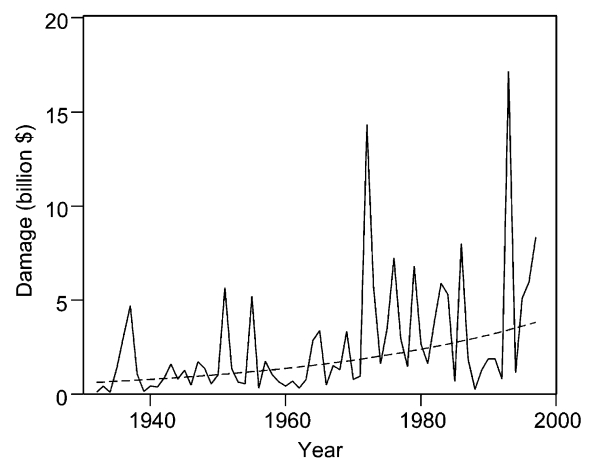


Fig. 5. US annual damage from floods, 1932–1997 (constant 1995 \$). Trend in median of fitted GP distribution given by dashed line.

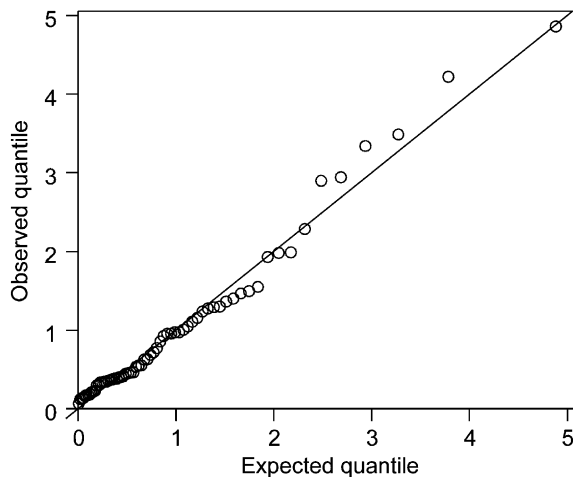


Fig. 6. Q - Q plot for fit of GP distribution to US annual damage from floods (line of equality indicates perfect fit).

Other functional forms of trend in the scale parameter could be considered, and the present linear one (in $\ln \sigma$) should be viewed as a simple starting point.

When Pielke and Downton [74] converted the data into damage per capita, there is still a statistically significant, but weaker increasing trend. For damage per unit wealth, the estimated trend is no longer even increasing. A more ideal, but also more elaborate, approach to the statistical modeling of such data would be predicated upon recognizing that total flood damage can be represented as a “random sum” (i.e., not only does the damage vary from flood to flood, but the number of floods varies from year to year as well). Katz [53] made use of this form of stochastic model in an application to economic damage caused by hurricanes.

4.2.2. Fremantle maximum sea level

Fig. 7 shows the time series of annual maximum sea level at Fremantle, Western Australia, for the time period 1897–1989 (source: Chapter 6 in [15] and <http://www.stats.bris.ac.uk/~masgc/isnev/datasets.html>). Only 86 yr of data are actually available (i.e., values for 1902, 1907, 1910–1911, 1924, and 1942 are missing). It should be noted that missing data pose no problem for ML in principle (in particular, there is no need to replace them with interpolated values). Perhaps an increasing trend is discernable from the plot of the raw time series, but it is not as obvious as in the flood data example just treated.

When the GEV distribution is fitted by ML to the sea level maxima, with a linear trend in the location parameter (i.e., $\mu(t) = \mu_0 + \mu_1 t$), an estimated slope of $\hat{\mu}_1 = 0.00232$ m per yr is obtained (P -value $< 10^{-3}$ for likelihood ratio test of $\mu_1 = 0$). Instead of being heavy tailed, the fitted GEV distribution has either a bounded tail or an unbounded, thin tail ($\hat{\gamma} = -0.125$, P -value ≈ 0.109 for likelihood ratio test of $\gamma = 0$). A Q - Q

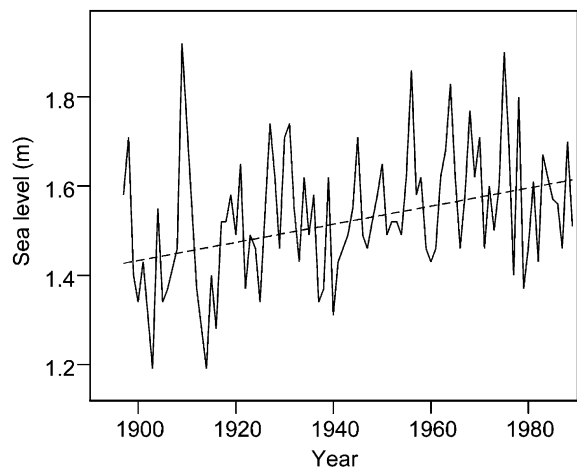


Fig. 7. Annual maximum sea level at Fremantle, Western Australia, 1897–1989. Trend in median of fitted GEV distribution given by dashed line.

plot (not shown) can be constructed on the basis of Eq. (6) to adjust for the trend in the location parameter and appears to be satisfactory. Also included in Fig. 7 is the median of the GEV distribution (i.e., quantile function given by Eq. (2) with $p = 0.5$) for the fitted trend in the location parameter, increasing from about 1.43–1.61 m over the period of record. Again, more complex forms of trend model for the location parameter could be considered.

Coles (Chapter 6 in [15]) introduced the Southern Oscillation into the model as an additional covariate, along with the trend term, on which the location parameter of the GEV distribution depends. This modification has little effect on the estimated slope of the trend term. The treatment of covariates within the framework of the statistics of extremes is the subject of Section 5.

5. Statistical downscaling of extremes

The term “statistical downscaling” [97] refers to empirical relationships between patterns in large-scale atmosphere–ocean circulation and smaller-scale climate (and related hydrologic) variables. In recent years, much attention has been devoted to this topic, with the impetus being that deterministic numerical models of the climate system predict large-scale patterns (e.g., of “smooth” variables such as atmospheric pressure) much better than they do regional or local weather or climate variables (especially, “erratic” variables like precipitation) [60,98].

Although the scope of statistical downscaling has included extreme events, generally the statistics of extremes have not been formally applied in this context (for one exception, see [6]). The dependence of the pa-

rameters of extremal distributions on covariates has occasionally been considered, but not in the context of downscaling (see Section 6.1). In most cases, the fitting of covariate relationships has been based on least squares regression, as opposed to ML for extremal distributions (for an exception, see [83]). Linking statistical downscaling and extreme value theory would be mutually beneficial. Replacing conventional regression analysis techniques with the statistics of extremes would make the treatment of downscaled extremes more statistically relevant, and the incorporation of covariates into the statistical models of extremes would also make this methodology more physically appealing.

5.1. Covariates

The appropriate methodology for the statistical downscaling of extremes can be achieved through the incorporation of covariates into the extremal distribution. As with the case of trends, the methodology for fitting extremal distributions by ML with time-dependent parameters (Section 3) applies. Unlike a deterministic trend variable, a covariate is itself a random variable. But by fitting the extremal distribution conditional on the values assumed by the covariate, the problem reduces to that of a time varying parameter. For instance, given the value of a covariate (say $Y = y$), the conditional distribution of the maximum could be assumed GEV with parameters $\mu(y)$, $\sigma(y)$, and $\gamma(y)$. A typical parameterization would be the same as in Eq. (9), replacing the time index t with the covariate value y . More generally, the covariate Y could actually be a vector (i.e., consisting of one or more covariates).

One natural candidate to serve as a covariate for hydrologic extremes would be the El Niño–Southern Oscillation (ENSO) phenomenon, the dominant mode in global climate variations on an annual time scale [73]. It has been associated with climate anomalies (such as droughts or floods) across large regions of the world. The ENSO phenomenon has also been directly linked to interannual variations in streamflow [27,47,76]. Although the relationship between extreme flows and ENSO has been occasionally examined (for low flows, see [69]; for high flows, see [45]), the statistics of extremes have not been used. More generally, both nonlinear dynamical theory and observational evidence support the existence of climate “regimes” [71,89]. So it might be anticipated hydrologic extremes ought to shift on at least an annual time scale in conjunction with patterns in large-scale atmosphere–ocean circulation. The term downscaling is appropriate because ENSO could be viewed as essentially a global phenomenon, whereas even streamflow for a large drainage basin reflects a much smaller than global scale.

5.2. Examples

To demonstrate that the statistics of extremes constitutes a viable methodology for the downscaling of extremes, two examples are treated: one concerned with estimating the distribution of the monthly maximum of daily precipitation amount conditional on monthly mean pressure, another with estimating the distribution of annual peak flow conditional on the state of ENSO.

5.2.1. Chico maximum precipitation

The maximum of daily precipitation amount for the month of January at Chico, CA, USA, for 78 yr is modeled (over time period 1907–1988, with 4 yr being eliminated because of missing values), with the covariate Y being the mean sea level pressure (mb – 1000) for January at a grid point (i.e., derived from observations within the grid box) off the Pacific Coast (40°N, 130°W). Fig. 8 shows a scatter plot of the January maximum of daily precipitation amount versus the pressure covariate, suggesting at least a weak inverse relationship.

Given a pressure value $Y = y$, the conditional distribution of the monthly maximum of daily precipitation is assumed GEV, with parameters related to the pressure value by

$$\mu(y) = \mu_0 + \mu_1 y, \quad \ln \sigma(y) = \sigma_0 + \sigma_1 y, \quad \gamma(y) = \gamma \quad (10)$$

(dependence of the shape parameter on pressure would be permissible as well). Katz and Parlange [54] originally analyzed this same precipitation and pressure data, but for conditional stochastic modeling of the time series of daily precipitation amounts (i.e., not just extremes). The following results can be viewed as an extension of those in Katz [51], in which extreme precipitation was modeled conditional on only two pressure states (below or above average).

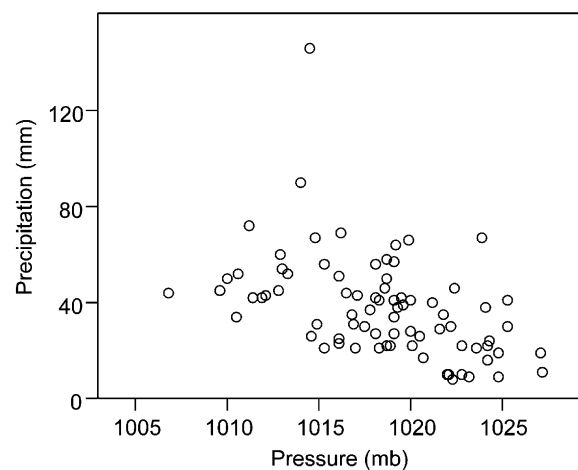


Fig. 8. Scatter plot of January maximum of daily precipitation amount at Chico, CA, USA, vs. January mean pressure.

Table 1

Fitted GEV distribution (by POT approach with threshold of 40 mm) to maximum daily precipitation (mm) in January at Chico, CA, USA, conditional on pressure covariate (mb – 1000)

Model	$\hat{\mu}_0$ (mm)	$\hat{\mu}_1$ (mm per mb)	$\hat{\sigma}_0$	$\hat{\sigma}_1$	$\hat{\gamma}$	$-\ln L$
$\mu_1 = 0, \sigma_1 = 0$	35.489	0	2.226	0	0.1984	244.10
$\mu_1 \neq 0, \sigma_1 = 0$	58.129	–1.361	2.315	0	0.1508	235.27
$\mu_1 \neq 0, \sigma_1 \neq 0$	58.146	–1.284	2.979	–0.045	0.1986	234.49

Table 1 gives the results of fitting by ML three possible conditional GEV distributions to the Chico precipitation maxima:

Model (i). No dependence of any parameters on pressure ($\mu_1 = 0, \sigma_1 = 0$ in (10));

Model (ii). Only location parameter depends on pressure ($\mu_1 \neq 0, \sigma_1 = 0$ in (10));

Model (iii). Both location and scale parameters depend on pressure ($\mu_1 \neq 0, \sigma_1 \neq 0$ in (10)).

Based on the POT approach with a threshold of $u = 40$ mm (because monthly maxima of daily data are involved, the time scaling constant is $h = 1/31$ in Eq. (5); no declustering was applied), the right-most column in Table 1 lists the negative of the maximized log likelihood function (labeled “ $-\ln L$ ”) for each of these three models. A likelihood ratio test for $\mu_1 = 0$ in Eq. (10) (i.e., model (i) vs. model (ii)) indicates strong evidence that the location parameter ought to be varied with pressure (P -value $< 10^{-4}$). The estimated slope parameter in model (ii) is $\hat{\mu}_1 = -1.361$ mm per mb, or higher precipitation extremes being associated with lower pressure (consistent with Fig. 8).

A likelihood ratio test for $\sigma_1 = 0$ in Eq. (1) (i.e., model (ii) vs. model (iii)) suggests only weak evidence that the scale parameter ought to be varied as well (P -value ≈ 0.209). There is also only weak evidence of Chico daily precipitation amount having a heavy-tailed distribution (model (ii): $\hat{\gamma} = 0.151$, P -value ≈ 0.160 for likelihood ratio test of $\gamma = 0$), most likely because only one month of data (i.e., January) is being analyzed. A residual $Q-Q$ plot (not shown) for the fitted GEV distribution in model (ii) based on Eq. (6) appears reasonably satisfactory.

To illustrate the magnitude of the effects of the pressure covariate on maximum precipitation, an “effective” return period is estimated on the basis of the best fitting model (ii). First, a design value is estimated for the fitted unconditional GEV distribution (i.e., model (i) with no conditioning on the pressure covariate) by Eq. (2) (e.g., 61.8 mm for a return period of 10 yr and 73.0 mm for 20 yr). Then the effective return period corresponding to this estimated design value is determined for the fitted GEV distribution (i.e., model (ii)), conditional on the value of the pressure covariate by Eq. (1). Given a 10-yr long-run return period, this effective return period varies from about 3.7 yr for the lowest

observed pressure to 23.6 yr for the highest; for a 20-yr return period, from about 8.1–44.9 yr.

5.2.2. Salt River peak flow

Fig. 9 shows the annual peak instantaneous flow (1 cfs = $0.028317 \text{ m}^3/\text{s}$; water year October–September) of the Salt River near Roosevelt, AZ, USA, for the time period 1924–1999 water year (1986 missing; source: US Geological Survey <http://water.usgs.gov/nwis/peak>). Streamflow at this location has been previously analyzed, with Anderson and Meerschaert [3] finding that the monthly mean flow is quite seasonal and possesses a heavy-tailed distribution (shape parameter $\gamma \approx 0.3$). Moreover, Dettinger and Diaz [27] detected statistically significant correlations between ENSO and monthly mean streamflow at many locations across the world, including higher flows being associated with El Niño events in the southwestern US. However, these studies did not focus on peak flow.

When the GEV distribution is fitted by ML to the Salt River peak flow data, a rather high estimated shape parameter is obtained ($\hat{\gamma} = 0.859$) and a $Q-Q$ plot (not shown) indicates very poor fit for the highest observations ($\hat{\gamma} = 0.677$ for penalized likelihood [16] and $\hat{\gamma} = 0.41$ for PWM, quite a bit smaller than ML but still relatively high values). Alternatively, a GP distribution is assumed instead, with ML yielding an estimated shape parameter of $\hat{\gamma} = 0.279$ (i.e., consistent with that obtained by Anderson and Meerschaert [3] for monthly

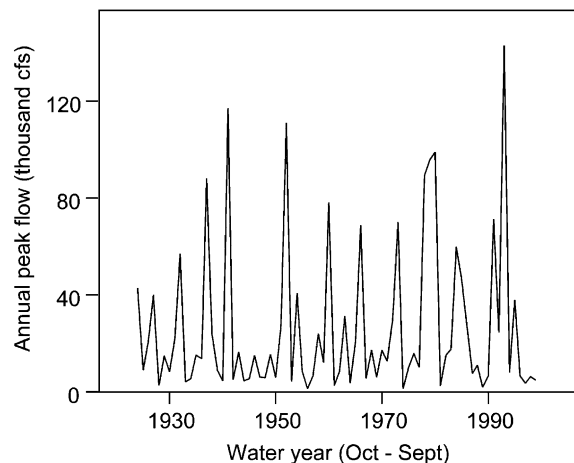


Fig. 9. Annual peak flow for Salt River near Roosevelt, AZ, USA, 1924–1999.

mean flow) and a $Q-Q$ plot (not shown) whose appearance is more satisfactory than that for the GEV distribution (irrespective of the estimation technique). Evidently, annual peak flow for the Salt River corresponds more closely to a single extreme event than to a maximum of a number of independent high values (the case for using the GP, instead of the GEV, distribution has been made by Smith [82] in conjunction with the Potomac River peak flow).

Next the Southern Oscillation is introduced as a covariate into the GP distribution. As an index of the Southern Oscillation, the fall (i.e., September–November) seasonal mean Darwin pressure (mb – 1000), overlapping with the start of the water season, is used (source: NOAA Climate Prediction Center <http://www.cpc.ncep.noaa.gov/data/indices/index.html>). Anomalously high Darwin pressure generally coincides with El Niño events. A scatter plot (not shown) of the peak flow versus Darwin pressure suggests at least a weak positive relationship (consistent with the results of Dettinger and Diaz [27]).

When the conditional GP distribution is fitted by ML to the Salt River peak flow, with the log-transformed scale parameter being linearly related to the Darwin pressure (i.e., $\ln \sigma(y) = \sigma_0 + \sigma_1 y$), the estimated slope of $\hat{\sigma}_1 = 0.0450$ (corresponding to an increase of about 4.61% per mb) is statistically significant (P -value ≈ 0.005 for likelihood ratio test of $\sigma_1 = 0$). In this model, the estimated shape parameter is smaller yet ($\hat{\gamma} = 0.156$) with only very weak evidence of a heavy tail (P -value ≈ 0.234 for likelihood ratio test of $\gamma = 0$). A residual $Q-Q$ plot (Fig. 10) for the GP distribution based on Eq. (8) appears satisfactory, even somewhat improved upon the corresponding $Q-Q$ plot without the Darwin pressure as a covariate. Given a 10-yr long-run return period, the effective return period varies from

about $2\frac{1}{2}$ to 101 yr over the historical range of the values assumed by Darwin pressure; for a 20-yr return period, from about $3\frac{1}{2}$ to 384 yr.

6. Future developments

In this section, the focus is on problems in hydrologic extremes whose solution would require advances in the methodology of the statistics of extremes. These problems include regional analysis of extremes, Bayesian quantification of uncertainty in extremal estimates, and statistics of multivariate hydrologic extremes. Finally, the issue of how to link the statistics of extremes with more conventional stochastic models of hydrologic time series is briefly discussed.

6.1. Regional analysis

Regional flood frequency analysis dates back many decades, revolving around the somewhat nebulous concept of an “index flood” [5,44]. In practice, it is assumed that if annual peak flows were scaled (e.g., dividing by the sample mean of annual peak flow for a particular site), then these index floods would have identical distributions across all sites within a given region. Estimation typically proceeds by averaging (e.g., weighted by number of observations at a site) individual site estimates (e.g., based on PWM) to obtain a single regional estimated flood distribution [42]. Local quantiles estimates can be obtained by applying the inverse operation to scaling (i.e., rescaling the quantile derived from the regional flood distribution). In practice, much evidence indicates that regional analysis produces more accurate site quantile estimates than those based on single site data alone [41]. Although regional analysis originated with floods in mind, the same approach has been applied to precipitation extremes [10]. Sometimes, covariates are incorporated into regional analyses. For example, Buishand [9] modeled the location and scale parameters of the Gumbel distribution for annual maximum snow depth as linear functions of the winter mean temperature; Smith [83] modeled the logarithm of the scale parameter of the GP distribution for high flows as a linear function of the basin size. More generally, the method of least squares (unweighted, weighted, or generalized) is used to relate site quantities (e.g., parameters or high quantiles of extremal distributions) to physical characteristics, such as catchment area and average annual precipitation (e.g., [63]).

Various limitations of regional analysis have been raised right from the start of its use [26], including the distortion that arises from dividing the peak flows by their sample mean instead of the unknown population mean [90,91], spatial heterogeneity of regions (i.e., the distribution of scaled flows not being identical across

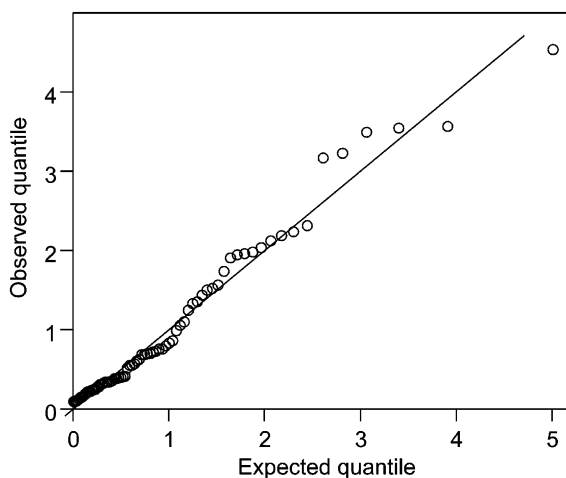


Fig. 10. $Q-Q$ plot for fit of GP distribution to annual peak flow of Salt River conditional on Darwin pressure (line of equality indicates perfect fit).

sites [61]), and spatial dependence of high flows [41,90]. Nevertheless, it is rare that regional analysis has been expressed in terms of a formal statistical model (e.g., including explicit assumptions about how the parameters of the GEV distribution vary within the region).

As an example of a formal model for regional analysis of extremes, consider the GEV distribution with parameters, $\mu(s)$, $\sigma(s)$, and $\gamma(s)$, where the two-dimensional index $s = (s_1, s_2)$ indicates the location of a site within a given region. In regional analysis, these parameters would not be estimated separately site by site, but jointly with the imposition of constraints on the parameters, such as $\mu(s)$ and $\sigma(s)$ varying with s but $\gamma(s) = \gamma$, for all s . Buishand [10] fit such a model by ML to annual maxima of daily precipitation amount at a number of sites, among other things obtaining stronger evidence of heavy tails than any single site analysis would permit. He also considered the case of the additional constraint of $\sigma(s)/\mu(s)$ being independent of s (i.e., consistent with the traditional form of index flood) (recently, Sveinsson et al. [91] fit the same form of model). Similarly, Smith [83] fit the GP distribution by ML, say with parameters $\sigma(s)$ and $\gamma(s)$, to high flows, likewise constraining the shape parameter (i.e., $\gamma(s) = \gamma$).

As an example of how regional analysis could be extended to incorporate a trend, let the time-dependent parameters of the GEV distribution be denoted by $\mu(s, t)$, $\sigma(s, t)$, and $\gamma(s, t)$, at site s and time t . A linear trend in the location parameter within the region could be represented as $\mu(s, t) = \mu_0(s) + \mu_1 t$. Here the intercept $\mu_0(s)$ depends on the site s , but the slope μ_1 is assumed constant within the region.

Why has this constrained ML approach not been applied more in practice for regional analysis? One reason is that the formulations to date have been predicated upon the assumption of spatial independence (for an exception, see [64]). Yet high flows are known to exhibit significant positive correlations across space (e.g., correlations of roughly 0.2 or higher [41]). While neglecting such dependence would have little, if any effect on point estimates of quantiles, it does result in an underestimation of the standard errors of such estimates. Nevertheless, even when spatial dependence is taken into account, a net gain in accuracy over single site analyses should be obtained.

Methodological approaches for dealing with spatial dependence of extremes are still in a rudimentary stage of development. Although some directions are promising, so far they leave something to be desired, either with respect to compromising the statistical theory of extremes or with respect to hydrologic realism. Within the hydrologic literature, one approach has been to examine the scaling properties of quantiles of peak flow, searching for some type of invariance principle as the basin size varies [36,67,84]. Such empirical scaling behavior is then used to infer the form of underlying sto-

chastic model (i.e., with the appropriate scaling properties). This approach does not make explicit use of the statistical theory of extremes; moreover, distributions such as the lognormal have been assumed that are not flexible enough for modeling extreme tails.

Within the statistics literature, models for how extremes ought to vary as a function of the spatial area have been devised through extension of extreme value theory [18]. The “max-stability” property, which can be viewed as an invariance principle for extremes, has been generalized from the univariate and multivariate cases to a form appropriate for spatial extremes. With the motivation being the extremal behavior of spatially aggregated precipitation, the model includes the property that the maximum value over time at an individual site within the region is approximately GEV distributed with parameters $\mu(s)$, $\sigma(s)$, and $\gamma(s)$. Under the simplifying assumption that the shape parameter is constant over all sites within the region (i.e., $\gamma(s) = \gamma$), the distribution of the maximum value over time of the areal-averaged process (say, averaged over region S with area A_S) is likewise approximately GEV with the same common shape parameter $\gamma(S) = \gamma$ but with location and scale parameters, $\mu(S)$ and $\sigma(S)$, that are not simply given by the spatial averages of the corresponding site values, $\bar{\mu} = (1/A_S) \int \mu(s) ds$ and $\bar{\sigma} = (1/A_S) \int \sigma(s) ds$, but also depend on a measure of the spatial dependence at high levels (somewhat analogous to the so-called “extremal index” that measures the temporal dependence of extremes [2]).

Before such an approach would be readily applicable to spatial hydrologic processes, the methodology needs to be further developed, including diagnostic checks on model assumptions. Still it would be useful to reconcile the scaling properties that arise from the statistics of extremes for spatial processes with those that have been empirically derived in hydrology.

6.2. Bayesian estimation

Consideration of Bayesian methodology might naturally arise in the estimation of hydrologic extremes for several reasons (Chapter 9 in [15]). As mentioned in Section 3.3, the large-sample approximate expressions for standard errors of ML estimates can be unreliable in practice. Moreover, in the estimation of events such as floods, historical information is often available that, while being of lower quality than more recent flow measurements, ought to be incorporated into the analysis. The Bayesian paradigm is well suited to deal with both of these circumstances. With recent advances in numerical methods for Bayesian estimation (i.e., Markov Chain Monte Carlo (MCMC)), the approach is feasible for a rich variety of model structures [13].

Yet the incorporation of Bayesian methods into the statistics of extremes has not been very prevalent so far,

with applications focused on hydrologic extremes being quite rare. Coles and Tawn [19], in a Bayesian analysis of extreme daily precipitation at a single site that makes use of prior information elicited from an expert hydrologist, obtained a 95% credible interval (i.e., Bayesian analogue to a confidence interval) for the 100-yr design value roughly half as wide as the corresponding confidence interval. Despite resistance to the use of historical information about floods [44], approximate Bayesian methods (termed “generalized” ML) that combine historical information with more recent high flow measurements have appeared in the hydrologic literature [46].

Bayesian estimation can also help deal with the spatial modeling of extremes, a topic just discussed in the context of regional analysis, enabling models to be fitted for which ML is intractable. Casson and Coles [13] employed the technique of MCMC to obtain Bayesian estimates for a model in which spatial dependence among high levels is induced solely by conditioning on a latent (“hidden”) process. For an application to high winds associated with hurricanes, as simulated by a numerical meteorological model, they showed that taking into account the spatial characteristics of extremes yields a substantial reduction in the length of confidence intervals for high quantiles. Ideally, actual observed covariates could be incorporated into such a statistical model (along with, or in place of, the latent process) to make the representation of spatial dependence more realistic.

6.3. Multivariate extremes

Within the last decade or so, the extension of the univariate statistics of extremes to the multivariate case has been well developed within the statistics literature (Chapter 8 in [15,17]). Yet the applications of the statistics of multivariate extremes to hydrology so far have been rather limited, in part because of the complexity of the theory. The central issue in multivariate extremes is how to measure the dependency among extreme levels of the variables involved [20]. This problem is somewhat analogous to that of temporal dependence at high levels in the case of univariate extremes.

As a compelling example of the need for a multivariate theory, consider the extreme event of the failure of a dike. As de Haan and de Ronde [37] pointed out, dike failure is a function two dependent variables, still water level and wave height, whose combination must be extreme for the dike to fail. Hydrologic applications have generally dealt with bivariate extremes, the most tractable situation. For example, Yue [100] used the Gumbel logistic model for bivariate extremes to model the joint distribution of annual maximum storm peak and total storm amount, observing that hydrologic engineering design and management requires more information about a storm than just its peak intensity. In a similar

vein, Adamson et al. [1] portrayed floods as “intrinsically multivariate random events, characterized not only by their peak flow, but also by their volume and the duration of discharge above critical thresholds” (p. 2825). Such information is needed to make realistic assessments of economic damage from floods.

6.4. Unified statistical modeling

Quite a large effort has been devoted to the development of rather elaborate stochastic models for hydrologic variables, such as precipitation or streamflow [80]. Yet such models are generally based on thin-tailed distributions (e.g., exponential, gamma, lognormal), and consequently fail to capture any heavy tails. For example, Rodríguez-Iturbe et al. [80] found that although a particular form of Poisson clustering (i.e., Bartlett–Lewis) model for precipitation has satisfactory aggregation properties, it still fails to produce a heavy enough tail for the distribution of the annual maximum of hourly or daily precipitation amounts. Similarly, Xu et al. [99] determined that although a chain-dependent Markov correlation pulse model for daily streamflow satisfactorily reproduces many short-term statistics, it still does not represent the distribution of annual peak flow well. One somewhat ad hoc approach to counteract this problem is to replace the upper tail of the model distribution with the GP (e.g., Cameron et al. [12] did so for hourly precipitation amount). Although such an approach is effective, it entails the estimation of additional parameters. Moreover, it is not clear how to extend this approach to allow for any temporal dependence of extremes (i.e., clustering of exceedances of a high threshold).

7. Discussion

The primary theme of this paper concerns how recent developments in the statistics of extremes can be applied not only to improve the rigor of hydrologic applications, but to make such analyses more physically meaningful. In terms of methodology, these developments primarily relate to ML estimation in the presence of covariates. From the perspective of hydrology, much remains to be done. In particular, we have not dealt with the issue of the physical basis for functional relationships between parameters of extremal distributions and covariates, as well as the physical interpretation of any such statistical models that are fitted to hydrologic extremes.

What has been done is to provide a more rigorous methodology by which trends in hydrologic extremes, with the anticipated intensification of the hydrologic cycle as part of global climate change, can be routinely incorporated in extreme value analyses. We have also linked the downscaling of hydrologic extremes with the statistical theory of extreme values. In the future, it is

anticipated that regional analysis of hydrologic extremes can be improved through use of constrained ML estimation, as well as explicit modeling of spatial dependence. It is also anticipated that uncertainties in the statistical estimation of hydrologic extremes will be more realistically quantified through the use of Bayesian methods. Finally, it is hoped that a unified statistical modeling approach can eventually be devised, in which the conventional stochastic models for hydrologic time series explicitly reflect what we know about the statistical characteristics of hydrologic extremes.

Acknowledgements

We thank Mary Downton for providing data on US flood damage, and three anonymous reviewers for providing helpful comments. This research was partially supported by NSF grant DMS-9815344 to the NCAR Geophysical Statistics Project. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Appendix A. Maximum likelihood estimation

Eq. (5) for the joint density of the occurrence of exceedances of a threshold and the amount of excess over the threshold follows directly from the point process representation. Here we only provide a heuristic argument (for a similar derivation, see Chapter 7 in [15]). The rate parameter of the Poisson process for the occurrence of exceedances of the high threshold u is now written as $h\lambda$, where the time scaling constant h is chosen so that λ corresponds to the rate for the time period over which the maximum is being taken. The distribution of the excess over u is GP with scale parameter σ^* and shape parameter γ on the (u, ∞) (i.e., distribution function in Eq. (4) with $x - u$ in place of x).

The joint density function $f(x, z; \lambda, \sigma^*, \gamma)$ can be expressed as:

$$\begin{aligned} -\ln f(x, z; \lambda, \sigma^*, \gamma) &= z\{\ln \sigma^* + (1 + 1/\gamma) \ln[1 + \gamma(x - u)/\sigma^*]\} \\ &\quad - z \ln \lambda + h\lambda, \quad 1 + \gamma(x - u)/\sigma^* > 0. \end{aligned} \quad (\text{A.1})$$

The first term in Eq. (A.1) is the negative log density function for an excess over threshold (i.e., differentiating GP distribution function in Eq. (4) with respect to x), multiplied by the indicator variable z (because an excess value only arises conditional on an exceedance having occurred). The remaining two terms in Eq. (A.1) constitute an approximate expression for the negative log probability function of the indicator variable z for the occurrence of threshold exceedances, except for a term

that does not depend on the parameter λ . That is, for the Poisson process governing occurrences:

$$-\ln[\Pr\{Z = z\}] \approx h\lambda - z \ln(h\lambda), \quad \text{as } h \downarrow 0. \quad (\text{A.2})$$

The relationships between the parameters, λ , σ^* , and γ , in the point process representation and the parameters, μ , σ , and γ , for the GEV distribution of the maximum is given by:

$$\ln \lambda = -(1/\gamma) \ln[1 + \gamma(u - \mu)/\sigma], \quad \sigma^* = \sigma + \gamma(u - \mu). \quad (\text{A.3})$$

Starting with the Poisson–GP model, an expression for the probability of no exceedances of the threshold u (i.e., equivalent to the maximum not exceeding u) can be derived, the GEV distribution function evaluated at u (but now expressed in terms of the parameters λ , σ^* , and γ). Equating the location and scale parameters (recall that the shape parameter is identical) gives Eq. (A.3). Reexpressing Eq. (A.1) in terms of the parameters of the GEV distribution, by means of Eq. (A.3), yields Eq. (5).

References

- [1] Adamson PT, Metcalfe AV, Parmentier B. Bivariate extreme value distributions: an application of the Gibbs sampler to the analysis of floods. *Water Resour Res* 1999;35:2825–32.
- [2] Ancona-Navarrete MA, Tawn JA. A comparison of methods for estimating the extremal index. *Extremes* 2000;3:5–38.
- [3] Anderson PL, Meerschaert MM. Modeling river flows with heavy tails. *Water Resour Res* 1998;34:2271–80.
- [4] Antal T, Droz M, Györgyi G, Rácz Z. $1/f$ noise and extreme value statistics. *Phys Rev Lett* 2001;87:240601–1–4.
- [5] Benson MA. Evolution of methods for evaluating the occurrence of floods. US Geological Survey water-supply paper 1580-A, 1962.
- [6] Bergström S, Carlsson B, Gardelin M, Lindström G, Pettersson A, Rummukainen M. Climate change impacts in Sweden—assessments by global climate models, dynamical downscaling and hydrological modelling. *Climate Res* 2001;16:101–12.
- [7] Brush GS. Rates and patterns of estuarine sediment accumulation. *Limnol Oceanogr* 1989;34:1235–46.
- [8] Brutsaert W, Parlange MB. Hydrologic cycle explains the evaporation paradox. *Nature* 1998;396:30.
- [9] Buishand TA. Statistics of extremes in climatology. *Stat Neerl* 1989;43:1–30.
- [10] Buishand TA. Extreme rainfall estimation by combining data from several sites. *Hydrolog Sci J* 1991;36:345–65.
- [11] Buishand TA, Demaré GR. Estimation of the annual maximum distribution from samples of maxima in separate seasons. *Stoch Hydrol Hydraul* 1990;4:89–103.
- [12] Cameron D, Bevin K, Tawn J. Modelling extreme rainfalls using a modified random pulse Bartlett–Lewis stochastic rainfall model (with uncertainty). *Adv Water Resour* 2001;24:203–11.
- [13] Casson E, Coles S. Spatial regression models for extremes. *Extremes* 1999;1:449–68.
- [14] Changnon SA. Comments on “Secular trends of precipitation amount, frequency, and intensity in the United States”. *Bull Am Meteorol Soc* 1998;79:2550–2.
- [15] Coles S. An introduction to statistical modeling of extreme values. London: Springer; 2001.
- [16] Coles SG, Dixon MJ. Likelihood-based inference for extreme value models. *Extremes* 1999;2:5–23.

- [17] Coles SG, Tawn JA. Modelling extreme multivariate events. *J Roy Stat Soc, Ser B* 1991;53:377–92.
- [18] Coles SG, Tawn JA. Modeling extremes of the areal rainfall process. *J Roy Stat Soc, Ser B* 1996;58:329–47.
- [19] Coles SG, Tawn JA. A Bayesian analysis of extreme rainfall data. *Appl Stat* 1996;45:463–78.
- [20] Coles S, Heffernan J, Tawn J. Dependence measures for extreme value analyses. *Extremes* 1999;2:339–65.
- [21] Cubasch U, Meehl GA. Projections of future climate change. In: Houghton JT et al., editors. *Climate change 2001: the scientific basis*. Cambridge: Cambridge University Press; 2001. p. 525–82.
- [22] Dalrymple T. Flood-frequency analyses, US Geological Survey water-supply paper 1543-A, 1960.
- [23] Davison AC, Hinkley DV. *Bootstrap methods and their application*. Cambridge: Cambridge University Press; 1997.
- [24] Davison AC, Ramesh NI. Local likelihood smoothing of sample extremes. *J Roy Stat Soc, Ser B* 2000;62:191–208.
- [25] Davison AC, Smith RL. Models for exceedances over high thresholds. *J Roy Stat Soc, Ser B* 1990;52:393–442.
- [26] Dawdy DR. Variation of flood ratios with size of drainage area, US Geological Survey professional paper, 424-C, C36, 1961.
- [27] Dettinger MD, Diaz HF. Global characteristics of stream flow seasonality and variability. *J Hydrometeorol* 2000;1:289–310.
- [28] Efron B, Tibshirani RJ. *An introduction to the bootstrap*. New York: Chapman and Hall; 1993.
- [29] Egozcue JJ, Ramis C. Bayesian hazard analysis of heavy precipitation in eastern Spain. *Int J Climatol* 2001;21:1263–79.
- [30] Faragó T, Katz RW. Extremes and design values in climatology, Report No. WCAP-14, WMO/TD-No. 386, World Meteorological Organization, Geneva, 1990.
- [31] Farquharson FAK, Meigh JR, Sutcliffe JV. Regional flood frequency analysis in arid and semi-arid areas. *J Hydrol* 1992;138:487–501.
- [32] Folland CK, Karl TR. Observed climate variability and change. In: Houghton JT et al., editors. *Climate change 2001: the scientific basis*. Cambridge: Cambridge University Press; 2001. p. 99–181.
- [33] Groisman PV, Knight RW, Karl TR. Heavy precipitation and high streamflow in the contiguous United States: trends in the Twentieth Century. *Bull Am Meteorol Soc* 2001;82:219–46.
- [34] Gumbel EJ. The return period of flood flows. *Ann Math Stat* 1941;12:163–90.
- [35] Gumbel EJ. *Statistics of extremes*. New York: Columbia University Press; 1958.
- [36] Gupta VK, Mesa OJ, Dawdy DR. Multiscaling theory of flood peaks: regional quantile analysis. *Water Resour Res* 1994;30:3405–21.
- [37] de Haan L, de Ronde J. Sea and wind: multivariate extremes at work. *Extremes* 1998;1:7–45.
- [38] Hall P, Tajvidi N. Nonparametric analysis of temporal trend when fitting parametric models to extreme-value data. *Stat Sci* 2000;15:153–67.
- [39] Hosking JRM. Maximum-likelihood estimation of the parameters of the generalized extreme-value distribution. *Appl Stat* 1985;34:301–10.
- [40] Hosking JRM. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J Roy Stat Soc, Ser B* 1990;52:105–24.
- [41] Hosking JRM, Wallis JR. The effect of intersite dependence on regional flood frequency analysis. *Water Resour Res* 1988;24:588–600.
- [42] Hosking JRM, Wallis JR. *Regional frequency analysis: an approach based on L-moments*. Cambridge: Cambridge University Press; 1997.
- [43] Hosking JRM, Wallis JR, Wood EF. Estimation of the generalised extreme-value distribution by the method of probability-weighted moments. *Technometrics* 1985;27:251–61.
- [44] Hosking JRM, Wallis JR, Wood EF. An appraisal of the regional flood frequency procedure in the UK Flood Studies Report. *Hydrol Sci J* 1985;30:85–109.
- [45] Jain S, Lall U. Floods in a changing climate: does the past represent the future? *Water Resour Res* 2001;37:3193–205.
- [46] Jin M, Stedinger JR. Flood frequency analysis with regional and historical information. *Water Resour Res* 1989;25:925–36.
- [47] Kahya E, Dracup JA. US streamflow patterns in relation to the El Niño/Southern Oscillation. *Water Resour Res* 1993;29:2491–503.
- [48] Karl TR, Knight RW. Secular trends of precipitation amount, frequency, and intensity in the United States. *Bull Am Meteorol Soc* 1998;79:231–41.
- [49] Karl TR, Knight RW, Plummer N. Trends in high-frequency climate variability in the twentieth century. *Nature* 1995;377:217–20.
- [50] Karr AF. Two extreme value processes arising in hydrology. *J Appl Probab* 1976;13:190–4.
- [51] Katz RW. Extreme value theory for precipitation: sensitivity analysis for climate change. *Adv Water Resour* 1999;23:133–9.
- [52] Katz RW. Techniques for estimating uncertainty in climate change scenarios and impact studies. *Climate Res* 2002;20:167–85.
- [53] Katz RW. Stochastic modeling of hurricane damage. *J Appl Meteorol* 2002;41:754–62.
- [54] Katz RW, Parlange MB. Effects of an index of atmospheric circulation on stochastic properties of precipitation. *Water Resour Res* 1993;29:2335–44.
- [55] Katz RW, Parlange MB. Generalizations of chain-dependent processes: application to hourly precipitation. *Water Resour Res* 1995;31:1331–41.
- [56] Kharin VV, Zwiers FW. Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere–ocean GCM. *J Climate* 2000;13:3760–88.
- [57] Lamoureux S. Five centuries of interannual sediment yield and rainfall-induced erosion in the Canadian High Arctic recorded in lacustrine varves. *Water Resour Res* 2000;36:309–18.
- [58] Landwehr JM, Matalas NC, Wallis JR. Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles. *Water Resour Res* 1979;15:1055–64.
- [59] Leadbetter MR, Lindren G, Rootzén H. *Extremes and related properties of random sequences and processes*. New York: Springer; 1983.
- [60] Lettenmaier D. Stochastic modeling of precipitation with applications to climate model downscaling. In: von Storch H, Navarra A, editors. *Analysis of climate variability: applications of statistical techniques*. Berlin: Springer; 1995. p. 197–212.
- [61] Lettenmaier D, Wallis JR, Wood EF. Effect of regional heterogeneity on flood frequency estimation. *Water Resour Res* 1987;23:313–23.
- [62] Lins HF, Slack JR. Streamflow trends in the United States. *Geophys Res Lett* 1999;26:227–30.
- [63] Madsen H, Rosbjerg D. Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modeling. *Water Resour Res* 1997;33:771–81.
- [64] Madsen H, Pearson CP, Rosbjerg D. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 2. regional modeling. *Water Resour Res* 1997;33:759–69.
- [65] Madsen H, Rasmussen PF, Rosbjerg D. Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events. 1. at-site modeling. *Water Resour Res* 1997;33:747–57.
- [66] Martins ES, Stedinger JR. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resour Res* 2000;36:737–44.
- [67] Morrison JE, Smith JA. Scaling properties of flood peaks. *Extremes* 2001;4:5–22.

- [68] Morrison JE, Smith JA. Stochastic modeling of flood peaks using the generalized extreme value distribution. *Water Resour Res*, in press.
- [69] Moss ME, Pearson CP, McKerchar AI. The Southern Oscillation index as a predictor of the probability of low streamflows in New Zealand. *Water Resour Res* 1994;30:2717–23.
- [70] Naveau P. Almost sure convergence of the maximum of a stationary sequence and asymptotic properties of probability weighted moments. PhD dissertation (unpublished), Colorado State University, Fort Collins, CO, 1998.
- [71] Palmer TN. A nonlinear dynamical perspective on climate prediction. *J Climate* 1999;12:575–91.
- [72] Petersen WA et al. Mesoscale and radar observations of the Fort Collins flash flood of 28 July 1997. *Bull Am Meteorol Soc* 1999;80:191–216.
- [73] Philander SG. El Niño, La Niña, and the Southern Oscillation. San Diego: Academic Press; 1990.
- [74] Pielke Jr RA, Downton MW. Precipitation and damaging floods: trends in the United States, 1932–97. *J Climate* 2000;13:3625–37.
- [75] Ramesh NI, Davison AC. Local models for exploratory analysis of hydrological extremes. *J Hydrol* 2002;256:106–19.
- [76] Redmond KT, Koch RW. Surface climate and streamflow variability in the western United States and their relationship to large-scale circulation indices. *Water Resour Res* 1991;27:2381–99.
- [77] Reiss R-D, Thomas M. Statistical analysis of extreme values. 2nd ed. Basel: Birkhäuser; 2001.
- [78] Revfeim KJA. Seasonal patterns in extreme 1-hour rainfalls. *Water Resour Res* 1982;18:1741–4.
- [79] Robson AJ, Jones TK, Reed DW, Bayliss AC. A study of national trend and variation in UK floods. *Int J Climatol* 1998;18:165–82.
- [80] Rodríguez-Iturbe I, Cox DR, Isham V. A point process model for rainfall: further developments. *Proc Roy Soc London, Ser A* 1988;417:283–98.
- [81] Shane RM, Lynn WR. Mathematical model for flood risk evaluation. *J Hydraul Engng* 1964;90:1–20.
- [82] Smith JA. Estimating the upper tail of flood frequency distributions. *Water Resour Res* 1987;23:1657–66.
- [83] Smith JA. Regional flood frequency analysis using extreme order statistics of the annual peak record. *Water Resour Res* 1989;25:311–7.
- [84] Smith JA. Representation of basin scale in flood peak distributions. *Water Resour Res* 1992;28:2993–9.
- [85] Smith RL. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. *Stat Sci* 1989;4:367–93.
- [86] Smith RL. Trends in rainfall extremes, 1999 (unpublished paper). Available at http://www.stat.unc.edu/faculty/rs/papers/RLS_Papers.html.
- [87] Smith RL. Extreme value statistics in meteorology and the environment. *Environmental Statistics*, 2001; Chapter 8:300–357 (NSF–CBMS conference notes). Available at <http://www.stat.unc.edu/postscript/rs/envstat/env.html>.
- [88] Smith RL, Shively TS. Point process approach to modeling trends in tropospheric ozone based on exceedances of a high threshold. *Atmos Environ* 1995;29:3489–99.
- [89] Smyth P, Ide K, Ghil M. Multiple regimes in Northern Hemisphere height fields via mixture model clustering. *J Atmos Sci* 1999;56:3704–23.
- [90] Stedinger JR. Estimating a regional flood frequency distribution. *Water Resour Res* 1983;19:503–10.
- [91] Sveinsson OGB, Boes DC, Salas JD. Population index flood method for regional frequency analysis. *Water Resour Res* 2001;37:2733–48.
- [92] Szilagyi J, Parlange MB. A geomorphology-based semi-distributed watershed model. *Adv Water Resour* 1999;23:177–87.
- [93] Szilagyi J, Katul GG, Parlange MB. Evapotranspiration intensifies over the conterminous United States. *J Water Resour Plan Manage* 2001;127:354–62.
- [94] Todorovic P, Zelenhasic E. A stochastic model for flood analysis. *Water Resour Res* 1970;6:1641–8.
- [95] Trenberth KE. Atmospheric moisture residence times and cycling: implications for rainfall rates and climate change. *Climatic Change* 1998;39:667–94.
- [96] Trenberth KE. Conceptual framework for changes of extremes of the hydrological cycle with climate change. *Climatic Change* 1999;42:327–39.
- [97] Von Storch H, Zorita E, Gubasch U. Downscaling of global climate change estimates to regional scales: an application to Iberian rainfall in wintertime. *J Climate* 1993;6:1161–71.
- [98] Wilby RL, Wigley TML, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS. Statistical downscaling of general circulation model output: a comparison of methods. *Water Resour Res* 1998;34:2995–3008.
- [99] Xu Z, Schumann A, Brass C, Li J, Ito K. Chain-dependent Markov correlation pulse model for daily streamflow generation. *Adv Water Resour* 2001;24:551–64.
- [100] Yue S. The Gumbel logistic model for representing a multivariate storm event. *Adv Water Resour* 2001;24:179–85.
- [101] Zhang X, Harvey KD, Hogg WD, Yuzyk TR. Trends in Canadian streamflow. *Water Resour Res* 2001;37:987–98.
- [102] Zwiers FW, Kharin VV. Changes in the extremes of the climate simulated by CCC GCM2 under CO₂ doubling. *J Climate* 1998;11:2200–22.