Authors' reply to Reviewers

We wish to thank both Reviewers for their useful comments that hopefully have helped us to improve our manuscript. We have compiled below a point-by-point reply to the Reviewers' individual comments and submitted a track-changed version of the manuscript so that edits and new additions are easy to find. Please note that there are few additions that we have made in response to additional comments we received through the NHESS interactive discussion, which we also thought useful addressing.

Response to Reviewer 1 (Dottori)

The manuscript by Pianosi et al. proposes a framework to analyse the sensitivity of large-scale flood models, taking into account the most relevant sources of uncertainty. In particular, the framework aims at reducing the complexity of the analysis while maintaining the significance of results.
I appreciated the opportunity to review this work. It's good to see a collaboration between academic and private sectors and I believe the topic is of high interest for the scientific community working on large-scale flood models.
Having said so, I have some remarks mainly regarding the application of GSA to the two case studies.

Reply: Thanks for appreciation of our paper and constructive criticisms that are very useful to improve our manuscript.

Main points
- The authors do not describe the JBA flood loss model but do provide some references to past works. However, my impression is that the outcomes of the application examples might be influenced by some modelling assumptions (see my following points) and therefore I would suggest including at least a short description of the JBA model.
- Uncertainty in hazard maps: previous studies have shown that hazard maps from large-scale flood models might show limited changes in flood extent and depths across return periods. For example, the 1-in-50-year and 1-in-100-year flood maps might be similar (see Bernhofer et al. 2018, and Aerts et al. 2018). If this is the case for JBA model, this might explain why this parameter has a limited importance in GSA outcomes. Moreover, the authors should explain how flood protection standards are considered in the model, given that they are a major component of hazard map uncertainty (Paprotny et al 2025).

Reply: Thanks for raising these important points, which we have added to the description of the model (Sec. 2) and the Rhine River case study section (Sec. 5.1) in the revised manuscript.

In brief, an uncertainty analysis and GSA of flood hazard maps was conducted internally, though not yet published, which guided our thoughts on the return period (RP) uncertainty bounds. The key result from the internal study was that maximum flow rate and DTM uncertainty were the biggest contributors to flood map depth uncertainty. Hazard maps also varied in extent with RP. How much the depths and extents increase with RP is largely determined by the underlying topography, but also other parameters play a role, including the hydrological inputs and surface roughness. As for the flood protection standards, these are taken into account within the model calculations. Specifically, if an asset falls behind a flood defence and is affected by an event with return period lower than the standard of protection (SOP) of the defence, the asset is considered fully protected and no flooding is expected to occur. If instead the return period exceeds the defence SOP, the flood model uses a growth curve approach (Kjeldsen, Jones and Bayliss, 2008) to calculate the severity of the potential flood overtopping the defence.

In the revised manuscript, we have expanded the description of the model (Sec. 2) to include the above points. In the Rhine River case study section (Sec. 5.1), we have clarified that our analysis does not consider uncertainty in the representation of flood defences (see edits to Table 1) and discussed how this choice may have led us to underestimate flood hazard uncertainty and its influence on risk estimates (see changes to end of Sec. 5.1).

- Uncertainty ranges in Table 1: the ranges of exposed values and damage functions applied in GSA come from a previous work by Sarailidis (2023). This is fair enough, however, the assumptions and limitations of the study by Sarailidis should be better explained. For instance, my understanding is that Sarailidis obtained the range of damage ratio reported in table 1 by perturbing the values from multiple damage functions found in literature, which in my opinion might lead to overestimate the uncertainty associated to this parameter. Given the outcomes of GSA for the Rhine case study, it is important to communicate that these outcomes might be influenced by this and other assumptions.

Reply: Thanks for raising this point, and yes the reviewer is right that we inferred the ranges using multiple damage functions, not only step functions. So, we agree we are possibly overestimating parameter uncertainty (for the given vulnerability function). In the revised manuscript, we have added more details to the description of how we built the uncertainty ranges (first paragraph in Sec. 5.1), and the impacts that our choices may have had on our results, particularly towards emphasising the role of vulnerability curves uncertainty (see edits to last paragraph before Figure 5). We also pick up again on this topic in the Outlook and Conclusion sections, to highlight once again that GSA results are conditional on the assumptions made on the uncertainty ranges.

Minor points:

- Abstract: The authors begin with "Flood loss models are increasingly used in the (re)insurance sector to inform a range of financial decisions." I would add that flood loss models are increasingly important also in research, (for instant, to understand present and future risk trends as discussed in Ward et al., 2015), as well as in policy support (see for instance the PESETA programme, https://joint-research-centre.ec.europa.eu/projects-and-activities/peseta-climate-change-projects/jrc-peseta-iv/river-floods_en). This would highlight the relevance of the

present study beyond the insurance sector.

Reply: Thanks, an important point that we have now added in the Abstract and the Introduction.

- One of the main outcomes is the relevance of vulnerability functions in driving model uncertainty, I believe this should be mentioned in the abstract.

Reply: We appreciate the Reviewer suggestion, but we are reluctant to add this point. We believe that the key contribution of our paper is not in providing insights about any of the specific analysed case studies but rather in discussing generic methodological issues and possible solutions when applying GSA to large-scale models. The Abstract is therefore focused on these methodological aspects, and we would not like to change this by adding a comment on the GSA application results. Still, we agree the consistent relevance of vulnerability functions is an interesting outcome and it is one of the key points we comment in the Outlook and Conclusion.

- Reference list: missing title for Galloway et al (2025)

Reply: Thanks for spotting, the reference was corrected.

References

Aerts, J. P. M., Uhlemann-Elmer, S., Eilander, D., and Ward, P. J.: Comparison of estimates of global flood models for flood hazard and exposed gross domestic product: a China case study, Nat. Hazards Earth Syst. Sci., 20, 3245–3260, https://doi.org/10.5194/nhess-20-3245-2020, 2020.

Bernhofer Bernhofen, M. V., Whyman, C., Trigg, M. A., Sleigh, P. A., Smith, A. M., Sampson, C. C., Yamazaki, D., Ward, P. J., Rudari, R., Pappenberger, F., Dottori, F., Salamon, P., and Winsemius, H. C.: A first collective validation of global fluvial flood models for major floods in Nigeria and Mozambique, Environ. Res. Lett., 13, 104007, https://doi.org/10.1088/1748-9326/aae014, 2018.

Paprotny, D., 't Hart, C.M.P. & Morales-Nápoles, O. Evolution of flood protection levels and flood vulnerability in Europe since 1950 estimated with vine-copula models. Nat Hazards 121, 6155–6184 (2025). https://doi.org/10.1007/s11069-024-07039-5

Ward, P. J., Jongman, B., Salamon, P., Simpson, A., Bates, P., De Groeve, T., Muis, S., de Perez, E. C., Rudari, R., Trigg, M. A., and Winsemius, H. C.: Usefulness and limitations of global flood risk models, Nat. Clim. Change, 5, 712–715, https://doi.org/10.1038/nclimate2742, 2015.

Response to Reviewer 2 (Paprotny)

The manuscript 'Global sensitivity analysis of large-scale flood loss models' applies the GSA method to two case studies. As the main author has published about GSA in the past, the main contribution are the case studies. However, I see several issues with how those are described and evaluated, as already also mentioned by other commenters here. It reduces the value of an otherwise interesting and potentially useful publication, but I believe the authors will be able to address all issues in the revision.

Reply: We wish to thank the reviewer for their constructive criticisms, which we hope will help improve the clarity of the paper. Detailed responses on how we did this are given below.

I would particularly stress the importance of comments by the other reviewer, Francesco Dottori. I concur that the lack of description of the JBA flood model is an issue to be addressed. No mention of flood protection levels is also troubling, given that it is often found to be the most sensitive parameter (see e.g. https://doi.org/10.5194/nhess-18-2127-2018).

Reply: We agree and in the revised manuscript we have now included this important point. In particular, in Section 2 we added a description of how flood defences are represented in the model and in Section 5.3 we have clarified that not varying the associated protection standard means we are likely underestimating the uncertainty in the hazard component, and used the references suggested by the reviewers to back up the point.

Also, the inconsistency of uncertainty ranges in Table 1 is problematic: vulnerability curves are assumed to have wide margins, much more than the flood event set and hazard maps (which are given arbitrary +/-50% range) or exposure (which has about +/-25% range). This rather unsurprisingly leads to conclusion that vulnerability is the most sensitive component for the Rhine. Better explanation of the choice of uncertainty ranges is needed.

Reply: We agree with the Reviewer that the use of wide uncertainty ranges for vulnerability curves may be a reason why this input uncertainty is dominant in the Rhine River case, although we disagree with the statement that these ranges are "arbitrary". Ranges were defined based on a systematic literature review, as described at the beginning of Sec. 5.1. In the revised manuscript, we have expanded this description and discussed some of the limitations of our approach. We clarified the impacts that our choices may have had on the results, particularly towards emphasising the role of vulnerability curves uncertainty (see edits to last paragraph before Figure 5). We also pick up again on this topic in the Outlook and Conclusion sections, to highlight once again that GSA results are conditional on the assumptions made on the uncertainty ranges.

I would highlight that the Rhine catchment was subject to a GSA study before, which is not mentioned by the authors (https://doi.org/10.5194/nhess-18-3089-2018). There, a major sensitivity of impacts to changes in flood protection, land use and reservoirs is highlighted. The authors of that analysis also used well-designed uncertainty ranges with a strong rationale from literature and empirical data. It would be important to compare and discuss those results in context of the results presented here.

Reply: Thanks for pointing us to the paper by Metin et al 2018, which is very relevant for our work, and we have now included in our literature review and discussion of the Rhine river application.

The paper resonates with our own work in two key points: it is motivated by the "*lack of knowledge of how and to what extent changes in influencing factors propagate through the chain and finally affect flood risk*" because most of previous research has focused on uncertainties in flood hazard assessment, instead of risk; it concludes that "*components that have not received much attention, such as changes in dike systems or in vulnerability, may outweigh changes in often investigated components, such as climate*". This is very much in line with our own findings that (1) GSA literature has focused on drivers of hazard rather than risk; (2) uncertainty in vulnerability is found to be very important in our case studies.

On the other hand, there are also some substantial differences between our work and Metin et al 2018:

1) Different conceptual framings. Metin et al 2018 use GSA as a tool to learn about drivers of changes in risk in the real system, under the (implicit) assumption that the model is a valid representation of that system. Instead, in our work we use GSA primarily as a tool to learn about the model, to understand the key drivers of uncertainty in model predictions and thus inform model validation and future improvements. Both are legitimate uses of GSA but in our revised literature review we have highlighted these different ways of using GSA (see edits on p. 3).

2) Different scale of investigation. Metin et al 2018 focus on a sub-basin of the Rhine, the Mulde catchment, covering an area of about 7,000 km2. This is much smaller than we investigate in our paper (185,000 km2 in the Rhine and 1,700,000 km2 in the Queensland application). Hence, we would not put their work under the umbrella of GSA application to "large-scale" flood risk modelling, which is our specific focus as increasing scale raises specific challenges that we aim to discuss in our paper.

In the revised manuscript, we have included the reference in the literature review (p. 3) and the Rhine river application section.

The case studies are implemented with very different methods and presented also completely differently. With the selection of model components and the choice of uncertainty distributions being quite arbitrary, the paper shows that there is a wide possibility of pre-selecting an outcome of any "validation test" (mentioned in the introduction), hence going against the authors' final sentence; "We hope this paper will provide motivation as well as practical ideas to foster the application of GSA to flood risk models and contribute to increasing their transparency and legitimacy." The challenge remains to fairly evaluate different models components with realistic uncertainty ranges and indeed identify the elements that matter: not only flood protection, reservoirs, land use, vulnerability beyond depth-damage functions, but also inundation modelling method (e.g. https://doi.org/10.1029/2024EF005164, model resolution (e.g. https://doi.org/10.1029/2023WR035100 or hydrodynamic parameters (e.g. https://doi.org/10.1016/j.coastaleng.2024.104541). An improved discussion on this would be beneficial, given that the title, abstract and introduction suggests a more comprehensive, overarching study.

Reply: We agree with the reviewer that GSA results depend on the selection of which input factors to vary and how. This is indeed a very important point that we stress multiple times in our paper while also giving a range of practical suggestions on how to tackle it (Sec. 4.3, Fig. 5b, Outlook and Conclusions).

However, we do not believe that "*there is a wide possibility of pre-selecting an outcome of any validation test*", as given the complexity of the models investigated here, reverse-engineering the problem to select input distributions that lead to a desired GSA outcome sounds implausible (if a modeller could do that, they would probably know the model behaviour so well that they would not need to embark in a GSA in the first place!).

This said, we reiterate that a degree of subjectivity in the definition of what qualifies as "realistic uncertainty range" is unavoidable, and GSA results should always be interpreted as answers to the *relative* question "what are the controls of model outputs *if these inputs are varied within these ranges?*". This point is also made in the paper suggested by the Reviewer (Metin et al 2018) who acknowledge: "*The presented results are subject to limitations related to the flood risk chain model and the subjective assumptions for the reasonable change scenarios*". In the revised paper, we have further highlighted this point – see new additions in Sec. 5.1 and the Outlook and Conclusions.

We have also made a small modification to the title, now changed to "Towards global sensitivity analysis of large-scale flood loss models", as we recognise that the original title (without "Towards") could be misleading in potentially suggesting that we can draw conclusions about the sensitivity of *all* large-scale flood loss models.
We hope this new title conveys the message that the paper aims at enabling and promoting the application of GSA to large-scale flood loss models, as clearly stated in the Introduction ("*The goal of this paper is to discuss how we can begin to tackle these challenges* [in applying GSA to large-scale models] *and showcase the type of results that could be obtained through two proof-of-principle applications to a commercial model – JBA Risk Management's (hereafter JBA's) flood loss model – applied to two large domains*").

Minor points:
Several figures include underlining of errors created by MS Office

Reply: Thanks, this has now been revised

CRESTA acronym is explained after one of the subsequent uses, not the first one.

Reply: Thanks, this has now been revised

L339: what level of CRESTA units is used here?

Reply: high resolution CRESTA, this has now been revised