Hydrol. Earth Syst. Sci. Discuss., 7, C5284–C5293, 2011
www.hydrol-earth-syst-sci-discuss.net/7/C5284/2011/

Hydrology and
Earth System
Sciences
Discussions

Interactive comment on "Multi-objective automatic calibration of hydrodynamic models utilizing inundation maps and gauge data" *by* N. V. Dung et al.

N. V. Dung et al.

dung@gfz-potsdam.de

Received and published: 1 April 2011

The authors wish to thank Guy J.-P. Schumann (Referee 1) for his fruitful review and the useful suggestions.
Before going into the specific replies, the authors would also take the chance to answer four general points, which were pointed out by two or more referees :

1. The justification of using a 1D hydraulic modeling approach
2. The method for constructing inundation maps from Radar satellite imagery
3. Optimization criteria F2 - definition and evaluation

4. The definition of the roughness parameter ranges

Ad 1: 1D hydraulic approach
The main justification for using a one-dimensional modeling approach for the simulation is the sheer dimension of the simulation domain, which has a size of more than 55000 km2. For this dimension, two-dimensional approaches are still very CPU-demanding, even with modern computational facilities and parallelization techniques. The only study of such large scale dimension published was by Wilson et al. (2007) for a reach of the Amazon. However, the simulation domain was about 13000 km2, i.e. less than a quarter of the size of the presented study, while size of the computational grid cells was 270 m square. The simulation time in this study for a 22 month dynamic simulation period was reported to be 14 days. Projecting this on the Mekong study, a single simulation for the Mekong would last about 360 hours, i.e. about 15 days. Clearly with this computational time the presented automatic calibration would be impossible. Moreover, a grid cell size of 270 m would results in badly represented inundation areas and inundation process. While this spatial discretization is acceptable for natural large scale floodplains as the Amazon, it is insufficient in a heavily controlled environment as the Mekong Delta, where the inundation is mainly controlled by dikes, respectively dike overflow. In order to capture this feature, the information of the dike lines and elevations would have to be mapped into the DEM, which itself is a major effort, and then represented as either 270 m wide dikes resulting in unrealistic inundation areas and processes, or with a higher spatial resolution of the DEM causing even longer computational times. Another important aspect is the high complexity of the channel network. Simplified 2D inundation model like LISFLOOD-2D are designed for tree-like channel topologies and are cannot cope with the multiple layered circle systems as present in the Mekong Delta.
Considering these computational cost implications and structural model demands, the choice of a sophisticated 1D approach is therefore without alternatives. In order to represent the important floodplain inundation within a 1D model, we developed the

presented approach of shallow cross sections for the floodplains, connecting via weir structures to the channels, representing the dikes. From our point of view, this approach is acceptable, because of the low topography and unstructured morphological characteristic of the floodplains. The main floodplain inundation features - the dike overflow and the flood propagation over an almost flat surface - can be represented in sufficient detail and resolution.

Ad 2. Derivation of inundation maps
The flood extent maps were provided by German Aerospace Center (DLR) utilizing ENVISAT ASAR radar imagery. The Advanced Synthetic Aperture Radar (ASAR) on the platform uses the C-Band. Data used for generation of the inundation maps are acquired in wide swath mode (image mode) with a geometric resolution of 90 meters to ensure the coverage of the whole Mekong Delta in one dataset for one point in time. The derivation of the inundation maps is performed by a histogram threshold based approach similar to the one described by Schumann et al. (2009). The implementation of the method used for inundation map generation (Huth et. al, 2009) was integrated in an automated processing chain for standardized and repeatable processing of inundation time series (Gstaiger et. al, 2010). The histogram threshold based method is based on the assumption that water surfaces are forward scattering the radar signal resulting in low backscatter signals to the sensor. It uses multiple grey level thresholds and image morphological operations. The derived inundation maps were validated by several field surveys in the Mekong Delta. The accuracy at the borders of the inundation maps is estimated as 1-2 pixels, i.e. 90-180 m for the ASAR derived inundation maps. Further details can be found in Gstaiger et al., which is now accepted for publication. Therefore more details cannot be given here for copyright reasons.

Ad 3. Optimization criteria F2 - definition and evaluation

As mentioned in the manuscript we used the most common measure for comparing simulated and observed inundation extends. The deficiencies of this measure, i.e. the bias towards large inundation extends, are known and reported. However, due to lack of better alternatives up to date, it is still the basic measure used and recommended for deterministic calibration (see Schumann et al. 2009 for a review). In our study we accept this limitation and put the focus more on the development and testing of automatic calibration routines, instead of improving the goodness of fit measure for the inundation extend. An advantage of the proposed approach is the assignment of weights to the individual inundation maps. By this this the deficiencies of the flood area index could be partially compensated by putting more weights e.g. on the small inundation areas at the beginning of the flood season. However, in the presented study the focus was on the large inundation extends. Therefore we put more weight on the maps during the high flood season. Testing the inverse way would surely be an option for another study.
Another issue in this field is the method of comparing the 1D simulated inundation extends with the 2D inundation maps. The straightforward approach, as correctly pointed out by the reviewers, would be to interpolate 2D inundation maps from the water levels at the 1D simulation nodes and intersection with the DEM. However, the main obstacle here are the dikes, respectively their absence in the DEM. Any automatic interpolation approach is therefore highly error prone: corners of the floodplain compartments will be filled erroneously, large areas may be filled despite their full protection by dikes, etc. Reliable derivation of inundation areas thus requires manual adjustments, which is the reason why we refrained from using interpolation in the automatic calibration I this study.
We therefore directly compared the inundation levels at the center nodes of the floodplain compartments of the model with the inundation maps. In order to consider the uncertainties associated with both inundation simulation and maps, we developed the proposed method. The fuzzy set defining the probability of the floodplain to be flooded is based on the actual micro-topography within the floodplain compartments.

As mentioned in the text, the overall flat area is divided into paddy fields enclosed by mini-dikes of about 30 cm height. Thus a complete inundation of the compartment, at least in the vicinity of the simulation node can only be assumed at water levels above 30 cm, whilst water levels below 5 cm are considered as inundations that a) cannot spread wide because of the micro-topography of the ground surface and b) can hardly detected by the ASAR radar satellite due to their highly intersected surface ("water in paddy fields" problem). By taking not only the single pixel of the inundation map covering the simulation node into account, we also consider the uncertainty in the inundation map generation. As mentioned before, the classification error especially along the border of the inundation area is 1-2 pixels. Thus we used the described buffer of inundation map pixels around the simulation node to derive an estimation of the uncertainty in the inundation map that covers both geo-referencing as well as classification errors, resp. accuracy of the inundation maps. The proposed method therefore does not introduce subjectivity into the evaluation of F2, it rather considers uncertainty aspects in a probabilistic way. But of course, the presented work is surely not the end of the story in spatial comparison of simulated and observed inundation areas. The presented calibration framework is open to any development in this field.

Ad 4. Definition of roughness parameter ranges
Regarding this point we got differing comments from the reviewers. While Guy Schumann asked for an extension of the range in order to allow for more compensation of model errors, reviewer 3 asked for a narrowing of the ranges in order to keep the values within typical roughness values. In principle we agree with Guy Schumann, that the ranges of the parameters should be selected as far as possible in order to compensate model errors through roughness calibration. When calibrating it is almost unavoidable that the roughness has to account for model deficiencies. This is often reported and accepted. Thus only an almost error free model will provide good results with typical parameter ranges. The reason why we did not set wider ranges for the parameters is model stability. In sensitivity runs we observed that the model got

unstable for values exceeding the defined ranges. Given the high non-linearity of the St.-Venant equations and the complexity of the model this is not surprising. So to sum this up, the ranges defined here were set as wide as possible to allow for maximum error compensation while still enabling stable model performance.

This part is used to reply to addressed points commented by G. J.-P. SCHUMANN (REFEREE 1)
RC1-j means the comment number j from the Referee 1
AC1-j means the corresponding answer to RC1-j

RC1-1: In the abstract, the authors claim that 'calibration of hydrodynamic models is still un-derdeveloped' - to avoid confusion and misleading readers, I suggest the authors elaborate this statement, probably emphasizing that they mean multi-objective or Monte Carlo based calibration as is often performed with hydrological models
AC1-1: We agree with the comment of the Referee. It should be rewritten as: "Automaticand multi-objective calibration of hydrodynamic models is. . ."

RC1-2: Please check reference Schumann et al., 2009b in line 25 on p. 9179. Should this be Schumann et al. 2009a?
AC1-2:Yes, the reference should readSchumann et al. 2009a.

RC1-3: L11P9180: How are these ?ood maps obtained? The authors should brie?y describe which sort of algorithms were used.
AC1-3:Please refer to the above general point 2.

RC1-4: L18P9180: How large is the model domain? More detail about why a 1D model is the only feasible option should be given. For example, what about a simple,

relatively coarse 2D model? Is it because of the complex micro-topography on the ?oodplain that the authors preferred a quasi 2D model built from a 1D structure?
AC1-4:see general points 1 above

RC1-5: Section 4.1: is 0.05 as an upper bound a reasonable value given the likely degree of compensation that might be needed to account for errors in the geometry and boundary conditions over this large model?
AC1-5:see general point 4

RC1-6: Section 4.3.1: the NS objective function performance depends largely on the quality of the model boundary conditions; can the authors comment on this effect?
AC1-6: In this study, we assume that there is no uncertainty in boundary conditions as well as in the hydrographs used for calibration.But it is planned to include both boundary as well as calibration uncertainty of F1 in the calibration routine in future. However, at the moment we consider the model uncertainties higher than those in the hydrographs because of the complexity and large scale of the model.

RC1-7: Section 4.3.2: the F measure (F2 in the paper) is known to be biased towards large inundation extent (the performance increases with larger inundation and thus may lead to unidenti?able parameter spaces). Although no alternative for a spatial performance measure has so far been proposed in the literature it is worth noting this.
AC1-7:see general point 3

RC1-8: An alternative way to assess a (1D) model with a satellite image - and in this study this would have facilitated the assessment as there would have been no need to transform from a 1D output to a pseudo 2D (P9188) - would have been to take the width of the inundation at each cross section and then measure the inundated width at

that same location on the SAR images thereby allowing a simple RMSE distance to be computed that is less biased and also less affected by a possible geopositioning error of the images. The use of F here (F2) is my main concern. The authors show in Table 5 that according to F1 (NS) their model calibrates well, so by trying to maximise F2 (the F measure) the authors may have introduced errors other than the model structural error given that F prefers larger inundations (which illustrates itself by lowering the dike elevations which increased F2). The ?nal calibration shown in Table 6 demonstrates that trying to reach a maximization for both measures very much degrades the performance of the water levels in the channel (from 0.93 to 0.759), which is the actual capability of the calibrated 1D model.Would the authors have gotten to the same conclusion about model structural errorsby using a less biased inundation width RMSE???
AC1-8: See general point 3.
Also we take this opportunity of this comment to clarify the representation of the floodplains in Vietnam following the short description (for the reason of keeping the manuscript not lengthy) stated in Section 3 (p9180). The floodplains in Vietnam are separated intoa multitude of dike enclosed compartments which is much higher than the number of modeled compartments.In setting up the model simplifications of the floodplains and their compartments were made. The floodplain compartments are modeled by artificial channels with low and wide cross sections extracted directly from the DEM. Figure 4 illustrates the approach for a typical rectangle compartment enclosed in dikes. The dikes' width mostly varying from 3m-6m which at present could not be described in any DEM for the whole Mekong Delta. The width of the cross sections of the artificial branches is set as to equalize the volume of the crossing artificial channels to the volume of the actual compartment in order to avoid volume errors. I.e. in most cases the cross sections of the artificial channels are about half the width of the floodplain compartment they represent. Therefore the method proposed by Guy Schumann does not work when the artificial channels are filled to their bank width.

Also it has to be noted that we do not transform the 2D results to pseudo-2D results. Instead we directly compare the water level at the center node representing the floodplain directly with the inundation map. In this procedure we also consider uncertainties involved, see general point 3 above.
In the other point made by the Referee 1, possibly wrong conclusionsfrom the values of the Table 6 were made. The results reported there were derived with the Pareto optimal solution which has the closest distance to the optimal pit of (1,1). This means that we selected a parameter set from the final Pareto optimal solutions which is the best compromise between performance for F1 and F2, which is a common choice from Pareto optimal solutions. Therefore the performance in F1 is lower than in the best solution for F1 reported in Table 5. We used this parameter set test the hypothesis of wrong dike elevations in the model. As Table 6 reports, the besides F2 also the performance of F1 is improving with lowered dike elevations using the same parameter set, although only slightly. The results presented in Table 6 are not the maximized performance for the individual performance criteria! We will make this clearer in the revised version of the manuscript.

RC1-9: P9187-9188: why this 'subjective' uncertainty introduction of modelled ?ooded cells in 2D? Why not simply use the DEM to create a 2D information on the number of inundated cells using the water level modelled in the channel, i.e. using a simple interpolation approach?
AC1-9: see general point 3

RC1-10: Section 5: the authors hardly mention the quality of the processed SAR data as a limitation in their calibration procedure
AC1-10:see general points 2 and 3

References

Wilson, M. et al., 2007.Modeling large-scale inundation of Amazonian seasonally flooded wetlands.Geophys. Res. Lett., 34(15): L15404.
Gstaiger, V., Gebhardt, S., Huth, J., Wehrmann, T., 2011. Multi-sensoral derivation of inundated areas using TerraSAR-X and ENVISAT ASAR data. International Journal of Remote Sensing, accepted for publication.
Huth, J. et al., 2009. Automated inundation monitoring using TerraSAR-X multitemporal imagery, European Geoscience Union Annual Assembly 2009, Vienna, Austria.
Schumann, G., Di Baldassarre, G., Bates, P.D., 2009. The Utility of Spaceborne Radar to Render Flood Inundation Maps Based on Multialgorithm Ensembles. IEEE Transactions on Geoscienceand Remote Sensing, 47(8): 2801-2807.
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. Rev. Geophys., 47.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 7, 9173, 2010.