Comment on bg-2023-108

Reply to RC1

The paper discusses the use of 3 different remote sensing products for the calibration and validation of 3 hydrodynamic models of different scales: AirSWOT to calibrate friction coefficient, UAVSAR to calibrate errors in the intertidal marsh topography and AVIRIS-NG to calibrate sediment properties.

The authors introduce the traditional calibration techniques for such a model, being the use of time series of observations (e.g. water levels, sediment concentration, etc.) and emphasise how remote sensing observations can complement such technique. I believe they innovative character of the research is clearly presented.

We want to thank the Reviewer for providing feedbacks and comments, which greatly helped better presenting our findings. In the following lines, we reply (in **black**) to each comment (in **blue**) and refer to the changes in the manuscript. Modifications are reported between "" and with *italicized* text. At the end, we also report all references to papers cited in the answers. We were not able to add all revisions. In these cases, we refer to the changes we will make in the manuscript instead of providing the actual modification.

Major comments:

Regarding the use of the AirSWOT data: both the calibration and validation are clearly presented. However, whether the use of AirSWOT actually leads to more accurate model results than calibration using time series is less obvious. Can you calibrate the model using the airswot and using time-series to show in a validation period that the airswot actually appears to be better? In the discussion, could you compare your evaluation statistics with other similar hydrodynamic models of intertidal areas? What are typical RMSE values? If needed, the authors could also calculate other evaluation statistics such as the Nash & Sutcliffe model efficiency to allow comparison with more existing studies.

We extensively updated this part of the paper. We followed the suggestion and performed two separated calibrations using AirSWOT and timeseries. By doing these we also realized that we could achieve better performance using a friction coefficient of 65 $m^{1/2}s^{-1}$. In the new version of the manuscript, we use the same ranges evaluated in the Atchafalaya model and compare results between 45 and 65 $m^{1/2}s^{-1}$. By doing so, we were also able to detect a few small calculation errors we previously made. Interestingly, we see that in the calibration with AirSWOT, 65 $m^{1/2}s^{-1}$ is the best value, while in the timeseries calibration we have similar results but slightly better in the 45 $m^{1/2}s^{-1}$ case. We then use the two coefficients in the Delta-X Fall campaign to validate the water levels and find that 65 $m^{1/2}s^{-1}$ works better by comparing water levels gauges. We used the RMSE and Model Efficiency index as suggested to compare performance. In the discussion, we added a few comparisons with existing studies to confirm the goodness of our results.

UAVSAR: The method is only very briefly explained how the use of water surface elevation changes can be used to calibrate errors in the marsh topography and the authors refer to another paper where more details can be found. As this is an essential part of the paper, I would like to suggest a slightly more extensive explanation of this method. Furthermore, could you address how the UAVSAR - topography calibration can be validated and why no direct validation of the calibrated parameter (being the marsh topography) is included in the paper? Finally, assuming a uniform friction coefficient is very likely not to be the case in reality and how do you ensure that the calibration of the marsh topography does not try to resolve these types of errors instead of errors in the marsh topography.

We agree on the fact that the procedure is important and with the suggestion of providing directly more details instead of referring to the paper that developed the method. To address this, instead of presenting the principle behind the method, we now explicitly explain the iterative procedure, so that it is more understandable to the reader. Furthermore, we added a new figure in the Supplementary Material, in which we provide a simple flow diagram showing the method (the figure is adapted from Zhang et al., (2022a)). The new method description reads: "*To correct marsh topography, the methodology proposed by Zhang et al. (2022a) was followed, in which errors in topography were removed by comparing modelled water-level changes with those observed via UAVSAR (see Figure S1). In this procedure, the first simulation is run using the original topography and the difference between modelled and observed water level change is computed in each cell. If the modelled water level change is larger than UAVSAR, the elevation of the cell is increased. In the opposite case, the elevation is decreased. The updated topography is used in the subsequent simulation. The procedure is run with the updated topography iteratively until the minimum RMSE is reached*".

One way to validate the modified bathymetry could be comparing the water level change in a different time using a different UAVSAR flight line. More specifically, one could consider a water level change measured between two different times. In our case, we chose a change in 3.5 hours. In the description of the procedure, Zhang et al. (2022a) point out that as long as the method converges to a unique solution with realistic values, using a different time would lead to similar results. Another way to validate is considering ground-truth values of elevation measured in-situ. In previous version of the manuscript we did not include any validation, however, in the new version we provide a validation of the results using RTK measurements collected during the Delta-X Spring 2021 campaign and RTK measurements collected by the CRMS. Despite some points show that the original topography was closer to the real value, the coupling with UAVSAR allowed us to reduce the error. It must be noted that the measurements cover a very limited area of the domain. A better validation would consider more spatially distributed point; however these are not available. We added a sentence in the subsection 2.5: "*To validate the topographic correction we used compared the original and calibrated topography with Real Time Kinematics elevation measurements collected during the Delta-X Spring 2021 campaign (Twilley and Rovai, 2022) and site 421 elevation provided by the CRMS*". We decided to show the validation figure in the Supplementary Material (Figure S2).

Regarding the friction we added a more detailed point in the discussion in the manuscript. Drag exerted by bottom and vegetation effects the flow. Zhang et al. (2020) showed that model performance did not significantly change for different Chezy coefficients, showing that elevation in these areas exerts a strong influence. They point out, as the Reviewer suggests, that the correction of the topography contains information of the friction. Moreover, they observed that if one would run the calibration process only on friction, this would lead to unrealistic values of frictions and unrealistic spatial distribution. Thus, a good practise would be to set an initial friction, calibrate the topography, and then adjust for friction by paying attention to stay reasonable ranges. In this case, we did not calibrated friction. We added a few sentences in the discussion of UAVSAR: "*As suggested by Zhang et al. (2022), friction plays a marginal role in affecting water levels on marsh platforms. They run a sensitivity analysis on the effect of variable friction by exploring a wide range of values finding little effect on model performance. The calibration of the topography inherently contains information of the friction, which can lead to its smaller effect on the computed flow. At the same time, applying the same procedure to the friction (without modifying the marsh elevation) would lead to unrealistic large spatial variation of the friction coefficient, with some unrealistic* *values. Therefore, it was decided to only change marsh elevation to match modelled water-level variations with those derived via UAVSAR*".

AVIRIS-NG: While the calibration process is clearly explained, I believe more emphasis should be put on validating whether the use of remote sensing indeed improves the model over the use of single point data. The author mentions in-situ observations. I would suggest calibrating the model using the in-situ measurements and independently calibrate the model with a AVIRIS-NG image. Then, independent validation (based on either in-situ measurements and/or AVIRIS-NG imagery of a different time period) could indicate whether the model calibrated with the AVIRIS-NG image indeed performs better.

We agree with the evaluation of the reviewer. More effort should be put in the validation of the suspended sediment. As suggested, we will use the TSS measurements that cover our area. We will add an independent calibration with in-situ measurements and validate the results of both calibration using measurements in the Delta-X Fall campaign. We will add the new figures in the Supplementary material.

Minor comments:

1. Line 134: Could you add the uncertainty on water level measurements.

   For each flight line there is an error file (named *err* in the Denbina et al. (2022) dataset) which contains an estimate of the vertical error for each pixel and provides a spatially-varying estimates based on the interferometric correlation. For both Delta-X campaigns, AirSWOT was validated using in-situ gauges and it was found a Root Mean Squared Error of 9 cm across the entire campaigns when AirSWOT was averaged over a 1 km^2 area. We added the last information in the manuscript at the end of the AirSWOT data presentation: "*Water surface elevation data were validated in both Delta-X campaigns using in-situ gauges and a root mean squared error of 9 cm was found when data were averaged on a 1 km^2 area*".

2. Line 108: Could you explain why small-scale Terrebonne model domain was chosen as a small-scale region of interest? Could you support that this is a representative area?

   Since the large-scale Terrebonne model covers a very large area (about 90 km east to west), it has a coarse resolution in order to be more computationally efficient. The downside is the inability to reproduce small features. Given the microtidal range, differences in water level change on the marsh depends on small topographic differences, that such a coarse grid cannot capture. Hence, the small-scale model allowed us to incorporate UAVSAR data in our numerical modelling framework. As we mentioned at the beginning of paragraph 2.1, the Terrebonne Basin has been constantly losing marshland. The small-scale domain is located in these degrading salt marshes. Thus, the processes in these areas can be considered representative for most of the Terrebonne Basin.

3. Line 164: could you support the decision for the chosen Chezy coefficients for ocean and marsh platform and on line 186 could you explain why these values differ from the large-scale Terrebone model? Same comment for the values mentioned in line 206.

   The Chezy coefficient range is chosen based on different modelling studies of coastal marshes and deltas. We forgot to add those references. We added a new sentence indicating this: "*The selected values fall within a range considered by several coastal marshes and deltas modelling studies (e.g. Edmonds and Slingerland (2010); Nardin eta al. (2013); Stark et al. (2015); Zhang et al. (2019))*". This applies for both Terrebonne and Atchafalya.

In the new version of the manuscript we improved the AirSWOT analysis and now we explore the same friction range in the two large-scale models. A different Chezy coefficient in the channels in the small-scale model in Terrebonne was used simply because the two models were developed at different time. However, the differences between using 55 or 65 $m^{1/2}s^{-1}$ are minimal and do not significantly affect model results.

1. Figure 4 & Figure 7: Spatial patterns are difficult to observe, could you make the map larger (for instance, by rotating the map so that the flight line is either horizontal or vertical)? Could you change the colour scale to stretch the values in the raster map? It seems both negative and positive values on the difference maps never - and + 0.75 m.

   We have modified both figures to better observe the patterns. We have enlarged the flight line in Figure 4 to focus on the south area of the flight line (similarly to Figure 5). Now the figure is clearer. We have also decided that in order to be consistent with the Figure 5, we show only one flight line instead of four (this allowed us a clearer figure). We included a figure with the four flight lines in the supplementary material so the reader can see a comparison with the same flight line in different time of the tidal cycle. We have also reduced the ranges of the error colorbars.

2. Figure 5 and 6: Could you make these maps also bigger? I would propose to only show the colorbar for elevation once and the error colorbar once and if needed, drop the grid labels and add a scale (like figure 4). The scale and zoom of figure 2 is a good example. The use of a single colorbar is shown very well in figure 7. In figure 5, subfigures are not labelled.

   We have modified Figure 5 and Figure 6 and followed the suggestions. Both figures are now larger and more readable. There is only one colorbar for the evelation and error in Figure 5 and only one colorbar for the water level change in Figure 6. In the latter, we also added a land/water mask to better separate channels and marsh and change the order of the subfigures with a better one. We added labels where missing.

3. Line 257: unclear which areas are considered critical and why.

   The critical areas are mentioned in the previous lines: the bottom an upper area. However, it was difficult to connect them to the ‘critical areas’ later mentioned. Those are considered as such because they show the higher disagreement between simulation and UAVSAR after the first run. We modified the text to correct the unclear sentences. Now the paragraph reads: “*Results from the first run (Figure 6e and 6g) highlight two critical areas. First, the model overestimates the water level change in the bottom area, indicating an error in the marsh elevation derived from LiDAR data. The opposite is verified in the northern area, where waters are found to recede too slowly and consequently underestimate the water level change*”.

4. Line 309: could you refer to the specific figure which supports this statement.

   We forgot to refer to the Figures in the Supplementary Information. In this case the correct reference would have been Figures S1 and S2. However, in the new version of the manuscript we removed these figures to show the results of the calibration with timeseries and validation of the calibrated Chezy in a different period. We will modify the text accordingly with the new references.

5. Line 267: switch order or ‘always’ and ‘to’.

   Corrected. The sentence now reads: “Here, the model tends to always overestimate sediment concentration and the error decreases as critical shear stress and settling velocity increases”.

6. Line 268: switch order of ‘better’ and ‘performs’.

   Corrected. The sentence now reads: “Overall, the model performs better in the more open sections of the flight line, while at the extremities, performance declines”.

7. Line 269: drop the first ‘best’ and the sentence misses an active verb (online 2017 change ‘as’ to ‘is’).

   Corrected. The sentence now reads: “The combination that provides the best comparison with measurements is a critical shear stress of 0.1 Pa and a sediment settling velocity of 0.25 mm/s”.

References

Denbina, M., Simard, M., and Rodriguez, E.: Delta-X: AirSWOT L2 Geocoded Water Surface Elevation, MRD, Louisiana, 2021, Version 2. ORNL DAAC, Oak Ridge, Tennessee, USA, https://doi.org/10.3334/ORNLDAAC/2128, 2022.

Zhang, X., Jones, C. E., Oliver-Cabrera, T., Simard, M., and Fagherazzi, S.: Using rapid repeat SAR interferometry to improve hydrodynamic models of flood propagation in coastal wetlands, Advances in Water Resources, 159, 104088, https://doi.org/10.1016/j.advwatres.2021.104088, 2022a.