Multifractal Rainfall Extremes: Theoretical Analysis and Practical Estimation

By

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

[1]Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts, 02139, U.S.A.

[2]Dipartimento di Ingegneria Civile
Università degli Studi di Salerno
Fisciano (SA) 84084, Italy

Submitted to

Chaos, Solitons, and Fractals

February, 2007

Correspondence: Andreas Langousis, Dept. of Civil and Environmental Engineering, MIT, Room 1-245, Cambridge, MA 02139. Phone : (617)407-0059, Fax: (440)756-2310, e-mail: andlag@mit.edu.

Abstract

We study the extremes generated by a multifractal model of temporal rainfall and propose a practical method to estimate the Intensity-Duration-Frequency (IDF) curves. The model assumes that rainfall is a sequence of independent and identically distributed multiplicative cascades of the beta-lognormal type, with common duration D. When properly fitted to data, this simple model was found to produce accurate IDF results [Langousis A, Veneziano D. Intensity–duration–frequency curves from scaling representations of rainfall. *Water Resources Research*, 2007; 43: doi: 10.1029/2006WR005245]. Previous studies also showed that the IDF values from multifractal representations of rainfall scale with duration d and return period T under either $d{\to}0$ or $T{\to}\infty$, with different scaling exponents in the two cases. We determine the regions of the (d, T)-plane in which each asymptotic scaling behavior applies in good approximation, find expressions for the IDF values in the scaling and non-scaling regimes, and quantify the bias when estimating the asymptotic power-law tail of rainfall intensity from finite-duration records, as was often done in the past. Numerically calculated exact IDF curves are compared to several analytic approximations. The approximations are found to be accurate and are used to propose a practical IDF estimation procedure.

Keywords: rainfall extremes, multifractal processes, scale invariance, IDF curves

1. Introduction

The estimation of rainfall extremes is a central problem of stochastic hydrology [2-4]. These extremes are typically quantified through $\varepsilon_{d,T}$, the average rainfall intensity in an interval of duration d with a T-year return period. The Intensity Duration Frequency (IDF) curves are plots of $\varepsilon_{d,T}$ against d for different values of T.

The IDF values $\varepsilon_{d,T}$ depend somewhat on the definition of the return period T. When the rainfall record consists of annual maximum values, T is defined as the reciprocal of the exceedance rate by the annual maximum and $\varepsilon_{d,T}$ is obtained as the upper $1/T$-quantile of that maximum for duration d; see for example [5]. When a continuous rainfall record is available, T may be taken as the reciprocal of the rate at which a threshold intensity is upcrossed (so-called peak-over-threshold or partial duration method; see for example [6]). Finally, when rainfall is represented as a random process, it is often more convenient to define T as the reciprocal of the marginal exceedance rate. In this last case $\varepsilon_{d,T}$ is the value exceeded with probability d/T by ε_d, the rainfall intensity in a generic d interval. Except for very short return periods T, these definitions produce similar results [6-9].

Here we use the last quoted definition of T to study rainfall extremes under the condition that rainfall has multifractal scale invariance. This invariance property has been found to apply in good approximation to rainfall, typically for durations d from about 1 hour to several days [10-13]. As we shall show, multifractality produces interesting theoretical insights into the structure of the IDF curves and suggests simple approximations, which may prove attractive in practice. Elucidating these theoretical and practical consequences of multifractality is our main objective.

Multifractality controls certain asymptotic scaling properties of $\varepsilon_{d,T}$, but does not completely characterize the rainfall process or the IDF values; hence any derivation of the IDF

curves must be in the context of a specific multifractal representation of rainfall. Recently, Langousis and Veneziano [1] have studied the IDF curves under various multifractal rainfall models. The preferred model (Model 1 in that reference) idealizes rainfall as a sequence of independent and identically distributed storms separated by dry periods. The duration D and average intensity I of each storm are jointly distributed random variables and rainfall in D is described by a stationary multifractal process (a multiplicative cascade).

Simpler models (referred to as Models 2 and 3) were also considered in [1]. In particular, Model 3 partitions the time axis into intervals of constant (long) duration D and represents rainfall inside different D intervals as independent realizations of a stationary multifractal process of the beta-lognormal type; see Section 2 for details. The beta component of the process models the alternation of dry and wet periods and the lognormal component describes the fluctuations of rainfall intensity when it rains (for previous applications of the beta-lognormal multifractal process to temporal rainfall, see for example [9] and [14]).   As defined above, $\varepsilon_{d,T}$ is a quantile of ε_d. However, in multifractal analysis what is fundamental is not the averaging duration d, but the resolution $r = D/d$ (of course r and d are equivalent once D is specified). This is why here we work with ε_r, the average rainfall intensity in $d = D/r$. With this notation, the IDF rainfall intensity $\varepsilon_{r,T}$ is the value exceeded by ε_r with rate $1/T$ and satisfies $P[\varepsilon_r > \varepsilon_{r,T}] = D/rT$.

An important fact about multifractal extremes is that, as $r \to \infty$ or $T \to \infty$, $\varepsilon_{r,T}$ has a power-law dependence on r and T, with different exponents in the two limiting cases. Most previous IDF analyses using multifractal rainfall models have focused on these asymptotic results [e.g. 9, 15, 16], but the asymptotic scaling properties are not sufficient for hydrologic applications where one needs the actual IDF values for finite r and T. The present analysis aims at finding the IDF values and covers both the asymptotic scaling regimes and the transient non-scaling range.

Exact calculation of the IDF curves under Model 3 requires numerical analysis. Section 2 uses such an approach to illustrate the behavior of the IDF curves under multifractality. However, numerical analysis is not convenient to study general properties of the IDF curves. For this purpose, Section 3 introduces and assesses several approximations. The approximations are then used in Section 4 to derive interesting properties of the IDF curves and in Section 5 to develop a practical IDF estimation method. Application to a 50-year simulated record illustrates the procedure and its accuracy. Conclusions are stated in Section 6.

Throughout the paper, we employ standard probability techniques with a minimum of multifractal formalism. This makes the approach more transparent than using large deviation theory and multifractal concepts like singularities and their fractal dimensions, although the latter tools are more elegant and general. Reference to large-deviation and multifractal theories is made when interpreting the results.

2. The Exact IDF Curves Under Model 3

Temporal rainfall is said to be multifractal if its statistics remain unchanged when time is contracted by a factor $r > 1$ and the intensity is multiplied by some non-negative random variable A_r [17, 18]. The distribution of A_r controls the scaling and other properties of the rainfall process including the marginal distribution, the spacing and duration of the wet and dry intervals, and the extremes. In the case of Model 3, the distribution of A_r has a probability mass $1-r^{-C_\beta}$ at 0 and $(A_r | A_r > 0)$ has lognormal distribution with log-mean $m_{\ln(A_r|A_r>0)} = (C_\beta - C_{LN})\ln(r)$ and log-variance $\sigma^2_{\ln(A_r|A_r>0)} = 2C_{LN}\ln(r)$, where C_β and C_{LN} are non-negative parameters such that $C_\beta + C_{LN} < 1$. We call this distribution of A_r a beta-lognormal (β-LN) distribution and the resulting rainfall process, a beta-lognormal multifractal process (the "beta" nomenclature is from the multifractal literature and has no relation with the beta distribution).

5

Model 3 has four parameters: the outer limit of multifractal scaling D (typically around 10-15 days; see [1]), the average intensity I, and the multifractal parameters C_β and C_{LN} with typical ranges $0.4 < C_\beta < 0.6$ and $0.05 < C_{LN} < 0.10$. The parameter C_β controls the fractal dimension (1-C_β) of the rain support, whereas C_{LN} controls the amplitude of the intensity fluctuations when it rains. C_β and C_{LN} also determine the scaling of the moments of the average rainfall intensity ε_r for different resolutions r. In fact $E[\varepsilon_r^q] \sim r^{K(q)}$ with moment-scaling function $K(q)= \log_r E[(A_r)^q]$ $= C_\beta(q-1)+C_{LN}(q^2-q)$. This moment-scaling property is often used to infer C_β and C_{LN} from data [19].

Since $\varepsilon_{r,T}$ is proportional to I, in the analysis that follows we set $I = 1$. Given ($I = 1$, C_β, C_{LN}), one can calculate the marginal distribution of ε_r and from that obtain the exact IDF values $\varepsilon_{r,T}$. Calculation of the marginal distribution uses the fact that $\varepsilon_r = A_r Z$, where A_r is the beta-lognormal variable introduced above and Z is the so-called dressing factor, which accounts for the rainfall intensity fluctuations at resolutions higher than r [19]. The distribution of Z does not have analytical form, but it can be calculated numerically from C_β and C_{LN} using an iterative procedure [20].

Important features of the distribution of Z are the asymptotic Pareto upper tail $P[Z > z] \sim z^{-q^*}$ with exponent $q^*=(1-C_\beta)/C_{LN}>1$ [21] and the fact that, like A_r, Z has a nonzero probability mass at zero. Due to these properties of A_r and Z, $\varepsilon_r = A_r Z$ has asymptotic Pareto upper tail $P[\varepsilon_r > \varepsilon] \sim \varepsilon^{-q^*}$ and a probability atom at zero. Except for resolutions r close to 1, the rest of the distribution of ε_r is dominated by ($A_r | A_r > 0$) and is approximately lognormal. As r increases, the probability mass at 0 increases and the range of near-lognormality of ($\varepsilon_r | \varepsilon_r > 0$) broadens. For a more detailed discussion of the distributions of Z and ε_r, see [20].

To exemplify, we have calculated the distribution of ε_r for C_β =0.4, C_{LN} =0.05, and different values of *r*. Figure 1 shows the resulting IDF curves as plots of $\varepsilon_{r,T}$ against *r* for different *T/D* in the top panel and as plots of $\varepsilon_{r,T}$ against *T/D* for different *r* in the bottom panel. The plots in the top-panel are nearly straight and parallel for either (*r* large, *T* small) or (*r* small, *T* very large). As we show in Section 4, these regions correspond to scaling ranges of $\varepsilon_{r,T}$. Consistently with this behavior, the growth curves in the lower panel of Figure 1 are power-law functions for very short and very long return periods *T*, with a transient region that widens as *r* increases. The overall shape of these curves resembles the logarithmic function that, based on asymptotic properties of the Gumbel distribution, is sometimes assumed in empirical IDF estimation; see for example [22].

3. Approximations

Exact calculation of the IDF curves is tedious, mainly because obtaining the distribution of the dressing factor *Z* requires multiple convolutions [20]. In addition, the exact results do not have an analytical form. Hence for further theoretical analysis as well as practical use, it is important to develop approximations to the distribution of ε_r. Three approximations are introduced and evaluated in this section.

As discussed in Section 2, the conditional rainfall intensity ($\varepsilon_r | \varepsilon_r > 0$) has a near-lognormal body and an asymptotic Pareto upper tail. In the first approximation presented below, we replace ε_r with ε'_r such that ($\varepsilon'_r | \varepsilon'_r > 0$) is exactly lognormal below some value ε_r^* and exactly Pareto above ε_r^*. This operation produces simpler but still non-analytical IDF curves and is the steppingstone for the other two approximations, which further simplify the distribution of ε'_r. Since the simplifications lead to results that are consistent with the so-called rough and refined

limits of large-deviation theory, we refer to these approximations as the rough and refined ε'_r approximations.

3.1 ε'_r Approximation

The distribution of ε'_r is obtained in two steps:

Step 1 (approximation of Z and distribution of ε'_r below ε_r^*). To approximate ε_r below ε_r^*, the dressing factor Z is replaced by a β-LN variable A_{r_Z} where r_Z is chosen to match some characteristic of Z. It is especially convenient to match an integer moment of Z, since calculation of these moments is rather straightforward [20] and the moments of A_{r_Z} are given by $E[(A_{r_Z})^q] = r_Z^{K(q)}$ with $K(q) = C_\beta(q-1) + C_{LN}(q^2-q)$. With Z replaced by A_{r_Z}, ε'_r has the same beta-lognormal distribution as A_{rr_Z}. By comparing the exact distribution of ε_r with the approximating distribution of A_{rr_Z}, we have found that a good criterion to select r_Z is to match a moment of Z of order q close to $q^*/2$, where $q^* = (1-C_\beta)/C_{LN} > 1$ is the order beyond which the moments of Z diverge. For example, for $C_\beta = 0.4$ and $C_{LN} = 0.05$ one finds $q^* = 12$ and matching $E[Z^6]$ gives $r_Z = 4.36$. Using the $q^*/2$-moment matching criterion, Figure 2 gives r_Z for C_β and C_{LN} in the range of interest for rainfall. The lowest values of r_Z, around 2.0-2.3, are found for purely lognormal cases with $C_\beta = 0$.

Step 2 (ε_r^* and distribution of ε'_r above ε_r^*). To improve the above approximation in the upper tail, we "graft" a power-law tail to the distribution of A_{rr_Z} above the value ε_r^* such that the

log-log slope $\dfrac{d \ln P[A_{rr_Z} > \varepsilon]}{d \ln \varepsilon}\bigg|_{\varepsilon_r^*} = -q^*$. Hence we assume that $P[\varepsilon'_r > \varepsilon] \propto \varepsilon^{-q^*}$ for $\varepsilon > \varepsilon_r^*$. The proportionality constant is used to enforce continuity of the CDF of ε'_r at ε_r^*.

Figure 3 compares the exact distribution of ε_r with the distribution of A_{rr_Z} after Step 1 and the distribution of ε'_r after Step 2, at resolutions $r = 1$ and $r = 100$. Notice that for $r = 1$, ε_r is distributed like Z. For both values of r, the ε'_r approximation is very accurate.

An expression for the change-point ε_r^* that separates the lognormal body from the Pareto tail of ε'_r is given in Appendix A, Eq. (A.2). Another important quantity derived in Appendix A [Eq. (A.4)] is T_r^*, the return period of the event $\varepsilon'_r > \varepsilon_r^*$.

Having determined ε_r^* and T_r^*, the IDF curves under ε'_r are given by

$$
\varepsilon_{r,T} = \begin{cases} (rr_Z)^{C_\beta - C_{LN}} \, e^{\sqrt{2 C_{LN} \ln(rr_Z)}} \; \Phi^{-1}[1 - \dfrac{(rr_Z)^{C_\beta} D}{rT}] & , \; T \le T_r^* \quad \text{(a)} \\[2em] \varepsilon_r^* \left(\dfrac{T}{T_r^*} \right)^{C_{LN}/(1-C_\beta)} & , \; T > T_r^* \quad \text{(b)} \end{cases}
\tag{1}
$$

where Φ^{-1} is the inverse of the standard normal CDF. While rather simple, Eq. (1) is not convenient to study the properties of the IDF curves due to the non-analytic function Φ^{-1}. The approximations in the next two sections are better suited for this purpose. These approximations also provide links to extreme-value results from multifractal and large deviation techniques [9, 15, 23].

3.2 Refined Approximation of ε'_r

A well-known property of the normal distribution [e.g. 24] is that $\lim\limits_{x \to \infty} \dfrac{x[1-\Phi(x)]}{\varphi(x)} = 1$ where $\varphi(x)$, $\Phi(x)$ and $h(x)=\varphi(x)/[1-\Phi(x)]$ are respectively the probability density function (PDF), cumulative distribution function (CDF) and hazard function of the standard normal variable. This limit suggests replacing the exceedance probability $1-\Phi(x)$ with $\varphi(x)/x$ and the hazard function $h(x)$

with *x*. With these replacements, the analysis in Appendix A gives the following simplified

expressions for ε_r* and T_r*:

$$\varepsilon_r* \approx (rr_Z)^{\gamma*} \qquad\qquad\qquad\qquad (a)$$

$$T_r* \approx \frac{D}{r}\left\{(2\pi)2\ln(rr_Z)\frac{(1-C_\beta)^2}{C_{LN}}\right\}^{1/2} (rr_Z)^{[\frac{(1-C_\beta)^2}{C_{LN}}+C_\beta]} \qquad (b)$$

$$(2)$$

where $\gamma*=2 - C_\beta - C_{LN}$.

The IDF curves may be characterized by $T_{r,\varepsilon}$, the return period of the event $(\varepsilon'_r > \varepsilon)$ for

different ε and r, but here it is more convenient to write ε as $(rr_Z)^\gamma$ and find the return period $T_{r,\gamma}$

as a function of r and γ. Note from Eq.(2.a) that the lognormal and Pareto ranges of ε'_r

correspond to $\gamma \leq \gamma*$ and $\gamma > \gamma*$, respectively. One finds (see Appendix B)

$$T_{r,\gamma} \approx \begin{cases} \frac{D}{r}\left[(2\pi)2C_{LN}\left(\frac{\gamma-C_\beta}{2C_{LN}}+\frac{1}{2}\right)^2\ln(rr_Z)\right]^{1/2} (rr_Z)^{C_{LN}\left(\frac{\gamma-C_\beta}{2C_{LN}}+\frac{1}{2}\right)^2+C_\beta} & ,\ \gamma \leq \gamma* \quad (a) \\[3mm] \frac{D}{r}\left[(2\pi)2\ln(rr_Z)\frac{(1-C_\beta)^2}{C_{LN}}\right]^{1/2} (rr_Z)^{[1+(\gamma-1)\frac{1-C_\beta}{C_{LN}}]} & ,\ \gamma > \gamma* \quad (b) \end{cases}$$

$$(3)$$

If needed, the inverse relation ($\varepsilon_{r,T}$ as a function of r and T) can be calculated numerically.

An interesting feature of Eq. (3.a) is that it can be obtained also from large deviation

theory, using Cramer's refined limit [23, 25-27],. This is why we refer to the present

approximation as the refined approximation of ε'_r.

3.3 Rough Asymptotic Approximation of ε'_r

Further simplification follows from approximating the factor $\left[(2\pi)2C_{LN}\left(\frac{\gamma-C_\beta}{2C_{LN}}+\frac{1}{2}\right)^2\ln(rr_Z)\right]^{1/2}$ in

Eq. (3.a) by a constant. Values of r and γ of interest for hydrologic applications are in the ranges

$1 \leq r \leq 100$ and $\gamma_1 < \gamma < \gamma*$, where $\gamma_1 = C_\beta - C_{LN} + 2\sqrt{C_{LN}(1-C_\beta)} < 1$ and $\gamma* = 2 - C_\beta - C_{LN} > 1$. For example,

for $C_\beta = 0.4$ and $C_{LN} = 0.05$ one finds $\gamma_1 = 0.7$ and $\gamma^* = 1.55$. The lower limit γ_1 is such that values of

ε_r in excess of $(rr_Z)^{\gamma_1}$ occur approximately once per year, whereas $(rr_Z)^{\gamma^*}$ marks the lognormal-

Pareto transition of ε'_r. Combinations (r, γ) with $r > 10$ and $\gamma > 1$ produce very small exceedance

probabilities $P[\varepsilon_r > (rr_Z)^\gamma]$ and should not be considered. Under these constraints, the factor

$\left[(2\pi)2C_{LN} \left(\frac{\gamma - C_\beta}{2C_{LN}} + \frac{1}{2} \right)^2 \ln(rr_Z) \right]^{1/2}$ has low sensitivity to r and γ and may be approximated by a

constant δ. For example, for $C_\beta = 0.4$, $C_{LN} = 0.05$ and $r_Z = 4.36$, this factor ranges from about 3

(both r and γ small) to about 10 (r or γ large) and one may take $\delta = 5$. Moreover, δ is insensitive

to C_β and C_{LN} in the range of interest for rainfall.

With the δ approximation, Eq. (3) becomes

$$T_{r,\gamma} \approx \begin{cases} \dfrac{D\delta}{r} (rr_Z)^{C_{LN}\left(\frac{\gamma - C_\beta}{2C_{LN}} + \frac{1}{2}\right)^2 + C_\beta} & , \ \gamma \leq \gamma^* \quad \text{(a)} \\[4mm] \dfrac{D\delta}{r} (rr_Z)^{[1+(\gamma-1)\frac{1-C_\beta}{C_{LN}}]} & , \ \gamma > \gamma^* \quad \text{(b)} \end{cases} \tag{4}$$

Also the expression in Eq. (4.a) (for $\delta = 1$) can be derived from large-deviation theory, in this

case using Cramer's rough limit [23, 25, 26].

Unlike Eq. (3), Eq. (4) can be inverted analytically to give

$$\varepsilon_{r,T} \approx \begin{cases} (rr_Z)^{C_\beta - C_{LN} + 2\sqrt{C_{LN}[\log_{rr_Z}(rT/\delta D) - C_\beta]}} & , \ T \leq T_r^* \quad \text{(a)} \\[4mm] (rr_Z)^{1 + \frac{C_{LN}}{1-C_\beta}[\log_{rr_Z}(rT/\delta D) - 1]} & , \ T > T_r^* \quad \text{(b)} \end{cases} \tag{5}$$

where $T_r^* \approx \dfrac{D\delta}{r} (rr_Z)^{[\frac{(1-C_\beta)^2}{C_{LN}} + C_\beta]}$.

Figure 4 compares the exact IDF curves with the approximations in Eqs. (1), (3) and (5) for $\log_{10}(T/D)=1(2)9$. In all cases $C_\beta = 0.4$, $C_{LN}=0.05$ and $r_Z = 4.36$. The results from Eqs. (1) and (3) are almost identical and all approximations are accurate over a wide range of r and T.

4. Properties of the IDF Curves

Next we use the approximation in Eq. (5) to discuss two properties of the IDF curves under multifractality conditions. One is the scaling of $\varepsilon_{r,T}$ (for what ranges of r and T the IDF values have a power-law dependence on r and T and with what exponents). The other issue is the relation between IDF scaling and the Pareto upper tail that is often observed in the empirical distribution of ε_r.

4.1 Ranges of IDF Scaling

Equation (5) implies that $\varepsilon_{r,T}$ has power-law dependence on r and T/D under two conditions:

1. *In the high-resolution limit $r \rightarrow \infty$ or more specifically when $\log_{rr_Z}(rT/D) \rightarrow 1$. Writing*

   $\log_{rr_Z}(rT/D) = 1+\log_{rr_Z}(T/Dr_Z)$, this limiting condition is equivalent to $\log_{rr_Z}(T/Dr_Z) \rightarrow 0$. Since in this case $T \leq T_r^*$, $\varepsilon_{r,T}$ must be calculated from Eq. (5.a) Noting that here $\log_{rr_Z}(T/Dr_Z) \ll 1$, one finds after some algebra

$$\varepsilon_{r,T} \approx (\delta^{-1/q_1} \, r_Z^{\gamma_1-1/q_1}) \, r^{\gamma_1} \, (T/D)^{1/q_1} \tag{6}$$

   where $\gamma_1 = C_\beta - C_{LN} + 2\sqrt{C_{LN}(1-C_\beta)}$ and $q_1 = \sqrt{(1-C_\beta)/C_{LN}}$. Hence, at high resolutions r the IDF curves have power-law dependence on r and T. Numerical validation of this result is shown in Figure 1, where the scaling exponents γ_1 of r and $1/q_1$ of T are indicated.

2. *In the long-return-period limit and more precisely when $\log_{rr_Z}(rT/\delta D) \geq C_\beta + (1-C_\beta)^2/C_{LN}$. In* this case Eq. (5.b) applies. This equation may be written as

$$\varepsilon_{r,T} \approx (\delta^{-1/q^*} \, r_Z^{\, 1-1/q^*}) \, r \, (T/D)^{1/q^*} \tag{7}$$

where $q^* = (q_1)^2 = (1-C_\beta)/C_{LN}$. Hence also in this case $\varepsilon_{r,T}$ is a power-law function of r and T/D, but the scaling exponents are quite different from those in Eq. (6). Also the scaling relation in Eq. (7) is verified numerically in Figure 1.

The constants γ_1, q_1 and q^* in Eqs. (6) and (7) have a simple geometrical interpretation in terms of the moment-scaling function $K(q) = C_\beta(q-1)+C_{LN}(q^2-q)$: γ_1 is the slope of the tangent to $K(q)$ with K intercept equal to -1, q_1 is the moment order q at the point of tangency, and $q^* > 1$ satisfies $K(q^*)=q^*-1$; see [9]. Equation (7) was first obtained by [15] and Eq. (6) was derived by [9] using large-deviation techniques.

It is interesting that dependence of $\varepsilon_{r,T}$ on $r = D/d$ and T displays two power-law regimes for small and large values of the parameters. In these regimes the function $\varepsilon_{r,T}$ is separable, a property that is often assumed in practical IDF estimation [e.g. 3-5]. However, in the transient non-scaling range the function is non-separable. This behavior is different from any suggested in the past.

To better characterize the shape of the IDF curves and see whether in practice either scaling regime is of interest, we study the range of r and T under which each scaling behavior may be considered to apply. We start with the second scaling behavior, given by Eq. (7). The exact IDF curves satisfy Eq. (7) asymptotically as $T\rightarrow\infty$, whereas in the ε'_r approximation, Eq. (7) applies for $\log_{rr_Z}(rT/\delta D) \geq C_\beta+(1-C_\beta)^2/C_{LN}$. One may take the latter inequality as the practical range of validity of Eq. (7); see dashed-dotted lines in Figure 1 delimiting the "high-T scaling region".

To determine the region of approximate validity of Eq. (6), we consider the partial derivatives $\frac{\partial \log(\varepsilon_{r,T})}{\partial \log r}$ and $\frac{\partial \log(\varepsilon_{r,T})}{\partial \log(T/D)}$ of $\varepsilon_{r,T}$ in Eq. (5.a). The ratios between these derivatives and their asymptotic values from Eq. (6) are

$$S_r = \frac{\frac{\partial \log(\varepsilon_{r,T})}{\partial \log r}}{\gamma_1} = \frac{C_\beta - C_{LN} + \left(\xi + \frac{1}{\xi}\right)\sqrt{C_{LN}(1-C_\beta)}}{C_\beta - C_{LN} + 2\sqrt{C_{LN}(1-C_\beta)}} \quad \text{(a)}$$

$$\text{(8)}$$

$$S_{T/D} = \frac{\frac{\partial \log(\varepsilon_{r,T})}{\partial \log(T/D)}}{1/q_1} = \xi \quad \text{(b)}$$

where $\xi = \sqrt{\frac{1-C_\beta}{\log_{rr_Z}(rT/\delta D) - C_\beta}}$. As $r \to \infty$, these ratios approach 1. The "high-r scaling regions"

in Figure 1 satisfy $S_{T/D} > 0.9$ (in these regions, S_r is very close to 1, indicating that scaling with r is achieved over a wider range of the parameters than scaling with T). As Figure 1 shows, there is a wide transient range of r and T/D in which neither scaling relation holds. In this range the IDF curves are non-separable.

4.2 Pareto Tail of ε_r

The fact that ε_r has asymptotic Pareto tail with exponent $q^* = (1-C_\beta)/C_{LN}$ is often used to estimate q^* from empirical plots of $\log(P[\varepsilon_r > \varepsilon])$ against $\log(\varepsilon)$ [10, 11, 15, 28]. According to the refined ε'_r approximation, the minimum record length needed to observe that power law is given by T_r^* in Eq. (2.b). T_r^* is large and increases very rapidly as r increases. For example, for $C_\beta = 0.4$, $C_{LN} = 0.05$, $r_Z = 4.36$ and $D = 15$ days one finds $T_r^* \approx 1.5 \ 10^4$ years at the lowest resolution $r = 1$ and $T_r^* \approx 1.5 \ 10^6$ years for $r = 2$. These minimum lengths are far longer than any available rainfall record, implying that the empirical slope underestimates q^*.

Suppose that the record has length $T < T_r^*$. One can use Eq. (5.a) to quantify the bias when q^* is estimated as the negative slope of the empirical plot of $\log(P[\varepsilon_r > \varepsilon])$ against $\log(\varepsilon)$. Near the maximum observed value of ε_r, the theoretical value of this negative slope is

$$-\left(\frac{\partial \log(\varepsilon_{r,T})}{\partial \log(T/D)}\right)^{-1} = \eta \frac{1-C_\beta}{C_{LN}} = \eta q^* \tag{9}$$

where $\eta = \dfrac{\sqrt{C_{LN}[\log_{rr_Z}(rT/\delta D) - C_\beta]}}{1-C_\beta}$ is a bias factor. Figure 5 shows how η depends on r and T/D for $C_\beta = 0.4$, $C_{LN} = 0.05$, $r_Z = 4.36$ and $\delta = 5$. For example, for $T = 50$ years and $D = 15$ days, $T/D = 1217$ and η ranges from 0.68 for $r = 1$ to 0.39 for $r = 1000$. Since the q^* analysis is typically done at resolutions r on the order of 100, underestimation is by a factor around 0.4. This explains why the literature often reports estimates of q^* around 3.5 [e.g. 10, 15, 28]. These estimates are actually closer to $q_1 = \sqrt{q^*}$ than to q^*. In fact, as r increases, ε_r develops a Pareto range with exponent $q_1 = \sqrt{(1-C_\beta)/C_{LN}}$ over a wider set of return periods; see Eq. (6). This is consistent with the fact that, as $r \to \infty$, η in Eq. (9) approaches $1/q_1$ for any finite T. A related observation is the "straightening" of the empirical moment scaling function $\hat{K}(q)$ for q larger than about 3-3.5 [e,g. 29]. For a detailed analysis of the root cause of these phenomena (the fact that high-order singularities are not observable in one-dimensional cascades), see [30, 31].

5. A Practical IDF Estimation Method

Langousis and Veneziano [1] have found that Model 3 produces accurate IDF curves over the scaling range of durations, which in many temperate climates extends from about 1 hour to several days. Figure 4 further shows that little accuracy is lost if one replaces the exact IDF curves with the approximations in Eqs. (1), (3) and (5). Here we devise an IDF estimation

procedure using these approximations, as an alternative to the standard method of fitting extreme-type distributions to annual rainfall maxima.

The proposed method consists of four steps. The first three steps estimate the parameters of Model 3 and the final step estimates the IDF curves:

1. Estimate the mean rainfall intensity I as the average rainfall intensity in the historical record;

2. Using the scaling range of the empirical moments, estimate the $K(q)$ function and the multifractal parameters C_β and C_{LN}. We recommend choosing C_β and C_{LN} to reproduce the scaling of the moments of orders 0 and 3. This gives $C_\beta = -K(0)$ and $C_{LN} = [K(3)+2K(0)]/6$. The IDF results are insensitive to the parameter r_Z. One may obtain this parameter from Figure 2 or just fix it to a constant such as $r_Z = 4$;

3. Fit the empirical third moments with a log-log linear function of d. The parameter D is estimated as the value of d for which this function equals the theoretical third moment $I^3 r_Z^{K(3)}$;

4. Estimate the IDF curves for different averaging durations $d = D/r$ and different return periods T using Eqs. (1), (3) or (5), the latter with $\delta = 5$. In all cases, multiply the calculated IDF values by the average rainfall intensity I.

The reason why the value of r_Z is not important is that for any given r_Z one estimates D to reproduce the third empirical moment. What matters to the IDF curves is that this moment is reproduced, not the specific combination of r_Z and D. Different approximations in Step 4 produce similar IDF estimates. Use of Eq. (5) is simplest, but also implementation of Eqs. (1) and (3) is rather straightforward. A simplifying feature when using Eq. (1) is that one is interested mainly in the lognormal range; hence one needs only work with Eq. (1.a).

Figure 6 illustrates the above procedure using a 50-year continuous simulation of Model 3 with parameters $I = 1$, $D = 15$ days, $C_\beta = 0.4$, and $C_{LN} = 0.05$. The empirical average intensity for this realization is 1.012. From the moment plot in Figure 6.a one estimates $K(0) = -0.397$ and $K(3) = 1.086$ (the exact values are -0.4 and 1.1); hence the multifractal parameters are estimated to be $C_\beta = 0.397$ and $C_{LN} = 0.049$. Taking $r_Z = 4$, one finds the theoretical third moment of the average rainfall intensity in D to be $I^3 4^{K(3)} = 4.506$. This is matched by taking $D = 9.24$ days; see Figure 6.a. Notice that the exact value of D is not retrieved from the synthetic record. This is due to the fact that the value of r_Z has been arbitrarily set to 4, and hence D is found such that the combination (r_Z, D) reproduces the third empirical moment.

Finally, Figure 6.b and 6.c compare the IDF curves from Eqs. (1) and (5) with the exact IDF curves for the parameters used in the simulation. Since the parameter estimates are very good, the accuracy of the IDF approximations is comparable to that already observed in Figure 4.

6. Conclusions

It has been known for some time [9, 15] that the IDF curves generated by multifractal rainfall models display two asymptotic scaling behaviors, one as the averaging duration $d \to 0$ (or more precisely as the resolution $r = D/d \to \infty$ where D is the upper limit of the scaling range) and the other as the return period $T \to \infty$. Without knowing the ranges of (r, T) for which these scaling relations apply in good approximation, these asymptotic results have limited practical value. By using certain approximations to the exact IDF curves, we have determined these ranges. The results show that the two scaling regimes are widely separated and the scaling for $T \to \infty$ holds only for very long return periods; see the example Figure 1.

A consequence of this fact is that the asymptotic power-law tail of the rainfall intensity at resolution r, ε_r, cannot be observed unless the rainfall record is unrealistically long. This result

shows that the common practice of estimating the power-law decay from the empirical upper tail of the distribution of ε_r produces biased (too low) estimates.

Another current deficiency that we have addressed is the calculation of the actual IDF values (as opposed to calculation of the asymptotic scaling exponents) over the entire range of resolutions r and return periods T. For this purpose we have developed three different approximations, which are far simpler than numerical calculation of the exact IDF values. We have derived these approximations using classical probability methods, which are more transparent than large deviation theory and singularity analysis.

Finally, we have proposed an especially simple method to estimate the IDF curves under the condition of multifractality. The method performs well when it is applied to simulations from multifractal models. A more extensive assessment of method and comparison with traditional IDF estimation procedures will be presented in a companion paper.

Acknowledgments

Appendix A: Derivation of ε_r^* and T_r^* in the ε'_r Approximation

The change-point ε_r^* that separates the lognormal body from the Pareto tail of ε'_r is found from the condition $\left. \dfrac{d \ln P[A_{rr_Z} > \varepsilon]}{d \ln \varepsilon} \right|_{\varepsilon_r^*} = -q^*$ by noting that $\ln(A_{rr_Z} | A_{rr_Z} > 0)$ is a normal random variable with mean value $(C_\beta - C_{LN}) \ln(rr_Z)$ and variance $2 C_{LN} \ln(rr_Z)$. Therefore

$$\frac{d \ln P[A_{rr_Z} > \varepsilon]}{d \ln \varepsilon}\bigg|_{\varepsilon_r^*} = \frac{d \ln\{1-\Phi(x)\}}{d \ln \varepsilon}\bigg|_{x^*} = -\frac{1}{\sqrt{2C_{LN}\ln(rr_Z)}}h(x^*) \tag{A.1}$$

where $\varphi(x)$, $\Phi(x)$ and $h(x)=\varphi(x)/[1-\Phi(x)]$ are respectively the probability density function (PDF),

cumulative distribution function (CDF) and hazard function of the standard normal variable and

$x^* = \dfrac{\ln(\varepsilon_r^*) - (C_\beta - C_{LN})\ln(rr_Z)}{\sqrt{2C_{LN}\ln(rr_Z)}}$. To find ε_r^* such that the slope in Eq. (A.1) equals $-q^* = -(1-C_\beta)/C_{LN}$,

one first obtains x^* from the condition $h(x^*) = (1-C_\beta)\sqrt{2\ln(rr_Z)/C_{LN}}$ and then one calculates ε_r^*

using

$$\varepsilon_r^* = (rr_Z)^{C_\beta - C_{LN}} e^{\sqrt{2C_{LN}\ln(rr_Z)}\, x^*} \tag{A.2}$$

The return period T_r^* of the event $\varepsilon'_r > \varepsilon_r^*$ is found as follows. Notice that $P[\varepsilon'_r > 0] = (rr_Z)^{-C_\beta}$.

Therefore

$$P[\varepsilon'_r > \varepsilon_r^*] = P[\varepsilon'_r > 0][1-\Phi(x^*)] = (rr_Z)^{-C_\beta}\frac{\varphi(x^*)}{h(x^*)}$$

$$= \left[(2\pi)2\ln(rr_Z)\frac{(1-C_\beta)^2}{C_{LN}}\right]^{-1/2}(rr_Z)^{-C_\beta}e^{-(x^*)^2/2} \tag{A.3}$$

Since $T_r^* = \dfrac{D}{rP[\varepsilon'_r > \varepsilon_r^*]}$, one obtains

$$T_r^* = \frac{D}{r}\left[(2\pi)2\ln(rr_Z)\frac{(1-C_\beta)^2}{C_{LN}}\right]^{1/2}(rr_Z)^{C_\beta}e^{(x^*)^2/2} \tag{A.4}$$

Appendix B: Return Period $T_{r,\gamma}$ in Eq. 3

In the lognormal range (for $\gamma \le \gamma^*$), the exceedance probability $P[\varepsilon'_r > (rr_Z)^\gamma]$ is given by

$$P[\varepsilon'_r > (rr_Z)^\gamma] = (rr_Z)^{-C_\beta}\left[1-\Phi\left(\frac{\gamma\ln(rr_Z) + (C_{LN}-C_\beta)\ln(rr_Z)}{\sqrt{2C_{LN}\ln(rr_Z)}}\right)\right]$$

$$\approx \left[(2\pi) 2C_{LN} \left(\frac{\gamma - C_\beta}{2C_{LN}} + \frac{1}{2} \right)^2 \ln(rr_Z) \right]^{-1/2} (rr_Z)^{-C_{LN} \left(\frac{\gamma - C_\beta}{2C_{LN}} + \frac{1}{2} \right)^2 - C_\beta} \quad , \gamma \leq \gamma^* \tag{B.1}$$

where for the last expression we have replaced $1 - \Phi(x)$ with $\varphi(x)/x$ and used the equality

$$\frac{\gamma \ln(rr_Z) + (C_{LN} - C_\beta) \ln(rr_Z)}{\sqrt{2C_{LN} \ln(rr_Z)}} = \left[2C_{LN} \left(\frac{\gamma - C_\beta}{2C_{LN}} + \frac{1}{2} \right)^2 \ln(rr_Z) \right]^{1/2} \tag{B.2}$$

In the Pareto range (for $\gamma > \gamma^*$),

$$P[\varepsilon'_r > (rr_Z)^\gamma] = P[\varepsilon'_r > (rr_Z)^{\gamma^*}] \left(\frac{(rr_Z)^\gamma}{(rr_Z)^{\gamma^*}} \right)^{-(1-C_\beta)/C_{LN}}$$

$$= \left[(2\pi) 2\ln(rr_Z) \frac{(1-C_\beta)^2}{C_{LN}} \right]^{-1/2} (rr_Z)^{-[1+(\gamma-1)\frac{1-C_\beta}{C_{LN}}]} \quad , \gamma > \gamma^* \tag{B.3}$$

where $P[\varepsilon'_r > (rr_Z)^{\gamma^*}]$ has been evaluated using Eq. (B.1). Equation (3) follows from Eqs. (B.1)

and (B.3) and the definition $T_{r,\gamma} = \dfrac{D}{r\, P[\varepsilon'_r > (rr_Z)^\gamma]}$.

References

[1]   Langousis A, Veneziano D. Intensity–duration–frequency curves from scaling representations of rainfall. *Water Resources Research*, 2007; **43**: doi: 10.1029/2006WR005245.

[2]   Eagleson PS. *Dynamic Hydrology*. New York: McGraw-Hill, 1970.

[3]   Chow VT, Maidment DR, and Mays LW. *Applied Hydrology*. New York: McGraw-Hill, 1988.

[4]   Singh VP. *Elementary Hydrology*. New Jersey: Prentice-Hall, 1992.

[5]   Koutsoyiannis D, Kozonis D and Manetas, A. A Mathematical Framework for Studying Rainfall Intensity-Duration-Frequency Relationships. *Journal of Hydrology*, 1998; **206**: 118-135.

[6]   Madsen H, Rasmussen PF and Rosbjerg D. Comparison of Annual Maximum Series and Partial Duration Series Methods for Modeling Extreme Hydrologic Events: 1. At-site Modeling. *Water Resources Research*, 1997; **33**(4): 747-757.

[7]   Rosbjerg D. Return Periods of Hydrological Events. *Nordic Hydrology*, 1977; **8**: 57-61.

[8]   Willems P. Compound Intensity/Duration/Frequency Relationships of Extreme Precipitation for Two Seasons and Two Storm Types. *Journal of Hydrology*, 2000; **233**: 189-205.

[9]   Veneziano D and Furcolo P. Multifractality of Rainfall and Intensity-duration-frequency Curves. *Water Resources Research*, 2002; **38**(12): 1306-1317.

[10]   Lovejoy S and Schertzer D. Multifractals and Rain, In: *New Uncertainty Concepts in Hydrology and Hydrological Modelling*. Edited by: Kundzewicz AW. Cambridge press, 1995, pp. 62-103.

[11]   Tessier Y, Lovejoy S and Schertzer D. Universal Multifractals in Rain and Clouds: Theory and Observations. *Journal of Applied Meteorology*, 1993; **32**: 223-250.

[12]   Olsson J. Limits and Characteristics of the Multifractal Behavior of a High-Resolution Rainfall Time Series. *Nonlinear processes in Geophysics*, 1995; **2**: 23-29.

[13]   de Lima M I P and Grasman J. Multifractal Analysis of 15-min and Daily Rainfall from a Semi-arid Region of Portugal. *Journal of Hydrology*, 1999; **220**: 1-11.

[14] Schmitt F, Vannitsem S and Barbosa A. Modeling of Rainfall Time Series Using Two-state Renewal Processes and Multifractals. *Journal of Geophysical Research*, 1998; **103** (D18)**92**: 23181-23193.

[15] Hubert P, Bendjoudi H, Schertzer D and Lovejoy S. A Multifractal Explanation for Rainfall Intensity-Duration-Frequency Curves. In: *Proceedings*, International Conference On Heavy Rains and Flash Floods. National Research Council, Group for prevention from Hydrological Disasters, Istanbul, Turkey, 1999.

[16] Veneziano D and Langousis A. The Areal Reduction Factor a Multifractal Analysis. *Water Resources Research*, 2005; **41**: W07008, doi:10.1029/2004WR003765.

[17] Gupta VK and Waymire EC. A Statistical Analysis of Mesoscale Rainfall as a Random Cascade. *Journal of Applied Meteorology*, 1993; **32**: 251-267.

[18] Veneziano D. Basic Properties and Characterization of Stochastically Self-similar Processes in R^D. *Fractals*, 1999; **7**(1): 59-78.

[19] Schertzer D and Lovejoy S. Physical Modeling and Analysis of Rain and Clouds by Anisotropic Scaling of Multiplicative Processes. *Journal of Geophysical Research*, 1987; **92**: 9693-9714.

[20] Veneziano D and Furcolo P. Marginal Distribution of Stationary Multifractal Measures and Their Haar Wavelet Coefficients. *Fractals*, 2003; **11**(3): 253-270.

[21] Kahane JP, and Peyriere J. Sur certaines martingales de Benoit Mandelbrot. *Advances in Mathematics*, 1976; **22**: 131-145.

[22] Raudkivi AJ. *Hydrology, An Advanced Introduction to Hydrological Processes and Modeling*. Oxford: Pergamon, 1979.

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

[24] Abramowitz M. and Stegun IA (Eds.) *Handbook of Mathematical Functions*. New York: Dover Publications, 1965, pp. 932.

[25] Cramer H. Sur un noveau theoreme-limite de la theorie des probabilites, *Actualites Scientifiques et Industrielles,* No. 736 of Colloque consacre a la theorie des probabilities, Paris : Herrman, 1938, pp. 5-23.

[26] Varadhan SRS. *Large Deviations and Applications*. Philadelphia: Society for Industrial and Applied Mathematics, 1984.

[27] Dembo A and Zeitouni O. *Large Deviations Techniques and Applications*. Boston: Jones and Bartlett Publishers, 1993.

[28] Chaouche K, Hubert P and Lang G. Graphical Characterisation of Probability Distribution Tails. *Stochastic Environmental Research and Risk Assessment*, 2002; **16**: 342-357.

[29] Lashermes B, Abry P and Chainais P. New Insights into the Estimation of Scaling Exponents, *International Journal of Wavelets*, *Multiresolution and Information Processing*, 2004; **2**: 497-523.

[30] Ossiander M and Waymire EC. Statistical Estimation for Multiplicative Cascades. *Annals of Statistics*, 2000; **28**: 1533-1560.

[31] Ossiander M and Waymire EC. On Estimation Theory for Multiplicative Cascades. *The Indian Journal of Statistics*, 2002; **64**: 323-343.

Figure Captions

Figure 1: (a) $\varepsilon_{r,T}$ as a function of r for $\log_{10}(T/D) = 1(1)10$. (b) $\varepsilon_{r,T}$ as a function of T/D for $\log_{10}(r) = 0(1)4$. The dashed lines delimit the regions of approximate scaling; see Section 2.

Figure 2: Values of r_Z that reproduce the moment of Z of order $q^*/2$.

Figure 3: Comparison of the exact distribution of ε_r with the distributions of A_{rr_Z} and ε'_r for $r = 1$ and $r = 100$. Other parameters are $C_\beta = 0.4$, $C_{LN} = 0.05$ and $r_Z = 4.36$.

Figure 4: Comparison of exact (solid lines) and approximate IDF curves (dashed lines) for $\log_{10}(T/D) = 1(2)9$. Other parameters are $C_\beta = 0.4$, $C_{LN} = 0.05$ and, for the approximations, $r_Z = 4.36$.

Figure 5: Bias factor when estimating q^* as the empirical slope of $\log(\varepsilon_{r,T})$ against $\log(T)$. Model 3 with $C_\beta = 0.4$, $C_{LN} = 0.05$, $r_Z = 4.36$ and $\delta = 5$.

Figure 6: Model fitting and IDF curve estimation using a 50-year simulation of Model 3 with parameters $I = 1$, $D = 15$ days, $C_\beta = 0.4$ and $C_{LN} = 0.05$. (a) Empirical moments (circles), fitted straight lines to estimate $K(q)$, and estimation of D. (b,c) Comparison of exact IDF curves (solid lines) with IDF values from either Eq. (1) or Eq. (5) (dashed lines). All curves are for return periods $\log_{10} T(yr) = 1(2)9$, increasing from below.
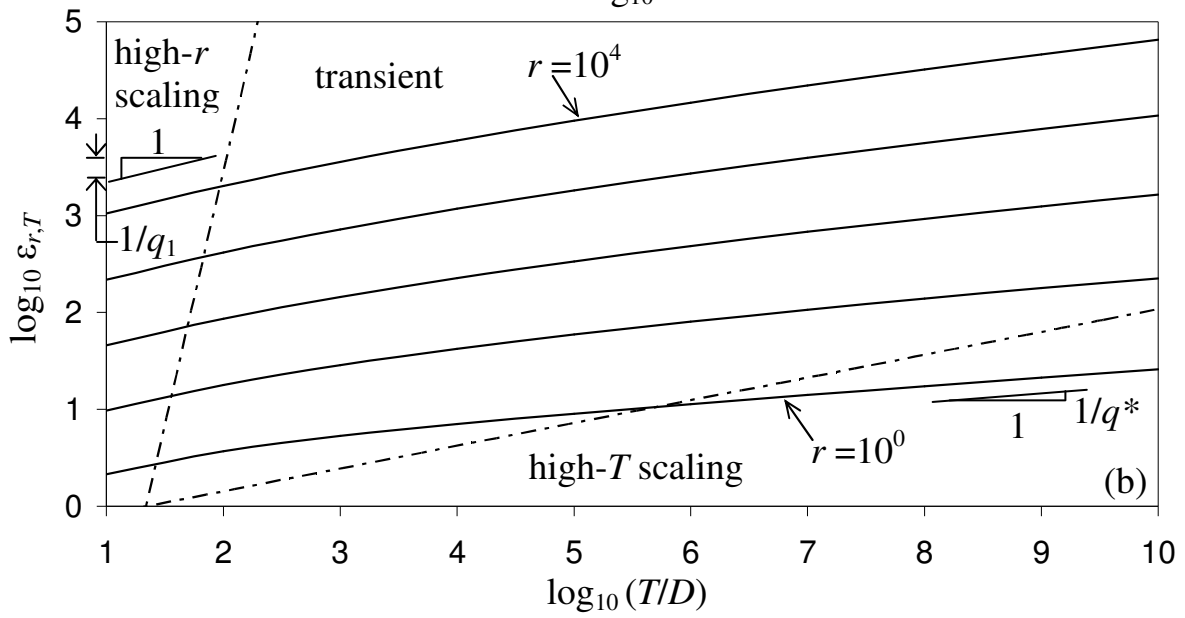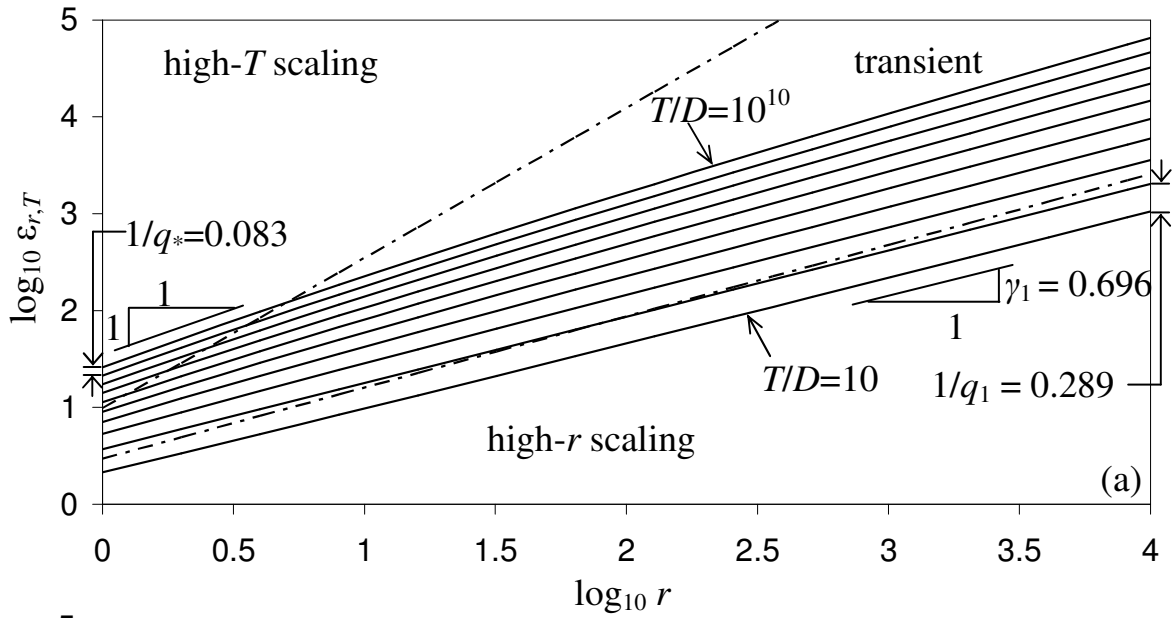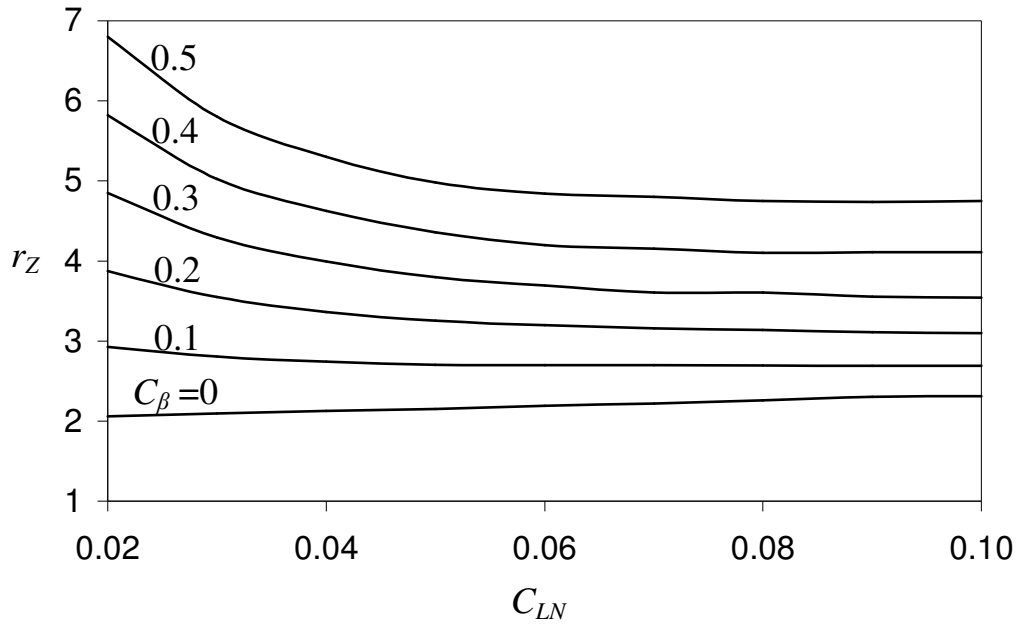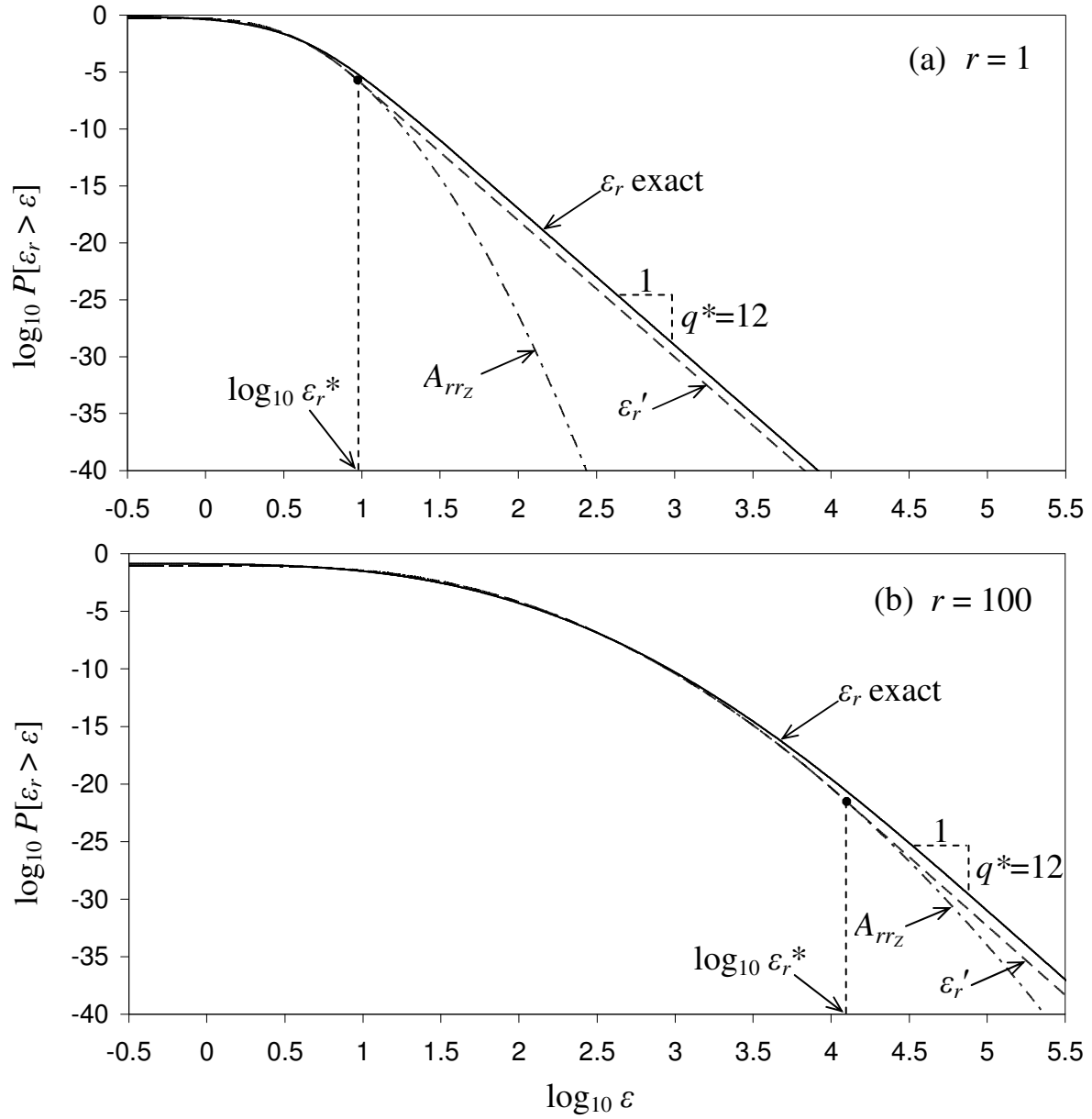
Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6