

Interactive comment on “Calibration of channel depth and friction parameters in the LISFLOOD-FP hydraulic model using medium resolution SAR data” by M. Wood et al.

M. Wood et al.

m.wood@bristol.ac.uk

Received and published: 26 September 2016

General comment 1:

I found many drawbacks mainly due to the paper organization and presentation that have to be solved by the authors before it can be considered adequate for a publication in the HESS journal. Moreover there are some key points that I would like to underline.

One is the justification of the small sensitivity of the roughness parameter with respect to the channel bathymetry that to me seems reasonable but I found some difficulty in understanding if this is supported by a robust analysis or it is not adequately shown in the manuscript. To this end, it would be interesting to see the DYNIA analysis carried

C1

out also for the channel roughness in order to recognize its information content and its value of identifiability. If this was already done, but not shown, some comments or an explaining figure would be very welcome.

Response:

We did indeed carry out the DYNIA analysis on the channel roughness ‘nc’ parameter for all the SAR images used in this study but we found the results for this parameter were not as sensitive as those for the depth parameter ‘r’. Variations in channel roughness value did not produce a strong enough response in the model to match observed flood extent. For this reason these ‘r’ results did not feature in the final version of the paper. The authors have amended the text in section 3.1 to explain this:

“Consequently an important result of this paper is that - in this particular experimental set up with channel roughness parameter ‘nc’ examined simultaneously with the channel depth parameter ‘r’ for the available ENVISAT SAR data - ‘nc’ has a much reduced sensitivity compared with the ‘r’ depth parameter response. It is observed that ‘nc’ will yield optimal results for as long as ‘r’ is also unknown. This lack of sensitivity of channel roughness in this and all subsequent results meant that ‘nc’ could not be identified with any real confidence with this methodology (while ‘r’ is also unknown). So while ‘nc’ analysis was carried out we present here onwards only those results from the more identifiable ‘r’ parameter. ‘nc’ results are now omitted (but can be provided upon request if of interest)”.

To illustrate we show some of the channel roughness results in Figures 1 and 2. Here we show firstly the results for the ‘r’ parameter (red), then ‘nc’ (blue) CSI results, against parameter value for a sample of the SAR images. These CSI plots generally reflect the results we observed in all SAR data and illustrate the greater sensitivity of the depth parameter over the channel roughness parameter in our experimental results.

Additionally, Figure 3 shows the cdf plot of the gradient of the cumulative distribution of rescaled support values for groupings of SAR data, against ‘nc’ value. This illustrates

C2

how different/lower the results are for 'nc' identifiability when compared with 'r' results (paper fig 7, inset).

Lastly, the Information Content (IC) results for 'groupings' of SAR images are also shown here for 'nc' and 'r' in Table 1. These illustrate how much lower the 'nc' information content can be compared with the IC of the depth parameter, for the same combination of data.

General comment 2:

A second point is the assumption that the error related to the processing of the SAR image will not affect the results (not considered for simplicity). I think that these errors are part of the procedure of identifiability and are able to affect the information content of the different images. For this point my question is: due to the different acquisition, times of the SAR images under different atmospheric and land conditions can be considered the error related to the image processing stationary? My opinion is that this error varies from image to image. This at least deserves some discussion. The authors could consider that aerial flood maps for analyzing this point or, since the area is very well instrumented, doing the same type of analysis with stage data.

Response:

This is a good point and certainly each SAR image will have a unique processing error associated with it and indeed the errors inherent in the processed SAR data will be passed on to the final identifiability and IC score for single images. These will not be stationary errors, they will vary between images. For example, while different atmospheric conditions will not significantly affect the radar signal, different incident angles can have an effect. In the paper these were considered to be so small compared with the errors associated with the assumptions around parameter identifiability that they were thought to be overshadowed. Furthermore the use of moderate resolution flood

C3

imagery for hydraulic model calibration may lead to inaccuracies but it was deliberately chosen because we want to understand the usefulness of this data for global locations where other data may quite simply not be available. And while the magnitude of flooding and the land surfaces affected can cause specific errors/uncertainties, this is somewhat mitigated by the larger spatial scales employed in this analysis.

Other errors in the processing of the SAR image can be evident, such as from bias - where in some areas radar data does not inform on the extent of flooding. Ideally, such non-informative areas would be masked out but this requires more comprehensive analysis and is currently an active area of research (e.g. Giustarini et al., submitted: puts forward the idea of flood 'probability' maps to illustrate confidence in the detected flood extent). There is now more discussion about this around Figure 3 in the paper which illustrates the CSI scores for the aerial flood map image that was acquired in July 2007. However since the authors already have some concerns about the length of the document, and the comments of reviewer #2 on the length of the paper also, we attempt to keep these discussions brief.

General Comment 3:

A final point is the quality of all the figures in the manuscript that I found very poor and such that to impede a proper understanding of the manuscript.

Response:

Thank you for this feedback, we have replaced the figures mentioned with the full colour versions and increased the image resolution to improve the quality.

Based on that I recommend publication after major revisions. In the following, the authors can find a list of comments with the associated relevance listed in order of appearance in the manuscript.

C4

COMMENTS:

1:

Pag. 2 Lines 53-76 MINOR: the authors may also cite the work of Moramarco et al. (2013) which uses an interesting method for identifying the flow depth distribution in natural channels.

Moramarco, T., Corato, G., Melone, F., Singh, V.P., 2013. An entropy-based method for determining the flow depth distribution in natural channels. *Journal of Hydrology*, 497,176-188.

Response:

This reference has been added to the paper.

#2:

Pag. 3 Lines 88-91 MINOR: Can you rephrase this sentence more clearly?

Response:

Lines 88-91 have been rephrased to: "In particular the methodology uses flood extent with an accuracy-scoring method that disregards the correct detection of 'no water' pixels"

3:

Pag. 4 Lines 120-128 MAJOR: If I understand correctly the channel depth is expressed as $H=r*B$ where H is the channel depth, and B is the width of the channel. Since the hypothesis of linear scaling is central in the study, I think this part deserves more profound discussion about: 1) How much it will affect the results of the study. 2) Which are the expected problematics associated with the uniform channel depth.

C5

Response:

Extra discussion has been inserted at the end of the manuscript around the limitations of our assumptions and what affect these might have on the results.

4:

See also the paper of Yan et al. (2014) where H is a free parameter of the model uniform along the river reach.

Yan, Kun, et al. "Exploring the potential of SRTM topography and radar altimetry to support flood propagation modeling: Danube case study." *Journal of Hydrologic Engineering* 20.2 (2014): 04014048.

Response:

This reference has been added to the paper.

5:

Pag 3 Section 1.2. MINOR: A scheme or figure of the method would significantly help to understand the image-processing algorithm.

Response:

A good point, the authors have inserted a new figure to illustrate the steps of the methodology: Figure 1.

6:

Pag. 6 Lines 211-213: MODERATE: it is not clear how the procedure is used with multiple images. Please provide more details.

C6

Response:

A description of the procedure for combining image results has been updated in section 1.4 of the paper: "These group scores are determined by multiplying each single model/SAR flood map CSI result with the CSI score of the next SAR flood map until all members of the particular group have been added. The unique combinations which comprise these groups are described in Table 3 below. This combining of CSI scores is done for results from each of the 1000 models/parameter scenarios. The next step is the same as for single CSI scores as described above – i.e. to rescale the objective function and compute the cumulative support".

7:

Pag 7 Figure 1 MINOR: The quality of this figure is very poor. Please provide a larger and cleared picture where the identification of the study area and the boundary conditions are more clearly visible.

Response:

Thank you for pointing this out, Figure 1 (now Figure 2) has now been amended to be clearer with boundary locations highlighted.

8:

Pag 11 Figure 2 MODERATE: the quality and the description of this figure is very poor. Also, ENVISAT and Aerial data seem to be a bit different although with this picture is very difficult to compare the results. I understand that the processing of the SAR image inherently contain errors, I am wondering if the results of the paper might be affected by these errors. The authors could test the procedure also on the aerial photograph to understand the effect of the errors in the processing of the image or on the observed

C7

stages.

The observed model which is expected to behave better than the test model seems to be worse than the test model? Do you have a justification for that? Does this depend on the calibration?

Response:

A well observed point for Figure 2 (now Figure 3) in the paper. The description of this figure in the text has been updated so it is clearer why it was inserted. Also the CSI scores embedded in the figure have been moved to a table for easier comparison/interpretation for the reader. The figures themselves (particularly the modelled and ENVISAT flood extents) have a coarse resolution that does not reproduce very nicely on the page unfortunately.

This particular aerial flood extent image was derived from a single aerial photograph of the flood of July 2007 on the Severn by manual delineation and so is restricted both by the limits of the photographs (cutting off the upper River Avon for example) and interpretation of the image in terms of the flood boundary through vegetation. In this case we used the aerial data here more for validation purposes rather than explicitly to test the calibration methodology. The aim of the paper was to test a series of more 'moderate' resolution imagery that is more extensively and frequently available.

However the authors have added some additional text around the new table of aerial versus ENVISAT CSI scores to explain further the test results we found when applying the methodology on the single aerial flood extent data. A good comparison could be made between the methodology applied to the ENVISAT observation (acquired 23rd July 2007, 10:27am) and the subsequent aerial image (acquired 24th July 2007, 11:30am) as they were observing the flood merely 24hrs apart. While it is quite likely that the impact of SAR processing errors would be manifest in the final results, and be most obvious when compared with better resolution data such as can be obtained from a gauge record or aerial imagery, the authors felt that an explanation of the impact of

C8

these errors was beyond the scope of the paper (which is rather long already).

In Figure 2 (now Figure 3), it is obvious that the observed model (constructed using surveyed cross sections) has not represented observed flood extent as well as the test model. This is most evident in the tributaries to the main River Severn. We updated the text in section 3.1 of the paper to explain this: "The scores and flood extent for the observed model are not better than the test model results as might be expected. This may be explained by the fact that while the bathymetry of the observed model does come from survey data, the (domain-average) channel roughness value is not calibrated in either model. While the test model had 1000 parameter-varying depth and roughness values, the observed model had a best estimate of domain-average channel roughness parameter (of 0.038). While appropriate for the main rivers, it is evident that the channel roughness value is not suitable for the narrower tributaries."

9:

Pag 11 Figure 3 MAJOR: Please provide a better figure with colors. It is very difficult (with this figure) to follow the authors' statements.

Response:

This contour plot has now been updated in black and white so it is easier to read with two single opposing colours than the original colour spectrum which made it difficult to see where CSI was highest and lowest, even in full colour (see fig: contour plots)

10:

Pag 11 Section 3.1 MODERATE: I found this section very difficult to follow and to read. I suggest to try to present it better.

Response:

C9

A re-write of this section has been carried out and the authors hope it is now easier to read and follow.

11:

Pag 12 lines 346-349, MAJOR: the authors concluded that nc is insensitive when estimated simultaneously with the channel depth. However, it seems that this was concluded based only on two images (23rd July 10:27, and 17 January 2008 21:55). Do the authors exclude that this is true in any case and there are not effects of the time of acquisition and the magnitude of the flood event? I think the authors should provide more proofs for this statement. Overall, I find this assumption reasonable however I think that including the DYNIA also for the parameter nc would add a lot of value to the paper.

Response:

This paragraph has been rewritten within the context of the above comment on section 3.1 and hopefully now better explains why 'nc' results were excluded from the final version of the paper. All available SAR data were analysed for 'r' and 'nc' though only two 'r' plots featured in the paper for simplicity. As the authors hope has been explained sufficiently within the general comments section above, the 'nc' parameter did indeed undergo the same DYNIA analysis as the depth parameter but results were not exceptional enough to be included within an already long paper. The same lack of responsiveness of this parameter in the results were observed for all SAR data analysed, and even when the SAR data were 'grouped' the identifiability and IC results did not greatly improve for 'nc'. With this available dataset of ENVISAT SAR images and hydraulic model set up, the authors have concluded that 'nc' is insensitive when estimated simultaneously with the channel depth parameter 'r'. However additional testing would need to be carried out to conclusively say that this insensitivity is true with other SAR data and magnitudes of flooding.

C10

12:

Pag. 13 line 361. MODERATE: It is not clear how the IC score is calculated for multiple images.

Response:

The grouping of SAR data occurs after each single model/SAR flood map is assessed and given a CSI score. It is the CSI scores which are multiplied together in particular combinations (as described in Table 3) before the DYNIA methodology is applied and identifiability and IC determined. A group IC score is derived from the grouped SAR data and CSI scores multiplied together. This explanation has been inserted in to the paper in section 3.2 and 1.4 for greater clarity.

13:

Figures 4, 5, 6, 7 and 8 MAJOR: The interpretation of these figures must be described in the method section. From the text it is very difficult to follow the authors' statements. Please provide a better quality figures as well. It is impossible to discriminate between the different lines. If I understand correctly these are the cumulative distribution of the rescaled support values and not the gradient. The gradient should refer to their slopes. Isn't? If so, I expect a figure like the one in the paper of Wagener et al. (2003), (see FIGURE 8 in their paper).

Response:

Thank you for pointing out how difficult these figures are to read. The original figure coloured lines have been reinstated in this revised version and an explanation of how the cdf plots should be interpreted is now inserted into the method section.

It is correct that Figures 4 – 8 in the paper are representative of the (gradient of the)

C11

cumulative distribution of rescaled support values. At the suggestion of Prof. Wagener the gradient plots were converted to cumulative distribution function (cdf) plots, and this is what was shown in the final version of the paper. This was a change introduced to make the plots easier to read together and normalised, when the original histogram plots were numerous and previously side-by-side for each grouping. For this experiment, the SAR data acquisitions are not regular or plentiful through time (as in the experiment of Wagener et al., 2003 which used 6 years of continuous daily flow data) and so it was more difficult to represent in a line plot. However for purposes of explanation, the last cdf plot in the paper is converted into a 'gradient' plot for parameter 'r' below (top figure) to mimic Figure 8 in the paper of Wagener et al., 2003. In the opinion of the authors this reconfiguration to 'gradient' appears less smooth but nevertheless still conveys the same information as the original figure (bottom) of 'cdf of gradient' that features in the paper. For this reason, we have not amended the Figures 4-8:

14:

Pag 18 lines 503. MODERATE: No colors can be seen in the figures.

Response:

Figure colours have been reinserted.

15:

Pag 19 lines 545-556. MODERATE: I expect here some discussions about the possible consequences of the assumptions made in the paper.

Response:

This consequences of our assumptions in the paper have been expanded within the results and discussion sections, as described above.

C12

C13

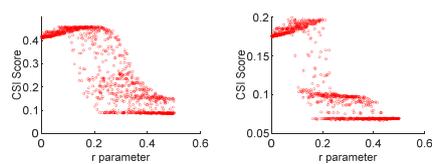


Fig. 1. Figure 1 - Plot of 'r' parameter against Critical Success Index score for left: 23rd July 2007 at 10:27 and right: 24th January 2008 at 10:12.

C14

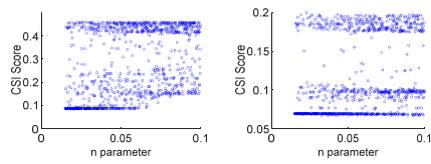


Fig. 2. Figure 2 - Plot of 'nc' parameter against Critical Success Index score for left: 23rd July 2007 at 10:27 and right: 24th January 2008 at 10:12

C15

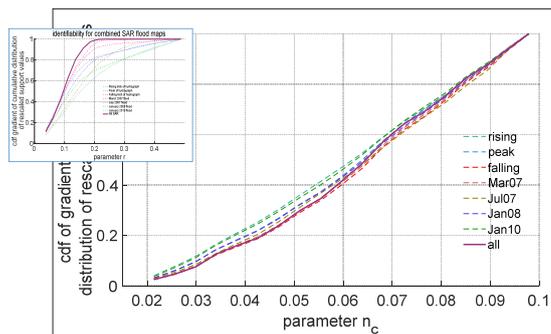
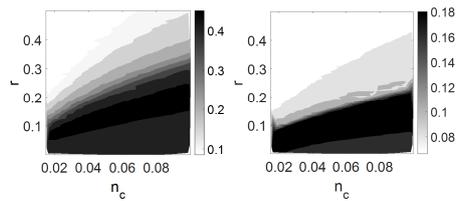


Fig. 3. Figure 3 - cdf of the gradient of the cumulative distribution of rescaled support values for each individual SAR image, against 'nc' parameter value.

C16



From the manuscript: figure 4 : Single SAR acquisitions are compared with LISFLOOD-FP modelled flood maps for the July 2007 flood event. Left: results from the SAR acquisition on 23rd July 2007 at 10:27, right: result from the SAR acquisition 24th January 2008 at 10:12.

Fig. 4. contour plots

C17

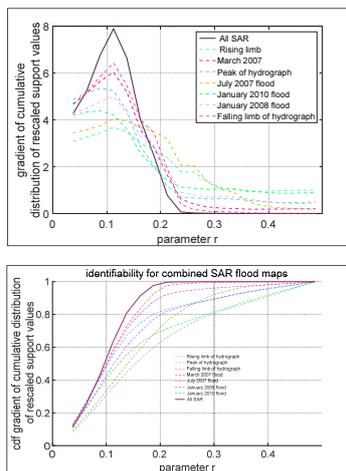


Fig. 5. Figure 4 - top: a new gradient of the cumulative distribution of rescaled support values for each SAR grouping, and bottom: the original cdf of the gradient of the cumulative distribution of rescaled

C18

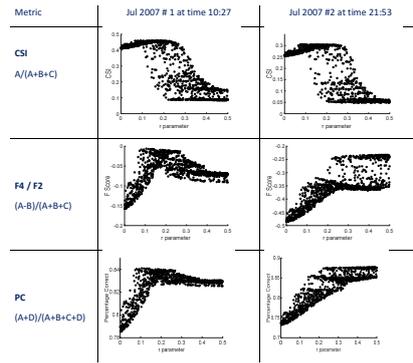


Fig. 6. Figure 5 - Skill Scores using a range of performance metrics. Top: CSI, Middle: F4 or F2 Score, Bottom: Percentage Correct

C19

| 'grouping' | Parameter 'r' | Parameter 'n _c ' |
|--------------------|---------------|-----------------------------|
| Rising limb | 0.13 | 0.11 |
| Peak of hydrograph | 0.23 | 0.12 |
| Falling limb | 0.64 | 0.16 |
| March 07 event | 0.50 | 0.14 |
| July 07 event | 0.37 | 0.14 |
| January 08 event | 0.25 | 0.12 |
| January 10 event | 0.13 | 0.11 |
| All SAR [1-11] | 0.68 | 0.16 |

Fig. 7. Table 1 - Information Content (IC) for groupings of SAR images, showing results for r and n_c.

C20

| IC for | CSI | F2/F4 | PC |
|--|---------------|---------------|---------------|
| Single SAR data (left: SAR1 at time 10:27, right: SAR2 at time 21:53) | 0.165 / 0.188 | 0.066 / 0.102 | 0.102 / 0.101 |
| July 2007 'flood event' | 0.37 | 0.079 | 0.105 |

Fig. 8. Table 2 - Information Content (IC) for parameter r . Top row: single SAR images from July 2007 flood event, and bottom row: the same 2 data, grouped into 'flood event'.