Reply to comments by Reviewer #1

We thank Reviewer #1 for the helpful comments and the positive view regarding our manuscript.

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

General comments

This manuscript examines changes in small and large floods in Brazil as well as explores potential causes of these changes. The manuscript finds that large floods are increasing more than small floods and that climate drivers are not able to fully explain the reasons for these changes – this is an important and interesting finding with implications for management of hazards and water supply. For these reasons, the manuscript is suitable for HESS.

I have a number of comments to improve support for the findings and one methodological question regarding the use of multiple linear regression to quantify the elasticity of different explanatory variables to changes in small and large floods. Overall, these comments tend towards Minor Revision but a few could be considered Major comments, in that new analysis may be needed rather than edits to the existing text.

I will note that the text of the flood event detection in Section 2.2 is thorough and thoughtful in the application of trend detection methods. In particular, it is worth noting the excellent attention to details regarding the assurance that methods to isolate large and small floods are robust to catchments of varying size. This is a careful and important consideration that is often overlooked.

R: We thank Reviewer #1 for the detailed read of our manuscript and constructive comments.

We address each point in detail below.

Major comments

Major#1. The one major comment that I have is with respect to the use of coefficients resulting from a linear regression model as elasticities. This approach is used to assess the sensitivity of changes in large and small floods to changes in rainfall and antecedent wetness, forming the foundation of these conclusions. Linear regression finds the optimal model coefficients that minimize the sum of square errors between a dependent variable and its predictor variables. The estimated coefficients are, therefore, subject to factors like omitted variable bias and multicollinearity thus making the coefficients a less desirable approach to understand the sensitivity of a dependent variable to changes in a predictor variable. The coefficients are appropriate for prediction but not very useful for physical interpretation. There are other regression approaches, namely panel regression, that can formally treat regression coefficients as elasticities. There is also a large literature on formal causal attribution approaches.

For this manuscript, I wondered why these approaches were not applied and how the results and interpretation might differ if these more formal methods were applied to assess causes for changes in large and small floods.

R: We agree that linear regression coefficients, whether estimated via ordinary least squares (OLS) or quantile regression, are optimal model parameters and may be affected by issues such as omitted variables and multicollinearity. In our case, coefficients are estimated using a quantile regression framework that minimizes asymmetrically weighted absolute deviations rather than the sum of squared errors, as in OLS.

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 L152-L158).

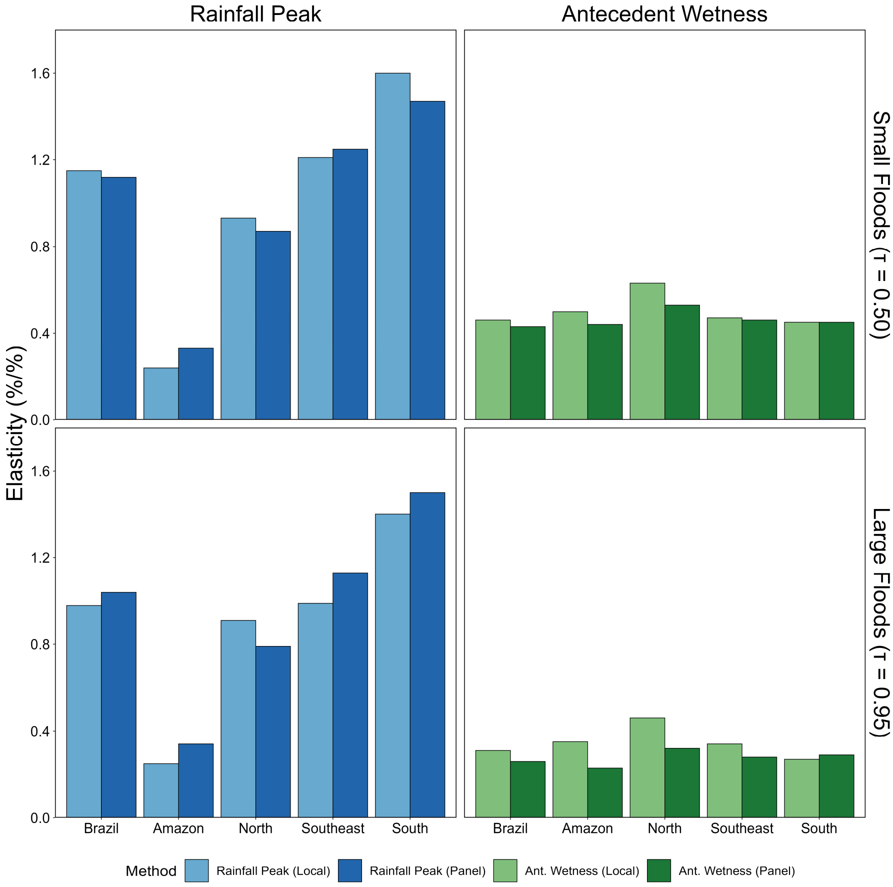

Figure R1: Comparison between elasticities estimation between local and panel quantile regression for small (0.50 quantile) and large floods (0.95 quantile).

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.

Major#2. In the current version of the manuscript, I would request that the introduction be expanded to discuss these approaches and provide justification, if possible, for why the application of linear regression to assess elasticity is appropriate. The introduction at present does not provide any justification and elasticity is only briefly mentioned in L40. The introduction should be expanded to include (1) a discussion of the use of causal attribution for the attribution of flood changes, (2) a discussion of panel regression and why it was not applied here, and (3) justification for the use of elasticities via linear regression to attribute reasons for flood changes.

R**:** We agree that the Introduction benefit from a clearer conceptual grounding in attribution and elasticity. To address this while maintaining the overall structure and conciseness of the Introduction, we have added a brief discussion of our attribution framework in the final paragraph of the Introduction of the revised manuscript (please, refer to L41-L46). Following the reviewer's suggestion, we provide a more comprehensive justification for our methodological choices, specifically regarding the use of multivariate quantile regression over panel regression and of elasticities in the Materials and Methods section of the revised manuscript (please, refer to L152-L158).

Major#3. Later in the text, section 2.3 explains that "the log-log regression allows us to interpret the model coefficients as elasticities" but I do not think this is correct given the discussion above. If the authors continue to use the linear regression approach, a much more detailed justification is needed as to how omitted variable bias and other limitations are addressed. The authors do note that they consider multicollinearity but not omitted variable bias which could also affect the coefficient estimates.

R: Regarding the definition of elasticities, assuming a log-log specification in the linear model, the regression coefficients mathematically correspond to elasticities. Therefore, the estimated coefficients can be interpreted as statistical elasticities describing the conditional sensitivity of floods to rainfall peak and antecedent wetness, rather than as estimates of a true causal effect. We have now explicitly clarified in Section 2.3 that our coefficients represent statistical elasticities (the conditional sensitivity of flood peaks) rather than absolute causal effects (please, refer to L147-L152).

The reviewer correctly points out that omitted variable bias was not directly considered in our modeling approach. Instead, we addressed this issue through the study design by focusing on the primary first-order controls on flood generation in Brazil: rainfall peaks (meteorological forcing) and catchment state variables (antecedent wetness), as previously investigated by Chagas et al. (2022a). Additionally, to reduce the influence of other potential confounding factors, such as human activities, we restricted our analysis to catchments with low levels of flow regulation (<30%). These catchments also typically have a very low fraction of impervious surfaces such as urban land cover; where only 0.2% of the catchments have more than 5% of their area classified as impervious (Chagas et al., 2020). Previous assessment on flood change attribution in Brazil suggest that other human influences, such as deforestation and water use, have played a minor role in the observed flood changes (e.g., Chagas et al., 2022b). Furthermore, the high consistency between our local estimates and the Fixed-Effects Panel model, which controls time-invariant omitted variables, provides additional confidence that our results are not biased.

REFERENCES

Chagas, V. B. P., L. B. Chaffe, P., Addor, N., M. Fan, F., S. Fleischmann, A., C. D. Paiva, R., and Siqueira, V. A.: CAMELS-BR: Hydrometeorological time series and landscape attributes for 897 catchments in Brazil, Earth Syst Sci Data, 12, 2075–2096, https://doi.org/10.5194/essd-12-2075-2020, 2020.

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

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

Major#4. A potential implication of the use of linear regression comes in the form of some counter-intuitive results. For example, L216 states that "there is a noticeable shift in the sign of rainfall changes (from negative to positive) between large and small events." This seems counter intuitive that the contribution of rainfall would not always be positive in terms of its effects on flood changes. A negative sign indicates that more rainfall would lead to less flooding, which does not make sense (assuming I am reading this statement correctly). Could this result be an artefact of the use of linear regression-estimated coefficients for the elasticities?

R: We would like to clarify that the elasticity (i.e., the sensitivity of flood peaks to rainfall) is positive for both small and large events in our model. Therefore, flood magnitudes scale with their drivers; for example, on average, higher rainfall peaks are associated with higher flood peaks.

The “shift in sign” mentioned in line 216 (now Line 236) refers specifically to temporal changes in rainfall peaks, not to the elasticity. In other words, the estimated elasticities remain positive. However, rainfall peak trends differ across quantiles: smaller rainfall peaks show a slight decrease over time, whereas larger rainfall peaks tend to increase.

As a result, when computing the contribution of rainfall to flood changes (i.e., elasticity × rainfall peak trend), the contribution can be negative for smaller events and positive for larger events. This does not imply that more rainfall reduces floods; rather, it reflects that smaller rainfall events are becoming slightly less intense over time, while extreme rainfall events are intensifying. We have revised the text to clarify this distinction and avoid potential misinterpretation (please, refer to L235-L240)

Major#5. Another potential indication that this may not be the appropriate approach is that you needed to apply a constraint on the parameters because you are getting negative elasticity values (L152).

R: The parameter constraints were introduced to ensure a hydrologically reasonable relationship between floods and their drivers and to ensure a straightforward interpretation of the results. The contribution of each driver is estimated by multiplying its elasticity (Eq. 3) by the change in the driver (Eq. 1). In this context, negative elasticities could lead to counterintuitive interpretations (e.g., suggesting that increasing rainfall reduces floods).

However, such cases were rare in our set of catchments, occurring in only 0.9% and 5.8% of the catchments for small and large floods, respectively. This indicates that, in most cases, the estimated relationships between floods and their drivers are hydrologically consistent.

Minor comments

Minor#1. The period of analysis ends in 2018, which is now almost 8 years ago. Is it possible to include the additional 7-8 years of data beyond the measurements included in the CAMELS-BR dataset? For example, can you use the stream gauges in the CAMELS-BR gauge list but obtain the additional years of data?

R: We fully agree that extending the analysis period and including more recent years of data would be beneficial and could potentially strengthen the robustness of the results.

Although more recent streamflow data are available for some gauges, data availability provided by the National Water and Sanitation Agency (ANA) is not uniform across Brazil. A complete and consistent extension of the time series up to 2025 is not available for all catchments included in the CAMELS-BR dataset. Extending the analysis period would therefore result in substantial variability in the final year of the series regions, reducing the comparability of results in terms of the analyzed period. To ensure temporal consistency and comparability across all catchments, we therefore retained the original CAMELS-BR period.

Minor#2. L78: What is the justification for using the multiplier of 1.11? Was a sensitivity analysis conducted here to assess this choice or was that part of your visual inspection and other checks on the approach? It might be helpful to clarify this.

R: This threshold of 1.11 follows the standard implementation of the nonparametric baseflow separation method based on the identification of turning points in the daily streamflow series, proposed by the Institute of Hydrology (1980). In this method, a point is identified as a turning point when it exceeds neighboring values by a factor of 1.11, which has been widely adopted in subsequent applications. Based on our experience implementing the method, the most influential parameter is the window size used to search for turning points. Therefore, this parameter was adjusted for each catchment based on the baseflow index.

REFERENCES

Institute of Hydrology: Low flow studies, Report No. 1, Wallingford, UK, 1980.

Minor#3. L75-85: Was a software package used to perform this analysis or did you use your own original code to make these calculations? If you did use an existing software package, please cite.

R: All analyses were performed with the R packages quantreg (Koenker, 2009), lmtest (Zeileis, 2002), sandwich (Zeileis et al., 2006), and ggplot2 (Wickham, 2016). We have now clearly stated this information in the manuscript. Accordingly, we have added a Data Availability Statement specifying the software environment and listing the R packages used in the analysis.

REFERENCES

Koenker, R. (2009). quantreg: Quantile regression. The R Project for Statistical Computing.

Zeileis, A., Köll, S., & Graham, N. (2020). Various versatile variances: An object-oriented implementation of clustered covariances in R. Journal of Statistical Software, 95, 1–36. https://doi.org/10.18637/jss.v095.i01.

Zeileis, A., Köll, S. & Graham, N. Various versatile variances: An object-oriented implementation of clustered covariances in R. J. Stat. Softw. 95, 1–36 (2020).

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis. Springer-Verlag Institute of Hydrology (1980). Low flow studies, Report No. 1, Wallingford, UK.

Minor#4. Figure 2: It is very difficult to distinguish between the grey and black lines. Perhaps consider another way to present this information.

R: We have now improved Fig. 2 in the revised manuscript by increasing the line width to make the lines easier to distinguish.

Minor#5. L142-143, the comment is made: "Assuming that the contribution of the drivers is additive…" I wondered if this is correctly stated because taking the logs of the data and using the log values in the linear regression would imply that the contribution of the drivers is multiplicative.

R: That is correct, the original statement that the contribution of the drivers is additive was inaccurate. Since the analysis is performed in log space, the contributions of the drivers are indeed multiplicative, not additive. We have corrected this statement in the revised manuscript to avoid confusion and to ensure consistency with the underlying model formulation (please, refer to L162).

Minor#6. L145: The text says "we check for these correlations [between the drivers]" however, I think this more precise to say that we "check the variance inflation factors for correlations between the drivers." This is an important point because variance inflation can lead to problematic coefficients and therefore unreliable elasticity estimates.

R: We agree that variance inflation can lead to unstable coefficient estimates in linear regression models.

In the context of quantile regression, multicollinearity is commonly assessed by examining the dependence structure among the explanatory variables themselves, as previously reported in the manuscript through pairwise correlations between rainfall peak and antecedent wetness.

Nevertheless, to address this issue in a more formal way, we performed an exploratory OLS regression using the same set of explanatory variables and computed the corresponding VIFs. We emphasize that this analysis is intended solely as a diagnostic assessment of potential multicollinearity among the predictors, and that VIFs derived from OLS are not directly transferable to or interpretable as variance inflation factors in quantile regression. The results indicate that multicollinearity among the drivers is low (VFI < 1.1 in most catchments; please see Fig. R2) and, therefore, unlikely to lead to unstable coefficient estimates in the quantile regression model.

Therefore, to maintain methodological consistency and avoid misinterpretation, we retained the collinearity assessment based on Spearman correlations in the revised manuscript.

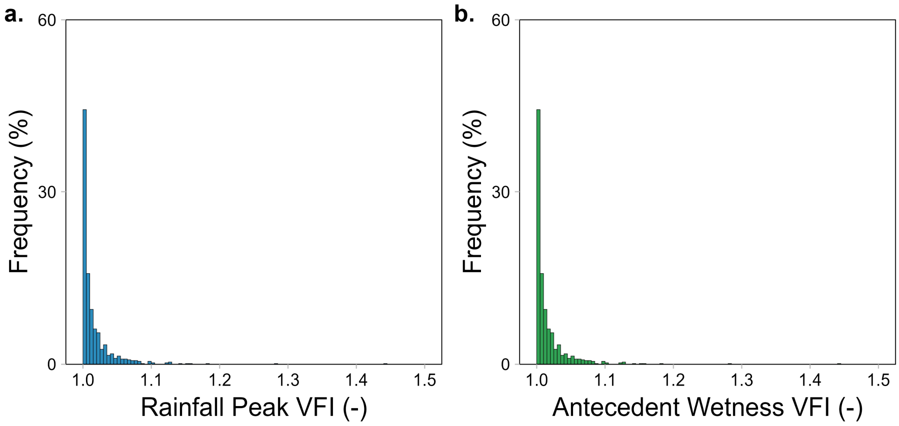

Figure R2. Frequency distribution of OLS-based VIF estimates for (a) rainfall peak, and (b) antecedent wetness. Note that all VFI values are in the range 1-1.1, suggesting that multicollinearity is not an issue in our model.

Minor#7. L171: Please include the percentage of small floods in parenthesis exactly as you showed for the proportion of large floods. You cannot assess the accuracy of this statement without the other percentage shown.

R: We have now included this in the revised manuscript (please, refer to L191).

Minor#8. L173-174 are confusing. The sentence reads that large events are increasing more and decreasing less; however in L172, it appears that large events are decreasing more and increasing less (the range shown is -76.6 to +59.2). Can you clarify this?

R: We have now clarified these points in the manuscript. Our intention in those lines was not to compare the full range of trends between large and small events, but rather the occurrence of substantial increases. Specifically, we define substantial flood increases as trends exceeding 13.5% per decade (one standard deviation of the flood-change distribution). Within this subset of substantial increases, 78% occur in large flood events, indicating that strong positive changes are more frequently associated with large floods.

The ranges reported in the previous sentence describe the overall distribution of trends, which includes both increases and decreases, and therefore should not be interpreted as evidence that large floods necessarily increase more in all cases. Instead, the key point is that when substantial increases occur, they are predominantly associated with large events (please, refer to L193-196).

Minor#9. L176-180 refers to "hotspots"; however, it is unclear what geographic region corresponds to these hotspots because there is no reference to a figure to support these statements. Please add supporting evidence.

R: We now explicitly refer to the corresponding figure, which allows the reader to clearly identify the geographic regions corresponding to the identified hotspots (please, refer to L197-198).

Minor#10. L200: The general comment is made that "The elasticity of floods to antecedent wetness is noticeably lower than that to rainfall peaks…" Clarification is needed as to whether you intend to say all floods (small and large floods). Later in the sentence you differentiate between large and small floods. I would change to read: "The elasticity of both large and small floods to antecedent wetness is noticeably lower than that to rainfall peaks…"

R: In the sentence "The elasticity of floods to antecedent wetness is noticeably lower than that to rainfall peaks…", we intended to refer to both small and large floods. Following the reviewer's suggestion, we have revised the sentence to: "The elasticity of both large and small floods to antecedent wetness is noticeably lower than that to rainfall peaks…" (please, refer to L221).

Minor#11. Please provide evidence to support your statements in L228-230.

R: This statement was based on the interpretation of the distribution of the relative contribution values shown in Fig. 7a. In the South hotspot, the boxplots representing the contributions of rainfall peak and antecedent wetness show similar median values and substantial overlap between their interquartile ranges. This overlap indicates that the relative importance of the two drivers varies considerably among catchments, with neither rainfall peak nor antecedent wetness consistently dominating across the region.

We have clarified this interpretation in the revised manuscript (please, refer to L249-L252).

Minor#12. Figure 7 is not standalone and the colors are not ideal. The hotspot abbreviations should be spelled out in the caption or on the figures. It is also confusing to use blue and green in the boxplots to differentiate between wetness and rainfall peaks and then re-use the same two colors to designate hotspots in panel c, which have no relation to the interpretation of those colors used in panels a and b.

R: We have improved the description of the figure to clarify the hotspot abbreviations in the caption. Additionally, we revised the color scheme to improve visual clarity. To avoid misinterpretation, we ensured that the colors used to represent rainfall peak and antecedent wetness in panels (a) and (b) are not reused for the hotspots in panel (c).

Minor#13. In Section 4.2, consider also citing Collins et al. (2022), who also show large floods in the US have not changed over the past 50 years.

Collins, M. J., Hodgkins, G. A., Archfield, S. A., & Hirsch, R. M. (2022). The occurrence of large floods in the United States in the modern hydroclimate regime: Seasonality, trends, and large-scale climate associations. *Water Resources Research*, 58, e2021WR030480. https://doi.org/10.1029/2021WR030480

R: We thank the reviewer for suggesting the inclusion of this relevant paper. Collins et al. (2022) have now been incorporated into the Discussion section of the revised manuscript (please, refer to L295).

Editorial comments

Editorial#1. L76: Consider rewording as: "To accomplish this, local minima are identified…"

R: We have corrected this statement in the revised manuscript (please, refer to L80).

Editorial#2. L255: Consider changing "timidly explored" to "explored only in a limited way in Brazil"

R: We agree that timidly explored is not appropriate. We have replaced it with "relatively underexplored" in the revised manuscript to improve clarity (please, refer to L276-L277).

Editorial#3. L259-270: Nice discussion section overall; however L265 contains an incomplete sentence. The text reads: "In Australia (Wasko & Nathan, 2019), and globally (Wasko et al., 2021)."

We thank the reviewer for the positive view about our discussion section. We have now corrected the incomplete sentence in the revised manuscript (please, refer to L296-L297).

Editorial#4. L284: Change to read: "...and driving mechanisms for these changes in Brazil."

R: We change this sentence to "...and driving mechanisms for these changes in Brazil" (please, refer to L315).

Editorial#5. L285 and L284: Change the word "sign" to "direction". Consider this change across all mentions of "sign" in the text to improve clarity.

R: Following the reviewer's suggestion, we have replaced the term "sign" with "direction" throughout the revised manuscript to improve clarity.