

Dear Trisha Atwood,

After thorough revision, we hereby resubmit our manuscript entitled " Sediment heterogeneity shapes spatial variability of resuspension-induced CO<sub>2</sub> production" for publication in Biogeosciences.

We thank you and the reviewers for investing your valuable time to review our manuscript.

Based on the reviewer's suggestions we thoroughly revised the manuscript and now provide:

- a more detailed description of the resuspension assay method and the boosted regression tree (BRT) analysis
- a revised result section that presents the BRT results in more detail, presents results from additional analyses using the original data set, and a clearer description what the results mean
- a revised discussion section where we address possible mechanistic explanations of the spatial variability of RCO<sub>2</sub>P, discuss the role of organic matter composition and origin, discuss methodological aspects of the assay, and provide a deeper evaluation of the potential implications of the role of sediment heterogeneity for resuspension impact assessments and fisheries management
- a revised introduction to align with the revised discussion

Below you will find the detailed point-by-point responses (blue). References to lines and sections refer to the revised manuscript without tracked changes.

We think the reviewers' insights have improved the manuscript significantly and we hope our changes are satisfactory. We are happy to discuss and revise any further aspects should they arise.

Sincerely,

Ines Bartl

(On behalf of all authors)

## Point by point response to reviewer 1:

Bartl and Thrush conduct experiments to gauge the amount of additional CO<sub>2</sub> released from sediment disturbance using incubations of natural sediments sampled from Hauraki Gulf, New Zealand. The authors analyse their results using a machine learning method and find non-linear relationships and interaction effects between additional CO<sub>2</sub> release and sediment characteristics. It is concluded that assessments of carbon storage vulnerability must account for sediment heterogeneity.

In general, the paper is well written, and the methodology clearly described. However, I found the current presentation and interpretation of results to be lacking. The fact that sediment heterogeneity needs to be accounted for when assessing carbon impacts is already well-established (and somewhat trivial), and the usefulness of the resuspension assay has already been introduced in the earlier work by Bartl et al. (2025). The results of the BRT model are interesting, but they are presented in a quite condensed manner, and it is not laid out clearly what exactly we can learn from them.

### Response:

We thank the reviewer for their thorough assessment of the manuscript which has drawn us to focus more clearly on the purpose of work presented in this manuscript. With our study we aim to identify sediment types that are at highest risk of producing CO<sub>2</sub> when resuspended and to identify what relationships with sediment characteristics influence the variability of resuspension-induced CO<sub>2</sub> production. This forms the basis for our next research step: upscaling the risk of resuspension-induced CO<sub>2</sub> releases across heterogeneous seafloor spaces to inform decision makers. We agree that for our purpose of this manuscript we need to explain more clearly what can be learnt from the BRT results in the context of identifying vulnerable sediment types. Modifications of the BRT analysis and results are addressed in the following responses.

I do think the data collected and the experiments done are valuable and useful, but the discussion focuses almost exclusively on the BRT results, which are difficult to interpret, since such ML methods tend to obfuscate possibly straight-forward interactions and relationships. I encourage the authors to dig a bit deeper into their data through additional analyses and/or to present the BRT results in more detail, and to discuss the possible mechanistic explanations for the observed patterns. I list some specific suggestions below, along with other comments.

### Response:

We thank the reviewer for highlighting this. We have deeply revised the result section and now present our BRT results in more detail and additionally investigate relationships using the original data and Pearson correlation analysis.

A note on the revised BRT analysis:

Upon further reading into interpreting BRT model outputs, we have concluded that it is necessary to exclude highly collinear features from the analysis. While multi-collinearity is not problematic for prediction, it is problematic for the interpretation of BRT results, particularly the ranking of individual feature importance and feature interactions as it can mask the importance of relationships and interactions of other individual features (Boulesteix et al., 2012; Dormann et al., 2013; Lucas, 2020). We checked collinearity between our features (sediment characteristics and water depth) and removed Mud and M-Sand from the analysis as they presented a strong correlation with OM ( $r > 0.8$ ). We decided to keep OM content as it is the substrate for organic carbon mineralization. The removal of Mud and M-Sand does not change the key results of this study and in addition now shows much clearer relationships which allows for a better interpretation of results.

General comments:

· The resuspension assay is meant to mimic trawling impacts, but it's not clear what the historical and current trawling intensity in the study area looks like, or what other bottom-disturbing activities (dredging, sand mining, ...) may occur in the study area. Can the authors give some additional information about this in 2.1? Right now, it is only briefly mentioned in the introduction.

Response:

Thank you for pointing this out. We now describe anthropogenic disturbances (trawling, dredging, sand mining) in Hauraki Gulf and provide additional information on trawling intensity from the past 5 years based on global fishing watch data (*Global Fishing Watch*, 2025). See **line 83-86 and Figure S1**.

· The assay is based on SOD measurements, which are then converted to CO<sub>2</sub> based on a constant RQ. The authors have justified this choice in Bartl et al. (2025), but I am not convinced that this should hold for this analysis as well. Is the value of RQ=0.9 valid for every sampling location? How can the authors be sure that their measured SOD is due purely to OC mineralisation, as opposed to aerobic oxidation of other species, which has been shown to be a larger oxygen sink in some muddy sediments compared to OC mineralisation (e.g. Kalapurakkal et al., 2025)? Perhaps the authors can give some information on oxygen penetration, redox-depth etc. in their samples to clarify this.

Response:

We thank the reviewer for pointing this out. We have checked our choice of an RQ=0.9 (Jørgensen et al., 2022) and identified, that for sites shallower than 50 m, the use of RQ=0.9 is correct but for sites deeper than 50 m, we should have used an RQ=0.85 for outer shelves (50-200m). In addition, we have conducted a full revision of our data quality

process and have identified that two sites need to be omitted from data analysis as they do not meet the quality assessment of the resuspension assay method (Bartl et al. 2025). We apologize for this mistake and have repeated the BRT analysis with the corrected data set. It did not change the key results of this manuscript but the order of most important features is slightly different (Table below).

Metric	BRT results original manuscript	Revised BRT results with corrected RCO2P (Mud and M-Sand excluded due to collinearity with OM)
R <sup>2</sup>	0.57 ± 0.1	0.57 ± 0.1
RMSE	0.56 ± 0.1 mmol CO <sub>2</sub> m <sup>-2</sup> h <sup>-1</sup>	0.54 ± 0.07 mmol CO <sub>2</sub> m <sup>-2</sup> h <sup>-1</sup>
feature importance by mean  SHAP	<ol style="list-style-type: none"> <li>1. OM, SHAP = 0.42</li> <li>2. Depth, SHAP = 0.13</li> <li>3. F-Sand, SHAP = 0.12</li> <li>4. M-Sand, SHAP = 0.12</li> <li>5. C-Sand, SHAP = 0.12</li> <li>6. OM:Phyto, SHAP = 0.08</li> </ol>	<ol style="list-style-type: none"> <li>1. OM, SHAP = 0.47</li> <li>2. C-Sand, SHAP = 0.13</li> <li>3. Depth, SHAP = 0.10</li> <li>4. F-sand, SHAP = 0.9</li> <li>5. OM:Phyto, SHAP = 0.07</li> <li>6. Chl.a, SHAP = 0.04</li> </ol>

It is true, a fraction of SOD may be due to aerobic oxidation of reduced species such as pyrite which can occur at similar time scales to OM mineralisation. Kalapurakkal et al. (2025) incubated slurries of 10 cm sediment depth from a eutrophicated bay where the oxygen penetration depth is <1.6 mm. In contrast, our incubation used the top 3 cm of sediments in an oligotrophic system with an oxygen penetration depth of 3 – 6 mm and nitrate penetration of 12 mm (Cheung et al. 2024). This aligns with findings of low acid volatile sulfide concentrations (AVS) in non-impacted muddy Hauraki Gulf sediments (0 – 1 µmol g<sup>-1</sup> ww) compared to higher AVS levels in muddy sediments impacted by a mussel farm (2 – 12 µmol g<sup>-1</sup> ww, Wilson and Vopel, 2015). While Kalapurakkal et al. (2025) mention the analysis of pyrite content in their methods, data was not provided so we couldn't compare it. We now discuss this methodological aspect at **lines 336-345**.

· Only absolute SHAP values are shown in Fig 2A+B even though, according to 169-172, positive and negative impacts of each feature can be distinguished by the sign. For example, I would expect OM:Phyto to have a negative sign. This important information is hidden by showing only the absolute values. Also, can metrics like SD or ranges, (based on the 50 iterations of the model and/or based on the mean of iterations but for all data points) be shown for the SHAP values to get a sense of their distribution/uncertainty?

Response:

We thank the reviewer for highlighting that information is hidden. We now present mean absolute SHAP to describe the global ranking of features (**Table 1 in revised manuscript**) and show the relationship between positive/negative SHAP and feature

values in partial dependence plots, for all features including OM:Phyto (**Figure 3 in revised manuscript**). We hope the revised description of the results is clearer.

We have added the 95% confidence interval to BRT metrics to provide a measure of model uncertainty (**Table 1 in revised manuscript**). In the partial dependence plot, we now show LOWESS smoothed functions of the mean across the 50 model iterations and its 95% confidence interval (**Figure 3 in revised manuscript**). While the variability of our data, i.e. the scatter of the points in the partial dependence plots is large the likelihood of the mean across the 50 model iterations, i.e. 95% confidence interval, falls in a small range.

· The authors focus their discussion on the results of the BRT model, but the interpretations of the resulting patterns and their mechanistic explanation is lacking. For example, two mechanisms that immediately come to mind are (1) decreased oxygen penetration in muddier sediments that could decrease OM degradation upon resuspension and could explain the disproportionate RCO<sub>2</sub>P compared to sandier sediments and (2) a decrease of terrestrial OM with distance from coast could explain the pattern in Fig 2D, assuming that terrestrial OM is generally less labile in marine environments compared to marine OM. Terrestrial vs. marine OM is mentioned briefly in 2.1, but it is never discussed afterward, though I assume it should have an impact on the degradability. Can the jumps/non-linearities (Fig. 2) be separated geographically, e.g. between firth, channels and offshore? This could be a straight-forward explanation for the patterns shown in the BRT results. Another question that is not addressed is why M-Sand should have such an important role (though it's not clear why the authors focus on this fraction in the first place, since it seems to be no more important than the other sand fractions; see specific comment).

Response:

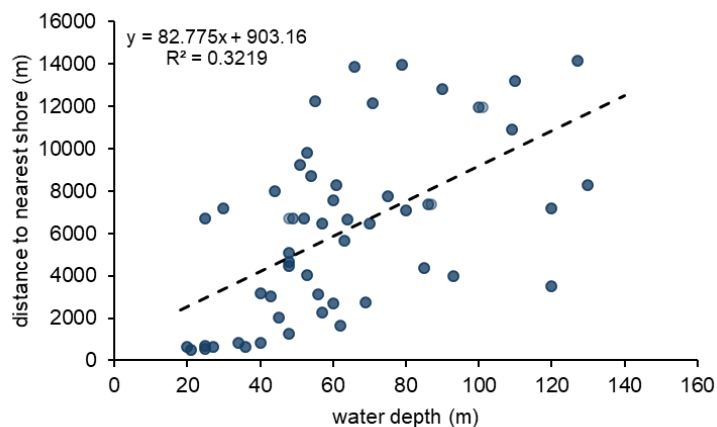
We thank the reviewer for their thoughts on potential mechanistic explanations. As mentioned by the reviewer earlier the critical issue is that the results, i.e. role of most important features, what is driving non-linear relationships, need to be discussed in more detail. We will address this by adding a paragraph in the discussion with the focus on what underlying mechanisms might be implied by the BRT results (**section 4.1 in revised manuscript**). As our study was aimed to observe and identify patterns between RCO<sub>2</sub>P and variable sediment characteristics, discussion around underlying mechanisms can be speculative and we will use our observations to suggest potential scopes for future (experimental) studies (**lines 297-300**).

Regarding the two suggested mechanisms:

(1) Decreased oxygen penetration in muddier sediments decreases OM degradation in undisturbed sediments, as the lack of oxygen limits the more effective oxic mineralization and allows for accumulation OM. However, when resuspended and in

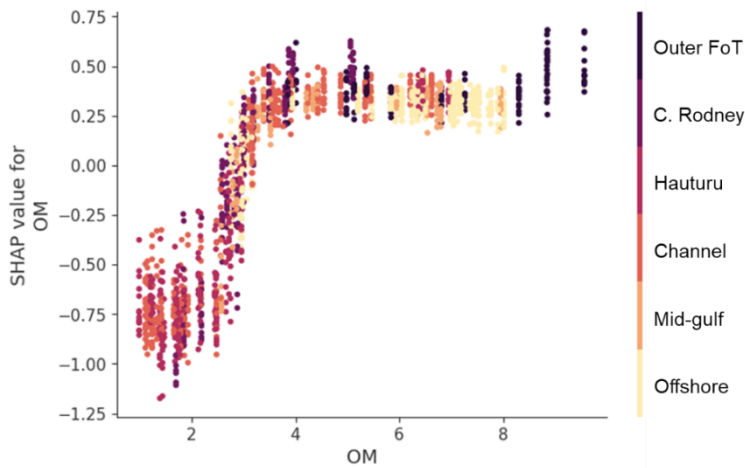
contact with oxygen, such sediments have been observed to experience a stimulation of OM degradation as both, labile and older/refractory OM are mineralised (Hulthe et al. 1998). The conditions in the resuspension treatment were oxic, irrespective of the sediment type and potential oxygen penetration depth that the sediments had.

(2) Yes, decrease of terrestrial OM/ increase of fresh marine-derived OM with distance from coast could explain the higher RCO<sub>2</sub>P with increasing water depth (Fig. 2D). This would assume though, that distance to coast correlates with water depth. We have derived distance to coast values for our sampling sites and while there is a positive trend, the relationship is governed by high variability due to the presence of multiple islands and steep slopes within the Hauraki Gulf (Figure below).



Non-linearities due to regional differences:

We overlaid different regions (offshore, Channels, Firth, mid-Hauraki, Cape Rodney, Hauturu) of the Hauraki Gulf in the partial dependence plot of OM (most important feature) to see if the non-linearity is linked to different regions (figure below). Overall, the non-linearity cannot be explained by region, but some regions, or the conditions there, might play a contributing role for some sites, e.g. negative SHAP values at low OM linked to sites from the channel and Hauturu Island. We think that the non-linearities and shifts are linked to multiple drivers, such as bottom water currents and redistribution of OM within the Gulf that differ locally within the Hauraki Gulf and we the revised discussion makes this clearer (**see line 275-280**).



Medium sand:

This has changed with the revised BRT method where M-Sand is removed due to collinearity with OM to make BRT results more interpretable.

Specific comments

· 1: The authors may consider adding a larger-scale, regional map to give readers unfamiliar with the area a better sense of where the area is located.

Response: : Larger-scale map added.

· 14-15: „...we quantified RCO2P *and it* with measurements of sediment grain size, organic...” missing word?

Response: Sentence corrected.

· 28: I suggest adding references to some other studies to give a sense of the large uncertainty in this number (e.g. Epstein et al., 2022; Hiddink et al., 2023; Zhang et al., 2024).

Response: We revised the introduction and this sentence plus references is now at **line 56-58**.

· 2: Though details may be provided in Bartl et al. (2025), a short description on how the resuspension assay was performed would be helpful (e.g. using only the top 3 cm, gentle shaking, ...)

Response: We added more detail to the description of the RA method (**section 2.2 in the revised manuscript**).

· 117-119: The naming of the combined sand size classes is a bit unfortunate as it could create confusion, so I suggest introducing the acronyms “F-Sand” and “C-Sand” here

and using those whenever referring to the combined classes, rather than “fine sand” and “coarse sand”.

Response: We renamed the sand size classes as: F-Sand for sand grain sizes of 63-250  $\mu\text{m}$ , M-Sand for sand grain sizes of 250-500  $\mu\text{m}$ , and C-Sand for sand grain sizes of 500-2000  $\mu\text{m}$ . Instead of “fine sand” or “coarse sand” we use terms “fine grained sand” or “coarse grained sand” in the discussion.

· 137+231: remove comma after „Both“

Response: Removed

· Tab 1: I don't see Phaeo:Chla in the table, though it's defined in the Table description.

Response: Apologies for this confusion. We have removed this from the table description.

· 175-177: Can the authors be more specific here? Which features could be omitted for the prediction? And doesn't this clustering imply that the interactions are not so important, and that OM itself is already a quite good predictor?

Response: The result section has been extensively revised. Omitted features are now described/discussed in the **method section 2.4**.

· 177-178: M-Sand does not seem to be more important than the other sand classes according to Fig. 2a. Why do the authors choose to focus on M-Sand, and how can the high interaction score of M-Sand with OM be explained? The short discussion in 249 (“...other environmental factors play a role for the reactivity of sediment OM and thus RCO<sub>2</sub>P”) is quite unspecific.

Response: This has changed with the revised BRT method where M-Sand is removed due to collinearity with OM to make BRT results more interpretable.

· Fig. 2C-E: It may be worth drawing a zero-line in these plots

Response: Zero-line added to the plots.

· 211: The authors may consider removing the linear regression formula here, since it disrupts the flow for the reader and is already included in the description of Fig. 3.

Response: Removed

· 240-241: Which range is being compared here to arrive at 2-88 mmol/m<sup>2</sup>/d? The range of undisturbed CO<sub>2</sub>P from Table 1 (0.1-1.0 mmol/m<sup>2</sup>/h or 2.4-24 mmol/m<sup>2</sup>/d) is different.

Response: The range of RCO<sub>2</sub>P is compare. Sentence has been clarified (**line 322-323**).

· 248: “median” should be “medium”

Response: Corrected

· 250: Looking at the supplemented maps, it doesn't seem like grain size is always decreasing along the depth gradient. Maybe the authors can rephrase to clarify what is meant here.

Response: Rephrased.

· 254: I don't think it's fair to say that the study has found any "interactive dynamics" at this stage, but rather relationships; the dynamic process understanding needs to be explored further.

Response: Agreed and removed "interactive dynamics".

#### References from reviewer

Bartl, I., Evans, T., Hillman, J., Thrush, S., 2025. Simple assay quantifying sediment resuspension effects on marine carbon storage. *Methods Ecol Evol* 16, 309–316. <https://doi.org/10.1111/2041-210X.14479>.

Epstein, G., Middelburg, J.J., Hawkins, J.P., Norris, C.R., Roberts, C.M., 2022. The impact of mobile demersal fishing on carbon storage in seabed sediments. *Glob Change Biol* 28, 2875–2894. <https://doi.org/10.1111/gcb.16105>.

Hiddink, J.G., van de Velde, S.J., McConnaughey, R.A., Borger, E. de, Tiano, J., Kaiser, M.J., Sweetman, A.K., Sciberras, M., 2023. Quantifying the carbon benefits of ending bottom trawling. *Nature* 617, E1-E2. <https://doi.org/10.1038/s41586-023-06014-7>.

Kalapurakkal, H.T., Dale, A.W., Schmidt, M., Taubner, H., Scholz, F., Spiegel, T., Fuhr, M., Wallmann, K., 2025. Sediment resuspension in muddy sediments enhances pyrite oxidation and carbon dioxide emissions in Kiel Bight. *Communications Earth & Environment* 6, 156. <https://doi.org/10.1038/s43247-025-02132-4>.

Zhang, W., Porz, L., Yilmaz, R., Wallmann, K., Spiegel, T., Neumann, A., Holtappels, M., Kasten, S., Kuhlmann, J., Ziebarth, N., Taylor, B., Ho-Hagemann, H.T.M., Bockelmann, F.-D., Daewel, U., Bernhardt, L., Schrum, C., 2024. Long-term carbon storage in shelf sea sediments reduced by intensive bottom trawling. *Nature Geoscience*. <https://doi.org/10.1038/s41561-024-01581-4>.

## Point-by-point response to Reviewer 2

Bartl and Thrush carried out replicated manipulation experiments of superficial (3 cm) sediments obtained from a total of 57 sites in the Hauraki Gulf (New Zealand) characterized by varying edaphic characteristics to estimate CO<sub>2</sub> release due to resuspension. The study poses foundation on an early manuscript (Barts et al. 2025 Meth Ecol Evol) describing the resuspension assay and how this could provide an estimate of the vulnerability of marine sediments ability to store carbon when exposed to bottom disturbance and applies it in the field at a basin scale.

Using machine learning methods, and after the removal of collinear explanatory factors, they identified organic matter (OM) content (but not its freshness), sand contents and water depth as the factors most influential on the CO<sub>2</sub> release from the manipulated sediments. Based on these results and using scatter heatmaps, the authors identified values of the explanatory factors representing thresholds of vulnerability to CO<sub>2</sub> release by sediment resuspension. According to those thresholds, a large portion of the sediments of the Gulf under scrutiny show from moderate to high attitude to releasing CO<sub>2</sub> when disturbed. The authors conclude that the assessment of the vulnerability of marine sediments ability to store carbon must account for sediment heterogeneity. Ultimately, the authors prompt that sediment heterogeneity must be accurately considered when deploying plans of spatial management of demersal fisheries (one of the most disturbing anthropogenic factors on the sea bottoms).

The paper is well written and easy to read. The premises in the introduction and the logical flow that brought the authors to the study design are both clear. The methodology is clear as well (but see below for some missing details) and the data analysis conducted rigorously. The results highlight the most important insights obtained from the experiments. Nonetheless, some issues emerge dealing with the description and reliability (or possible biases) of the resuspension assay and with the discussion, which is limited and provides the analysis only a small portion of the other (not considered factors) possibly explaining (>40%) of the CO<sub>2</sub> release variability. The discussion, as it is presented now, opens more questions than the answers (though credible) the study provides. I'm convinced that the effort in doing this study and the potential of the obtained results can be conceivably more properly and deeply discussed.

I'm convinced that this study deserves publication and the information it includes could become a starting point to include sediment heterogeneity among the factors to be considered when planning demersal fisheries, but it needs a larger effort to discuss more affordably the strengths and the weaknesses of the presented results.

**Response:**

We thank the reviewer for their thorough assessment of the manuscript which has drawn us to focus more deeply on purpose of our study. We aim to identify sediment types that are at highest risk of producing CO<sub>2</sub> when resuspended and to identify what

relationships with sediment characteristics influence the variability of resuspension-induced CO<sub>2</sub> production. We will address the reviewer's suggestions by adding more details to the description of the resuspension assay, discussing the observed patterns of our results in more depth with more balance on the strengths and weaknesses, and highlight the implications for spatial fisheries management more clearly.

Below I provide some suggestions to improve the manuscript.

The putative role of sediment heterogeneity on the effects of impacts (including the natural ones) affecting the sea bottom is somewhat obvious. This, however, has been not considered in deep by science nor policy making. Despite the authors declare it, a deeper and more accurate description of the potential implications of considering it is not anticipated in the introduction, nor extensively addressed in the discussion.

Response:

Thank you for highlighting this gap. We agree with the reviewer – it is a critical issue is that the role of spatial sediment heterogeneity is not deeply considered in seafloor disturbance studies, and it is far from being adequately integrated in regulations of demersal fishing impacts. We addressed at **lines 312-317** and **lines 348-358** of the revised discussion.

The manipulation method used for resuspending the sediment in this study derives from an early paper published by the same scholars' team. Despite they prompted in the former article "Depending on the research question, we recommend trials with local sediments to determine. How optimal incubation times, core sizes, and sediment to water ratios", no mention of these important data is provided in the manuscript much this affects the observed interactions? Moreover, other important details about the resuspension assay are missing, which obliges the reader to jump between the two papers for these details, that could be added in a supporting methodological file. This, obviously, does not prejudice the validity of this manuscript. Nonetheless, either in the former paper or in this manuscript, there is a possibly relevant bias: the energy employed to resuspend the sediment was apparently the same for all samples ("The jar was topped with filtered seawater, sealed airtight and gently inverted for 30 s. The sediment-water mixture was left to settle for 30 s, the jar was re-opened and the initial oxygen concentration was measured"). Considering the ample variety of grain size composition of the manipulated sediments, this could represent a possible uncontrolled simplification. This appears a possible not irrelevant bias: since sediment grain size is tightly related with its compactness, different sediments are differently impacted by the same amount of energy. This means that the simulated resuspension cannot represent the mechanic impacts of trawling on sediments with different characteristics. In this sense, the authors can use their results only to discuss the effects of a severe resuspension on sediments, but without providing a comparison of

the energy used to simulate resuspension and that causing the resuspension by the “average” trawl, the results cannot be effectively translated into possible demersal fishery management plans accounting for sediment heterogeneity.

Response:

We apologize for the lack of information and have added information on incubation time, core size and sediment to water ratio for the resuspension assay in a more detailed description of the resuspension assay method (**section 2.2**).

We also thank the reviewer for their insights into the relationship of grain size (compactness), energy and amount of resuspended sediment. Yes, a trawler with the same gear will cause more resuspension disturbance in muddier/less compact sediments than in sandier/more compact sediments and there exist useful modelling approaches estimating the amount of resuspended sediment and organic C mineralized by different trawling gear at different grain sizes (O’Neill & Ivanović, 2016; Porz et al., 2024). The critical problem that we aim to solve is how spatial sediment heterogeneity influences resuspension impacts on organic C mineralisation to provide spatially more detailed information to policy makers and fishery management. So instead of developing a model based on assumptions of degradation rate constants, trawling gear, penetration depths of trawling gear, hydrodynamic drag etc., we have taken a different approach to solving the problem using the resuspension assay method. The empirical resuspension assay results may not be representative of the mechanic impact of individual trawls, but they do inform decision makers on where sediments are at most risk of losing stored organic C when disturbed (**now discussed at lines 333-336**). By adding knowledge on the role of sediment heterogeneity in resuspension impacts, our approach can contribute to the translation into demersal fishery management plans.

In the methodological paper the authors identify different RQ constants, but here they used only one. Is this correct? Is there any possibility to distribute better this constant?

Response:

This has been corrected to differentiate between inner shelf (<50m – RQ = 0.9) and outer shelf (50-200m – RQ = 0.85) (Jørgensen et al., 2022). The BRT analysis has been revised (**lines 150-159**) and corrected RCO2P have been used. This has not changed the key results of this study.

Metric	BRT results original manuscript	Revised BRT results with corrected RCO2P (Mud and M-Sand excluded due to collinearity with OM)
R <sup>2</sup>	0.57 ± 0.1	0.57 ± 0.1
RMSE	0.56 ± 0.1 mmol CO <sub>2</sub> m <sup>-2</sup> h <sup>-1</sup>	0.54 ± 0.07 mmol CO <sub>2</sub> m <sup>-2</sup> h <sup>-1</sup>
feature importance	7. OM, SHAP = 0.42 8. Depth, SHAP = 0.13	7. OM, SHAP = 0.47 8. C-Sand, SHAP = 0.13

by mean  SHAP	9. F-Sand, SHAP = 0.12 10. M-Sand, SHAP = 0.12 11. C-Sand, SHAP = 0.12 12. OM:Phyto, SHAP = 0.08	9. Depth, SHAP = 0.10 10. F-sand, SHAP = 0.9 11. OM:Phyto, SHAP = 0.07 12. Chl.a, SHAP = 0.04
------------------	---	--

The lack of a control with OC-free sediments would have helped to clarify whether the CO<sub>2</sub> production measured in the resuspended microcosms depends only from OC mineralization. I see this now is unfeasibly corrected, but a possible bias from this should be acknowledged.

Response:

To our knowledge, it is not possible to create an OC-free sediment sample without interfering with the redox conditions and the composition and activity of microbial community that carries out OC mineralisation. But we agree with the reviewer that it is critical to acknowledge possible biases. We will address this by adding discussion on the link between SOD and organic carbon mineralisation and potential biases in SOD quantification due the potential oxidation of reduced species (**lines 336-345**).

The authors used the OM:Phyto ratio as a proxy for OM freshness. They referred to a paper in which the authors, however, used the percentage fraction of phytopigments and biopolymeric C (which does not account for total OM). This, honestly, does not modify the significance of its “new” use, but should be acknowledged.

Response: The reference is updated.

The lack of information about the composition (and origin) of the OM in the sediments makes difficult making any generalization. The authors acknowledge that different proportions in labile and refractory components could lead to different reactivity to oxygen of OM. Nonetheless, this issue needs to be discussed more deeply.

We thank the reviewer for pointing this out. Our aim was to use measures that are simple, affordable and quickly to derive, e.g. organic matter content based on LOI, OM:Phytopigment ratio, so that this approach can find wide application, both in research and for other stakeholders. We have deepened the discussion around the potential influence of OM composition and origin but kept it concise as we do not have measurements of composition and origin (**lines 286-301**).

## Response References

- Bartl, I., Evans, T., Hillman, J., & Thrush, S. (2025). Simple assay quantifying sediment resuspension effects on marine carbon storage. *Methods in Ecology and Evolution*, 16(2), 309–316. <https://doi.org/10.1111/2041-210X.14479>
- Boulesteix, A. L., Janitza, S., Kruppa, J., & König, I. R. (2012). Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(6), 493–507. <https://doi.org/10.1002/WIDM.1072>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36(1), 27–46. <https://doi.org/10.1111/J.1600-0587.2012.07348.X>
- Global Fishing Watch*. (2025). [www.globalfishingwatch.org](http://www.globalfishingwatch.org)
- Jørgensen, B. B., Wenzhöfer, F., Egger, M., & Glud, R. N. (2022). Sediment oxygen consumption: Role in the global marine carbon cycle. *Earth-Science Reviews*, 228, 103987. <https://doi.org/10.1016/J.EARSCIREV.2022.103987>
- Kalapurakkal, H. T., Dale, A. W., Schmidt, M., Taubner, H., Scholz, F., Spiegel, T., Fuhr, M., & Wallmann, K. (2025). Sediment resuspension in muddy sediments enhances pyrite oxidation and carbon dioxide emissions in Kiel Bight. *Communications Earth and Environment*, 6(1), 1–14. <https://doi.org/10.1038/S43247-025-02132-4>;SUBJMETA=106,4112,4113,47,704;KWRD=BIOGEOCHEMISTRY,CARBON+CYCLE,ELEMENT+CYCLES
- Lucas, T. C. D. (2020). A translucent box: interpretable machine learning in ecology. *Ecological Monographs*, 90(4), e01422. <https://doi.org/10.1002/ECM.1422>
- O'Neill, F. G., & Ivanović, A. (2016). The physical impact of towed demersal fishing gears on soft sediments. *ICES Journal of Marine Science*, 73(suppl\_1), i5–i14. <https://doi.org/10.1093/ICESJMS/FSV125>
- Porz, L., Zhang, W., Christiansen, N., Kossack, J., Daewel, U., & Schrum, C. (2024). Quantification and mitigation of bottom-Trawling impacts on sedimentary organic carbon stocks in the North Sea. *Biogeosciences*, 21(10), 2547–2570. <https://doi.org/10.5194/BG-21-2547-2024>
- Wilson, P. S., & Vopel, K. (2015). Assessing the Sulfide Footprint of Mussel Farms with Sediment Profile Imagery: A New Zealand Trial. *PLOS ONE*, 10(6), e0129894. <https://doi.org/10.1371/JOURNAL.PONE.0129894>