Response to review comments (R1). Our response in bold.

Review of: “Wave-induced sediment resuspension in the Finnish Archipelago, Baltic Sea: Combining small-scale in situ measurements and large-scale numerical model simulations”

Dear Editor,

The authors present an interesting case study on wave resuspension of sediment in the Finnish Archipelago. The authors have used a spectral wave model to calculate bottom orbital velocities and bed shear stresses, and use this to estimate the threshold of motion for the sediments in the region. Interestingly, the authors have also measured the threshold of motion from sediment samples using a laboratory which has been published (Joensuu et al. 2020). As I have little experience of spectral wave modelling, I will keep my reviews focused on the sediments. However I would ask if wind forcing at a temporal resolution of 3 hours is enough to adequately resolve the storms?

Our response: If a 3 hour wind forcing resolution is high enough to resolve the storms is a valid question. Björkqvist et al. (2018) made Baltic Sea wide simulations using a 6 h temporal resolution in the wind, and that model validated well when compared to extensive measurements. However, in this study the 3 hour resolution was only used for 2017 (based on available data at the time). The model was run for 2017 mainly to compare with available wave measurements, and this validation was good. For the times actually overlapping with the sediment samples, we used an 1 h resolution. This means that the examples that are seen in Figures 7 and 8 are generated using a 1 h temporal resolution in the forcing wind field.

The authors find a large discrepancy between the predicted and measured critical shear stresses, and attribute this to biological and/or biogeochemical properties of the sediments. Concerningly, there is no reporting of the biogeochemical properties of the sediments, so it is difficult assess how accurate their claim is.

The authors seem to think this is the only way to explain the differences in measured and predicted threshold of motion for the sediments – but this is unfounded. Bulk density of sediment has been shown to have the largest effect on the threshold of motion (Thompson et al. 2019). Mixed grain sizes (for example, a bimodal distribution) can increase the threshold of motion (Staudt et al. 2017; McCarron et al. 2019). In short, the manuscript is, at present, not representing the state of the science. The authors need to address their simplification of the problem.

Upon reading (Joensuu et al. 2020), I see that all the relevant information has been collected, but not used in the present work. I encourage the authors to utilise this information, particularly to follow the work of Thompson et al. (2019), who were able to adjust the threshold of motion for their sediment based upon an array of sedimentary and biological variables. As such, I recommend major revisions as a reanalysis of the data is necessary for the manuscript to be up to date with the state of the science.

Our response: Thank you for your constructive review. Based on your suggestions we have made significant changes to how we conduct our study. In summary, we tested (and rejected) the model from Thompson et al. (2019), but instead determined our own similar model using both physical and biological properties directly from the in situ samples. This serves to directly quantify the role of biology based on the samples, and also allows for the biological component to vary in time when the model is applied. It also opens up for a discussion on why the model from Thompson et al. (2019) doesn't necessarily translate to our data. For more details, please see our comments below.

Comments to the authors.

Dear authors, I read your manuscript with interest. This is tricky subject, and I think you need to include some of the sedimentary information from (Joensuu et al. 2020) into your estimate of the threshold of motion of the sediments and use the work of (Thompson et al. 2019) to recalculate your thresholds of motion - it seems like an approach similar to their "model 1" would work well. You seem to have limited your explanation of the difference between modelled and measured threshold of motion to only biological processes, yet there is little reason given in the manuscript to justify this. There are numerous reasons why such a difference could occur, much of this is covered in Thompson et al. 2019, and other papers.

Our response: Thank you for taking the time to review our manuscript. Following your suggestions we have gone back to the raw data of Joensuu et al. (2018, 2020) and used this to both test the model of Thompson et al. (2019) to the degree it was possible, but, more importantly, constructed our own similar model for the critical shear stress based on the data.

Some specific comments:

Equation 8 – the Soulsby Whitehouse equation is incorrectly written, it should be:

$$\theta^*_{crit} = \frac{0.30}{1+1.2D^*} + 0.055\left(1 - \exp^{-0.02D^*}\right)$$

$$\theta^*_{sus} = \frac{0.30}{1+D^*} + 0.1\left(1 - \exp^{-0.05D^*}\right),$$

(see attached image)

I also highlight to the authors that this is the equation for initiation of bedload, not suspended load. There is a Soulsby-Whitehouse like equation for this, fitted by Van Rijn (unpublished, the source is his website):

(see attached image)

I have calculated the suspension values for their grain sizes (Their table 1, my numbers in red). It would appear they have used the correct version of the Soulsby Whitehouse equation, and it was just written wrong in the manuscript.

Our response: We appreciate the attention to detail. This indeed looks like a notational issue and will be corrected in the manuscript.

Sediment class	Grain size (mm)	theta crit (N m^{-2})	theta sus(N m^{-2})
Mud to muddy sand	0.09	0.14	0.16
Sand	0.34	0.21	0.37
Coarse sediments	2.00	1.20	3.06
Mixed sediments	0.15	0.16	? mixed ? (0.21)
Boulders	200.00	178.10	323

The difference between theta crit and theta sus could account for most of the discrepancy the authors have found, especially for the larger grain sizes. It is unclear how they have arrived at their bed shear stress estimate for “mixed” sediments, it appears to be just using the median particle size, which is unjustified. This is unwise as larger clasts (such as gravels and boulders) can “armor” the bed and reduce mobility of all sediments (Wiberg et al. 1994; Vericat et al. 2006). Likewise fine sediments (< sand size) can add cohesion to the bed and reduce hyporheic flow, impacting sediment mobility (Blois et al. 2014; Fox et al. 2014; Parsons et al. 2016; Perret et al. 2018; 2023).

Our response: We have modified our approach based on the extensive comments we got in the review stage. All grain sizes are now based on the in situ measurements only, and the modelled data from EMODNET is only used as a classification. Since no in situ data is available for coarse sediments and gravel, no results are given for those areas. Since we don’t have the grain size distribution, we had no choice but to rely on the median grain size also for the mixed sediments class. However, since our model now also includes the bulk density and chlorophyll a, the results are not only reliant on that particular measure for the grain size.

In particular, i recommend to the authors that they read the paper by Thompson *et al.,* (2019), as this paper works through many of the issues with defining a threshold of motion in complicated sediments, including those with biological controls (for instance, Thompson *et* *al.,* include the concentration chlorophyll-A in their “Model 1”). Moreover, that work found that the bulk density and porosity of the sediment was the overriding control on benthic resuspension, I suggest you try and include this in your work based on the data available in (Joensuu et al. 2020).

Our response: We have tested Model 1 of Thompson et al. (2019) to the degree it was possible. For the parameters not available to us we used similar realistic values to those found in Thompson et al. (2019). This model did not work well for our data. Especially, the authors found that bulk density explained a lot of the variation in critical stress, and show a model (their Fig. 3) with a positive quadratic relationship. Nonetheless, in their Model 1, the coefficient for the bulk density is *negative.* Although their model fit their data, we got unphysical negative values when applying it on our data, because our bulk density was higher. The conclusion that biological factors are not that important also doesn’t necessarily translate between their and our dataset, since their data was collected at significantly larger depths.

We therefore fit our own model to our dataset using three variables: grain size, bulk density and Chlorophyll *a*. We made a fit for each available bottom type (mud, sand, mixed sediment) separately (Figure 1, all except top left), and a fit to all the data (Figure 1, top left). Here, the grain size and bulk density represents the physical conditions, while Chlorophyll A is meant to capture the biological activity. We can see from Fig.1 (top left) that the model can especially explain the variability between the groups.

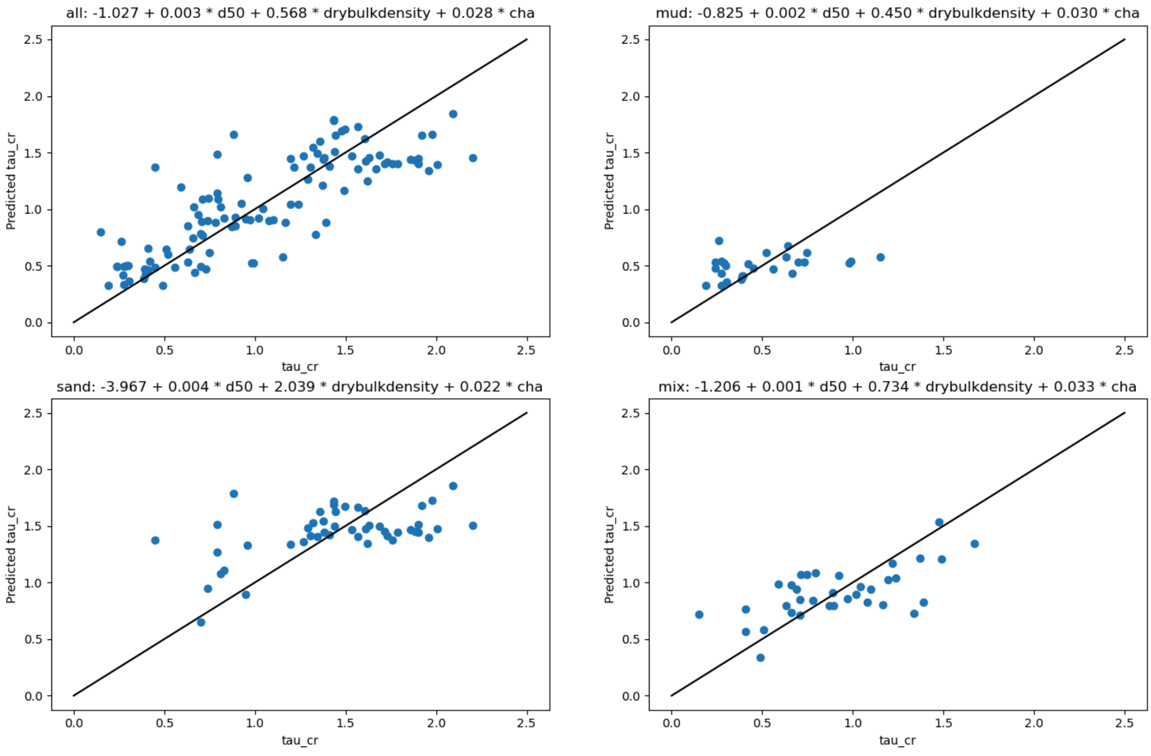

Figure 1. Our three variable model (grain size, dry bulk density and ChlorophyllA) to model the critical shear stress. The model can especially explain the variations between the different classes (top left).

After we fit the model we then determined representative values for grain size and bulk density for each class based on the in situ measurements. While we assumed that the physical properties are constant, we allowed for the Chl A to vary over time. These values were determined as monthly values based on the in situ data (made possible by the temporal dataset of Joensuu et al. 2020).

This spatio–temporal model can now be used for predicting the critical threshold for each class, and these thresholds can be compared to the shear stresses from the numerical model. This significant modification accomplishes several things: a) it puts our work in respect to other similar work, especially that of Thompson et al. (2019), b) it makes a stronger connection between using a combination of measurements and models, where measurements are directly used to determine a critical threshold, while models are used to generalize these using the EMODNET seabed model, and compared to the stresses from the numerical simulations by SWAN, c) it highlights the difference between the physical and biological properties, by allowing the latter to vary in time.

References.

Blois, Gianluca, James L. Best, Gregory H. Sambrook Smith, and Richard J. Hardy. 2014. ‘Effect of Bed Permeability and Hyporheic Flow on Turbulent Flow over Bed Forms’. *Geophysical Research Letters* 41 (18): 6435–42. https://doi.org/10.1002/2014GL060906.

Fox, Aryeh, Fulvio Boano, and Shai Arnon. 2014. ‘Impact of Losing and Gaining Streamflow Conditions on Hyporheic Exchange Fluxes Induced by Dune‐shaped Bed Forms’. *Water Resources Research* 50 (3): 1895–907. https://doi.org/10.1002/2013WR014668.

Joensuu, Mari, Conrad A. Pilditch, and Alf Norkko. 2020. ‘Temporal Variation in Resuspension Potential and Associated Nutrient Dynamics in Shallow Coastal Environments’. *Estuaries and Coasts* 43 (6): 1361–76. https://doi.org/10.1007/s12237-020-00726-z.

McCarron, Connor J., Katrien J.J. Van Landeghem, Jaco H. Baas, Laurent O. Amoudry, and Jonathan Malarkey. 2019. ‘The Hiding-Exposure Effect Revisited: A Method to Calculate the Mobility of Bimodal Sediment Mixtures’. *Marine Geology* 410 (April): 22–31. https://doi.org/10.1016/j.margeo.2018.12.001.

Parsons, Daniel R., Robert J. Schindler, Julie A. Hope, et al. 2016. ‘The Role of Biophysical Cohesion on Subaqueous Bed Form Size’. *Geophysical Research Letters* 43 (4): 1566–73. https://doi.org/10.1002/2016GL067667.

Perret, Emeline, Céline Berni, Benoît Camenen, Albert Herrero, and Kamal El Kadi Abderrezzak. 2018. ‘Transport of Moderately Sorted Gravel at Low Bed Shear Stresses: The Role of Fine Sediment Infiltration’. *Earth Surface Processes and Landforms* 43 (7): 1416–30. https://doi.org/10.1002/esp.4322.

Perret, Emeline, Benoit Camenen, Céline Berni, Kamal El kadi Abderrezzak, and Benjamin Renard. 2023. ‘Uncertainties in Models Predicting Critical Bed Shear Stress of Cohesionless Particles’. *Journal of Hydraulic Engineering* 149 (4): 04023002. https://doi.org/10.1061/JHEND8.HYENG-13101.

Staudt, Franziska, Julia C. Mullarney, Conrad A. Pilditch, and Katrin Huhn. 2017. ‘The Role of Grain-Size Ratio in the Mobility of Mixed Granular Beds’. *Geomorphology* 278 (February): 314–28. https://doi.org/10.1016/j.geomorph.2016.11.015.

Thompson, C.E.L., M.E. Williams, L. Amoudry, et al. 2019. ‘Benthic Controls of Resuspension in UK Shelf Seas: Implications for Resuspension Frequency’. *Continental Shelf Research* 185 (September): 3–15. https://doi.org/10.1016/j.csr.2017.12.005.

Vericat, Damia, Ramon J. Batalla, and Celso Garcia. 2006. ‘Breakup and Reestablishment of the Armour Layer in a Large Gravel-Bed River below Dams: The Lower Ebro’. *Geomorphology* 76 (1–2): 122–36. https://doi.org/10.1016/j.geomorph.2005.10.005.

Wiberg, Patricia L., David E. Drake, and David A. Cacchione. 1994. ‘Sediment Resuspension and Bed Armoring during High Bottom Stress Events on the Northern California Inner Continental Shelf: Measurements and Predictions’. *Continental Shelf Research* 14 (10–11): 1191–219. https://doi.org/10.1016/0278-4343(94)90034-5.

Response to review comments (R2). Our response in bold.

General comments

Mixed sediments are a complex problem and not well integrated into numerical models of sediment dynamics and transport. In situ data are limited and often site specific due to the variability of biogeochemical properties of mobile sediment. Improving understanding of the differences between modelled and observed sediment erosion and resuspension should lead to the development of better models of sediment processes. This work uses a novel high spatial resolution model to highlight the amount of uncertainty in coastal models of hydrodynamics and sediment processes, when compared to in situ samples, and variability on small spatial scales.

The figures could be better thought out, to make it easier for the reader to compare and interpret them. Many statements are given without references to support them and the reader cannot follow them up. The work of Joensuu et al. is relied on for the in-situ data, but their contribution and the authors contribution is not always clear. Key information on the source of the bathymetry and its accuracy and quality is missing.

Our response: Thank you for taking the time to review our manuscript. Based on the comments you and the second reviewer have given, we have heavily edited the manuscript. The main addition is that we have now used the data from Joensuu et al. (2018, 2020) to make our own model for the critical shear stress. This makes it evidently clear what is the new contribution of this manuscript, lifts the scientific novelty and better connects the measurements with the numerical results.

Specific comments

Line 26: Exceeding the critical shear stress for motion doesn't always lead to resuspension, it might just be bedload. It implies here that once the critical limit is exceeded suspension will happen. Please clarify.

Our response: We have edited the text to reflect that exceeding the threshold creates a potential for resuspension.

Section 2.1.2 mentions 'variation in sediment surface characteristics (e.g. bedforms, biofilms) across sites', but this data is not shown or used. Grain size class % and biological factors are key factors in sediment resuspension, which were measured by Joensuu et al. (2018; 2020), but not fully integrated into this work.

Our response: You are right that there was a lack of connection between the in situ data, and how the critical threshold were modelled in our study. Based on the review comments we have made significant changes to our study. In summary, we tested (and rejected) the model from Thompson et al. (2019), but instead determined our own similar model using both physical and biological properties directly from the in situ samples. This serves to directly quantify the role of biology based on the samples,

and also allows for the biological component to vary in time when the model is applied. It also opens up for a discussion on why the model from Thompson et al. (2019) doesn't necessarily translate to our data. For more details, please see our comments below.

Section 4 – Discussion: What was the bathymetry source used? And were there any problems with it? Also, the section could be better referenced, many statements are made that aren't supported by the results. Joensuu et al. give % clay fraction but there is no mention of how cohesive sediment could affect these results. Biological factors are highly variable, and effects difficult to account for, but as the clay has been measured could it be used to reduce uncertainty in the model results? The EMODNET data set is not explicitly mentioned in the discussion and, given that this represents the greatest uncertainty, I wanted to know more about where the data came from and how it was put together. Although it is out of scope of this work, examining or re-analysing the raw data used by EMODNET could result in a better understanding of the problem of sediment variability. What survey work would be needed to get a data set that could significantly reduce the uncertainty of a resuspension model?

Our response: The bathymetric data has been compiled by FMI using nautical charts and the Velmu depth model (which is a part of the Baltic Sea Bathymatric Database). In additions, the depth measurements made at the sites (Joensuu et al. 2018; 2002) were incorporated into the wave model grid. Because of defence purposes, the accuracy of the depth information in the area cannot be guaranteed. While there is also an uncertainty in the EMODENET data, our updated methodology only relies on the EMODNET data for classification, while the actual properties for the classes are based directly on the in situ measurements. As mentioned, a more in depth review of the quality of the EMODNET data falls without the scope of our work. Connecting the in situ measurements and the numerical data is a challenging work in itself, so we have to rely on taking the EMODNET data "as is". However, we do find som shortcoming in the classification based on the samples we used, and raise this in the discussion.

Access to the in-situ sediment sample data set was not made available for the reviewers to check, though the Zenodo record exists. The details of where the wave buoy data is archived were not given.

Our response: These data sets will be opened up upon the acceptance of the manuscript.

Technical corrections

Lines 17-20: citations?

Our response: Thank you for pointing this out. We have added appropriate citations to the lines 17-20.

Line 80: Citation for DWR-G4 Directional Waverider? Is it the same for the Waverider Mk-III?

Our response: Thank you for pointing this out. The two buoys do have different manuals. The reference to Mk-III has been updated and the reference to G4 has been added.

Figure 2: What is the data source of the bathymetry contours and coastline? The area of this figure doesn't correspond to the bottom classification and model figures, which is a bit confusing when trying to link the bathymetry to the bed type and model results. Site 13 isn't listed in Table 2.

Our response: The source for the bathymetry is nautical charts and the Velmu depth model (https://ckan.ymparisto.fi/dataset/velmu-syvyysmalli). The coastline data has been rasterized from the polygon data in the dataset (https://ckan.ymparisto.fi/dataset/ranta10-rantaviiva-1-10-000). We have now added tick marks to the bottom classification figure which makes it easier to compare to the bathymetry. We will remove site 13 from the map.

Line 93: citations here for those unfamiliar with EROMES coring.

Our response: Thank you for the suggestion. We have added a citation that describes the EROMES system to provide context for readers unfamiliar with it.

Lines 109-110: What is the data source of the bathymetric data and land-sea mask?

Our response:The source for the bathymetry is nautical charts and the Velmu depth model (https://ckan.ymparisto.fi/dataset/velmu-syvyysmalli). The coastline data has been rasterized from the polygon data in the dataset (https://ckan.ymparisto.fi/dataset/ranta10-rantaviiva-1-10-000)

Figure 4: where are the latitude and longitude tick marks, and land-sea mask? It is difficult to relate this to figure 2. Marking the sites of the cores would be useful to understand how good the EMODNET data is?

Our response: We have redrawn the figure of the classes to have tick marks. We also marked the region used in the other figures. We did not add the stations to all figures, but we added them to figure 7. This shows some limitations in the accuracy of the EMODNET classification, and this is raised in the edited discussion.

Line 141: Citation to back this up. Finer grains are generally the first to be suspended, but coarser grains will be mobilized first as bedload. Please clarify that you are writing about wave suspension. Is this just for waves and non-cohesive sediment? This statement seems a bit over simplified without a citation to back it up.

Our response: Thank you for your valuable feedback. We agree that particle motion is influenced by grain size, however, our study focuses on particles that can be resuspended rather than coarser sediments or boulders, which were excluded from our dataset as they fall outside the scope of this work. The critical shear stress values obtained using the EROMES system are based on the suspended motion of the particles.

While the primary focus of this study is on wave-induced sediment resuspension, we also aim to cover a range of sediment types. Natural sediments typically consist of a mixture of particle sizes, and the erosion process can vary depending on sediment properties (e.g. floc erosion versus suspension). In addition to grain size, other factors such as cohesion and organic content influence erosion behavior.

Despite these complexities, finer particles are generally the first to be resuspended under wave action, which is the phenomenon we sought to capture in natural submerged sediments. We have clarified this point in the manuscript and added appropriate citation.

Line 149: citation for this equation.

Our response: We assume this refers to Eq (1). Citation was added.

Line 165: citation please.

Our response: We assume this refers to Eq (7). Citation was added.

Line 168: Soulsby and Whitehouse improved Shield's curve with extra data, not the parameter.

Our response: This was changed to "The parameterization of the Shields parameter, as improved by Soulsby and Whitehouse (1997),

reads:"

Line 193: Tvärminne research station is not marked on the model data figures. 23° 06' -23° 12'is easier to find. Please mark the areas you are writing about on the figures so that readers can understand better. Bathymetry contour lines would help with the interpretation of shallow and deep areas, though this could clutter up the figure, contours every 10- 20 m depth could make the interpretation easier than switching between figures 1 & 4 and model figures.

Our response: We appreciate that the figures are not that easy to interpret. We have made an effort to harmonize the presentation and marking the main area in the figure presenting the seabed classes. We however decided not to include contour lines, since it adds too much information to the figures. In the new figures it should be clearer which areas are e.g. of type sand, which makes it easier to follow the results.

Figure 6: The figure has a title of Tvärminne but it is not shown in the figure. Place labels, e.g. for Tvärminne and Täktbukten could be used to highlight interesting areas where the exceedance is high (as mentioned in line 200) and discuss these individual sites.

Our response: In the edited manuscript we have moved away from discussing individual sites. However, we have marked the Tvärminne research station on the figure, since it is mentioned in the text.

Lines 252-4: Were these measurements carried out by Joensuu et al.? Cite them if so.

Our response: Thank you for pointing this out. We confirm that these measurements were carried out by Joensuu et al., and we have added the appropriate citations to acknowledge their work.

Line 259: The sample sites aren't marked on figures 7 or 8, so it is hard to see this.

Our response: This sentence has been rewritten in the revised manuscript.

Table 2: the grain size measurements are from Joensuu et al. (2018; 2020) and should be cited here.

Our response: Thank you for pointing this out. We have added the appropriate citations.

Line 299: remove ID from 'stIDress'.

Our response: This typo has been corrected

Lines 340-367: Appendix A - Citations please.

Our response: We have added citations to the Appendix

Line 442: 'https://doi.org/' is duplicated.

Our response: Thank you for catching that duplication. We've reviewed and cleaned the references to remove the extra `https://doi.org/` entries.

Line 445: Full reference so that it may be found. An English translation can be found here: http://authors.library.caltech.edu/25992/1/Sheilds.pdf

Our response: Thank you for the suggestion. We have added a citation to the original Shields (1936) work and included a reference to the English translation for accessibility. We have noted the translation link in the reference for readers who wish to access it.