Hydrol. Earth Syst. Sci. Discuss., 5, S2619–S2633, 2009
www.hydrol-earth-syst-sci-discuss.net/5/S2619/2009/

Hydrology and Earth System Sciences Discussions

HESSD

5, S2619–S2633, 2009

Interactive Comment

Interactive comment on "Improved estimation of flood parameters by combining space based SAR data with very high resolution digital elevation data" *by* H. Zwenzner and S. Voigt

H. Zwenzner and S. Voigt

Received and published: 27 February 2009

First of all, the authors wish to thank the referee for the in-depth analysis of our paper and the constructive comments, criticism and suggestions which will definitely improve this paper. In the following pages we try to answer the issues raised by the referee in sequence, in which the referee comments are printed in italic and the answers are printed in bold.

One concern I have with this paper is that, in my opinion, the methodological approach is not sufficiently well explained and that part of the argumentation appears to be illogic. Many processing steps remain very vague and unclear (e.g. certain segments with a *mean elevation above a certain elevation are excluded). The key sentence of this paper (Differences are balanced...) leaves room to several interpretations. What do you mean by "' the lowest flood profile elevation"'? Is this the mean elevation extracted at the left and right edge of each profile segment? Would the result be the same if you would minimize the differences between water stages extracted on the two edges of each profile segment? I would ask the authors of this paper to clarify the methodology in order to avoid any misinterpretation.*

1. We agree that some passages of our methodology section are very briefly written and some matters could be misinterpreted. In our revision the methodological approach is explained more precisely and in greater detail. Additionally, we will provide a flowchart to illustrate the logical sequence and the relevance of each processing step, i.e. if it applies to a single cross section profile or to the longitudinal profile of the river reach. The more specific questions raised by the referee will be answered below.

What appears to be rather illogic in the methodological approach is that in the first processing step the authors have to assume that there are no classification errors, because if there are classification errors, there is no point in shifting the flood segment. By following this procedure, the relocation of flood segments that are influenced by other errors than positional errors will lead to an over- or underestimation of the true water levels. By trimming and extending the flooded segments once they are repositioned, the method merely helps to make sure that there is a coherence between RS-derived flood extents and topography (the elevation will remain the same). Hence I dont agree with the authors that this second processing step enables the correction of classification errors.

2. First of all, we totally agree that for larger areas that were misclassified as flooded or non-flooded, there is no use in shifting the respective profile segment because then a meaningful water level can not be derived. Actually, this was the main problem in the Elbe case study and explains why we did not derive

flood depth in contrast to the Severn case study where we followed the whole processing chain. Initially, we made several assumptions in our methodological approach. Two major assumptions for the derivation of flood depth are that each cross section has a horizontal water level and that only minor thematic errors (in the order of some pixels) exist in the flood profiles. Assuming there are also minor geometric errors in the remote sensing data (in the order of few pixels again) we have the problem that we can not disentangle thematic and geometric errors within one single cross section profile. These errors might even be superimposed. Thus, we developed a method which allows for the compensation of small scale thematic and/or geometric errors. This compensation is based on a moving average which is applied on the sequence of cross sections (i.e. water levels) over the longitudinal profile. We used a moving average over 31 cross sections to get a smooth water surface and to eliminate outliers. In the resulting water surface of the longitudinal profile (see figure 7 in our paper) variations in the successive water levels due to small scale thematic/geometric errors in some particular cross sections are averaged (or leveled) by means of the adjacent cross sections. Afterwards, each of the cross section flood profiles are trimmed/extended according to the modified water level from the longitudinal profile, thereby fitting the remotely sensed flood profiles onto the underlying topography. Although we see the problem of underestimating the true water stage in case of flooded vegetation and urban areas, we do not think that the exclusion of all urban settlements and vegetated areas from the analysis is applicable, especially if we lack a proper land use classification. Instead one could exclude certain cross section profiles showing obvious anomalies in the water level and which can be explicitly attributed to vegetation or urban structures. The resulting gaps in the longitudinal profile (not more than 2-3 cross sections) would then be interpolated by the averaging process.

In conclusion, I am rather skeptical about this methodology's capability to correct for classification errors and I would even say that the processing steps might cause gross

HESSD

5, S2619–S2633, 2009

Interactive Comment

errors if areas with classification errors are considered. I would appreciate if the authors would develop a discussion around some of the limitations of this method and I would recommend them not to claim that the method can correct thematic errors. It is just an assumption that is not proven in this paper (see comment below). I agree with the authors that their method allows to achieve a coherence with the topographic data (which is a good thing).

3. It is true that we can not claim that the proposed method can correct large thematic errors. The wording in our paper might have been misleading. We would rather use the formulation: the method allows for the compensation of small scale thematic and/or geometric errors (see explanation above). However, there are a number of preconditions and limitations of the approach which will also be described in detail in our paper. For the adequate estimation of the water level it is necessary to determine the accurate position of the flood water boundary. Since the water level is a function of the water extent, the terrain which confines the flood water should have a moderate slope. A very steep slope would lead to increasing errors, because small changes in flood extent would lead to high variation in the water level. This means that this method should only be applied for large flood events in which the normal river channel is overtopped. Apart from dense vegetation and urban structures at the flood boundary, also very flat terrain seems to cause problems with the accurate identification of the flood extent and thus can lead to errors in the derived water levels. As it is stated in 2, our method is based on a sequence of cross sections along the main river course. The occurrence of the above mentioned errors and uncertainties in deriving the water level can be compensated for a number of single cross sections and even a few adjacent cross sections. However, large scale classification errors as well as geometric errors, which are greater than a few pixels or range over more than a few cross sections, have to be excluded or reduced by further preprocessing. The proposed method is a relatively simple method which should be considered in the context of the rapid mapping of flood parameters,

such as flood extent and flood depth, as contribution for disaster management operations or the rapid estimation of flood damages. The requirements of these application fields are fulfilled by the fact that computation time is negligible and data requirements are low, i.e. only high resolution remote sensing data and LiDAR elevation data are needed. The proposed method stands in contrast to hydrological modeling approaches which are more complex with respect to data requirements, parameterization and computation time. However, we believe that both approaches can benefit from each other in terms of cross-comparison and validation.

One of the main differences between this method and the one given by Schumann et al. (2006) is related to the choice of the authors to independently shift the cross sections. Whereas Schumann et al. (2006) applied a same dx,dy shift to the entire flood mask, here the authors look for a different "'optimal"' shift for each flood segment. I would be very interested in knowing if the same (or at least similar) dx,dy couple (inferred from the shift along the cross section) was found for every section (which would speak in favour of a systematic geocoding error) or if rather different values were found. If one could see how the dx,dy values are distributed along the reach one could probably distinguish between positional and thematic errors.

4. Schumann et al. (2006) present an interesting approach for compensating a geometric shift and improving the positional accuracy of flood masks. In our paper we decided to treat each cross section individually for the compensation of small scale thematic and geometric errors. In our analysis we also mapped the shift of each individual flood profile based on a color scheme. The results indicate no systematic shift for the whole flood mask and the variability in different sections makes it difficult to distinguish between geometric and thematic errors. However, we assume that thematic classification errors are predominant.

Another rather annoying thing about this paper is that there is no means to verify the reliability and plausibility of the methodology. In fact, in both case studies, the reference

HESSD

5, S2619–S2633, 2009

Interactive Comment

data is provided by two optical images that were acquired several hours before or after the SAR acquisition. Without being able to see how the situation evolved between the two acquisitions, the comparison is pretty much useless. I would appreciate if the authors would provide the flood hydrographs that are depicting the timing of the two acquisitions. Also, for a real validation it would be necessary to have some ground truth data with respect to water stages. The validity of the methodology cannot be verified by the means of flood extent imagery alone. The authors claim that their approach can correct thematic errors. With the data at hand they cannot proof their a priori assumption (especially in the light of the serious doubts that I rose in the previous section).

5. Fortunately, we were able to get hold on data from the water gauge at Mythe Bridge (Tewkesbury, UK). Thus we can present the hydrograph and illustrate how the flood situation of the Severn evolved between the two acquisitions of TerraSAR-X and aerial photography. For validation purposes we use water gauge data together with flood extent from aerial photos. As it can be seen from the vast majority of flood related studies, it is very difficult to acquire ground truth data during a catastrophic flood situation such as it was the case in Tewkesbury. For such a task a large amount of manpower needs to be available and the campaign had to be well prepared, both being more than unrealistic during of a catastrophic flood event. Furthermore, the accessibility of the flooded area is rather limited. Having the water gauge data at hand we will present a comparison of the gauge data and the water level derived from TerraSAR-X data and aerial photography at Mythe Bridge. The results will be discussed in the final version of our paper.

Specific comments: p. 2955 l. 2 the underlying assumption is that the water stage is horizontal along a given profile. The profiles need to be chosen in such a way that this assumption remains valid

6. Initially, the terrain profiles (i.e. cross sections) are generated perpendicularly

HESSD
5, S2619–S2633, 2009

to the centerline of the river. Via intersection of the terrain profiles with the flood mask, the flood profiles are derived. The assumption is that the water stage of an individual flood profile is horizontal. However, due to small geometric or thematic errors, a difference in the water level of the left and right bank would be immanent. In this case, the profile is shifted to diminish this variation.

p. 2955 l. 15 very vague! Please explain in more detail how you choose these thresholds. I assume that this threshold changes in a downward direction along the river flow line because you cannot choose a single mean elevation for the entire reach

7. Each cross section flood profile consists of one or more flooded segments. In order to exclude segments which were misclassified as water (e.g. on backward slopes) we used a threshold based on the mean elevation of the flooded segment. The threshold was 3 meters in the Tewkesbury case study. That means that all flood segments which are more than 3 meters above the lowest flood segment of this profile are excluded. This procedure is computed for each cross section, and the threshold is held constant over the whole reach. We will clarify this in the revised paper.

p. 2955 l. 21 what is the "'lowest flood profile elevation"'? How do you compute it? (average between the elevations extracted at the left and right border of each segment?)

8. As it was described above (see paragraph 6), we shifted the flood profile in the direction of the lower side until the "'minimum mean elevation"' was reached and the difference in the elevation between the left and right border was leveled. The direction and the dimension of the shift will be shown in a new figure described above (see paragraph 4). However, large thematic errors (e.g. by flooded vegetation) might cause a large shift which can lead to a significant underestimation of the water level (see paragraph 2). In this case the profile should be excluded. In case of minor geometric errors the shift can help to improve the positional accuracy. Variability due to small scale thematic errors will be compensated during a

HESSD

5, S2619–S2633, 2009

later step.

p. 2955 l. 21 I am not sure about the term "‘moving window"’? Can you develop the explanation of this processing step?

9. We admit that the term "‘moving window"’ is probably not the exact term. We will change it into "‘shift of the flood profile along the cross section"’ as it has been described in 8.

p. 2955 l. 28 this processing step was not applied in the Elbe case study. Why?

10. The elevation was also derived for the Elbe case study and can be seen in figure 2 and 3 in our paper. However, the results are very weak and no further processing (i.e. mapping of flood depth) was carried out. We will make this clear in the revised version.

p 2956 l. 3 I dont agree (see comments above), I agree though that this step ensures a coherence with the underlying DEM that might in some cases compensate minor classification errors.

11. Indeed, establishing a coherence with the underlying DEM and compensating possible small scale geometric or thematic errors is exactly what we intend with our proposed method. We will modify the formulation according to paragraph 3.

p. 2956 l. 7 - l. 10) this is a preliminary discussion/conclusion

12. The sentence will be moved to the discussion chapter.

p. 2956 l. 22 "‘gradient is about 7m"’. Meaning?

13. Gradient means the difference between the elevation of the upstream and the downstream end of the river reach shown in figure 1. Accordingly, the longitudinal slope is 0.0005 m/m.

p 2956 l. 2 how did you geocode the image? What is the accuracy with respect to *ground control points?*

14. The geocoding of the Radarsat image was done by means of the software package SARscape and the incorporation of one ground control point to improve the positional accuracy. The accuracy assessment based on 8 independent and evenly distributed GCPs gave a total RMSE of 16.59 meters. Considering a pixel size of 12.5 meters this is a more than satisfying value. We will include the information about the geometric accuracy in our paper.

p. 2957 l. 5 does this filtering step over 7x7 pixels preserve the edge of the flood boundary?

15. The adaptive Lee-Sigma filter is a standard edge-preserving filter in most image processing software packages, and it is specifically used for speckle suppression in SAR images (see Sheng and Xia, 1996).

p. 2957 l. 10 rather confusing sentence! What reference data was used to guide your trial an error procedure? How did you evaluate the errors?

16. The resulting classified flood extents were visually compared against official municipality flood maps derived from aerial survey, ground truth information and citizen interviews. These official flood maps showed good agreement with the IKONOS image. Details of the reference flood maps will be given in the final text version. The errors were evaluated qualitatively on the basis of a visual interpretation and fine tuning of the threshold was done at a number of selected points where the water boundary could be reliably identified from the reference data.

p. 2957 l. 18 contradiction with Fig. 1: 3 or 5 hours?

17. There was a mistake in the caption of Figure 1. It has been corrected: >three< hours instead of >five<.

p. 2958 l. 1 why Fig.4 before Fig. 2?

HESSD

5, S2619–S2633, 2009

18. To avoid confusion we positioned Figure 4 between Figure 1 and 2.

p. 2958 l. 11 please comment the differences that are shown on Fig. 3 , also I would suggest to maintain the same axes Why did you not apply the moving average?

19. Assuming that you refer to Figure 2, it illustrates the shift of the flood profiles onto the slope of the left side of the river. Misclassification mainly caused by radar shadow is responsible for the large outliers in the elevation of the water level at the left river bank. This could be corrected by means of a lateral shift of the flood profiles. The results are shown in Figure 2. However, the results of the corrected profiles are very disappointing with respect to the derivation of the water levels. Nearly half of the profiles exhibit just a minor increase of the water level or no increase at all compared to the mean water level of the river. This is shown by the values which are located on the blue line (see Figures 2 and 3). The remaining profiles show a tremendous underestimation of the water levels when compared to IKONOS (Figure 3). This indicates that flood masks from Radarsat-1 are inappropriate for the derivation of the water level. As we will elaborate in the methodology section of our revised paper version, the initial assumptions and boundary conditions for the derivation of the water level are not fulfilled, and thus the following processing steps (such as the moving average) are dismissed.

p. 2958 l. 17 nothing is proven here! This is a mere visual analysis. After the processing the water surface line appears to be more plausible than before but the "‘hydraulical plausibility"’ cannot be proven with the data at hand. In fact the lines are still not hydraulically plausible (the data points are still scattered)!

20. The "‘hydraulic plausibility"’ refers to the above mentioned correction of the misclassification due to radar shadow and the resulting shift of the profile. But we agree that the term "‘hydraulic plausibility"’ and the discussion are misleading. In the respective paragraph of our paper, the focus lies on the fact that Radarsat-1 data do not allow a proper extraction of the water level, and the hy-

draulic circumstances can not be represented by these data.

p. 2958 l. 17 This is not a confirmation! At best it is another indication! "‘normal water level"’ ?

21. We agree! It is another indication of what is illustrated in Figure 2. In Figure 4 the laterally shifted profiles are drawn in XY-direction into a map and can be compared to the original flood mask and the elevation data, which allows the detection of problematic areas. "‘Normal water level"’ is a synonym for "‘mean water level"’ of the river, meaning the pre-flood average water level as it is displayed by the LiDAR data.

p. 2958 l. 25 in this case the shifting method doesnt give reasonable results. Please comment. p. 2960 l. 3 "‘ cannot be guaranteed"’ This is a nice euphemism! If you consider IKONOS as a valuable reference, a 2m difference cannot be tolerated. This would prove that the method failed. How do you explain this systematic underestimation of the water surface line.

22. If thematic classification errors are greater than a few pixels and the flood extent is not expressed by the flood profiles (e.g. they are too short), no reliable water level can be derived. This has already been elaborated in paragraph 3. Because of their agreement with official high resolution flood maps, we considered the IKONOS data as a valuable reference. The comparison of IKONOS with Radarsat-1 data rather intended to show that Radarsat is not suitable to derive meaningful water levels and NOT to prove that the method failed. The method works quite well if certain boundary conditions (see paragraph 3) are fulfilled. The longitudinal profile (Figure 3) is a key instrument of our methodology since the derived water levels of the individual profiles can be compared against reference data, water gauge data, etc.. This information can be used as an indicator of how reliable the water level can be derived from the SAR data and where problems (e.g. classification errors) occur. Of course, 2 m difference is a clear

indication that Radarsat-1 derived water levels are of no use and the processing is stopped at this step. We will clarify this in the final text version.

p. 2959 l.9 "‘previous"’ instead of "‘last highest"’

23. Corrected.

p.2959 l. 11 - 13 this seems to be a contradiction: "‘the flood situation was stable with two local maxima"’. Can you show the hydrograph?

24. Please see paragraph 5.

p. 2960 l. 2 what is the multiresolution segmentation?

25. Image segmentation techniques automatically group neighboring pixel in contiguous regions based on similarity criteria of the pixels properties. In multiresolution segmentation algorithms hierarchical data structures are used for the segmentation (see Baatz and Schäpe, 2000).

p. 2960 how did you geocode the TSX image? What is the accuracy with respect to ground control points?

26. We did not assess the geometric accuracy by means of GCPs. The TerraSAR-X image was delivered in the standard high precision geocoded Enhanced Ellipsoid Corrected (EEC) format. The "‘rapid"’ orbit type (GPS orbit determination) was chosen, which exhibits a pixel localization accuracy of 2m. This value is predefined rather conservatively, reflecting the possibility of a low incidence angle and errors within the DEM used for geocoding, which both have a negative effect on the pixel localization accuracy (see TerraSAR-X Ground Segment - Basic Product Specification Document). The incidence angle of the TerraSAR-X scene used for our study was 24 degree. In rather flat terrain such as the Tewkesbury area the pixel localization accuracy is much better than 2m.

p. 2960 l. 17 does this mean that the shift that is applied is always an integer number?

Why did you not apply floating numbers? I mean that the geometric error probably doesnt correspond to an integer number.

27. The sampling distance of our cross-section profiles is given in integer numbers which are aligned to the pixel size of the flood mask. Since the shifting and correction measures are bound to the integer sampling points, the shift is always given in integer numbers.

p. 2963 l. 17 is it not possible to detect flooding inside urban areas with TSX data?

28. It is possible to detect flooding inside urban areas with TerraSAR-X data, however it is problematic because of the incidence angle and shadow/layover effects and the strong scattering (double bounce) caused by urban structures. As we indicated in our paper, the high resolution SpotLight mode of TerraSAR-X is of advantage to detect flooding in urban areas. But this is not in the scope of our paper!

p. 2961 l. 26 in my opinion a likely reason could be that your approach didnt allow to detect the flooding within the urban settlement. Hence you relocate the flooded segment to a place where the intersection with the underlying DEM provides water stages that underestimate the true water stages.

29. This might be one possible explanation for the depression of the water level in the middle of the river reach illustrated in Figure 7. We agree that misclassifications due to urban settlements may have caused this underestimation of the flood extent in part of the river reach. This would then also have a diminishing effect on the water levels. We will elaborate on this in our final text version.

p. 2962 l. 11 delete "'Discussion of Results and Conclusions"'

30. Corrected.

p. 2963 l. 5 not only with the R-1 image! The results obtained with TSX are similarly disappointing (cf. Fig. 6 am Fig. 7)

HESSD

5, S2619–S2633, 2009

31. We do not agree with the Referees opinion that the results from TerraSAR-X are actually disappointing. In combination with high resolution SAR data the proposed method rather appears to yield promising results within the context of a rapid mapping of flood parameters such as flood depth. However, there is of course enough potential for improvement and further development of the method, as these are the first results that are being published. As we pointed out the robustness of the method has to be proven and more validation and parameters studies would be more than desirable in order to make more quantitative statements instead of qualitative ones.

p. 2963 l. 8.- l. 13 This paragraph is overly optimistic. Since there was apparently no data available that would allow a true validation, I cant see how you can conclude this. How can you claim the "'methodology allows the generation of reliable and hydraulically sound maps"'? Please change the wording in this paragraph because in my opinion it gives the impression that you gloss over the results that were shown in this paper.

32. We agree that the wording of the paragraph mentioned by the Referee could well be improved and parts of the conclusion might have been misleading. It was not our intention to gloss over any of our results and we will reformulate our conclusions in the final paper version: "'The proposed methodology enables the improvement of the mapping of flood depth, thereby enhancing the hydraulic plausibility of flood masks from high resolution SAR data based on LiDAR elevation data."'

Fig. 1: 5 hours or 3 hours? Fig.2 meaning of "'mean water level"' ? Fig.3 meaning of "'normal water level"' ? Fig. 4 meaning of second sentence in the legend is rather unclear to me Fig. 6 meaning of "'mean water level"'?

33. Figure 1: caption corrected (see paragraph 17). "'Mean water level"' (or as a synonym "'normal water level"') specifies the pre-flood average water level of the river which is represented by the LiDAR elevation model. We will change it

HESSD

5, S2619–S2633, 2009

Interactive Comment

in our paper into: "‘pre-flood average water level"’. It is displayed in Figures 2, 3, and 6 for comparison with the flood water levels. In this way, not only the total flood water level is displayed in the longitudinal profile but also the relative flood water level (with respect to the mean water level from the LiDAR DEM). This helps to identify and evaluate possible errors in the derived water levels at their particular position in the longitudinal profile. Fig. 4: Second sentence was changed into: The cross section flood profiles with a spacing of 100 meters were horizontally shifted according to the underlying terrain data and are drawn in yellow.

References:

CAF - Cluster Applied Remote Sensing, 2006. TerraSAR-X Ground Segment - Basic Product Specification Document, Public Document, Date: 06.10.2006, 45 pages, German Aerospace Center (DLR)

Sheng, Y., and Xia, Z.-G. 2006. A comprehensive evaluation of filters for radar speckle suppression, In: Remote Sensing for a Sustainable Future, International Geoscience and Remote Sensing Symposium (IGARSS), Lincoln, Nebraska, USA, 1996, 1559-1561

Schumann, G., Black, A., Cutler, M., Henry, J.-B., Hoffmann, L., Matgen, P., Pfister, L. 2006. Hydraulic and event knowledge to reduce the positional uncertainty in SAR flood images for improved flood model calibration and development, In: Proceedings of Accuracy 2006, 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, Lisboa, Portugal, 5-8 July 2006, pp. 633-642

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 5, 2951, 2008.