A Simple Approximation to Multifractal Rainfall Maxima using a Generalized Extreme Value Distribution Model

Andreas Langousis[1], Roberto Deidda[2], and Alin Carsteanu[3]
[1] Department of Environmental Engineering, Technical University of Crete, Chania, Greece
[2]Dipartimento di Ingegneria del Territorio, University of Cagliari, Italy
[3]School of Physics and Mathematics, National Polytechnic Institute (ESFM-IPN), Mexico City, Mexico

Abstract

Define I_d to be the average rainfall intensity inside an interval of duration d, and denote by $I_{yr,d}$ the maximum of I_d in a year. Based on a standard asymptotic result from extreme value (EV) theory, assuming independence and distributional identity between the variables I_d, $I_{yr,d}$ is typically assumed to follow a generalized extreme value (GEV) distribution with shape parameter $k(d)$ that depends on the extreme upper tail of the distribution of I_d.

Estimation of $k(d)$ from either at-site or regional rainfall data is generally difficult for two reasons. The first is the poor knowledge of the upper tail of the distribution of I_d, even for long rainfall records. The other is more theoretical and it is related to the applicability of the asymptotic EV result, when the number n = 1yr/d of the d-intervals in a year (or, equivalently, the number n of the I_d variables over which the maximum $I_{yr,d}$ is taken) is finite.

In a recent study, Veneziano *et al.* (2009) showed that for multifractal rainfall and typical values of d, 1yr/d is too small for convergence of $I_{yr,d}$ to a GEV distribution. Hence, $k(d)$ cannot be derived from asymptotic arguments and it is influenced by a region of the distribution of I_d that is close to the body thereof, rather than its extreme upper tail.

Here, we propose a simple method to theoretically calculate the shape parameter $k(d)$ of a GEV distribution model fitted to $I_{yr,d}$, as a function of the averaging duration d. We do so by assuming that rainfall is stationary multifractal below some maximum temporal scale D, and estimate $k(d)$ by fitting a GEV distribution to the maximum of n = 1yr/d independent and identically distributed I_d variables.

1. Simple multifractal model for temporal rainfall

We approximate temporal rainfall as a sequence of independent pulses with constant duration D and average rainfall intensity $I_D = m\,Z$ (see Figure 1), where m is the long-term mean rainfall intensity, and Z is a unit-mean random variable; see below. For simplicity we set w.l.o.g. m = 1.

For temporal scales $d < D$ rainfall is assumed to be stationary multifractal. In this case:

$$I_d \sim A_r\, Z \quad (1)$$

where ~ denotes equality in all finite dimensional distributions, $r = D/d \geq 1$ is a contraction factor and A_r is a unit-mean random variable, independent of Z.

The distribution of Z, usually referred to as the *dressing factor*, can be calculated from the distribution of A_r through multiple (numerical) convolutions; see Section 2 below.

A frequent choice to realistically model the alternation of wet (intra-storm) and dry (inter-storm) periods is to use a beta-lognormal distribution model for A_r (Langousis and Veneziano, 2007; Langousis *et al.*, 2007). In this case the unit mean random variable A_r in equation (1) is the product of two independent random factors: a factor $A_{ln,r}$ with lognormal distribution [specifically, ln $A_{ln,r} \sim N(\mu,\sigma^2)$, where $\mu = -C_{ln}\ln r$ and $\sigma^2 = 2C_{ln}\ln r$], and a discrete factor $A_{\beta,r}$ such that $P[A_{\beta,r} = 0] = 1 - r^{-C\beta}$ and $P[A_{\beta,r} = r^{C\beta}] = r^{-C\beta}$. Hence,

$$A_r = 0, \text{ with probability } 1 - r^{-C\beta}; \text{ and}$$
$$A_r = r^{C\beta}\exp(\mu+Q\sigma), \text{ with probability } r^{-C\beta} \quad (2)$$

where Q is a standard normal variable, and C_β and C_{ln} are parameters that satisfy $C_\beta + C_{ln} < 1$. The parameter C_β controls the alteration of wet and dry intervals within D-periods, whereas C_{ln} is responsible for the intensity fluctuations when it rains. It follows from equations (1) and (2) that the moments of I_d satisfy

$$E[(I_d)^q] \propto d^{-K(q)}, \text{ for } q < q^* \quad (3)$$

where $q^* = (1-C_\beta)/C_{ln} > 1$ is the lowest moment order greater than 1 such that $E[(I_d)^q]$ diverges, and $K(q) = \log_r E[A_r^q] = C_\beta(q-1)+C_{ln}(q^2-q)$ is the associated moment-scaling function; see Figure 2.

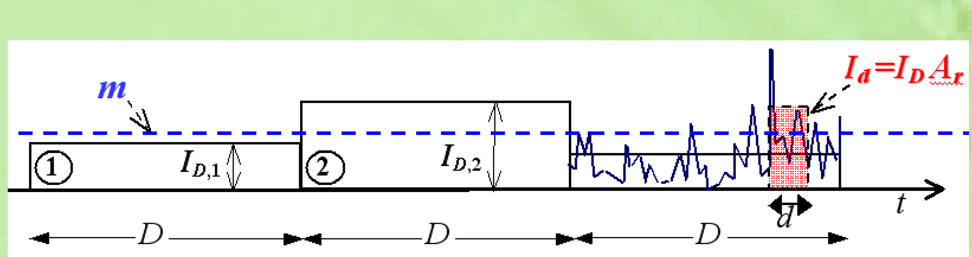

Figure 1: Rainfall as a sequence of pulses with constant duration D and internal multifractal scale invariance.

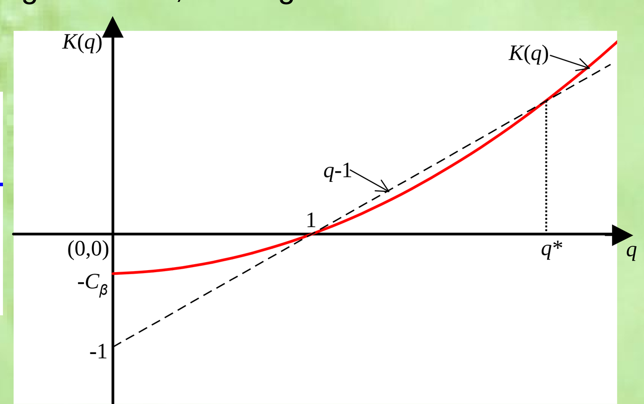

Figure 2: Schematic illustration of the moment scaling function, $K(q)$, of a beta-lognormal process.

2. Simple approximation to multifractal rainfall maxima

Calculation of the exact distribution of $I_{yr,d}$ is tedious, because it requires multiple numerical convolutions. Specifically, one needs to

1) calculate the distribution of Z following the iterative procedure of Veneziano and Furcolo (2002),
2) obtain the distribution of I_d from equation (1),
3) calculate the distribution of the maximum rainfall intensity averaged over duration d, $I_{max,d} := \max\{I_{d,1}, I_{d,2}, \ldots, I_{d,D/d}\}$, inside a generic D-interval, following the iterative procedure of Veneziano and Langousis (2005) (note that the variables $I_{d,j}$, j = 1, 2, .., D/d are identically distributed but not independent), and
4) calculate the distribution of $I_{yr,d}$ as the maximum of n = 1yr/D independent and identically distributed (iid) variables $I_{max,d}$. In this case $P[I_{yr,d} \leq i] = P[I_{max,d} \leq i]^{1/D}$, where D is in years.

In an effort to avoid tedious numerical operations, Langousis *et al.* (2007) found an analytical approximation to the T-year return period value of $I_{yr,d}$ (i.e. the value of $I_{yr,d}$ exceeded with probability 1yr/T, under ergodicity).

3. GEV(k) parameter estimation for multifractal rainfall maxima

A common method to estimate $i_{d,T}$ for different durations d and return periods T, even beyond the range of the available data, is to fit a generalized extreme value (GEV) distribution function directly to the empirical annual maxima:

$$F_{Id,T}(i) = \exp\{-[1+k(d)(-\psi(d)+i)/\lambda(d)]^{-1/k(d)}\} \quad (4)$$

where $k(d)$, $\psi(d)$ and $\lambda(d)$ are the shape, scale and location parameters of the distribution, respectively.

A limitation of the method, usually referred to as the *annual maximum* (AM) method of estimation, is that it uses only the annual rainfall maxima and discards the largest portion of the available hydrologic information. Consequently, independently of the method used for parameter estimation (i.e., maximum likelihood or probability weighted moments), the obtained estimates of the shape parameter k, which is informative for the upper tail of the fitted distribution model, are sensitive to outliers and exhibit significant variability.

When continuous rainfall data are available, an alternative method to estimate $k(d)$ is to

1) fit the multifractal model described in Section 1 to the continuous rainfall record,
2) calculate the distribution of $I_{yr,d}$ following Langousis *et al.* (2007), and
3) estimate $k(d)$ by fitting the GEV model in equation (4) to an appropriate range of exceedance probabilities $[1/T_2, 1/T_1]$ (say T_1= 2 years and T_2 = 100 years), of $I_{yr,d}$.

Figure 3 shows plots of the distribution of $I_{yr,d}$ on a GEV(k) paper, for C_β = 0.4, C_{ln} = 0.05 and r = 1 and 512, where k has been calculated from steps 1-3. The fact that both plots are very close to a straight line, even beyond the fitting range $[T_1, T_2]$, indicates that the distribution of $I_{yr,d}$ is well approximated by a GEV model with shape parameter k, which is insensitive to the fitting range $[T_1, T_2]$. Similar results have been obtained, also, for different combinations of the parameters C_β and C_{ln} and resolutions r.

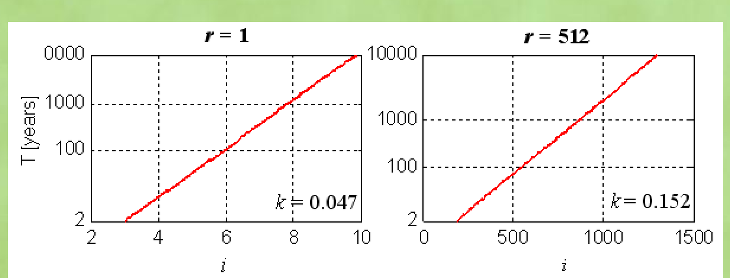

Figure 3: Plots of the distribution of $I_{yr,d}$ in equation (4), for C_β = 0.4, C_{ln} =0.05 and r = 1, 512, on a GEV(k) paper.

4. Robust estimation of the GEV(k) parameter for multifractal rain

For the range of temporal scales d where rainfall exhibits multifractal scale invariance, one can use the results presented in Sections 2 and 3 to propose a robust method to estimate the shape parameter $k(d)$ of a GEV distribution model fitted to annual rainfall maxima. The method includes the following steps:

1) Calculate the mean rainfall intensity m as the average rainfall intensity in the historical record.
2) For q = 0, 2 and 3, calculate the moments $E[(I_d/m)^q]$of the standardized rainfall intensity (I_d/m) as a function of the averaging duration d; see Figure 4a.
3) Use the scaling range of the empirical moments to estimate $K(q)$ as the negative slope of ln $E[(I_d/m)^q]$ against ln$d$; see Figures 4a and 4b.
4) Choose C_β and C_{ln} to reproduce the scaling of the moments of orders 0 and 3. This gives $C_\beta = -K(0)$ and $C_{ln} = [K(3)+2K(0)]/6$.
5) Use the values of C_β and C_{ln} to estimate the parameter D as the value of d for which $E[(I_d/m)^3] = r_0^{K(3)}$, where r_0 can be calculated as a function of C_β and C_{ln} from Figure 2 in Langousis et al. (2007); see Figure 4a.
6) Follow steps 2-3 in Section 3 to calculate the shape parameter k of a GEV distribution model that best fits the distribution of $I_{yr,d}$, in the exceedance probability range $[1/T_2, 1/T_1]$ (say T_1 = 2 years and T_2= 100 years).

To test our approach to estimate k, we used the multifractal model in Section 1 with parameters C_β = 0.4, C_{ln} = 0.05 and D = 15 days, to simulate a 4000-year realisation at approximately 30 min resolution. Figure 4 shows plots of the empirical moments $E[(I_d/m)^q]$ of order q = 0, 1, 2 and 3 as a function of the averaging duration d, the associated $K(q)$ function and the upper scale D of multifractal scale invariance, calculated using a sub-sample of 100 years.

For durations d = 1 hour and 1 day, Figure 5 shows estimates of the GEV(k) parameter calculated for 40 sub-samples of different sizes n = 5, 10, 25, 50 and 100 years. Estimation is done 1) using the suggested multifractal (MF) approach (red circles), and 2) by fitting a GEV distribution model directly to the empirical annual maxima (AM) using the method of probability weighted moments (PWM) (black circles). One sees that both methods are asymptotically unbiased, but the proposed MF method of estimation produces estimates of k with much lower variance relative to those of the traditional AM method. Figure 6 shows the standard deviation, σ_k, of the obtained estimates of k(d), for d = 1 hour and 1 day, as a function of the record length n in years. <u>One sees that the suggested approach has variance that is more than a factor of 10 lower relative to that of the traditional AM method.</u>

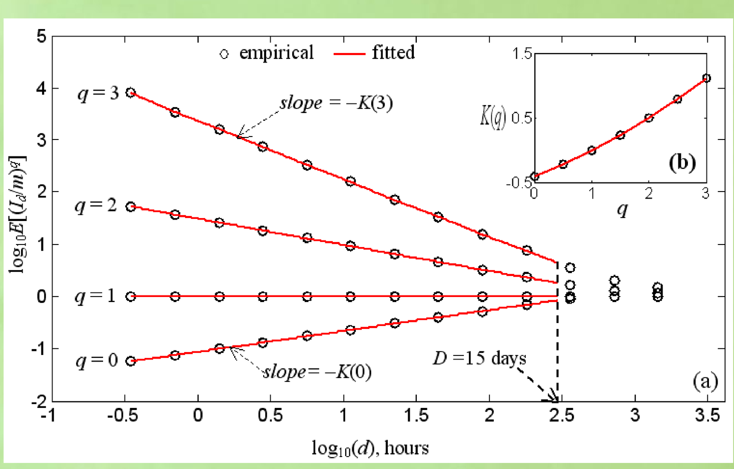

Figure 4: (a) Plot of the empirical (black circles) and fitted (red lines) moments $E[(I_d/m)^q]$ of order q = 0, 1, 2 and 3, as a function of the averaging duration d. (b) Plot of the empirical and fitted $K(q)$ functions. Fitting is done by calculating the parameters $C_\beta = -K(0)$ and $C_{ln}=[K(3)+2K(0)]/6$.

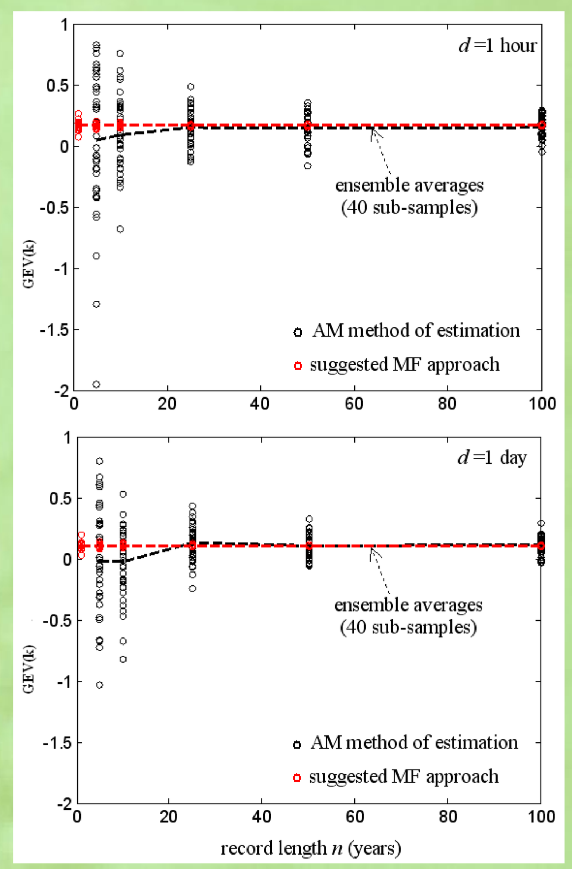

Figure 5: Estimates of the GEV(k) parameter for durations d = 1 hour and 1 day, calculated for 40 sub-samples of different sizes n = 5, 10, 25, 50 and 100 years.

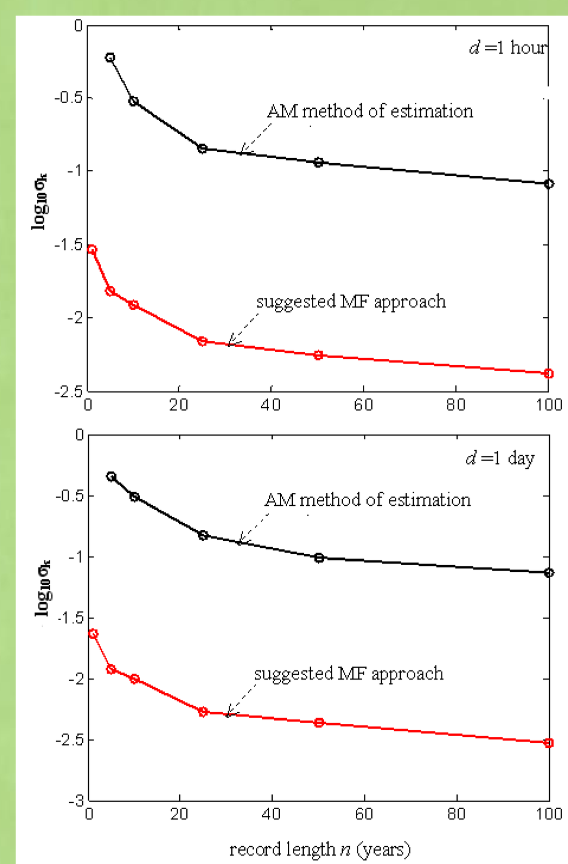

Figure 6: Standard deviation, σ_k, of the obtained estimates of k, for d = 1 hour and 1 day, as a function of the record length n in years.

5. Conclusions

We used multifractal theory to develop a method to estimate the shape parameter $k(d)$ of a GEV distribution model that best fits the annual rainfall maxima for different averaging durations d. Similarly to the traditional method of estimation, based annual maxima (AM), the suggested approach is asymptotically unbiased, but has variance that is more than 10 times lower. Since the parameter k specifies the shape of upper tail of the fitted distribution model, reducing its estimation variance is of critical importance for hydrologic applications, especially when one needs to extrapolate to return periods T beyond the range of the available data. The method applies for the range of durations d where rainfall exhibits multifractal scale invariance, but its applicability can be extended also beyond that range.

References
•Langousis A, D. Veneziano, P. Furcolo, and C. Lepore (2007), Chaos Solitons and Fractals, doi:10.1016/j.chaos.2007.06.004.
•Langousis, A. and D. Veneziano (2007), Wat. Resour. Res., 43, doi: 10.1029/2006WR005245.
•Veneziano, D. and A. Langousis (2005), Fractals, 13(4), 311-324.
•Veneziano, D. and P. Furcolo (2002), Wat. Resour. Res., 38(12): 1306-1317.
•Veneziano, D., A. Langousis and C. Lepore (2009), Wat. Resour. Res., 45, doi:10.1029/2009WR008257.

Acknowledgements

The first author, who is currently in Greece doing his 1-year mandatory military service, would like to thank his superior officers for granting him 3 full days and nights to think some aspects of the presented work and prepare parts of this poster.