Anonymous Referee #2, 08 Jun 2023

The study uses a modeling approach aiming to assess the flood risk at the sub-basin scale in India and the impact of reservoirs on the flood risk. My recommendation is rejection due to several significant issues in the methodology that seem inadequate for addressing the posed questions. Further justification is also necessary for the conclusions made. Here are my main concerns:

The understanding of flood risk seems to focus too heavily on the worst flood event in history. To understand flood risk, it requires examination of a large number of flood events over a range of conditions and incorporating uncertainties.

Thanks. We appreciate insightful comments and suggestions that have been addressed in the revised manuscript. We have included more observed flood events of varying intensities in the revised manuscript. By incorporating multiple flood events, our revised analysis captures the variability in flood characteristics and associated uncertainty. In addition, our revised analysis examines the flood risk at the sub-basin scale in a robust manner and provides a comprehensive understanding of the potential impacts. We would like to highlight that there is a lack of observations at an appropriate spatial and temporal resolution that can be used for flood risk assessment in India at the sub-basin scale. This is probably the first study that attempts to reconstruct the flood risk at the sub-basin scale in India based on the last 120 years of data.

I question the suitability of a large-scale model like H08-CaMaFlood for flood risk assessment, which typically requires higher-resolution models that can accurately capture local topography and features. Given the shown substantial bias in simulated flood occurrences, I am unconvinced of the model's efficacy in predicting flood water depth at the event scale. The downscaling approach, which scales simulated flood depth from a 0.1 degree to a 200 m resolution within CaMa-Flood, compounds this uncertainty. Ultimately, the rationale behind the choice of CaMa-Flood for localized flood risk assessment is unclear to me, as its resolution seems too coarse for the purpose.

Thanks for the comment and critical insights. We are aware that flood risk at the local scale requires a high-resolution (sub-meter) scale of hydrological and hydraulic modeling. However, the aim of the current work is not to provide a flood-risk assessment for the entire country based on high-resolution modeling and data. This would be beyond the scope of the work to reconstruct flood maps at high resolution using the database for more than 120 years. In the revised manuscript, we have clearly mentioned the scope of the current work and its limitations.

We appreciate the comments and suggestions on the use of the H08-CaMa-Flood model for flood risk assessment. While it is true that flood risk assessment requires higher-resolution models to accurately capture local topography and features, it is also important to consider the specific context and objectives of the assessment.

The suitability of a large-scale model like H08-CaMa-Flood for flood risk assessment can depend on several factors. One such factor is the scale of the study area. In our case, the assessment is focused on a broader region where simulating high-resolution inundation dynamics for the entire area may not be feasible due to computational limitations and high computation time. In such situations, a large-scale model like H08-CaMa-Flood can provide a valuable overview and identify areas of higher flood risk that can be further investigated using more detailed methods, if necessary.

While local-scale models may be preferred for predicting flood water depth at the event scale (Bates et al., 2010), the aim of our study is not to precisely estimate water depths but rather to identify areas at risk and assist in prioritizing resources for mitigation efforts. In such cases, a large-scale model can still provide valuable insights by indicating areas that are more susceptible to flooding, allowing decision-makers to allocate resources and develop appropriate strategies. Therefore, we believe that our study will provide crucial insights into large-scale drivers and patterns of flood risk to develop more informed adaptation and mitigation measures. We have revised the manuscript to clarify these issues and highlighted the novel contribution of our study, along with potential limitations. While H08-CaMa-Flood exhibited some bias, its performance against other available models cannot be ignored (Hirabayashi et al., 2013; Yamazaki et al., 2011). Therefore, our modeling framework for the sub-basin scale flood risk assessment in India can provide important insights that can help in flood mitigation.

For the downscaling approach within the CaMa-Flood model, it is true that there are uncertainties associated with the process of scaling simulated flood depth from a larger grid resolution to a finer resolution (Yamazaki et al., 2017). However, the downscaling approach is often utilized to make the outputs of large-scale models more applicable at local scales. We have compared the flood risk maps at the original model resolution against the downscaled maps/observational datasets and discussed the limitation in the revised manuscript.

The authors' claim of an acceptable model skill is unconvincing to me. For river flooding, they set a NSE threshold of 0.5, which is questionable since a score of 0.6 is generally considered the minimum for model adequacy. Even then, some stations fail to meet this lowered threshold. There is also a lack of flood-relevant metrics, such as bias in peak discharge of flood events.

While different thresholds for NSE (Nash-Sutcliffe Efficiency) for satisfactory model performance are available in the published literature, it's important to note that there is no universally agreed-upon threshold for model adequacy in streamflow prediction. Regarding the note that an NSE threshold of 0.5 for river flooding is questionable, although it is arguable, that this threshold is commonly used and accepted in the field, especially for daily streamflow prediction (Dakhlalla & Parajuli, 2019; Leta et al., 2018). However, it's true that higher NSE values are generally desired for more accurate predictions.

The mention of some stations failing to meet even the lowered threshold is indeed a valid concern. We have improved the model calibration for these stations, re-evaluated the model parameters and input data, and conducted a comprehensive analysis to improve the model's performance in the revised manuscript. We thank the reviewer for highlighting the need for flood-relevant metrics, specifically the bias in peak discharge of flood events. In the revised manuscript, we have included an additional plot showing the bias and timing error in peak discharge for each river basin.

I would suggest evaluating the worst flood event selected as well. Concerning flood inundation modeling, it would be beneficial if flood extent data were used to evaluate the model's skill. With respect to flood occurrences, I noted previously that the bias seems significant even before the application of downscaling. Despite these evident issues, no discussions on the uncertainties present in this study are included. This omission casts further doubt on the reliability of the results and necessitates a comprehensive review of the methodology.

Thanks. In terms of flood inundation modeling, incorporating flood extent data for evaluating the model's skill is an excellent suggestion. Flood extent data can offer a more direct measure of the model's performance in simulating flood dynamics and spatial patterns. Including such an evaluation would strengthen the study's methodology and provide a clearer understanding of the model's capabilities. We have re-examined the model's ability in simulating the flood extent as well for some selected flood events. In addition, we have discussed the causes of the bias and highlighted potential sources (input data, model parameterization) of uncertainty. Further, the omission of discussions on uncertainties is indeed a noteworthy point. We have addressed the suggestions related to limitations and uncertainty in flood risk assessment in the revised manuscript and provided a separate section in the discussion of the manuscript.

The use of the C-ratio to assess the role of reservoir operations in flood risk is confusing. The C-ratio, defined as the ratio of a reservoir's total maximum storage capacity to the mean annual discharge at the sub-basin outlet, is essentially a constant that doesn't account for variability in reservoir outflow resulting from operations serving different objectives. The mean annual discharge also seems irrelevant when examining a record flood event at a much shorter timescale. Consequently, I find the results based on C-ratio to be lacking in significance. Certain fundamental details that could aid in interpreting the results are missing, such as a clear definition of how a flood event is defined.

Thanks. We appreciate the suggestion and acknowledge that the C-ratio does not capture the full complexity of reservoir operations and their impact on flood risk. However, we would like to highlight that incorporating the reservoir operations related complexities is challenging due to lack of observational datasets related to reservoir operations. Therefore, our aim was to provide an overview of the sub-basins that have high flood risks and affected substantially by the reservoir operations. While the C-ratio may have limitations in assessing the influence of reservoir operations during extreme flood events at shorter timescales, it can still provide a useful measure of the potential storage capacity available in a reservoir relative to the average discharge over a longer term. We have provided more details on the utility of C-ratio in the revised manuscript.

References

Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. Journal of Hydrology, 387(1–2), 33–45. https://doi.org/10.1016/J.JHYDROL.2010.03.027

Dakhlalla, A. O., & Parajuli, P. B. (2019). Assessing model parameters sensitivity and uncertainty of streamflow, sediment, and nutrient transport using SWAT. Information Processing in Agriculture, 6(1), 61–72. https://doi.org/10.1016/J.INPA.2018.08.007

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., Kim, H., & Kanae, S. (2013). Global flood risk under climate change. Nature Climate Change, 3(9), 816–821. https://doi.org/10.1038/nclimate1911

Leta, O. T., El-Kadi, A. I., & Dulai, H. (2018). Impact of Climate Change on Daily Streamflow and Its Extreme Values in Pacific Island Watersheds. Sustainability 2018, Vol. 10, Page 2057, 10(6), 2057. https://doi.org/10.3390/SU10062057

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. Geophysical Research Letters, 44(11), 5844–5853. https://doi.org/10.1002/2017GL072874

Yamazaki, D., Kanae, S., Kim, H., & Oki, T. (2011). A physically based description of floodplain inundation dynamics in a global river routing model. Water Resources Research, 47(4), 4501. https://doi.org/10.1029/2010WR009726