Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2019-388-RC2, 2019

Hydrology and Earth System Sciences
Discussions

Interactive comment on "Historical and future changes in global flood magnitude – evidence from a model-observation investigation" *by* Hong Xuan Do et al.

Anonymous Referee #2

Received and published: 24 September 2019

This manuscript describes the work of an extensive model study and in its final version will be for sure appreciated by the readers of HESS.

General Comments:

As the study presented is quite extensive, it is sometimes difficult to follow study setup and all the analysis steps. Therefore, the authors should provide a detailed schematic, showing the main building blocks of their study and the different steps of analysis (preferably showing the section numbers in the schematic as well) to allow the reader to have a complete 'picture' of the study design, before embarking on the details in the main text.

Additionally, due to the complexity of the study and details provided in the result section, I think a summary table or bullet points at the end of the study would be helpful for the reader to get a better overview of the key results obtained.

Another important point is that the study uses 7-day annual maximum as a surrogate for 'food'. This fact needs to be made more explicit throughout the study to avoid misunderstandings from the general perception of flood, which would shorter (e.g. often 1-day). This is of importance, as the results might be quite different. I.e. a single day peak value trend study will show different results, not only in terms of magnitude of change, but also in terms of the flood hydrograph shape. E.g. if floods would become flashier in some location in future, it might look as if the trend of a 7-day maximum might not change at all or get smaller, but the peak day could be of much higher magnitude. The authors need to make sure they call the variable under investigation for what it is, i.e. not calling it 'flood', 'peak discharge' or 'streamflow maximum' to avoid misunderstanding of the results. Along this line, I also think that the title '. . . changes in global flood magnitude . . . ' is also misleading. The study shows rather an 'global assessment of the 7-day annual maximum average value'. Please consider changing the title to better represent the content of the study. Additionally, to avoid misinterpretations of your results please avoid using the term 'hazard' in its current form in the manuscript, as hazard means: hazard=risk*exposure (which is not the correct terminology here). The same also applies to the term 'risk' which is related to 'probability and consequences'.

In this manuscript I feel that the GHM are used by the authors as 'black-box' that give some output. However, for this study to be valuable, it would be important that the authors would try to relate the observed differences/deviations in the outputs to the actual differences in the hydrological model setup. The authors just state ". . . there are potential effects of technical discrepancies to the findings which cannot be checked in the context of this study" (L 126). However, I think based on the model selection, the authors should have a notion of why they selected certain models and what the

key differences are. Hence, the authors should at least try to come up (also based on past literature) with some sort of reasoning for model selection and also more importantly an interpretation of their findings... For example, are the changes the models are giving as an output considered in line with the current understanding of the effects of climate change on floods or are there surprising results? I think this could be done in a separate paragraph discussing/comparing with previous literature.

In several instances in the manuscript, the authors are highlighting the 'substantial influence of the atmospheric forcing in driving the spatial structure of the simulated trend'. I think this is another important point that needs to be discussed in more detail in the discussion section, i.e. why to the hydrological models have little influence....

Overall, I think a new separate discussion section of the results of such a complex analysis would be beneficial, as this would free up the room for a better refined summary and conclusion section, that focused on the key results and the overall implications of the results not just for the scientific world but also for the 'end-users', such as decision makers etc.

Specific Comments:

L37: For clarity, please provide significance level used in this study in parentheses.

L38: replace the term 'high-risk location'.

L54: Please provide reverence to this statement

L77: What is 'factorial evidence' in this regard? Please elaborate.

L121-122: Please elaborate why the authors think that the 'naturalised runs and the human impact runs exhibit similar characteristics of trend' Would one not expect considerable differences?

L126: What are the 'potential effects'. Can you briefly elaborate.

L127: Please also elaborate what the effects/impacts of this on the results are.

L158: What is the rationale of 335 days. Please explain briefly.

L172: Fig1: These colours are not 'safe' for colour-blind readers. Please use different colour combination

L184: 'Our preliminary analysis... did not lead to substantial changes'. So what were the 'not so substantial changes' one is wondering?

L192: Can you please name the 'three identified objectives' again as it is quite difficult to keep up with this extensive work.

L210: To spare the reader from having to go to the original reference, please name the field significance test used and elaborate briefly what exactly is evaluated.

L211: What 'Pearson's (spatial) correlation' was used? Reference? What variables are correlated?

L220: Please replace the term 'flood hazard' with something more appropriate to what has been done. This also applies to the subsequent usage, as well as the term 'flood-risk' later used in the manuscript.

L245 & 493:to me it does not look like norther Europe has increasing trends. Scandinavia etc looks decreasing.... Please check.

L258: I agree, very much with this point. The study analyses 'extremes (i.e. floods) but then model 'averages' are provided. His is counter intuitive. This can lead to strong underestimation of the actual changes. The usage of averages vs individual models that show extremes should be better discussed in the discussion section. Hence, I also agree with L 419.

L281: is this really 'the spatial pattern of trends' that is evaluated or is it a cell by cell comparison? Please elaborate and have in mind that although a correlation is it can still mean that the overall spatial pattern (i.e. approximate location of increasing and decreasing trends) might still be correct.

L 370-384: The authors mention ‘a significant difference between trend characteristics from all model grid cells compared to those obtained from the observation locations’ and the conclude that ‘’ that trends exhibited from observation locations are not a representative sample of trends obtained from all simulation grid cells‘’ (L379-380) And then call ‘’to improve data accessibility and expand streamflow observational networks‘’. However, if there are such ‘’significant difference even in data rich regions, how can one justify expanding the network based on the previous finding? Instead to me this reasoning would rather require the need to improve our models instead (notwithstanding the fact that I agree with the data needs mentioned by the authors.)

L 460: Maye the authors can elaborate a little more what an ‘flexible adaptation strategy’ entails in terms of flood mitigation. Any suggestion on how this can be achieved under tight budgets. Can we as scientists not provide any guidance than just saying ‘stay flexible’ to those who have to take decisions know?

L531 & 534: Along the lines of improved GHM: It is not only important that the spatial patterns are being reproduced correctly but also that the timing of the high-flows/floods are being modeled correctly. I.e. ‘the ïňĆood seasonality patterns can be used as an additional metric to test large-scale hydro-logical models for their ability to reproduce the spatial and temporal ïňĆood characteristics.’ (Hall and Bloschl, 2018, HESS). As this would give more confidence that the models actually get the flood generation processes correctly.

L 538: What does ‘constraining ’ entail? Please briefly elaborate. Would this prevent the model to adjust to changes in the flood generating processes, as one would expect to happen in some regions of the world. E.g. from snow-melt floods to rainfall-generated floods?

L 550-559: I agree with this call, as this is very important. However, one needs to keep in mind that in many countries maintaining monitoring networks and data curation is/is considered too expensive. Hence it needs to be made clear to decision makers that

such data is of importance. However, I know of cases where countries/agencies have been or are currently considering discontinuing their data networks, as they don’t see the benefit or don’t see their data being used (partly lack of proper citation of the (often freely available) original data source). This implication needs to be kept in mind when large datasets of observational data are being compiled and subsequently only credit is given to the compiled data... This hides to the funding/responsible agencies the usage of their data (i.e. the original data source) and might lead to the misconception that their data is not being needed/downloaded and hence the data network can be discontinued and to allocate funds to more (perceived) useful sectors...

Fig S5: Suggest using same y-axis scale for all panels on the left/right to be able to compare the regions better with each another.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., https://doi.org/10.5194/hess-2019-388, 2019.