Hydrol. Earth Syst. Sci. Discuss., 9, C4757–C4761, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C4757/2012/

Hydrology and Earth System Sciences Discussions

Interactive comment on "How extreme is extreme? An assessment of daily rainfall distribution tails" *by* S. M. Papalexiou et al.

S. M. Papalexiou et al.

smp@itia.ntua.gr

Received and published: 14 October 2012

CC: Commenter's comments
AC: Authors' comments

We thank Santiago Begueria for his useful comments. We are surprised that the Commenter is surprised because our paper is not about extreme value theory (EVT) as EVT is not identical with the study of the distribution tails. The Commenter may be interested in other works we have contributed as listed in the reference list of the paper (e.g., Koutsoyiannis, 2004) and as well as a new study under review (Papalexiou and Koutsoyiannis, 2012) (a predecessor presentation of a paper under review) where we analysed the same dataset but in the context of EVT.

CC1. *What the authors describe in their methods section is basically a peaks-over-threshold (POT) sampling, also termed exceedance series, partial duration series (as mentioned in the manuscript) or censored series. The authors cite correctly the article of Cunnane (1973) as one of the first studies to make a comprehensive use of this technique. A high number of authors have studied the statistical properties and methodologies associated to POT variables afterwards, two fundamental modern references being those of Coles (2001) and Katz et al. (2002).*

After defining the sampling of the extremes of the distribution, the authors concentrate on how to model the distribution of the magnitudes of the exceedances. However, a very relevant question is the probability distribution of the inter-event arrival times, i.e. wether the resulting process is random in time or not. Usually when modeling POT variables one or several tests such as the dispersion index (Cunnane, 1979) are performed to ensure that the process can be described by a Poisson (random) process. Leadbetter and co-authors (1983) demonstrated that a POT variable with random occurrence times belongs to the Generalized Pareto family, providing strong theoretical support for using this model. This was a fundamental milestone in POT analysis, and the Poisson/ Generalized Pareto (P/GP) model has been used afterwards by many authors for extreme value analysis (e.g. van Montfort and Witter, 1986; Hosking and Wallis, 1987; Wang, 1991; Martins and Stedinger, 2001). Numerical methods for obtaining sample estimates of the P/GP model parameters have been reviewed, among others, by Rao and Hamed (2000) and Coles (2001).

Although it is difficult at this stage, I strongly believe that the manuscript would greatly increase its interest if at least some relation to this very developed branch of the EVT would be considered in the introduction and discussion sections.

AC1. There are many relevant questions, so many that they cannot all be hosted in this paper. Regarding the question whether the process of rainfall arrival is random in time, again the Commenter is referred to other studies (see e.g., Koutsoyiannis, 2006). We believe though that the probability distribution of the inter-event arrival times is irrelevant to our present study. In our present study we clearly describe our method and obviously there are other approaches to follow. Regarding the theory that supports the use of the Generalized Pareto which begins with the results by Balkema and de Hann (1974) and Pickands III (1975), we may note that infinity is a tricky objects of mathematics. If the parent distribution is of power-type or of exponential type then the theory is applicable even for not so large threshold values as the tail obviously converges fast. In other cases, e.g., Lognormal or stretched exponential the convergence is only in theory. The same applies to the classical extreme value theory, which predicts that the distribution of maxima converges to one of the three extreme value distributions. For some examples exhibiting the slow convergence to the classical limiting distributions of EVT (the same philosophy applies for Balkema-de Haan-Pickands theorem) the Commenter can see slide no. 3 in Papalexiou and Koutsoyiannis (2012) and Fig. 1 in Koutsoyiannis (2004). We will add a comment about this in the revised manuscript.

CC2. *The selection of the threshold value for the POT sampling is still an open question, which has been addressed among others by Valadares-Tavares and Evaristo da Silva (1983), Coles (2001), and Beguería (2005). On page 5762 the authors propose what seems to be a new method of threshold definition. As this is most interesting to the POT community, I suggest that more information be provided. For example, knowing some statistics regarding the magnitude of the thresholds and their frequency (e.g. as quantiles in the EPF, or the average number of events per year) would allow comparing with other methods. Also, performing some test of time randomness such as the dispersion index, or at least indicating the mean and variance of the number of events per year would also be very relevant for interpreting the results, since this affects directly the assumption of time independence and affects the distribution of the extremes.*

AC2. We are not sure that the way we select the threshold constitutes a new method; for a time series of *N* years we simply study the *N* largest values. Now, regarding the

statistical properties of the threshold that emerges in this way, the Commenter's remark gave as the opportunity to study this a little further. Particularly, we deem that the most important issue here is to study the relationship of the threshold and its exceedence probability, i.e., to verify that indeed values above the threshold x_{L} can be regarded as extreme values for which the tail distribution applies. We have searched this further and we will add our analysis in the revised manuscript. Nevertheless, we note here that the exceedence probability $\bar{p}(x_{\mathrm{L}})$ of x_{L} can be proved that it is essentially related only to the probability dry, i.e.,

$$\bar{p}(x_{\mathrm{L}}) \approx \frac{1}{(1-p_0)\,\mathrm{CST}} \tag{1}$$

where p_0 is the probability dry and CST = 365.25 is the average number of days in a year (given that the time scale of our study is daily and we chose *N* equal to the number of years of the record). Interestingly, for the average probability dry of the records analysed in this study, approximately 0.8, the exceedence probability of x_{L} is as low as 0.014, while even for $p_0 = 0.95$ its value is 0.055. We believe that values above this threshold can be assumed as tails. We also note that the Commenter in one of his studies (Beguería et al., 2009) chose the threshold to correspond to the 90th percentile, a value much larger than the one corresponding to our choice of threshold. Anyhow, we thank the Commenter for this comment as it gave us the opportunity to explore this issue further.

References

Balkema, A. A. and de Haan, L.: Residual Life Time at Great Age, Ann. Probab., 2(5), 792–804, doi:10.1214/aop/1176996548, 1974.

Beguería, S., Vicente-Serrano, S. M., López-Moreno, J. I. and García-Ruiz, J. M.: Annual and seasonal mapping of peak intensity, magnitude and duration of extreme precipitation events across a climatic gradient, northeast Spain, International Journal of Climatology, 29(12), 1759–1779, 2009.

Koutsoyiannis, D.: Statistics of extremes and estimation of extreme rainfall, 1, Theoretical investigation, Hydrological Sciences Journal, 49(4), 575–590, 2004.

Koutsoyiannis, D.: An entropic-stochastic representation of rainfall intermittency: The origin of clustering and persistence, Water Resources Research, 42(1), doi:10.1029/2005WR004175 [online] Available from: http://www.itia.ntua.gr/en/docinfo/675/ (Accessed 3 October 2012), 2006.

Papalexiou, S. M. and Koutsoyiannis, D.: A global survey on the distribution of annual maxima of daily rainfall: Gumbel or Fréchet?, in European Geosciences Union General Assembly 2012, p. 10563. [online] Available from: http://www.itia.ntua.gr/en/docinfo/1207/ (Accessed 3 October 2012), 2012.

Pickands III, J.: Statistical inference using extreme order statistics, the Annals of Statistics, 119–131, 1975.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 9, 5757, 2012.