Hydrol. Earth Syst. Sci. Discuss.,
https://doi.org/10.5194/hess-2020-403-AC4, 2021

Hydrology and Earth System Sciences
Discussions

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

Concetta Di Mauro et al.

concetta.dimauro@list.lu

Received and published: 8 February 2021

Reply to the review nr.2

February 2021

We would like to thank referee for the careful reading and the very useful comments. Below we address the referee comments and explain how the manuscript will be updated. Comments of the referee are in italic and bold font.

This is a highly technical manuscript focused on assimilation of many different data sources using multiple techniques to predict flood extent and depth. I think this is an interesting study, but I overall I think it needs major improvements before it can be published. The science is sound and interesting, but the manuscript could be clarified and re-vised throughout to make this easier for the reader to understand. I summarize my major comments and minor comments below.

We will clarify the manuscript as suggested by the referee.

1 INTRODUCTION

1. ***My main recommendation to the authors is to clearly clarify the contribution of this study to the literature. The manuscript incorporates many technical methodological assessments, but it is not always clear why these as-***

sessments are being conducted, and what they help us learn about flood modelling. The authors should clearly state their contributions in the introduction and clarify in a discussion section how their findings advance those conducted by other studies.

This study is a follow up of Hostache et al. (2018) who applied a Particle Filter (PF) assimilation to a real case study at the River Severn which however, resulted in an overestimation of streamflow in some cases. In this study we identify the reasons for this overestimation using synthetic experiments in order to be able to exactly control the assumptions, and more generally assess the strengths and limitations of the method. Additionally, we improve their method to overcome degeneracy issue, and evaluate the effect of pixel misclassification in SAR observations on the DA performance. We will now more clearly state the contributions in the introduction and add a discussion section to spell out the advances of the paper. Further explanations are provided later in this document (paragraphs 4.1 and 4.2).

2. ***The introduction should be revised and reorganized. At current, the introduction is very technical, and describes a lot of the existing literature. However, I had a hard time following the common threads and major points being made across the arc of the introduction. Many individual references are described but aren't necessarily connected to the bigger picture of flood modelling.***

   We thank referee for these relevant remarks and the manuscript will be updated accordingly.

   (a) In the first part a brief introduction to DA of satellite observations is made with the following key points: the importance of flood forecasting (19 -24 lines of the manuscript), the advantages of using satellite observation and adequate DA techniques (25 -32 lines of the manuscript), the SAR image

acquisition characteristics and the information on flood extent that can be extracted and assimilated into a model (33 -36 lines of the manuscript).

   (b) In the second part of the introduction we focus more specifically on the existing methods for assimilating flood extent maps into forecasting models. Most of them (36 -41 lines of the manuscript) transform the flood extent information into state variables of the model such as water levels or discharge [e.g. García-Pintado et al. (2015), Matgen et al. (2010), Revilla-Romero et al. (2016), Giustarini et al. (2011), Hostache et al. (2010)], other do not require this transformation into a model state variable and thus allow directly assimilating backscatter or probabilistic flood maps [Lai et al. (2014), Revilla-Romero et al. (2016), Cooper et al. (2018), Cooper et al. (2019), Hostache et al. (2018)] (lines 41-57of the manuscript). These direct assimilation techniques have been proposed only very recently and there are still numerous open research questions to answer before they can be applied routinely.

   (c) In the third part of the introduction, we indeed focus on the study of Hostache et al. (2018) that requires further investigations to enable a better understanding of current limitations and strengths. Some main differences between the Particle filter and other assimilation techniques are defined (lines 58-61 lines of the manuscript) to explain the choice of the Particle Filter in Hostache et al. (2018) and in the current study. More details (lines 68 -75 of the manuscript) are provided on the Hostache et al. (2018) method and on the assumptions made on the rainfall as the main source of uncertainty. We think that such information is relevant in the introduction as it introduces the reasons that have led the authors of this manuscript to conduct a synthetic experiment. The remaining lines define two main issues that are being addressed in the manuscript: lines 76 - 80 introduce the degeneracy problem and the different methods adopted in the paper (standard and adapted methods), lines 81-90 describe the issue of misclassification of SAR pixels.

The introduction will be shortened in accordance with the referee's recommendations. Less detail on the existing literature (from 42 – 57 lines of the manuscript) will be given. Moreover, we agree that the paragraph on lines 58-61 could create some confusion as it is not completely aligned with the rest of the introduction and will therefore be removed. With these adjustments, we believe that the reader can better follow the main reasoning and the major points of the introduction.

3. ***More synthesis is needed across these references and paragraphs to highlight the major knowledge gaps. Furthermore, I'd recommend shortening the introduction.***
In the literature, there are very few studies directly assimilating SAR-derived flood inundation information into a forecasting model. The most commonly used method is the transformation of SAR-derived information into a state variable of the model, namely into water levels. We argue that the methodology here presented is very novel and that the method proposed by Hostache et al. (2018) needs further investigation to better understand its current limitations and strengths and to therefore assess its applicability at large scale. We agree that the introduction should be shortened. Revilla-Romero et al. (2016), Lai et al. (2014), and Cooper et al. (2019) are some examples of studies where the information derived from SAR data is not transformed into a state variable of the model before being assimilated. To make the introduction more focused, we will condense this part and remove unnecessary details about these methods. From line 58 up to line 61, some differences between PFs and other assimilation techniques are mentioned. This part will be removed.

4. ***Finally, the introduction section normally concludes with a statement about the novelty of the study, the scope, and the objectives. These are instead***

first introduced on line 75, then again later in the introduction. I'd recommend consolidating these statements into a coherent paragraph at the end of the introduction.
We fully agree that objectives are defined with two separate statements. The paper will be modified accordingly with a "coherent paragraph" at the end of the introduction pointing out these objectives:

   (a) The main objective of the manuscript is to evaluate the strengths and limitations of the DA framework proposed by Hostache et al. (2018) with a synthetic experiment where rainfall, together with SAR observations, are the only sources of uncertainty.
   (b) The second objective is to further develop this DA framework for combating degeneracy more efficiently.
   (c) The last objective is to evaluate the effects of misclassification in the SAR-derived observations on the performances of the PF.

5. ***At the end of the introduction, I am left unsure of the scope and objectives of the manuscript (for instance, nothing about SAR or flooding is mentioned). These three concluding sentences could benefit from more specifics as to what will be tested and explored in this particular article. Specifics, such as types of model used, data resolution, etc could be specified here, to more clearly articulate to the reader the framing of your particular work.***
We will add a brief statement at the end of the introduction section, where we state some specifics as pointed out by the referee, but more details are given in the methods section: "the proposed forecasting system consists in a loose coupling of a hydrological (SUPERFLEX) and a hydraulic model (LISFLOOD-FP). The meteorological data are derived from the ERA-5 archive with a spatial resolution of 25 km and a temporal resolution of 1 hour. The SAR data are synthetically generated with a resolution pixel spacing of 75 m. Experiments are carried

out to evaluate the standard and the enhanced version of PF, and the effects of pixels misclassification of SAR observations on the DA."

2 METHODS

1. ***The methods section is very detailed (which I appreciate). Yet, I had a hard time understanding the major comparisons to be made in the results/discussion section Could you more clearly summarize these and why you are comparing these methods at the start of this section?***
   In the section 2 (from line 98) we will add a paragraph summarizing the different experiments:
   (a) The standard filter where degeneracy occurs;
   (b) The adapted method where a tempering coefficient is used to avoid degeneracy. A sensitivity analysis of the tempering coefficient is realized. Different tempering coefficients based on the desired effective ensemble size after the assimilation (5%-10%-20% and 50%) are used;
   (c) The proposed methods are also applied with known errors in SAR image classification in order to evaluate and understand the impact of these errors on the DA.
2. ***The workflow is helpful, but with the number of methods and acronyms, I had a hard time following this.***
   The workflow, (figure 1) is the same for all these tests. The "assimilation" represented by a blue circle is the only element changing between the standard filter and the adapted filters. We will explain the acronyms in the flow chart caption to make it more easily readable.

3. ***The Study Area section comes after the methods section – this was a little confusing to me, because the nuances of this are discussed in the methods section. Is it worth switching the order of these?***
   The reason why we have put the study area after the methods section is due to the fact that in the methods section we give a more general overview of the methodology (not related to the study area) and to show that it is applicable to many cases, whereas the study area is more specific, closer to our particular situation. We will ensure that the method part is free from site-related discussion of information.
4. ***Please ensure that all paragraphs are 3+ sentences, and ensure that these are appropriately combined throughout the text.***
   We will pay attention to this and ensure that all paragraphs contain at least 3 sentences.

3 RESULTS

1. ***I would recommend relabelling sub-sections within the results to separate out the different comparisons and techniques you are making – organizing these headings would help me connect what you do to your methods section.***
   We will restructure the results section in order to match the structure of the methods section.
2. ***For instance, I had a hard time connecting these results to the stated objective of detecting uncertainty in precipitation, and then to the conclusions section. It could also help to start each sub-section by describing what methods/approaches you are testing and why, given there are many comparisons.***

As recommended by the referee, we will re-organize the results section and will also carefully explain at the beginning of each subsection of the results section what will be tested and compared, in order to make the section clearer.

4 CONCLUSIONS

1. ***This may be a personal preference, but I would recommend shortening the conclusions section, and moving much of what is in there now to a discussion section.***
   We agree with this suggestion and we will revise the paper accordingly (see paragraphs 4.2 of this document).

2. ***Within this discussion section, the main piece I don't see is a discussion of the limitations of this approach.***
   The limitations, that are currently reported in the conclusions section, will be moved to the discussion section:

   (a) "Even though larger effective ensemble size prevents degeneracy, results are at the same time less accurate and performances of the predictions are degraded (line 387 of the manuscript)" . . . which is maybe due the fact that. . .. "in this study the tempering coefficient is applied only to flatten the likelihood" and the full tempering scheme in not applied (line 416 of the manuscript).

   (b) The study shows the validity of the DA framework if the uncertainty derives only from the precipitation and the SAR observations. Additional sources of uncertainties could be possibly considered for some real cases: it could be the case of Hostache et al. 2018 (line 397-399 of the manuscript).
   In the discussion section, we will also follow the referee's suggestions and

add the following remarks:

   (c) ***Is there a reason to think that this approach is transferable? Why or why not?***
   The proposed DA framework can be applied to a variety of flood inundation forecasting chains since the model updating is carried out via a sequential importance sampling only (i.e. importance weights).

   (d) ***In what scenarios is this approach more useful (i.e. at what scale)?***
   We argue that the method used in the manuscript has the potential to support EO-based modelling in sufficiently large floodplains where flood inundations remain present over a sufficiently long time period to be detectable to satellite sensors given their revisit interval. Indeed, this constraint must be satisfied to enable the application of the proposed framework and to make use of the analysis carried out in this manuscript.

   (e) ***Given that rainfall is the main source of uncertainty, what does this mean for future work?***
   For those cases where rainfall represents the main source of uncertainty, for example in poorly gauged or ungauged catchments or in forecasting models, our study results indicate that the application of the approach described in the manuscript will lead to improved results of the model simulations. Some modifications of the DA framework are still required to fully overcome the issue of degeneracy. For those cases where the uncertainty of other sources is more relevant, these sources need to be taken into account explicitly. Possible ways to adapt and advance the proposed DA framework are currently under development (e.g. updating a state variable of the model, using an enhanced version of the adapted filter).

   (f) ***Can this work improve forecasting?***
   It is shown in the paper that the assimilation is beneficial as it reduces fore-

cast errors not only at the assimilation time steps, but also for subsequent times steps. The persistence in time of these "improvements depends on the flashiness of the flood event, i.e. "the rapidity with which hydrological conditions change" (line 389 of the manuscript). More frequent image acquisitions could help in keeping model predictions on track, especially when the dynamics of the system are fast. In future studies, we will evaluate if additionally updating the state variable helps in obtaining longer-term positive impacts on the simulations.

5 MINOR POINTS

1. ***Line 42 – "used" is repeated Line 48-49 – I had trouble understanding this sentence – could you rephrase? It was not clear to me what 'the latter' referred to:***
   We will change this sentence as follows: "In the existing literature only few studies have used DA for assimilating flood extent maps into flood forecasting models [e.g. Lai et al. (2014), Revilla-Romero et al. (2016), Cooper et al. (2018), Cooper et al. (2019), Hostache et al. (2018)]. The reason is the difficulty of directly assimilating flood extent since this is not a state variable of the model. Consequently, in many assimilation studies the flood extent information is transformed into water level as this is a state variable of most hydraulic models."

2. ***Line 42 – I am missing the connection from this paragraph to the next - why would one want to use a KP, 4DVar, or PF technique for assimilation of flood information? Can you connect these thoughts to the previous sentence?:***
   These assimilation techniques are mentioned because they are used in the cited references [Lai et al. (2014), Revilla-Romero et al. (2016), Cooper et al. (2018), Cooper et al. (2019), Hostache et al. (2018)] (lines 43 -44 of the manuscript). The

paragraph could be restructured as follows: "In the existing literature, only few studies have used DA techniques, such as Kalman Filter (KF), Four-Dimensional Variational (4DVar) and Particle Filter (PF), for directly (without any transformation into a model variable) assimilating flood maps into flood forecasting models".

3. ***Line 52 – is there a reason to have a whole paragraph focused on this particular article? Is it most similar to what is done in this study? Do you improve on their work? If not, I'd recommend shortening the description of this article:***
   The reason why we are mentioning the article is to give an example of studies where the information derived from the SAR image is not transformed into a state variable of the model before the assimilation. We want to highlight the fact that "the new observation operator performs well compared to the assimilation of flood-edge water elevation observations". In the context of our study, it means that using information derived from SAR without transforming it into a model variable is an interesting option. The advantage of this technique is its applicability for operational uses in near- real time assimilation. Moreover, we agree that this paragraph should be reduced in the revised version of the paper. Therefore, it will be changed as follows: "Cooper et al. (2019) have used an EnKF to update a 2D hydrodynamic model. In this study, the backscatter values are directly assimilated into the model. The study has shown that the SAR backscatter-based assimilation method performs well compared to the EO-derived water levels assimilation."

4. ***Lines 76 – 90 – this is very detailed, to the point where I am unsure if this is helpful in the introduction. Would you be able to shorten this section and distil a few key messages? Could this be moved to the methods section?:***
   As suggested, we will shorten this paragraph and move some technical details to the methods section as follows: "Moreover, Hostache et al. (2018) highlighted that degeneracy may be a major issue of PFs: after the assimilation, the number

of particles with significant weight reduces significantly to few or one particle so that the ensemble loses statistical significance. To overcome this issue Hostache et al. (2018) used a site-dependent tempering coefficient which inflates the posterior probability. In our study, we adopt an enhanced tempering coefficient defined as a function of the desired effective ensemble size after the assimilation. Moreover, there could be errors in the detection of flooded areas in SAR images. Detecting and removing these errors represents one of the main scientific challenges of using SAR data for a systematic, fully automated, and large-scale flood monitoring. In Hostache et al. (2018), speckle errors are therefore taken into account, but no conclusions are given on the effect of misclassified pixels in the SAR observations. Consequently, in this synthetic experiment, misclassification errors are artificially added to the SAR-derived flood extent with the aim to assess the robustness of the proposed method with respect to this type or errors."

5. ***Line 178: "supposed to be uniform" – do you mean assumed to be uniform? Sampled as uniform? Please clarify:***
   Yes, we meant "assumed to be uniform", meaning that each particle has the same weight before the assimilation. This will be corrected.

6. ***Section 2.3 – please weave the equations into the text, instead of listing them after the text here:***
   We will do so.

7. ***Section 2.4 – please do a thorough read to ensure that all variables in the contained equations are clearly defined in this section:***
   We will do so.

8. ***Lines 228 – 232 – this reads as 'results' – should this be moved to the results section?:***
   Lines 231 - 232 will be rephrased and moved to the conclusions section: "This

methodology leads to slightly biased estimates because the observation are down-weighted."

9. ***Section 3.0 – please capitalize 'area':***
   We will take this into account.

10. ***Line 276: The plots in this section show four time points –why did you select these time points? Please introduce the time points in this section:***
    Over Europe, the current revisit time is around 3- 4 days, which means 2 Sentinel-1 satellite images are acquired on average every week. The revisit time for a single orbit is 6 days but in our case study we are considering many orbits (ascending and descending) so the revisit time will be 3-4 days as shown in the enclosed image. The acquisition dates used in the manuscript were derived from a realistic acquisition plan of SENTINEL-1 over the area. They correspond to: 22-07-2007 10:00, 24-07-2007 17:56, 25-07-2007 17:49, 28-07-07 07 09:00. Please note that with another satellite being added to the constellation and other satellite missions being considered the revisit time as well can be further shortened.

11. ***Line 285: You show a sub-section of the result area multiple times – please introduce this area and why you selected it in the text. Also – are you computing results for just this section of the river or the entire watershed? I wasn't sure from the methods and study area section. Please clarify:***
    This sub-area represents the domain of the hydraulic model. We compute and compare results along the main River Severn within this sub-area of the watershed which represents the flood-prone zone. The hydrological model with which the boundary conditions are evaluated covers the contributing upstream catchment shown in figure 2 of the manuscript.

12. ***Line 274 – 284 – should this be in methods?:***
    Yes, we agree and will move this paragraph in the method section.

13. ***279 – 281 – what is the significance of this? Could you explain more why you mention this here? This again seems like 'methods' – should this be moved to the methods section, or is it a 'result' of your investigation?:***
Yes, this will be moved to the methods section. In Giustarini et al. (2016), the prior probability is assumed to be 0.5 since no information on the prior can be obtained. In this paper, however, because of the synthetic nature of the experiment, the prior is known as it can be derived from the true binary flood extent maps. We have carried out experiments with a default value of the prior (0.5) and with the estimated prior and found that its value has no significant effect on the results of the experiment, as explained in lines 279 -283 of our manuscript.

14. ***Line 284 – Figure 3 and Figure 5 are mentioned – figures should be listed in order. Figure 4 is not cited in the text. Should this be removed or moved to Supporting Information?:***
This is a typo because it should be written as follows: "the PFMs are shown in figure 4 and the corresponding reliability plots in Figure 5". This will be corrected.

15. ***Line 315 – Please do not start a sentence with a number:***
We will take this into account.

16. ***Line 387 – 399 – could you rephrase this sentence? I don't understand what it is saying:***

    (a) Although the use of a smaller tempering coefficient leads to a larger effective ensemble size (e.g. 50 %) and helps avoiding degeneracy, the results are less accurate compared to the standard method or a 5% EES method.

    (b) The persistence in time of the beneficial effects of the assimilation varies according to the rapidity of variations of flood extent; a more frequent image acquisition could help in better keeping the predictions on track.

    (c) Our study further shows that it is important to characterize and mask errors in the SAR observations. A large number of misclassified pixels substantially

degrades DA performance. In our study, the improvement of model simulations (water levels and streamflow) and performances (CSI and RMSE) after the assimilation is still possible if the errors in the SAR observations are rather limited (not more than the 20% of the pixels). However, if the misclassification goes beyond 40% of the pixels, the assimilation has no effect or even degrades the model predictions.

    (d) The results confirm the validity of the DA framework when the hypothesis of the rainfall as main source of uncertainty is verified. This confirms that the limitations identified in the previous real case study by Hostache et al. (2018) could be explained by additional sources of uncertainties that were not taken into account.

17. ***Figure 2: Could you highlight on this figure the places you select for Figure 3 and Figure 4? Figure 3 and 4: The legend is hard to see, and there is no label of what 'value' is being shown (and its associated units). Figure 3 and 4: What are the four assimilation time steps? Please label these figures as (a)-(d) or on the figure to indicate this. Figure 3 and 4: Should these be combined to enable comparison? It is not entirely clear from the results text what these images show and how these connect to the workflow. Table 1 and Table 2: Please direct readers to Figure 6 in the captions for these:***

    The remaining corrections of the figures will be made according to the referee's suggestions.