New Asymptotic and Pre-Asymptotic Results on Rainfall Maxima from Multifractal Theory

By

Daniele Veneziano[1], Andreas Langousis[1] and Chiara Lepore[2]

(1) Department of Civil and Environmental Engineering, MIT, Cambridge, MA, 02139, USA

(2) Parsons Laboratory for Environmental Science and Engineering, MIT, Cambridge, MA 02139, USA

Submitted to

Water Resources Research

Submitted: May, 2009

Revised: August, 2009

---

*Corresponding author: Andreas Langousis, Dept. of Civil and Environmental Engineering, MIT, Room 1-245, Cambridge, MA, 02139. Tel: 617-407-0059, andlag@gmail.com.

Abstract

Contrary to common belief, Fisher-Tippett's extreme value (EV) theory does not typically apply to annual rainfall maxima. Similarly, Pickands' extreme excess (EE) theory does not typically apply to rainfall excesses above thresholds on the order of the annual maximum. This is true not just for long averaging durations *d*, but also for short *d* and in the high-resolution limit as $d \to 0$. We reach these conclusions by applying large deviation theory to multiplicative rainfall models with scale-invariant structure. We derive several asymptotic results. One is that, as $d \to 0$, the annual maximum rainfall intensity in *d*, $I_{yr,d}$, has generalized extreme value (GEV) distribution with a shape parameter *k* that is significantly higher than that predicted by EV theory and is always in the EV2 range. The value of *k* does not depend on the upper tail of the marginal distribution, but on regions closer to the body. Under the same conditions, the excesses above levels close to the annual maximum have generalized Pareto distribution with parameter *k* that is always higher than that predicted by Pickands' EE theory. For finite *d*, the distribution of $I_{yr,d}$ is not GEV, but in accordance with empirical evidence is well approximated by a GEV distribution with shape parameter *k* that increases as *d* decreases. We propose a way to estimate *k* under pre-asymptotic conditions from the scaling properties of rainfall and suggest a near-universal $k(d)$ relationship. The new estimator promises to be more accurate and robust than conventional estimators. These developments represent a significant conceptual change in the way rainfall extremes are viewed and evaluated.

Keywords: rainfall maxima, extreme value theory, extreme excess theory, large deviations, IDF curves

1. Introduction

This paper deals with the classical problem of characterizing the distribution of annual rainfall maxima. Let I_d be the average rainfall intensity in an interval of duration d and $I_{yr,d}$ be the maximum of I_d in one year. A long-standing tenet of stochastic hydrology is that, at least for d small, the distribution of $I_{yr,d}$ is of the generalized extreme value (GEV) type; see e.g. Chow *et al*. (1988), Singh (1992), and Stedinger *et al*. (1993). This belief stems from the fact that, if under suitable normalization the maximum of n independent and identically distributed (*iid*) variables is attracted as $n \to \infty$ to a non-degenerate distribution $G_{\max}$, then $G_{\max}$ must have the GEV form

$$G_{\max}(x) = \exp\left\{-\left[1 + k\left(\frac{x-\psi}{\lambda}\right)\right]^{-1/k}\right\} \qquad (1)$$

where λ, ψ and k are scale, location and shape parameters, respectively. Methods to estimate extreme rainfall intensities from recorded annual maxima (e.g. Koutsoyiannis *et al*., 1998; Martins and Stedinger, 2000; Gellens, 2002; Overeem *et al*., 2008) are generally based on this result.

The specific form of the distribution (EV1 when the shape parameter k = 0, EV2 when k > 0 and EV3 when k < 0) depends on the upper tail of the parent distribution, in our case the distribution of I_d (Gumbel, 1958). For k = 0, equation (1) reduces to the Gumbel (EV1) form $F(x) = \exp\{-\exp(-(x-\psi)/\lambda)\}$ with an exponential extreme upper tail, whereas for positive k the distribution is Frechet (EV2) whose upper tail behaves like a power function with exponent -1/k. Thus, for the same probability of exceedance, larger values of k are associated with higher rainfall intensities and more extreme behavior of the rainfall process. For negative k the distribution is Weibull (EV3), with a finite upper bound.

Another pillar of extreme rainfall modelling is extreme excess (EE) theory. Let X be a random variable with distribution F. The excess of X above u, $X_u = (X - u \mid X \geq u)$, has distribution $F_u(x) = \dfrac{F(u+x) - F(u)}{1 - F(u)}$. Pickands (1975) derived limiting properties of F_u that parallel the results of extreme value theory for the maxima. He found that, as u increases and $F(u) \to 1$: (1) the distribution of X_u converges to a non-degenerate distribution G_{exc} if and only if the maximum of n *iid* copies of X converges to a non-degenerate distribution $G_{\max}$; (2) G_{exc} has generalized Pareto (GP) form; and (3) G_{exc} has the same shape parameter k as $G_{\max}$ in equation (1).

An important property of the GP distribution is that the maximum of a Poisson number of *iid* GP(k) variables has GEV(k) distribution with the same k (e.g. Stedinger *et al.,* 1993). In conjunction with Pickands' results, this property has been extensively used in Peak-over-Threshold (PoT) and Partial-Duration-Series (PDS) methods of extreme rainfall analysis. Peak-over-Threshold methods generally assume that the peak of I_d above some high threshold u has GP distribution and find the (GEV) distribution of the annual maximum assuming that I_d up-crosses level u at Poisson times; see e.g. Smith (1985), Leadbetter (1991) and, Madsen *et al*. (1997). Partial-Duration-Series methods do the same using the marginal excesses of I_d above u; see e.g. Stedinger *et al*. (1993), Beirlant *et al.* (1996) and Martins and Stedinger (2001a,b).

We question whether the distribution of the annual maximum $I_{yr,d}$ is in fact GEV and has the shape parameter k of $G_{\max}$ in equation (1). This is clearly not the case for long durations d, say $d > 1$ week, because n = (1 year)/d is too small. However, extreme value (EV) theory might become relevant to $I_{yr,d}$ as $d \to 0$, since then $n \to \infty$. Similarly, we question whether as $d \to 0$ the excesses of I_d above thresholds on the order of the annual maximum have GP distribution with the same k as $G_{\max}$. We address these issues by using stationary models of rainfall in which rainfall intensity at different scales satisfies a scale invariance condition. These (multifractal) models have been found to accurately predict rainfall extremes (Veneziano *et al.*, 2006a; Langousis and Veneziano, 2007).

We find that, under stationarity and multifractality, EV theory does not apply to the annual maximum, because for any given d the block size n needed for reasonable convergence to the asymptotic GEV distribution far exceeds (1 year)/d. We are especially interested in the annual maxima at small scales, for which an appropriate framework is provided by large deviation (LD) theory (on LD theory, see e.g. Dembo and Zeitouni, 1993 and Den Hollander, 2000). Using LD tools, we obtain several new asymptotic results. One is that, as $d \to 0$, the annual maximum $I_{yr,d}$ approaches an EV2 distribution with a shape parameter k that is always higher than that predicted by extreme value theory. Interestingly, k does not depend on the upper tail of I_d but on regions of the distribution closer to the body and can be obtained in a simple way from the scaling properties of the rainfall process. Similarly, as $d \to 0$, the excess of I_d above thresholds on the order of $I_{yr,d}$ has GP(k) distribution, where k is the same as for $I_{yr,d}$ and therefore is always higher than the value from Pickands' theory.

We also study the distribution of $I_{yr,d}$ under pre-asymptotic conditions (d finite). These are the conditions of greatest interest in practice. In this case the distribution of $I_{yr,d}$ is not GEV and in fact may differ significantly from any EV or LD asymptotic distribution, but over a finite range of quantiles is accurately approximated by a GEV distribution with parameter k that decreases as d increases. This dependence of k on d is in accordance with much empirical evidence; see e.g. Asquith (1998), Mohymont *et al*. (2004), Trefry *et al*. (2005), Veneziano *et al*. (2007) and Section 4 below. We propose a method to estimate $k(d)$ from the scaling properties of the rainfall process and the range of quantiles (or return periods) of interest. The multifractal parameters provide a linkage between k and the local precipitation climate. We also suggest a near-universal default $k(d)$ relationship for use at non-instrumented sites.

Section 2 describes the rainfall model (a simple sequence of discrete multifractal cascades) and recalls results on the upper tail of I_d for such cascades from LD theory. Section 3 derives asymptotic properties of the N-year maximum $I_{Nyr,d}$ in the small-scale limit $d \to 0$ for cases with N fixed and N that varies as a power law of the averaging duration d. Section 3 derives also corresponding properties of the excess of I_d above thresholds on the order of $I_{Nyr,d}$. Section 4 focuses on the distribution of the annual maximum under pre-asymptotic conditions and Section 5 summarizes the main conclusions and outlines future steps.

In subsequent sections we make a change of notation, as follows. An important parameter of stationary multifractal processes is the upper limit D of the durations d for which the process displays scale invariance (e.g. Schertzer and Lovejoy, 1987; Gupta and Waymire, 1990; Veneziano, 1999; Langousis *et al*., 2007). In the analysis of such processes, what matters is not the duration d but the resolution $r = D/d$ relative to D. Accordingly, we use $I_r, I_{yr,r}$ and $I_{Nyr,r}$ in place of $I_d, I_{yr,d}$ and $I_{Nyr,d}$, respectively. Since the analysis is confined to the scaling range, we only consider resolutions $r \geq 1$.

2. Multiplicative and Multifractal Rainfall Models

There is ample evidence that the fluctuations of rainfall intensity at different scales combine in a multiplicative way; see e.g. Over and Gupta (1996), Perica and Foufoula-Georgiou (1996), Veneziano *et al*. (1996), Venugopal *et al*. (1999), Deidda (2000), and Veneziano and Langousis (2005a). Multiplicative models represent rainfall intensity $I(t)$ as

$$I(t) = m \prod_{j=1}^{\infty} Y_j(t) \tag{2}$$

where m is the mean rainfall intensity and the processes $Y_j(t)$ are non-negative, independent, with mean value 1. These processes contribute fluctuations at characteristic temporal scales d_j or equivalently at resolutions $r_j = D/d_j > 1$ relative to some large reference scale D. Since for our analysis the mean value does not matter, in what follows we set $m = 1$.

In the case of multifractal models, the resolutions r_j satisfy $r_j = b^j$ for some $b > 1$ and $Y_1(t), Y_2(t),...$ are contractive transformations of the same stationary random process $Y(t)$, meaning that $Y_j(t)$ is equivalent to $Y(r_j t)$; see e.g. Veneziano (1999). An important special case is when $Y(t)$ is a process with constant *iid* values inside consecutive D intervals and b is an integer ≥ 2. Then equation (2) generates a sequence of *iid* discrete multifractal cascades of multiplicity b within consecutive D intervals (on discrete multifractal cascades, see e.g. Schertzer and Lovejoy, 1987; Gupta and Waymire, 1990; and Evertsz and Mandelbrot, 1992). Discrete-cascade sequences of this type have been found to reproduce well the intensity-duration-frequency (IDF) curves extracted from historical records or generated by more sophisticated rainfall models (Langousis and Veneziano, 2007).

In a discrete-cascade representation of rainfall, the average rainfall intensity in a generic cascade tile at resolution r_j, I_{r_j}, satisfies

$$\begin{aligned} I_{r_j} &= A_{r_j} Z \\ A_{r_j} &= Y_1 Y_2 \cdots Y_j \end{aligned}, \qquad j = 0, 1, \ldots \tag{3}$$

where $A_{r_0} = 1$, the factors $Y_1,...,Y_j$ are independent copies of a non-negative variable Y with mean value 1, and Z is a mean-1 "dressing factor." Each Y_i, $i \leq j$, models the effect on I_{r_j} of the rainfall intensity fluctuations at resolution r_i, while Z captures the combined effect of all multiplicative fluctuations at resolutions higher than r_j; see Kahane and Peyriere (1976) and Schertzer and Lovejoy (1987).

An important feature of the distribution of Z is the asymptotic Pareto upper tail (i.e. $P[Z > z] \sim z^{-q^*}$) where $q^* > 1$ is the order at or beyond which the moments of Z diverge. The distribution of Z does not have analytical form, but it can be calculated numerically using the procedure of Veneziano and Furcolo (2003), or approximated analytically; see Langousis *et al*. (2007).

To realistically represent rainfall, one must model both the alternation of dry and wet conditions and the fluctuations of rainfall intensity during the rainy periods. This requires Y to have a non-zero probability mass at zero. A frequent choice is $Y = Y_\beta Y_{LN}$, where Y_β is a discrete random variable with probability mass P_0 at zero and probability mass 1- P_0 at $1/(1\text{-}P_0)$ and Y_{LN} is a lognormal variable with mean value 1 (e.g. Over and Gupta, 1996; Langousis *et al*., 2007). In the multifractal literature, processes with $Y = Y_\beta$ are called "beta" processes, while those with $Y = Y_{LN}$ are referred to as "lognormal" processes, although the marginal distribution is not exactly lognormal due to the dressing factor Z; see equation (3). When $Y = Y_\beta Y_{LN}$, we say that the process is "beta-lognormal" (beta-LN) and refer to the distribution of Y as a beta-LN distribution. The scaling properties of a beta-LN process depend on the probability P_0 and the variance of $\ln(Y_{LN})$ (see below for an alternative parameterization).

Later sections make frequent use of the moment-scaling function

$$K(q) = \log_{r_j}(E[A_{r_j}^q]) = \log_b(E[Y^q]) \tag{4}$$

and its Legendre transform $C(\gamma)$ given by

$$C(\gamma) = \max_q \{\gamma q - K(q)\}, \qquad K(q) = \max_\gamma \{\gamma q - C(\gamma)\} \tag{5}$$

In the beta-LN case, these functions are

$$\begin{aligned} K(q) &= C_\beta (q-1) + C_{LN}(q^2 - q), \qquad q \geq 0 \\ C(\gamma) &= \frac{C_{LN}}{4}\left(\frac{\gamma - C_\beta}{C_{LN}} + 1\right)^2 + C_\beta, \qquad \gamma \geq \gamma_{\min} \end{aligned} \tag{6}$$

where $C_\beta = -\log_b(1-P_{0)}$ and $C_{LN} = 0.5Var[\log_b(Y_{LN})]$ provide an alternative parameterization of the distribution of Y and $\gamma_{\min} = C_\beta - C_{LN}$ is the slope of $K(q)$ at 0. For example, in fitting a beta-LN model to a rainfall record from Florence, Italy, Langousis and Veneziano (2007) found $D \approx 15$ days, $C_\beta \approx 0.4$ and $C_{LN} \approx 0.05$. Figure 1 shows qualitative plots of the $K(q)$ and $C(\gamma)$ functions and indicates quantities of interest for the analysis that follows. Although for the present analysis the values of C_β and C_{LN} and more in general the distribution of Y do not matter, we use these settings to exemplify the theoretical results.

In the next section we need to evaluate how, in the small-scale limit $j \to \infty$, exceedance probabilities of the type $P[I_{r_j} > r_j^\gamma]$ depend on the resolution r_j and the exponent γ. For this we turn to large deviation (LD) theory (e.g. Dembo and Zeitouni, 1993). Specifically, Cramer's Theorem (Cramer, 1938) gives an asymptotic expression for the probability with which the sum of j *iid* variables exceeds levels proportional to j, as $j \to \infty$. One might think that as $j \to \infty$ the sum should have a normal distribution, but as j increases the quantiles of interest move into more extreme tail regions where the sum has not yet converged to the normal distribution. If for the moment one neglects the dressing factor Z in equation (3), then $I_{r_j} = A_{r_j}$ and Cramer's Theorem is directly relevant to our problem because $P[A_{r_j} > r_j^{\gamma}] = P[\sum_{i=1}^{j} \log_b(Y_i) > \gamma j]$. One can extend Cramer's results to include the dressing factor *Z*; see Veneziano (2002). This extension gives

$$P[A_{r_j} Z > r_j^{\gamma}] \sim \begin{cases} r_j^{-C(\gamma)}, & \gamma_{\min} \le \gamma \le \gamma^* \\ r_j^{-C(\gamma^*) - q^*(\gamma - \gamma^*)}, & \gamma > \gamma^* \end{cases} \tag{7}$$

where ~ denotes equality up to a factor $g(r_j, \gamma)$ that varies slowly (slower than a power law) with r_j at infinity, $C(\gamma)$ and $K(q)$ are the functions in equation (5), $q^* > 1$ is the moment order such that $K(q^*) = q^* - 1$, and γ^* is the slope of $K(q)$ at q^*. For $C(\gamma)$ and $K(q)$ in equation (8), $q^* = (1 - C_\beta)/C_{LN}$ and $\gamma^* = 2 - C_\beta - C_{LN}$. The asymptotic behavior of $g(r_j, \gamma)$ as $j \to \infty$ is known (Veneziano, 2002), but for the present objectives it is sufficient to work with the "rough limits" in equation (7).

The result in equation (7) for $\gamma < \gamma^*$ is also the limiting behavior of $P[A_{r_j} > r_j^{\gamma}]$ produced by Cramer's Theorem. The reason is that, for $\gamma_{\min} \le \gamma \le \gamma^*$ and *j* large, the dressing factor *Z* contributes a factor to the probability $P[A_{r_j} Z > r_j^{\gamma}]$ that does not depend on *j* and therefore can be absorbed into the function $g(r_j, \gamma)$. By contrast, for $\gamma > \gamma^*$ and *j* large, the probability $P[A_{r_j} Z > r_j^{\gamma}]$ is dominated by the Pareto tail of I_{r_j}, which has the form $P[I_{r_j} > i] \propto i^{-q^*}$ and starts at $i^* \sim r_j^{\gamma^*}$ (Langousis *et al.*, 2007). This power-law tail originates from the Pareto tail of the dressing factor *Z*; see comments following equation (3).

3. Asymptotic Analysis

In practice, one is interested in the distribution of the annual maximum $I_{yr,r}$ for finite resolutions *r* and the distribution of the excess $I_{r,u}$ for finite *r* and thresholds *u* on the order of

$I_{yr,r}$. Before studying these pre-asymptotic properties (see Section 4), we examine the behaviour of the *N*-year maximum $I_{Nyr,r}$ and the excess $I_{r,u}$ for thresholds *u* on the order of $I_{Nyr,r}$ under various asymptotic conditions. This asymptotic analysis produces extensions of extreme value (EV) and extreme excess (EE) results and clarifies why those theories do not apply to the annual rainfall maxima. We consider two cases: the classical limit (*r* fixed, $N \to \infty$) and the non-classical limit ($r \to \infty,\ N = cr^{\alpha}$) for any given $c > 0$ and α. When $\alpha = 0$, the latter limit becomes ($r \to \infty$, $N = c$ fixed) and thus characterizes the distribution of the *c*-year maximum of I_r at small scales. To simplify notation, we denote the resolution by *r*, with the understanding that in a discrete cascade model *r* is constrained to have values $r_j = b^j$. An important property of multifractal cascades that we use below is that, for resolutions *r* larger than about 2 and return periods *T* of practical interest (say $T/D \approx 10^2 - 10^6$), the distribution of $I_{Nyr,r}$ is accurately approximated by the distribution of the maximum of *rN*/*D* independent copies of I_r, where *D* is in years; see Langousis *et al*. (2007).

Consider first the limiting case (*r* finite, $N \to \infty$). As we have noted at the end of Section 2, the dressing factor *Z* causes I_r to have an algebraic upper tail of the type $P[I_r > i] \propto i^{-q^*}$, with q^* in equation (7). It follows from classical extreme value theory that, as $N \to \infty$, $I_{Nyr,r}$ is attracted to EV2(1/q^*), an EV2 distribution with shape parameter $k^* = 1/q^*$. It also follows that the excess above thresholds on the order of the *N*-year maximum is attracted to GP(1/q^*), a generalized Pareto distribution with the same shape parameter k^*.

The case ($r \to \infty,\ N = cr^{\alpha}$) is more interesting and produces new results. Our first step is to investigate the asymptotic behavior of the distribution of I_r for intensities in the range of the cr^{α}-year maximum. By this we mean the range between the ε- and (1-ε)-quantiles of $I_{cr^{\alpha}yr,r}$, where ε is a positive number arbitrarily close to 0. We denote these quantiles by $i_{\max,\varepsilon}$ and $i_{\max,1-\varepsilon}$, respectively. To examine the distribution of I_r within this range in the small-scale limit, we need the exceedance probabilities $P[I_r > i_{\max,\varepsilon}]$ and $P[I_r > i_{\max,1-\varepsilon}]$ as $r \to \infty$. Under the assumption that rainfall intensities in non-overlapping (*D/r*)-intervals are independent (as indicated above, this assumption produces accurate approximations of the maximum distribution), these probabilities are given by

$$\begin{aligned} P_\varepsilon &= P[I_r > i_{\max,\varepsilon}] = 1 - \varepsilon^{1/n} \\ P_{1-\varepsilon} &= P[I_r > i_{\max,1-\varepsilon}] = 1 - (1-\varepsilon)^{1/n} \end{aligned} \tag{8}$$

where $n = cr^{1+\alpha}/D$, with D expressed in years, is the number of (*D*/*r*)-intervals in cr^{α} years. Considering that ε is very small, $P_{1-\varepsilon} \approx \varepsilon/n$. One can further show that, for any given ε, $P_\varepsilon = 1 - \varepsilon^{1/n} \to \dfrac{\ln(1/\varepsilon)}{n}$ as $n \to \infty$. Therefore, for any given $\varepsilon > 0$, as ($r \to \infty$, $N = cr^{\alpha}$) the range $[i_{\max,\varepsilon}, i_{\max,1-\varepsilon}]$ corresponds to intensities i with exceedance probabilities $P[I_r > i] = \eta/n = \eta D/(cr^{1+\alpha})$, where $\varepsilon < \eta < \ln(1/\varepsilon)$ is positive and finite.

Appendix A uses equation (7) and the above results to show that, in the ($r \to \infty$, $N = cr^{\alpha}$) limit and for $\varepsilon < \eta < \ln(1/\varepsilon)$, the intensity i that is exceeded by I_r with probability $\eta D/(cr^{1+\alpha})$ varies with r, η and α as

$$i \sim \begin{cases} r^{\gamma_{1+\alpha}} \eta^{-1/q_{1+\alpha}}, & \alpha_{\min} < \alpha < \alpha^* \\ r^{\gamma^* + (\alpha-\alpha^*)/q^*} \eta^{-1/q^*}, & \alpha \ge \alpha^* \end{cases} \tag{9}$$

where $\gamma_{1+\alpha}$ satisfies $C(\gamma_{1+\alpha}) = 1+\alpha$, $q_{1+\alpha}$ is such that the slope $K'(q_{1+\alpha}) = \gamma_{1+\alpha}$, q^* and γ^* are the same as in equation (7), $\gamma^* = K'(q^*)$, $\alpha^* = C(\gamma^*) - 1 = q^*(\gamma^* - 1)$, and $\alpha_{\min} = -K(0) - 1$.

Some of these quantities are illustrated in Figure 1. Note that the results in equation (9) do not depend on the outer scale of multifractal behavior D or the constant c.

What is important for our analysis is that i in equation (9) varies with η like η^{-k_α} with

$$k_\alpha = \begin{cases} 1/q_{1+\alpha}, & \alpha_{\min} < \alpha < \alpha^* \\ 1/q^*, & \alpha \geq \alpha^* \end{cases} \tag{10}$$

From this power-law behavior of I_r in the range of the cr^α-year maximum we conclude that the maximum itself must be attracted to an EV2(k_α) distribution with k_α in equation (10). It also follows that, in the range of thresholds and intensities that satisfy $[\, i_{\max,\varepsilon} < u, I_{r,u} + u < i_{\max,1-\varepsilon}]$, the excess $I_{r,u}$ is attracted to a GP(k_α) distribution (generalized Pareto, with the same shape parameter k_α). Note that $k^* = 1/q^*$, the value of k_α for $\alpha \geq \alpha^*$, coincides with the shape parameter of the asymptotic GEV distribution from EV/EE theory.

For example, in the case of beta-LN processes, the parameters in equation (9) are

$$\begin{aligned} &\gamma_{1+\alpha} = C_\beta - C_{LN} + 2\sqrt{C_{LN}(1+\alpha-C_\beta)}, \qquad q_{1+\alpha} = \sqrt{(1+\alpha-C_\beta)/C_{LN}} \\ &\gamma^* = 2 - C_\beta - C_{LN}, \qquad q^* = (1-C_\beta)/C_{LN} \\ &\alpha_{\min} = C_\beta - 1, \qquad \alpha^* = (1-C_\beta)(q^*-1) \end{aligned} \tag{11}$$

and the shape parameter k_α in equation (10) is

$$k_\alpha = \begin{cases} \sqrt{C_{LN}/(1+\alpha-C_\beta)}, & C_\beta - 1 < \alpha < \alpha^* \\ C_{LN}/(1-C_\beta), & \alpha \geq \alpha^* \end{cases} \tag{12}$$

The value $\alpha = 0$ is of special interest, as in this case the maximum is over a constant number of years N (including N = 1 for the annual rainfall maxima). For α = 0, equation (12) gives $k_0 = \sqrt{C_{LN}/(1-C_\beta)} = \sqrt{k^*}$, where k^* is the value of k for $\alpha \geq \alpha^*$ (as well as the value of k from EV theory).

Figure 2 shows how k_α in equation (12) varies with α for beta-LN processes. The expressions in the figure are generic for any scaling parameters C_β and C_{LN} , but the plot is for $C_\beta = 0.4$ and $C_{LN} = 0.05$, which are realistic values for rainfall. As one can see, for all $\alpha < \alpha^*$ the parameter k_α exceeds the value k^* from EV theory and diverges as $\alpha \rightarrow \alpha_{\min} = C_\beta - 1$. For $\alpha = 0$, the constraint $C_\beta + C_{LN} < 1$ implies $k_0 < 1$. For the specific values of C_β and C_{LN} used in the figure, $k^* = 0.083$ and $k_0 = 0.289$. Hence EV theory severely under-predicts the shape parameter k of the annual maximum in the small-scale limit. This under-prediction results in unconservative intensity-duration-frequency (IDF) values for long return periods.

The main conceptual results of this section are illustrated in Figure 3. The coordinate axes are the resolution $r = D/d$ and the number of independent I_r variables over which the maximum is taken. For the N-year maximum, this number is $n(r) = Nr/D$, where D is in years. The scale is logarithmic in both variables. Extreme value (EV) analysis gives that, for any given r, as $N \rightarrow \infty$ the distribution of the maximum converges to an EV2(k^*), where $k^* = 1/q^*$. The frequent use of this result for the annual maximum ($N = 1$) is based on the implicit assumption that a relatively low block size n_0 (see dashed horizontal line in Figure 3) is sufficient for convergence of the maximum to EV2(k^*). If this is not true for low r because $n(r) = r/D$ is too small, the distribution of the maximum should be EV2(k^*) at higher resolutions for which $r/D >> n_0$. Figure 3 shows that (a) when r is relatively small, reasonable convergence of the maximum to EV2(k^*) requires block sizes $n(r)$ that are $10^3 - 10^4$ times the annual block size r/D; hence, unless $N \approx (10^3\text{-}10^4)$ years, the N-year maximum cannot be assumed to have EV2(k^*) distribution, and (b) the threshold n_0 is not constant, but increases with increasing r as $n_0 \sim r^{1+\alpha^*}$, with $\alpha^* \approx 7$; the latter value of α is obtained from equation (12), using realistic values of

C_β and C_{LN} from Figure 6.b; see Section 4 below. Since $1+\alpha^* >> 1$, as r increases the threshold on *n*(*r*) above which EV theory applies moves farther away from the available block size *r/D*. This makes the EV results even less relevant at high resolutions. Based on these results, we conclude that, under multifractality, EV theory (and for the same reasons EE theory) does not apply to annual rainfall extremes.

For a number of years $N = cr^\alpha$, the block size is $n(r) = cr^{1+\alpha}/D$, where D is in years. Therefore, as r increases, one moves in Figure 3 along straight lines with slope $(1+\alpha)$. For $\alpha > \alpha^*$, one eventually enters the region where EV theory holds and, as $r \to \infty$, the maximum becomes EV2(*k**); see equation (10). It follows from the same equation that, for $\alpha < \alpha^*$ and as $r \to \infty$, the cr^α-year maximum is attracted to an EV2(k_α) distribution with k_α in equation (12).

Summarizing, in the context of multifractal models, large deviation (LD) theory extends the results on rainfall extremes beyond the classical context of extreme value (EV) and extreme excess (EE) theories. Specifically, the latter theories deal with the maximum of I_r at fixed resolution r over an infinitely long period of time, whereas LD theory produces results for $r \to \infty$ and periods of time that are either constant or diverge as power laws of *r*.

4. Pre-Asymptotic Distribution of the Annual Maximum and GEV Approximations

In practice, one is interested in the annual maximum rainfall $I_{yr,r}$ over a finite range of resolutions. The associated points (*r*, *r/D*) in Figure 3 are typically far from the regions where the EV, EE and LD theories apply. For these (*r*, *r/D*)-combinations the distribution of $I_{yr,r}$ is not GEV, but over a finite range of exceedance probabilities *P* or equivalently of return periods *T* = 1/*P*, it may be accurately approximated by a GEV distribution. Indeed, one often finds that GEV(*k*) distributions fit well annual maximum data, with *k* being an increasing function of *r*. If one could relate the best-fitting k to the resolution r and the multifractal parameters C_β and C_{LN}, then one could develop a new estimator of k based on scaling theory: i.e. based on the estimates of the multifractal parameters C_β and C_{LN} from empirical records. This would be a valuable finding, since k is notoriously difficult to infer directly from annual maxima; see e.g. Mohymont *et al*. (2004) and Koutsoyiannis (2004). Moreover, linking $k(r)$ to C_β and C_{LN} would shed light on what rainfall-climate factors control the shape of the annual maximum distribution.

First we investigate whether, over a range of return periods T, the theoretical distribution of $I_{yr,r}$ from the multifractal model in Section 2 is approximately GEV. For this purpose, we calculate the exact distribution of $I_{yr,r}$ for various (C_β, C_{LN})-combinations and different resolutions r using the method of Langousis *et al*. (2007) assuming independence of rainfall in different D intervals within a year. Then we plot this exact distribution on GEV(k) paper, varying k until the resulting plot in a given range of T is closest to a straight line in a least-squares sense. As an example, the top row of Figure 4 shows these best linear fits for a beta-lognormal cascade with parameters ($C_\beta = 0.4$, $C_{LN} = 0.05$, D = 15 days) and gives the associated values of k for r = 1 and 512 in the return-period range $2 < T < 10\,000$ years. For comparison, the lower rows in Figure 4 show similar plots on GEV(k) paper for $k = 0$ (EV1 distribution), $k^* = 1/q^* = 0.083$ (EV2 distribution predicted by EV and EE theories), and $k_0 = 1/q_1 = 0.289$ (EV2 distribution from LD theory under $r \to \infty$). It is clear that when k is optimized (top row), GEV(k) distributions provide accurate approximations to the exact distribution, whereas fixing k to 0, $1/q^*$ or $1/q_1$ generally produces poor fits. We have repeated the analysis using different ranges of return periods, a denser set of resolutions r and different multifractal parameters. In all cases the quality of the best fit is comparable to that in the top row of Figure 4. As one may expect from the top-row panels of Figure 4, the least-squares k is insensitive to the range of return periods used in the least-squares fit. For example, k is almost the same when best fitting a GEV(k) distribution in the ranges from 2 to 100, 2 to 1000, or 2 to 10 000 years.

For the same multifractal process as in Figure 4, Figure 5.a shows plots of the best-fitting k (in the $2 < T < 100$ years range) against r. The vertical bars are $(m \pm \sigma)$ intervals for the probability weighted moments (PWM) estimator of k applied to 60 series of 100 annual maximum values, each extracted from a 100-year continuous multifractal process simulation; on the PWM method of parameter estimation, see e.g. Hosking (1990, 1992), Koutsoyiannis (2004) and Trefry *et al*. (2005). For reference, the values $k^* = 1/q^*$ and $k_0 = 1/q_1$ are shown as dashed horizontal lines. As $r \to \infty$, k approaches $1/q_1$, but over the range of resolutions considered, k remains far from this limit. The mean of the estimator follows closely the least-squares k line, except for a slight negative bias at low resolutions. As one can see, even with 100 years of data the PWM estimator has high variability. Figure 5.b compares the least-squares k from Figure 5.a with values of k from the literature. These values were obtained from annual maximum rainfall records of different lengths using the probability weighted moment (PWM) method. The empirical values have a wide scatter, which is broadly consistent with the sampling variability in Figure 5.a. The theoretical best-fitting k values (for $C_\beta = 0.4$, $C_{LN} = 0.05$ and D = 15 days) are generally higher, but have a dependence on r similar to the empirical values. Larger values of k correspond to a thicker upper tail and, hence, higher upper quantiles of the annual maximum distribution. Possible reasons for the theoretical values being higher are negative bias of the empirical estimators and deviations of actual rainfall from the multifractal model used to produce the theoretical estimates. The latter include variations in the multifractal parameters (C_β, C_{LN}, D) and deviations from strict scale invariance; see e.g. Menabde *et al.* (1997), Schmitt *et al.*

(1998), Olsson (1998), Güntner *et al*. (2001), Veneziano *et al*. (2006b) and Veneziano and Langousis (2009). These sources of discrepancy will be the subject of future investigations. It is remarkable (but possibly coincidental) that the only empirical results based on a very extensive data set [169 daily records, each having 100-154 years of data (Koutsoyiannis, 2004); see "*K*" point in Figure 5.b] are almost identical to the theoretical values.

Figure 6.a compares the variation of the least-squares *k* value with *r* for selected combinations of C_β and C_{LN}. Generally, *k* increases as either parameter increases. However, if one considers the relative small spatial variation of these parameters (see Figure 6.b where C_β and C_{LN} estimates from different rainfall records are plotted against the local mean annual precipitation $\bar{I}_{yr}$), the sensitivity of *k* in Figure 6.a is modest. As Figure 6.b shows, C_{LN} may be considered constant around 0.053, whereas C_β has a linear decreasing trend with $\bar{I}_{yr}$. The default *k* curve in Figure 6.c has been obtained by using the (C_β, C_{LN}) combinations in Figure 6.b and ensemble averaging the results. The dashed lines in the same figure are bounds considering the variability of C_β in Figure 6.b. If one uses higher values of C_β in more arid climates, as suggested by Figure 6.b, *k* would be slightly higher.

The solid line in Figure 6.c is close to the following analytical expression:

$$k = 2.44\ [\log_{10}(r) + 0.557]^{0.035} - 2.362 \qquad (13)$$

whereas the dashed lines deviate by approximately ±0.03-0.05 (depending on the resolution $r = D/d$) from the default *k* values in equation (13).

5. Conclusions

A long tradition links the modeling and analysis of rainfall extremes to Fisher-Tippett's extreme-value (EV) and Pickands' extreme-excess (EE) theories. This includes methods that use annual-maximum and peak-over-threshold rainfall information. However, for realistic rainfall models, neither theory applies. The basic reason is that the annual maxima depend on a range of the marginal distribution much below its upper tail. This realization has profound consequences on the distribution of the annual maxima and on methods for its estimation.

To prove these points and obtain new results on rainfall extremes, we have used stationary rainfall models with multifractal scale invariance below some temporal scale *D*. This scale may be seen as the time between consecutive synoptic systems capable of generating rainfall; see Langousis and Veneziano (2007). Stationary multifractal models are non-negative random processes in which the fluctuations at different scales combine in a multiplicative way and for equal log-scale increments have statistically identical amplitude. These models have received significant attention in the precipitation literature, including rainfall extremes. For multifractal models, one can use a branch of asymptotic probability theory known as large deviations (LD) to extend the limiting results from EV and EE theories. Specifically we have found that, as the averaging duration $d \to 0$ or equivalently the resolution $r = D/d \to \infty$, the distribution of the annual maximum $I_{yr,r}$ is GEV with shape parameter *k* in the EV2 range. Under the same asymptotic conditions, the excess of the marginal rainfall intensity I_r above thresholds *u* on the order of the annual maximum $I_{yr,r}$ has generalized Pareto (GP) distribution with the same shape parameter *k*. The value of *k* is much higher than that produced by EV and EE theories and can be found theoretically from the scaling properties of the rainfall process. These asymptotic results hold also for the distribution of the *N*-year maximum $I_{Nyr,r}$ for any finite *N* and the excesses of I_r above thresholds on the order of $I_{Nyr,r}$.

With added generality, LD theory gives the asymptotic distribution of $I_{cr^{\alpha}yr,r}$, the (cr^{α})-year maximum, for any $c > 0$ and $\alpha \geq \alpha_{\min}$ where $\alpha_{\min} < 0$ is a certain lower bound. As $r \to \infty$, the distribution of $I_{cr^{\alpha}yr,r}$ is again EV2, with shape parameter k_{α} that: 1) is always higher or equal to the value $k = k^*$ predicted by EV and EE theories, 2) depends only on α and 3) can again be found from the scaling properties of the rainfall process. The excess of I_r above thresholds on the order of the (cr^{α})-year maximum has GP(k_{α}) distribution with the same value of k_{α}. The value $k = k^*$ from classical EV and EE analysis is recovered for α larger than a critical value α^*. Therefore, in the context of multifractal models, our analysis generalizes the results of classical EV and EE theories. Note that using k^* instead of k_{α} would result in underestimation of the probability of extreme rainfalls.

At the root of the differences between our results and those of classical EV theory is that the settings under which the results are obtained are different: In EV analysis one fixes the resolution r and considers the distribution of the maximum of n independent copies of I_r as $n \to \infty$. The asymptotic EV results are commonly assumed to apply to the annual maxima, at least at high resolutions r. By contrast, in the LD analysis one lets $r \to \infty$ while setting n to the number of resolution-r intervals in one year. In the latter formulation, n varies with r in a way that makes sense for the study of the annual maxima at small scales.

Other important results we have obtained concern the distribution of the annual maximum $I_{yr,r}$ for finite r. In this case the distribution is not GEV, but over a range of quantiles of practical interest can be accurately approximated by a GEV(k) distribution. We have found that the best-fitting shape parameter k increases with increasing resolution r, in a way consistent with findings from directly fitting GEV distributions to annual maximum data; see Section 4. The best-fitting k generally remains within the EV2 range, but at large scales it is close to zero (EV1 fit). This finding is important, as it explains why an EV2 distribution often fits well the annual maximum data and why the shape parameter depends on the resolution (in contrast with the asymptotic EV prediction that k is constant with r). The best-fitting k depends little on the range of quantiles used in the fit and is not very sensitive to the scaling parameters, within the range of values that are typical for rainfall (except that k tends to be somewhat higher in dry than in wet climates). Taking advantage of this lack of sensitivity, we have obtained default values of k as a function of r, which can be used at non-instrumented sites or in cases of very short rainfall records.

The above results are significant in several respects. The asymptotic findings (1) show that large-deviation theory should find a place in stochastic hydrology at least as prominent as EV and EE theories and (2) indicate that what matters for the annual maximum rainfall is usually not the upper tail of the parent distribution, but a range of that distribution closer to the body. In addition, the pre-asymptotic analysis (1) shows that GEV models accurately approximate the non-GEV distribution of the annual maximum, (2) indicates that the shape parameter k of the approximating GEV distribution varies with resolution r, and (3) produces new ways to estimate k, from the scaling properties of rainfall.

This line of inquiry should continue. There is evidence that rainfall satisfies multifractal scale-invariance only in approximation, over a finite range of scales (typically between about 1 hour and several days) and under certain conditions (for example only within rainstorms); see e.g. Schmitt *et al.*, (1998), Sivakumar *et al*. (2001), Veneziano *et al*. (2006b) and Veneziano and Langousis (2009). It would be interesting to examine the sensitivity of our results to the structure of the rainfall model. Specific alternatives to our multifractal representation are bounded cascades (see e.g Menabde *et al.*, 1997 and Menabde, 1998), which retain the multiplicative structure but allow the intensity of the fluctuations to vary with scale, and models that explicitly recognize rainstorms and dry inter-storm periods and assume scale invariance (or bounded-cascade behavior) within the storms (e.g. Langousis and Veneziano, 2007).

A notoriously difficult problem is to estimate the shape parameter k of the annual maximum distribution from at-site information (see e.g. Koutsoyiannis, 2004). This is why one often resorts to regionalization. The finding that k is determined not by the upper tail of I_r but by regions of the distribution closer to the body and can be calculated from the scaling properties of rainfall opens new possibilities for both at-site and regionalized estimation of this parameter. Developments in this direction will be the subject of follow-up communications.

Appendix A: Small-Scale Behavior of Certain Quantiles of I_r

Let i be the value exceeded by I_r with probability $\eta D/(cr^{1+\alpha})$, where c and D are given positive constants. We are interested in how, as the resolution $r \to \infty$, i varies with r and $0 < \eta < \infty$, for different α. For this purpose, we write i as r^{γ} and use equation (7) to find γ such that $P[I_r > r^{\gamma}]$ $= \eta D/(cr^{1+\alpha})$.

Suppose first that $\gamma \le \gamma^*$, where γ^* is the slope of $K(q)$ at q^* (as we shall see, γ does not exceed γ^* if α does not exceed a related threshold α^*). Then equation (7) gives

$$P[I_r > r^{\gamma}] \sim r^{-C(\gamma)} \tag{A.1}$$

We want γ such that the right hand side of equation (A.1) equals $\eta D/(cr^{1+\alpha})$. Therefore γ must satisfy

$$C(\gamma) = (1+\alpha) + \log_r(\frac{c}{\eta D}) \tag{A.2}$$

For any finite c, b and D, $\log_r[c/(\eta D)] \to 0$ as $r \to \infty$. Hence one may replace $C(\gamma)$ in equation (A.2) with its linear Taylor expansion around the value $\gamma_{1+\alpha}$ such that $C(\gamma_{1+\alpha}) = 1+\alpha$. Using equation (5), this gives

$$C(\gamma) = (1+\alpha) + q_{1+\alpha}(\gamma - \gamma_{1+\alpha}) \tag{A.3}$$

where $q_{1+\alpha}$ is the moment order at which the slope of $K(q)$ in equation (4) equals $\gamma_{1+\alpha}$ and is also the derivative of $C(\gamma)$ at $\gamma_{1+\alpha}$; see Figure 1. Equating the right hand sides of equations (A.2) and (A.3), one obtains

$$\gamma = \gamma_{1+\alpha} + \frac{1}{q_{1+\alpha}} \log_r\left(\frac{c}{\eta D}\right) \tag{A.4}$$

We conclude that, for large r and any given c and D, $i = r^{\gamma}$ satisfies

$$i \sim r^{\gamma_{1+\alpha}} \eta^{-1/q_{1+\alpha}} \tag{A.5}$$

Equation (A.5) holds for $\gamma_{1+\alpha} \le \gamma^*$, or equivalently for $\alpha \le \alpha^*$, where $\alpha^* = C(\gamma^*) - 1 = q^*(\gamma^* - 1)$.

For $\alpha > \alpha^*$, γ exceeds γ^* and one must use the second expression in equation (7). Therefore γ must satisfy

$$C(\gamma^*) + q^*(\gamma - \gamma^*) = (1+\alpha) + \log_r\left(\frac{c}{\eta D}\right) \tag{A.6}$$

Solving for γ and using $C(\gamma^*) = 1 + \alpha^*$ gives the following expression for $i = r^{\gamma}$:

$$i \sim r^{\gamma^* + (\alpha - \alpha^*)/q^*} \eta^{-1/q^*} \tag{A.7}$$

The results in equations (A.5) and (A.7) are reproduced in equation (9).

Acknowledgments

This work was supported by project RISK (High Speed Rail) of the M.I.T.-Portugal Program (MPP), project POCI/GEO/59712/2004 of the Portuguese Foundation for Science and Technology (FCT), and the NSF Award # EAR-0910721. The second author was also partially supported by the Alexander S. Onassis Public Benefit Foundation under Scholarship No. F-ZA 054/2005-2006. We thank Bellie Sivakumar and two anonymous reviewers for their comments and suggestions.

References

Asquith, W.H. (1998) Depth-Duration Frequency of Precipitation for Texas, Water-Resources Investigations Report 98–4044, U.S. Geological Survey, http://pubs.usgs.gov/wri/wri98-4044/pdf/98-4044.pdf.

Beirlant, J., J.L. Teugels and P. Vynckier (1996) *Practical Analysis of Extreme Values,* Leuven University Press Blijde-Inkomststraat 5, B-3000 Leuven (Belgium), pp. 137.

Caporali, E., E. Cavigli and A. Petrucci (2006) Regional Frequency Analysis of Rainfall extremes in Tuscany (Italy), International Workshop on Hydrological Extremes "Observing and Modelling Exceptional Floods and Rainfalls", Rende (CS), 3-4 May 2006.

Chow, V.T., D.R. Maidment and L.W. Mays (1988) *Applied Hydrology*, McGraw-Hill, New York.

Cramer, H. (1938) Sur un Nouveau Theoreme-limite de la Theorie des Probabilites, *Actualites Scientifiques et Industrielles,* No. 736 of Colloque consacre a la theorie des probabilities, Herrman, Paris, pp. 5-23.

Deidda, R. (2000) Rainfall Downscaling in a Space-time Multifractal Framework, *Wat. Resour. Res.*, **36**(7), 1779-1794.

Dembo, A. and O. Zeitouni O (1993) *Large Deviation Techniques and Applications*, Jones and Barlett, Boston.

Den Hollander, F. (2000) *Large Deviations*, American Mathematical Society.

Evertsz, C.J.G. and B.B. Mandelbrot (1992) Appendix B of *Chaos and Fractals*, eds. H.-O. Peitgen, H. Jurgens, and D. Saupe, Springer-Verlag.

Gellens, D. (2002) Combining Regional Approach and Data Extension Procedure for Assessing GEV Distribution of Extreme precipitation in Belgium, *J. Hydrol.*, **268**, 113-126.

Güntner, A., J. Olsson, A. Calver and B. Gannon (2001) Cascade-based Disaggregation of Continuous Rainfall Time Series: The Influence of Climate, *Hydrol. Earth Syst. Sci.*, **5**, 145-164.

Gumbel, E.J. (1958) *Statistics of Extremes,* Columbia University Press, New York.

Gupta, V. K. and E. Waymire (1990) Multiscaling Properties of Spatial Rainfall and River Flow Distributions, *J. Geophys. Res.,* **95** (D3), 1999-2009.

Hosking, J.R.M. (1990) L-moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *J. Roy. Statist. Soc. Ser. B*, **52**, 105–124.

Hosking, J.R.M. (1992), Moment or L Moments? An Example Comparing two Measures of Distributional Shape, *The American Statistician*, **46**(3), 186-189.

Kahane, J.P. and J. Peyriere (1976) Sur certaines martingales de Benoit Mandelbrot, *Adv. Math.,* **22**: 131-145.

Koutsoyiannis, D. (2004) Statistics of Extremes and Estimation of Extreme Rainfall: II. Empirical Investigation of Long Rainfall Records, *Hydrolog. Sci. J.*, **49**(4), 591– 610.

Koutsoyiannis, D., D. Kozonis and A. Manetas (1998) A Mathematical Framework for Studying Rainfall Intensity-Duration-Frequency Relationships, *J. Hydrol.,* **206**, 118-135.

Langousis A, D. Veneziano, P. Furcolo, and C. Lepore (2007) Multifractal Rainfall Extremes: Theoretical Analysis and Practical Estimation, *Chaos Solitons and Fractals*, doi:10.1016/j.chaos.2007.06.004.

Langousis, A. and D. Veneziano (2007) Intensity-duration-frequency curves from Scaling Representations of Rainfall, *Wat. Resour. Res.*, **43**, doi: 10.1029/2006WR005245.

Leadbetter, M.R. (1991) On a Basis for “Peak over Threshold” Modeling, *Statistics & Probability Letters*, **12**, 357-362.

Lepore, C., D. Veneziano and A. Langousis (2009) Annual Rainfall Maxima: Practical Estimation Based on Large-Deviation Results, European Geosciences Union, Vienna, Austria, April 2009.

Madsen, H., P.F. Rasmussen and D. Rosbjerg (1997) Comparison of Annual Maximum Series and Partial Duration Series Methods for Modeling Extreme Hydrologic Events 1. At-site Modeling, *Wat. Resour. Res.*, **33**(4), 747-757.

Martins, E.S. and J.R. Stedinger (2001a) Historical Information in a Generalized Maximum Likelihood Framework With Partial Duration and Annual Maximum Series, *Water Resour. Res.*, **37**(10), 2559–2567.

Martins, E.S. and J.R. Stedinger (2001b) Generalized Maximum Likelihood Pareto-Poisson Estimators For Partial Duration Series, *Water Resour. Res.*, 37(10), 2551–2557.

Martins, E.S. and J.R. Stedinger (2000) Generalized Maximum-likelihood Generalized Extreme-value Quantile Estimators for Hydrologic Data, *Wat. Resour. Res.*, **36**(3), 737-744.

Menabde, M. (1998) Bounded Lognormal Cascades as Quasi-Multiaffine Random Fields, *Nonlinear Processes Geophys.* **5**, 63-68.

Menabde, M., D. Harris, A. Seed, G. Austin and D. Stow (1997) Multiscaling Properties of Rainfall and Bounded Random Cascades, *Wat. Resour. Res*., **33**(12), 2823-2830.

Mohymont, B., G.R. Demareé, D.N. Faka (2004) Establishment of IDF-curves for Precipitation in the Tropical Area of Central Africa: Comparison of Techniques and Results, *Nat. Haz. Ear. Sys. Sci.*, **4**, 375-387.

Olsson, J. (1998) Evaluation of a Scaling Cascade Model for Temporal Rain-fall Disaggregation, *Hydrol. Earth Syst. Sci.*, **2**, 19-30.

Over, T.M. and V.K. Gupta (1996) A Space-Time Theory of Mesoscale Rainfal Using Random Cascades, *J. Geophys. Res.*, **101**(D21), 26 319- 26 331.

Overeem, A., A. Buishand, and I. Holleman (2008) Rainfall Depth-duration-frequency Curves and their Uncertainties, *J. Hydrol.*, **348**, 124-134.

Perica, S. and E. Foufoula-Georgiou (1996) Model for Multiscale Disaggregation of Spatial Rainfall Based on Coupling Meteorological and Scaling Descriptions. *J. Geophys. Res.*, **101**(D21), 26347-26361.

Pickands, J. (1975) Statistical inference using extreme order statistics, *Ann. Statist.* **3**, 119-131.

Schertzer, D. and S. Lovejoy (1987) Physical Modeling and Analysis of Rain and Clouds by Anisotropic Scaling of Multiplicative Processes, *J. Geophys. Res.*, **92**, 9693-9714.

Schmitt, F., S. Vannitsem, and A. Barbosa (1998) Modeling of Rainfall Time Series Using Two-state Renewal Processes and Multifractals, *J. Geophys. Res.*, **103**(D18)**92**, 23181-23193.

Singh, V.P. (1992) *Elementary Hydrology*, Prentice-Hall, New Jersey, U.S.A.

Sivakumar, B., S. Sorooshian, H.V. Gupta and X. Gao (2001). A Chaotic Approach to Rainfall Disaggregation, *Wat. Resour. Res.*,, **37**(1), 61-72.

Smith, R.L. (1985) Threshold Methods for Sample Extremes, in J. Tiago de Oliveira, ed., *Statistical Extremes and Applications*, NATO ASI Series (Reidl, Dordrecht), pp. 623-638.

Stedinger, J.R., R.M. Vogel and E. Foufoula-Georgiou (1993) Frequency Analysis in Extreme Events, in *Handbook of Applied Hydrology*, edited by D. A. Maidment, Ch. 18, pp. 18.1-18.66, McGraw-Hill, New York.

Trefry, C.M., D.W. Watkins Jr. and D. Johnson (2005) Regional Rainfall Frequency Analyses for the State of Michigan*, J. Hydrol. Eng.*, **10**(6): 437-449.

Vaskova, I. (2005) Rainfall Analysis and Regionalization Computing Intensity-Duration-Frequency Curves, www.lyon.cemagref.fr/projets/floodaware/report/05upv1.pdf

Veneziano D. and A. Langousis (2009) Scaling and Fractals in Hydrology, In: *Advances in Data-based Approaches for Hydrologic Modeling and Forecasting*, Edited by: B. Sivakumar, World Scientific (in press).

Veneziano, D. (1999) Basic Properties and Characterization of Stochastically Self-similar Processes in R^d *Fractals,* **7**, 59.

Veneziano, D. (2002) Large Deviations of Multifractal Measures, *Fractals*, **10**, 117-129. Erratum in *Fractals*, 2005; **13**(2), 1-3.

Veneziano, D. and A. Langousis (2005a) The Areal Reduction Factor a Multifractal Analysis, *Wat. Resour. Res.*, **41**, W07008, doi:10.1029/2004WR003765.

Veneziano, D. and A. Langousis (2005b) The Maximum of Multifractal Cascades: Exact Distribution and Approximations, *Fractals*, **13**(4), 311-324.

Veneziano, D., A. Langousis and P. Furcolo (2006a) Multifractality and Rainfall Extremes: A Review, *Wat. Resour. Res.*, **42**, W06D15, doi:10.1029/2005WR004716.

Veneziano, D., C. Lepore, A. Langousis and P. Furcolo (2007) Marginal Methods of IDF Estimation in Scaling and Non-Scaling Rainfall, *Water Resour. Res.*, **43**, W10418, doi:10.1029/2007WR006040.

Veneziano, D., P. Furcolo and V. Iacobellis (2006b) Imperfect Scaling of Time and Space-Time Rainfall, *J. Hydrol.*, **322**(1-4), 105-119.

Veneziano, D., R.L. Bras and J.D. Niemann (1996) Nonlinearity and Self-similarity of Rainfall in Time and a Stochastic model, *J. Geophys. Res.*, **101**(D21), 26371- 26392.

Venugopal, V., E. Foufoula-Georgiou and V. Sapoznhnikov (1999) Evidence of Dynamic Scaling in Space Time Rainfall, *J. Geophys. Res.*, **104**(D24), 31599-31610.

Figure Captions

Figure 1: Illustration of the moment scaling function $K(q)$ and its Legendre transform $C(\gamma)$ in equation (6).

Figure 2: Shape parameter k_α of the N-year maximum of I_r under $(r \to \infty,\ N = cr^\alpha)$. Beta-lognormal rainfall process with C_β = 0.4 and C_{LN} = 0.05. Larger values of k correspond to higher probabilities of exceedance of extreme rainfalls.

Figure 3: Schematic illustration of asymptotic results on rainfall maxima from extreme value (EV) and large deviation (LD) theories.

Figure 4: GEV(k) approximations to the exact distribution of the annual maximum $I_{yr,r}$ at resolutions r = 1 and 512, in the return-period range from 2-10 000 years. The top row shows the best least-squares fit on GEV(k) paper and gives the associated value of k. The lower rows show plots on GEV(k) paper for k = 0 (EV1 paper), $k^* = 1/q^*$ (value predicted by EV and EE theories), and $k_0 = 1/q_1$ (value predicted by LD theory for $r \to \infty$). Deviations of the plots from a straight line indicate lack of fit for the selected value of k.

Figure 5: Dependence of the least-squares shape parameter k on the resolution $r = D/d$. (a) Theoretical values of k for C_β = 0.4 and C_{LN} = 0.05 when fitting is over the return period range from 2-100 years. The vertical bars are $(m \pm \sigma)$ intervals for the probability weighted moments (PWM) estimator of k using the annual maxima from 100-year continuous multifractal process simulations. The values $k^* = 1/q^*$ and $k_0 = 1/q_1$ are shown for reference. (b) Comparison of the theoretical values of k from (a) with empirical estimates from the literature assuming an average value of D = 15 days.

Figure 6. (a) Best-fitting shape parameters k at different resolutions r for selected combinations of C_β and C_{LN}. The range of return periods T used for fitting is from 2 - 100 years. (b) Estimates of C_β and C_{LN} from different rainfall records plotted against the mean annual precipitation. (c) Suggested default values of k as a function of the resolution $r = D/d$.

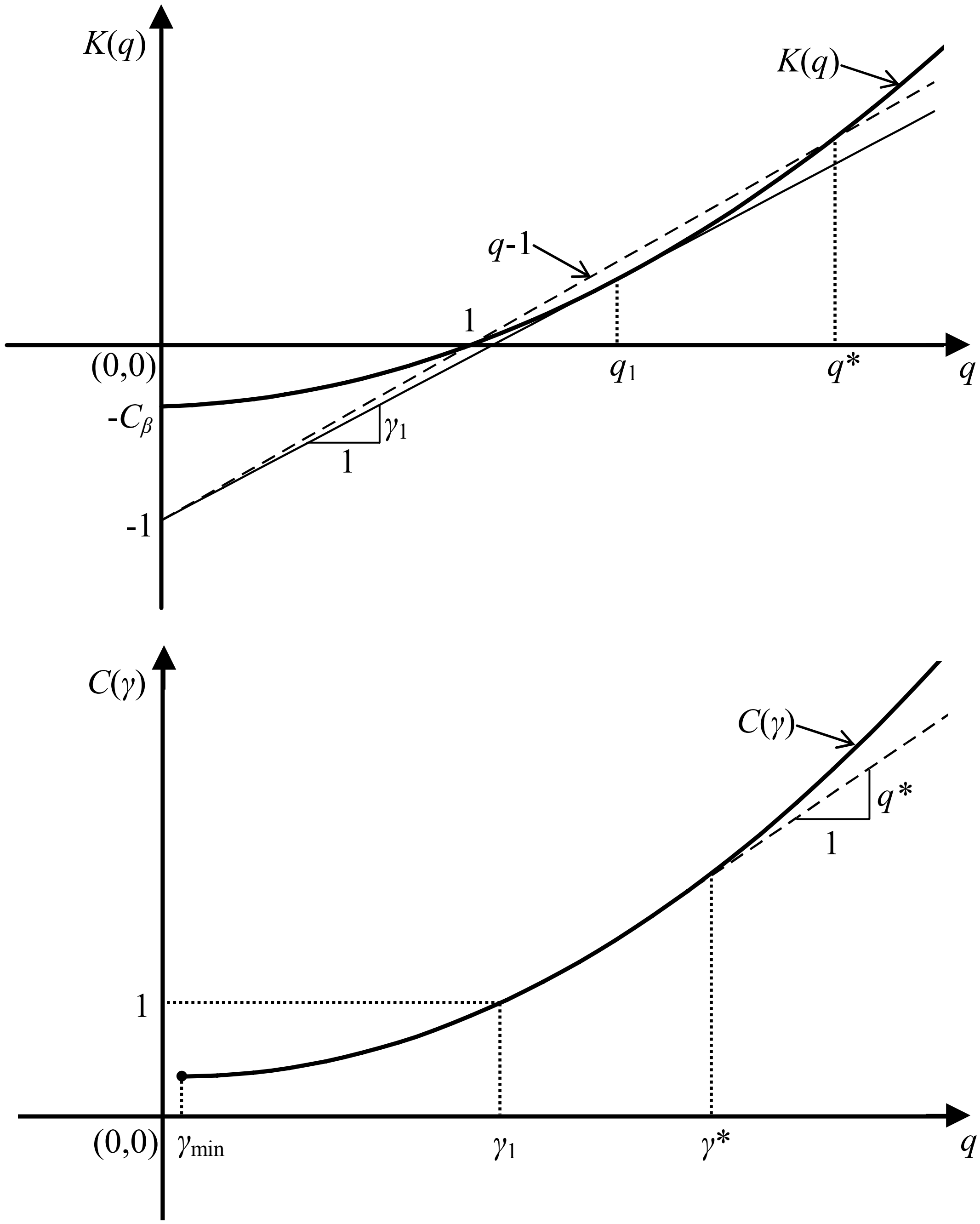

Figure 1: Illustration of the moment scaling function $K(q)$ and its Legendre transform $C(\gamma)$ in equation (6).

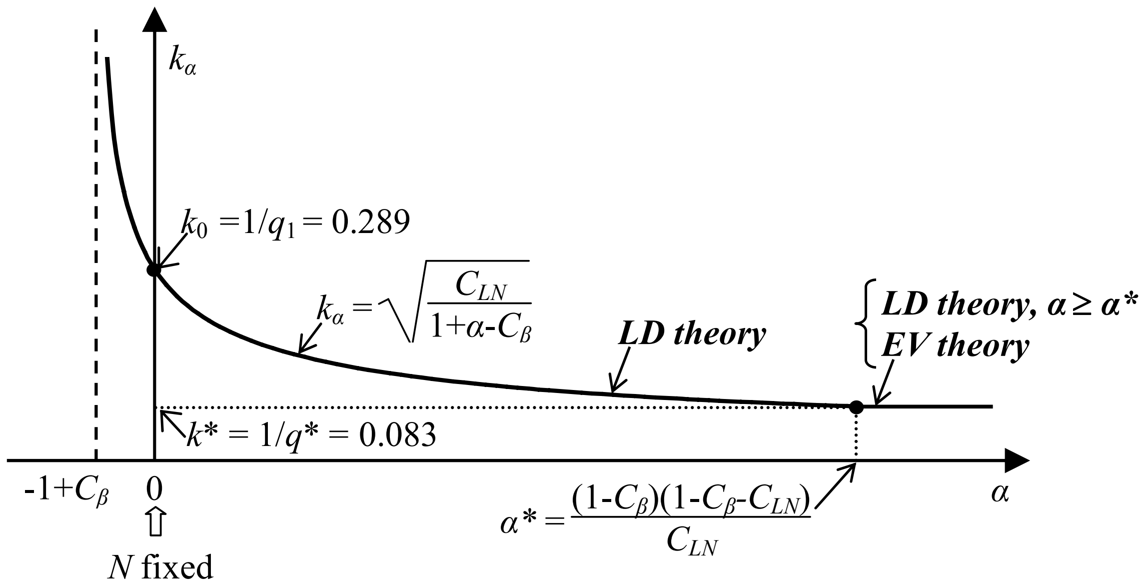

Figure 2: Shape parameter k_α of the N-year maximum of I_r under ($r \to \infty$, $N = cr^\alpha$). Beta-lognormal rainfall process with C_β = 0.4 and C_{LN} = 0.05. Larger values of k correspond to higher probabilities of exceedance of extreme rainfalls.

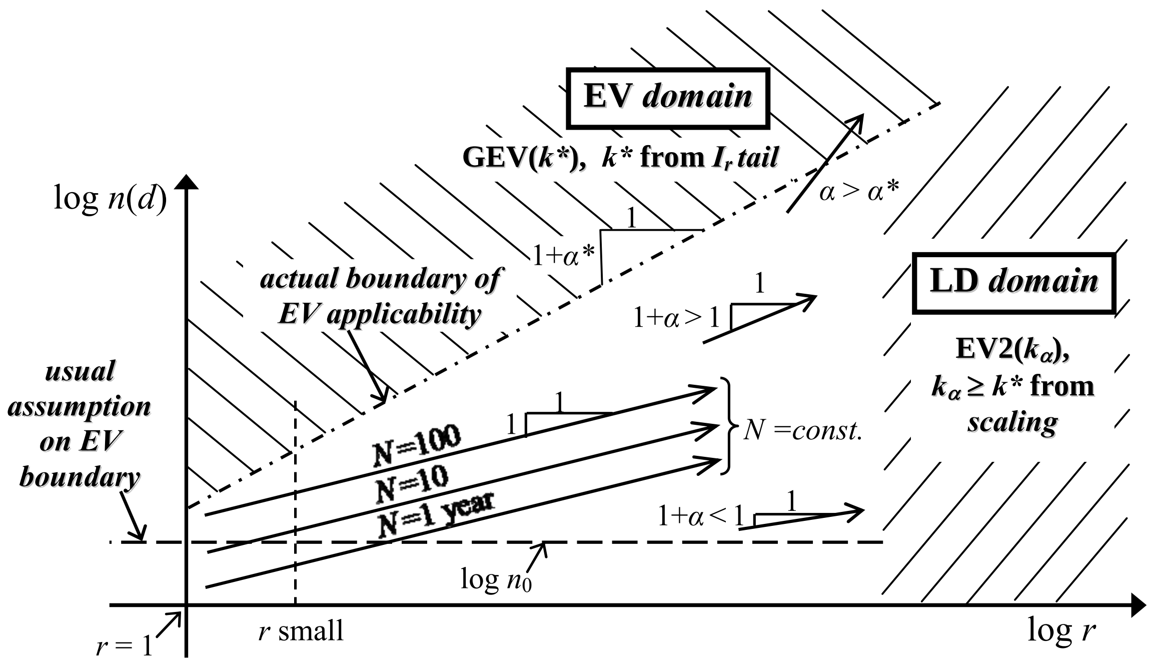

Figure 3: Schematic illustration of asymptotic results on rainfall maxima from extreme value (EV) and large deviation (LD) theories.

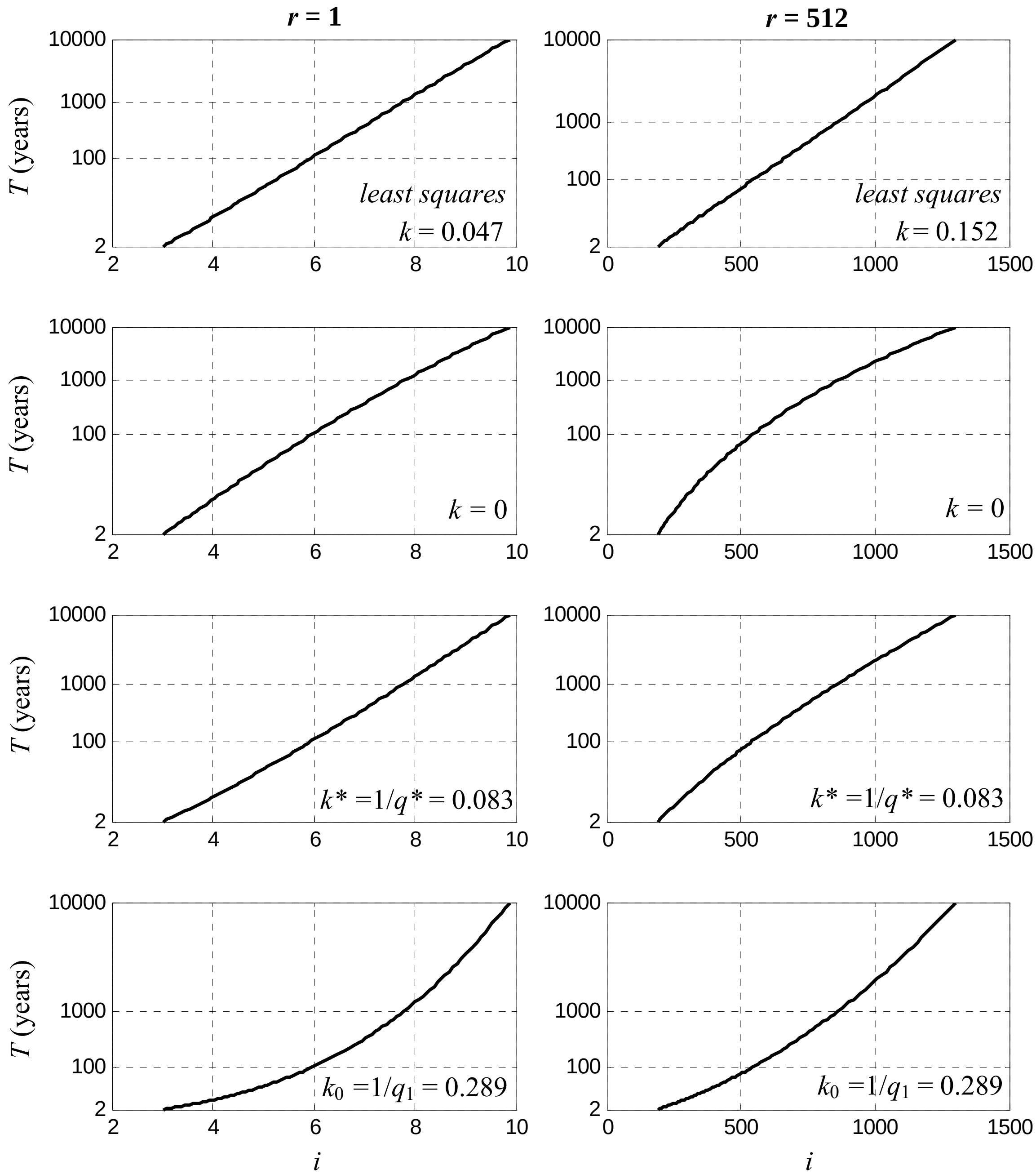

Figure 4: GEV(k) approximations to the exact distribution of the annual maximum $I_{yr,r}$ at resolutions $r = 1$ and 512, in the return-period range from 2-10 000 years. The top row shows the best least-squares fit on GEV(k) paper and gives the associated value of k. The lower rows show plots on GEV(k) paper for $k = 0$ (EV1 paper), $k^* = 1/q^*$ (value predicted by EV and EE theories), and $k_0 = 1/q_1$ (value predicted by LD theory for $r \to \infty$). Deviations of the plots from a straight line indicate lack of fit for the selected value of k.

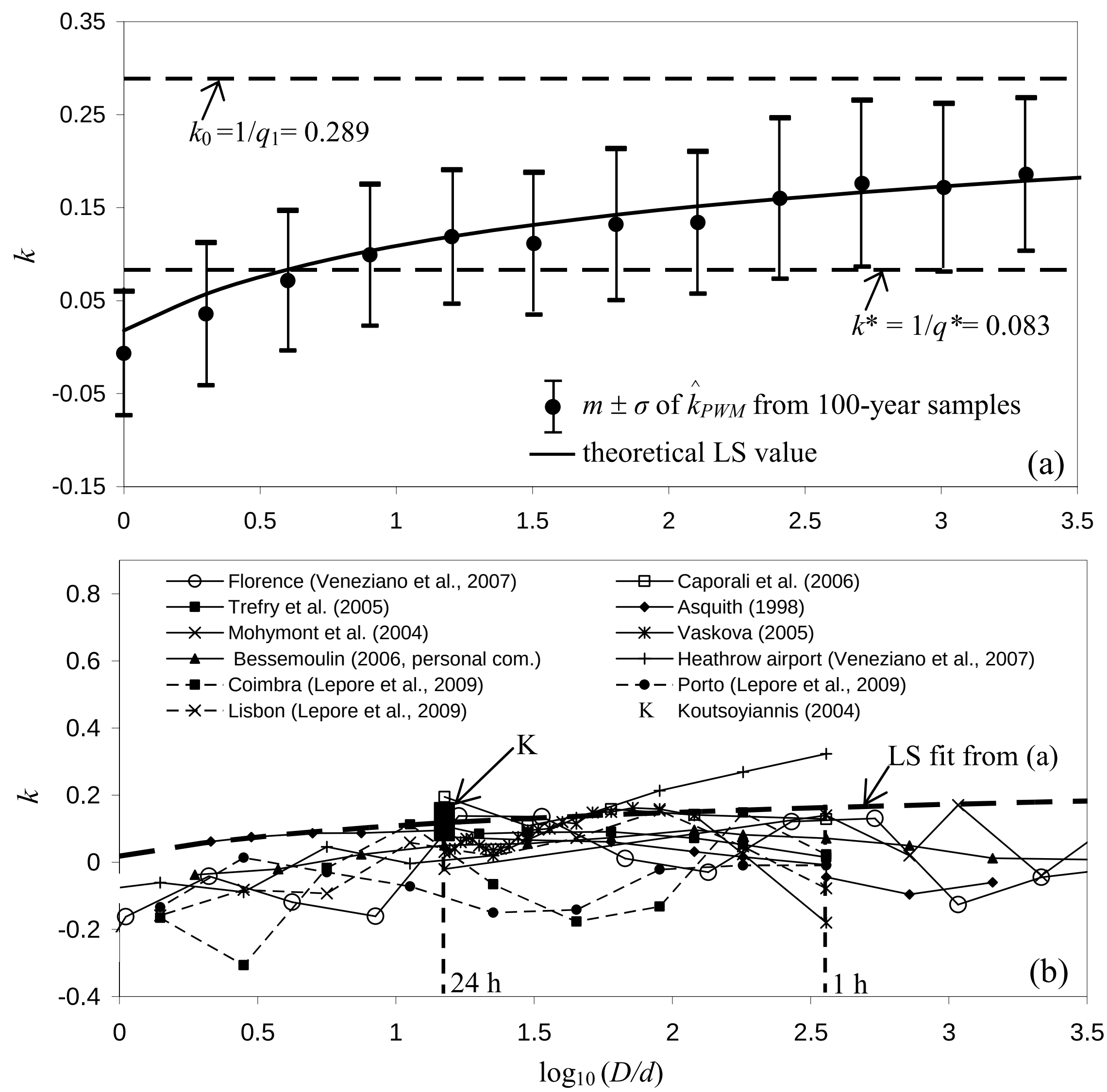

Figure 5: Dependence of the least-squares shape parameter k on the resolution $r = D/d$. (a) Theoretical values of k for $C_\beta = 0.4$ and $C_{LN} = 0.05$ when fitting is over the return period range from 2-100 years. The vertical bars are $(m \pm \sigma)$ intervals for the probability weighted moments (PWM) estimator of k using the annual maxima from 100-year continuous multifractal process simulations. The values $k^* = 1/q^*$ and $k_0 = 1/q_1$ are shown for reference. (b) Comparison of the theoretical values of k from (a) with empirical estimates from the literature assuming an average value of D = 15 days.

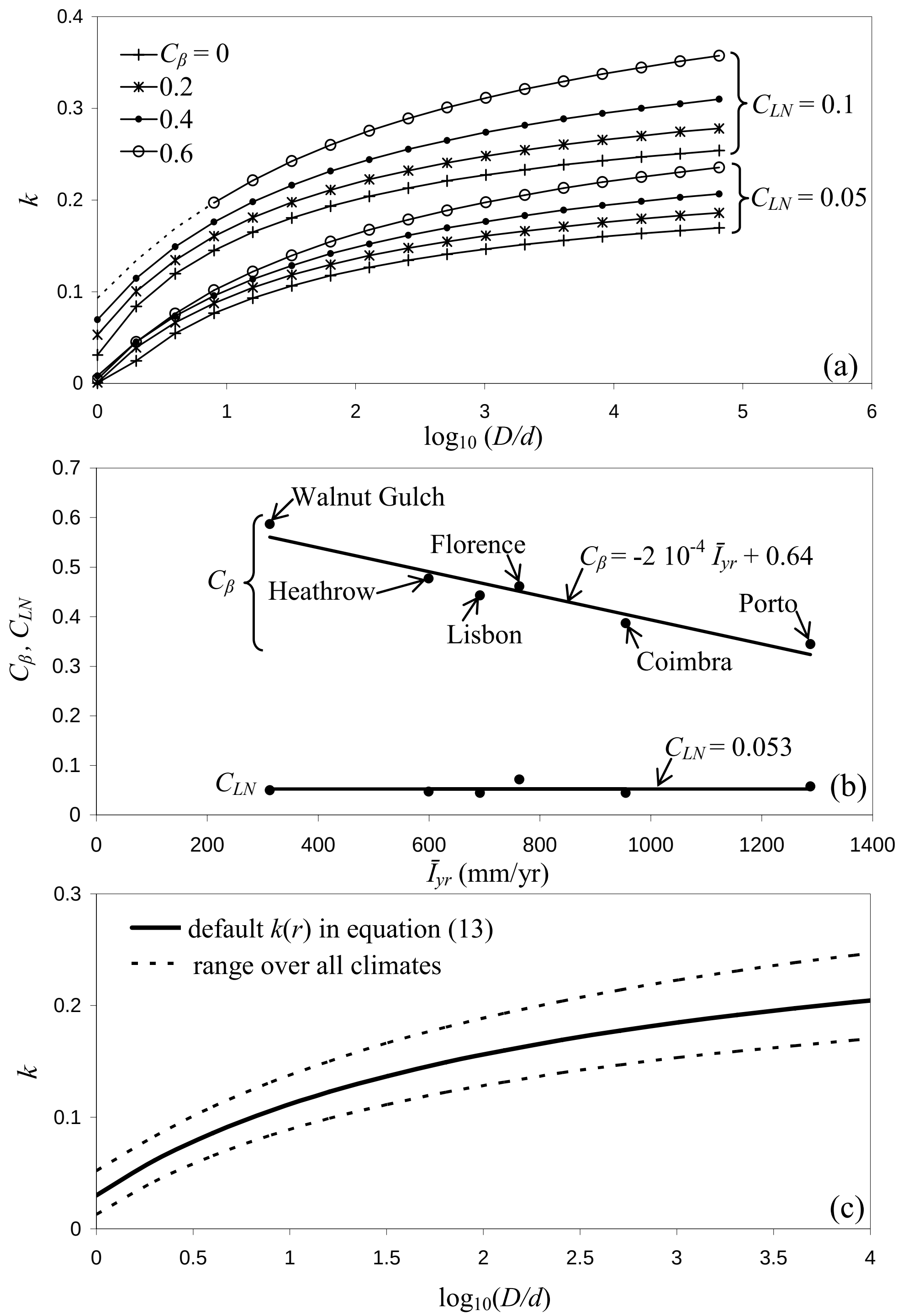

Figure 6. (a) Best-fitting shape parameters k at different resolutions r for selected combinations of C_β and C_{LN}. The range of return periods T used for fitting is from 2 - 100 years. (b) Estimates of C_β and C_{LN} from different rainfall records plotted against the mean annual precipitation. (c) Suggested default values of k as a function of the resolution $r = D/d$.