In this document, the review comments are in black, our responses is in red and the revised text are in blue.

This paper discusses the generation of future hazard maps in a global inundation modelling context. As present day global hazard models have become more common and more detailed interest has been building in creating future hazard maps from these data. This study is a timely analysis examining some of the decisions involved in propagating GCM analysis into these hazard data sets and thus a valuable addition to the literature.

Overall the analysis has been done well and the paper is clearly written with an easy to follow structure. The novelty of the work could be more precisely defined in the induction because many local scale climate conditions hazard maps have been created in the past (see specific comments below). Furthermore, although I appreciate why the authors have developed a method to ensure monotonically decreasing water levels, I don't think the rational for doing this (backwater effect) is a clear cut as suggested in the text because a hazard map is not a physically coherent event. If the authors agree with my comments then they should be able to modify the arguments in the text and publish the research.

Reply:

We would like to express our gratitude to referee #1. We will address all the comments in the revised manuscript, and comprehensive explanations are provided below.

Backwater_Modificiation is applied in order to revise the spatial inconsistency that is due to the distribution fitting (fitting Gumbel distribution) at each grid (unit-catchment). It is possible that it would be overestimated by applying Backwater_Modification because they are corrected at all reverse slope occurrence points. Reversed water surface slopes might exist in conventional flood hazard maps, given that they are not always constructed by one flood event simulation as pointed out by referee #1.

In light of the above perspectives, we have conducted validation to investigate whether Backwater_Modification improves the consistency to national hazard maps or not. In result, we have found that the Backwater_Modification brought us closer to the hazard map prepared by Japanese goverment, indicating that it is important to apply Backwater_Modification.

Specific points:

Line 58: It would be good to add some references on in country hazard mapping. In general as this paragraph is an overview of local-scale flood hazard mapping efforts I found it a bit citation lite. It would be nice to signpost some other review articles as De Moel et al 2009 was published some time ago.

At the risk of self-promotion (no need to include any of this unless it's useful to support your later arguments) we included a UK focused review of local-scale modelling efforts that highlights some of the challenges of using these data sets in a climate conditioned application.

Bates, P. D., Savage, J., Wing, O., Quinn, N., Sampson, C., Neal, J., and Smith, A.: A climate-conditioned catastrophe risk model for UK flooding, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2022-829, 2022.

Reply: We agree with referee #1 suggestion, so we will add "Bates et al., 2022" and "Wing et al 2022" as a reference.

Bates, P. D., Savage, J., Wing, O., Quinn, N., Sampson, C., Neal, J., and Smith, A.: A climate-conditioned catastrophe risk model for UK flooding, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2022-829, 2022.

Wing, O. E., Lehman, W., Bates, P. D., Sampson, C. C., Quinn, N., Smith, A. M., Neal, J. C, Porter, J. R, and Kousky, C.: Inequitable patterns of US flood risk in the Anthropocene. Nature Climate Change, 12(2), 156-162, https://doi.org/10.1038/s41558-021-01265-6. 2022.

"In the research field, for example, Bates et al., 2022 estimated annual flood damage in UK and Wing et al., 2022 also estimated flood damage in US by using local-scale flood-hazard maps. "

"Line 81:assessment of future flood risks based on the spatial distribution of inundation depths has not yet been established" I agree at the global scale but there are many local scale examples of such studies that have had to consider the GCM bias issue. I think that needs to be acknowledged explicitly in this section or the previous paragraph.

Reply: We would like to express our gratitude to referee #1. We will revise the sentence as follows. We will clarify that in this paper we described a “global” study. We will also clarify t hat flood “risk” has been assessed, but flood “hazard” itself has not been sufficiently verified in previous global-scale studies.

“While assessment of climatic and meteorological hazard under climate change (e.g. extrem e temperatures, droughts and heavy-rainfall events) has been widely performed using direct output variables of general circulation models (GCMs), such as precipitation and temperatur e (Li et al., 2021, Lu et al., 2019), at present, no global high-resolution flood hazard (i.e. in undation depth) in a future has been sufficiently verified. Even though some studies assesse d future flood risks (e.g. affected population and GDP) at the global scale (Ward et al., 2020 b), it is important to analyze global future flood-hazard (i.e., inundation depth distributions), and also important to assess uncertainties such as those caused by different bias correction s.”

For example, this sentence “While bias-correction methods for precipitation and temperature have been studied in detail (Watanabe et al., 2012, Hempel et al., 2013 and Lafond et al., 20 14), such methods have not been established for runoff data for use as inputs to flood models “ is not correct without a more precise focus on the novelty of this study because there are m any examples of localized climate conditions hydrological cascades all the way to inundation h azard and impact. These studies have some of the same issues you so I would have thought t here were some useful conclusions from the literature that might feed into this section.

Reply: We will revise the sentence to clarify and focus on the novelty of this study. Specifica lly, we will change from “such methods have not been established for runoff data for use as inputs to flood models” into the following sentence.

“such methods have not been established for runoff data for use as inputs to **global** flood m odels to **construct large-domain** future flood **hazardmap.**"

Line 113: “Although CaMa-Flood is a global model, it is unique in that it represents the physic al processes necessary to reproduce floodplain inundation dynamics” I appreciate there are un ique aspects to CaMa-Flood, however there are several global flood inundation models so perhaps remove the word “unique” or be more specific about the combination of capabilities that are unique.

Reply: In response to referee #1 comments, we will remove "unique" and revise the sentence as follows.

“Although CaMa-Flood is a global model, it **has characteristics** that it represents the physical processes necessary to reproduce floodplain inundation dynamics.”

Line 137: could you be specific about the simulation length – I suspect its shorter than most readers will assume by the wording long term. Is it 35 years?

Reply: As you understand, the simulation length is 35 years. Based on referee #1 comment, we will remove "long-term" and added "(time period: 1980-2014)" to the end of the sentence to specify about the simulation length.

“we conducted a historical river hydrodynamics simulation with a daily time step using observation-based runoff data (Reanalysis_Runoff) as an input to CaMa-Flood **(time period: 1980-2014)**.”

Line 150: I have no issue with this pragmatic method for getting a spatially consistent hazard map, but I’m not convinced by the description. The hydrodynamic model simulated the backwater effect so I assume that this effect is captured in the hazard/level simulations. Could it be the case that the spatial inconstancy is therefore due to the distribution fitting and other data processing factors rather than a physical backwater effect being omitted? I appreciate the maps will look better with this correction but do you risk biasing the levels higher in the process? On a more fundamental level a hazard map is not an event, so does the water level even need to monotonically decrease downstream since you simple seek to best simulate the hazard rather than a plausible physical water surface which a hazard map will never be?

Reply: We thank referee #1 for comment. we will revise the sentence to make it clear that it is the case that the spatial inconstancy is due to the distribution fitting at each unit-catchment scale (grid-scale) as follows.

"If reverse-slope revision is not conducted, **reverse slope occurring through fitting of Gumbel distributions remain** and the inundation-depth distribution may not be physically reasonable. For this reason, a novel reverse-slope revision method was applied in this study **in the purpose of revising the spatial inconstancy that is due to the distribution fitting at each unit-catchment scale (grid-scale)**."

In response to referee #1 comment " I appreciate the maps will look better with this correction but do you risk biasing the levels higher in the process? ", we will add the notes of overestimation to the last paragreaph in section 4.2 (Specifically, the following sentence will be added).

"Please note that it is possible that it would be overestimated by applying Backwater_Modification because they are corrected at all reverse slope occurrence points. Reversed water surface slopes might exist in conventional flood hazard maps, given that they are not always constructed by one flood event simulation. In light of the above perspectives, we have conducted validation to investigate whether Backwater_Modification improves the consistency to national hazard maps or not."

Section 2.3.1: Could you cite some alternatives to the bias correction method and explain why this one was chosen. I'm not an expert in these methods, but my assumption is that the results of this method would be sensitive to the choice of bias correction. Depending on your response above you might need to caveat the conclusions drawn around line 336.

Reply: We thank referee #1 for comment. We will cite some alternatives to the bias correction method and explain why this one was chosen.

"As an alternative to addictive correction method, there are multiplicative correction method, which multiplies the ratio with the reanalysis data, and Quantile Based Mapping method (Panofsky and Brier 1968, Watanabe 2020), which is to obtain ordinal statistics from the reanalysis data and the GCM, and to create an equation relating these statistics. In the future cl imate, the average monthly runoff may fluctuate significantly due to changes in the arid zones caused and changes in the timing of the wet and dry seasons by increasing temperatures. Based on the above, this study used addictive correction method, which is relatively insensitive to the above fluctuations."

Did you consider looking at more frequent hazards? I assume the runoff bias correction method would work better the closer you were to the mean annual flood because the bias would be better corrected? Or maybe the distribution fitting is still impacted…

Reply: We looked at more frequent hazards, the runoff bias correction method worked closer to reanalysis data than RP100 in some cases. For example in GRDC BAN_BANG_KAEO station (100.4533°E, 14.5847°N), as below figures are seen, the runoff bias correction method worked closer to reanalysis data at more frequent hazards (e.g. RP2-10) than RP100 in cases.

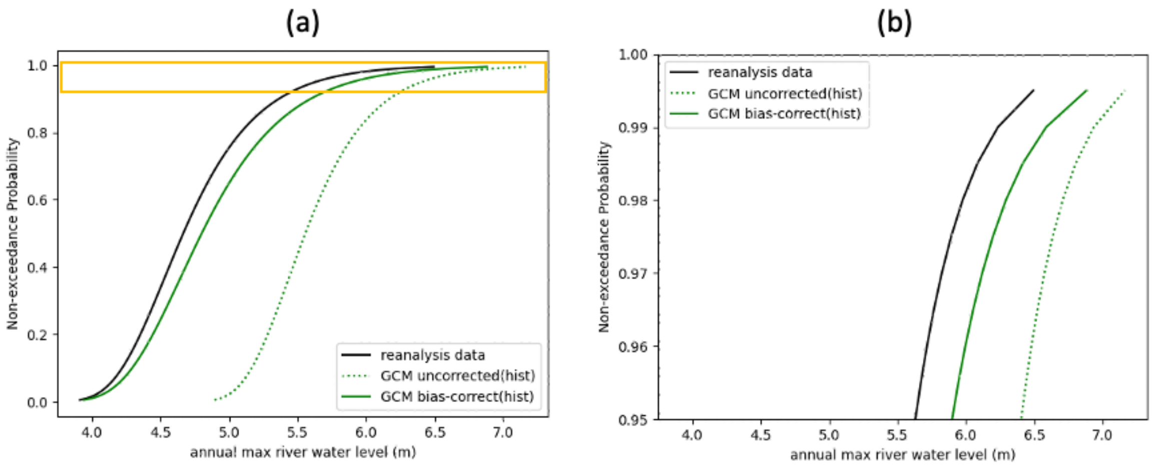

Figure : Comparison of reanalysis data, uncorrected GCM, bias-corrected GCM at GRDC BAN_BANG_KAEO station (100.4533° E, 14.5847° N). (a) Gumbel distribution of annual maximum river water levels and (b) enlarged view of the orange square in (a).

Line 383: "On the other hand, the lookup method did produce hazard maps consistent with the changes in flood risk under the future climate projected by the GCMs." I found this a bit confusing. Do you mean flood risk here? I assume you can only derive risk from the hazard maps, do you mean the changes in GCM runoff or precipitation? I think this section needs to be more precise about what you are looking for constancy with because I don't think its risk – or at least I didn't understand how it could be from the text.

Reply:

As you pointed out, because we use “flood risk” uncorrectrly in the sentence, we will revise the sentence to clarify “the changes in flood **hazard** under the future climate projected **by CaMa-Flood simulations**“

Specifically, we will change from “the changes in flood risk under the future climate projected by the GCMs” into as follows.

“the changes in flood **hazard** under the future climate projected **by CaMa-Flood simulations with input of GCMs runoff**”

Line 390: “4.1 Which method is more convenient for generating future-hazard maps?” Convenient is the wrong word here it only applies to some of your arguments.

Reply: We agree with your comment. we will change from "convenient" into "reasonable" in the title of section 4.1.

“4.1 Which method is more **reasonable** for generating future-hazard maps?”

Line 394: I think you could also note here that this approach can also be used with hard mapping methods based on regional flood frequency analysis or machine learning from gauging station data, which at least in historically well monitored river reaches can be more accurate than rainfall-runoff reanalysis based methods (perhaps this is contested) for present day hazard mapping and extreme flow estimation because the modelling chain is much shorter (e.g. Laura Devitt et al 2021 Environ. Res. Lett. 16 064013 DOI 10.1088/1748-9326/abfac4)

Reply:

Thank you very much for your suggestion.

As you pointed out, there is no need to limit it to reanalysis-basis as a “historical hazard map". Based on the above, we have added the following sentence to the the end of 1st paragraph in section 5 “Conclusion”.

“ This implies that combining accurate historical hazard maps with information on future frequency changes of floods is considered optimal in general for generating future hazard map. Please note that the historical flood hazard map does not have to be reanalysis-based simulation using GFM, and the proposed method can be also applicable to gauge-based or machine-learning based historical hazard map.”

Line 407 I think this is a separate point from the one above “The lookup method also has the advantage of facilitating research on efficient construction of future climate hazard maps, as it allows for improvement of the reanalysis hazard map by upgrading the model, and the estimated changes due to climate change can be considered separately” and you could expand to discuss using multiple flood hazard models for the analysis as several are introduced in the preceding section.

Reply: We would like to express our gratitude to referee #1 suggestion. We will add one point to advantage of the the lookup method

The lookup method also has the advantage of facilitating research on efficient construction of future climate hazard maps because preparation of historical hazard map and estimation of future frequency change can be separated. This is beneficial for two aspects: 1) it allows for improvement of the reanalysis hazard map by upgrading the model; **2) it allows for use of multiple reference hazard maps by using different reanalys-based simulations.**

Section 4.2: I think it’s valuable to discuss this correction (so I’m not suggesting a major revision). But in my opinion it’s debatable whether this correction is desirable for a hazard map which is not an event and thus not a physically possible water surface. I think it potentially double counts for the backwater effect and over-predicts the flood inundation levels by biasing locations upstream to any over-prediction errors in the distribution fitting downstream. That said the hazard mapping is improved relative to some high quality validation data so I don’t dispute that some form of postprocessing of the levels to aid spatial constancy is the wrong thing to do … but you could smooth for example. Personally I would slightly modify the discussion in section 4.3 to present the results with and without the backwater modification as an indicator of th e sensitivity to this issue – which is less than the climate change signal and impact of different GCM treatments. This would also require a small edit to the conclusion ~line 532.

Reply: We would like to express our gratitude to referee #1.

Backwater_Modificiation is applied in order to revise the spatial inconsistency that is due to the distribution fitting(fitting Gumbel distribution) at the unit-catchment scale. In response to referee #1 comment "I think it potentially double counts for the backwater effect and over-predicts the flood inundation levels by biasing locations upstream to any over-prediction errors in the distribution fitting downstream." we will add the notes of overestimation (Specifically, the following sentence will be added) in the last paragraph on section 4.2.

"Please note that it is possible that it would be overestimated by applying Backwater_Modification because they are corrected at all reverse slope occurrence points. Reversed water surface slopes might exist in conventional flood hazard maps, given that they are not always constructed by one flood event simulation. In light of the above perspectives, we have conducted validation to investigate whether Backwater_Modification improves the consistency to national hazard maps or not."

Line 465: why choose lower resolution population data rather than say Worldpop? Could this bias the results?

Reply: We would like to express our gratitude to referee #1 question.

In this study, as to the estimates of future flood risk change, we compared the differences between the two methods to construct a future hazard map or with/without implementation of bias correction. Since the purpose is **to analyze the impact of correction at the global scale**, we think that 30 arcsec resolution was enough.

We used Gridded Population of the World; (CIESIN, 2018), which has data that are stored together for the entire world, rather than Worldpop, whose data are broken down by country.

In addition to above, we will add the following notes regarding uncertainty in spatial resolution in section 4.3.

"As per Zhou et al., 2021 discussed, the spatial resolution of the flood hazard map is a particularly important determiner of impact assessment. Smith et al. (2019) evaluated the popul ation exposure to a 1-in-100 year flood in 18 developing countries, showed that decreasing the spatial resolution of flood hazard map from 90 to 900 m increases the exposure by 51 % to 94 % for different population products. Although there is uncertainty involved in the choice of the spatial resolution of the flood hazard map as per above, we used 30 arcsec resolution instead of 3arcsec resolution for the purpose of comparison between methods to construct a future flood hazardmap on a global scale in this section.”