Reply to comments by Reviewer #3 Bailey Anderson

We thank Bailey Anderson for the helpful comments, and for the positive view of our manuscript.

For clarity, the authors' responses are inserted as blue text.

General comments

Review of the manuscript titled: "Attribution of changes in small and large floods across Brazil"

This manuscript investigates the drivers of observed changes in small and large floods across 765 catchments in Brazil by combining analyses of event-scale flood elasticities with trends in event rainfall peaks and antecedent wetness. The authors show that increases in large floods are primarily linked to changes in rainfall intensity, whereas changes in smaller floods are more strongly associated with shifts in antecedent wetness conditions. They further demonstrate that the relative importance of these drivers depends on catchment storage capacity, with antecedent wetness dominating in high-storage basins and rainfall peaks playing a larger role in low-storage systems.

The manuscript is clear, easy to follow, and I find the quality of the research to be generally high. The results are interesting and, in my opinion, useful to the community. For these reasons, I feel that the manuscript can be published subject to some revisions. My concerns are outlined below.

R: We thank Bailey Anderson for the detailed read of our manuscript and constructive comments.

We address each point in detail below.

Major comments

Major#1. Single linear regression model coefficients might be unstable. Why was this approach chosen over other methods and was any sensitivity analysis performed?

R: We agree that single linear regressions might be unstable. This important point was also highlighted by Reviewer #1 and Reviewer #2, who suggested that techniques such as a panel regression model may be more appropriate.

To assess the robustness of our results with respect to our modeling choice, we conducted an additional analysis using panel quantile regression. We run a panel regression model with fixed effects by polling catchments across Brazil and across our four hotspots (Fig. R1). We find that the regional median elasticities derived from the catchment-specific regressions are broadly consistent with those obtained from the panel quantile regression model. In our interpretation, the consistency between the estimates indicates that the results are unlikely to be methodological artifacts but rather reflect an intrinsic characteristic of the data. We have added a brief discussion of these methodological choices and their implications in the Material and Methods section of the revised manuscript (please, refer to L154-L158).

We have decided to retain the local quantile regression approach, as adopting a panel regression approach would require assuming that, within a given region, all catchments share a common average sensitivity (i.e., elasticity) of flood magnitudes to rainfall peak and antecedent wetness, potentially hiding important catchment-to-catchment heterogeneity in hydrological response. Our current framework is based on catchment-specific (local) quantile regressions, which allow us to explicitly account for spatial variability in elasticities and assess how sensitivities differ across individual catchments.

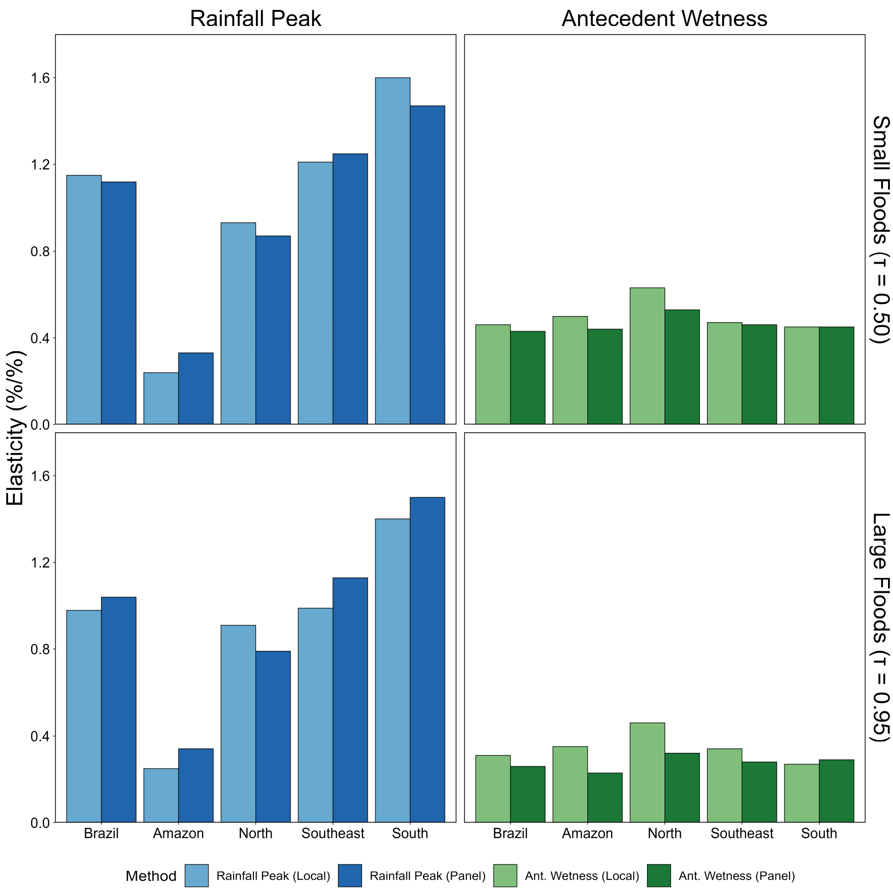

Figure R1: Comparison between elasticities estimation between local and panel quantile regression for small (0.50 quantile) and large floods (0.95 quantile).

Major#2. I am not saying that I disagree with the event selection approach (i.e. the focus on rainfall events) but I find it unusual, and I would appreciate a bit more explanation of the reasoning earlier in the paper. At present there is a bit in the discussion, but I would like to see it in the methods.

R: We agree that the event selection strategy deserves a clearer explanation in the Methods section. This point was also highlighted by (please, see Major Comment 2 from Reviewer #2)

We have therefore expanded the description of the sampling procedure earlier in the manuscript (please, refer to L74-L78).to clarify the reasoning behind the rainfall-based event selection. We now explain that the approach aims to capture flood responses across a wider range of antecedent wetness conditions, which is essential for disentangling the respective roles of rainfall peaks and antecedent soil moisture.

Minor comments

Minor#1. Abstract: It would be useful to clarify that by "flood changes" you mean changes in magnitude only.

R: We have now clarified that flood change is related to flood magnitudes (please, refer to L54-L58).

Minor#2. L33 remove "-dominated"

R: We have now removed "-dominated" from this sentence magnitudes (please, refer to L34).

Minor#3. L42 what exactly is meant by "frequency-based…" in the context of a POT approach?

R: Frequency-based POT refers to selecting events by prescribing an average number (frequency) of events per year, rather than defining a threshold based on a fixed quantile of the data series (e.g., the 95th percentile; Wasko et al. 2019, 2021). For instance, assuming an average sampling frequency of 3 events/year, a 30-year time series would yield the 90 largest rainfall peaks. This approach ensures that exceedance percentiles are consistent across catchments.

We have now clarified this point in the revised manuscript (please, refer to L120-124).

REFERENCES

Wasko, C. and Nathan, R.: Influence of changes in rainfall and soil moisture on trends in flooding, J Hydrol (Amst), 575, 432–441, https://doi.org/10.1016/j.jhydrol.2019.05.054, 2019.

Wasko, C., Nathan, R., Stein, L., and O'Shea, D.: Evidence of shorter more extreme rainfalls and increased flood variability under climate change, J Hydrol (Amst), 603, 126994, https://doi.org/10.1016/j.jhydrol.2021.126994, 2021.

Minor#4. L51 point i How do you define "data gap" here? Is this really no missing days whatsoever? Or did you define some threshold?

R: We define a data gap as any missing value in the daily streamflow time series. Therefore, we retained only complete water years, i.e., water years with no missing daily streamflow observations. We have clarified this statement in the revised manuscript (please, refer to L53-L54).

This selection criterion may be considered conservative; however, it is necessary, as gaps in the data can compromise the separation of hydrographs and the identification of flood events.

Minor#5. L52 point ii if "spurious values" were present in the record, did you remove the entire record for that gauge, only the spurious value, the year that the values were in, or something else?

R: We removed the entire record. We have now clarified this statement in the revised manuscript (please, refer to L52)

This selection criterion follows the original quality control procedures of the CAMELS-BR dataset. Specifically, streamflow gauges with issues such as repeated zero values used instead of missing data, abrupt changes due to modifications in measurement instruments or rating curves, annual streamflow exceeding annual precipitation, and unrealistic daily streamflow values (e.g., > 1000 mm d^{-1}) are excluded. Gauges affected by such errors were therefore not included in the dataset.

Minor#6. L76 Remove "the procedure is as follows."

R: We have removed "the procedure is as follows" in the revised manuscript.

Minor#7. L116-117 why three events per year?

R: We select up to three events per year as a trade-off to capture flood occurrence across a broad range of hydrological regimes in Brazil. For instance, in fast-responding catchments with low seasonality (e.g., in the South), floods may occur at any time of the year. In contrast, slow-responding and strong seasonal catchments (e.g., in the North and Amazon regions) exhibit fewer events annually. Consequently, some basins experience multiple flood events per year, whereas others have relatively few (see hydrographs in Figure 2 for illustrative examples).

We have now clarified the choice of the average number of sampled events in the revised manuscript (please, refer to L124-L127).

Minor#8. L120-122 how do these quantiles relate to return periods?

R: This important point was also highlighted by Reviewer #2 (please, see Major Comment 4). Providing an equivalent return period is helpful to make our results more comparable with previous findings. Following Stedinger et al. (1993), the non-exceedance quantile (τ) associated with an annual recurrence interval T is given by:

$$\tau = 1 + \frac{1}{\lambda}\log_e\left(1 - \frac{1}{\mathrm{T}}\right) \quad \text{(Eq. R1)}$$

where λ is the average sampling frequency. Here, we sample our events using a frequency-based rainfall peak-over-threshold approach, ensuring, on average, three events/year/catchment ($\lambda = 3$). Therefore, small events (0.50 quantile) correspond to an equivalent return period of approximately 1.3 years, while large events (0.95 quantile) correspond to an equivalent return period of approximately 7.2 years. These values are consistent with the relatively high sampling frequency of our peak-over-threshold approach, which allows us to characterize both frequent and relatively rare flood responses within each catchment. We have now included the equivalent return periods of our selected quantiles in the revised manuscript (please, refer to L132-L133).

REFERENCES

Stedinger, J. R., Vogel, R. M., and Foufoula-Georgiou, E.: Frequency analysis of extreme events, in: *Handbook of Hydrology*, edited by: Maidment, D. R., McGraw-Hill, New York, 18.1–18.66, 1993.

Minor#9. L145-148 These spearman correlations are extremely weak. I would call this, effectively, no correlation.

R: We agree that the correlations are extremely weak. We have clarified this statement in the revised manuscript (please, refer to L166-168).

Minor#10. L154-155 Rephrase this statement about hydrological consistency with their drivers in order to clarify what you mean.

R: We have now rephrased this statement to make it clearer (please, refer to L173-L176).

Minor#11. Figures 3,5, and 6 using border thickness and size like this does not make the difference between the points clear enough. This is especially difficult in figure 5. I cam not sure how to fix this given the multi-dimensionality of the figures though.

R: We agree that in regions with a high density of stations, the visualization can indeed become less clear, particularly in Figure 5. This is a common challenge when representing multi-dimensional information.

We explored alternative visualization strategies; however, these either reduced the amount of information or made the interpretation less straightforward.

The current representation was therefore chosen as a compromise between readability and information content and is consistent with standard practices commonly adopted in the literature to represent spatial patterns and trends (e.g., Blöschl et al., 2019; Chagas et al., 2022b).

REFERENCES

Blöschl, G., Hall, J., Viglione, A., et al.: Changing climate both increases and decreases European river floods, Nature, 573, 108–111, https://doi.org/10.1038/s41586-019-1495-6, 2019.

Chagas, V. B. P., Chaffe, P. L. B., and Blöschl, G.: Climate and land management accelerate the Brazilian water cycle, Nat Commun, 13, https://doi.org/10.1038/s41467-022-32580-x, 2022b.

Minor#12. Figure 4 The figure indicates “1 standard deviation,” but the mean (μ) is not shown. Since standard deviation is defined relative to the mean, it would be helpful to either indicate the mean in the figure or clarify whether the marked value corresponds to $\mu + \sigma$. Alternatively, showing both $\mu - \sigma$ and $\mu + \sigma$ would make the interpretation clearer. For panel a, consider plotting the hot spots separately to allow this to be more clearly interpreted.

R: We have now included the mean value in the scatterplot.

Regarding the suggestion to show both $\mu - \sigma$ and $\mu + \sigma$, our intention is to focus on the right tail of the flood change distribution (i.e., large flood increases). Therefore, we chose to display only $\mu + \sigma$ threshold to avoid potential misinterpretation.

Regarding plotting each hotspot separately, this would require substantial modifications to the original figure structure and compromise its overall consistency. Therefore, to improve visualization while maintaining the original design, we refined the color scheme to enhance clarity and readability.

Minor#13. Figure 5 consider using the same scale for both elasticities

R: In our initial attempt, we used the same color scale for both rainfall peaks and antecedent wetness. However, because these variables have substantially different value ranges, the shared scale hides the spatial patterns of antecedent wetness elasticities.

Minor#14. Figure 7 Don't use the same colors for the boxplots as for the hotspots. Currently, AM and S seem to match the boxplot colors.

R: We revised the color scheme to improve visual clarity. To avoid misinterpretation, we have now ensured that the colors used for peak rainfall and antecedent wetness in panels are not reused for representing the hotpots (please, refer to Fig. 1, 4, and 7).

Minor#15. L289 Discuss in more detail how water "storage capacities" were defined in a previous study. This could just be one sentence.

R: We defined water storage capacity based on the root zone water storage capacity, that is, the maximum volume of hydrologically active soil water available for plant transpiration (e.g., Chagas et al., 2022a; Wang-Erlandsson et al., 2016). We have now clarified this statement in the revised manuscript (please, refer to L321-L322).

REFERENCES

Wang-Erlandsson, L., Bastiaanssen, W. G. M., Gao, H., Jägermeyr, J., Senay, G. B., van Dijk, A. I. J. M., et al. (2016). Global root zone storage capacity from satellite-based evaporation. Hydrology and Earth System Sciences, 20(4), 1459–1481. https://doi.org/10.5194/hess-20-1459-2016

Chagas, V. B. P., Chaffe, P. L. B., and Blöschl, G.: Process Controls on Flood Seasonality in Brazil, Geophys Res Lett, 49, https://doi.org/10.1029/2021GL096754, 2022a

Minor#16. L253-255 Why do these results highlight the importance of using a broader range of events? Please elaborate on this thought.

R: We intended to highlight that relying exclusively on annual maximum streamflow series may hide patterns occurring in more frequent events. Analyzing a broader set of events allows the inclusion of smaller floods that may respond differently to rainfall peaks and antecedent wetness conditions. To clarify this reasoning, we have expanded the text in the manuscript (please, refer to L280-L286).

Minor#17. L255 “timidly” is too casual of a word

R: We have replaced the expression “timidly explored” with “Changes in large floods have been relatively underexplored in Brazil so far” in the revised manuscript to improve clarity (please, refer to L276-L277).

Minor#18. L259-260 specify that you mean changes in flood magnitude

R: We have now clarified this sentence in the revised manuscript (please, refer to L287).