Authors: We do sincerely appreciate the detailed and very good comments from the reviewers', and we have acknowledged their contribution in improving the quality of this paper. We hope that we have now addressed all the reviewers' comments satisfactorily in the revised manuscript.

After recommendation from other reviewer, we have updated the title

"Buildings hazard maps with differentiated risk perception for flood impact assessment"

Please find our point-by-point responses to the comments

Reviewer: General comments:

- The paper is very interesting and the main objective to create hazard maps or doing impact assessment with a transparent declaration and consideration of uncertainties is desirable. Especially the presented approach using confidence intervals of a hydrological forecast ensemble is interesting and has potential. However, there are doubts about some major points in this study.
- It is questionable whether the approach presented in this study "inherently communicates the underlying uncertainties", as stated in the conclusion. Looking at the final map presented in figure 9a), the meaning of Scenario I, II and III is not directly derivable and the coherence of all classifications and the different combinations of hydrograph scenarios with building types is not easily comprehensible. Even if the names of scenarios in the map would be changed to "high exceedance probability" (=S I), "average exceedance probability" (=S II) and "low exceedance probability" (=S III) – what would already improve the understanding - it's still questionable whether the multi-model combination presented is the right way to deal with uncertainties in hydrological forecasting for flood impact assessment. To combine the exposure for different confidence interval hydrographs in a new scenario defined by the same exceedance probability is not very elegant. According to the output of the ensemble members, the M50% confidence interval hydrograph as used in the paper (= best-model = median) is the hydrograph that is forecasted as the most likely one and therefore, to define scenarios with low, average or high exceedance probabilities based on multi-model combinations referring to different confidence intervals is misleading. I try to exemplify this issue on table 3: The way the M%-hydrographs from the ensemble forecast are used would imply that the higher the confidence interval, the lower the exceedance probability of an event. But this is not the correct way to implement the confidence intervals here. At the time of the forecast and according to the model ensemble, it is as likely that a M25% or a M75% (= 50% confidence interval) event appears (when I

Authors: Thank you for pointing out the error, as suggested we have replaced the term percentile in the revised manuscript and corrected as in Figure 1 (Figure 3 in Manuscript). We also agree that the exceedance probability is not the right choice of word to describe the scenarios, hence we have replaced it by "risk perception" and have added the scenario description in Figure 2 (Figure 9a in manuscript).

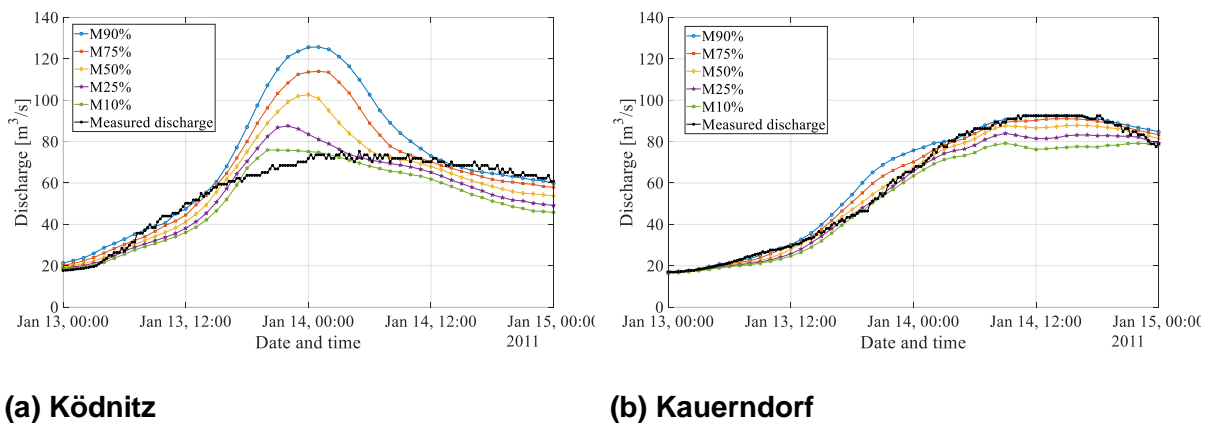**(a) Ködnitz**   **(b) Kauerndorf**

Figure 1. Hindcasted flood event of January 2011: measured discharge hydrograph along with 95%, 90% and 75% percentile discharges for gauges a) Ködnitz and b) Kauerndorf (Data based on Beg et al., 2018).
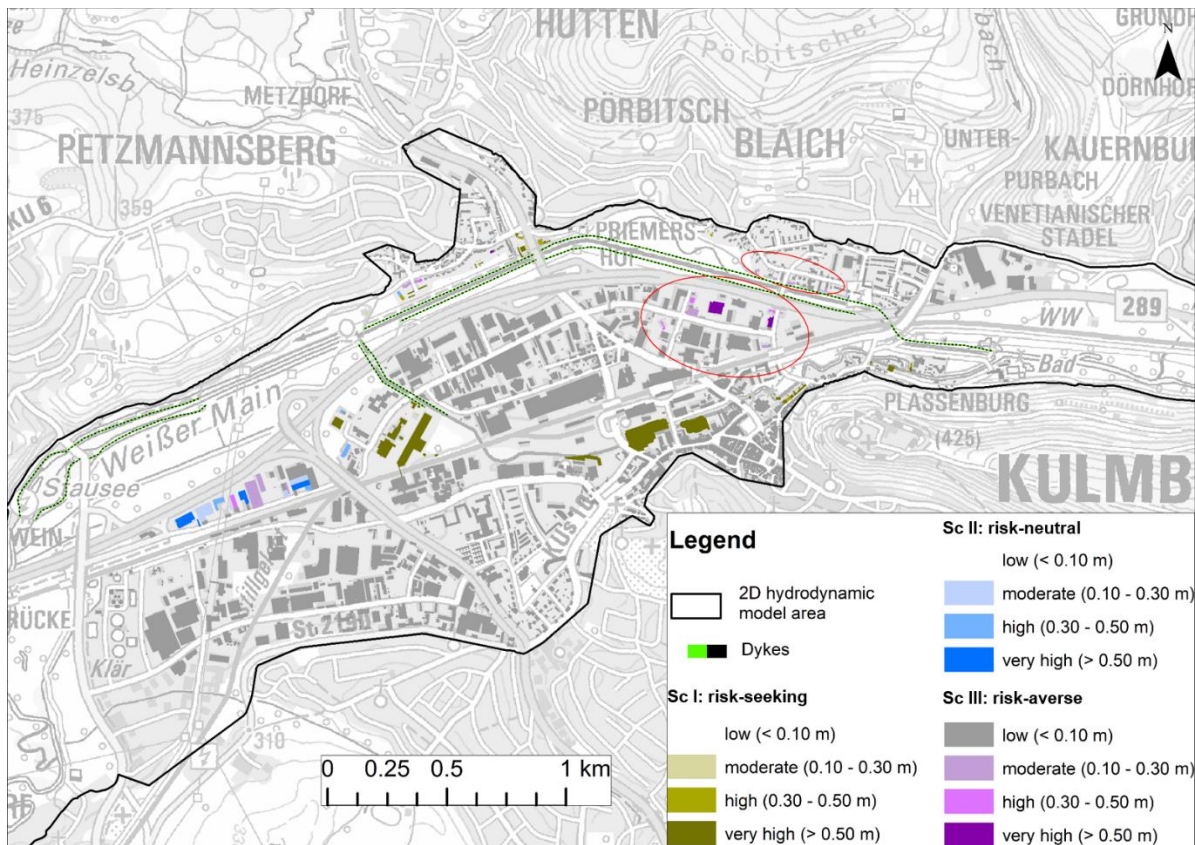
Figure 2: Multi-model combination

In addition, we stated that these maps offer an alternative way to communicate the underlying uncertainties. It will be interesting for authors to measure how the scientific community and practitioners will receive the methodology. In conclusions, Page 22: Line 21-24

"In summary, we have presented a new methodology for flood impact assessment using a multi-model combination in the form of *building hazard maps with differentiated risk perceptions.* These maps offer an alternative way to communicate the underlying uncertainties in forecasting models and are ready-to-use for decision-makers in the field of flood risk management.".

- In addition, I don't see the a fundamental novelty of this approach compared to the cited study by Zarzar et al. (2018), who already presents the use of a multi-model ensemble framework based on hydrological ensemble members for visualizing flood inundation uncertainty. I agree with the authors that the use of confidence intervals in the case of a high number of hydrological forecast members can support a transparent declaration of forecast uncertainties.

Authors: We have now differentiates the work done by Zarzar et al. (2018), they have used an average of HD model raster outputs to obtain the percentage of ensemble agreement. Whereas, we have exported HD models on buildings as urban features and proposed a new methodology of multi-model combination. In Introduction, Page 3: Line 4-7

"Zarzar et al. (2018) have used a multi-model combination framework consisting of hydro-metrological and HD models to visualise flood inundation uncertainties in which they have

used an average of HD model raster outputs to obtain the percentage of ensemble agreement."

- Besides all of this, to my point of view the display of the inundation pathways and extend is necessary for a use in early warning systems, emergency planning or flood impact assessment. The approach of a building hazard map doesn't provide enough details to intervene (where does the water come in contact with the building, where are the "weak spots" in the river- and floodplain-system).

Authors: We acknowledge the importance of the inundation extent; we have proposed to use hazards maps in addition to other layers of information, such as inundation map for planning. In conclusion, Page 23-Line 12-16

"Finally, the output of the framework can be extended to hazard maps uploaded in a web-based GIS system to improve visualization, along with providing layers of additional information, such as inundation pathways and weak spots in the river and floodplains to provide sufficient details to intervene (aid in planning). This additional information would enhance the usefulness to different target users, such as planners, decision-makers and flood forecasting agencies."

Reviewer: In addition to these major points, there are some further remarks in the specific comments.

Specific comments

- Section 2.1, 2.2 + supplement tables: It is mentioned that the particular focus of this study is on the development of the post-processing part (classification and multi-model combination, mapping). Therefore, the two parts of the framework that are already developed and explained similarly in Beg et al. (2018) and Bhola et al. (2018a, 2018b) should be shortened, as for example a detailed explanation of the model LARSIM, FloodEvac and HEC-RAS is not needed to understand the context of this study. Nevertheless, Figure 3 helps to understand the setup of the multi-model approach. It should be better explained, where for example the M10% hydrograph can be found in such a graphic (consistent use of confidence intervals). It would be helpful if you could add the forecast ensemble for the virtual station that was used as upper boundary condition and maybe mark the used hydrographs.

Authors: We agree with the reviewer comment regarding shortening the details, however, additional details are provided in order to reproduce the methodology, as suggested by the editor.

In the revised manuscript, we have added data labels in Figure 1 to clearly present the percentiles. These two stations where used as the input boundary conditions upstream of the model. The virtual gauge is nothing but the addition of these two hydrographs, we present here the data at gauges upstream of the city.

- Section 2.3: The main literature cited (Krieger et al., 2017) is specific to Germany and is not peer-reviewed. In which way are the classification methods for buildings and hazard types comparable with international, peer-reviewed literature? E. g. Jonkman et al. (2008), Dutta et al. (2003), as well as Thieken et al. (2008) and Kreibich et al.

Authors: Thank you for providing the studies, we have compared their classification to ours in the revised manuscript. We also stated further break down of class III in conclusion. In Section 2.3.1, Page 8: Line 8-18

"There are various classifications of land use features available in the literature. Dutta et al. (2003) have used direct and indirect damages as the basis of their classification and classified their study area in residential and non-residential categories. Jonkman et al. (2008) have classified urban features in residential, businesses, commercial and public property and agricultural to estimate flood loss. Furthermore, the vulnerability was the basis of classification in residential (Thieken et al., 2008) and industrial and commercial sectors (Kreibich et al., 2010) in order to estimate flood losses. We have used the damage potential of a building as a basis for classification in order to focus on the flood impact assessment. Building damage potential is required for a variety of flood mitigation planning activities including flood damage assessment, multi-hazard analyses and emergency measures (Shultz, 2017). The buildings were classified into four classes based on their function following the recommendation of the German standard for risk management in urban areas in the case of flash floods (Krieger et al., 2017)"

And, in Conclusion, Page 23: Line 11-12

"In addition, other classification methods for buildings and hazard types should be evaluated, especially to further dissect the impact of class III in commercial and industrial."

Authors: Thank you for pointing out the error, we have corrected in the revised manuscript. The building class I belongs to buildings that are in open green area such as small park and garden building. In Section 2.3.1, Page 8: Line 29-30

"There are a total of 2695 buildings in Figure of which 1, 958, 1716 and 20 buildings were classified in classes I, II, III and IV respectively."

Regarding the building classification, Page 8: Line 15-18

"The buildings were classified into four classes based on their function following the recommendation of the German standard for risk management in urban areas in the case of flash floods (Krieger et al., 2017)."

We acknowledge that depending on the aim the classification can be adjusted. Page 8: Line 21-22

"In any case, we acknowledge that the number of classes or criteria can be changed/adapted depending on the aim of the forecast."

We have also added in Page 8: Line 31-32

"The nature of the data in this case study leads to a differentiated representation of the classes. It should be noted that the classification aims at creating classes based on damage potential, and not on generating clusters with similar sizes."

- Figure 8 + 9, p. 15 line 1 - 3: In this context, you should maybe explain somewhere, how you assign the hazard for the same building (finally in figure 9) that is in a moderate hazard zone for scenario I but then rise to a very high hazard in scenario III (is the potential in the forecasted event to have very high hazard at a particular building somehow considered?).

Authors: Thank you for your comment. This was done based on Table 1, in Results Section 3.2, Page 16: Line 4-11

"The main objective of the combination is to differentiate the impact of water depths on building classes. Therefore, to design the combinations, a high percentile was assigned to the buildings with a high damage potential class. Each scenario presents a given risk perception of a decision-maker that can be easily adjusted depending on the perception of different stakeholders. Hence, different risk perceptions will lead to different exceedance probability maps. A risk-averse person will likely select low exceedance probabilities for the hazard maps, whereas a risk-seeking person will likely develop risk management strategies based on higher exceedance probabilities. The hazard maps for the three scenarios are shown in **Fehler! Verweisquelle konnte nicht gefunden werden.**."

Table 1. Scenarios of multi-model combinations based on risk perception.

Scenario	Risk perception	Building class			
		I	II	III	IV
I	Risk-seeking	$M_{10\%}$	$M_{10\%}$	$M_{25\%}$	$M_{50\%}$
II	Risk-neutral	$M_{10\%}$	$M_{25\%}$	$M_{50\%}$	$M_{75\%}$
III	Risk-averse	$M_{25\%}$	$M_{50\%}$	$M_{75\%}$	$M_{90\%}$

- "p. 16, lines 4 – 11: In this part, the time issue in real-time assessment of the framework is discussed. As well in the introduction it is mentioned, that flood forecasts might be restricted to computational time of the models. Please provide information about the lead- and run-time of the hydrological forecast model, the computational time for the HD model with each hydrograph and, therefore, how much time would be left for

Authors: We have provided information on the run-time of the entire map production. In Discussion, Page 22: Line 3-7

"In this study, a 50-member ensemble forecast was used from Beg et al. (2018) where the entire process took 25 minutes in a 3-core desktop in parallel mode to generate a forecast of 12 hours. Various percentile discharges were then run simultaneously in the HD model, which required 30 minutes to simulate a 12-hour event on an 8-core, 2.4 GHz (Intel E5-2665), including the initial start (Bhola et al. 2018a). Post-processing of the model results would consume an additional 15 min. Therefore, real-time hazard maps are delivered to decision-makers in 70 min.

Technical corrections

- In general: As already mentioned, the use of the term confidence interval discharge in combination with the M% HD outputs seems not to be correct. Please replace confidence interval with "percentile" when referring to M%-levels – or change these levels accordingly.

Authors: The term is changed to percentile in the revised manuscript.

- p. 1, line 21: It would maybe help if you would explain a bit more in detail, what in this study is meant as multi-model combination. Based on this first explanation, one would assume that the building classification is not part of the multi-model combination and this consists only of the HD and hydrological model (which, according to figure 1, is not the case). The term is also used in various ways: multi-model, multi-model combination, multi-model ensemble combination, multi-model combination scenarios. E.g. in the first sentence of the discussion section, the context is given only to hydrology. If this is the case, then figure 1 should be adapted. This can be a bit confusing.

Authors: We have used multi-model combination consistently in the revised manuscript. The term is changed to percentile in the revised manuscript. Building classification is part of the entire framework but independent of multi-model combinations. We have focused this study on the multi-model combination. In Methodology, Page 3: Line 23-24.

"The particular focus of this study is on the development of the framework of a multi-model combination in the post-processing component"

- p. 8, line 17 + 18: As mentioned, check the numbers for each class and compare with figure and other statements.

Authors: Thank you for pointing it out, we confirm that the numbers provided are correct in revised manuscript.

- p. 11, line 11/12: As mentioned in special comments, to my point of view the confidence intervals cannot be used to describe exceedance probabilities in the way it was done here.

Authors: We have corrected the term and used percentile consistently.

- p. 11, line 22: You refer to post-event information that "no serious damage was reported" -> p. 14 line 4: "figure 7c is in agreement with the post-event information" -> 7c = M50% discharge scenario. According to figure 6, in this scenario 126 buildings are exposed, 67 classified with high hazard -> How does that fit to "no serious damage was reported"?

Authors: This is explain in the discussion. The high hazard at $M_{50\%}$ is due to their proximity to the Mühl canal. In Discussion, Page 18: Line 8-14

"It should be noted that few buildings show *very high* hazards due to their proximity to the Mühl canal (Figure7**Fehler! Verweisquelle konnte nicht gefunden werden.**a). Even though there was no over-topping of water from the canal, because of buildings geolocation being near to the canal, these were assigned automatically with the highest hazard, starting with a discharge of $M_{10\%}$. Ideally, this should be prevented by removing the river channel elements from the dataset before to assign the water depths to the buildings as in Bermúdez and Zischg (2018). However, and without retract to our conclusions, it was decided not to include it in this work in order to keep the automation process simple."

- Figure 7, 8 and 9 b-d: Legends (and building numbers in 9b) are not readable. Also the maps themselves are at the edge of readability. The red circles in figure 9 are not in the legend and are never mentioned in the text (see comments below)? Is figure 9d really necessary?

Authors: The figures have been resized so the legends are readable, in addition all the figures are provided to the editorial so the final print will be of high-quality and readable. We have omitted figure 9d from the revised manuscript. The red circles are defined in the figure caption as well as in the text.

- p. 13, line 12 – 15: This error should be eliminated by either using water depth derived from water surface elevation [m a. s. l.] minus ground level of building [m a. s. l.] or by removing the river channel elements from the dataset you used to assign to the buildings. Check out Bermúdez and Zischg (2018).

Authors: Thank you for your comment. We have added this limitation of our study to the article. In Discussion, Page 18: Line 11-14.

"Ideally, this should be prevented by removing the river channel elements from the dataset before to assign the water depths to the buildings as in Bermúdez and Zischg (2018). However, and without retract to our conclusions, it was decided not to include it in this work in order to keep the automation process simple."

- p. 14, line 17 – 22: Please add the information that you in this specific case compare the red circles from M50% and the multi-model map. ID 1393 is not recognizable in figure 9d. p. 16. Line 3 + 4: What is the computational time of the whole framework? What's the lead-time used here?

Authors: We have added the information and stated the computational time of whole framework. Please see in specific comments.

Publication bibliography

Thank you for providing further literature in helping the quality of te paper, we have added the recommended citations

- Bermúdez, María; Zischg, Andreas Paul (2018): Sensitivity of flood loss estimates to building representation and flow depth attribution methods in micro-scale flood modelling. In Nat Hazards 92 (3), pp. 1633–1648. DOI: 10.1007/s11069-018-3270-7.
- Dutta, Dushmanta; Herath, Srikantha; Musiake, Katumi (2003): A mathematical model for flood loss estimation. In Journal of Hydrology 277 (1-2), pp. 24–49. DOI: 10.1016/S0022-1694(03)00084-2.
- Jonkman, S. N.; Bockarjova, M.; Kok, M.; ˇBernardini, P. (2008): Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. In Ecological Economics 66 (1), pp. 77–90. DOI: 10.1016/j.ecolecon.2007.12.022.
- Kreibich, Heidi; Seifert, Isabel; Merz, Bruno; Thieken, Annegret H. (2010): Development of FLEMOcs – a new model for the estimation of flood losses in the commercial sector. In Hydrological Sciences Journal 55 (8), pp. 1302–1314. DOI: 10.1080/02626667.2010.529815.
- Thieken, A. H.; Olschewski, A.; Kreibich, H.; Kobsch, S.; Merz, B. (2008): Development and evaluation of FLEMOps – a new F lood L oss E stimation MO del for the private s ector. In D. Proverbs, C. A. Brebbia, E. Penning-Rowsell (Eds.): Flood Recovery, Innovation and Response I. FLOOD RECOVERY, INNOVATION AND RESPONSE 2008. London, England, 02.07.2008 - 03.07.2008. Southampton, UK: WIT Press, pp.315–324.

Authors:

We appreciate the constructive feedback from the reviewer in improving the quality of the manuscript. We have addressed all the review comments in the revised manuscript.

After revision of reviews 1, we have changed the term "confidence interval" to percentile and "exceedance probability" to risk perception.

The followings are our point-by-point responses to the reviewer's quotes

Authors: We have updated the title

"Buildings hazard maps with differentiated risk perception for flood impact assessment"

Authors: We have omitted the term from methodology. The real-time refers to producing the maps as the discharge are forecasted. We have provided more provided information on the run-time of the entire map production. We have provided information on the run-time of the entire map production. In Discussion, Page 22: Line 3-7.

> "In this study, a 50-member ensemble forecast was used from Beg et al. (2018) where the entire process took 25 minutes in a 3-core desktop in parallel mode to generate a forecast of 12 hours. Various percentile discharges were then run simultaneously in the HD model, which required 30 minutes to simulate a 12-hour event on an 8-core, 2.4 GHz (Intel E5-2665), including the initial start (Bhola et al. 2018a). Post-processing of the model results would consume an additional 15 min. Therefore, real-time hazard maps are delivered to decision-makers in 70 min."

Authors: We added references to other classification in literature.

In methodology section 2.3.1, Page 8: Line 8-18

> "There are various classifications of land use features available in the literature. Dutta et al. (2003) have used direct and indirect damages as the basis of their classification and classified their study area in residential and non-residential categories. Jonkman et al. (2008) have classified urban features in residential, businesses, commercial and public property and agricultural to estimate flood loss. Furthermore, the vulnerability was the basis of classification in residential (Thieken et al., 2008) and industrial and commercial sectors (Kreibich et al., 2010) in order to estimate flood losses. We have used the damage potential of a building as a basis for classification in order to focus on the flood impact assessment. Building damage potential is required for a variety of flood mitigation planning activities including flood damage assessment, multi-hazard analyses and emergency measures (Shultz, 2017). The buildings were classified into four classes based on their function following the recommendation of the German standard for risk management in urban areas in the case of flash floods (Krieger et al., 2017)."

And we added also in Page 8: Line 20-22

> "In the authors' opinion by keeping our classification simple will likely fit a vast majority of cities regardless of their size. In any case, we acknowledge that the number of classes or criteria can be changed/adapted depending on the aim of the forecast."

For applicability in other study areas, In discussion, Page 22: Line 1-2

> "Overall, the methodology is independent of the choice of models, i.e. hydrological and HD, and is transferable to other study areas."

And In conclusion, Page 23: Line 4-8

> "Further research investigating multi-model combinations and validation in other study areas may be beneficial. A more extensive study on the validation of the multi-model combination may be required, possibly by using measuring gauges, post-event survey (as conducted in Thieken et al., 2005), satellite images (as in Triglav-Čekada and Radovan, 2013), and/or crowd-sourced data (Bhola et al., 2018b)."

Authors: We have added the information on handling the buildings and assigning hazard. In 2.2 Hydrodynamic modelling Page 7: Line21-23

> "The buildings are explicitly included using their shape in the mesh and are excluded from the flow calculation by assigning a high roughness value. To assign hazard to a building, the maximum water depth of all the neighbouring cells was used."

Authors: Thank you for pointing it out, we have changed the color so they are visible in Figure 3.

Figure 3. City of Kulmbach and building use classification. (Data source: Bavarian Ministry of the Interior, for Building and Transport and Water Management Authority Hof).

Authors: The classification is general and can be applied in any other city. In this study area there is no metro but there are class IV buildings.

Authors: We have replaces = by -> in Figure 4

Figure 4: An example of a multi-model combination in which the four building classes I, II, III and IV are assigned to the 2D HD model results of 25%, 50%, 75% and 90% respectively.

Authors: Thank you for the comment. The following paragraph was added/changed in the discussion, Page 22: Line 13-19.

> "Therefore, our methodology would allow the target users to benefit from hazard maps enabling them to better prioritise and coordinate evacuation planning based on the stakeholder perception to risk. The maps could further serve as a tool for flood risk assessment. The methodology can be used for flood mitigation and flood forecast planning in the form of emergency management training, where forecasted hazard scenarios can be presented to the training groups. By visualising inundation scenarios, potential damage at the building's level which has been prioritized based on the desired classification can be estimated with this methodology and made available together with each forecasted scenario."

And in outlook, Page 23: Line 8-17

> "In future, damage potential classification can further be improved by including additional criteria, such as population density or water quality, and with it extend the applicability of this method. For example, the assessment of the damage potential of commercial enterprises, substances or machinery containing elements that which could be a source of water pollution could be included (Krieger et al., 2017). In addition, other classification methods for buildings and hazard types should be evaluated, especially to further dissect the impact of class III in commercial and industrial. Finally, the output of the framework can be extended to hazard maps uploaded in a web-based GIS

system to improve visualization, along with providing layers of additional information, such as inundation pathways and weak spots in the river and floodplains to provide sufficient details to intervene (aid in planning). This additional information would enhance the usefulness to different target users, such as planners, decision-makers and flood forecasting agencies."

Buildings ~~hH~~azard maps with differentiated ~~exceedance probability~~risk perception for flood impact assessment

Punit K. Bhola[1], Jorge Leandro[1], Markus Disse[1]

[1]Chair of Hydrology and River Basin Management, Department of Civil, Geo and Environmental Engineering, Technical University of Munich, Arcisstrasse 21 80333 Munich, Germany

Correspondence to: Punit K. Bhola (punit.bhola@tum.de)

Abstract. In operational flood risk management, a single *best-model* is used to assess the impact of flooding, which might misrepresent uncertainties in the modelling process. We have used quantified uncertainties in flood forecasting to generate flood hazards maps that were combined based on different ~~exceedance probability~~risk perception scenarios. The purpose is to differentiate the impact of flooding depending on the building use~~,~~ enabling ~~therefore~~, therefore, more flexibility for ~~stakeholder~~stakeholder's ~~ss~~ variable risk-averse and risk-seeking profiles. The aim of the study is thus to develop a novel methodology that uses a multi-model combination of flood forecasting models to generate flood hazard maps with differentiated ~~exceedance probability~~risk perception. These maps take into account uncertainties steaming from the rainfall-runoff generation process and could be used by decision-makers for a variety of purposes in which the building use plays a significant role, e.g. flood impact assessment, spatial planning, early warning and emergency planning.

1 Introduction

Floods are one of the most destructive natural hazards and lead to severe social and economic impacts (European Union, 2007; Alfieri et al., 2016). The number of people exposed to recent flooding occurred in many Central European countries highlights the importance of assessing flood hazards. During the extensive June 2013 floods in Germany, for example, more than 80,000 people in eight federal states had to be evacuated (Thieken et al., 2016). The vulnerability of settlements calls for an improved flood forecasting, which includes underlying uncertainties and impacts.

In this study, we present a novel methodology that uses a multi-model combination of two-dimensional (2D) hydrodynamic (HD) models to assess the impact of flooding based on water depth hazards. These hazards can be evaluated for key urban features, such as buildings, roads, bridges and green spaces (Leandro et al., 2016). This study focusses in particular on buildings. Furthermore, the hazard maps serve a variety of purposes, e.g. flood impact assessment, spatial planning, early warning and emergency planning (Hammond et al., 2013) for target users. For this paper, the users consist of a group of decision-makers, such as the Bavarian Water Authorities and disaster relief organizations in Germany, the Federal Agency for Technical Relief or the German Red Cross.

In deterministic flood forecasting, the predictions of forecasting models, precipitation forecasts, hydrological models and HD models, are used to generate flood hazard maps. These maps form the basis of flood risk management and are utilised to assess the impact of floods (Schanze, 2006; Hagemeier-Klose and Wagner, 2009). Although advances are continually being made in real-time forecasting, they are still inherently uncertain (Meyer et al., 2009; Bates et. al., 2014; Beven et al., 2018). The decision-making process based on uncertain predictions can have a huge economic impact and possibly lead to life and death situations (Leedal et al., 2010). Thus, a thorough assessment is required of the extent to which uncertainties affects the predictions of flood hazards. In addition, forecasting predictions that inform policy or risk management decisions should include major sources of uncertainty and communicate them coherently (Todini, 2017).

Researchers have addressed various sources of uncertainties in flood modelling, such as precipitation measurements, spatial interpolation of the precipitation, model parameter, model structure (Nester et al., 2012; Leandro et al., 2013), discharge data, measured discharge and uncertainty estimation techniques (Dotto et al., 2012). Although uncertainties arising from precipitation and HD models are significant, the generation of discharges using a hydrological model is considered as one of the most uncertain steps in flood forecasting (Di Baldassarre and Montanari, 2009). Substantial research has been dedicated to the field of discharge forecasting and reducing uncertainties by using methods, such as Generalized Likelihood Uncertainty Estimation (Beven and Binley, 2014), Global Sensitivity Analyses (Pappenberger et al., 2008) and the Shuffled Complex Evolution Metropolis Algorithm (Dotto et al., 2012). To find the appropriate method, Pappenberger et al. (2006) have provided a decision tree that helps users select a suitable method for a given solution. Furthermore, in a recent study Boelee et al. (2018) reviewed uncertainty quantification methods to provide practitioners with an overview of ensemble modelling techniques. An overview of existing ensemble forecasts in operational use can be found in Cloke and Pappenberger (2009) and Todini (2017).

Most notably, in the federal states of Rhineland-Palatinate (Bartels et al., 2017) and Bavaria (Laurent et al., 2010) discharge ensembles is are generated using the COSMO-DE-EPS precipitation ensemble as input to a distributed hydrological model LARSIM (Large Area Runoff Simulation Model). These and similar developments offer a potential framework for quantifying uncertainties. A challenging issue in natural hazards, however, remains the effective communication of the quantified uncertainties to decision-makers (Doyle et al., 2019). Researchers have questioned how uncertainties should be communicated to reduce the risk of wrong or inappropriate decisions (Bruen et al., 2010; Todini, 2017).

In operational flood forecasting, hazard maps are provided in the form of exceedance probability scenarios and generally, only one scenario is considered for emergency planning. Normally, a 50% exceedance probability scenario (or median) is expected to be close to the deterministic *best-model* approach (Di Baldassarre et al., 2010). In other examples (Beven et al., 2014; Beven et al., 2015; Disse et al., 2018), model results of various exceedance probabilities are provided on separate or combined maps. Kolen et al. (2019) stated that there is a need for new methodologies that employ a multi-model combination approach by including several scenarios for improving decision making. A multi-model combination is based on the results of several models and creates a more robust forecasting system with a better representation of uncertainties (Kauffeldt et al., 2016). Although the multi-model combination approach has been used widely in the field of discharge forecasting (Shamseldin et al., 1997; See and Openshaw, 2000; Oudin et al., 2006; Weigel et al., 2008), the approach is not commonly used in the field of

real-time flood hazard forecasting. The high-computational time required by the HD models restricts the use of such an approach in real-time forecasting. However, the use of a simple model structure and/or high-performance computing makes it possible to simulate HD models in real-time; thus, making it feasible to use multi-model combination approaches. Zarzar et al. (2018) have used a multi-model ~~ensemble~~ combination framework consisting of hydro-metrological and ~~hydraulic~~ HD

5   models to visualise flood inundation uncertainties in which they have used an average of HD model raster outputs to obtain the percentage of ensemble agreement.~~.~~

We develop a methodology for obtaining a multi-model combination as an effective alternative to traditional *best-model* approach for producing detailed hazard maps, which are termed as *building hazard maps*. This term can be defined as a map

10   that highlights buildings that are affected by or are vulnerable to flooding with differentiated ~~exceedance probabilities~~risk perceptions (Paulsen et al., 2012)~~-~~ of flood inundation extents projected on ~~as a function of the~~ building use. Risk perception is defined as the subjective judgement of a decision-maker about the severity of the risk. In this manuscript, this concept is utilized to derive three maps with differentiated exceedance probabilities, each referring to the subjective classification of buildings with varying flood impact. To the best of our knowledge, this combination approach has yet not been used to assess

15   the impact of flooding. The maps help prevent serious damage to buildings and aid in evacuation planning in the case of flooding. The methodology is applied for the flood event of January 2011 in the city of Kulmbach, Germany.

2 Methodology

The framework to generate *building hazard maps* ~~in real-time~~ (as shown in Figure 1~~Figure 1~~) consists of three components:

(1) Hydrological modelling – discharge ensemble forecasts were produced using forecasted precipitation; (2) HD modelling –

20   the water depths were simulated using a pre-calibrated 2D HD model; (3) Post-processing of the model results – a multi-model combination was used to produce flood hazard maps based on a classification of buildings. The framework was tested for the flood event of January 2011 in the city of Kulmbach, Germany. The first two components of the framework were developed in previous studies (Beg et al., 2018; Bhola et al., 2018a, Bhola et al., 2018b). The particular focus of this study is on the development of the framework of a multi-model combination in the post-processing component. For the sake of clarity, each

25   component is described in detail in chronological order.

3

Figure 1. Schematic view of the methodology used to generate building hazard maps. The major components consist of the operational hydrological ensemble forecasts (Beg et al., 2018), the hydrodynamic model and post-processing that includes the multi-model ~~ensemble~~ combination. M$_{x\%}$ denotes the HD model results generated using x% percentile~~confidence interval~~ discharge.

2.1 Hydrological modelling

5   2.1.1 Hydrological model - LARSIM

The conceptual hydrological model LARSIM (Large Area Runoff Simulation Model) was used to study the hydrology of the model area and to generate discharge forecasts. In the model, the hydrological processes are simulated in a series of subarea elements connected by flood routing elements in a pre-determined sequence. LARSIM simulates the hydrologic processes for one element for a defined period and passes the resulting output hydrographs information to the next element (Figure 2). The

10   model structure can be both grid-based or based on hydrologic sub-catchments. The model uses a soil module with storage capacities in considering infiltration, evapotranspiration and runoff generation. The discharge generation consists of three components: runoff generation, runoff concentration and river component. In addition to simulating hydrological processes, LARSIM is most suitable in operational flood forecasting (Demuth and Rademacher, 2016). It deals with the gaps in hydrometeorological input data and allows for the correction/manipulation of numeric weather forecasts (e.g. external forcing

15   parameters). Furthermore, the model automatizes processes for the assimilation of hydrological data, which is crucial in flood forecasting (Luce et al., 2006; Haag and Bremicker, 2013).

4

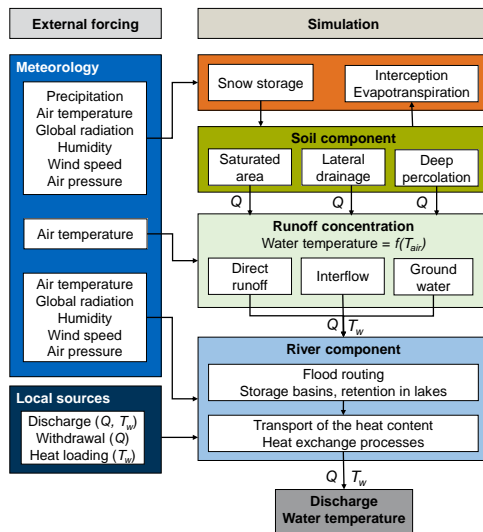**Figure 2. LARSIM water balance model. Source based on Ludwig and Bremicker (2006).**

For this study, a pre-setup model for the study area was provided by the Bavarian Environment Agency and this model is operationally used in the Flood Forecasting Centre for the river Main (Laurent et al., 2010). The model uses a grid-based structure with a resolution of 1 km^2 and a temporal resolution of 1 hour. This LARSIM model considers a soil module with

5   storage capacities in considering the water balance, which consists of three parts: upper, middle and lower soil storages that contribute to the discharge components, modelled as a linear storage system. The model includes 34 parameters that allow the modelling of different processes, such as direct discharge, interflow and groundwater flow. A complete description of calibration parameters is not the scope of this study and has been elaborated on by Ludwig and Bremicker (2006) or Haag et al. (2016). Nevertheless, Table 1 in the Supplement presents is provided for a comprehensive description of important

10   parameters along with eight most sensitive parameters identified in Beg et al. (2018), which were considered in generating the discharge ensemble forecasts.

2.1.2   Discharge ensemble forecasts

The winter flood event of January 2011 was hindcasted to test the framework. The event was one of the largest in terms of its magnitude and corresponds to a discharge of 100-year return period at gauge Kauerndorf (river Schorgast) and 10-year return

15   period at gauge Ködnitz (river White Main). Intense rainfall and snow melting in the Fichtel mountains caused floods in several rivers of Upper Franconia. Within five days, two peak discharges were recorded. The first one occurred on 9th January 2011,

and the second peak measured five days later (on 14th January 2011) caused even higher discharges and water levels. The maximum discharge of 92.5 m³/s was recorded at gauge Kauerndorf and 75.3 m³/s at gauge Ködnitz (Figure 3).

To automatize the generation of forecasts, a tool *FloodEvac* was developed in MATLAB® R2018a (Disse et al., 2018). The tool considers model input and model parameter uncertainty in simulating flood scenario combinations. The tool generates rainfall spatial distributions using sequential conditional geospatial simulations and model parameter uncertainty using Monte-Carlo sampling. The uncertainties in the discharge hydrographs were quantified in Beg et al. (2018) using this FloodEvac tool. In their study, the forecast was performed using 50 ensemble members. Parameter uncertainty module was used to generate 50 different parameter sets (for eight sensitive parameters). In addition, geostatistical simulation for rainfall was implemented using two different R-packages, namely *gstat* and *RandomFields*. The rainfall data was available at an hourly interval at 50 gauges in the catchment. Each forecast was simulated for 61 hours: 49 hours of observed hourly rainfall and 12 hours of forecast rainfall data. To hindcast the event of January 2011, 10 different raster dataset of rainfall uncertainty ~~were~~ was generated for the catchment. The 50 parameter sets were combined with the 10 rainfall uncertainty cases, linking one rainfall scenario with every 5-parameter sets in ~~a~~ sequential order, thus, making 50 sets of hydrological models for the Upper Main catchment. These 50 models were then simulated, and the results of discharge ensembles were stored.

Figure 3 shows the ~~confidence intervals~~percentiles of 75%, 90% and 95% for the January 2011 flood event at two gauging stations upstream of the city, Ködnitz and Kauerndorf. Uncertainty bands are much wider at gauge Ködnitz (Figure 3a) than at gauge Kauerndorf, which suggests that the model parameters are more sensitive in the catchment of White Main than Schorgast. In addition, the peak of the measured discharge at gauge Ködnitz was well over-predicted, which suggests that the uncertainty of the discharges is higher in the catchment of White Main than Schorgast. While the peak of the measured discharge at Kauerndorf is very well predicted, the one at the gauge Ködnitz is over-predicted. Nevertheless, it can be seen from Figure 3 that the ensemble of these 50 members could effectively bracket the observed discharge data.
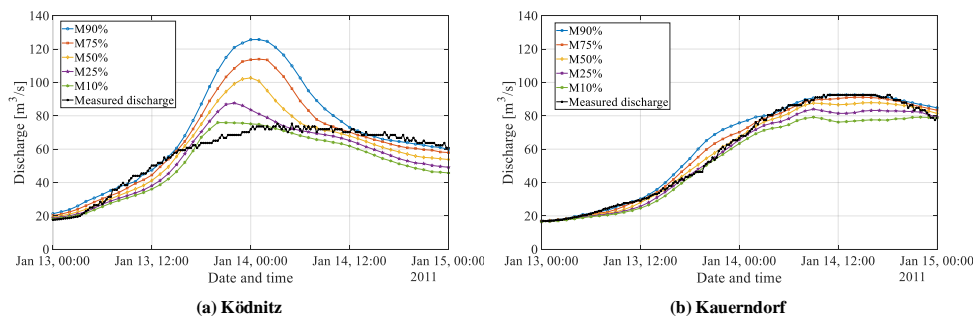

(a) Ködnitz   (b) Kauerndorf

Figure 3. Hindcasted flood event of January 2011: measured discharge hydrograph along with 9~~0~~5%, 75%, 50%, 25%~~90%~~ and 10~~75~~% percentile~~confidence interval~~ discharges for gauges a) Ködnitz and b) Kauerndorf (Data based on Beg et al., 2018).

2.2 Hydrodynamic modelling

HEC-RAS was used as the 2D HD model to quantify uncertainties in flood inundation. It is a non-commercial hydrodynamic model developed by the U.S. Army Corps of Engineers and has been used widely for various flood inundation applications (Moya Quiroga et al., 2016; Patel et al., 2017). The implicit method allows for larger computational time steps compared to an explicit method. HEC-RAS solves either 2D Saint Venant or 2D diffusion-wave equations. The latter allows faster calculation and has greater stability due to its complex numerical schemes (Martins et al., 2017). Due to these advantages and suitability for use in real-time inundation forecast (Henonin et al., 2013), we have used the diffusive-wave model that was previously set-up, calibrated and validated in Bhola et al. (2018a) and Bhola et al. (2018b). For the diffusive-wave approximation, it is assumed that the inertial terms are less than the gravity, friction, and pressure terms. Flow movement is driven by a barotropic pressure gradient balanced by bottom friction (Brunner, 2016). The equations of mass and momentum conservation are as follows:

$$\frac{\partial H}{\partial t} + \frac{\partial(hu)}{\partial x} + \frac{\partial(hv)}{\partial y} + q = 0 \qquad (1)$$

$$g\frac{\partial H}{\partial x} + c_f u = 0 \qquad (2)$$

$$g\frac{\partial H}{\partial y} + c_f v = 0 \qquad (3)$$

$$c_f = \frac{g|V|}{M^2 R^{4/3}} \qquad (4)$$

Where H is the surface elevation (m); h is the water depth (m); u and v are the velocity components in the x- and y- direction respectively (ms^{-1}); q is a source/sink term; g is the gravitational acceleration (ms^{-2}); cf is the bottom friction coefficient (s^{-1}); R is the hydraulic radius (m); |V| is the magnitude of the velocity vector (ms^{-1}); and M is the inverse of the Manning's n (m$^{(1/3)}$s^{-1}).

Table 2 in Supplement summarises the model properties, such as the model size and mesh size, and model roughness in the domain. The model parameter consists of the roughness coefficient Manning's M for five land use classes. The buildings are explicitly included using their shape in the mesh and are excluded from the flow calculation by assigning a high roughness value. To assign hazard to a building, the maximum water depth of all the neighbouring cells was used. Sensitivity analysis of the model was performed using one thousand uniformly distributed model parameter sets for the flood event of 2011.

Although uncertainties arise in the HD modelling, we have considered discharges in hydrological modelling as the sole source of uncertainties in this paper as we have assumed them to be more significant. Various HD simulations were conducted based on confidence intervalspercentiles of the discharges (Figure 3) as upstream boundary conditions at river gauges Ködnitz and Kauerndorf.

2.3 Post-processing

2.3.1 Building use classification

In this study, we have considered only buildings as urban features to access the flood impact and in preparation of flood hazard maps. The shape and use of the buildings were provided by the Bavarian Ministry of the Interior, for Building and Transport (Figure 4). ~~Building damage potential is required for a variety of flood mitigation planning activities including flood damage assessment, multi-hazard analyses and emergency measures (Shultz, 2017).~~

There are various classifications of land use features available in the literature. Dutta et al. (2003) have used direct and indirect damages as the basis of their classification and classified their study area in residential and non-residential categories. Jonkman et al. (2008) have classified urban features in residential, businesses, commercial and public property and agricultural to estimate flood loss. Furthermore, the vulnerability was the basis of classification in residential (Thieken et al., 2008) and ~~i~~Industrial and~~&~~ commercial sectors (Kreibich et al., 2010) in order to estimate flood losses. We have used the damage potential of a building as a basis for classification in order to focus on the flood impact assessment. Building damage potential is required for a variety of flood mitigation planning activities including flood damage assessment, multi-hazard analyses and emergency measures (Shultz, 2017). The buildings

~~The shape and use of the buildings were provided by the Bavarian Ministry of the Interior, for Building and Transport (Figure 4). They~~ were classified into four classes based on their function following the recommendation of the German standard for risk management in urban areas in the case of flash floods (Krieger et al., 2017). According to this standard, building use is one of the important criteria for assessing the damage potential of a building. In this study, four damage potential classes for each building use were taken into consideration as presented in Table 1. In the authors' opinion by keeping our classification simple will likely fit a vast majority of cities regardless of their size. In any case, we acknowledge that the number of classes or~~/~~ criteria can be changed/adapted depending on the aim of the forecast.

The damage potential varies from *low* to *very high* based on the building use, for example, residential buildings with a basement, industries and schools need special protection and thus were rated with a correspondingly high damage potential (class III). In addition, nursery, hospitals as well as low-lying facilities, such as traffic underpasses, driveways to underground garages and other entrances require greater protection and are thus categorised as having the highest damage potential (class IV). Residential buildings and retail businesses were classified as having moderate damage potential (class II), and gardens and parks relatively *low* damage potential (class I). Figure 4 shows the city centre, where buildings were classified according to Table 1. It can be seen that most of the buildings belong to class III as the area is industrial. There are a total of 2695 buildings in Figure 4 of which 1, ~~20,~~ 958~~, and~~ 1716 and 20 buildings were classified in classes I, II, III and IV respectively. The nature of the data in this case study lead~~s~~ to a differentiated representation of the classes~~.~~. It should be noted that~~,~~ the classification aims at creating classes based on damage potential, and not on generating clusters with similar sizes
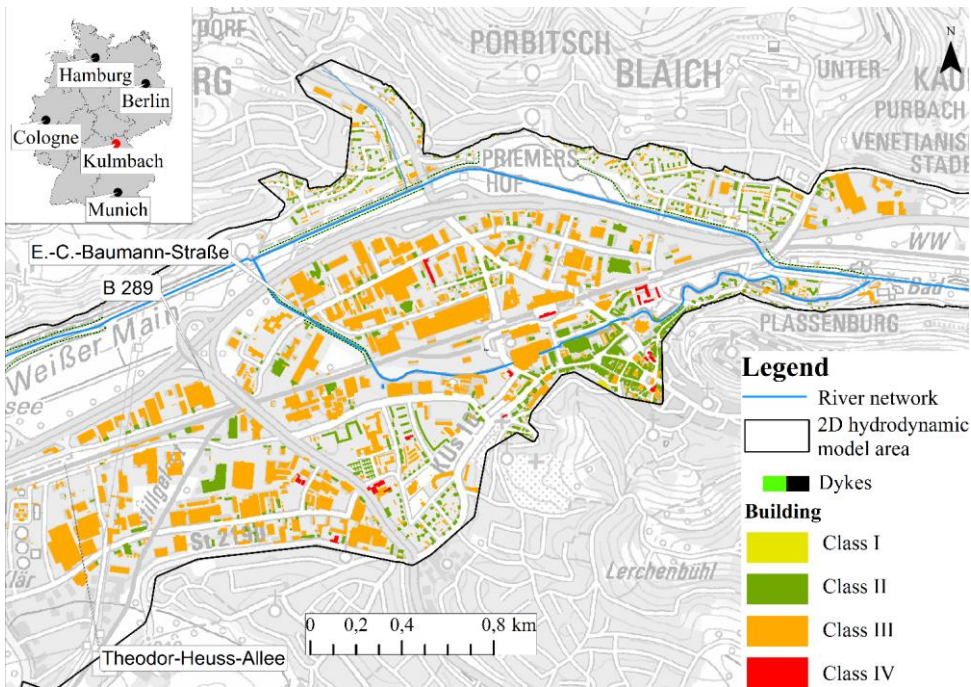
8

Figure 4. ~~City~~ The city of Kulmbach and building use classification. (Data source: Bavarian Ministry of the Interior, for Building and Transport and Water Management Authority Hof).

5

10

Table 1. Building use classification based on the guidelines of Krieger et al. (2017).

Class	Building use	Damage potential
I	Garden buildings Parks / green areas	low
II	Residential building without a basement Retail / small business	moderate
III	Residential building with basement (inhabited) Industry / Trade School / College	high
IV	Nursery / hospital / nursing home / emergency services Energy / telecommunications Underground car park Metro access / Subways	very high

2.3.2 Hazards classification

In this study, hazard classification was based on the recommendations given in the German standard for risk management in urban flood prevention (Krieger et al., 2017). The classification was based on the estimated water depths of the 2D HD model.

5   Table 2 shows the four categories of flooding hazards, which consider water flow in urban areas and vary from *low* to *very high*. It should be noted that in individual cases, the risk may also arise at lower water depths (<0.10 m) for buildings, such as underground parking and metro stations, which are classified as the building class IV in the previous section.

Table 2. Hazard classification used in this study based on water depths. Classification source Krieger et al. (2017)

Hazard class	Flooding hazard	Water depth [m]
1	low	< 0.10 m
2	moderate	0.10 – 0.30 m
3	high	0.30 – 0.50 m
4	very high	> 0.50 m

2.3.3 Multi-model combination

10   The multi-model combination of the 2D HD model results was based on considerations of evacuation planning and gives priority to buildings with higher damage potential. In order to prioritise, it is important to differentiate the impact of water depths on building classes. A certain water depth might have a different impact on a building with higher damage potential. For example, there is more threat for a low water depth in underground metro access that the same water depth to a residential

building. Therefore, buildings classified to higher damage potential class relates to model results of a higher ~~confidence interval~~percentile. Each building class corresponds to a certain discharge percentile~~confidence interval~~ and the resulting damage potential assessment can be visualised and presented as a building hazard map.

Figure 5 shows an example of a multi-model combination in which the four building classes were assigned four different ~~confidence intervals~~percentiles. The simulation results (water depth in this case) obtained from the HD model with 25%, 50%, 75% and 90% ~~confidence interval~~percentile discharges were assigned to the building classes I, II, III and IV respectively. The novelty of the multi-model combination approach is that the flood inundation uncertainty is coupled with the building use. As such evacuation planning or investment planning can take the information of uncertainties in the water depths into consideration.

Figure 5. An example of a multi-model combination in which the four building classes I, II, III and IV are assigned to the 2D HD model results of 25%, 50%, 75% and 90% respectively.

3 Results

In this section, we present the results of five percentiles~~confidence intervals~~ and the performance of the multi-model combination. To assess the methodology, the flood event of January 2011 was used to quantify uncertainties in discharge hydrographs. The forecasts corresponding to 10%, 25%, 50%, 75% and 90% percentiles~~confidence intervals~~ were further used as input boundary conditions to the 2D HD model and water depths were stored. Furthermore, the flood inundation maps and building hazards were then classified.

3.1 Flood inundation maps and building hazards

The number of affected buildings in each hazard class for all five HD models are presented in Figure 6. As the discharge ~~exceedance probability~~percentile increases, the number of affected buildings in each hazard class increases. The maximum flood inundation of the five models is presented in Figure 7. The figures present both the inundation extent and building hazards based on the classification discussed in section 2.3.2.

Post-event binary information of the flood extent was collected from newspaper articles and press releases published by the Bavarian Water Authority. The information shows that the dykes were at their full capacity and most of the floodplains and traffic routes were flooded, but no serious damage was reported (Hof, 2011). The streets Theodor-Heuss-Allee and E.-C.-Baumann-Straße were flooded and some flooding was observed on motorway B289 (see Figure 4 for locations).

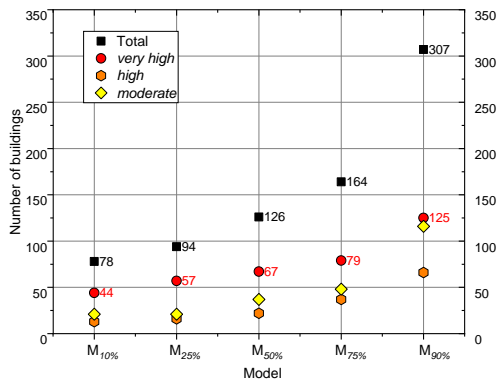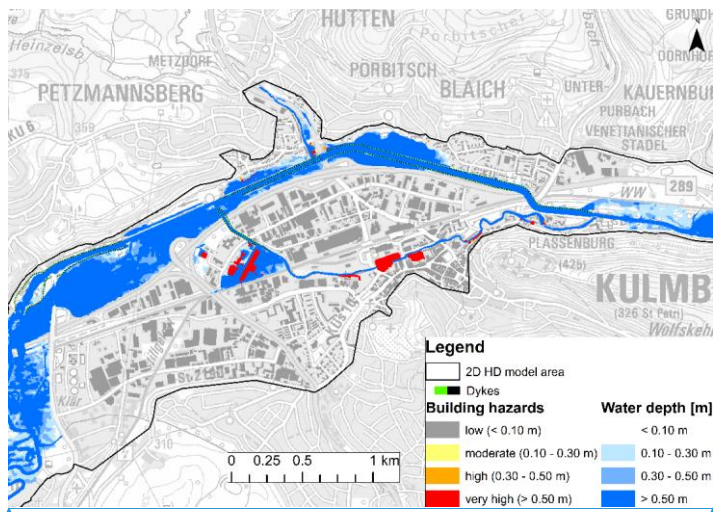**Figure 6.** ~~Number~~ **The number** of affected buildings in each hazard class for 2D HD model results using five ~~confidence interval~~**discharge percentiles** ~~discharges~~.
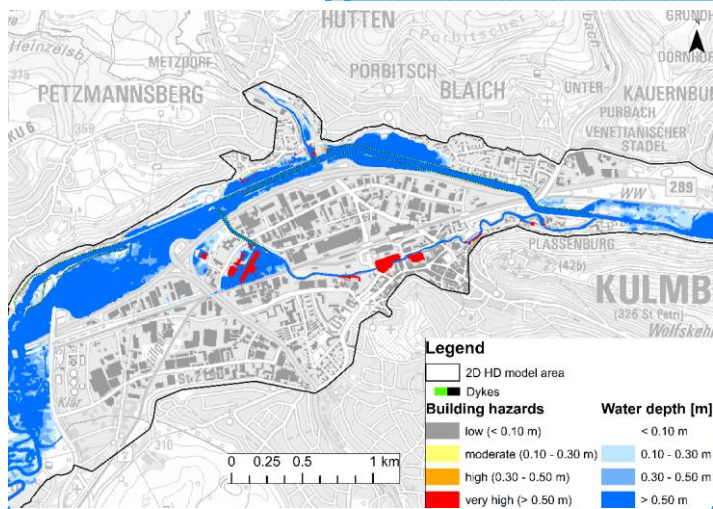
(aa) M$_{10\%}$

(bb) M$_{25\%}$

(ce) M50%

(a) M10%

(d) M$_{75\%}$

(ee) M$_{90\%}$

Figure 7. Flood inundation and building hazard maps for five ~~confidence intervals~~percentiles discharge hydrographs.

3.2 Multi-model combination

Three combination scenarios based on risk-seeking, risk-neutral and risk-averse *high, average* and *low* risk perceptionsexceedance probability approach were were designed to illustrate the methodology developed in this for this study and and are presented in Table 3. The main objective of the combination is to differentiate the impact of water depths on building classes. Therefore, to design the combinations, a high confidence intervalpercentile was assigned to the buildings with high damage potential class. Each scenario presents a certaingiven risk perception of a decision-maker that. This can be easily, which can be adjusted depending on the needperception of different stakeholders. Hence, Variousdifferent risk perceptions will lead to different exceedance probability mapsies,. e.g. aA person with risk-averse person will likely select a hazard map with a low exceedance probabilitiesy for the hazard maps, whereas a for his ris. A risk-seeking person will likely develop risk management strategies based on higher exceedance probabilities management. The hazard maps for the three scenarios are shown in Figure 8.

Table 3. Scenarios of Mmulti-model combinations scenarios based on exceedance probabilityrisk perception.

Scenario	Exceedance probabilityRisk perception	Building class			
		I	II	III	IV
I	HighRisk-seeking	$M_{10\%}$	$M_{10\%}$	$M_{25\%}$	$M_{50\%}$
II	AverageRisk-neutral	$M_{10\%}$	$M_{25\%}$	$M_{50\%}$	$M_{75\%}$
III	LowRisk-averse	$M_{25\%}$	$M_{50\%}$	$M_{75\%}$	$M_{90\%}$

(a) Scenario I: risk-seeking

(b) Scenario II: risk-neutral

(ce) Scenario III: risk-averse

Figure 8. Building hazard maps for the three scenarios, the numbers of affected buildings are 84, 107 and 142 respectively. Hazard classification is based on Krieger et al. (2017).

4 Discussion

Prior work in hydrology has demonstrated the effectiveness of multi-model combinations in improving flood forecasts as compared to the *best-model* approach (Weigel et al., 2008). However, these methodologies were previously limited to discharge ensemble forecasts and were not researched for hazard maps. In this study, we extend the use of multi-model combinations to produce flood hazard maps for buildings depending on their use.

First, the five simulation results are presented in Figure 7 as inundation and building hazard maps. It should be noted that few buildings show *very high* hazards due to their proximity to the Mühl canal (Figure 7a). Even though there was no over-topping of water from the canal, because of buildings geoGEO location being the buildings near to the canal, these were assigned automatically with the highest hazard, starting with a discharge of M$_{10\%}$. Ideally, This limitation can potentially be tackledthis should be prevented by removing the river channel elements from the dataset before to assign the water depths to the buildings by excluding the buildings as in Bermúdez and Zischg (2018). However, and without retract to our conclusions, it was decided not to include it in this work in order to keep the automation process simple. Up to a discharge of M$_{50\%}$, no inundation in the city centre was observed as the dykes were not breached. It can be observed in Figure 6, that the increment in the number of affected buildings is gradual, especially the buildings belonging to *very high* hazard class. As the peak discharge increases in

18

$M_{75\%}$, the dykes at the B289 road were breached and water entered in the city centre and more buildings were affected. Most damages were observed in $M_{90\%}$ with 307 affected buildings, out of which 125 buildings show *very high* hazard, an increment of 46 from $M_{75\%}$. The affected buildings were located in the city centre (Figure 7e), mainly in industrial and commercial areas. Similarly, the streets Theodor-Heuss-Allee and E.-C.-Baumann-Straße were inundated starting from a discharge of $M_{50\%}$.

5   In operational use, the mean of the discharge ensemble or $M_{50\%}$ would normally have been used as the *best-model*, which according to Figure 7c, is in agreement with the post-event information. However, this match might not always be representative, especially in the case of an event of different or higher magnitude, as discussed in Di Baldassarre et al. (2010). They argued that visualising flood hazards as a probability is a more accurate representation as compared to a single *best-model*, which might misrepresent the uncertainty in the modelling process.

10   With the objective of visualising uncertainties, three scenarios based on risk perception~~exceedance probability~~ were used to combine HD model results and are presented in Figure 8. In scenarios I and II, 84 and 107 buildings were affected, which shows that the impact of risk-seeking and risk-neutral ~~*high* and *average* exceedance probability~~risk perception scenarios was less as compared to $M_{50\%}$ in which a total of 126 buildings were affected, out of which 67 buildings were classified in *very high* hazard class.

15   Further, as the majority of the buildings were classified in class II and III, the resulting map of a risk-averse~~low exceedance probability~~risk perception (~~scenario III~~) corresponds closely with $M_{50\%}$ and $M_{75\%}$, with 142 affected buildings. In scenario II, 63 buildings were classified in the *very high* hazard class, which increased to 71 in scenario III. Similarly, 22 buildings belonged to both *moderate* and *high* hazard classes, and shifting to scenario III, the number increased to 33 and 38 in the *moderate* and *high* classes respectively.

20   In Figure 9, a comparison is presented between the *best-model* ($M_{50\%}$) and the multi-model ~~approach~~combinations and the areas with prominent changes are highlighted in red circles. The figure presents building hazards resulting from the combination of ~~exceedance probability~~risk perception scenarios~~,~~ and locates 16 buildings that are affected as compared to $M_{50\%}$. ~~The areas with prominent changes are highlighted in red circles and here we compare the red circles~~compared from~~using the M50% t~~of the multi-model combination. The buildings that belong to class III (Figure 9b) were assigned the results of

25   $M_{75\%}$, and show a *very high* hazard. Figure 9a~~d~~ shows that an adjacent building belonging to class II (ID 1393) was not flooded. This demonstrates that the methodology was implemented accurately and prioritised measures such as flood impact assessment, spatial planning, early warning and emergency planning, according to the damage potential of a building. The prioritisation is important in order to focus on a combination of various evacuation strategies to prevent damage and save lives (Kolen et al., 2010). Hence, decision-makers must be made aware of the impact associated with a ~~lower probability risk~~risk-

30   averse perception to improve their planning strategies (Pappenberger and Beven, 2006; Uusitalo et al., 2015).

(a) Multi-model combination

(b) Building use classification

(c) M_{50%}

(b) Building use classification

Figure 9. Comparison of building hazard maps between *best-model* (M_{50%}) vs. multi-model ~~approach~~**combinations. The areas with prominent change are highlighted in red circles.**

A potential drawback of the combination is that the hazard classification may shift from *low* to *very high* in two adjacent buildings belonging to different classes. This might confuse evacuation planners by presenting inconsistent information. To tackle this issue, more information and specific guidelines should be provided to them on how to use the maps. In addition, continuous flood inundation maps are hard to obtain, especially at the boundaries of two combinations. There might be a step rise in the water depths while shifting from the results of one model to another. To address this issue, future research should be conducted to provide consistency in interpolation and in combining models (see Zazar et al., 2018). In addition, in order to avoid the confusion, these maps could be forecasted for a regular interval of 3-4 hours.

Overall, the methodology is independent of the choice of models, i.e. hydrological and HD, and is transferable to other study areas. In order to use the methodology in real-time, the run-time of the flood forecasting modelling should be below the flow travel time. In this study, a 50-member ensemble forecast was used from Beg et al. (2018) where the entire process took 25 minutes in a 3-core desktop in parallel mode to generate a forecast of 12 hours. Various percentile discharges were then run

5   simultaneously in the HD model, which required 30 minutes to simulate a 12-hour event on an 8-core, 2.4 GHz (Intel E5-2665), including the initial start (Bhola et al. 2018a). Post-processing of the model results would consume an additional 15 min. Therefore, real-time hazard maps are delivered to decision–makers in 70 minutes. A faster run-time ~~, which~~ can be ensured by either using a simple model structure (Leandro et al., 2014) and/or high-performance computing (Kuchar et al., 2015). In the absence of such infrastructures or a very large catchment size, HD models can be replaced with alternatives, such

10   as terrain-based models (Zheng et al., 2018) and satellite images (Voigt et al., 2007). In addition, a database of pre-recorded inundation scenarios as shown in Bhola et al. (2018a) can expand the application of this methodology.

Molinari et al. (2014) have stated that a comprehensive uncertainty assessment improves emergency responses by assessing the potential consequences of flood events. Therefore, our methodology would allow the target users to benefit from hazard maps enabling them to better prioritise and coordinate evacuation planning based on the stakeholder perception to risk. The

15   maps could further serve as a tool for flood risk assessment. The methodology can be used for flood mitigation and flood forecast planning in the form of emergency management training, where forecasted hazard scenarios can be presented to the training groups. By visualising inundation scenarios, potential damage at the building's level which ha~~ve~~s been prioritized based on ~~a~~the desired classification~~,~~ can be estimated with this methodology and made available together with each forecasted scenario.

20   **5 Conclusions**

In summary, we have presented a new methodology for flood impact assessment using a multi-model combination in the form of *building hazard maps* with differentiated ~~exceedance probability~~risk perception*s*. These maps ~~inherently~~ offer an alternative way to communicate the underlying uncertainties in forecasting models and are ready-to-use for decision-makers in the field of flood risk management. The entire forecasting framework consists of three stages: (i) generation of discharge ensemble

25   forecasts, (ii) 2D HD simulations using the generated forecasts and (iii) hazard maps using multi-model combinations. The framework was applied to the city of Kulmbach and three multi-model combination*s* ~~scenarios~~ were designed based on ~~exceedance probability~~risk perception. The model results of $M_{50\%}$ show a good match with binary information collected after the flood event. The risk-seeking ~~low~~*high* risk perception~~exceedance probability~~ scenario corresponds closely with $M_{50\%}$ and $M_{75\%}$. We expect this multi-model combination to improve the current visualisation techniques in operational flood risk

30   management and evacuation planning.

In this study, we have considered only buildings as a feature; additional urban features, such as bridges (Gebbeken et al., 2016) and roads (Goerigk et al., 2018), should be included in future to extend the methodology. Furthermore, other sources of

22

uncertainty, such as HD model parameters, model structures and measured data should also be incorporated for a comprehensive assessment. In addition, the economic, social and hazardous effects of carrying out an evacuation in the case of false alarm must be considered. Hence, a validation of the combination is crucial to building trust in its prediction in real-time. Further research investigating multi-model combinations and validation in other study areas may be beneficial. A more

5 extensive study on the validation of the multi-model combination may be required, possibly by using measuring gauges, post-event survey (as conducted in Thieken et al., 2005), satellite images (as in Triglav-Čekada and Radovan, 2013), and/or crowd-sourced data (Bhola et al., 2018b).

In future, damage potential classification can further be improved by including additional criteria, such as population density in urban areas or water quality, and with it extend the applicability of this method. For example, the assessment of the damage

10 potential of commercial enterprises, substances or machinery containing elements that which could be a source of water pollution could be included (Krieger et al., 2017). In addition, other classification methods for buildings and hazard types should be evaluated, especially to further dissect the impact of class III in commercial and industrial. Finally, the output of the framework cancanould be extended to upload hazard maps uploaded in a web-based GIS system to improve visualization, along with providing layers of additional information, such as inundation pathways and weak spots in the river and floodplains

15 to provide sufficient details to intervene (aid in planning). such as inundation pathways and hot spot to aid in planning. Thise additional information latter would enhance the usefulness to different target users, such as planners, decision-makers and flood forecastersing agencies. In addition, analysing additional model outputs, such as flow velocity, flood duration and contamination of flood water, should improve the existing forecasting framework by incorporating flood risk assessments.

20 **Author contribution**

Punit K. Bhola conceptualised and completed the formal analysis of uncertainty analysis. Punit K. Bhola wrote the original draft and subsequently reviewed and edited by all co-authors. All authors contributed to writing the paper.

Competing interests

The authors declare that they have no conflict of interest.

25 **Acknowledgements**

23

in Hof for providing us with the quality data to conduct the research. We would also like to thank the language centre of the Technical University of Munich for their consulting in improving the language of the manuscript.

References

Alfieri, L., Feyen, L., Salamon, P., Thielen, J., Bianchi, A., Dottori, F., and Burek, P.: Modelling the socio-economic impact of river floods in Europe, Nat. Hazards Earth Syst. Sci., 16, 1401-1411, 10.5194/nhess-16-1401-2016, 2016.

Bartels, J., Seidel, J., Bardossy, A., Bliefernicht, J., Kunstmann, H., Kunstmann, H., Johst, M., and Demuth, N.: Bewertung von Ensemble-Abflussvorhersagen für die operationelle Hochwasserwarnung, Hydrol. Wasserbewirtsch. Hydrologie und Wasserbewirtschaftung, 61, 297-310, doi: 10.5675/HyWa_2017,5_1 2017.

Bates, P. D., Pappenberger, F., and Romanowicz, R. J.: Uncertainty in Flood Inundation Modelling, in: Applied Uncertainty Analysis for Flood Risk Management, edited by: Beven, K., and Hall, J., Imperial College Press, London, UK, 232-269, 2014.

Beg, M. N. A., Leandro, J., Bhola, P., Konnerth, I., Amin, K., Koeck, F., Carvalho, R. F., and Disse, M.: Flood Forecasting with Uncertainty Using a Fully Automated Flood Model Chain: A Case Study for the City of Kulmbach, HIC 2018. 13th International Conference on Hydroinformatics, Palermo, Italy, 2018.

Bermúdez, María; Zischg, Andreas Paul (2018): Sensitivity of flood loss estimates to building representation and flow depth attribution methods in micro-scale flood modelling. In Nat Hazards 92 (3), pp. 1633–1648. DOI: 10.1007/s11069-018-3270-7.

Beven, K., and Binley, A.: GLUE: 20 years on, Hydrol. Process., 28, 5897-5918, 10.1002/hyp.10082, 2014.

Beven, K., Leedal, D., and McCarthy, S.: Framework for assessing uncertainty in fluvial flood risk mapping, CIRIA, London, UK, 58 pp., 2014.

Beven, K., Lamb, R., Leedal, D., and Hunter, N.: Communicating uncertainty in flood inundation mapping: a case study, Intl. J. River Basin Management, 13, 285-295, 10.1080/15715124.2014.917318, 2015.

Beven, K. J., Almeida, S., Aspinall, W. P., Bates, P. D., Blazkova, S., Borgomeo, E., Freer, J., Goda, K., Hall, J. W., Phillips, J. C., Simpson, M., Smith, P. J., Stephenson, D. B., Wagener, T., Watson, M., and Wilkins, K. L.: Epistemic uncertainties and natural hazard risk assessment – Part 1: A review of different natural hazard areas, Nat. Hazards Earth Syst. Sci., 18, 2741-2768, 10.5194/nhess-18-2741-2018, 2018.

Bhola, P., Leandro, J., and Disse, M.: Framework for Offline Flood Inundation Forecasts for Two-Dimensional Hydrodynamic Models, Geosciences, 8, 346, doi: 10.3390/geosciences8090346, 2018a.

Bhola, P. K., Nair, B. B., Leandro, J., Rao, S. N., and Disse, M.: Flood inundation forecasts using validation data generated with the assistance of computer vision, J. Hydroinform., 21, 240-256, 10.2166/hydro.2018.044, 2018b.

Boelee, L., Lumbroso, D. M., Samuels, P. G., and Cloke, H. L.: Estimation of uncertainty in flood forecasts—A comparison of methods, J. Flood Risk. Manag., 0, e12516, 10.1111/jfr3.12516, 2018.

Bruen, M., Krahe, P., Zappa, M., Olsson, J., Vehvilainen, B., Kok, K., and Daamen, K.: Visualizing flood forecasting uncertainty: some current European EPS platforms—COST731 working group 3, Atmos. Sci. Lett., 11, 92-99, 10.1002/asl.258, 2010.

Brunner, G. W., United, S., Army, Corps of, E., Institute for Water, R., and Hydrologic Engineering, C.: HEC-RAS river analysis system: hydraulic reference manual Version 5.0, US Army Corps of Engineers, Institute for Water Resources, Hydrologic Engineering Center, Davis, CA, 2016.

Cloke, H. L., and Pappenberger, F.: Ensemble flood forecasting: A review, J. Hydrol., 375, 613-626, doi: 10.1016/j.jhydrol.2009.06.005, 2009.

Demuth, N., and Rademacher, S.: Chapter 5 - Flood Forecasting in Germany — Challenges of a Federal Structure and Transboundary Cooperation, in: Flood Forecasting, edited by: Adams, T. E., and Pagano, T. C., Academic Press, Boston, 125-151, 2016.

Di Baldassarre, G., and Montanari, A.: Uncertainty in river discharge observations: a quantitative analysis, Hydrol. Earth Syst. Sci., 13, 913-921, 10.5194/hess-13-913-2009, 2009.

Di Baldassarre, G., Schumann, G., Bates, P. D., Freer, J. E., and Beven, K. J.: Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches, Hydrolog. Sci. J., 55, 364-376, 10.1080/02626661003683389, 2010.

Disse, M., Konnerth, I., Bhola, P. K., and Leandro, J.: Unsicherheitsabschätzung für die Berechnung von dynamischen Überschwemmungskarten – Fallstudie Kulmbach, in: Vorsorgender und nachsorgender Hochwasserschutz: Ausgewählte Beiträge aus der Fachzeitschrift WasserWirtschaft Band 2, edited by: Heimerl, S., Springer Fachmedien Wiesbaden, Wiesbaden, 350-357, 2018.

Dotto, C. B. S., Mannina, G., Kleidorfer, M., Vezzaro, L., Henrichs, M., McCarthy, D. T., Freni, G., Rauch, W., and Deletic, A.: Comparison of different uncertainty techniques in urban stormwater quantity and quality modelling, Water Res., 46, 2545-2558, doi: 10.1016/j.watres.2012.02.009, 2012.

Doyle, E. E. H., Johnston, D. M., Smith, R., and Paton, D.: Communicating model uncertainty for natural hazards: A qualitative systematic thematic review, Int. J. Disast. Risk Re., 33, 449-476, doi: 10.1016/j.ijdrr.2018.10.023, 2019.

Dutta, Dushmanta; Herath, Srikantha; Musiake, Katumi (2003): A mathematical model for flood loss estimation. In Journal of Hydrology 277 (1-2), pp. 24–49. DOI: 10.1016/S0022-1694(03)00084-2.

European Union: Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. Union, E. (Ed.), 2007.

Gebbeken, N., Videkhina, I., Pfeiffer, E., Garsch, M., and Rüdiger, L.: Risikobewertung und Schutz von baulichen Infrastrukturen bei Hochwassergefahr, Bautechnik, 93, 199-213, 10.1002/bate.201600003, 2016.

Goerigk, M., Hamacher, H. W., and Kinscherff, A.: Ranking robustness and its application to evacuation planning, Eur. J. Oper. Res., 264, 837-846, doi: 10.1016/j.ejor.2016.05.037, 2018.

Haag, I., and Bremicker, M.: Möglichkeiten und Grenzen der Schneesimulation mit dem Hochwasservorhersagemodell LARSIM, Forum für Hydrologie und Wasserbewirtschaftung, 33, 47-58, 2013.

Haag, I., Johst, M., Sieber, A., and Bremicher, M.: Guideline for the Calibration of LARSIM Water Balance Models for operational Application in Flood Forecasting, LARSIM Entwicklergemeinschaft – Hochwasserzentralen LUBW, BLfU, LfU RP, HLNUG, BAFU, 2016.

Hagemeier-Klose, M., and Wagner, K.: Evaluation of flood hazard maps in print and web mapping services as information tools in flood risk communication, Nat. Hazards Earth Syst. Sci., 9, 563-574, doi: 10.5194/nhess-9-563-2009, 2009.

Hammond, M. J., Chen, A. S., Djordjevi, S., Butler, D., and Mark, O.: Urban flood impact assessment: A state-of-the-art review, Urban Water J., 12, 14-29, doi: 10.1080/1573062X.2013.857421, 2013.

Jonkman, S. N.; Bockarjova, M.; Kok, M.; ˇBernardini, P.: Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. In Ecological Economics 66 (1), pp. 77–90. DOI: 10.1016/j.ecolecon.2007.12.022, 2008.

Wasserwirtschaftsamt Hof: Gebiet des Mains: https://www.wwa-ho.bayern.de/hochwasser/hochwasserereignisse/januar2011/main/index.htm, access: 27.03.2019, 2011.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P., and Thielen, J.: Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level, Environ. Modell. Softw., 75, 68-76, doi: 10.1016/j.envsoft.2015.09.009, 2016.

Kolen, B., Kutschera, G., and Helsloot, I.: A comparison between the netherlands and germany of evacuation in case of extreme flooding, 2019.

Kreibich, Heidi; Seifert, Isabel; Merz, Bruno; Thieken, Annegret H.: Development of FLEMOcs – a new model for the estimation of flood losses in the commercial sector. In Hydrological Sciences Journal 55 (8), pp. 1302–1314. DOI: 10.1080/02626667.2010.529815, 2010.

Krieger, K., Schmitt, T. G., and Illgen, M.: Risikomanagement in der kommunalen überflutungsvorsorge nach DWA-Merkblatt M 119, GWF Wasser Abwasser, 158, 34-39, 2017.

Kuchar, S., Golasowski, M., Vavrik, R., Podhoranyi, M., Sir, B., and Martinovic, J.: Using High Performance Computing for Online Flood Monitoring and Prediction, Int. J. Env. Chem. Ecolog. Geolog. Geophy. Eng., 101, 432 - 437, 2015.

Laurent, S., Hangen-Brodersen, C., Ehret, U., Meyer, I., Moritz, K., Vogelbacher, A., and Holle, F.-K.: Forecast Uncertainties in the Operational Flood Forecasting of the Bavarian Danube Catchment, 2010.

Leandro, J., Leitão, J. P., and de Lima, J. L. M. P.: Quantifying the uncertainty in the Soil Conservation Service flood hydrographs: a case study in the Azores Islands, J. Flood Risk. Manag., 6, 279-288, doi: 10.1111/jfr3.12010, 2013.

Leandro, J., Chen, A. S., and Schumann, A.: A 2D parallel diffusive wave model for floodplain inundation with variable time step (P-DWave), J. Hydrol., 517, 250-259, doi: 10.1016/j.jhydrol.2014.05.020, 2014.

Leandro, J., Schumann, A., and Pfister, A.: A step towards considering the spatial heterogeneity of urban key features in urban hydrology flood modelling, J. Hydrol., 535, 356-365, doi: 10.1016/j.jhydrol.2016.01.060, 2016.

Leedal, D., Neal, J., Beven, K., Young, P., and Bates, P.: Visualization approaches for communicating real-time flood forecasting level and inundation information, J. Flood Risk. Manag., 3, 140-150, doi: 10.1111/j.1753-318X.2010.01063.x, 2010.

Luce, A., Haag, I., and Bremicker, M.: Daily discharge forecasting with operational water-balance models in Baden-Württemberg, Hydrol. Wasserbewirtsch., 50, 58-66, 2006.

Ludwig, K., and Bremicker, M.: The water balance model LARSIM : design, content and applications, Inst. für Hydrologie der Univ., Freiburg i. Br., 2006.

Martins, R., Leandro, J., Chen, A. S., and Djordjević, S.: A comparison of three dual drainage models: shallow water vs local inertial vs diffusive wave, J. Hydroinform., 19, 331-348, 10.2166/hydro.2017.075, 2017.

Meyer, V., Scheuer, S., and Haase, D.: A multicriteria approach for flood risk mapping exemplified at the Mulde river, Germany, Nat. Haz., 48, 17-39, 10.1007/s11069-008-9244-4, 2009.

Molinari, D., Ballio, F., Handmer, J., and Menoni, S.: On the modeling of significance for flood damage assessment, Int. J. Disast. Risk Re., 10, 381-391, doi: 10.1016/j.ijdrr.2014.10.009, 2014.

Moya Quiroga, V., Kure, S., Udo, K., and Mano, A.: Application of 2D numerical simulation for the analysis of the February 2014 Bolivian Amazonia flood: Application of the new HEC-RAS version 5, RIBAGUA - Revista Iberoamericana del Agua, 3, 25-33, https://doi.org/10.1016/j.riba.2015.12.001, 2016.

Nester, T., Komma, J., Viglione, A., and Blöschl, G.: Flood forecast errors and ensemble spread—A case study, Water Resour. Res., 48, doi: 10.1029/2011wr011649, 2012.

Oudin, L., Andréassian, V., Mathevet, T., Perrin, C., and Michel, C.: Dynamic averaging of rainfall-runoff model simulations from complementary model parameterizations, Water Resour. Res., 42, doi: 10.1029/2005wr004636, 2006.

Pappenberger, F., and Beven, K. J.: Ignorance is bliss: Or seven reasons not to use uncertainty analysis, Water Resour. Res., 42, doi: 10.1029/2005WR004820, 2006.

Pappenberger, F., Harvey, H., Beven, K., Hall, J., and Meadowcroft, I.: Decision tree for choosing an uncertainty analysis methodology: a wiki experiment, Hydrol. Process., 20, 3793-3798, 10.1002/hyp.6541, 2006.

Pappenberger, F., Beven, K. J., Ratto, M., and Matgen, P.: Multi-method global sensitivity analysis of flood inundation models, Adv. Water Resour., 31, 1-14, doi: 10.1016/j.advwatres.2007.04.009, 2008.

Patel, D. P., Ramirez, J. A., Srivastava, P. K., Bray, M., and Han, D.: Assessment of flood inundation mapping of Surat city by coupled 1D/2D hydrodynamic modeling: a case application of the new HEC-RAS 5, Nat. Haz., 89, 93-130, 10.1007/s11069-017-2956-6, 2017.

Paulsen, D., Platt M., Huettel, S. and Brannon E.: From risk-seeking to risk-averse: the development of economic risk preference from childhood to adulthood, Frontiers in Psychology, https://doi.org/10.3389/fpsyg.2012.00313, 2012.

Schanze, J.: A hybrid multi-model approach to river level forecasting, in: Flood Risk Management: Hazards, Vulnerability and Mitigation Measures, edited by: Schanze, J., Zeman, E., and Marsalek, J., Flood Risk Management: Hazards, Vulnerability and Mitigation Measures, Springer Netherlands, Dordrecht, 1-20, 2006.

See, L., and Openshaw, S.: A hybrid multi-model approach to river level forecasting, Hydrolog. Sci. J., 45, 523-536, doi: 10.1080/02626660009492354, 2000.

Formatiert: Englisch (Vereinigte Staaten)

Shamseldin, A. Y., O'Connor, K. M., and Liang, G. C.: Methods for combining the outputs of different rainfall–runoff models, J. Hydrol., 197, 203-229, doi: 10.1016/S0022-1694(96)03259-3, 1997.

Shultz, S.: The Extent and Nature of Potential Flood Damage to Commercial Property Structures in the Midwestern United States, J. Contemp. Water Res. Educ., 161, 81-91, doi: 10.1111/j.1936-704X.2017.3253.x, 2017.

Thieken, A. H., Müller, M., Kreibich, H., and Merz, B.: Flood damage and influencing factors: New insights from the August 2002 flood in Germany, Water Resour. Res., 41, doi: 10.1029/2005wr004177, 2005.

Thieken, A. H.; Olschewski, A.; Kreibich, H.; Kobsch, S.; Merz, B.: Development and evaluation of FLEMOps – a new F lood L oss E stimation MO del for the private s ector. In D. Proverbs, C. A. Brebbia, E. Penning-Rowsell (Eds.): Flood Recovery, Innovation and Response I. FLOOD RECOVERY, INNOVATION AND RESPONSE 2008. London, England, 02.07.2008 - 03.07.2008. Southampton, UK: WIT Press, pp.315–324, 2008.

Thieken, A. H., Bessel, T., Kienzler, S., Kreibich, H., Müller, M., Pisi, S., and Schröter, K.: The flood of June 2013 in Germany: how much do we know about its impacts?, Nat. Hazards Earth Syst. Sci., 16, 1519-1540, doi: 10.5194/nhess-16-1519-2016, 2016.

Todini, E.: Flood Forecasting and Decision Making in the new Millennium. Where are We?, Water Res. Manag., 31, 3111-3129, doi: 10.1007/s11269-017-1693-7, 2017.

Triglav-Čekada, M., and Radovan, D.: Using volunteered geographical information to map the November 2012 floods in Slovenia, Nat. Hazards Earth Syst. Sci., 13, 2753-2762, doi: 10.5194/nhess-13-2753-2013, 2013.

Uusitalo, L., Lehikoinen, A., Helle, I., and Myrberg, K.: An overview of methods to evaluate uncertainty of deterministic models in decision support, Environ. Modell. Softw., 63, 24-31, doi: 10.1016/j.envsoft.2014.09.017, 2015.

Voigt, S., Kemper, T., Riedlinger, T., Kiefl, R., Scholte, K., and Mehl, H.: Satellite Image Analysis for Disaster and Crisis-Management Support, IEEE T. Geosci. Remote., 45, 1520-1528, doi: 10.1109/TGRS.2007.895830, 2007.

Wasserwirtschaftsamt Hof: Gebiet des Mains: https://www.wwa-ho.bayern.de/hochwasser/hochwasserereignisse/januar2011/main/index.htm, access: 27.03.2019, 2011.

Weigel, A. P., Liniger, M. A., and Appenzeller, C.: Can multi-model combination really enhance the prediction skill of probabilistic ensemble forecasts?, Q. J. Roy. Meteor. Soc., 134, 241-260, doi: 10.1002/qj.210, 2008.

Zarzar, C. M., Hosseiny, H., Siddique, R., Gomez, M., Smith, V., Mejia, A., and Dyer, J.: A Hydraulic MultiModel Ensemble Framework for Visualizing Flood Inundation Uncertainty, J. Am. Water Resour. As., 54, 807-819, doi: 10.1111/1752-1688.12656, 2018.

Zheng, X., Tarboton, D. G., Maidment, D. R., Liu, Y. Y., and Passalacqua, P.: River Channel Geometry and Rating Curve Estimation Using Height above the Nearest Drainage, J. Am. Water Resour. As., 54, 785-806, doi: 10.1111/1752-1688.12661, 2018.

Jonkman, S. N.; Bockarjova, M.; Kok, M.; ˇBernardini, P. (2008): Integrated hydrodynamic and economic modelling of flood damage in the Netherlands. In Ecological Economics 66 (1), pp. 77–90. DOI: 10.1016/j.ecolecon.2007.12.022.Kreibich, Heidi; Seifert, Isabel; Merz, Bruno;

29