Hydrol. Earth Syst. Sci. Discuss., 9, C2287–C2290, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C2287/2012/

Hydrology and Earth System Sciences Discussions

HESSD

9, C2287–C2290, 2012

Interactive Comment

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

S. Beguería

santiago.begueria@csic.es

Received and published: 18 June 2012

This is a well-written article on a highly relevant topic: how to best model the right tail of rainfall and other climatic variables. The authors used a vast dataset and conducted interesting analyses. However, I was very surprised for the lack of references to now very well established fundamentals of the extreme value theory (EVT). Also, I believe that a bit more information about the results of the threshold selection method would greatly increase the interest of the manuscript. I develop these comments in the following paragraphs.

1. Fundamentals of the extreme value theory

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

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.

2. Selection of the threshold value

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

HESSD

9, C2287–C2290, 2012

(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.

References:

Coles S. 2001. An Introduction to Statistical Modeling of Extreme Values. Springer-Verlag: London. Katz RW, Parlange MB, Naveau P. 2002. Statistics of extremes in Hydrology. Advances in Water Resources 25(8–12): 1287–1304. Cunnane C. 1973. A particular comparison of annual maxima and partial duration series methods of flood frequency prediction. Journal of Hydrology 18(3/4): 257–271. Cunnane C. 1979. A note on the Poisson assumption in partial duration series models. Water Resources Research 15: 489–494. Leadbetter MR, Lindgren G, Rootzen H. 1983. Extremes and Related Properties of Random Sequences and Series. Springer Verlag: New York. van Montfort MAJ, Witter JV. 1986. The generalized Pareto distribution applied to rainfall depths. Hydrological Sciences Journal 31(22): 151–162. Hosking JRM, Wallis JR. 1987. Parameter and quantile estimation for the Generalized Pareto distribution. Technometrics 29(3): 339–349. Wang QJ. 1991. The POT model described by the Generalized Pareto distribution with Poisson arrival rate. Journal of Hydrology 129: 263–280. Martins ES, Stedinger JR. 2001. Generalized maximum likelihood Pareto-Poisson estimators for partial duration series. Water Resources Research 37(10): 2551–2557. Rao AR, Hamed KH. 2000. Flood Frequency Analysis. CRC Press: Boca Raton, FL. Valadares-Tavares L, Evaristo da Silva J. 1983. Partial duration series method revisited. Journal of Hydrology 64: 1–14. Beguería S. 2005. Uncertainties in partial duration

HESSD

9, C2287–C2290, 2012

Interactive Comment

series modelling of extremes related to the choice of the threshold value. Journal of Hydrology 303: 215–230.

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

HESSD

9, C2287–C2290, 2012

Interactive Comment