

HESS-2024-188 Review

This manuscript titled Integrating historical archives and geospatial data to revise flood estimation equations for Philippine rivers is a revision of Flood estimation for ungauged catchments in the Philippines manuscript that has been submitted to Hydrology and Earth System Sciences.

The authors have addressed the comments raised by the reviewers in the previous round of review.

Therefore, it was time consuming to compare the previous and current versions of the manuscript.

The revised manuscript presents an ambitious effort to combine short-duration historical records from diverse sources into a single, coherent analysis for deriving flood design equations at a national scale.

I suggest the authors consider the following points in their revision:

Major Comments

Although the authors mentioned uncertainty many times in the manuscript, they did not provide any quantification of the uncertainty in their results. This is a major issue that needs to be addressed.

Also, in general, flood frequency analysis is not a proper method to estimate flood magnitude when you have limited data. Fitting a curve to 7–10 data points is not a reliable method to estimate flood magnitude.

Abstract

While the abstract effectively conveys the general research objective and findings, in my opinion it may need some revisions to improve clarity and precision. When you start to mention R^2 and then express the added value of including the new variables, the sentence is not clear (L25-27). I suggest you revise it. It would benefit from a clearer statement regarding the limitations of the low R^2 values and the implications for design uncertainty.

Introduction

The introduction provides an overview of the study. It briefly describes the importance of catchment area and mean annual rainfall as predictors of flood magnitude. The authors tried to highlight the impact of pooling data from available sources to improve flood estimation when the data are limited in time and space.

I suggest merging the two middle paragraphs of the introduction to make it more concise and clear.

Also, the hypothesis and research questions of the study are not clearly stated in the introduction. It would be better to state them explicitly.

Methodology

- In the section Data sources, the authors provide a detailed description of the data sources and the process of data collection. Different sources introduce distinct uncertainties and biases into the analysis. For example, in Figure 1, some sources (red and blue dots) are more concentrated in regions such as the north of the country,

while in the west and south—where there is lower rainfall and lower contribution from tropical cyclones—we have no or only one source of data. This may introduce bias in the analysis. The authors should discuss this issue in the manuscript.

- More details regarding the screening criteria for data quality and the rationale behind the selected catchment properties would improve transparency. For example, three sources of data are used in the analysis; while they were recorded differently, in different periods of time, and likely with different measurement techniques, the method of merging these data should be discussed in the manuscript. It is highly likely that the quality of measurements before the 1980s is lower.

Analysis Methods

- I am curious to know whether you ever tried to employ two peaks per year or any POT analysis to identify the peaks in the data, instead of only using the annual maximums. This approach would give you more freedom not only to select the highest peak in the year but also the second highest independent peak in the year. This could help you better understand the flood frequency in the region, as the second peak may occur in another season and allow you to better capture your basin's behavior. Then, you could continue to determine Q_{med} of the new series of peaks.
- The manuscript provides a thorough description of the curve fitting using L-moments and the subsequent regression analyses. Yet, the discussion on the potential biases arising from combining data of varying quality and the choice of best-fit distributions (with respect to low R^2 values) deserves further elaboration. Moreover, since the study aims to estimate extreme floods, linear regression may not be the best approach. The authors should consider using a more robust method, such as quantile regression, to account for the non-linear relationship between the predictors and the response variable.
- What is the set threshold of low CvM p-values used to exclude data from the analysis in L183?

Results

- The correlation approach in Table 4 does not lead to a new conclusion. The fact that a larger catchment area leads to a higher correlation is not a new finding. It is the same with the DPLBAR variable, the length of the streamflow network, and the mean annual rainfall. Therefore, your addition in Table 5 should be highlighted. I suggest restructuring the results section to emphasize the new findings of the study.
- Perhaps testing and illustrating your approach on only the new dataset as a test case would be a good idea to show the robustness of your approach. This will also help in understanding the uncertainty in the results.
- What would be the expected best R^2 value by adding the new variables? It would be better to have a benchmark to compare the results. What is the ideal R^2 value for flood frequency analysis in the region? Is the benchmark 0.92 in Papua New Guinea? You could randomly generate some synthetic data and try to estimate the flood frequency analysis to see the ideal R^2 value.
- As you mentioned, land use change is a major factor in flood frequency analysis, and you employed almost current land use data in the analysis. This is a significant challenge and limitation of the study.
- The abbreviations in this study are not mathematically scientific, such as AREA or RMED. It would be better to use the full names of the variables in the text and use better letters for the variables. For example, A for area, and R_m for RMED, and so on.

- Since the results are mainly presented on Q10 and they are not significantly appropriate for flood control and design, it would be better to include a discussion on the results and the limitations of the study.

Discussion

- It would be valuable to discuss the limitations (e.g., stationarity assumptions, data quality issues, and land use change) more explicitly and to outline potential paths for future improvement, such as incorporating non-stationary models or enhancing continuous monitoring.
- Tropical cyclones were not part of your investigation; however, they play a role in the discussion.
- Climate change and spatiotemporal variability in the region are not discussed in the manuscript at all, despite the merged data varying over time.
- The discussion section is generally long. I suggest revising it to be brief, more concise, and clear. However, the current form is good for readers to understand the results and limitations of the study.
- The comparison with HEC-HMS modeling lends additional credibility, though the discussion might be expanded to explain the practical implications of the observed discrepancies between instantaneous peak flows and daily mean flow estimates.

Conclusion

- The conclusion is well-structured and effectively summarizes the key findings of the study.

Minor Comments

Abstract:

- L18: Split the long sentence `However, the global ...` into two sentences for clarity. The current sentence contains four commas.
- L23: What does `national and regional scales` mean? Are they two different scales?
- L25: The term `GIS-derived` is not needed here. You can simply say `geospatial catchment characteristics`.
- L30: There is a redundancy with the term `predictive equation` in the same sentence.

Introduction:

- L43: The sentence `The resulting equations ...` is not well connected to the previous sentence. The starting lines are quite good, but there is a gap between the first and second parts of the first paragraph.
- L79: A reference to Figure S1 is needed.
- Please provide a map of the available length of time series in the Philippines. This will help in understanding the data availability in the country (L80). Although the time period is indicated in Table 1, it is not clear whether the records are continuous or if there are gaps in the data. Alternatively, you can provide some sentences in the text to explain this issue.
- How do you define short time series? Is it less than 35 years? (L80). It would be better to provide a definition for short time series or a reference for the definition.

- L84: The FEH abbreviation has already been defined previously in L50.

Data sources:

- Figure 1: In the caption, it is mentioned that the four climate types that have been identified for the Philippines (Coronas, 1920). Since the climate types were identified in the 1920s, is there any more recent climate type identification for this region? Given global warming and climate change, the climate types may have changed or been better defined in recent years.
- Figure 1: Please replot panels b and c and use discrete colors instead of gradient colors. Also, Figure 1C does not support any of your results except for a sentence in the conclusion. It would be better to remove it from the manuscript or integrate its insight into your interpretation.
- I suggest moving Table 1 to the supplementary material, as it is not necessary in the main text.

Analysis Methods:

- To achieve more consistency in the manuscript text, I suggest adding Q5 and Q50 in Figure 2, and so on, in your text.
- Since Table 2 does not show any relation between the size of the catchment, climate type, and the best-fit curve, is there any geographical pattern in the best-fit curve? For example, do catchments in the north of the country have the same best-fit curve? What if you plot the best-fit curve on the map of the country? Usually, subcatchments in the same basin may have the same best-fit curve since they are flow-connected.
- L204: The phrase (Figure S1) show this pattern is unclear. I have not seen this pattern in Figure S1. Please revise the text. The mentioned figure is Administrative regions of the Philippines. Since the numbering starts from north to south, it would be better to reorder the legend of the figure to follow the same order instead of alphabetical order.
- L208: The reference to Figure S2 is incorrect. It is currently written as (Figures 2, S1); it should be (Figure S2), as seen in the supplementary material. Also, the figure itself is not well plotted.
- The quality of Figure S3 is too low. It is not readable. Please revise it. The current figure overlays the main curves on top of each other. The area of concentration should be zoomed in to see the differences between the curves on the right part of the x-axis.
- The same applies to Figures S4 and S5. However, they are slightly better. I think these figures can be highlighted for regionalization since they are important for understanding how the curves differ by region, climate type, and catchment size. The current format does not help in understanding the differences between the regions.
- L247: As far as I know, we have free global DEMs with 30 m resolution. So why did you resample the DEM?

Results:

- L287: The phrase This contrasts with Meigh's should be moved to the conclusion.
- Table 4: Instead of NA, write – in the table.
- Theoretically, your Q_med should be equivalent to your Q2 when you have a limited length of time series. If you look at Table 4, the columns Q_med and Q2 are almost identical. Also, the correlation is sensitive to the number of data points.
- Table 5: The alignment of the table is not correct. Please revise it, and make it more readable.
- Figure 5 and subsequent figures: Please elaborate on "normally distributed" in the figure caption, especially for subplots b, c, e, and f.
- Set a fixed significance level for the p-values in the text. In section 5.3.3 it is 0.05, while previously it was 0.01.

- Figure S7 must be revised. The current figure is not readable enough; it is a bit small, and the selected colors do not help the readers understand it. Also, since this figure has three parts, the main body of the manuscript does not support it well.

Discussion:

- Figure 8: It seems that the x-axis of panel b is not correct. Please revise it.

This study contributes to hydrological modeling by demonstrating how pooling individually short historical flood records—combined with high-resolution geospatial data—can produce nationally applicable flood estimation equations even in data-sparse tropical regions. The `Recommended design equations` section is a part that the authors may consider including in their analysis.