Response to review comments (R2). 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 authors
I now have two referee reports on your first revision. Both consider the manuscript much improved but also raise substantial comments that should be addressed and I expect to send it to them again after your further revision taking account of this round of comments. I am sorry this is taking time but I think it worthwhile in view of the progress already made and still hoped for. I am attaching the comments, a pdf file and also "Additional private note" with comments from both referees.
Yours sincerely
John Huthnance (editor)

Our response: Thank you. We appreciate the time and effort that is invested in the review process to improve our manuscript. We have carefully addressed the referees comments and prepared a revised manuscript accordingly.

Additional private note (visible to authors and reviewers only).

RefA
General comments
Major changes and improvements to the model used for estimating the critical threshold have been carried out since the initial submission, and the manuscript re-written. Monthly variation in threshold due to biological cohesion has been included in a new model for critical stress. In addition, the manuscript is much better referenced.

I found it interesting that wave stress is focused by channels between islands leading to peaks in stress, which once passed allow re-deposition of fine sediment resulting in mud beds. While near the shore higher mean/background stress in shallower water results in coarse sediment at the bed, as would be generally expected. The re-suspension events and re-deposition seem to be controlled by the coastal bathymetry and morphology. For me, to ‘demonstrate that sediment mobilization is highly episodic, concentrated in exposed shallow areas and constricted channels, with implications for turbidity, nutrient fluxes, and benthic habitats’ is a key result made possible by the high spatial resolution of their model. The integration of the high-resolution model with the EMODnet data and in-situ measurements, factoring in the sediment classes, is also a highlight of the work.

Our response: Thank you for the positive and encouraging comments on our revised manuscript, and for acknowledging the progress we have made.

Specific comments
Section 4.3
The non-stationary physical–biological model could be better described, with more statistical details about the multiple linear regression please. The number of measurements used to create the model, RMS error/difference and the overall p-value.

These would help the reader assess the uncertainty and significance of the model. The uncertainty isn't mentioned and if there is a wide range of possible critical stress values within the physical–biological model, then the spatial model maps could look different at the extremes of the range. In particular, Chl a data can have a lot of scatter, due to patchiness over the sediment bed, and this could propagate to a large uncertainty in the critical stress values.

Correlation between variables (collinearity) can be tested for (see Zuur et al. 2010, step 5, https://besjournals.onlinelibrary.wiley.com/doi/10.1111/j.2041-210X.2009.00001.x), but understanding the variables and keeping the number to a minimum few key variables reduces its likelihood. 'The choice of which covariates to drop can be based on the VIFs, or perhaps better, on … knowledge.' (Zuur et al, 2010).

Our response: We thank you for these constructive suggestions. We have expanded the description of the non-stationary physical-biological model in Section 4.3 to include the requested statistical details, including the number of measurements used to construct the model, the overall model significance, and the root-mean-square error.

Following the recommendation, we evaluated collinearity among predictor variables using the approach described in Zuur et al. 2010. We calculated pairwise Pearson correlations and variance inflation factors (VIF). As expected, the two physical variables were strongly correlated (Pearson R = 0.86), whereas Chl a showed only weak correlations with the physical variables (R = 0.1 and 0.34). VIF indicated moderate collinearity for the physical variables (VIF = 4.96 and 4.44) and low collinearity for Chl a (VIF = 1.33). Removing either median particle size or dry bulk density lowered the explained variance and reduced predictive power, confirming that each physical variable contributes unique explanatory information. In addition, both physical variables have strong justification for inclusion, so we retained them while acknowledging the associated statistical uncertainty.

Section 4.3 has now been revised and provides a clearer and more transparent description of the model, the uncertainty, and rationale for variable selection.

Technical corrections
Figure 5 caption: remove 'mathrm' from 'mathrmChlA'

Our response: This has been removed.

L223-4: 'While the areas with sand bottom are highlighted also in the maximum shear stress values, the highest values (exceeding 2 N m−2) can be found at mud bottoms' where is this shown? Is this comparing figures 4a and 6c? Please clarify.

Our response: Thank you for pointing this out. We have now clarified the description in the revised manuscript (Section 5.1). The sentence refers specifically to the comparison between the sediment class map (Fig. 4a) and the maximum near-bottom shear stress values (Fig. 6c). The revised text now states that the highest instantaneous shear stress peaks (> 2 N m^{-2}) shown in Fig. 6c

occurs in areas identified as mud or muddy-sand in Fig. 4a. This comparison is now stated directly in the manuscript to make the interpretation clear.

Conclusions: ‘This study underlines the importance of incorporating both physical and biological factors – and their temporal dynamics– into sediment transport models to achieve reliable predictions of erosion thresholds and resuspension potential.’ could be added to the end of the abstract as a key conclusion.

Our response:
Thank you for the suggestion. We have revised the abstract accordingly.

RefB
The manuscript “Wave-induced sediment resuspension in the Finnish Archipelago, Baltic Sea: Integrating field measurements with large-scale numerical model simulations” by Björkqvist et al is much improved from the initial submission, but there are still some refinements to the paper and clarity needed around modelling validation for it to be accepted.

The spurious bed shear stresses and suspension of boulders have been removed from the results, primarily as the authors do not have field measurements of their suspension. This focusing of the model results onto what relevant field data they have has greatly improved the scope of the manuscript. The authors have ingested the paper of (Thompson et al., 2019) well, and have created their own model for the initiation of motion for sediments with differing biological properties and compaction which now explains a larger proportion of the sediments’ mobility. In doing this they have also been able to define the difference between sediment mobility with and without the biological properties, which is a good new result for this work.

There are still questions as to other processes which are controlling the threshold of motion for their sediments – questions which the wider literature also has. As such the paper is not much more up to date with the state of the science.

Our response: We thank the reviewer for regarding our revised manuscript improved from the first version. We hope that our response below will satisfactorily answer the reviewer’s remaining concerns about our manuscript.

I currently still have reservations about two aspects of the manuscript
1) The lack of any distinction between bedload and suspended load in the calculations. The authors are only using the threshold of motion, and not suspension, yet always talk about resuspension. I advise them to either adjust their method or change the focus to initiation of motion. They may well compare both methods – which would also be fine.

Our response: Since our focus is on how the biological properties affect the critical threshold and how it manifests itself on a wider scale under intermittent wave action, our study handles the initiation of the sediment movement. Including a detailed analysis of our data to include critical suspension and the different criteria for it (e.g. Sun et al. 2025 and the references within) as well as the transport connected to the intermittent forcing is an important subject, but it

would deserve a study of its own. We have gone through the manuscript and changed the word "resuspension" to "critical shear stress" or "resuspension potential" when appropriate to avoid any misunderstanding of the focus of our study.

Lixia Sun, Zhongwu Jin, Zhilin Sun, Guiying Shen, Haolei Zheng, Chao Guo, Lingyun Li,
Critical criteria for sediment suspension derived from suspension probability,
Engineering Science and Technology, an International Journal,
Volume 64,2025,102034, ISSN 2215-0986,
https://doi.org/10.1016/j.jestch.2025.102034

2) The poor validation of wave period in their model compared to observation (Figure 3). Whilst bottom orbital velocity is well matched with observation, wave period is not well compared to the point where wave period is often> 50% lower in the model compared to the field measurement. Considering the orbital velocities are good, this suggests wave heights are not well modelled - presumably during period of low wave height conditions. Figure 3 needs to be expanded to include wave height observations and model prediction. An estimate of wave energy as well would let readers have a clear reason why the authors are confident in ignoring this discrepancy in their model – as presumably the wave energy during the periods of low correlation is also low.

Our response: As explained in Section 2.2, lines 93-95 (lines 95-97 in the revised manuscript), the near-bottom periods calculated from the measured wave spectra were spurious during lower sea states (Fig. 3b). During these cases the wave action had strongly attenuated before reaching the bottom (Fig. 3a), and the low frequency noise in the measured wave spectra was dominating resulting in spurious, high near-bottom periods. For keeping the focus of the manuscript on near-bottom wave action, we added a new Appendix C to show the comparison of the main modelled and measured wave parameters. The comparison shows that the modelled significant wave height was slightly overestimated but the bias and RMSE of the significant wave height were comparable to those obtained in an earlier comparison (Björkqvist et al. 2020). In low sea states with variable winds, the wave spectrum can have several peaks of corresponding magnitudes which results in variable peak period which can be seen in the comparison: Figure C1 in Appendix C show that when the significant wave height was over 0.30 m the agreement between the modelled and measured peak period was good. Please also note that Figure 3 shows the more robust mean near-bottom period integrated over the spectrum (Appendix A) while the peak period is a single period of the spectral peak.

Figure C2 in Appendix C also shows that as suspected by the reviewer, the cases with high measured near-bottom periods were connected to low sea states and consequently the wave action did not reach/was minimal at the bottom. The energy of these low sea states were also satisfactorily simulated by the model (Appendix C). However, we decided to leave the spurious values to Figure 3 instead of selecting a threshold for reliable near-bottom velocities to filter the data: the measurements were used solely for validation purposes. We also added

a sentence to the caption of Figure 3 to remind the reader of these less defined values of near-bottom periods in the measured data.

There is also a concern that the period of model validation does not appear to be used in the rest of the paper. The 2017 validation period is unused in section 5. The 2017 model coincides with available wave measurement data – whilst the sediment collection data (and model runs for these) cover periods in 2014 and 2015. Figures 6 -7 (modelling period undefined) and Figures 8-9 (a storm event on 28 November 2015) are presumably all from the same 2015 model run. All things being equal, this shouldn't be an issue. A validation period of a different year should also work, my concern however is that the range of bottom orbital velocities shown in the November 2015 storms is far higher than any during the validation period; 0.8 m/s (Figure 8) in 2015 vs a maximum of 0.1 m/s in 2017 (Figure 3). Given the absence of observations for this period, additional justification is needed to demonstrate that the model is operating in a physically realistic way under these higher-energy conditions. A comparison of offshore wave climate between years and a discussion of whether the validated conditions adequately represent the extremes simulated would be useful.

Given these concerns I still think the manuscript requires a major revision.

Thompson, C.E.L. et al. (2019) 'Benthic controls of resuspension in UK shelf seas: Implications for resuspension frequency', Continental Shelf Research, 185, pp. 3–15.

Our response: All the available wave model data was used in Figures 6 and 7 (see also the first paragraph of section 5.1, lines 217-218 (lines 237-238 in the revised manuscript)). The modelled time periods are given in the first paragraph of section 2.2, lines 63-65 (lines 65-67 in the revised manuscript): they cover the time periods of the field measurement campaigns in 2014 and 2015 as well as the wave model validation period in 2017. We have reformulated the captions of the Figure 6 (and rewrote the caption of Fig. 7) to be clearer on the used data, now it just mentions "for all the available model data".

Like the reviewer pointed out, the validation period in 2017 did not include the highest sea states encountered in the study period of 2015. The wave model SWAN is, however, extensively validated in the Baltic Sea and its complex archipelago (e.g. Björkqvist et al. 2018, Björkqvist et al. 2020, Giudici et al. 2023). There are always some inherent uncertainties connected to modelled wind fields, accuracy of the water depth information, implementation, and numerical solutions of the wave model, but based on the above mentioned previous model validations, we are confident in the wave model's ability to model the wave conditions accurately enough for our purposes, also in high wind events. The validation of the high resolution SWAN in 2017 with wave measurements from the near-shore locations in the archipelago further supports the usability of the wave model results used in this study. We have added this discussion of the previous validation results to section 2.2 where SWAN is described:

«The wave model SWAN has been extensively used and verified in the Baltic Sea. In the open sea areas, the accuracy of the model was found good (e.g. Björkqvist

et al. 2018, Giudici et al. 2023). Closer to the shore, and especially in the archipelago in the northern parts of the Baltic Sea, the accuracy of the numerical wave models depends on the grid size, reflecting the capability of the model to resolve smaller islands and variable bottom topography that influence the wave energy dissipation (e.g. Tuomi et al., 2014). A detailed analysis of the performance of three different numerical wave models, including SWAN, in the archipelago off Helsinki in the Gulf of Finland showed that the models on 0.1 nmi grid resolution are capable of simulating the wave field with good accuracy (Björkqvist et al., 2020). In the present study, our grid resolution is 20 m, and the islands and bottom topography are well resolved as the comparison with measurements show (Appendix C and Fig. 3). The validation period in 2017 did not include the highest sea states during the study periods in 2014 and 2015, but according to the studies mentioned above, also the high wind events can be expected to be reasonably well simulated.

Björkqvist, J.-V., Lukas, I., Alari, V., van Vledder, G. Ph., Hulst, S., Pettersson, H., Behrens, A., Männik, A. (2018): Comparing a 41-year model hindcast with decades of wave measurements from the Baltic Sea, Ocean Engineering, 152, pp. 57-71. https://doi.org/10.1016/j.oceaneng.2018.01.048

Björkqvist, J.-V., Vähä-Piikkiö, O., Alari, V., Kuznetsova, A., Tuomi, L. (2020): WAM, SWAN and WAVEWATCH III in the Finnish archipelago - the effect of spectral performance on bulk wave parameters. J. Operational Oceanography, 13, pp. 55 70. https://doi.org/10.1080/1755876X.2019.1633236

Giudici, A., Jankowski, M. Z., Männikus, R., Najafzadeh, F., Suursaar, Ü., Soomere, T. (2023): A comparison of Baltic Sea wave properties simulated using two modelled wind data sets, Estuarine, Coastal and Shelf Science, 290. https://doi.org/10.1016/j.ecss.2023.108401

Tuomi, L., Pettersson, H., Fortelius, C., Tikka, K., Björkqvist, J.-V., Kahma, K.K. (2014): Wave modelling in archipelagos. Coastal Engineering, 83, pp, 205-220. http://dx.doi.org/10.1016/j.coastaleng.2013.10.011.»

Unfortunately our high resolution model data do not allow for the comparison between 2014, 2015 and 2017, nor comparison to long-term statistics, due to the scarcity of overlapping seasonal coverage. Existing long-term wave model hindcasts (1980-2025 EU Copernicus Marine Service Product, 2025) show that in March-April 2017, near the location of the outer wave buoy, the sea state was rougher than in the mean (the Figure below). Since the wave model is a different one, WAM (Komen et al. 1994), and the grid resolution is 1 nmi, we do not think these model hindcasts would significantly support the validation of the model used in this study - but they do tell us that the sea state in the mean during our validation period in 2017 was higher than is typical for the season. We prefer to rely on the validation of the SWAN model done in the earlier papers cited in the discussion now added to Section 2.2.

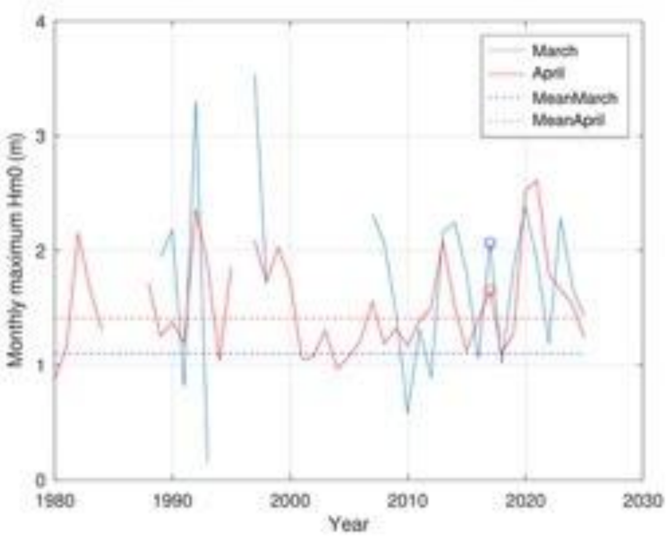


Figure. CMEMS hindcasts of monthly maxima and means of significant wave heights at the location of the outer buoy in March and April in 1980-2025. The year 2017 is marked with circles. The missing values indicate the presence of ice.

EU Copernicus Marine Service Product: Baltic Sea Wave Hindcast, FMI, Mercator Ocean international, [data set], 2025 https://data.marine.copernicus.eu/product/BALTICSEA_MULTIYEAR_WAV_003_015/

Komen G.J., Cavaleri L., Donelan M., Hasselmann K., Hasselman S., Janssen P.A.E.M. 1994. Dynamics and modelling of ocean waves. Cambridge University Press, Cambridge.

There are still questions as to other processes which are controlling the threshold of motion for their sediments – questions which the wider literature also has. As such the paper is not much more up to date with the state of the science.

Our response: We appreciate the reviewer's comment regarding the broader processes that may influence sediment motion thresholds. However, we would like to clarify that identifying or quantifying all possible mechanisms controlling sediment stability is outside the scope of the present study. Our work does not attempt to provide a comprehensive process-level explanation of sediment erodibility. Instead, as stated from the title and abstract onward, our objective is to use field measurements to construct a simple and practical model that can be integrated with sparse spatial datasets and numerical wave simulations in an archipelago environment.

The reviewer's statement that the paper is "not much more up to date with the state of the science" appears to evaluate only one component of our work (the empirical description of sediment thresholds) in isolation. We respectfully emphasize that this component is not intended to stand alone, nor do we claim that it represents a state-of-the-art advance on its own. The novelty and contribution of our study come

from the integration of several elements; linking biological and physical controls on sediment thresholds, combining small-scale field measurements with large-scale wave modelling, and applying this integrated framework across four different sediment types under intermittent wave forcing.

Thus, the manuscript contributes by connecting multiple dimensions rather than by resolving all open questions on sediment stability, many of which indeed remain open in the wider literature. To prevent misunderstandings, we have revised terminology throughout to focus clearly on initiation of motion and critical shear stress, clarified the datasets used in each figure, and added further validation and discussion of the SWAN model, including references to earlier Baltic Sea validation studies. These revisions should help situate our work more clearly within the existing scientific field.

Responses to the additional private note:

Dear Authors
I think your manuscript has been greatly improved, and it is clear you have ingested the paper of Thompson et al., (2019) well, and this focusing of the results onto sediment sizes for which you have measurements of their mobility has clarified the manuscript. I also appreciate breaking down your predictive ability for sediment motion into individual classes as well as the entire suite of grain sizes you have data for – the distinction of variability within and between sediment classes is well taken. I still think you need to switch to using a threshold for suspension, rather than initiation of motion, because at present (and particularly for your coarser sediments) the thresholds are quite different for initiation of motion and suspension. Hopefully you can include both the initiation of motion and suspension in your analysis, and consequently a (hopefully!) improved predictive ability when the threshold for suspension is used. I also have some concerns about how well the storm in November 2015 is modelled – primarily as your validation year doesn’t appear to have a storm of such magnitude. You need to demonstrate that the model is operating in a physically realistic way under these higher-energy conditions (perhaps by comparing offshore forcings between the validation year and 2015). Lastly, (and not for this manuscript!) I hope you will be able to expand on this “natural laboratory” with further field measurements and modelling, particularly in incorporating the grain size distributions, and potentially EPS measurements to (hopefully) further improve the correlation between predicted and observed threshold of motion – information of that kind is vital for understanding what is needed to accurately model the threshold of motion in real environments.

Our response:
We thank the reviewer for regarding our revised manuscript much improved from the previous version. Below are our responses to the concerns raised by the reviewer.
We fully agree with the reviewer that the complex processes involved in modelling the erosion threshold need further studies - our manuscript being one step forward in better understanding these processes.

Abstract.
Please differentiate between critical threshold of motion, and suspension. These are two separate parameters. The final sentence: “Based on the numerical model data, the critical shear stresses from the newly proposed model, τcr (d50, ρB , ChlA(t)), were rarely exceeded based on only wave induced motions in most of the model grid, but could, nonetheless, be exceeded to up around 10%

of the times in smaller areas..” needs reworking, it is a confused sentence. - “were rarely exceeded based on only wave-induced motions in most of the model grid” this is one result, and valid on its’ own (with added quantification). - “but could, nonetheless, be exceeded to up around 10% of the times in smaller areas..”, what are these smaller areas? What’s different about them relative top the rest of the model? Is this result expected or unexpected? What interesting thing do we learn from this?

Our response:
Our focus is on the initiation of the motion. We have gone through the manuscript and changed the word "resuspension" to "critical shear stress" or "resuspension potential" when appropriate to avoid any misunderstanding of the focus of our study.
In this sentence we are referring to the wave-motion as such (e.g wave induced orbital velocities at the sea-bed) which should (at least implicitly) imply that we do not cover wind-driven currents or other motions of the fluid in this work. Also we make the distinction that these motions can be high enough not at the whole modelling domain scale, but at some certain locations, due to the complex bathymetry. Our predictive model for the threshold of sediment motion on the other hand is novel and therefore we added the sentence “This study highlights the importance of incorporating both physical and biological factors – and their temporal dynamics – into sediment transport models to achieve reliable predictions of critical shear stresses and resuspension potential” at the end of the abstract.

Introduction.
The point about models typically using the median grain size to define a threshold of motion as a limitation to many models is well taken, but perhaps laboured. Plenty of models use multiple grain sizes, or distributions, if the detail is available. You are still not differentiating between threshold of motion and suspension, these are well known to be different.

Our response:
As mentioned above, we have gone through the text and used the correct term for initiation of motion.

Methods
It doesn’t appear that there’s any infragravity waves (20-300 second periods), but these have been found to be important for near shore sediment transport in the Baltic sea (see referances in Weisse et al., 2021). Can you explain why your site doesn’t appear to have any? Long period waves are often crucial for sediment transport in coastal environments.

Our response:
According to the CMEMS long-term wave hindcast covering 45 years, the peak period exceeds 10 s 0.23 % of the time at the outer boundary of our modelling domain, 11 s 0.02 % and 12 s (maximum 12.285 s) has been exceeded only 1 h. Therefore there is not any probable source for the bound infragravity waves to occur at the area, and we are not aware of any studies in the Baltic Sea of free infragravity waves radiated from other shorelines in the Baltic Sea towards the Finnish coast. However, as the absence of evidence is not the evidence of absence, we cannot state 100 % that there is no infragravity wave energy in our study of interest.

Figure 3 is concerning. The wave period from the SWAN model and wave buoy does not compare well, and suggests that wave height is poorly represented. Wave heights need to be shown in figure 3 as well. I would also include wave energy to help the reader see why any periods of poor correlation can be discounted (i.e. Low energy, not important).

Our response:
As explained in Section 2.2, lines 93-95 (lines 95-97 in the revised manuscript), the near-bottom periods calculated from the *measured* wave spectra were spurious during lower sea states (Fig. 3b). During these cases the wave action had strongly attenuated before reaching the bottom (Fig. 3a), and the low-frequency noise in the measured wave spectra was dominating resulting in spurious, high near-bottom periods. For keeping the focus of the manuscript on near-bottom wave action, we added a new Appendix C to show the comparison of the main modelled and measured wave parameters. The comparison shows that the modelled significant wave height was slightly overestimated but the bias and RMSE of the significant wave height were comparable to those obtained in an earlier comparison (Björkqvist et al. 2020). In low sea states with variable winds, the wave spectrum can have several peaks of corresponding magnitudes which results in variable peak period which can be seen in the comparison: Figure C1 in Appendix C show that when the significant wave height was over 0.30 m the agreement between the modelled and measured peak period was good. Please also note that Figure 3 shows the more robust mean near-bottom period integrated over the spectrum (Appendix A) while the peak period is a single period of the spectral peak.

Figure C2 in Appendix C also shows that as suspected by the reviewer, the cases with high measured near-bottom periods were connected to low sea states and consequently the wave action did not reach/was minimal at the bottom. The energy of these low sea states were also satisfactorily simulated by the model (Appendix C). However, we decided to leave the spurious values to Figure 3 instead of selecting a threshold for reliable near-bottom velocities to filter the data: the measurements were used solely for validation purposes. We also added a sentence to the caption of Figure 3 to remind the reader of these less defined values of near-bottom periods in the measured data.
Björkqvist, J.-V., Vähä-Piikkiö, O., Alari, V., Kuznetsova, A., Tuomi, L. (2020): WAM, SWAN and WAVEWATCH III in the Finnish archipelago - the effect of spectral performance on bulk wave parameters. J. Operational Oceanography, 13, pp. 55-70. https://doi.org/10.1080/1755876X.2019.1633236

There is also a concern that the period of model validation does not appear to be used in the rest of the paper. The 2017 validation period is unused in section 5, due to the lack of wave data for validation during the 2014 and 2015 measurement campaign. As such Figures 6-7 (modelling period undefined) and Figures 8-9 (a storm event on 28 November 2015) are presumably all from the same 2015 model run.

I have two issues here:
1) Results in section 5 – which model runs, and which time periods are shown need to more clearly labelled for each figure.
2) All things being equal, the different model validation period to those shown in the results section should be fine. My concern however is that the range of bottom orbital velocities shown in the November 2015 storms is far higher than any during the validation period; 0.8 m/s (Figure 8) in

2015 vs a maximum of 0.1 m/s in 2017 (Figure3). Given the absence of observations for this period or set of conditions, additional justification is needed to demonstrate that the model is operating in a physically realistic way under these higher energy conditions. A comparison of offshore wave climate between years and a discussion of whether the validated conditions adequately represent the extremes simulated would be useful.

Our response:
All the available wave model data was used in Figures 6 and 7 (see also the first paragraph of section 5.1, lines 217-218 (lines 237-238 in the revised manuscript)). The modelled time periods are given in the first paragraph of section 2.2, lines 63-65 (lines 65-67 in the revised manuscript): they cover the time periods of the field measurement campaigns in 2014 and 2015 as well as the wave model validation period in 2017. We have reformulated the captions of the Figure 6 (and rewrote the caption of Fig. 7) to be clearer on the used data, now it just mentions "for all the available model data".
Like the reviewer pointed out, the validation period in 2017 did not include the highest sea states encountered in the study period of 2015. The wave model SWAN is, however, extensively validated in the Baltic Sea and its complex archipelago (e.g. Björkqvist et al. 2018, Björkqvist et al. 2020, Giudici et al. 2023). There are always some inherent uncertainties connected to modelled wind fields, accuracy of the water depth information, implementation, and numerical solutions of the wave model, but based on the above mentioned previous model validations, we are confident in the wave model's ability to model the wave conditions accurately enough for our purposes, also in high wind events. The validation of the high resolution SWAN in 2017 with wave measurements from the near-shore locations in the archipelago further supports the usability of the wave model results used in this study. We have added this discussion of the previous validation results to section 2.2 where SWAN is described:

The wave model SWAN has been extensively used and verified in the Baltic Sea. In the open sea areas, the accuracy of the model was found good (e.g. Björkqvist et al. 2018, Giudici et al. 2023). Closer to the shore, and especially in the archipelago in the northern parts of the Baltic Sea, the accuracy of the numerical wave models depends on the grid size, reflecting the capability of the model to resolve smaller islands and variable bottom topography that influence the wave energy dissipation (e.g. Tuomi et al., 2014). A detailed analysis of the performance of three different numerical wave models, including SWAN, in the archipelago off Helsinki in the Gulf of Finland showed that the models on 0.1 nmi grid resolution are capable of simulating the wave field with good accuracy (Björkqvist et al., 2020). In the present study, our grid resolution is 20 m, and the islands and bottom topography are well resolved as the comparison with measurements show (Appendix C and Fig. 3). The validation period in 2017 did not include the highest sea states during the study periods in 2014 and 2015, but according to the studies mentioned above, also the high wind events can be expected to be reasonably well simulated.

Björkqvist, J.-V., Lukas, I., Alari, V., van Vledder, G. Ph., Hulst, S., Pettersson, H., Behrens, A., Männik, A. (2018): Comparing a 41-year model hindcast with decades of wave measurements from the Baltic Sea, Ocean Engineering, 152, pp. 57-71. https://doi.org/10.1016/j.oceaneng.2018.01.048

Björkqvist, J.-V., Vähä-Piikkiö, O., Alari, V., Kuznetsova, A., Tuomi, L. (2020): WAM, SWAN and WAVEWATCH III in the Finnish archipelago - the effect of spectral performance on bulk wave parameters. J. Operational Oceanography, 13, pp. 55-70. https://doi.org/10.1080/1755876X.2019.1633236
Giudici, A., Jankowski, M. Z., Männikus, R., Najafzadeh, F., Suursaar, Ü., Soomere, T. (2023): A comparison of Baltic Sea wave properties simulated using two modelled wind data sets, Estuarine, Coastal and Shelf Science, 290. https://doi.org/10.1016/j.ecss.2023.108401

Tuomi, L., Pettersson, H., Fortelius, C., Tikka, K., Björkqvist, J.-V., Kahma, K.K. (2014): Wave modelling in archipelagos. Coastal Engineering, 83, pp, 205-220. http://dx.doi.org/10.1016/j.coastaleng.2013.10.011

Unfortunately our high resolution model data do not allow for the comparison between 2014, 2015 and 2017, nor comparison to long-term statistics, due to the scarcity of overlapping seasonal coverage. Existing long-term wave model hindcasts (1980-2025 EU Copernicus Marine Service Product, 2025) show that in March-April 2017, near the location of the outer wave buoy, the sea state was rougher than in the mean (the Figure below). Since the wave model is a different one, WAM (Komen et al. 1994), and the grid resolution is 1 nmi, we do not think these model hindcasts would significantly support the validation of the model used in this study - but they do tell us that the sea state in the mean during our validation period in 2017 was higher than is typical for the season. We prefer to rely on the validation of the SWAN model done in the earlier papers cited in the discussion now added to Section 2.2.

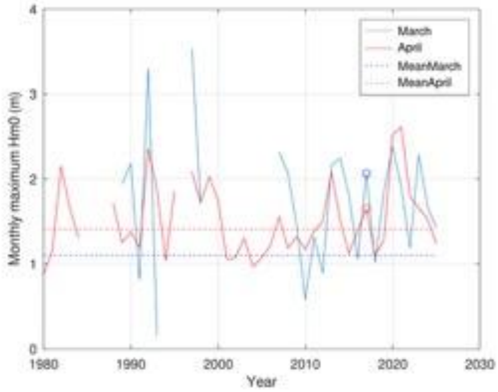
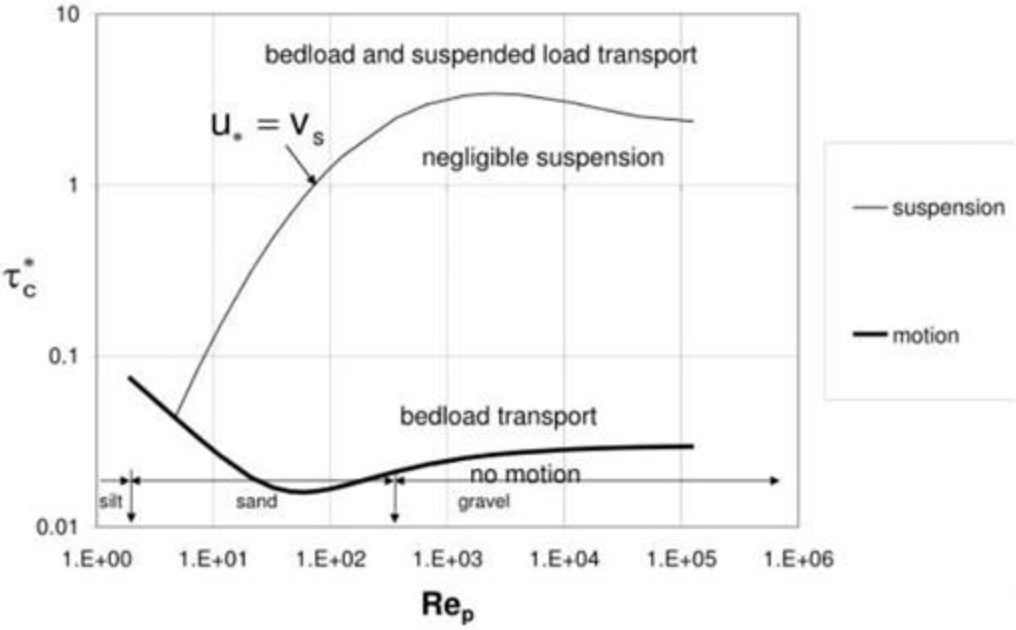


Figure. CMEMS hindcasts of monthly maxima and means of significant wave heights at the location of the outer buoy in March and April in 1980-2025. The year 2017 is marked with circles. The missing values indicate the presence of ice.

EU Copernicus Marine Service Product: Baltic Sea Wave Hindcast, FMI, Mercator Ocean international, [data set], 2025
https://data.marine.copernicus.eu/product/BALTICSEA_MULTIYEAR_WAV_003_015/

Komen G.J., Cavaleri L., Donelan M., Hasselmann K., Hasselman S., Janssen P.A.E.M. 1994. Dynamics and modelling of ocean waves. Cambridge University Press, Cambridge.

I think you have to switch to using theta_sus over the current threshold of motion. These two thresholds are not the same especially for gravels and sands – see figure below.



At the moment your use of the Soulsby-Whitehouse equation is giving you the initiation of motion, not suspension. You need to also use the van rijn (or an equivalent equation) for the threshold of suspension: https://www.leovanrijn sediment.com/papers/Threshsandmud2020.pdf

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

You could add this threshold into your current work, and see how much more of the variability in your results is explained by using a threshold for suspension over a threshold for motion.

Our response: We thank the reviewer for highlighting the important distinction between the threshold of sediment motion and the threshold of suspension. We agree that these represent different physical processes, particularly in mixed sediments where the transition from bedload transport to suspension can be substantial.

The EROMES device is designed to determine critical shear stress for the initiation of sediment motion under in situ conditions, integrating the effects of natural sediment composition and biological activity. Accordingly, the thresholds used in this study reflect erosion and incipient motion rather than suspension. Since our primary objective is to investigate how biological and physical sediment properties influence the probability of sediment mobilization across natural seabed types under intermittent wave forcing, the use of initiation thresholds is fully consistent with the scope and aims of this work.

We did consider incorporating formulations for the threshold of suspension, such as those proposed by van Rijn and other authors. However, such formulations are typically based on idealized sediment properties (e.g. grain size, settling velocity) and do not explicitly account for biological effects or the sediment heterogeneity that is inherently captured by EROMES measurements. Combining empirically derived in situ erosion thresholds with theoretically derived suspension criteria would therefore introduce conceptual inconsistencies into the analysis. Moreover, suspension threshold formulations remain less well validated for

biologically active, heterogeneous seabeds, and their applicability to our study conditions is uncertain.

As the reviewer correctly notes, initiation of motion and suspension are distinct parameters. Including a comprehensive analysis of suspension thresholds, their probabilistic characterization (e.g. Sun et al., 2025), and the associated transport processes under intermittent forcing would require additional assumptions and data and would constitute a substantial extension beyond the present study. Such an analysis would merit a dedicated investigation of its own.

Nevertheless, we acknowledge that suspension processes may contribute to the observed variability in critical shear stress, particularly for coarser sediments. We have therefore revised the manuscript to consistently use terminology appropriate to incipient motion and to clarify this limitation and its implications in the Discussion.

Lixia Sun, Zhongwu Jin, Zhilin Sun, Guiying Shen, Haolei Zheng, Chao Guo, Lingyun Li, Critical criteria for sediment suspension derived from suspension probability, Engineering Science and Technology, an International Journal, Volume 64,2025,102034, ISSN 2215-0986, https://doi.org/10.1016/j.jestch.2025.102034

Figure 5. There seems to be far greater variance for EROMES data than for the predicted (Width of scatter far higher). Can you suggest why this is?

Our response:
We agree with the reviewer that the EROMES-derived critical shear stress exhibits a larger spread than the predicted values in Figure 5. This difference is expected and can be attributed to several factors.
First, the EROMES tau_cr values represent empirically derived estimates based on in situ observations, which integrate substantial natural variability. These include spatial heterogeneity in sediment composition, grain-size distribution, consolidation state, biological effects, and bedform roughness, all of which can vary at scales smaller than the observational footprint. Such variability contributes to a broader scatter in the EROMES data.
Second, the predictive approach is based on a functional relationship between tau_cr and a limited set of controlling parameters, effectively smoothing variability through parameterisation and averaging. As a result, the model captures the central tendency of the observations but does not explicitly resolve fine-scale heterogeneity, leading to a narrower distribution of predicted values.
Third, uncertainty in the EROMES estimates themselves (e.g. due to measurement noise and threshold detection) further increases the apparent spread of the observed tau_cr values. In contrast, the predicted values are internally consistent by construction and therefore less dispersed.
All in all, the wider scatter of the EROMES data reflects real environmental variability and observational uncertainty, whereas the reduced spread of the predicted tau_cr primarily reflects model regularisation. Importantly, despite these differences in variance, Figure 5 shows that the predictions reproduce the overall magnitude and sediment class-dependent trends of tau_cr reasonably well.

Lines 280 to 292 – needs references

Our response: References were added.

Lines 293- “the three parameter linear model could explain most of the variation”. Please give us some numbers!

Our response: We have revised the text in the Discussion to quantify model performance.

Line 297 onwards: “and we surmise that a single variable (ChlA) might not be enough the capture the biological effects well enough to sufficiently model the variations in critical stress on muddy seabeds”. You could certainly expand here, what about EPS?

Our response: We thank the suggestion and agree that a brief extension is justified. Extracellular polymeric substances (EPS) are known to play an important role in biostabilization by increasing sediment cohesion. However, EPS measurements were available for the same study area in Joensuu et al. (2020). Although EPS concentrations exhibited clear seasonal variability and differences between sediment types, they did not emerge as a significant explanatory variable for variations in sediment erodibility in that dataset. Consequently, EPS was not included in the present model. We now clarify this point in the revised manuscript and note that while EPS can be a key control in other environments (e.g. intertidal mudflats), its explanatory power appears to be site-specific.

Line 306- “1) Our in situ data could not provide representative values for all of the classes.” -A Value of….what exactly

Our response: The sentence has been clarified.

Line 317- “as horizontal wave motion does not attenuate significantly when the wavelength is at least 20 times the water depth.” -Reference needed
Our response: Reference to Holthuijsen (2007) added.

Thompson, C.E.L. et al. (2019) ‘Benthic controls of resuspension in UK shelf seas: Implications for resuspension frequency’, Continental Shelf Research, 185, pp. 3–15. Available at: https://doi.org/10.1016/j.csr.2017.12.005.

Weisse, R. et al. (2021) ‘Sea level dynamics and coastal erosion in the Baltic Sea region’, Earth System Dynamics, 12(3), pp. 871–898. Available at: https://doi.org/10.5194/esd-12-871-2021.