Response to reviewer 1:

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. **I would like to suggest some clarification on the methodology and a further discussion on the potential practical use of the observation operator (3).**
   We will make clarifications to the methodology as described in our responses to the reviewer's specific comments 2 and 4. We will add a new 'Discussion' (section 6) to the paper to address practical application of the new observation operator.

   6. Discussion

   'In this study we have chosen to use a small number of backscatter observations for our experiments. This allowed us to compare updates between the three observation operators when the observation operators were all given equivalent information; in this way we can draw conclusions about the physical mechanisms responsible for the different updates. In a real case, one of the major advantages of using our new backscatter observation operator is that it would be possible to use a large number of backscatter observations compared to the number of water level observations that are typically available. The availability of a large number of observations may be a major strength of our new approach; in our simple experiments (not shown) we found that assimilating a larger number of observations with the backscatter operator provided a better analysis than using only a few. Another merit of the backscatter operator is that there is less processing involved in using backscatter observations directly, potentially reducing the amount of time between acquisition of a SAR image and its use to update an inundation forecast. The backscatter operator also removes the need for locating the 'nearest wet pixel' in the model forecast, which can be computationally costly.

   There are a number of potential problems with practical implementation of the backscatter operator. One is that using histograms to produce SAR-derived inundation maps can lead to errors in assigning pixels to wet/dry categories. One way to deal with this would be to use region growing techniques (see e.g. Horrit et al (2001)) or change detection techniques (see e.g. Hostache et al (2012)) to produce robust wet/dry maps for SAR images, and then perform a quality control procedure to discard any backscatter observations that would lead to mis-classification due to e.g. emergent vegetation. This procedure would remove the advantage of fewer processing steps for the backscatter operator, but may not be necessary. Further research is required to understand how robust the method is to the proportion of misclassified SAR pixels in a real case study. We note that the backscatter operator would not generate an update to the forecast in model cells that all the ensemble members predicted to be dry (or wet) as discussed in the last paragraph of section 5.1.2. This means that SAR pixels far from the river wrongly classified as wet, or SAR pixels in the river channel wrongly classified as dry would not degrade the forecast through an erroneous update.

The new backscatter operator is likely to work well in cases where good separation of the wet/dry distributions can be obtained through a histogram, and less well in cases where the distributions overlap. The new observation operator does not require a digital elevation model to generate forecast-observation equivalents, although the hydrodynamic model would require topography information to generate a forecast. Water level observations cannot be accurately determined in areas with high slope, whereas backscatter observations will be unaffected. Like the other observation operators, the new operator will likely provide better results in rural settings than urban settings; double-bounce and layover effects due to buildings are potential sources of problems for all of the operators (Mason et al. (2018)).

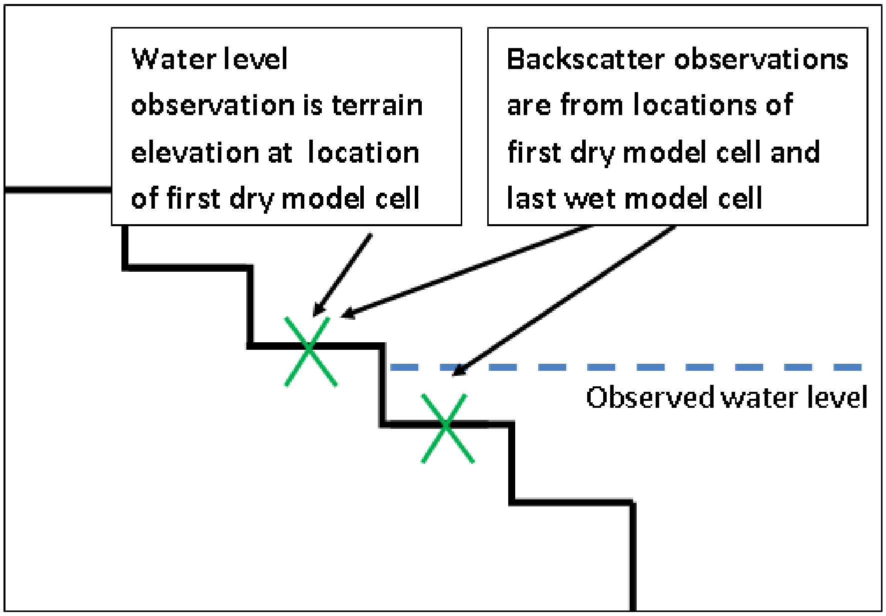
Reference: Mason, D. C., Dance, S. L., Vetra-Carvalho, S. and Cloke, H. L. (2018) Robust algorithm for detecting floodwater in urban areas using Synthetic Aperture Radar images. Journal of Applied Remote Sensing, 12 (4). 045011. ISSN 1931-3195 doi: https://doi.org/10.1117/1.JRS.12.045011

2. **'Firstly, I would like to recommend rewording the sentence in lines 4-9 page 14. The spatial distribution of the backscatter observations is a crucial aspect of this study and I think that this sentence is a bit confusing. The authors might (or might not) consider adding a graphic explanation (maybe by adding details to one of the figures of the manuscript).**

   We will add a schematic to make the observation locations relative to the flood edge clearer and reword p14. Lines 4-9:

   'In this study we wish to investigate the differences in the updates generated by different observation operator approaches. We therefore use equivalent observation information for each of the operators. In the case of the water level observation operators we have used flood edge water level observations at six locations, where the flood edge location is defined as the position of the first dry model cell (see section 4.4). For the new operator we use two backscatter observations for each transect.

   Figure 5 shows a schematic of the locations of the observations we have used in this study, relative to the edge of the flood. In practical applications of the backscatter operator, observations could be used from any location covered by the SAR image.'



Caption: Figure 5: Schematic of observation locations used in this study for each transect in cross section. The thick black line shows the discretised domain elevation, the dashed blue line shows the observed flood water level. The arrows and green crosses show the locations of the observations.

3. **Do the authors think that the merits of observation operator (3) are due to "increased data availability" and/or to the different physical mechanism underpinning the assimilation of backscatter rather than SAR-derived water level values?**
   We will address this question in the new 'Discussion' section – see first paragraph of response to comment 1.

4. **One thing that I could not understand from the manuscript is whether the methodology implied that a pixel with backscatter lower than the mode mw is more likely to be flooded than a pixel having backscatter mw.**
   The histograms give us information about the distribution of backscatter values for wet and dry pixels. The two distributions represent the probability that a pixel has a backscatter value, b, given that the pixel is wet, i.e. p(b|w), and the probability that a pixel has a backscatter value, b, given that the pixel is dry, i.e. p(b|d) . We do not compute the probability distribution that the reviewer asked about, i.e. the probability that a pixel is wet(dry) given its backscatter value, p(w|b), (or p(d|b)) in this method.

5. **As the authors underlined, the number and spatial distribution of backscatter observations used in the data assimilation approach has to be carefully defined. I think that this is a critical aspect for the practical application of the methodology. SAR-derived inundation maps are affected by a large number of uncertainties. For instance, the histogram analysis used here might not be a reliable**

approach in catchments with emerging flooded vegetation where double bounce effects are not limited to the flood edge. The flood edge itself is often the area of largest uncertainty. Furthermore, SAR-derived inundation maps computed using histogram thresholding are often fragmented and further analysis steps such as region growing or use of ancillary data are required to produce a "continuous" inundation layer. These steps are not included in the methodology presented in the manuscript. Would the authors recommend adding these (or similar) steps in a real case scenario?

These questions will be addressed in the new 'Discussion' section – see response to comment 1.

6. **Pixels within flooded areas might have large backscatter due to double bouncing effects, speckle, and other uncertainties. If such pixels are used as backscatter observations, the data assimilation approach will degrade the performance of the flood forecasting model. Is this a possible scenario? If so, how do the authors recommend avoiding this problem?**

   These points will be addressed in the new 'Discussion' section – see response to comment 1.

7. **I understand that this paper focused on a synthetic experiment and I agree with the authors that a detailed investigation of physical mechanisms within a simplified context is essential to develop knowledge and to explore the feasibility of proposed techniques. However, my main comment concerns the effectiveness of the proposed method in a real world scenario. I think that a further discussion on the potential hurdles and possible solutions for the implementation of observation operator (3) in a real case study would facilitate the reception of the proposed approach and encourage its application. The authors might consider adding a short description of the overall characteristics of the real world scenarios for which they think that their method could provide reliable results. An overall description of the characteristics of the real world scenarios for which the backscatter observation operator is not recommended could also provide useful information to the readers.**

   We will address issues relating to practical application of our new backscatter operator in the proposed 'Discussion' section – see response to comment 1. The authors plan to publish results from applying the new observation operator to a real world case study in a separate paper.

8. **Page 9, line 21: I think the second parenthesis after George (2018) should be removed.**

   This will be removed.

9. **Page 10, line 18: a full stop should be added before "Inflows".**

   This will be added.

10. **Page 11, line 19: "ch" in the symbol for channel roughness should be a subscript**

    This will be corrected.

11. **Page 12, line 3: are the intervals of roughness values correct?**

    We will clarify the intervals for the channel friction parameter distributions at p11, line 20:

'The value of n_{ch} for each forecast ensemble member was initially drawn from a normal distribution with mean, μ, that is different to the true value and standard deviation σ......... Forecast channel friction parameters are randomly drawn from a normal distribution with μ = 0.05 and σ = 0.01 for experiments with positive bias in n_{ch} and μ = 0.03 and σ = 0.01 for experiments with negative bias in n_{ch}.'

12. **Page 13, Figure 3: are the values in the x-axis correct?**
    The values will be corrected.

13. **Page 13, Figure 3: low backscatter values are usually represented with dark grey to black, high backscatter values are usually represented with light grey to white. In the colour scale used in this figure, the higher the backscatter value, the darker the pixel. Despite this is just a cosmetic detail, I was wondering whether the authors are willing to reverse the current colour scale to allow a more straightforward interpretation of the figure.**
    The colour scale on figure 3 will be corrected to match real SAR images as suggested.

14. **Page 15, line 12: "growns" should be corrected.**
    This will be corrected.

15. **Page 18, Figure 7c: the authors might consider adding a description of the red arrow.**
    We will add the following text to the caption:
    'The red arrow shows the difference between the observation location and the nearest wet pixel location.'

16. **Page 19, line 12: the authors might consider rewording the sentence "the water level predicted by the observation at the observation location" to improve the readability of the paragraph. More specifically, "predicted" might not be the most appropriate word in this context.**
    We will replace this sentence with:
    'Figure 7 illustrates the fact that the simple flood edge operator cannot produce a useful update when the mean of the forecast ensemble is shallower than the observed water level.'