Hydrol. Earth Syst. Sci. Discuss., 9, C2924–C2928, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C2924/2012/

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.

P. Willems

Patrick.Willems@bwk.kuleuven.be

Received and published: 11 July 2012

I agree with other reviewers that the paper is well written and deserves publication in HESS after revision to meet the comments of the different referees and commenters.

1) The comment by Referee Clauset on the possibility that the heavy-tailed pattern observed may be due to non-stationary light-tailed processes is valid. In my paper (Willems, 2000) I have shown that POT extremes of rainfall intensities follow a mixed or two-component exponential distribution. However, when these extremes are studied per storm type or limited to a season where one storm type dominates, one-component exponential distributions were found. Combining rainfall extremes from different seasons and/or storm types may lead to the wrong conclusion that the distribution is heavy

tailed. Calibration of a "wrong" heavy tailed distribution may in that case lead to a close match of the calibrated theoretical distribution and the empirical quantiles or exceedance probabilities, but may not lead to reliable extrapolations beyond the range of the empirical data.

2) To meet the concern raised in previous comment, next to matching the empirical quantiles or exceedance probabilities, reliable representation of the tail's shape and the asymptotic distribution properties towards higher return periods is of equal importance. This is certainly the case when the objective of the extreme value analysis is extrapolation of the distribution beyond the range of empirical return periods (as is the case in many engineering applications, and which is also the main focus of this paper). Note that in the statistical literature several methods have been proposed to directly estimate the distribution's shape parameter; see e.g. Beirlant et al. (1996), Kratz and Resnick (1996). Other methods are based on the analysis of asymptotic distribution properties in quantile plots (Willems et al., 2007). For heavy tailed datasets, the distribution's tail appears asymptotically linear towards the higher quantiles or return periods in a Pareto quantile plot (plot of the logarithmically transformed rainfall intensity versus logarithmically transformed exceedance probability). The asymptotic linear slope equals the (inverse of the) shape parameter. For datasets with exponential tails, asymptotic linear tail behaviour is observed in an exponential quantile plot (same as the Pareto quantile plot, but no logarithmic transformation applied to the rainfall intensity in ordinate). See also the similar comment by Referee Deidda (his comment 4).

3) As Referee Laio, I was surprised to read that the performance of the different theoretical distribution tails was evaluated based on the error on the exceedance probability. In engineering design applications, quantiles are indeed estimated for given exceedance probabilities or return periods rather than exceedance probabilities estimated for given rainfall intensities. In extreme value analysis based on the analysis of the tail behaviour in quantile plots (e.g. Csörgo et al., 1985; Beirlant et al., 1996), it is common to apply weighting factors to the MSE computation (e.g. Willems et al., 2007). Most common

are the weighting factors by Hill (1975).

4) Rather than prior fixing the number of extremes or the POT threshold, the threshold could be optimized by minimizing the MSE. The MSE will increase for the smaller exceedance probabilities due to the increased variance when the parameter estimation is based on a lower number of observations (increased statistical uncertainty). When more extremes are considered, the bias in the asymptotic distribution's tail may increase and consequently the MSE may increase. In the intermediate range, the optimal threshold can be selected at the threshold with minimum MSE. Statistically principled determination of the threshold was also proposed by Referee Clauset.

5) Rather than separating distribution tails in two categories, heavy and light tails, it is more common to use three classes of tails: heavy, normal and light. The shape parameter γ, also called 'extreme value index', is positive for heavy tails, zero for normal tails, negative for light tails. According to the sign of the extreme value index, the following three classes are traditionally considered for extreme value distributions: class I (for $\gamma > 0$), class II (for $\gamma = 0$), and class III (for $\gamma < 0$, having upper bound). The Generalized Pareto Distribution (GPD) (for PDS/POT extremes) but also the Generalized Extreme Value (GEV) distribution explicitly considers these three types for the same distribution. These types correspond with the three distribution families defined by the authors on p.5765 lines 14-15: sub-exponential, exponential and hyper-exponential.

6) I agree with Referee Deidda that when the shape parameter is close to 1, the most parsimonious model can be preferred because of the reduced variance in the parameter estimation.

7) It is indeed surprising that the GPD distribution was not considered by the authors given that the distribution of excess values over a threshold (PDS/POT extremes) converges to the GPD, as was shown by Pickands (1975). This distribution includes the Pareto type II distribution (heavy tailed for $\gamma > 0$) used by the authors, the exponential distribution ($\gamma = 0$, normal tailed) and light tailed distribution ($\gamma < 0$). Same comment

was made be other referees or commenters.

8) I am not sure that lognormal distributions are heavy tailed (p. 5765 line 5). As also indicated by Referee Laio in his comment 1, the lognormal distribution has an exponentially decaying tail.

9) I think the authors made a mistake on p.5764 line 18. When the shape parameter or extreme value index γ converges to zero, the Pareto type II distribution's tail degenerates to the exponential tail, and not for γ towards infinity as the authors write.

References:

Beirlant, J., Teugels, J.L., Vynckier, P., 1996. Practical Analysis of Extreme Values. Leuven University Press, Leuven.

Csörgo, S., Deheuvels, P., Mason, D., 1985. Kernel estimators of the tail index of a distribution. Ann. Statist. 13, 1050–1077.

Hill, B.M., 1975. A simple and general approach to inference about the tail of a distribution. Ann. Statist. 3, 1163–1174.

Kratz, M., Resnick, S., 1996. The qq-estimator and heavy tails. Commun. Statist. Stochastic Models 12, 699–724.

Pickands, J., 1975. Statistical inference using extreme order statistics. Ann. Statist. 3, 119–131.

Willems, P., 2000. Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types. Journal of Hydrology, 233, 189 – 205

Willems, P., Guillou, A., Beirlant, J., 2007. Bias correction in hydrologic GPD based extreme value analysis by means of a slowly varying function. Journal of Hydrology, 338, 221-236

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