Hydrol. Earth Syst. Sci. Discuss., 9, C2414–C2416, 2012
www.hydrol-earth-syst-sci-discuss.net/9/C2414/2012/

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.

A. Clauset (Referee)

Aaron.Clauset@Colorado.EDU

Received and published: 21 June 2012

The authors investigate the distributional structure of heavy rainfall events by fitting and comparing several "tail" models to the upper end of the rainfall distributions. The goal here is to better understand and estimate the probability of extremely large events, which are also extremely rare. Their rarity makes the problem of model fitting and testing difficult because precisely where we would like the most statistical power, we have the least empirical data.

In setting up their analysis, the authors assume that the underlying distribution generating rainfall events is stationary and therefore all events are drawn independently from some unknown underlying distribution. This is a common and reasonable assumption,

but it also raises the possibility that the heavy-tailed pattern observed is not due to hydrological processes that produce stationary heavy-tailed distributions but rather due to non-stationary light-tailed processes. Testing this hypothesis is an important open question given the authors' results favoring heavy-tailed distributions. However, it may not be necessary to explore this question within this particular publication, but it should at least be discussed as another possible explanation for the observed patterns. Since the data are timestamped, I expect a number of tests of non-stationarity would yield interesting results without much additional work.

Although the authors do not cast their work within the modern literature on extreme value theory in statistics (a comment made by another referee), I'm not too worried about this. In fact, there must be a physically imposed upper limit on the largest possible rainfall, which means the extreme tail of the distribution must be truncated by finite-size cutoff (exponential tail). The scientifically relevant questions, however, are whether this physical limit is low enough to impact any of the empirical data and what the shape of the distribution is below that cutoff. In this sense, many of the stronger results from extreme value theory may not apply and the central question of tail-fitting remains reasonable. Some points, however, do remain relevant, e.g., the classification of general tail structures, and the manuscript would be improved by at least briefly discussing these connections relative to the authors' stated goals.

As with many studies of rare events in empirical data, the authors are faced with the question of how to quantitatively identify the value above which the "tail" of the distribution may be modeled separately from its body. In their notation, this is the question of choosing x_L. I agree that there is currently no universally accepted method for choosing x_L; however, there are (more objective) methods with advantages over heuristic of choosing the largest N values that the authors employ. The issue is that choosing x_L too small means including some of the distribution's body in the empirical data, inducing bias in the subsequently estimated tail model parameters if the body follows a different structure than the tail, while choosing it too large means reducing the sample

size and the statistical power of any model comparison technique. An arbitrary choice of x_L will lead to an uncontrolled tradeoff between bias and variance, and the resulting conclusions may not be trustworthy. Although there is no single best way to objectively solve this problem, one increasingly popular approach is described in *SIAM Review* **51**(4), 661-703 (2009), which chooses x_L automatically and in a statistically principled manner for each data set.

Finally, one choice by the authors did mystify me: why use what is essentially a least-squares regression approach to fitting the distributional models when one could instead use the more universally accepted and more statistically principled approach of maximum likelihood? Using maximum likelihood to estimate the model parameters would also allow the comparison of models using powerful techniques like the Vuong likelihood ratio test. This would provide much stronger evidence in favor of one model over another, and would also allow the decision that two or more models are statistically indistinguishable given the current data. One approach to conducting this kind of test is described in the same *SIAM Review* article mentioned above. For the scientific questions being addressed here, likelihoods seem like a superior methodological approach and I would encourage the authors to consider them. Now, it may be that the authors' existing results would continue to stand under the likelihood approach, but they may not. Either way, the results and conclusions would be placed on more firm methodological footing.

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