Response to reviewer 2:

The authors thank the reviewer for their useful comments, which will help improve the manuscript. Below we give each comment in bold (abridged where appropriate) and describe how we plan to alter the manuscript to address the reviewer's concern. We give suggested changes to the manuscript in italic font.

1. **References from Oubanas should be included in the state of the art when referring to hydrodynamics data assimilation, with remote sensing data.**
   The authors thank the reviewer for pointing out these papers and will add these references to section 1 of the manuscript.

2. **Front like observations were assimilated in the framework of wildfire propagation. Coordinates (x- and y-) of along edge markers were used as observation, avoiding the non-gaussian issue of binary data (burn or unburned area, dry or wet area). This work should be cited in the references as it proposes an alternative to the 3 options presented here**
   We will add the references as suggested to section 3.3.

3. **ETKF algorithm should be presented in more details even though this is a classical algorithm and references are given. The choice of the perturbation matrix is essential in this deterministic filtering algorithm and for the present paper to be self-dependent, a short description of how this is done should be included.**
   We will add further details of the matrix used to update the perturbations from p.5 line 23:
   'The perturbation matrix is updated by the matrix $\boldsymbol{T} \in \boldsymbol{R}^{M\times M}$. We use an unbiased, symmetric square root formulation of the matrix $\boldsymbol{T}$, constructed in a way that ensures that the analysis state error covariance, $\boldsymbol{P}^a = \boldsymbol{X}^a (\boldsymbol{X}^a)^T$ is the same as the analysis error covariance calculated in the Kalman covariance update (in e.g. Kalman (1960)). The formulation makes use of a singular value decomposition (Golub and Van Loan, 1996),

   $$(\boldsymbol{R}^{-1/2}\boldsymbol{Y}^f)^T = \boldsymbol{U\Sigma V}^T,$$

   where $\boldsymbol{U} \in \boldsymbol{R}^{M\times M}$ and $\boldsymbol{V} \in \boldsymbol{R}^{p\times p}$ are orthogonal. The columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ are the left and right singular vectors of $(\boldsymbol{R}^{-1/2}\boldsymbol{Y}^f)^T$ respectively. The diagonal elements of the matrix $\boldsymbol{\Sigma} \in \boldsymbol{R}^{M\times p}$ are the singular values of $(\boldsymbol{R}^{-1/2}\boldsymbol{Y}^f)^T$. A solution for $\boldsymbol{T}$ is then

   $$\boldsymbol{T} = \boldsymbol{U}(\boldsymbol{I} + \boldsymbol{\Sigma\Sigma}^T)^{-1/2}\boldsymbol{U},$$

   where $\boldsymbol{I}$ is the identity matrix. See Livings et al. (2008), Cooper et al. (2018) for further details of how $\boldsymbol{T}$ is computed.'

4. **Please justify why using a deterministic filter ETKF instead of a stochastic EnKF?**
   The authors follow the approach of Garcia-Pintado et al. (2013), Garcia-Pintado et al. (2015) and Cooper et al. (2018) in using an ETKF for a similar application.

5. **Part 3.1 should include Figures 7 and 9.**
   Figures 7 and 9 contain information relevant to the Results and Discussion section which is not relevant to section 3. Nevertheless we will add text in section 3.2 to cross-reference the figure:

   'Figure 7c shows the location of the observed flood edge and the corresponding nearest wet pixel for a simplified case. '

6. **Assuming that the water level is constant perpendicularly to the flow is essential for the second observation operator hnp. While this was mentioned at the end of 3.2 relating to other published papers, this should be mentioned earlier when presenting the operator along with the related difficulties (such as finding the nearest point).**
   In our view it is important to describe the operator in section 3.2 before discussing problems with its practical implementation (these are already discussed later in section 3.2).

7. **For the backscatter approach, operator h_b associates either m_d or m_w to the model equivalent at the observation point. This means that operator h_b only returns 2 possible values that are compared to the entire range of backscatter values. As a consequence, any wet pixel in the model state (for instance different WLO values for different members) would return the same equivalent, and the difference between members is lost. I feel, we are losing information in the ensemble here. Yet, I may be missing a point here, so please clarify.**
   It is indeed a feature of the binary backscatter observation operator that ensemble predictions of water depth are converted only to predictions of wet or dry; this is because they correspond to backscatter observations which contain no information about water depths. This is already mentioned as a potential drawback on p. 21 lines 15-29; in situations where all the ensemble members agree a cell is wet (or dry) no update can be generated even when the observed water level is different to the mean forecast level. However, as mentioned in the new Discussion section (see response to reviewer 1, comment 1) this feature is potentially beneficial as it means the method is robust to outliers and will not update the forecast when, for example, pixels very far from the river are wrongly classified as wet based on backscatter value.

8. **While I am not questioning the use of the Clawpack model in this work, I am curious to know why the authors did not use a community model such as LISFLOOD, MIKE or TELEMAC.**
   We chose Clawpack as the code is open source, available for Linux, and uses robust, accurate and efficient numerical solvers which are able to deal effectively with shocks in the solution. In addition, this work builds on our previously published study Cooper et al (2018).

9. **Some details on Clawpack model may be included here again for the paper to be a little more self-dependent: is it a full 2D model? (is the water level constant perpendicular to the flow direction ?)**
   We already stated in section 4.1 that the 2D shallow water equations are solved, and will add *'everywhere in the domain'*

10. **How are the limits of the simulation domain prescribed? (solid boundaries?)**
    We will add the following to section 4.1:
    'In our simulations the boundary condition is extrapolating (outflow) on the y=0 boundary; all other boundaries are solid wall. '

11. **The ensemble construction relies on the perturbation of a true inflow with additive time correlated signal, assuming that the correlation length is large. Why not simply using a scalar additive perturbation constant over time, as it basically comes down to the same result?**
    This would give a similar result; we chose the approach used here, in which perturbations depend on the flow, based on a similar method used to generate inflow ensembles in Garcia-Pintado et al. (2013), Garcia-Pintado et al. (2015) and Cooper et al. (2018).

12. **Cycling of the analysis requires explanations on how the friction coefficient is updated along the analysis. First, it is not clear to me whether the analysis is carried out at an instant of observation or over an assimilation window (like a smoother would be).**
    We already stated on p4, line 6 that the ETKF is a sequential technique. Assimilations are therefore carried out at observation times. We will add the following sentence on p4 line 11 to clarify:
    'We use the ETKF in its standard application as a sequential filter. As such we perform an update step at the time of each available observation.'

13. **Secondly, it is mentioned that the friction coefficient is drawn from a normal distribution with mean different from the true friction coefficient. But how is the analysed value of the friction used for the following cycle?**
    We already stated in section 2.2 that the friction parameter follows the same update-forecast cycle on p.6 line 13: 'The augmented state vector is updated by the ETKF algorithm through equations (10) and (14). Parameter value(s) are updated according to the observations due to covariances between errors in the model state and errors in the parameter(s).' We will change the text at p11, line 19 to make this clearer:
    'For the initial forecast step, a value of n_{ch} for each forecast ensemble member was drawn from a normal distribution with mean, μ, that is different to the true value and standard deviation σ. This imposed bias in the forecast ensemble channel friction parameter means that we can test how well data assimilation with different observation operators can correct the forecast state and parameter value towards the truth. In our state estimation experiments, the value of n_{ch} assigned to each ensemble member remained constant throughout the simulation. For joint state-parameter experiments, the values of n_{ch} were updated at each assimilation time through

the ETKF equations, as described in section 2.2. Using an incorrectly specified channel friction parameter in the forecast is realistic, as the true value is unlikely to be known in operational situations. Initial forecast channel friction parameters are randomly drawn.....'

14. **Is the friction drawn from a normal distribution with a mean equal to the mean analysis? How about the standard deviation?**
We will make this clearer in the text - see response to comment 13. The initial values for the ensemble parameter values are drawn from a Gaussian distribution – see response to reviewer 1, comment 11, for clarification of the distribution characteristics. For analysis values, the perturbation matrix for the augmented case includes the friction parameter perturbations; this follows from equations (15) and (2). The friction parameter perturbations (and therefore the standard deviation of the parameter distribution) are therefore updated through equation (14) in the same way as for the state perturbations.

15. **Is there any inflation on the model parameter to avoid ensemble collapse?**
We already state on p. 15 line 5 that we have not used any inflation; this applies to both parameter and state perturbations. As in *Cooper et al. (2018)* we did not observe ensemble collapse in this simple system.

16. **Is the corrected value of the friction kept persistent for the forecast?**
Equation (16) already showed that the friction values are not changed during the forecast step.

17. **I suggest adding a scheme to properly explain the ensemble cycling in part 4.3.**
The standard ETKF scheme is used for the ensemble cycling. Our responses to comments 12 – 16 clarify this.

18. **In the synthetic observations part 4.4, I understand that given the WLO in a flood edge pixel, a backscatter value is drawn from a normal distribution centred in md or mw. Why bother computing the Gaussian fit and new Gaussian values when these observations are going to be compared to binary values (equivalent model state values that are either md or mw)?**
We use the variance of the distributions to provide information about observation uncertainty as already stated in section 4.5.

19. **The location of the observation is not clear to me: in 4.4, it is said that the flood edge is defined to be the elevation at the first 'dry' pixel encountered when moving in a perpendicular direction from the centre of the channel along one of the defined cross sections then it is said that two observations per flood edge are considered. Please clarify and locate the observations in Fig 3.**
Reviewer 1 made the same point. We will add a schematic to make the observation locations relative to the flood edge clearer and reword p14. Lines 4-9 – see response to reviewer 1, comment 2.

20. **I couldn't find information of observation frequency while it is mentioned that the assimilations are carried out every 12 hours. This goes back to my previous question on instantaneous or time window assimilation.**
    The observation frequency is the same as the assimilation frequency. We will clarify this at p15 line 4:
    'Assimilations are carried out at 12 hourly intervals. This is currently the shortest likely time between observations due to return times for satellites equipped with SAR instruments.'

21. **I suggest adding rank diagram to check the validity of the ensemble with regards to the observation. This would be a starting point to identify cases where all members WLO are lower than observation as this case leads to a problematic zero correction in the analysis. This would allow for correcting the ensemble (and its spread) beforehand applying data assimilation while being aware of a problem.**
    For a binary operator a rank histogram in observation space would not give meaningful information, as the value of the forecast-observation equivalent for each ensemble member can only have two values (wet or dry). We agree that it would be useful to have a method of checking for filter divergence so that the user can check the observations and model forecasts. The best approach to this would need to be determined by experience with real case studies but this is not within the scope of this study.

22. **The computation of the ensemble mean at the flood edge illustrated in Fig 8 causes a negative effect of the analysis because the members that are shallower than the observation are associated with a zero WLO at the flood edge, thus not contributing to the mean computation. I regret that no solution was proposed in the paper. A suggestion would be to compute the mean WLO at the centre of the river and assume it is constant perpendicularly to the flow. This assumption is made already for the second operator solution**
    We agree that the flood edge operator gives poor results; we consider the nearest wet pixel approach already discussed in section 3.2 to be a solution to this problem. This improved version effectively uses the approach suggested by the reviewer.

23. **I have doubts about Figure 9: the observation is located at the true flood edge, where the innovation is computed. I guess the arrow in the observation space should be translated on the left, above the flood edge. Plus, md and mw are mixed in the right hand side legend**
    We will correct the m_d and m_w labels and move the arrow to the left.

24. **The results for hb are satisfying while difficulties occur when all members are shallower than the truth or reversely. I regret the authors did not propose an alternative to this while being aware of it. I suppose that in a real case scenario, this situation may occur depending on how the ensemble is generated. Thus, I suggest adding ensemble validity check as well as reconsidering the computation of the model equivalent that binaiyly returns md or mw indistinctly of the water level value.**

See response to comment 21 for ensemble validity check. We agree that there are potential problems with applying the backscatter operator to a real case. We propose adding a new ‘Discussion’ section after section 5 to address the problems noted here. See response to reviewer 1, comment 1.

25. **Locating the flood edge seems a difficult task for a real case with a randomly shaped flood surface, also, locating the nearest wet pixel is a complex task in non-idealized cases. I suggest to investigate this topic that is central to all 3 observation operator proposed here.**
We agree that locating the flood edge is a complex task; one of the advantages of the backscatter approach over the nearest wet pixel approach is that this step is not necessarily needed in order to perform data assimilation. We would make this clearer in the new ‘Discussion’ section – see response to reviewer 1, comment 1.