Anonymous Referee 3

In their paper Blum et al. applied some well-known methodologies for finding suitable probability distributions for both period-of-record (POR) and median annual (MA) Flow Duration Curves (FDCs) in a very large area, such as the conterminous US. The authors found that, for the huge number of gauges analyzed, both the 4-parameter kappa and 3-parameter generalized Pareto distributions can reasonably simulate MA-FDC, while on the contrary even more complex distributions are unable to fit completely the very complex behavior of POR-FDCs, which explicitly accounts for extreme values. Furthermore, the authors also provide an example on possible application of their results for predicting FDC in ungauged sites, by means of the linear regression technique. While the paper does not present in my opinion any relevant novelty from the methodological point of view, the effort of the authors to fit FDCs to such a large dataset has to be underlined.

I have few minor comments about the manuscript, that I list below. I hope my comments can help to improve further the paper. Since the research does not deal with intermittent streams, and a relevant percentage of sites (170 on 590, almost 30%) was not considered into the analysis, I would suggest to slightly modify the title of the contribution, in order to make it more fitting with the content. I suggest something like this: "The probability distribution of daily streamflow in the perennial rivers of conterminous United States". Furthermore, some words would be appreciated about future research concerning intermittent streams in conterminous US.

The title will be modified as suggested and the revised manuscript will include a discussion on the need for future research on intermittent streams.

Paragraph 3.1 and Figure 3: due to the huge extension of the study area and the number of catchments analyzed, it would be interesting to verify if specific distributions fit better to specific regions or other climate/catchment features. I suggest to go at least a bit into details with this point. For example (but it's just an idea) points in Figures 3A and 3B can have different colors depending on different regions (and/or other climate/catchment distinctive features).

Excellent suggestion. We are adding an analysis that assesses the goodness-of-fit of the FDCs within each of 19 major hydrologic regions of the United States to supplement the nationwide results. We also intend to evaluate how climate and catchment characteristics can help explain variations in the goodness-of-fit of the various distributions. We will make a figure as you suggest for the revised manuscript and add additional text to compliment the figure.

P 11 l 18-20: I would rather say that "the selection [...] may be as challenging as [...]". However, among the theoretical advantages associated to the index flow method, there is the fact that complexity of Kappa and GPA distributions applied to the dimensionless daily streamflow is reduced, since the parameter alpha can be achieved as a combination of the other distribution parameters (please refer to Castellarin et al., 2007). This is a very important feature for regionalization studies. I would include this comment in the discussion.

Noted, we will revise the manuscript to acknowledge this point.

P 16 Eq. 7: I'm confused about using BFI as an explanatory variable, since to my knowledge it should be calculated/estimated from observed/estimated streamflow. Perhaps this variable can be replaced by some others accounting for the influence of lithplogical features on streamflows

This will be removed along with the entire case study due to various concerned raised by several reviewers. The BFI_AVE was employed because it is available as a grid generated by inverse-distance-weighting of base-flow index values computed at USGS streamgages for the entire United States (Falcone et al., 2010).

Falcone, J. A., D. M. Carlisle, D. M. Wolock, and M. R. Meador (2010), GAGES: A stream gage database for evaluating natural and altered flow conditions in the conterminous United States, *Ecology, 91*(2), 621–621, doi:10.1890/09-0889.1.

P 18 l 6-14: I acknowledge limitations and drawbacks of using POR-FDCs, but the discussion seems to me too 'biased' towards MA-FDCs. I suggest a more detailed discussion, so that also the final sentence (l 19: "MA-FDCs [...] should not be used when severe floods and droughts are of interest") is better contextualized.

We did not intend for the discussion to come across as promoting MA-FDCs and will revise the manuscript accordingly.

Finally, please consider to edit the text following the suggested corrections:
P 4 l 12: "When additional goodness-of-fit (GOF) metrics..." so that you can use the acronym later (from P 5 l 17 onwards)
Paragraphs 2.1 and 2.2: please correct the numbering
P 6 l 20: "Figure 6 illustrates the differences..."
P 7 l 16: I think that the sentence "where log represents the natural log" should be moved to line 11.
P 11 l 3: maybe it could be useful for the reader if authors comment a little bit more the figure, highlighting briefly why L-moment ratios simulated from WAK are less consistent than those simulated from KAP.
Captions Fig. 7 and Fig. 8: it is useful to highlight that lowest, median and highest NSE values are referred to GPA probability.
Figure 9 caption: I guess one number is missing concerning the number of outliers for HUC 10

Thank you very much for these comments. We will address these issues in the revised manuscript.