Hydrol. Earth Syst. Sci. Discuss.,
doi:10.5194/hess-2015-511-AC2, 2016

Hydrology and Earth System Sciences
Discussions

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

Comment 1:

The IC approach is really nice and gives an objective assessment of the value of a flood image for calibration. However, what we are still missing in the literature is to find a way that gives an objective IC of a SAR flood map without the need to calibrate first. In other words, in this paper, which I think has a lot of merit, IC is built up based on parameter identifiability rather than for instance inter-comparing each SAR image and applying the score and identifiability that way, so without the need of a model and its parameter but I understand that this is outside the scope of this paper.

Response:

It is an appealing idea to attribute IC to individual SAR data, without need of using models and parameters for calibration. At the moment the authors are not sure how to inter-compare SAR data to reveal information content but it could be an interesting topic of further study.

Comment 2:

I also think that what is innovative here is the analysis of IC and identifiability in relation to what stage in the hydrograph we are looking at and what type of data we use (single image, combined images, gauge data). I wonder if the title and the introduction should better reflect that since to me this is one of the first papers to try and answer these questions using real data.

Response:

Thank you for raising this point. The authors have updated the title of the paper and reworded the abstract, so that these reflect more accurately the unique points within the paper.

Comment 3:

My biggest reservation in this study lies with the choice of performance metric used, which may explain in my opinion why the greatest information content is in the SAR images closest to peak flow. Stephens et al. (2014) showed that the performance measure used here is particularly biased towards largest flooded area (in other words, it always gives the highest score to the biggest flooded area). This is significant in this study and could lead to an unwanted "bias" in the calibration. I suggest the authors repeat the exercise offline with the "F2" measure for instance ((A-B)/(A+B+C)) or an

area in error index ((B+C)/(A+B+C+D)) to see if the same SAR images give the highest sensitivity still.

Response:

A valid point is raised here. In the preparation for this paper the authors did indeed prepare a number of 'skill score metrics' before deciding on the CSI skill score in preference over 'F' measures and other promising metrics such as Percentage Correct. A sample of these initial plots are shown here in Figure 5 to explain why in the end we decided to use the CSI results.

These are just a sample of the results available and show the results from analysis of 2 SAR data from the flood of July 2007, but they are fairly representative of all the results for the full range of SAR-derived flood maps which were analysed. The F2 (aka F4) and CSI plots in particular gave a more sensitive response through the changing parameter value 'r'. This was no doubt due to the 'white space' of no-water cells being absent from the skill score equation. Considering all SAR results together, the CSI scores were observed to provide a more responsive and consistent result with changing parameter value than the other metrics which were assessed.

Concerning the comment regarding CSI skill scoring usually providing a better result for fuller flood extents due to the dominance of 'water' pixels, as pointed out by Stephens et al. (2014). This point is well raised but the authors would respond that the CSI skill score is secondary to the shape of the CSI peak itself in this particular case. This identifiability methodology looks at how sensitive the model is for different parameters around this CSI peak and gives little significance to the CSI scores themselves. To illustrate also the greater sensitivity/information for 'r' seen in the images when using CSI, the IC scores using PC and F4/F2 as the central metrics are shown in Table 2, against the original scores obtained using CSI.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., doi:10.5194/hess-2015-511, 2016.

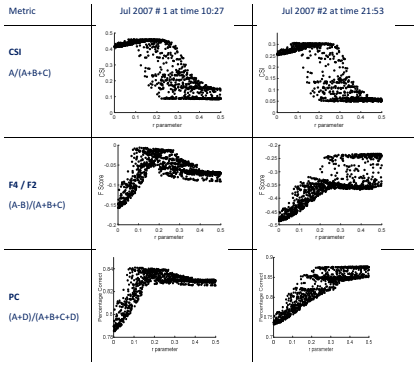

Fig. 1. Fig 5 - Skill Scores using a range of performance metrics. Top: CSI, Middle: F4 or F2 Score, Bottom: Percentage Correct

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. 2. 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'.