

Hydrol. Earth Syst. Sci. Discuss., referee comment RC3  
<https://doi.org/10.5194/hess-2021-570-RC3>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on hess-2021-570

Anonymous Referee #3

---

Referee comment on "Technical Note: Flood frequency study using partial duration series coupled with entropy principle" by Sonali Swetapadma and Chandra Shekhar Prasad Ojha, Hydrol. Earth Syst. Sci. Discuss., <https://doi.org/10.5194/hess-2021-570-RC3>, 2022

---

The manuscript presents an approach based on entropy for choosing the most suitable statistical models to represent partial duration series of streamflow. In particular, it proposes to evaluate the combined entropy of the statistical models describing the arrival of peaks above a certain threshold and the magnitudes above this threshold.

The idea is interesting, especially because it advocates using additional criteria (i.e., the capability to represent occurrences of events exceeding a threshold) to the goodness of fit of theoretical distribution calibrated to magnitude exceedances above the threshold. However, the study has some issues which prevent reaching substantial conclusions. They are described below.

- Calculating entropy constitutes an additional step to the usually applied procedure in this field. The value of performing this additional step should be made clear. As I stated above, I see value in the evaluation of an additional criterion to the goodness of fit of theoretical distribution calibrated to magnitude exceedances above the threshold. However, what advantage does it actually bring with it? Does this method for choosing the most suitable statistical model improve its predictive power? The authors claim it does, but the support of this claim is not clear to me (see the next comment).
- The authors claim to discuss the predictive ability of the statistical model selected by means of the proposed entropy-based approach in Figure 9. The figure shows the flow value associated to 50 and 100 years return period, calculated by means of a generalized Pareto distribution calibrated to exceedances above a set of different thresholds. Confidence intervals of the estimates are also displayed. I do not understand what this figure tells about predictive ability. I would be happy to hear about it, in case I am missing something obvious. First of all, Log Pearson 3 is the most suitable statistical distribution according to the values of entropy, whereas results for generalized Pareto are shown here. Then, where do we see in Figure 9 a better predictive performance of the distribution suggested by the entropy metric? Also, its predictive performance is better compared to what? I guess it should be better compared to the performance of the statistical model that would have been chosen based on goodness-of-fit metrics displayed in Figure 7 (see the next comment about

the interpretation of those results). I also do not understand how the bootstrapping was performed: could you provide a number for the length of data used for each resampling (line 344)? In addition, the authors state at line 365 that the proposed method "gives more accurate optimum threshold values". Based on what facts do they claim the threshold identified from the entropy metric to be more accurate? What is their reference value? Lines 358-363 simply discuss thresholds identified by means of alternative methods. If the value from the operational guidelines of Lang et al. (1999) (i.e., 730 m<sup>3</sup>/s, line 359) is used as reference (although this is also just another method) then Langbein (1949) would still provide a more accurate threshold (716 m<sup>3</sup>/s) than the proposed method (710 m<sup>3</sup>/s). Please clarify.

#### Additional points

- The proposed approach involves several steps which rely on visual observations and graphical analyses. These usually imply a high degree of subjectivity and difficulties to apply them to large datasets. It occurred to me that the approach described in section 2.4 to identify independent peaks is the same adopted by recent papers which leverage the Metastatistical Extreme Value framework to estimate flood magnitude and frequency from the whole series of ordinary peaks (i.e., with no need to define a threshold). Given that this novel statistical approach is gaining momentum, and that differently from the approach proposed here it can be completely automatized, it may be good to spend some words to justify the importance of identifying partial duration series by means of the classical peak over threshold methods.
- Some suggestions concerning the structure of the paper:
  - - More precise explanations of what is shown in the figures and how it enables to reach the stated results are needed. Just to give two examples: line 240: how did Kendall's Tau verified the independence of the series? How do we see it?; line 286: why finally the Poisson and not the Binomial distribution is chosen for the arrival of events above a threshold?
  - - Figures 1 to 5 display results of standard procedures which could be easily summarized with a few words in the text. Although this is a Technical Note where technical details shall be provided, Figure 2b-d only shows examples of results for arbitrarily chosen thresholds and Figure 4b is simply a zoom of Figure 4a. These figures could be deleted, which would help highlighting the actual results of the approach proposed in the paper (Figure 6).
  - - Figures with several panels could be condensed. For example, Figure 6 could be condensed to Figure 6f only; Figure 7 can be condensed in one single panel displaying total rank only.
- Several minor issues exist in the paper, especially related to correct and precise use of language (e.g., 22, 68, 163, 167, 174, 128-129, 130, 212, 216, 255), definition of symbols and units (symbols shall be introduced the first time a variable is named, e.g.,  $t$  is only defined at line 274 although appearing in Figure 1), differences between statements on the same subject (e.g., lines 88 and 314), motivations for showing these specific plots, given that many are examples for, e.g, different threshold (e.g., Figure 2 and 7). I do not detail them all here given the prior need to address the major issues described above. A carefully revision of the manuscript is however recommended.