Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2018-105-RC1, 2018

Hydrology and Earth System Sciences
Discussions

Interactive comment on "An objective cross-validation framework for mapping rainfall hazard based on rain gauge data" *by* Juliette Blanchet et al.

R. Katz (Referee)

rwk@ucar.edu

Received and published: 26 April 2018

GENERAL COMMENTS

This paper deals with the statistical modeling of the distribution of rainfall amounts within a region, especially focusing on extremes. The approach is computationally intensive, with the parameters of the rainfall distribution at individual sites being mapped across the region using (e.g.) Kriging or splines. Cross validation is applied to evaluate the performance of the candidate models (e.g., form of distribution and spatial interpolation technique). Challenges include the attempt to model the entire range of rainfall amounts (i.e., from near zero to the most extreme) with a single distribution.

It is claimed that the proposed approach to rainfall modeling is "general and could be applied to any region of the world" (p. 24, line 29). Yet some aspects of the approach seem tailored to the application to a specific region in France. In particular, seasonality is treated by dividing the year into two seasons, one in which extremes typically occur. Plus the model is fitted conditional on one of three possible weather patterns (WPs), based on the spatial correlation of rainfall for the region. Although the number of seasons and WPs could certainly be varied to model rainfall for other types of climate, it is not clear that the constraint of being limited to a quite small number of seasons and WPs could always permit an adequate fit.

Another potential limitation concerns the performance of the different forms of distributions fitted to rainfall amounts, particularly for extremes. Conclusions are drawn about "heavy tails" that could benefit by relying more on extreme value theory. The restriction to a single distribution may have distorted the performance for extremes, with some of the conclusions conflicting with results in the literature when only extremes are modeled.

For these reasons, I recommend that the manuscript be accepted for publication subject to revision.

SPECIFIC COMMENTS

(1) Generality of proposed approach

It seems like a crude approximation to consider only two seasons and assume stationarity within a given season. More realistic approaches include allowing the parameters of the rainfall distribution to gradually change depending on the time of year. Some regions of the world even have more than one wet season, indicating a limitation of the proposed approach.

Conditioning on a few WPs based on the degree of spatial correlation of rainfall is an intriguing and not very common approach. Alternatives in the literature have included either introducing a hidden state variable (likewise assuming only a few possible states), which would require much more involved calculations, or including an observed covariate (such as an index of atmospheric circulation), which would require its identification for a given region but could assume effectively infinitely many possible states. Other than convenience, the advantages of the proposed approach are not clear.

(2) Using extreme value theory to interpret results

It is concluded that a mixture of gamma distributions provides the best fit, especially for extreme high precipitation amounts. Yet a gamma distribution has a light tail, well known to not be heavy enough for precipitation extremes. Still it is argued that allowing the gamma distribution to vary depending on the season and on the WP induces a heavier tail (Figure 5).

If this claim were correct for seasonality, then it would appear that the apparent heavy tail is at least partly an artifact of ignoring seasonality. Yet there is some evidence in the literature (e.g., by explicity allowing some of the parameters in an extreme value analysis to vary within the year) that this is not necessarily the case.

Concerning conditioning on WPs, it is well known that a mixture of gamma distributions can induce a heavier tail than a single gamma. Yet I wonder whether a mixture involving only a few gamma distributions (i.e., only three for the wet season) is sufficient to produce a truly heavy tail (in the sense of extreme value theory).

So it may be informative to examine how well the gamma distribution (and the mixture of three gamma distributions depending on the WP) fits precipitation amounts in the wet season alone. As it stands, I worry that the results for extremes may have been distorted by the constraint of fitting a single distribution to all rainfall amounts.

(3) Assumption of temporal independence

It is effectively assumed that the rainfall amounts at an individual site, especially extreme high values, are temporally independent (e.g., second displayed equation on p.

8 and p. 9, line 7). But this assumption never appears to be explicitly stated or verified.

There is some evidence in the literature of "clustering" at high levels for time series of daily rainfall amounts at individual sites. Cross validation, depending on how it is implemented, would not properly account for the effects of such temporal dependence.

EDITORIAL COMMENTS

(1) p. 3, lines 16-17

Not clear how the "factor" is defined or calculated.

(2) p. 25, Figure 10

Three of the graphs are for the same quantity, mean of non-zero rainfall for different weather patterns. But the color coding varies making comparisons difficult.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2018-105, 2018.