Dear Editor,

We appreciate comments and suggestions from the reviewer/editor. We have carefully addressed the comments in the revised manuscript. Please see the details below.

Best regards
Vimal

Response to the reviewer's comments

1. The large-scale H08-CaMaFlood model used by the authors successfully (NSE>0.6) reproduced hydrographs of daily flow over a long-term period in almost all river basins under consideration. For me, this result is quite unexpected, since, as a rule, global hydrological models poorly reproduce the seasonal variation of river flow, even with flow averaging over larger time intervals than a day (Hattermann et al., 2017; Krysanova et al., 2018). Specifically, H08 model, as far as I can judge from publications, has not yet given such good results for river basins located in monsoon climates (see, for example, Yoshida et al., 2022). The result obtained by the authors is important because it can expand our understanding of the effectiveness of global hydrological models at the scale of river basins. Taking this into account, I would like to see a more detailed description of the methods for setting the model's parameters, its calibration and verification, and other important, from the authors' point of view, details that made it possible to achieve this result.

Thanks. We appreciate the comment. The model calibration is crucial for the performance of the any hydrological models. Most global scale hydrological models are often not calibrated due to lack of data and/or the effort that is needed. Moreover, our modeling setup has been well calibrated for streamflow and reservoir storage. Moreover, we have also evaluated the performance of the H08 model against the satellite-based ET and soil moisture (Kushwaha et al., 2021; Journal of Hydrology). Therefore, the model shows relatively a better performance compared to global scale studies.

We have added the following text in the revised manuscript (Lines 131-141):

"Large-scale global hydrological models do not perfectly capture the observed trends and variations as these are often not well calibrated at river basin scale (Krysanova et al., 2018). The H08 model performs well when calibrated at the river basin scale rather than coarser domains such as climate zones (Chuphal & Mishra, 2023; Yoshida et al., 2022). Here, we manually calibrated the H08 model by adjusting four key parameters that considerably influence streamflow for each river basin, which include single-layer soil depth, gamma, bulk transfer coefficient, and tau (Hanasaki et al., 2008; Raghav & Eldho, 2023). A more detailed discussion about the calibration parameters of H08 are discussed in Dangar & Mishra (2021). Different sets of combinations of calibration parameters within a range were used to calibrate the H08 model. The employed sets of parameters for the 18 river basins in the Indian sub-continent are listed in Table S2. The calibrated parameters account for the effect of human interventions because the model calibration is performed against the observed streamflow rather than the naturalized streamflow (Duc Dang et al., 2020)."

2. The presented results for reproducing high flow and flood inundation (Fig. 2, 3, S2-S4) do not convince one of the possibility of using them to answer the research question formulated in the article: "How does the flood risk vary at the sub-basin scale in India for the observed worst floods that occurred during the 1901-2020 period?". This is not surprising since flood risk assessments require the use of rainfall-runoff and hydrodynamics models of much greater spatial resolution, which are able to take into account local runoff formation mechanisms, local topography, etc. The authors are well aware of the limitations of the model used and clearly articulate this in the Discussion section. At the same time, the authors pay excessive, in my opinion, attention to the analysis of specific catastrophic floods and the comparison of their simulated and observed characteristics, including using satellite-based flood extent data. Given such high uncertainty in modeling results, their agreement with observed data may be coincidental. I recommend shifting the focus of the article to the analysis of characteristics averaged over a long-term period (such an analysis is illustrated in Figures 6-9) by changing the research question as: "How does the flood risk vary at the sub-basin scale in India during the 1901-2020 period?"

Thanks. We acknowledge the limitations of our current model, which have been highlighted in the discussion section of the revised manuscript. We also understand the need for more sophisticated spatial resolution in rainfall-runoff and hydrodynamics models for accurate flood risk assessments, considering local runoff formation mechanisms and topography. However, the choice of hydrological/hydrodynamic models depends on the objective of the study. Our main aim is to identify the sub-basins in India that have high flood risk based on the long-term (1901-2020) observational record, for which, we feel that our hydrological modeling framework has satisfactory performance. Moreover, the smallest sub-basin in our study is larger than 5000 km^2 and at this scale the model performance to examine the flood risk can be considered reasonable. We are aware of the high resolution (meter or sub-meter) scale hydrological/hydrodynamic models that need comprehensive data inputs (elevation, cross sections etc) and can be applied at sub-daily time scale to examine the impacts of flooding at a local scale (within a urban area). Since we provide regional scale assessment, the future work can use the high-resolution models to estimate flood risks within sub-basins at the selected stretches. As per your suggestion, we have revised the question from "How does the flood risk vary at the sub-basin scale in India for the observed worst floods that occurred during the 1901-2020 period?" to "How does the flood risk vary at the sub-basin scale in India during the 1901-2020 period?".

We have added the following text in the revised manuscript (Lines 382-402):

"While mapping the flood risk at appropriate spatial resolution is complex and challenging, it is vital for disaster risk reduction. Flood inundation mapping that provides the spatial extent of flooding is crucial as the first responders use it during a flood emergency (Apel et al., 2009). There are several approaches to mapping flood inundation (Teng et al., 2017). Various hydrological models have been employed for conducting flood risk assessments at a global scale (Dottori et al., 2018; Gu et al., 2020; Tabari et al., 2021). For instance, Dottori et al. (2018) used the H08 model combined with CaMa-Flood model to estimate losses resulting from river flooding at the country level. Additionally, the LISFLOOD model (van der Knijff et al., 2010) at 5 km spatial resolution was used to estimate the river flood risk in Europe (Alfieri et al., 2018). Flood risk assessment at relatively larger scales are conducted using the coarse resolution land surface hydrological models. The objective of these large scale flood risk assessment is to identify regions that are flood-prone (Dottori et al., 2018; Gu et al., 2020; Tabari et al., 2021). On the other hand, high resolution flood inundation mapping is needed to understand the local flood risk and damage caused to particular infrastructure. For the analysis of flood inundation during a particular floodat a local scale, high-resolution models such as HEC-RAS and Mike FLOOD can be employed (Khalaj et al., 2021; Nguyen et al., 2016). High resolution flood risk mapping requires comprehensive information of high-resolution topography, cross-sections of channels,

and data associated to structural measures of flood protection. However, the smallest subbasin considered in our study has more than 5000 km² area (Fig S7), while most subbasins have area between 10,000 and 50,000 km², with Lower Yamuna being the largest subbasin, with an area of 124,867.19 km². Therefore, the performance of our modelling framework against the satellite and other observations can be considered satisfactory to provide a sub-basin scale flood risk assessment. Moreover, we used hydrodynamic modelling to develop long-term flood inundation maps for the Indian sub-basins. The long-term data (1901-2020) provides us a record of several floods, which can help in robust estimates of flood risk in different sub-basins.“

3. The article must be formatted in accordance with the requirements of the journal.

Thank you. We have formatted the article according to the journal requirements.

Hattermann, F.F., et al. (2017) Cross-scale intercomparison of climate change impacts simulated by regional and global hydrological models in eleven large scale river basins. Climatic Change, 141 (3), 561–576. doi:10.1007/s10584-016-1829-4

Krysanova, V., Donnelly, C., Gelfan, A., Gerten, D., Arheimer, B., Hattermann, F., Kundzewicz, Z.W. (2018) How the performance of hydrological models relates to credibility of projections under climate change. Hydrological Sciences Journal, 63(5), 696-720 DOI: 10.1080/02626667.2018.1446214

Yoshida, T., Hanasaki, N., Nishina, K., Boulange, J., Okada, M., & Troch, P. A. (2022). Inference of parameters for a global hydrological model: Identifiability and predictive uncertainties of climatebased parameters. Water Resources Research, 58, e2021WR030660. https://doi.org/10.1029/2021WR030660

Thanks for the suggestions. We have cited these in the revised manuscript.