Interactive comment on "Assimilation of probabilistic flood maps from SAR data into a hydrologic-hydraulic forecasting model: a proof of concept" *by* Concetta Di Mauro et al.

Arnoud Goossen

arnoud.goossen@wur.nl

This review was prepared as part of graduate program course work at Wageningen University and has been produced under supervision of dr Ryan Teuling. The review has been posted because of its potential usefulness to the authors and editor. Although it has the format of a regular review as was requested by the course, this review was not solicited by the journal, and should be seen as a regular comment, we leave it up to the author's and editor which points will be addressed.

Floods represent one of the major natural disasters with a global annual average loss of US $104 billion, which emphasizes the need for reliable and cost-effective flood fore-

casting models. In this manuscript the authors aim to understand the main strengths and limitations of a previously proposed data assimilation framework in a fully controlled environment, in order to improve the quality of flood forecasting models. To do so, they performed synthetic twin experiments; At first, following the study from Hostache et al. (2018) only speckle uncertainty of the SAR image has been taken into account in the Probabilistic flood maps. In a second step, a bias to reproduce misclassified pixels is introduced in the synthetic SAR observations. The experimental results show that the assimilation of SAR probabilistic flood maps significantly improves the predictions of streamflow and water elevation, thereby confirming the effectiveness of the data assimilation framework.

This study is a follow-up of the previous real case study from Hostache et al., (2018). Major issues found in this study have led to this follow-up study, where two major things have been carried out differently. Firstly, Hostache et al. (2018) used a variant of the Particle Filter with Sequential Importance Sampling (SIS), to assimilate probabilistic flood maps (PFMs) derived from SAR data into a coupled hydrologic-hydraulic model with the assumption that the rainfall is the main source of uncertainty. This resulted in a reduction of the forecast errors; however, the improvements were not systematic: for some cases the updated hydraulic output deviates from the observations. The reason for such outliers could be the assumption that rainfall represents the dominating source of uncertainty together with satellite observation errors, excluding other possible sources of uncertainty in the model system. In this study the authors have decided to carry out a similar experiment but this time in a controlled environment so that rainfall is actually the only source of uncertainty.

Secondly, Hostache et al. (2018) highlighted that degeneracy may be a major issue of PFs. To overcome issue Hostache et al. (2018) used a site-dependent tempering coefficient which inflates the posterior probability. In this study, the authors adopted an enhanced tempering coefficient. The latter is a function of the desired effective ensemble size after the assimilation. The adapted method is compared to the standard

method where only one particle is left after the assimilation.

The manuscript is of societal significance, as floods represent one of the major global natural disasters. Therefore, the importance of this topic is to develop a reliable and cost-effective flood forecasting model, however, it is questionable whether this is achieved. In general, the improvement of model predictions is critical in reducing future material and immaterial damage caused by flooding. However, I found the relative contribution of this study in the improvement of flood model predictions unclear. It is indistinct how significantly this study contributes to improvements in global flood predictions. Whether the findings of this research are useful in other catchment areas over the globe is not defined.

Overall, I found the methodology and the results of this study worked out well. The stepwise approach is clear, and the representation of the results is very interesting to read. In the contrary, I found the introduction quite long, and very detailed. In my opinion, it difficult to find out what the broadly interesting knowledge gaps are. I suggest the introduction to be focused on the importance and contribution of this study to the appliance of flood model prediction on a global scale, and how that is achieved. Overall, this study fits well with the scope of the journal and can be published when several relatively minor issues are addressed. Below I provide more detailed comments.

Firstly, the introduction is very detailed on background information, but lacking in the focus of this specific research. The background information consists of a description of Data Assimilation (DA) and Synthetic Aperture Radar (SAR) images which is an essential part of the study. Different assimilation methods are discussed in detail, but it is not clear which and why this method is expected to be used. The detailed treatment of all the background information causes an unclear overview of what the study actually is about, as several DA methods are mentioned (KF, 4DVar, PF, EnKF (line 42-58). The introduction does not lead to a specific research question or clear objective. The objective is mentioned twice (line 74 and line 91), but the choice of words is different, which causes confusion. The objective is mentioned, but a briefly description on how

the objective is achieved is missing. The referencing of previous studies is used to discuss this, but it is unclear which references are really used for the methodology. Now, only the last sentence of the introduction is stating that 'a sensitivity analysis of the DA framework with respect to the tempering coefficient is conducted', which is rather vague.

For this major argument I would recommend being more specific with referring to previous studies, and to have a critical look on the broad background information. Especially, some of the referred studies concerning KF, 4DVar and PF seem unnecessary to me. Try to aim for a narrowing of the introduction, so that the introduction leads to the objective and research questions of this study. I would remove the objective mentioned in line 74, as it is to subtle and not agree with the objective stated in line 91.

The second argument concerns the conclusions of this study. Overall, I am very content with the conclusions. The conclusions of this research focus on the specific study area of the River Severn (UK), which is logically in line with the objective of this study. In line 376 it is stated that the main issue of using SAR observations in flood forecasting models is the difficulty of detecting flooded areas for specific cases, such as urban or vegetated areas. For now, this study only seems to be applicable for this specific study area, but I wonder if that is really the case. The societal significance of this study would be large if it contributes to a global improvement of flood modelling. I would recommend discussing the use of the findings of this study on a global scale. As a reader I would like to know how these results improve the appliance of Sar observations for different types of land use, or if a significant error increase is expected when the analysis is performed for different types of landscapes or land uses.

Thirdly, I did not really understand what is described in paragraph 2 of the methodology (line 139-144). In a previous study by Giustarini et al. (2016) the prior probabilities were proposed to be 0.5 as default value. The new methodology will probably lead to more area specific, but this is not confirmed by a reference. I am wondering if this method has been used before, or how the authors came up with this approach. The

difference between the 0.5 default value defined by Giustatrini et al. (2016) and the derivation from the true binary map in this study, could lead to significant differences in the chances of pixels being flooded or non-flooded. The improvement of the calculations by these methods would really confirm if it is valid, yes or no. To be short, I agree with the use of true binary maps (from true rainfall) to make a better estimation of the probability for each pixel of a SAR image, as true data is directly used to validate the SAR observations. However, I would recommend giving further details about the reasoning for choosing this methodology. There is no clear reason for using a more complicated method instead of the method by Giustarini et al. (2016). Below, some minor arguments and/or issues that I found in the manuscript have been described.

Minor argument 1: In line 68-69 it is mentioned that there could be other possible sources of uncertainty in the model system. I suggest some examples, as for now it is unclear in what direction these other sources uncertainties have to be found.

Minor argument 2: The spilling of water into the floodplain is modelled with a 2D diffusion wave scheme neglecting the convective acceleration (line 114). I would like to a reference or reasoning for the neglection of the convective acceleration. Even though it is logical, an assumption has been made about whether this term can be neglected, yes or no.

Minor argument 3: In line 118 the true meteorological is defined as temperature and rainfall. However, it is mentioned that the true rainfall data is used in the hydrological and hydraulic model. It is therefore unclear why temperature is taken into account in the true meteorological model. I would recommend giving a clarification on this by indicating how temperature is used in the model or remove it if it has not been used at all.

Minor argument 4: In this study, the alpha value is based on the desired effective ensemble size (EES) (line 222). It is unclear if this method has been used in previous studies and what the expected outcome of the use of this EES would be. I recom-

mend adding some detail about this in the method section as it is a major change in comparison to the previous study by Hostache et al. (2018).

List of minor issues Introduction Minor issue 1: Line 68-69: It is mentioned that other sources of uncertainties could influence the model system. I think it is interesting to mention what these other uncertainties are.

Methods Minor issue 2: Line 116: "No later inflow in...". I assume this is incorrect. Shouldn't this be 'latent inflow'?

Minor issue 3: Line 216: "Since ïĄą and weights have values are lower than one", missing "that" before "are".

Results Minor issue 4: Line 274-276: It is unclear if this is correct. The revisit time is around 3-4 days, which means 2 satellite images per week. I do not understand how this results in four assimilated synthetic observations in a period of 10 days. 2 images per week $\approx$ 3 images per 10 days?

Minor issue 5: Fig 3 & 4: I have not read what the pixel size of the SAR observation is. I think it is important to mention the pixel size of the SAR observations by the Sentinal-1 satellite.

Minor issue 6: Fig 6: Labels incorrect. In the graph of streamflow time series (left) the assimilation of image I is indicated four times, while in the image on the right the labels assimilation of image I, II, III, IV are given.

Minor issue 7: Line 301 & Fig 7: "higher than the orange ones...". As the comparison between over detection (red) and under detection (black) is probably meant, "orange" should be "black".