Hydrology and Earth System Sciences Discussions

Interactive comment on "Automated global water mapping based on wide-swath orbital synthetic aperture radar" *by* R. S. Westerhoff et al.

P. Matgen (Referee)

matgen@lippmann.lu

General comments:

Westerhoff et al. introduce the different elements of an operational SAR-based flood mapping service. The core component of this service is a statistical method that computes the probability of individual pixels being covered with water. The probability density functions are conditional on the backscatter value and the incidence angle. The resulting flood maps are qualitatively evaluated against MODIS-derived flood inundation maps. The authors conclude that their automatic and systematic method provides "hydrologically correct" results.

We discussed this interesting manuscript in our research group that is working on a similar topic. We all agree that the method that is presented in this manuscript is very innovative and of high interest for the community. As of today, flood mapping methods are mostly based on one or, at best, a couple of remote sensing images. On the other hand, the approach that is shown here makes use of gradually increasing image archives in order to build region-specific empirical histograms of typical "land" and "water" backscatter values. These functions are used in combination with recently observed backscatter values in order to compute posterior distributions in the framework of Bayes' theorem. This is clearly the strong point of this contribution and in the light of recent and upcoming satellite missions we believe that the presented method can indeed change the way flood inundation maps are computed from SAR observations.

Having said this, we have three main concerns regarding this manuscript. First, we regret that the authors made the choice to present an operational service rather than a scientific method. This means that arguably there are some parts in this paper that are not very relevant in the context of HESS (e.g. production of Google maps, discussion on the handling of current and future data amounts, speed of image processing etc.) and that blur the paper's significant scientific message.

It also means that in our opinion the authors did not carry out a convincing quantitative evaluation and sensitivity analysis of their scientific method. This is a pity as in our opinion this would be of much higher interest for the community than the more technical aspects of the operational service (at least in the context of HESS). At one point there is a visual assessment of the flood probability maps. This purely qualitative evaluation does not provide much insight into the performance of the method. The title of section 10 ("Validation of the method ...") is misleading and does not seem appropriate. At no point the authors compute performance indices. Is there no data set available that would allow for a quantitative evaluation of the algorithm's performance? Would it not be possible to compare the maps to those obtained with other methods?

The third concern relates to the way the authors approximate the prior knowledge that

a pixel within the SAR scene is covered with water or not (see Equations 1 & 2). In fact the authors state that they have no prior knowledge and that therefore the probability is set on 0.5. Clearly, this prior largely overestimates the probability of a pixel being covered by water (basically this would mean that half of the Earth's terrestrial surface is covered by water). A much better prior would be in our opinion to consider the number of permanent "water" pixels (e.g. using the same mask that was used to build the empirical histogram of water surfaces) and to divide it by the number of pixels covering a tile. Under these assumptions the simplifications that lead to equation 7 would be no longer valid and, as a result, the posterior probabilities of the individual pixels being covered by water would change. Moreover, the fact that the authors compute histograms of water and dry pixels on different areas needs more explanation and discussion. Indeed, the fact that these two histograms are not computed on the same population (global scale for water and tile scale for dry land) could render the Bayes' law not applicable. The statistical distribution of water is probably stable over the globe but then authors have to demonstrate that it is the case.

In general, it would be also beneficial for the manuscript if the authors would be more critical about their own approach. What are the inherent weaknesses of their method if it is compared to other established SAR-based flood mapping methods? For example, we would hypothesize that it is difficult to apply this method to very high resolution SAR imagery (because of the difficulty to have a sufficiently large samples of images and because of the speckle). Does this method enable the detection of water within urban settlements or under vegetated canopies? Or when wind roughens the water surface?

For all these reasons, we recommend accepting this contribution, subject to some moderate revisions.

BTW: Another recommendation would be to make two papers based on this material: the first paper could deal with the scientific method and its critical evaluation and the second paper could deal with the operational service.

Specific comments:

Introduction: the state of the art of existing SAR-based flood mapping methods is not exhaustive. Some important papers seem to be missing (e.g. Mason et al., 2010, Pulvirenti et al., 2011, Martinis et al., 2009, 2011).

p.7808 l.25 we cannot understand the range of values given here. The values in dB have to be much lower (i.e. negative values for backscatter on open water). This seems to be a mistake.

p.7809 l.20 We think it is no justifiable to have the same priors. Your water mask (p.7807) enables you to compute more adequate priors per 1 degree tile.

p.7810 l.13 Do you have to parameterize the pdfs of the land and water empirical pdfs? The subplots in figure 2 do not seem to support your assumption of a normal distribution.

p.7811 How did you select the threshold value of the HAND index? Since you are not carrying out a quantitative evaluation of your output maps, it is difficult to evaluate the correctness of your value. At the very least a sensitivity analysis should be carried out.

p. 7812 l.5 What happens if smaller tiles are considered? Is there any advantage to consider smaller tiles? Does it have any impact on the results (other than computation time)?

P.7813 l.2 and in general: in the context of a systematic SAR-based flood mapping application it might be useful to indicate the number of images that are available across the globe. We assume that in some areas the sample size is too small to carry out the statistical analysis. Is this something one has to take account of? You mention that during the flooding event ESA "switched on" the image mode. But do you have enough historical data to build the training data set for this image mode?

p.7813 l.11-18: in our opinion, these observations cannot be considered as "validation". The same comments could be made with respect to many other SAR-derived flood

inundation maps. The fact that some elevated features can be identified on the map and that according to the algorithm's result the city of Bangkok did not suffer extensive flooding hardly validates the method. Do you not have any other study area with more useful ground information, aerial photographs or very high-resolution satellite imagery that could be used for a more meaningful evaluation?

p.7814 The authors conclude that their algorithm is useful to detect floods in cloud-covered areas. However, we do not see any hard evidence that would support this conclusion.

Fig. 2 A legend is missing. Moreover, we doubt that on the y-axis these are backscattering values in dB.

Fig. 3 A legend is missing here to understand the colour coding.

Fig. 4 The units are missing.

Fig. 7 The font size should be increased.

Minor comments:

p.7802 l.7 please reformulate this sentence as there can be no "probability distribution function of a pixel" p.7802 l.23 I would add "accurately". In general I would put more emphasis on "accuracy" and "reliability" than on the fastness of the service (also in the conclusion). p.7804 l.6-7 The meaning of this sentence is not very clear. It would be useful to add some examples of how the maps can be used in the context of flood inundation modelling (see for example: Schumann et al., 2008 for an extensive review on the subject). p.7804 l.1-16 It is also worth mentioning the problem of shadow and layover hampering the detection of water (see for example Mason et al., 2010) p.7805 please reformulate "an army of human operators" p.7805 l.13 The acronym NRT should be introduced here rather than in l. 17 p.7806 l.18 distribution p.7806 l.23 "degree" is missing here p.7806 l.27 "permanently dry" does not seem to be an appropriate term in this context. These pixels can be flooded. p.7807 l.2 the processing of smoothing

needs to be described somewhere. p.7808 l.10 Could you state that the underlying working hypothesis of lumping all freshwater values is that local effects are neglected (e.g. some areas are more exposed to wind than others, some water bodies have more or less sediments or contain more or less salt – all these factors may influence the shape of the backscatter histogram). p. 7810 l.3: it might be useful to formulate this condition as this helps the comprehension of Eq. 6 p.7813 l.6 This is not clear as in Fig. 9 all the water probabilities are provided. In general we do not understand why the authors use a 70% probability for computing the binary map. Should it not be 50 % (i.e it is more likely that the pixel is flooded than not flooded) ? p.7813 l. 9 delete "is" p.7813 l.20 it might be useful to introduce the "Global Flood Observatory" and its acronym (is given later on p. 7814) p.7814 l.22 What is the "complementarity" of the two mapping methods?

References:

Martinis, S., Twele, A., & Voigt, S. (2009). Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data, Nat. Hazards Earth Syst. Sci., 9, 303-314.

Martinis, S., Twele, A., & Voigt, S. (2011). Unsupervised extraction of flood-induced backscatter changes in SAR data using Markov image modeling on irregular graphs, IEEE Trans. Geosci. Remote Sens., vol. 49, no. 1, pp. 251-263.

Mason, D. C., Devereux, B., Schumann, G., Neal, J. C., & Bates, P. D. (2010). Flood detection in urban areas using TerraSAR-X, IEEE T. Geosci. Remote Sens., 48(2), 882-894.

Pulvirenti, L., Pierdicca, N., Chini, M., & Guerriero, L. (2011). An algorithm for operational flood mapping from Synthetic Aperture Radar (SAR) data using fuzzy logic, Nat. Hazards Earth Syst. Sci., vol. 11, pp. 529–540.

Schumann, G., Bates, P. D., Horritt, M. S., Matgen, P., & Pappenberger, F. (2009).

Progress in integration of remote sensing derived flood extent and stage data and hydraulic models, Review of Geophysics, 47, RG4001.

Patrick Matgen Laura Giustarini Renaud Hostache

C2934