Hydrological Risk: recent advances in peak river flow modelling, prediction and real-time forecasting. Assessment of the impacts of land-use and climate changes
(Proceedings of the ESF LESC Exploratory Workshop held at Bologna, Italy, October 24-25 2003)

ON THE APPROPRIATENESS OF THE GUMBEL DISTRIBUTION IN MODELLING EXTREME RAINFALL

Demetris Koutsoyiannis

Department of Water Resources, Faculty of Civil Engineering, National Technical University of Athens, Heroon Polytechneiou 5, GR-157 80 Zographou, Greece (dk@itia.ntua.gr)

ABSTRACT

For half a century, the Gumbel distribution has been the prevailing model for quantifying risk associated with extreme rainfall. Several arguments including theoretical reasons and empirical evidence are supposed to support the appropriateness of the Gumbel distribution. These arguments are examined thoroughly in this work and are put into question. Moreover, it is shown that the Gumbel distribution may misjudge the hydrological risk as it underestimates seriously the largest extreme rainfall amounts. Besides, it is shown that the three-parameter extreme value distribution of type II is a more consistent alternative and it is discussed how this distribution can be applied even with short hydrological records.

1 INTRODUCTION

Almost a century after the empirical foundation of hydrological frequency curves known as "duration curves" (*Hazen*, 1914) and the theoretical foundation of probabilities of extreme values (*von Bortkiewicz*, 1922a, b; *von Mises*, 1923), and half a century after the convergence of empirical and theoretical approaches (*Gumbel*, 1958) the estimation of hydrological extremes continues to be highly uncertain. This has been vividly expressed by *Klemeš* (2000), who argues that

> "... the increased mathematisation of hydrological frequency analysis over the past 50 years has not increased the validity of the estimates of frequencies of high extremes and thus has not improved our ability to assess the safety of structures whose design characteristics are based on them. The distribution models used now, though disguised in rigorous mathematical garb, are no more, and quite likely less, valid for estimating the probabilities of rare events than were the extensions 'by eye' of duration curves employed 50 years ago."

Twenty years earlier, a similar critique was done be *Willeke* (1980; see also *Dooge*, 1986), who, among several common myths in hydrology, included the "Myth of the Tails", which reads

"Statistical distributions applied to hydrometeorological events that fit through the range of observed data are applicable in the tails",

and emphasises the fact that the tails of distributions are highly uncertain.

Among the most common probabilistic models used in hydrological extremes is the Gumbel distribution. This has been especially the case in modelling rainfall extremes. It is well known that estimation of rainfall extremes is very important for major hydraulic structures, given that design floods are generally estimated from appropriately synthesised design storms (e.g. *U.S. Department of the Interior, Bureau of Reclamation*, 1977, 1987; *Sutcliffe*, 1978).

Recently, several studies have shown that floods seem to have heavier tails than a Gumbel distribution (*Farquharson et al.*, 1992; *Turcotte*, 1994; *Turcotte and Malamud*, 2003). Other studies (*Wilks*, 1993; *Koutsoyiannis and Baloutsos*, 2000) have extended the scepticism for the Gumbel distribution to the case of rainfall extremes, showing that it underestimates the largest extreme rainfall amounts.

The adequacy of the Gumbel distribution for hydrological extremes, with emphasis of rainfall extremes, is the subject of this study. After a brief review of basic concepts of extreme value distributions (section 2) some theoretical arguments are provided, which explain that the Gumbel distribution is quite unlikely to apply to hydrological extremes (section 3). A case study based on a historical long hydrological record demonstrates how the Gumbel distribution may appear as an appropriate model for rainfall extremes while it is not (section 4). As alternative to the Gumbel distribution, the three-parameter extreme value distribution of type II is proposed and its use with typical short rainfall records is discussed (section 5).

2 BASIC CONCEPTS OF EXTREME VALUE DISTRIBUTIONS

It is recalled from probability theory that the largest of a number n of independent identically distributed random variables, i.e.,

$$X := \max \{Y_1, Y_2, \ldots, Y_n\} \tag{1}$$

has probability distribution function

$$H_n(x) = [F(x)]^n \tag{2}$$

where $F(x) := P\{Y_i \le x\}$ is the common probability distribution function of each of Y_i. Herein, $F(x)$ will be referred to as parent distribution. If n is not constant but rather can be regarded as a realisation of a random variable with Poisson distribution with mean v, then the distribution of X becomes (e.g. *Todorovic and Zelenhasic*, 1970; *Rossi et al.*, 1984),

$$H'_v(x) = \exp\{-v[1 - F(x)]\} \tag{3}$$

Since $\ln [F(x)]^n = n \ln \{1 - [1 - F(x)]\} = n \{-[1 - F(x)] - [1 - F(x)]^2 - \ldots\} \approx -n [1 - F(x)]$, it turns out that for large n or large $F(x)$, $H_n(x) \approx H'_n(x)$. Numerical investigation shows that even for relatively small n, the difference between $H_n(x)$ and $H'_n(x)$ is not significant (e.g., for $n = 10$, the relative error in estimating the exceedence probability $1 - H_n(x)$ from (3) rather than from (2) is about 3% at most).

In hydrological applications concerning the distribution of annual maximum rainfall or flood, it may be assumed that the number of values of Y_i (e.g., the number of storms or floods per year), whose maximum is the variable of interest X (e.g. the maximum rainfall intensity or flood discharge), is not constant. Besides, the Poisson model can be regarded as appropriate for such applications. Given also the small difference between (3) and (2), it can be concluded that (3) should be regarded as an appropriate model for every practical hydrological application.

However, the exact distributions (2) or (3), whose evaluation requires the parent distribution to be known, have not been used in hydrological statistics. Instead, hydrological applications have made wide use of asymptotes or limiting extreme value distributions, which are obtained from the exact distributions when n tends to infinity. *Gumbel* (1958), following the pioneering works by *Fréchet* (1927), *Fisher and Tippet* (1928) and *Gnedenco* (1941) developed a comprehensive theory of extreme value distributions. According to this, as n tends to infinity $H_n(x)$ converges to one of three possible asymptotes, depending on the mathematical form of $F(x)$ (*Gumbel*, 1958, p. 157). Obviously, the same limiting distributions may also result from $H'_\nu(x)$ as ν tends to infinity. All three asymptotes can be described by a single mathematical expression introduced by *Jenkinson* (1955, 1969) and become known as the General Extreme Value (GEV) distribution. This expression is

$$H(x) = \exp\left\{-\left[1 + \frac{\kappa\,(x-\varepsilon)}{\lambda}\right]^{-1/\kappa}\right\} \qquad \kappa\,x \geq \kappa\,\varepsilon - \lambda \tag{4}$$

where ε, $\lambda > 0$ and κ are location, scale and shape parameters, respectively. (Note that the sign convention of κ in (4) is opposite to that most commonly used in hydrological texts).

When $\kappa = 0$, the type I distribution of maxima (EV1 or Gumbel distribution) is obtained. Using simple calculus it is found that in this case, (4) takes the form

$$H(x) = \exp\{-\exp\,[-(x-\varepsilon)/\lambda]\} \tag{5}$$

which is unbounded from both left and right ($-\infty < x < +\infty$).

When $\kappa > 0$, $H(x)$ represents the extreme value distribution of maxima of type II (EV2). In this case the variable is bounded from the left and unbounded from the right ($\varepsilon - \lambda/\kappa \leq x < +\infty$). A special case is obtained when the left bound becomes zero ($\lambda = \kappa\,\varepsilon$). This special two-parameter distribution has been known as the Fréchet distribution and has the simplified form

$$H(x) = \exp\left\{-\left(\frac{\varepsilon}{x}\right)^{1/\kappa}\right\} \qquad x \geq 0 \tag{6}$$

with ε becoming a scale parameter.

When $\kappa < 0$, $H(x)$ represents the type III (EV3) distribution of maxima. This, however, is of no practical interest in hydrology as it refers to random variables limited to the right ($-\infty < x \leq \varepsilon - \lambda/\kappa$).

The simplicity of the above mathematical expressions is remarkable. This extends to the inverse function $x(H) \equiv x_H$ that is used to estimate a distribution quantile for a given non-exceedence probability H. This is

$$x_H = (\lambda/\kappa)\,[\exp(\kappa\, z_H) - 1] + \varepsilon \tag{7}$$

where z_H is the so called Gumbel reduced variate, defined as

$$z_H := -\ln(-\ln H) \tag{8}$$

For the Gumbel distribution, (7) takes the special form

$$x_H = \lambda\, z_H + \varepsilon \tag{9}$$

which implies a linear plot of x_H versus z_H (a plot known as the Gumbel probability plot). For the Fréchet distribution, (7) takes the form

$$x_H = \varepsilon \exp(\kappa\, z_H) \tag{10}$$

which implies a linear plot of $\ln x_H$ versus z_H (a plot referred to as the Fréchet probability plot).

Due to their simplicity and generality, the limiting extreme value distributions have become very widespread in hydrology, whereas the exact distributions (2) and (3) are used only rarely. In particular, EV1 has been by far the most popular model. In hydrological education is so prevailing that most textbooks contain the EV1 distribution only, omitting EV2. In hydrological engineering studies, especially those analysing rainfall maxima, the use of EV1 has become so common that its adoption is almost automatic, without any reasoning or comparing it with other possible models. There are several reasons for this:

1. **Theoretical reasons.** Most types of parent distributions functions that are used in hydrology, such as exponential, gamma, Weibull, normal, lognormal, and the EV1 itself (e.g. *Kottegoda and Rosso*, 1997, p. 431) belong to the domain of attraction of the Gumbel distribution. In contrast, the domain of attraction of the EV2 distribution includes less commonly met parent distributions like Pareto, Cauchy, log-gamma, and the EV2 itself.
2. **Simplicity.** The mathematical handling of the two-parameter EV1 is much simpler than that of the three-parameter EV2 (see also point 4 below).
3. **Accuracy of estimated parameters.** Obviously, two parameters are more accurately estimated than three. For the former case, mean and standard deviation (or second L-moment) suffice, whereas in the latter case the skewness is also required and its estimation is extremely uncertain for typical small-size hydrological samples.
4. **Practical reasons.** Probability plots are the most common tools used by practitioners, engineers and hydrologists, to choose an appropriate distribution function. As explained earlier, EV1 offers a linear Gumbel probability plot of observed x_H versus observed z_H (which is estimated in terms of plotting positions, i.e. sample estimates of probability of non-exceedence). In contrast, a linear probability plot for the three-parameter EV2 is not possible to construct. That is, equation (7) cannot be linearized; in fact the plot of x_H versus z_H is a curved convex function (unless $\kappa = 0$). This may be regarded as a primary reason of choosing EV1 against the three-parameter EV2 in practice. For the two parameter

EV2 (Fréchet) distribution, a linear plot (ln x_H versus z_H) is possible as discussed earlier. However, empirical evidence shows that, in most cases, plots of x_H versus z_H give more straight-line arrangements than plots of ln x_H versus z_H.

However, EV1 has one disadvantage, which is very important from the engineering point of view: For small probabilities of exceedence (or large return periods $T = 1/(1-H)$) it yields the smallest possible quantiles x_H in comparison to those of the three-parameter EV2 for any (positive) value of the shape parameter κ. This means that EV1 results in the highest possible risk for engineering structures. Normally, this would be a sufficient reason to avoid the use of EV1 in engineering studies. However, EV1 has been the prevailing model for rainfall extremes as discussed above.

Obviously, this disadvantage of EV1 would be counterbalanced only by strong empirical evidence and theoretical reasoning. In practice, the small size of common hydrological records (e.g. a few tens of years) cannot provide sufficient empirical evidence for preferring EV1 over EV2. This will be discussed further in section 4 using an appropriate real-world example. In addition, the theoretical reasons, exhibited in point 1 above, are not strong enough to justify the automatic adoption of the Gumbel distribution. This will be discussed in the section 3.

3 THEORETICAL STUDY OF THE APPROPRIATENESS OF THE GUMBEL DISTRIBUTION

To begin the theoretical discussion, it will be assumed that the events, whose maximum values are studied, can indeed be represented as independent identically distributed random variables Y_i (Assumption 1). Further, it will be assumed that the (unknown) parent distribution $F(y)$ belongs, with absolute certainty, to the domain of attraction of EV1 (Assumption 2). Are these rather oversimplifying and implausible assumptions sufficient to justify the adoption of EV1? The answer is clearly, No. This answer is demonstrated in Figure 1, which depicts Gumbel probability plots of the exact distribution functions of maxima $H_n(x)$ for $n = 10^3$ and 10^6 for two parent distribution functions. The first (upper panel) is the standard normal distribution and the second (lower panel) is the Weibull distribution ($F(y) = 1 - \exp(-y^k)$) with shape parameter $k = 0.5$. Both parent distributions belong to the domain of attraction of the Gumbel limiting distribution, so it is expected that the Gumbel probability plot tends to a straight line as $n \to \infty$. However, the tendency is remarkably slow, and even for n as high as 10^6 the curvature of the distribution functions is apparent. Obviously, in hydrological applications, such a high number of events within, say, a year, is not possible (it can be expected that the number of storms or floods in a location will not exceed the order of 10-10^2). Thus, the limiting distribution for $n \to \infty$ is not useful at all.

When studying storms and floods at a fine time scale, the parent distribution has typically a positively skewed, J-shaped density function. Thus, the normal distribution is not relevant in this case, but the Weibull distribution with shape parameter smaller than 1 (e.g. $k = 0.5$ as in the example of Figure 1) can be an appropriate parent distribution. In this case, it is observed in Figure 1, that the probability plots are convex functions, which indicates that, for a specified n, a three-parameter EV2 distribution may approximate sufficiently the exact distribution. Thus, even if the parent distribution belongs to the domain of attraction of the Gumbel distribution, an EV2 distribution can be a choice better than EV1.

Now, the Assumption 1 set above will be relaxed, forming the more plausible Assumption 1A. According to this, the events whose maximum values are studied are independent random variables Y_i but not identically distributed ones. Instead, it is assumed that all Y_i have the same type of distribution function $F_i(y)$ but with different parameters. This distribution function belongs to the domain of attraction of the Gumbel distribution, i.e., Assumption 2 is valid for each $F_i(y)$.

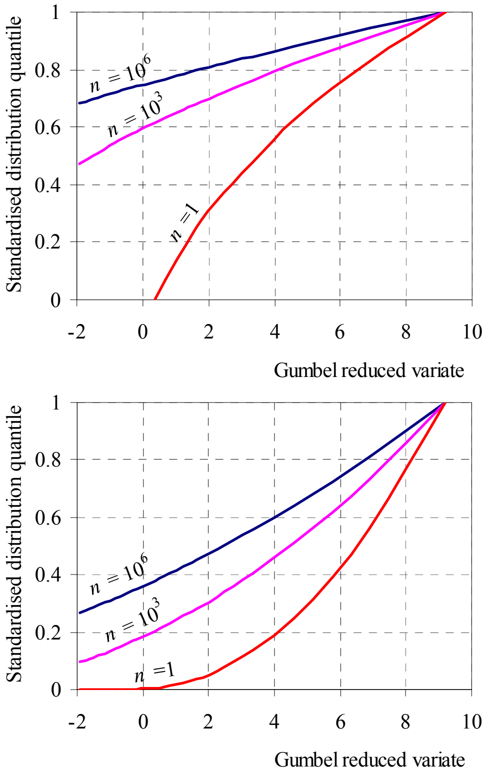

Figure 1. Gumbel probability plots of exact distribution function of maxima $H_n(x)$ for $n = 10^3$ and 10^6, also in comparison with the parent distribution function $F(y) \equiv H_1(y)$, which in the upper panel is standard normal and in the lower panel Weibull with shape parameter $k = 0.5$. The distribution quantile has been standardised by $x_{0.9999}$ corresponding to $z_H = 9.21$.

The relaxed assumption 1A is more consistent with hydrological reality. The statistical characteristics (e.g., averages, standard deviations etc.) and, consequently, the parameters of distribution functions exhibit seasonal variation. In addition, evidence from long geophysical records shows that there exist random fluctuations of the statistical properties on multiple large time scales (e.g., tens of years, hundreds of years, etc.). It has been proposed that these fluctuations constitute the physical basis of the well-known Hurst phenomenon (*Koutsoyiannis*, 2001).

The consequences of Assumption 1A are demonstrated by examples in which the parent distribution is specified to be the gamma distribution (which belongs to the domain of attraction of EV1) with varying scale parameter. More specifically, it may be assumed that during some ‘epoch’ (e.g. a specific month of a year through one or more years) the scale parameter is fixed to some value $\alpha_i > 0$. In this case, the probability density function of Y_i, conditional on α_i, is

$$f_i(y|\alpha_i) = \alpha_i^{\theta}\, y^{\theta - 1}\, \mathrm{e}^{-\alpha_i y} / \Gamma(\theta) \tag{11}$$

where the shape parameter $\theta > 0$ was kept constant for all ‘epochs’. In the first example it will be assumed that α_i varies randomly following a gamma distribution itself with scale parameter $\beta > 0$ and shape parameter $\tau > 0$, so that its density is

$$g(\alpha_i) = \beta^{\tau}\, \alpha_i^{\tau - 1}\, \mathrm{e}^{-\beta \alpha_i} / \Gamma(\tau) \tag{12}$$

If one is interested on the unconditional distribution of the variable Y, that is valid over all epochs, instead of a specified epoch, then one should determine from (11) and (12), the marginal density of Y, which is

$$f(y) = \int_0^{\infty} f_i(y|\alpha_i)\, g(\alpha_i)\, d\alpha_i = \{\beta^{\tau}\, y^{\theta - 1} / [\Gamma(\theta)\, \Gamma(\tau)]\} \int_0^{\infty} \alpha_i^{\theta + \tau - 1}\, \mathrm{e}^{-(y + \beta)\, \alpha_i}\, d\alpha_i \tag{13}$$

After algebraic manipulations it is obtained that

$$f(y) = \frac{1}{\beta\, B(\theta, \tau)} \frac{(y/\beta)^{\theta - 1}}{(1 + y/\beta)^{\tau + \theta}} \tag{14}$$

which shows that the marginal distribution of Y/β is beta of the second kind (*Kendal and Stuart*, 1963, p. 151; *Yevjevich*, 1972, p. 149). Consequently, the marginal probability distribution function of Y is

$$F(y) = B_{y/(y + \beta)}(\theta, \tau) / B(\theta, \tau) \tag{15}$$

where $B_z(\theta, \tau)$ and $B(\theta, \tau)$ denote respectively the incomplete beta function and the Euler (complete) beta function, i.e.,

$$B_z(\theta, \tau) := \int_0^{z} t^{\theta - 1} (1 - t)^{\tau - 1}\, dt, \quad B(\theta, \tau) := \int_0^{\infty} t^{\theta - 1} (1 - t)^{\tau - 1}\, dt \tag{16}$$

Thus, the exact distribution of maxima for constant and variable n is respectively

$$H_n(x) = [B_{x/(x + \beta)}(\theta, \tau) / B(\theta, \tau)]^n, \quad H'_v(x) = \exp\{-v[1 - B_{x/(x + \beta)}(\theta, \tau) / B(\theta, \tau)]\} \tag{17}$$

For $\theta = 1$, the parent distribution (15) simplifies to

$$F(y) = 1 - (1 + y/\beta)^{-\tau} \tag{18}$$

which is the Pareto distribution. Clearly, this belongs to the domain of attraction of EV2 with zero lower bound, i.e., the limiting distribution of maxima $H(x)$ is the Fréchet distribution. In the general case, it can be shown that

$$\lim_{y \to \infty} \frac{y\, f(y)}{1 - F(y)} = \tau > 0 \tag{19}$$

which is a sufficient condition for convergence of $H_n(x)$ to the EV2 distribution (e.g. *Kottegoda and Rosso*, 1997, p. 430).

In Figure 2 it is demonstrated how the exact distribution tends to the Fréchet distribution as n increases. In this case the shape parameter θ was assumed 0.5 and the exact distribution was calculated from (17). For n as high as 1000 the Fréchet probability plot becomes almost a straight line. However, as in the cases of Figure 1, for smaller values of n, which are more relevant in hydrological applications, the Fréchet plot of the exact distribution appears to be curved, so a three-parameter EV2 would yield a better approximation to the exact distribution than the two-parameter Fréchet distribution. (It is noted that the concave curvature appearing in the Fréchet plot of Figure 2 would be convex in a Gumbel plot).

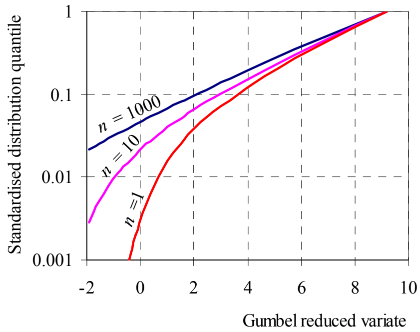

Figure 2. Fréchet probability plot of the exact distribution function of maxima $H_n(x)$ for $n = 1$, 10 and 1000, as this results assuming a gamma parent distribution with shape parameter $\theta = 0.5$ and scale parameter randomly varying following a second gamma distribution with shape parameter $\tau = 3$ and scale parameter $\beta = 1$. The distribution quantile has been standardised by $x_{0.9999}$ corresponding to $z_H = 9.21$.

A more specific numerical experiment is depicted in Figure 3. Here the exact distributions of maxima $H_5(x)$ (for $n = 5$), based on assumptions 1 and 1A, are compared. In case 1A, a variable parameter gamma distribution was assumed, with parameters $\theta = 0.5$, $\tau = 5$ and $\beta = 1$. In case 1, the variable scale parameter is replaced by a constant parameter $\alpha = \tau/\beta = 5$ (equal to the mean of the scale parameter of case 1A). The exact distribution of maxima for case 1 is almost a straight line on the Gumbel probability plot whereas that of case 1A is a convex curve. In addition to the theoretical distribution functions, empirical ones were also plotted, based on 4000 synthetic maxima. To these synthetic data series the EV1 and the three-parameter EV2 (GEV) distributions were fitted and were also plotted in Figure 3. As expected, EV1 is in good agreement with the exact distribution of case 1 but departs significantly from the exact distribution in case 1A, especially in the tail that corresponds to large return periods. In contrast, the three-parameter EV2 (GEV) is almost indistinguishable from the exact distribution.

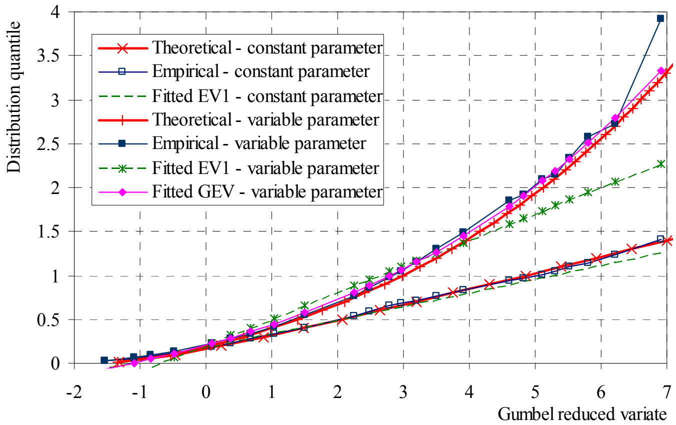

Figure 3. Gumbel probability plot of exact distribution function of maxima $H_5(x)$, as this results assuming a gamma parent distribution with shape parameter $\theta = 0.5$ and scale parameter either constant $\alpha = 5$ (case 1) or randomly varying following a second gamma distribution with shape parameter $\tau = 5$ and scale parameter $\beta = 1$ (case 1A). The additional plotted curves are empirical distribution functions from synthesised series of length 4000, and fitted to these series EV1 and EV2 distribution functions.

A second simpler example was based again on gamma parent distribution function with constant shape parameter $\theta = 0.5$ and scale parameter shifting between two values, $\alpha_1 = 2$ and $\alpha_2 = 6$ which are sampled at random with probabilities 0.25 and 0.75, respectively. For comparison, a gamma distribution with constant parameter $\alpha = 5$ (again equal to the mean of α_1 and α_2) was used. Here, the theoretical distributions were not determined but rather empirical ones were plotted, based on 4000 synthetic maxima.

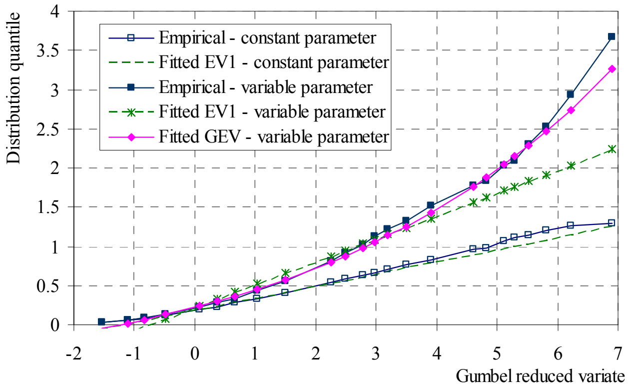

Figure 4. Gumbel probability plot of the empirical distribution functions of maxima $H_5(x)$ and fitted EV1 and EV2 distribution functions, as they result from synthesised series of length 4000 assuming gamma parent distribution with shape parameter $\theta = 0.5$ and scale parameter either constant $\alpha = 5$, or shifting at random between the values $\alpha_1 = 2$ and $\alpha_2 = 6$ with probabilities 0.25 and 0.75, respectively.

To these synthetic data series, the EV1 and the three-parameter EV2 (GEV) distributions were fitted and were also plotted in Figure 4. As in Figure 3, EV1 is in good agreement with the empirical distribution of the constant parameter case but departs significantly from the empirical distribution of the variable parameter case. Again, the departure is greatest in the tail, i.e. in large return periods. In contrast, EV2 (GEV) agrees well with the simulated distribution.

All this theoretical discussion and the examples show that the theoretical reasons, which have endorsed the use of the Gumbel distribution for hydrological extremes, are not strong enough to compensate the high risk it implies. In addition, they demonstrate that the GEV distribution bounded from the left (i.e. the three-parameter EV2) is a much better choice in comparison with the Gumbel distribution.

4 EMPIRICAL STUDY OF THE APPROPRIATENESS OF THE GUMBEL DISTRIBUTION

In this section it will be demonstrated how weak the empirical evidence that supports the choice of the Gumbel distribution as an appropriate distribution for rainfall maxima may be. As a case study, an annual series of maximum daily rainfall in Athens, Greece, extending through 1860-1995 (136 years), is used. This is the longest rainfall record available in Greece and its analysis was done by *Koutsoyiannis and Baloutsos* (2000). The annual series is shown graphically in Figure 5.

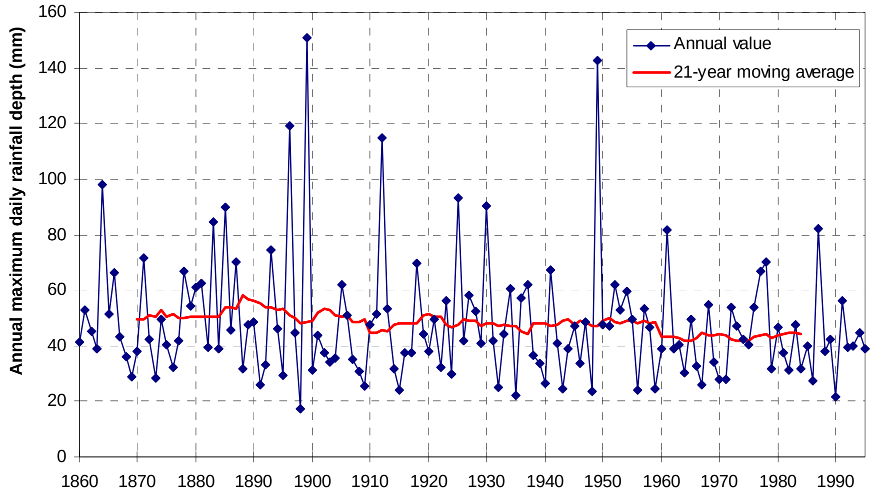

Figure 5. Plot of the time series of the annual maximum daily rainfall depth at Athens, Greece (from *Koutsoyiannis and Baloutsos*, 2000).

The empirical distribution function estimated using the Gringorten plotting positions is shown in the Gumbel probability plot of Figure 6. Clearly, this figure shows that the EV1 distribution departs significantly from the empirical distribution, whose points form a convex plot. The inappropriateness can be verified by a statistical test based on the L-moments estimator of κ. As shown by *Hosking et al.* (1985; see also *Stedinger et al.*, 1993, p. 18.18) when m data values are drawn from an EV1 distribution (κ = 0) this estimator has mean 0 and variance $0.5633/m$. This allows the construction of a test whether κ = 0 (i.e., appropriateness of the EV1 distribution; null hypothesis) or not (alternative hypothesis). Applied to the data of the case study, the test results in rejection of the null hypothesis at an attained significance level (i.e., probability of type I error) as low as 0.2%.

On the contrary, in Figure 6 the GEV distribution with positive shape parameter (κ = 0.185) fits well the empirical distribution. Furthermore, the goodness of fit of the GEV distribution with its three parameters estimated by the method of L-moments was tested using the χ^2 test, applied several times with a number of classes varying from 5 to 20. In all cases the null hypothesis (that the GEV distribution is consistent with the data) was not rejected at the typical 5%-10% (even at a non-typical 40%-50%) significance level. So, these statistical analyses provide evidence that the GEV distribution is a consistent probabilistic model for the annual maximum series under study, whereas the EV1 model is inconsistent with the data.

However, a record length of 136 years is exceptionally unusual; typically, sample sizes vary between 10 and 50 years. Thus, the question arises whether a sample of this typical small size can lead to conclusions similar with those of the 136-year record, or it delineates a different (and misleading) picture of the distribution function of maximum rainfall.

To answer this question, four sub-series, each corresponding to one quarter of the record length (34 years) were analysed. Figure 6 depicts, in addition to the empirical distribution of the complete series, the Gumbel probability plot of the empirical distribution of the fourth sub-series (last 34 years).

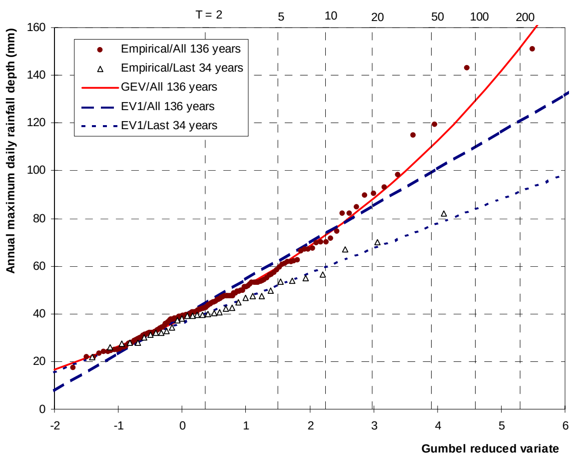

Figure 6. Comparison of the distribution function of the complete series (136 years) of annual maximum rainfall depths at Athens, Greece, and the fourth sub-series corresponding to one quarter (last 34 years) of the record length (from *Koutsoyiannis and Baloutsos*, 2000).

It is observed that, in this case, the empirical distribution forms a straight-line plot, which implies the appropriateness of the EV1 distribution for this sub-series. This EV1 distribution fitted by the method of L-moments for the fourth sub-series is also shown in Figure 6. The EV1 distribution of the 34-year sample departs significantly, particularly in the upper tail, from those of both the GEV and the EV1 distributions of the 136-year sample. Thus, the picture acquired from the sub-series of the last 34 years is deforming: the inappropriate EV1 distribution appears as appropriate and also shifted towards lower values of rainfall amounts in the upper tail.

Similar are the results for two other sub-series. In summary, the EV1 distribution tested by the κ-test described above is not rejected for the three out of four sub-series. Only the second sub-series results in a high value of the shape parameter κ, and consequently, a statistically significant departure from the EV1 distribution.

One may wonder whether this result is a peculiarity of the examined maximum rainfall record or it is a generalised behaviour of small versus large sample sizes, i.e., a purely statistical effect. Clearly, the answer is the latter. To demonstrate this, simulation experiments were performed, assuming that the true distribution of maximum rainfall is the GEV distribution with papameters equal to those estimated from the available 136-year record by the method of L-moments. With this assumption, 1000 34-year synthetic records were generated. For each of the 1000 records the κ-test was applied. Only in 241 out of 1000 cases (24.1% or, roughly, one out of four cases) the EV1 distribution was rejected, a figure quite the same with that already found from the analysis of the historic record. Note that the percentage 100% – 24.1% = 75.9% expresses the Type II error of the test, i.e., the probability of not rejecting a false null hypothesis. When this simulation experiment was repeated with 1000 136-year synthetic records, the Type II error fell off to 23.4%.

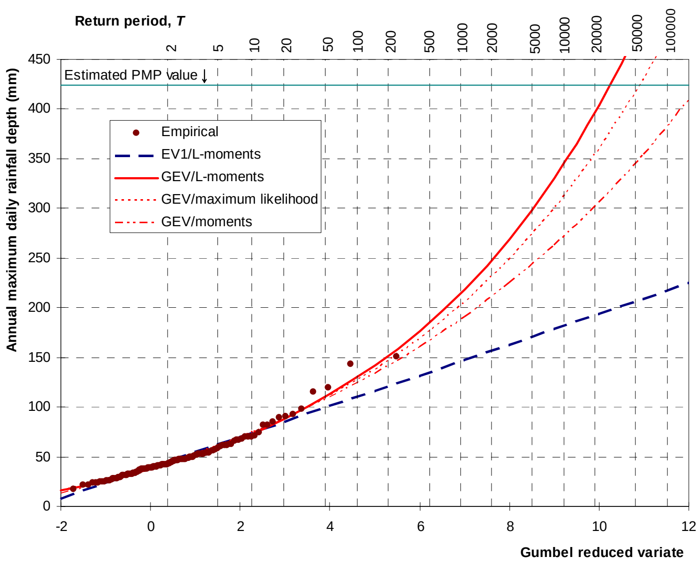

Figure 7. Plots of theoretical distribution functions in the area of low probabilities of exceedance, and comparison with the empirical distribution of the maximum daily rainfall in Athens, Greece, and the estimated PMP value (adapted from *Koutsoyiannis and Baloutsos*, 2000).

The findings of this investigation show that the empirical evidence supporting the wide applicability of the Gumbel distribution may in fact be the result of too small sample sizes rather than a manifestation of the real behaviour of rainfall maxima. To acquire an idea of the implications of an improper adoption of the EV1 distribution, this distribution along with the three-parameter EV2 (GEV) distribution fitted by different methods, was re-plotted in Figure 7, where emphasis was given to the tail of the distribution, for probabilities of exceedance less than 1/200. Clearly, the EV1 distribution, even though estimated from the complete 136-year record, underestimates seriously the maximum rainfall for small probabilities of exceedance. For instance, at the return period 10 000 years the EV1 distribution results in a value of rainfall depth half that obtained by the GEV distribution. It is noted that 10 000 years is not an unusual return period for the design of major flood protection works; for example most dam spillways in Greece were designed adopting this value.

5 RECOVERY FROM THE HIGH-RISK ESTIMATIONS OF THE GUMBEL DISTRIBUTION

The above discussion showed that the EV1 distribution may underestimate significantly hydrological extremes. Obviously, the uncertainty of the estimations of extremes cannot be eliminated. However, a proper step to avoid underestimation of extremes is to replace the EV1 distribution in typical applications with the three-parameter EV2 distribution (GEV bounded from the left). However, in applications involving small sample sizes, the three parameters of the latter may be a serious problem, given the well known disability of reliable parameter estimation of three-parameter distributions. The solution to this problem can be the so-called "substitution of space for time" (*National Research Council*, 1988), that is, the incorporation in the analysis of information from other rainfall data sets from other locations in the same region.

Although one must recognise that local factors play an important role in the distribution of rainfall extremes, results from global studies provide some guidance for a general behaviour. To the author's knowledge, the most comprehensive global study of rainfall extremes is the old work by *Hershfield's* (1961, 1965), which offered a basis for statistical estimation of probable maximum precipitation (PMP). Hershfield initially analyzed a total of 95 000 station-years of annual maximum rainfall belonging to 2645 stations, of which about 90% were in the USA. He standardised the maximum rainfall depth defining the standardised variate $k := (x - \mu)/\sigma$, where μ and σ are respectively (estimates of) the mean the standard deviation, and he found that the maximum observed value of k was 15. Then, he concluded that an estimate of the PMP amount can be determined by simply setting $k = 15$. Subsequently, *Hershfield* (1965) proposed that the maximum k varies with the rainfall duration d and the mean μ. More specifically, he found that the value of $k = 15$ is too high for areas with heavy rainfall and too low for arid areas, whereas it is too high for rain durations shorter than 24 hours. Therefore, he constructed an empirical nomograph indicating that the maximum k varies between 5 and 20.

Koutsoyiannis (1999) revisited on a probabilistic basis the analysis of the data published by Hershfield and concluded that they do not suggest an upper limit for k. Rather, they suggest a GEV distribution of the k values with shape parameter for the union of all records $\kappa = 0.13$, as shown in Figure 8. From this, it follows that the maximum observed value k = 15 is statistically expected for 95 000 data values. The conclusions of this study may be summarised as follows:

(1) The GEV distribution is appropriate for annual maximum rainfall series.
(2) The value of the standardised annual maximum rainfall k = 15, which was considered by Hershfield as defining PMP, corresponds to a return period of about 60 000 years.
(3) The shape parameter κ of the GEV distribution is given as a function of the mean value of annual maximum daily rainfall series μ, by

$$\kappa = \max(0.183 - 0.00049\,\mu, 0) \qquad (\mu \text{ in mm}) \tag{20}$$

The empirical equation (20) may be used to estimate the shape parameter of the GEV distribution thus avoiding the involvement of the highly uncertain (for short records) coefficient of skewness and making the estimation procedure of the three-parameter GEV similar to that of a two-parameter distribution like Gumbel. The fact that no straight line plot can be constructed for the GEV distribution (unless κ is fixed) should not be regarded as an obstacle to use it, given the great easiness to construct curved plots with simple computer tools such as spreadsheet packages.

It is emphasised that equation (20) applies to daily rainfall. However, given the scaling similarities of extreme rainfall in several temporal scales (*Koutsoyiannis et al.*, 1998) the shape parameter κ estimated from daily series can be applied also to greater and shorter rainfall durations.

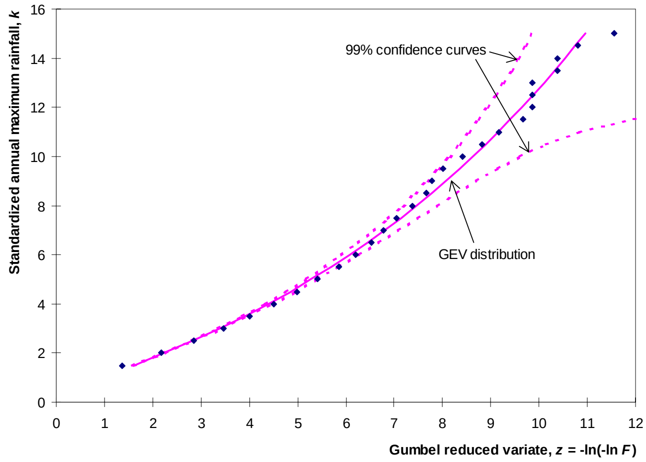

Figure 8. Gumbel probability plots of the empirical (rhombi) and GEV (continuous line) distribution functions of standardised rainfall depth k for all *Hershfield*'s (1961) data (from *Koutsoyiannis*, 1999).

6 SYNOPSIS AND CONCLUSION

The Gumbel or EV1 distribution is the asymptotic extreme value distribution for a wide range of parent distributions that are common in hydrology. However, the theoretical investigation of this study shows that convergence of the exact distribution of maxima to the asymptote may be extremely slow, thus making the Gumbel distribution an inappropriate approximation of the exact distribution of maxima. Besides, the attraction of parent distributions to this asymptote relies on a stationarity assumption, i.e. the assumption that parameters of the parent distribution are constant in time, which may not be the case in hydrological processes. Slight relaxation of this assumption may result in the EV2 rather than the EV1 asymptote.

The empirical investigation of the study, based on a 136-year record of extreme rainfall, shows that the EV1 distribution is inappropriate for the examined record (especially in its upper tail), whereas this distribution would seem as an appropriate model if fewer years of measurements were available (i.e., part of this sample were used). This allows the conjecture that the broad use of the Gumbel distribution worldwide may in fact be related to small sample sizes rather than to the real behaviour of rainfall maxima. In addition, the simplicity of the calculations of the Gumbel distribution along with its geometrical elucidation through a linear probability plot may have contributed to its popularity in hydrologists and engineers.

Interestingly, EV1 has been the prevailing model for rainfall extremes despite of the fact that it results in the highest possible risk for engineering structures, i.e. it yields the smallest possible design rainfall values in comparison to those of the three-parameter EV2 for any value of the shape parameter. The empirical investigation of this study demonstrates that the underestimation of design rainfall by the EV1 distribution is quite substantial (e.g. 1:2) for large return periods and this fact must be considered as a warning against the adoption of the EV1 distribution for rainfall extremes.

On the contrary, the three-parameter EV2 distribution (the GEV distribution bounded from the left) does not have the theoretical and empirical disadvantages of the EV1 distribution. Even though it is still a limiting distribution, yet away from the limit it can yield good approximations to the exact distribution of maxima, and it is not very sensitive to changes of parameters in time. In the long record of extreme rainfall of this study, the three-parameter EV2 distribution appears to be a suitable model. For rainfall records with smaller length EV2 may be not so easy to fit accurately. However, information from nearby records and an empirical relationship from a global study can assist in parameter estimation.

REFERENCES

Dooge, J. C. I., Looking for hydrologic laws, *Water Resour. Res.*, 22(9) pp. 46S-58S, 1986.

Farquharson, F. A. K., J. R. Meigh, and J. V. Sutcliffe, Regional flood frequency analysis in arid and semi-arid areas, *J. Hydrol.*, 138, 487–501, 1992.

Fisher, R. A., and L. H. C. Tippet, Limiting forms of the frequency distribution of the largest or smallest member of a sample, *Proc. Cambridge Phil. Soc.*, 24, 180-190, 1928.

Fréchet, M., Sur la loi de probabilité de l'écart maximum, *Ann. de la Soc. Polonaise de Math.*, Cracow, 6, 93-117, 1927.

Gnedenco, B.V., Limit theorems for the maximal term of a variational series, *Doklady Akad. Nauk SSSR*, Moscow, 32, 37 (in Russian), 1941.
Gumbel, E. J., *Statistics of Extremes*, Columbia University Press, New York, 1958.
Hazen, A., Storage to be provided in impounding reservoirs for municipal water supply, *Trans. ASCE*, ASCE, New York, 77, 1539–1640, 1914.
Hershfield, D. M., Estimating the probable maximum precipitation, *Proc. ASCE, J. Hydraul. Div.*, 87(HY5), 99-106, 1961.
Hershfield, D. M., Method for estimating probable maximum precipitation, *J. American Waterworks Association*, 57, 965-972, 1965.
Hosking, J. R. M., J. R. Wallis and E. F. Wood, Estimation of the generalized extreme value distribution by the method of probability weighted moments, *Technometrics*, 27(3), 251-261, 1985.
Jenkinson, A. F., The frequency distribution of the annual maximum (or minimum) value of meteorological elements, *Q. J. Royal Meteorol. Soc.*, 81, 158-171, 1955.
Jenkinson, A. F., Estimation of maximum floods, *World Meteorological Organization, Technical Note No 98*, ch. 5, 183-257, 1969.
Kendall, M.G. and Stuart, A., *The advanced theory of Statistics, Vol.1, Distribution theory*, 2nd edition, C. Griffin & Co., London, 1963.
Klemeš, V., Tall tales about tails of hydrological distributions, *J. Hydrol. Engineering*, 5(3), 227-231 & 232-239, 2000.
Kottegoda, N. T., and R. Rosso, *Statistics, Probability, and Reliability for Civil and Environmental Engineers*, McGraw-Hill, New York, 1997.
Koutsoyiannis, D., A probabilistic view of Hershfield's method for estimating probable maximum precipitation, *Water Resources Research*, 35(4), 1313-1322, 1999.
Koutsoyiannis, D., The Hurst phenomenon and fractional Gaussian noise made easy, *Hydrological Sciences Journal*, 47(4), 573-595, 2002.
Koutsoyiannis, D., and G. Baloutsos, Analysis of a long record of annual maximum rainfall in Athens, Greece, and design rainfall inferences, *Natural Hazards,* 22(1), 31-51, 2000.
Koutsoyiannis, D., D. Kozonis, and A. Manetas, A mathematical framework for studying rainfall intensity-duration-frequency relationships, *Journal of Hydrology*, 206(1-2), 118-135, 1998.
National Research Council, *Estimating Probabilities of Extreme Floods: Methods and Recommended Research*, National Academy Press, Washington, D.C., 1988.
Rossi, F., M. Fiorentino and P. Versace, Two-component extreme value distribution for flood frequency analysis, *WaterResour. Res.*, 20(7), 847-856, 1984.
Stedinger, J. R., R. M. Vogel, and E. Foufoula-Georgiou, Frequency analysis of extreme events, ch. 18 in *Handbook of Hydrology,* edited by D. R. Maidment, McGraw-Hill, New York, 1993.
Sutcliffe J.V., *Methods of Flood Estimation, A Guide to Flood Studies Report*, Report No. 49, Institute of Hydrology, UK, 1978.
Todorovic, P., and E. Zelenhasic, A stochastic model for flood analysis , *Water Resour. Res.*, 6(6), 1641-1648, 1970
Turcotte, D. L., Fractal theory and the estimation of extreme floods, J. Res. Natl. Inst. Stand. Technol., 99(4), 377-389, 1994.

Turcotte, D. L., and B. D. Malamud, Applicability of fractal flood-frequency statistics, *Hydrofractals '03, An international conference on fractals in hydrosciences*, Monte Verita, Ascona, Switzerland, August 2003, ETH Zurich, MIT, Université Pierre et Marie Curie, 2003.

U.S. Department of the Interior, Bureau of Reclamation, *Design of Arch Dams*, US Goverment Printing Office, Denver, Co., 1977.

U.S. Department of the Interior, Bureau of Reclamation, *Design of Small Dams*, 3rd edition, US Goverment Printing Office, Denver, Co., 1987.

von Bortkiewicz, L., Variationsbreite und mittlerer Fehler, *Sitzungsberichte d. Berliner Math. Ges.*, 21, 3, 1922a.

von Bortkiewicz, L., Die Variationsbreite bein Gauss'schen Fehlergesetz, *Nordisk Statistik Tidskrift*, 1(1), 11 & 1(2), 13, 1922b.

von Mises, R., Über die Variationsbreite einer Beobachtungsreihe, *Sitzungsber, d. Berliner Math. Ges.*, 22, 3, 1923.

Wilks, D. S., Comparison of three-parameter probability distributions for representing annual extreme and partial duration precipitation series, *Water Resour. Res.*, 29(10), 3543-3549, 1993.

Willeke, G. E., Myths and uses of hydrometeorology in forecasting, in *Proceedings of March 1979 Engineering Foundation Conference on Improved Hydrological Forecasting – Why and How*, pp. 117-124, American Society of Civil Engineers, New York, 1980.

Yevjevich, V., *Probability and Statistics in Hydrology*, Water Resources Publications, Fort Collins, Colorado, 1972.