Comment on bg-2023-108

Reply RC3

I strongly agree with the central argument of this paper, that spatial information from remote sensing provides critical data for calibrating and testing numerical models of coastal wetlands. The paper reports on results of an experiment in which 3 remote sensing products were used in conjunction with modeling of a region of coastal Louisiana to evaluate the value of combining these approaches. Overall I think the analysis is thorough and informative, but the text would benefit from editing for clarity. Most of my comments are editorial.

We want to thank the Reviewer for positive support to our study. We are grateful for all comments, and editorial suggestions, which allowed us to better show our analysis. In the following lines, we reply (in **black**) to each comment (in **blue**) and refer to the changes in the manuscript. We report the related modification in the manuscript 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.

Comments:

L18: "Peter Sheng et al." should just be "Sheng et al."

We doubled checked and the correct citation is Peter Sheng (see https://www.nature.com/articles/s41598-022-06850-z#citeas)

L19: There is a recent paper by Temmerman et al (2023; Reviews of Marine Science) that would be appropriate to cite here.

The suggested paper is very appropriate here, thus we added the citation.

L48: Does Gross Primary Productivity need to be capitalized?

No it does not. We corrected the mistake.

L60: change to: ", and coastal wetland above-ground biomass …"

Changed.

L106-112: The statement is made here that there are 3 models at different scales, but it seems that the scale of the large-scale Terrebonne model and the Atchafalaya model are essentially the same.

That is corrected. The large-scale Terrebonne model and the Atchafalaya model are developed at basin scale, and they are comparable (they also have the same spatial resolution). The paragraph was not precise, hence we added an additional sentence to clarify: "*Two models were set-up at the basin scale, while the third one was developed at smaller scale*".

L144: "TSS concentration from the …"

Corrected.

L145: Should the phrase "and vegetation structure are produced" be deleted? It seems the same point is made in the next sentence.

Yes. These lines were badly constructed. We reformulated and added more information regarding AVIRIS-NG and the algorithm to retrieve TSS: "*Local empirical algorithms derived using in-situ measurements were used to derive TSS concentration from the Rrs(λ) in the visible/near-infrared region and generate maps of TSS from the AVIRIS-NG imagery (Gao et al., 1993; Bue et al., 2015; Jensen et al., 2019). The in-situ samples were collected in both Terrebonne and Atchafalaya basins during the Delta-X 2021 Spring and Fall campaigns in order to capture high and low flow conditions. The algorithm to retrieve TSS from AVIRIS-NG performed well (Median Absolute Percent Difference 13.7\% and Median bias 6.71 mg/l) across a wide range of TSS concentrations (0.1–154.5 mg/l) (Fichot and Harringmeyer, 2021, 2022). AVIRIS-NG images were also used to produce maps of vegetation structure (Jensen et al., 2021)*".

L162: What is the resolution of the usSEABED database used to specify the initial sediment distribution in the model? Why use mud and sand (2 classes) in the large-scale Terrebonne model and mud, silt and sand (3 classes) in the large-scale Atchafalaya model?

The usSEABED provides sampling points data where the percentage of the different fractions are indicted. We took these points and interpolated on the 90 m grid. Most of the sampling points simply separate mud and sand, which is the separation we adopted in the large-scale Terrebonne model. Note that the interpolation approach was already taken by Liu et al. (2018), which modelled the same area. The Atchafalaya model was subsequently and separately developed from the Terrebonne one, and we simply took a different approach. In this area, there is a much smaller number of samples, thus we retrieved an average fraction weight (in percentage) and set a uniform initial concentration. From the measurements we got that the 22% is sand and 78% is mud. The 78% was then divided in half between clay and silt. In this case, we performed a spin-up of the bed by running a one-year simulation with a morphological factor of 50 to reach equilibrium.

Figure 2: Perhaps the caption should refer to Figure 3 for the spatial relationship among the 3 imaged regions.

Yes, this could help the reader better locating the flight lines. We added the reference in the caption: "*Refer to Figure 3 for the spatial relationship*".

L172-176: What value was used for the erosion parameter? Why wasn't this parameter varied in the calibration runs? Its effect is different from the critical shear stress.

The erosion parameter was fixed at 0.00001 kg $m^{-2}s^{-1}$. We agree with the reviewer. The erosion parameter has its own effect in the Partheniades-Krone formulation (Partheniades, 1965). In this case, we did not consider variability due to the fact that, despite having two classes of sediment, most of the bed is composed by mud (sand is located more of shore along the barrier island). The value is within the typical range suggested by Winterwerp et al. (2012). The same authors highlight the fact this parameter is more or less constant in the case of homogeneous beds. We acknowledge that this is a limitation. The bed is unlikely that homogeneous and some grain size variability should be expected, thus the erosion parameter could be different from what we assumed. The one selected is a very common value within the range $1\text{-}5\cdot10^{-5}$ kg $m^{-2}s^{-1}$ for modeling studies of coastal bays (e.g.Ganju and Schoellhamer (2010); van der Wegen et al., (2011); Wiberg et al., (2015)). We added this information into the methods section.

L178: CRMS station 421 is referred to in this paragraph, but the CRMS acronym is not introduced until the end of section 2 (L226).

We corrected an introduced the acronym earlier when we mention station 421 in the small-scale model description: "*Delft3D was utilized to simulate the hydrodynamics in one of the Delta-X intensive field study sites near station 421 of the Coastwide Reference Monitoring System (CRMS) network (blue rectangle in Figure 1) from 25 March 2021 to 18 April 2021*".

L182: Here and elsewhere, rather than refer to the bottom and upper boundaries of the model grid it would be better to use terms south and north boundaries.

We corrected the inaccuracy everywhere and used the term north and south as suggested.

L183: Where is Trouble Bayou and how far is that from CRMS station 421?

Trouble Bayou is located about 5 km north of station 421 and about 3 km north of the northern boundary of the small-scale model. We added this information in the text: "*The former were the water levels measured by the CRMS site 421, while the latter were computed using water levels measured by CRMS at site 421 and Trouble Bayou located about 5 km north of station 421*".

L186: Why is the Chezy coefficient for the channels set here (55) while the Chezy coefficient for the channels in the large-scale Terrebonne model setup allowed to vary (between 40-45) and over a smaller range of values?

This was a limitation in the large-scale Terrebonne model in the previous version of the manuscript. We now explore a wider range (the same as the Atchafalaya model) and found better results. Now, like in the Atchafalaya model, the best Chezy coefficient is 65 $m^{1/2}s^{-1}$. This is still different from the 55 $m^{1/2}s^{-1}$ used in the small-scale model, however differences between 55 and 65 $m^{1/2}s^{-1}$ are minimal, thus the slightly different values do not affect the results in the small-scale model.

L189-190: Couldn't a change in vegetation roughness affect the water-level changes in addition to marsh topography?

Yes. This an important point raised by another reviewer. In this case the Chezy coefficient does not account for vegetation but only the bottom of the marsh. The 35 $m^{1/2}s^{-1}$ is a typical value used to represent bottom roughness in modelling studies (e.g. Zhang et al., 2018; Passeri et al., 2018). In order to include vegetation a more correct value would range between 10 and 20 $m^{1/2}s^{-1}$ (e.g. Augustijn et al., 2008; Stark et al., 2015) or implement formulation such as Baptist (2005). In this case, our objective was to calibrate elevation and not friction. Calibrating only friction instead of wetland elevation would have led to unrealistic spatial distribution and values of Chezy (Zhang et al., 2022). Indeed, the authors suggest to first calibrate boundary conditions and elevation. Only after these steps, a calibration of the friction would provide more realistic values of friction. For this reason, we simply set an initial and homogeneous friction coefficient. Note that, the calibration of the elevation inherently contains information of vegetation, however, when Zhang et al. (2022) ran a sensitivity analysis on marsh Chezy coefficient found non-significant variation of model performance for all Chezy values (range 8-40 $m^{1/2}s^{-1}$).

Table 1: It could be helpful to provide values of the calibrated parameters (where appropriate) in this table.

Instead of adding the calibrated values in Table 1, we created a second table (Table 2 in the revised manuscript), where we report the best set of parameters for the two large-scale models.

Figure 3: It would be helpful if boxes were used to show the domains of the 3 models in relation to the remote sensing.

We modified Figure 3 and added the boundaries of the three models.

L214: "and the corresponding remote sensing …"

Corrected.

L216: "allow us to tune model parameters …"

For this paper, we decided to use the impersonal sentence structure. Thus, in this case we prefer to keep 'allowed to'.

L241: Delete "For the" before "Atchafalaya model results …"

Corrected.

L246: Why was a Chezy coefficient of 65 not considered in the Terrebonne model? Are any time series available for comparison with the model?

As mentioned in a previous answer, this was a limitation of the model. Now, we have explored a wider range and found that result improve when 65 $m^{1/2}s^{-1}$ is used. We have also added a separate calibration with timeseries and validation using a different period.

L252-3: "water level change in the southern area"?

Yes. Using 'southern' is more appropriate. We corrected.

Fig. 7: It would be helpful to add the date of the image in the figure (as in, e.g., Fig, 8) or in the caption. For the model calibration runs shown in this figure, what settling velocity was used for the results shown in Fig. 7A and what critical shear stress was used for the results shown in Fig. 7B?

We modified Figure 7 accordingly. We added the date and time of the AVIRIS-NG flight. The settling velocity used in Figure 7A, and critical shear stress used in Figure 7B were noted in the caption. However, we added the indication in the figure, so that it is easier for the reader to get the information.

L271: Add a reference to Fig. 8 after noting the date/time of the imagery.

Reference added.

L274: Add reference to Fig. 8B after providing the RMSE of 35.88. Are there any in situ measurements of SSC available for checking model values?

Reference added. Yes, we have available in-situ TSS measurements. We will use them to perform a second calibration that we will compare to the AVIRIS-NG one using a validation period (Delta-X Fall campaign).

Section 4.1-4.3: Are there any reasons to think that the Chezy coefficient might not be uniform?

It is very likely that the Chezy coefficient is not uniform. The area is characterized by patched of *Spartina alterniflora,* thus is it likely to have variable friction. In our case, we are more interested in calibrating the friction. Zhang et al. (2022) suggest to first calibrating the topography using an initial uniform friction. Then calibrating the friction to account for vegetation. They suggest that because a direct calibration of the friction would lead to unrealistic distribution and values. Thus, for the purposes of our analysis, using a uniform friction is admissible. We specified this in the discussion of the UAVSAR coupling results.

L285: Does "lower half of the flight line" = "southern half of the flight line"? If so, it would be clearer to refer to southern half.

Yes. We clarified the expression using southern half.

L285-290: Could the comparison with Manning values be included in results instead of discussion, or is the Manning value so well known that the agreement is a useful point of discussion?

We moved the comparison with Manning in the results.

L306: "The validation of the water levels across the domain …further confirms the goodness of the calibrated friction coefficient" seems like an overstatement given that the previous paragraphs described portion of the model domain where the modeled and remotely sensed water levels do not agree – for reasons that may be unrelated to friction coefficient, but the disagreement makes it impossible to evaluate how well the friction coefficient works in those regions.

In the previous version of the manuscript we were not precise enough. In the new version we have added a new calibration using both AVIRIS-NG and timeseries from gauges, from both calibration and a validation in a different period of the year we show that AirSWOT lead to a better result. The tidal gauges are spread over the domain. We will modify the text and be more precise. We will specify that the validation showed that AirSWOT lead to better results, but there are areas where it is not possible to make this evaluation due to the limitations imposed by the model resolution

L316-327: The argument that friction plays w marginal role in affecting water levels on marsh platforms merits more discussion, since otherwise that seems like a reasonable alternative to arguing that marsh platform topography is poorly quantified. Is there a correlation between where the topography must be adjusted and vegetation characteristics on the marsh? Is it also possible that the model isn't resolving microtopography that affects water fluxes? How much was the topography altered to improve fluxes. Were the values realistic?

This is an excellent point that was not discussed in the previous version of the manuscript. The statement about the marginal role of the friction was not justified. This is a point that was tackled by Zhang et al. (2022) when the method was developed. The calibration of the topography inherently carries information on the friction. The authors ran a sensitivity analysis on the friction coefficient by testing different values (range 8-40 $m^{1/2}s^{-1}$), showing that a variation in friction has no effect on model performance. They point out that it is necessary to first adjust the topography. In the hypothesis of applying the calibration with UAVSAR only on friction, the final friction map would present unrealistic distribution and values because. Thus, they recommend to first calibrate the elevation and then the friction. We added this information in the discussion: "*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*". The modification of the topography produced realistic values. On average it was necessary to decrease the elevation by 1.3 cm with a standard deviation of 16.3 cm. Regarding the point on the microtopography effect on water fluxes, we added additional information, as it was worth of discussion. The method depends on the ability of the model to solve the water fluxes. Therefore, if some areas are not well solved, it could be that the topography is changed even if it is not necessary. For instance, let's suppose there is a small channel that the 10 m mesh cannot represent. During falling tides, the areas near the channel drain faster, hence, they generate a larger water level change compared to internal areas. Since the model cannot solve the channel, the procedure will deepen the topography to allow water to flow and match the water level change, even if the topography is correct. We added this discussion by leveraging on some RTK measurements that we used to validate the calibrated topography: "*this iterative procedure might introduce errors in areas where the topography is correct. This effect can be noted for the three points before mentioned where the marsh is deepened. The methodology depends on how well the model is able to solve all channels. In this particular case, the points are located in proximity to a narrow channel with a 1.5-2 m cross section. The 10 m resolution represents a limitation because features such as channels and levees that are smaller than 10 m cannot be captured by the mesh. Yet, they affect water flow. In this example, we used a UAVSAR captured during falling tide, and in this phase, areas of the marsh close to these channels drains water faster than areas located internally. Since the model does not capture these channels, the method tries to compensate by lowering the marsh to achieve the measured water level change even if the elevation is correct*".

L333: change "worst" to "worse"

Corrected.

L337: Are any data available to evaluate this possibility?

Yes, there are other data available since there were repeat passes of the same flight line. Thus, with an extended analysis this aspect could be evaluated. Certainly, this is an interesting analysis that could be done in the future. For instance, it would be interesting to compare if the bathymetric correction changes (and how much) depending on the use of water level change during rising and falling tides. However, this goes beyond our scope, thus it is not included in the present analysis.

L349: Was TSS sampled just one time and at one location? If so, this doesn't seem like a robust enough test to declare the in situ sample to be more in error than the modeled value. What is the TSS value at that time and place in the AVIRIS-NG data? How much spatial and temporal variability do the model and remote sensing suggest?

There are many samples of TSS done at different times and well distributed in the domain. In order to overcome this limitation, we will perform a second calibration using only in-situ point measurements. Then, we will compare the two calibrations in a validation period (Delta-X Fall campaign). In this way, we will be able to answer the questions and provide a more robust evaluation of the calibration performed with AVIRIS-NG.

L356: Consider revising to "due to flocculation which increases settling velocities compare to …" In any case, the settling velocities used in the model for mud are effectively floc settling velocities.

We revised the sentence. It is correct, in the model these are flocs parameters. Indeed, our goal with the sentence is to remark that the process of flocculation introduces high uncertainty in the settling velocity and critical shear stress because we do not deal with single particles. The emphasis is on the flocs values. We recognized that in the previous version of the manuscript this was not clear. The sentence now reads: "*Especially in coastal areas, cohesive sediment properties are highly affected by flocculation. In this process, fine sediment aggregate to form flocs, for which both settling velocity and bed shear strength of cohesive particles are highly uncertain and difficult to predict*".

L360: Not clear what is meant by: "errors might be related to some the bathymetric modification"

This error is likely related to previous modifications of the text that were not corrected. We eliminated the 'the'. The sentence now reads: "*errors might be related to some bathymetric modification*".

L361: The channels were enlarged in the model? "might have generated"

Corrected.

L363: "inherent to the model. Although Delft-3D has 3-dimensional …"

Corrected.

L365-366: replace "tri-dimensional" with 3D and "bi-dimensional" with 2D.

We corrected all instances of 'tri-'and 'bi-dimensional' with 3D and 2D respectively.

L371: This paragraph might be better combined with the one before.

We agree with the suggestion as the two paragraphs are related. We merged them.

Fig. 9: It would be much better to use the same vertical scale for both profiles. Were the wave conditions much higher on Aug 19 (red curve) than on Aug 17? It seems important to recognize here that optical properties of sediment in suspension is strongly controlled by sediment size, and that the vertical profile of the finest, most optically active fractions might be more uniform that that of coarser bed fractions.

In the first version of Figure 9, we tried to use the same scale. However, the values of water depth (y-axis) have very different ranges. For instance, if we use the red curve range (0-3 m), the blue curve would be squeezed at the bottom of the figure due to its 0-0.5 m range. For this reason, we opted for a double y-axis.

Regarding wave conditions, these were wind speed and direction in the two sampling days:

- 17 August 2021 18:00 UTC: 1.25 m/s 146 degrees north
- 19 August 2021 18:00 UTC: 2.03 m/s 108 degrees north

The sampling point is located within the small-scale model. We do not have measurements of wave conditions in the sampling point (see red dot in Figure 3). Thus, we looked at model results from the large-scale Terrebonne model. Wave conditions were higher on 19 August, but with very small wave height (Figure 1) and orbital velocity near bottom. This seems to indicate that waves do not have an effect on the differences in sediment profiles.

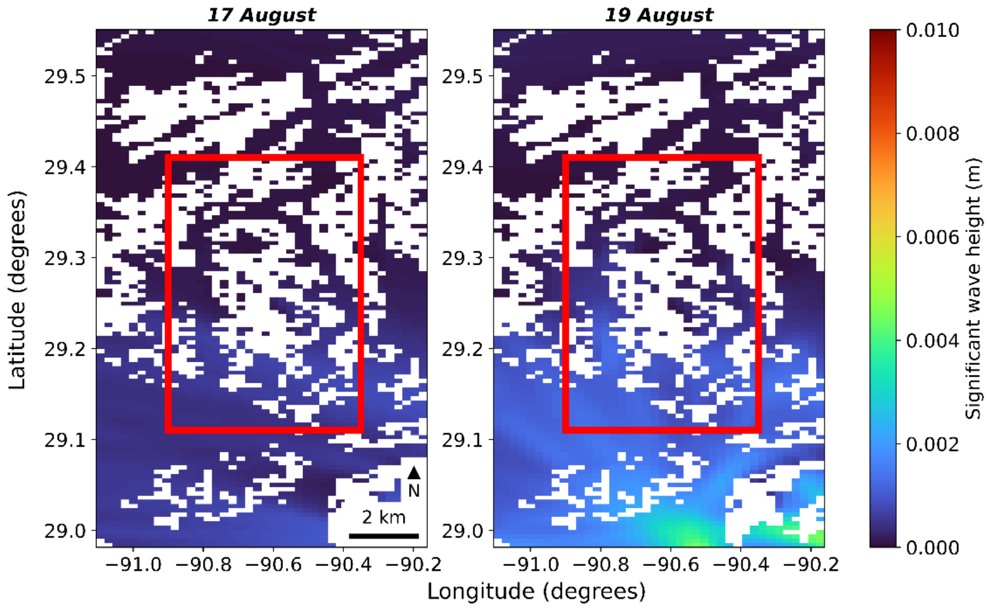

Figure 1. Modelled wave significant height on 17 and 19 August 18:00 UTC. The red rectangle identifies the small-scale domain.

However, there are differences in bottom shear stress, which is likely connected to different water velocity (see Figures 2 and 3). On 19 August both depth-averaged water velocity and bed shear stress are much higher. Interestingly, if we observe the sediment concentration profiles (Figure 4), we can see that only on 19 August the coarser fraction on bed sediment is entrained in the water column, while on 17 August only the finer fraction is measured.

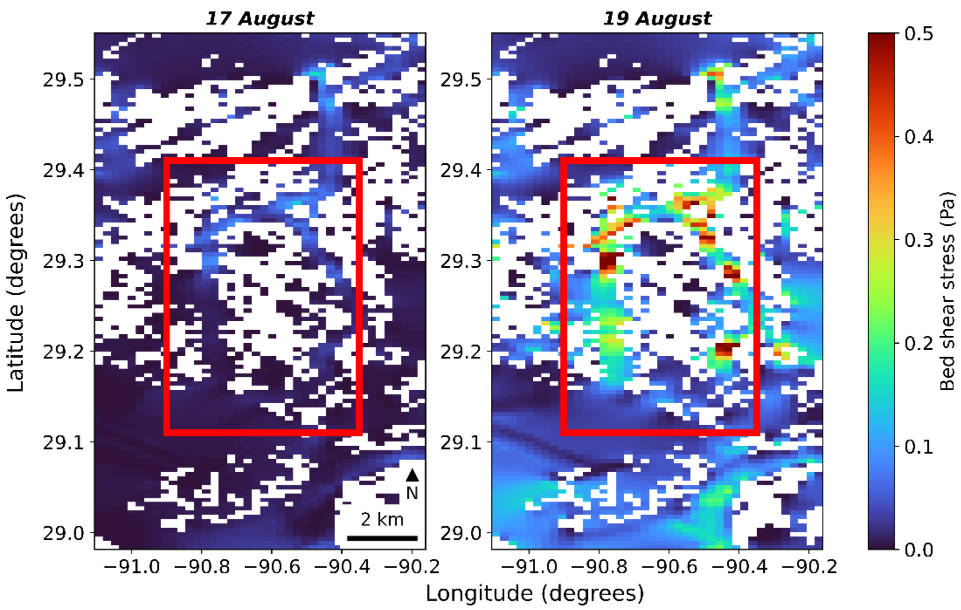

Figure 2. Modelled bed shear stress on 17 and 19 August 18:00 UTC. The red rectangle identifies the small-scale domain.

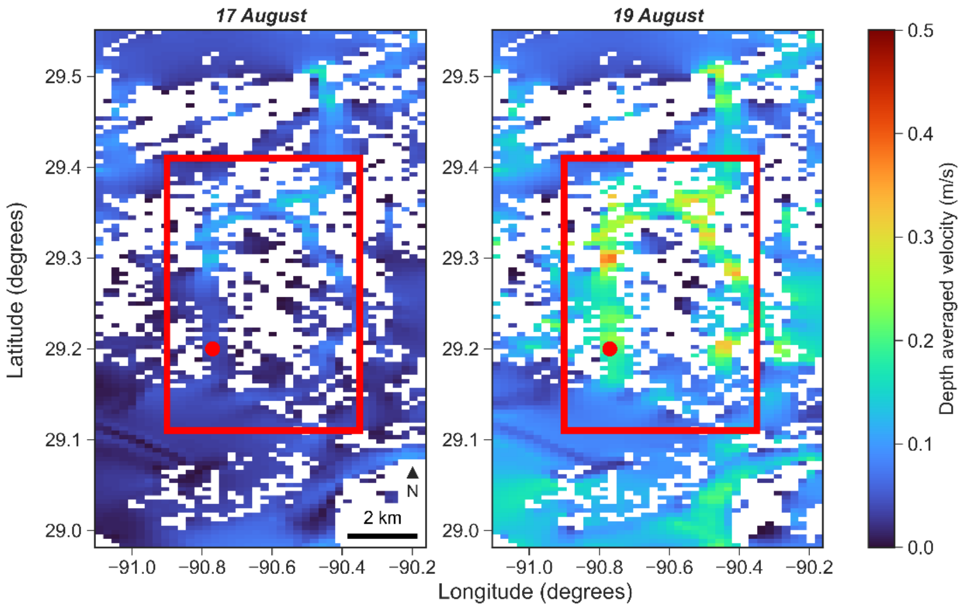

Figure 3. Modelled bed shear stress on 17 and 19 August 18:00 UTC. The red rectangle identifies the small-scale domain. The red dot the sampling point.

The reviewer raises an excellent point. We will add the importance of sediment size in the discussion of the AVIRIS-NG results. We will also modify Figure 9 using the figure below (Figure 4 in this document), to make a stronger case.

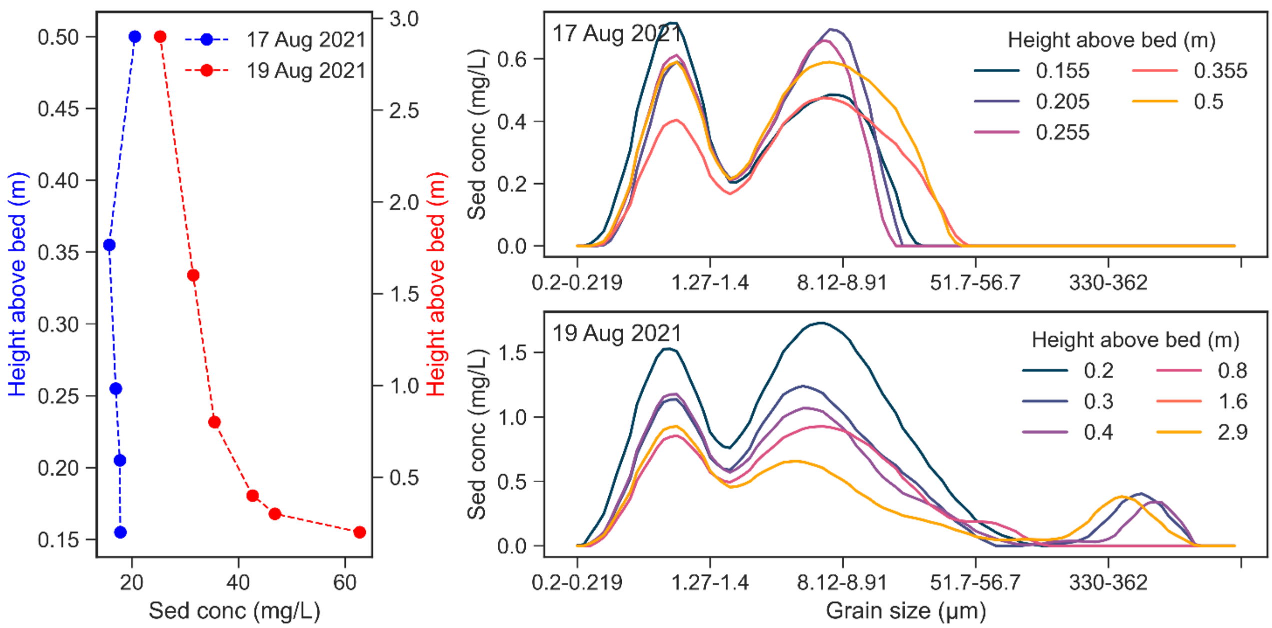

Figure 4. Total sediment concentration profile on the left on 17 and 19 August 2021. On the right the distribution of grain sizes at each depth for both days.

References

Feng, Z., Tan, G., Xia, J., Shu, C., Chen, P., & Yi, R. (2020). Two-dimensional numerical simulation of sediment transport using improved critical shear stress methods. *International Journal of Sediment Research*, *35*(1), 15-26.

Ganju, N. K. and Schoellhamer, D. H.: Decadal-timescale estuarine geomorphic change under future scenarios of climate and sediment supply, Estuaries and Coasts, 33, 15–29, https://doi.org/https://doi.org/10.1007/s12237-009-9244-y, 2010.

Liu, K., Chen, Q., Hu, K., Xu, K., and Twilley, R. R.: Modeling hurricane-induced wetland-bay and bay-shelf sediment fluxes, Coastal Engineering, 135, 77–90, https://doi.org/10.1016/j.coastaleng.2017.12.014, 2018.

Partheniades, E.: Erosion and deposition of cohesive soils, Journal of the Hydraulics Division, 91, 105–139, https://doi.org/10.1061/JYCEAJ.0001165, 1965.

Wiberg, P. L., Carr, J. A., Safak, I., and Anutaliya, A.: Quantifying the distribution and influence of non-uniform bed properties in shallow coastal bays, Limnology and Oceanography: Methods, 13, 746–762, https://doi.org/https://doi.org/10.1002/lom3.10063, 2015.

van der Wegen, M., A. Dastgheib, B. E. Jaffe, and J. A. Roelvink. 2011. Bed composition generation for mor-phodynamic modeling: Case study of San Pablo Bay in California, U.S.A. Ocean Dyn.61: 173–186.

Winterwerp, J. C., Van Kesteren, W. G. M., Van Prooijen, B., & Jacobs, W. (2012). A conceptual framework for shear flow–induced erosion of soft cohesive sediment beds. *Journal of Geophysical Research: Oceans*, *117*(C10).

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.