

The following document is structured as follows: the comments of each reviewer are presented, followed by our response comments directly after each point. Subsequently, the main revisions to the manuscript are summarized in relation to the reviewers' comments.

5 Review 1

10 The manuscript presents an interesting extension to the existing flood risk estimation and uncertainty quantification methods, with the aim of quantifying flood risk in economic terms for large river reaches. It applies a concept of simple surrogate models to the required flood model chain (rainfall-runoff simulation, river routing, dike breach/overtopping, inundation and damage estimation). It builds on the probabilistic flood risk assessment concepts developed about 20 years ago, including an uncertainty estimation using the concept of aleatory and epistemic uncertainty. The main advancement of the work is the inclusion of dike failures in a probabilistic manner and to include this in the risk and uncertainty estimation by a 2-staged Monte Carlo analysis. Moreover, the work also presents the framework for a cost-benefit analysis of large-scale flood protection measures, like flood retention basins. This is possible because the model concept enables the downstream effect of flood protection measures in a river reach. The proposed concept is valid and the work takes large scale flood estimation a step further, mainly by including probabilistic dike failures in the risk assessment. The manuscript is properly structured and the work is well presented. Overall, I recommend the publication in NHESS, but some issues should be addressed beforehand:

RC1.1: Emphasize in the introduction that this is the concept presentation, and not the final risk assessment for the river reach.

We will make sure to point that out in the abstract and introduction of the paper.

25

RC1.2: The inundation modelling is based on non-hydraulic methods, only volume-depth and volume-damage relationships are derived based on GIS analysis. In the light of now-existing fast GPU hydraulic models and AI flood modelling, this appears to be somewhat outdated and needs to be justified. The non-hydraulic nature of the inundation modelling is thus the major limiting factor in this study, because flood dynamics are not captured. Highly dynamic flood, where flow velocities and/or inundation duration has an impact on flood damage, cannot be properly captured with the presented approach. Elaborate on this limitation and the resulting limit of applicability of the approach, which is from my point of view large river floods.

30

Good point — We will state the limitations more clearly in the paper. The large uncertainties introduced by this simplification are, however, accounted for in the uncertainty analysis. The rationale for this model choice is that we wanted to run a very large number of simulations (which requires runtimes far shorter (~1sec per simulation) than any hydraulic model can provide, as far as we can judge) in order to holistically quantify the influence of epistemic and aleatory uncertainties (and therein also dike breaches) and to use this model in subsequent studies, f.e.x in optimization problems. Independent of that, we found your judgement very interesting: We had considered all model components to be in a roughly comparable degree of process representation...

35

40

RC1.3: The epistemic uncertainty analysis is quite encompassing, but it surely does not capture all possible epistemic uncertainty sources. This is hardly ever possible in practice. Please mention this aspect in the discussion. Also, all the epistemic uncertainty sources should be shown with their assumed distributions and impacts on the risk estimate. Since you don't show the final risk assessment and cost-benefit analysis for the river reach, but rather demonstrate the concept, this should be included. The readers can also learn from this about the uncertainty range and the impact of the different uncertainties on the risk assessment. Showing only the Sobol Index is not sufficient from my point of view.

Yes, we agree and struggled with this while writing the paper. On the one hand, the uncertainty and sensitivity analysis deserves a much larger share of the manuscript, since this is one of the main contributions of our work. On the other hand, the description of the flood risk model has already become quite long, and we wanted to avoid further inflating the paper. As a consequence, we chose to illustratively focus on only three epistemic uncertainties and therewith intended to show how the model — and the hierarchical structure — can incorporate and support UA and SA, without going into all modeling choices and background assumptions. I will try to shorten the flood risk model chapter (if possible) and expand the UA section.

RC1.4: The work limits the risk assessment for events with $T > 100$ a. Please explain why this limitation was taken. This is important, because also less extreme / more frequent floods can cause damage (see e.g. DOI: 10.5194/nhess-9-1033-2009)

In our case study, we limited the assessment to scenarios with recurrence intervals > 100 years because the mitigation measure we assess is only activated beyond that threshold. I realize this was not clearly described in the paper: we will add a description of the detention basin to the study area section and emphasize it again in the hydrological load and results sections. If other mitigation measures are evaluated, the catalogue of flood scenarios has to, of course, be extended.

RC1.5: I also miss a short discussion of the proposed model concept with existing fully coupled model chains (e.g. Sairam et al. 2021, where only the probabilistic dike breach modelling is missing, but the models are all the actual hydrological/hydraulic models, i.e. no surrogates), or a notion to the increasing number of AI approaches. Your point is here the fully probabilistic approach with uncertainty and the very quick model chain execution. This should be emphasized, but also set in relation to the limitations introduced by the simplified models you developed.

We agree and will add a short discussion in the end. Despite the more advanced approaches pursued in more recent studies (e.g. fully coupled model chains, like the benchmark study by Sairam et al. 2021), we believe that there are use cases where the development or application of these models is not feasible or, in the case of AI approaches, more useful. The model was developed in a fairly short time frame and with limited data available. Applying more advanced (and thus computationally more demanding) approaches to problem settings like feasibility studies, cost benefit analyses or optimization problems could be, out of our perspective, challenging. For these use cases, a simpler, computationally cheaper model based on surrogates can be advantageous. We will emphasize the model's intended use but also its limitations in the paper.

Review 2

90 The paper presents a flood risk model to enable quick appraisal of flood mitigation measures. The model calculates expected benefits of a measure as a reduction in flood damage with respect to the 'no measure' baseline. The key feature of the model is its computational efficiency, which is obtained through an implementation that allows to pre-calculate a range of systems relationships, so that run time is faster. Such computational efficiency in turn enables a comprehensive quantification of output
95 uncertainty and sensitivity through Monte Carlo simulations. I think the paper is potentially an interesting and valuable contribution, but it needs revision in order to clarify its novelty, scalability and transferability.

General remarks

100 RC2.1: The Introduction does not do enough to clarify the key focus and contribution of the study. The description of the "three key necessary developments" on page 1-2 is very broad and potentially misleading. It raises a wide set of issues but after reading the full paper I think only the first one is addressed in this work. For example, the conclusions of point 2 and 3 that "integration into official practice is still lacking" and "practical implementation often remains fragmented across various
105 professional sectors that focus on local flood protection planning" are certainly true but there is no follow up on them in the remainder of the paper. Point 3 on page 2 talks about "cross-border approaches" and "large spatial scales" but the proposed method is demonstrated here on a 380 km section of the Danube river in Germany, which is neither "cross-border" or necessarily "large scale". The literature review on L. 57 onwards includes studies at very different scales, from national flood risk assessment (in the UK and Germany) to regional studies within Germany or even reach-level
110 studies, for example the 50 km reach of the Po River of Domeneghetti et al. (2013). Hence, throughout the Introduction it is unclear what the paper focus is, in terms of spatial scale. The statement of the study contribution (L.85-90) could be better connected to the previous literature review, so to clarify how exactly the proposed model differs or advance on the methods
115 reviewed in previous lines 57-84. A much clearer statement is given in Sec 3.6: "The developed flood risk model differs from other methodologies, such as those presented by Vorogushyn (2008) and Domeneghetti et al. (2013), primarily in that" and in the Conclusions: "A key element enabling this independence is the developed flood routing method, along with segment-specific lookup tables that store the relationships between inundation elevation, floodwater volume, and damages.". It might be
120 useful to have these or similar statements made earlier on in the Introduction and not left at such later stages into the paper.

The description of the "three key necessary developments" on page 1-2 is very broad and potentially misleading. It raises a wide set of issues but after reading the full paper I think only the first one is addressed in this work. For example, the conclusions of point 2 and 3 that "integration into official
125 practice is still lacking" and "practical implementation often remains fragmented across various professional sectors that focus on local flood protection planning" are certainly true but there is no follow up on them in the remainder of the paper. Point 3 on page 2 talks about "cross-border approaches" and "large spatial scales" but the proposed method is demonstrated here on a 380 km section of the Danube river in Germany, which is neither "cross-border" or necessarily "large scale".
130 The literature review on L. 57 onwards includes studies at very different scales, from national flood risk assessment (in the UK and Germany) to regional studies within Germany or even reach-level studies, for example the 50 km reach of the Po River of Domeneghetti et al. (2013). Hence, throughout the Introduction it is unclear what the paper focus is, in terms of spatial scale. The

statement of the study contribution (L.85-90) could be better connected to the previous literature review, so to clarify how exactly the proposed model differs or advance on the methods reviewed in previous lines 57-84. A much clearer statement is given in Sec 3.6: “The developed flood risk model differs from other methodologies, such as those presented by Vorogushyn (2008) and Domeneghetti et al. (2013), primarily in that” and in the Conclusions: “A key element enabling this independence is the developed flood routing method, along with segment-specific lookup tables that store the relationships between inundation elevation, floodwater volume, and damages.”. It might be useful to have these or similar statements made earlier on in the Introduction and not left at such later stages into the paper.

We acknowledge that the introduction was not sufficiently regarding the actual novelty/contribution of our work. To address this, we propose to rewrite the main paragraph in the introduction, reframe some other points in the introduction and clarify the spatial scale in the conclusions. You find the proposed new introduction below. We do think that our flood risk model can be used to support all three necessary developments that we outlined in the beginning of the introduction. The results we show demonstrate that the model can be used to estimate the flood risk (plus the uncertainties involved), which is necessary for point 1), but that it can also quantify the expected benefit resulting from mitigation measures, which in turn can be used in benefit-cost analyses and decision making contexts, hence point 2). It is very helpful that you pointed out that the relation between point 3) and our model is not really clear – we agree and will amend the manuscript to clarify this point. The strong advantage of our model (in our view) is, besides the ability to perform extensive uncertainty and sensitivity analyses, that it can be applied to large rivers. It is true that the case study is not very large, but we will clarify in the manuscript that the methodology can be applied to rivers which are much longer. Together with that, we will also discuss the question of how transferable the model is to other case studies (easily), and how it could, if applied to cross-border river sections, help to increase system thinking. Furthermore, we will discuss how the computational efficiency of the model can help to identify global optimal mitigation strategies (rather than, what is currently done in practice, implement local strategies on municipal/regional level).

Updated introduction, starting from line 57:

“Indispensable for these three developments are flood-risk models that can represent interactions in flood processes, provide uncertainty-aware risk estimates, and operate at river or catchment scale. In the scientific literature, different models have been presented that meet one or more of these criteria. [summary of literature without changes, lines 58-82]

Overall, this comparative analysis of existing large-scale flood models highlights the necessary trade-off between modeling complexity, spatial extent, and comprehensiveness of the uncertainty analysis. In this work we introduce a probabilistic flood-risk model that seeks an innovative balance between these three objectives. The developed model achieves computational efficiency by replacing runtime coupling of computationally expensive hydraulic and hydrodynamic simulations with pre-processed surrogates and lookup tables, while still dynamically coupling flood-process components (Fig. 2). This allows the model to represent the key process interactions, without the prohibitive runtime associated with coupled simulation models. Crucially, the model explicitly captures downstream propagation of impacts following dike failures or activation of mitigation measures via an efficient flood-routing scheme. Computational efficiency supports embedding the risk model within a two-level Monte-Carlo framework: aleatory and epistemic uncertainties are introduced and

180 propagated separately, thereby enabling a more comprehensive uncertainty and sensitivity analysis than is typically feasible at river scales.

The reduction in modeling accuracy compared with flood risk models coupled with high-resolution hydraulic simulations is, as we argue, acceptable for many decision making contexts. Moreover, the resulting uncertainty is explicitly quantified within the uncertainty framework. The model is applied to the Bavarian Danube to evaluate the benefit of flood mitigation measures. The successful application in the case study proves that the model can (i) deliver a large-scale probabilistic risk assessment, (ii) support benefit-cost based decision making in practice by providing uncertainty-aware estimates of risk reduction, and (iii) support the identification of jointly optimized mitigation portfolios at river scale by explicitly modeling interactions of mitigation measures and downstream effects.

The remainder of the paper is organized as follows. Section 2 introduces the study area; Section 3 outlines the flood risk model components; Section 4 describes the two-level Monte-Carlo framework; and Section 5 presents results from the application of the model to the study area. The paper concludes with a final discussion and remarks in Section 6.”

RC2.2: Section 3 is very long and it is difficult to grasp which parts are a description of approaches that have been previously used and published by others, and which are the truly novel steps. Considering that the key contribution is the way the model is implemented (=with pre-computed look-up tables instead of dynamic coupling at runtime) it seems strange that “Sec. 3.6 - Model Implementation” takes only 8 lines out of a 12-pages “Sec 3 – Flood risk model” section! Indeed, novel ideas on how to perform pre-calculations are also spread out across Secs 3.1-3.5, but they are less obvious to find out. In summary, I suggest re-organising materials in this Section so that the key differences from previous studies and novel ideas are easier to identify.

We carefully reconsidered the paper’s organization and concluded that aggregating all novelties in a single section while placing the remaining model components elsewhere does not serve the presentation of the developed model well. We do not believe that a paper must always follow that structure, particularly when the aim is to present a new flood-risk model in which many components are at least partly novel or innovative. As you noted, it is not only the routing mechanism that is new, but rather a combination of different modelling approaches. Our intention with this manuscript is to provide a comprehensive overview of the model for researchers and practitioners who need a computationally inexpensive yet effective flood-risk framework that does not depend on coupling with external simulation software. For this target audience, it is important to understand the model configuration and the interactions among modules, rather than to itemize which elements are novel. Since the individual model components are described in detail in Sections 3.1–3.5, an extensive implementation description is not necessary. Moreover, aside from the eight lines you highlight, the principal content of this subsection is Figure 10, which we consider highly informative.

RC2.3: The transferability and scalability of the proposed approach should be discussed. Could the approach sketched out in Figure 10 be applied to the entire Danube river or Germany-wide? What would be the challenges in terms of data availability, software implementation, run-time, etc.? Could the approach be applied in other regions of the world with different climatic and landscape

characteristics? What are the critical assumptions underpinning the approach that may hinder its application elsewhere?

We fully accept this point. We added a new paragraph to the discussion that explicitly addresses these issues and more clearly emphasizes the model's limitations. The new paragraph is inserted at line 519:

"The flood risk model is applied to the 380 km of Bavarian Danube in Germany to assess the potential risk reduction of mitigation measures. Computational time scales approximately linearly with the number of dike segments; an extension of the model to the entire Danube, which spans roughly 2,850 km across ten countries, is thus expected to increase runtime by no more than a factor of ten. On this basis, we argue that the model can be applied in large-scale studies and support increasing system thinking. More critical for large cross-border applications are the pre-processing steps, which entail acquiring the necessary data, models and large-scale simulation outputs from multiple stakeholders. We acknowledge that this can be challenging in practice, but this challenge will be faced by any modeling approach. The extent to which uncertainties propagate and accumulate downstream over such long river stretches would require careful examination. Applying the model to a new study area follows a similar process, provided that region-specific climatic, hydrological and hydraulic simulation models are available. The pre-processing steps are repeated to capture relevant regional characteristics.

To further support system thinking, the model can be extended to include the entire catchment, rather than representing only the main river. Although tributary contributions are accounted for in the flood scenarios, potential protection failures, local mitigation measures, and other dynamic effects along the tributaries are not explicitly modeled; this limitation should be addressed in future work. Explicitly modeling tributaries in the same manner as the main river is straightforward and, if implemented with efficient parallelization, would not substantially increase computational cost. Such an extension would enable catchment-wide optimization of mitigation measures and improve the accuracy of risk estimates.

A further limitation of the model is due to the hydrological simulation of catchment runoff, which is performed during pre-processing in the hydrological load module. While the implemented design choice permits short run times of the main flood simulation module, it does not allow the assessment of interventions influencing the hydrological flood-forming processes. Consequently, many nature-based solutions cannot be evaluated using the present model implementation.

Notwithstanding, the developed model provides a solid foundation for decision-making, such as in benefit-cost analyses, which is demonstrated in a practical example at the case study. Therein, the benefit of a flood mitigation measure, along with its associated uncertainties, is quantified. In the future, the model can support assessments of various mitigation strategies, help identify high-risk areas, or optimize the combination of mitigation measures. In this way, we demonstrate how dynamic, probabilistic, and integrated flood risk management can be implemented in practice. We hope that our work encourages and supports decision-makers in moving toward a more holistic approach to flood risk management."

RC2.4: I am also confused by the organisation of materials through sections 3 and 5. “Sec. 5 - Results” only includes Figures of the aggregate loss estimates and their uncertainty and sensitivity to three selected parameters. This draws the reader’s attention to the UA&SA part of the paper, which is the least “interesting” as there is no novel contribution here (all the definitions and methods used for the UA&SA have been published before, as also acknowledged in Sec. 4). If the paper’s key contribution is the computationally-efficient implementation of the flood risk model, I would have expected the Results to mainly report on the behaviour of the model, and how its predictions compare to those of a dynamically coupled but more computationally expensive model - maybe at different river sections and under different scenarios with or without a dyke breaches. I suspect Fig. 7 and. 8 partially do this, but oddly they are embedded in Section 3 (“Flood risk model”), which somehow becomes both a description of the method and (implicitly) results section. In summary, I would suggest a better organisation of the materials and expanding results that demonstrate the “success” of the proposed approach for efficient implementation of flow routing and flood risk calculations, and less on the UA&SA results – where the authors apply standard methods instead of suggesting new ones.

We would welcome a comparison of our results with those from more sophisticated flood-risk models; however, developing such a model for the sole purpose of comparison is beyond the scope of the present study. While it is not possible to compare the entire model as a whole, we have evaluated all model components against hydraulic or hydrodynamic simulations, available post-event data, or other established models (some of which are presented in Section 3). With respect to uncertainty and sensitivity analysis, we believe that conducting an in-depth, two-level UA/SA represents a substantive extension beyond much of the existing literature, which typically performs UA/SA but not within a two-level framework. We agree that the Results section should be revised to more clearly demonstrate the model’s strengths and capabilities; we will make these revisions in the final manuscript. Your detailed comments will also be incorporated into the revised version.

300 **Dear Reviewers,**

Thank you for your detailed and highly valuable feedback. We have carefully revised the manuscript in response to your comments. Point-by-point responses were already provided in the comment responses (pasted above), where we addressed each remark in detail. In the following, we only
305 summarize the main revisions undertaken based on your review comments (RC).

RC1.1 and RC2.1

We revised the introduction to more clearly highlight the novelty and specific contribution of our work, as well as to sharpen the overall focus of the study.

310

RC1.2, RC2.3, RC1.5

We revised the final section of the manuscript and substantially expanded the discussion on model limitations, transferability, and scalability.

315 RC1.3, RC2.4

We note that your comments regarding the presentation of the uncertainty analysis pointed in different directions. We decided to revise the uncertainty analysis and results section, clarifying and consistently presenting the parameterization of all considered uncertainty sources. While the applied methods themselves are not novel, we believe that their transparent presentation improves the
320 overall understanding and interpretability of the developed model.

RC1.4

We addressed this comment throughout the manuscript, including the study area description, the results section, and other relevant parts where implications arose.

325

RC2.2

We carefully reconsidered the overall structure of the manuscript. We concluded that grouping all novel aspects into a single section, while separating other model components, would not improve the clarity of the presentation. Instead, we revised the introduction to more clearly communicate the key
330 ideas and contributions of the study. We hope that now through the revised introduction, the overall structure of the paper provides a more coherent presentation of the developed model.

In addition, we have incorporated your detailed comments throughout the manuscript.

335 We hope that our revisions adequately address your concerns and we appreciate the opportunity to improve the manuscript. We look forward to your further feedback.