Novel extensions to the Fisher copula to model flood spatial dependence over North America

Reply to reviewer

1. Abstract highlights XGBoost as a key contribution for predicting discharge Kendall's τ coefficients, but the literature review lacks discussion or references to studies using XGBoost or similar machine learning models. It is suggested to include a review of relevant studies.

Studies using XGBoost or similar machine learning models are plentiful, but those which utilize machine learning methods to predict parameter values of a copula model are virtually inexistent. A reference is added in section 4.1 to a study showing that boosting models (especially XGBoost) are outperforming neural network on tabular data (line 223).

2. Methodology is complex, involving steps like the Fisher Copula, the XGBoost model, Kendall's τ interpolation, Conditional simulation, GEV parameter estimation, and Back-transformation. To help readers better understand the interrelationships and sequence of these steps, a flowchart is recommended.

Thank you for the suggestion. A flowchart has been included in the main article.

3. In Figure 9, the x-axis only shows the AAL and AML for return periods of 3, 10, and 30 years, while the focus of this study is on extreme discharge. It is suggested to explain why these specific return periods were chosen and whether observed and simulated data are available for longer return periods, such as 50 or 100 years.

The observed records have a limited length of 30 years, so the empirical return period levels can only be calculated up to return period 30 years.

4. In Figure 10, the discharge levels for return periods of 2, 20, and 100 years are all concentrated within 50-5000. It is suggested to explain why these discharge for different return periods appear to have similar ranges and to add units to the axes.

Some points corresponding to return period 100 years were left out in figure 10 due to the axis value limits. The axis value range has been corrected to include all points. The appearance of similar ranges is due to the logarithmic scale, but the leftmost points do correspond to return period 2 and rightmost points to return period 100 years. Units have been added to the axes. Figure 10 has been updated.

5. It is suggested to explain the reason behind choosing the number of catchments impacted by the event with a return period flow greater than 5 years as the denominator in the formula (π) and whether different thresholds were considered during the analysis.

The number of catchments impacted by the event with a return period flow greater than 5 years is used to define an event's footprint size, as floods smaller than the 1 in 5-year return period are unlikely to lead to significant losses. This choice is also made in [1], and we tried to use the same metric in order to compare our results with theirs. Different thresholds were not considered, as the event footprint size was defined that way earlier in the article.

6. In Fig. S3, it's puzzling why the same axis values (1e01, 1e01, 1e03, 1e05) are used across all subplots for return periods of 2, 5, 10, 50, 100, and 500 years. Typically, shorter return periods correspond to smaller discharges, while longer periods correspond to larger discharges.

The value range used to plot the discharges is wide enough to cover all the return periods from 2 to 500 years. A gradual shift toward higher values is visible as the return period gets higher, although this is small due to the logarithmic scale. The choice of the same axis values aims for a better homogeneity between the subplots.

References

[1] Quinn, N., Bates, P. D., Neal, J., Smith, A., Wing, O., Sampson, C., Smith, J., and Heffernan, J.: The Spatial Dependence of Flood Hazard and Risk in the United States, Water Resources Research, 55, 1890–1911, https://doi.org/10.1029/2018WR024205, 2019.