Reviewer 1:

This is an interesting paper that addresses an important question concerning the importance of physical and biological processes in sediment mixing. I think the Introduction of the paper does not match the title very well. This initially mis-led me. Firstly, except for the extremes of physical activity the issue is about the importance of both physics and biology not one over the other. Secondly, the paper needs to be clearer on prioritising the importance of physical and biological processes to sediment ecosystem processes or the methods used to make the measurements. Both are important but the text jumps around, which does not help with the clarity of the message. The end of the Introduction states that the goal of the study is to "unravel the relative contributions of bioturbation and physical dynamics to sediment" and to do this appropriately a range of tracers that work over different times scales are used. But the Introduction then jumps to more methodological questions "(i) can single-grain luminescence be meaningfully applied as a bioturbation tracer in such dynamic environments? (ii) how do short- and long-term tracers aid in distinguishing bioturbation from physical mixing?" The ecosystem significance is the results are returned to at the end of the Discussion (c L580) but this is not well developed.

The central hypothesis of the paper is stated as "if biotic mixing dominates, we will observe tracer-dependent mixing and, depending on the benthic community, diffusive or advective transport". While the paper is based on a field study in mud and sand habitats – the assumptions of how physics and biology interact in mixing sediment is not framed in a generalisable way. Storms (extreme events) are not the only source of sediment transport. In many intertidal flat environments, there is an important role for locally generated wind waves in mixing sediment and generating ripples. This can involve an interaction between tides, waves and biological sediment stabilisation or destabilisation. The description of processes around L 65 is too narrow – it may well represent the specific of this study site but think how this is framed. This interaction of physical and biological phenomena does not lead to the assumption of event driven sediment horizons vs deep mixing by animals (around L425).

To clarify that one process does not exclude the other, and to specify the intertidal context of our results, we have rephrased the title: "*Worms and storms: shedding light on bioturbation and physical mixing on an intertidal flat by combining multiple tracers*".

Secondly, we revised the last paragraph from the introduction, already indicating that this is a case study, the physical processes that play a role here, and to clarify the twofold aim of this study (line 152-166): "*The goal of this case study is twofold: firstly we aim to unravel the relative contributions of bioturbation and physical dynamics to sediment mixing in a (micro)tidal flat environment, using tracers that act on different timescales and on different sediment fractions. Secondly, we test the application of single-grain luminescence to study sand grain reworking in the intertidal environment, and compare tracer profiles with traditional short-term (Chl-a, luminophores) and long-term (^{210}Pb) tracers. We aim to answer the questions: (i) how do short- and long-term tracers aid in distinguishing bioturbation from physical mixing? And (ii) can single-grain luminescence be meaningfully applied as a bioturbation tracer in such dynamic environments? We hypothesise that, if biotic mixing dominates, we will observe tracer-dependent mixing and, depending on the benthic community, diffusive or advective transport. With intense bioturbation, the slow-to-bleach luminescence signals will also be more fully bleached. If physical dynamics, such as wind and tidal reworking, are responsible for the deposition and sediment mixing, we expect a clearer stratification. Slow-to-bleach luminescence signals will be*

poorly bleached. We compare two adjacent intertidal locations: a sandy and a muddy flat in the Dutch Wadden Sea."

Furthermore, we acknowledged that the results are site-specific in the introduction (line 84-85: "*Furthermore, the relative magnitude of these processes is highly site-specific.*") and the methodology (line 176: "*Our findings apply therefore to similar microtidal sites.*") and the discussion, where we also describe how the methodology is more generally applicable than the findings (line 576-577: "*Although our findings on sediment reworking are site-specific, the tracers used here are more broadly applicable, provided that the limitations of each tracer are considered.*")

Minor points:

- Abstract: "Luminescence age distributions suggest that quartz luminescence signals were fully reset upon recent deposition, while bioturbation enhanced resetting of feldspar luminescence signals" this is not a very clear statement for the Abstract.
    - We have replaced this sentence by: *"The combination of luminescence signals suggested that after deposition, not all sand grains did resurface repeatedly and for longer time periods through bioturbation."*
- "sedimentary habitats by so-called 'bioturbation" – omit so-called.
    - Done
- The paper is methodologically appropriate in terms of the tracers, albeit I have not experience with one set of these. However, the sediment long cores with an n=2 will underestimate biogenic effects in heterogeneous sediments. This needs to be acknowledged as a limitation and used in the interpretation of the results particularly around the relative importance of sediment mixing and episodic deposition.
    - Although we found high spatial heterogeneity in the short-term tracer reworking, on the long term this heterogeneity is evened out. Although current biological community is stochastic, over the longer term the biota effects on sediment reworking will be averaged over all organisms that have been present. Long-term tracers therefore give an integrated outcome rather than a snapshot of particle distribution. In long-term sedimentology and geology, such a low replication is therefore common. To clarify, we added this sentence to the discussion (line 573-575): *"For the long-term tracers, the number of replicates was lower, however because the distribution of these tracers captures the integrated activity of the macrozoobenthic community over a longer time period, we would expect similar results even if we had used a higher number of replicates.*"
- "de Boer et al., in review" do not cite unless in press/published
    - This manuscript is currently in press in Scientific Reports, we adjusted this in the reference list.
- Please explain how the sand bleaching rates are linked to levels of radiation – are these a constant or do they vary with latitude or air pollution?
    - For clarification, we added: *"The light exposure needed to fully reset luminescence signals varies for minerals and the luminescence signal used. For quartz OSL a few tenths of seconds of direct sunlight is sufficient for resetting (Lindvall et al., 2017) while feldspar post IR IRSL signals reset much slower (de Boer et al., 2026b). Moreover, ambient light levels vary in time and space as they are affected wide range of aspects including time of day, latitude and inundation (de Boer et al., 2026b)."* (line 118-122)

- Fig 2 -depth scale on the long cores is not visible
    - We adapted figure 2.
- Around L210: "The downcore distribution of 210Pb in the mud fraction (<63 µm) of the sediment was determined indirectly through alpha spectrometry measurement of its grandchild isotope 210Po. Spell out Polonium-210 on first use. There is no reference cited for this method.
    - We added "*pollonium-210 (^{210}Po) (de Stigter et al., 2011)*"; for consistency, we also added "lead-210" to line 109.
- "For cosmic dose rate calculation (Prescott et al., 1994), we assumed gradual burial of the samples to the present depth. Why given intro text where you introduce episodic events.
    - We modified the text to indicate that this is a crude assumption, but that it has little implications as the cosmic dose is a minor part of the total dose rate (line 259-261): *"This is a crude assumption given the episodic nature of deposition and potential reworking of grains after deposition in this environment. However, the effects of this assumption are minor as the cosmic dose constitutes a minor fraction of the total dose rate experienced by the grains (< 25 % even for samples closest to the surface)."*
- "Macrozoobenthic community biomass and abundance were higher on the muddy site …" acknowledge that this does not include the *Arenicola*.
    - We added "(excluding *Arenicola marina*)" (line 312).

Reviewer 2

General comments

The ms by Kooistra et al. presents an interesting work that aims studying both physical and biological particles mixing in intertidal environment. The *in situ* work carried out in qualitative comparison between different experimental short-term and natural long-term tracers, including a tracer (luminescence signal) that, to my knowledge, has never been used in such a multi-tracer approach in the marine environment.

The experimental protocol used in such natural conditions and had its limitations. They were mentioned by the authors, for example the relating to the size of the core samplers used for Chl-a and luminophores, which were insufficient to capture the biological heterogeneity. Adressed also by the authors in their comparison of tracers, the specific characteristics that need to be taken into. For example, the luminescence signal can be reset, unlike other signals.

Based on the combination of all the information provided by the various tracers, the authors proposed scenarios for the history of deposits in the two studied sandy and muddy sites. But also a smart decision tree guiding the choice of different tracers to study the transport capacity of intertidal environmental particles, whether the movements generated by abiotic or biotic processes are of interest. It also takes into account the grain size distribution of the site under study. A very interesting tool for geochemists and/or biologists interested in sediment transport.

I recommend this manuscript for publication after minor corrections.

Specific comments

- L77: perhaps write "(per)turbation" instead, in order to recall the etymology and meaning of the word bioturbation?
    - Done
- Table 1, it is indicated an Investigation timescale of 15 days - 3 years for Chl-a. Theoretically I agree, but practically I think it should be rather considered as a non-conservative short time tracer. Also this will depend if experimental pulse protocols (like with luminophores) or natural measurements are implemented.
    - We added *"non-conservative short term"* in line 104.
- L140: "Moreover, biotic mixing may occur during low tide...". Is this a general fact whatever the species involved? Any references to support this statement?
    - We added (line 150): "*.. biotic mixing also occurs in the moist tidal flat during low tide ... (Cadee, 1976).*"
- Figure 2. One cannot read the depth indications on Fig 2d. Increase the font size and darken the characters, or indicate a general scale?
    - We adapted figure 2.
- L175: indicate the range of depths
    - We added the approximate core depth (line 187: "(depth approximately 0.25 m)", and also added area (line 192: "area and depth"), which was another reason why we did not manage to sample *A. marina*, that we forgot to mention here.
- L178: I guess the surface area covered by the spade was the same as that covered by the hand corer?
    - Unfortunately, the sampling area was too low to capture a realistic lugworm density—on average this would have resulted in less than 1 individual at the sandy site. Also at the muddy site, this area would have been too low to get a representative number of individuals to get reliable individual biomass estimates. Instead, we determined the number of individuals was per surface area through cast counts. We added the following to clarify (line 192-195): *"Due to the limited area and depth of our samples and the lower density of larger Arenicola marina individuals, we did not capture this species in our samples. From the field observations of large A. marina casts at the sandy location, we expected this species to dominate bioturbation."*
- L258: why not have applied the same type of protocol (i.e. vertical slicing, homogenisation then measurement) for the luminophores than for the Chl-a? You could have obtained results for both tracers within the same cores.
    - We chose to analyse the luminophore cores following a similar protocol as had been used in previous comparable studies to allow for analysing a higher number of cores. Due to the high spatial heterogeneity that we observed already in the *Chl-a*, we did not expect to find exactly the same patterns in each core. The aim of analysing the luminophore cores was thus to find evidence of recent bioturbation of fine particles, rather than quantifying the reworking, for which the horizontal slicing protocol would have been more suitable. Lastly, the luminophore layer covered the sediment surface (including the microphytobenthos) and would thus probably have resulted in a periodical decrease of surface Chl-*a* supply. We added the following to the stress the aim of this analysis (line 268–269): "*In order to find evidence of recent bioturbation of the fine particles, the frozen cores were sliced..*". In the discussion we added (line

582-583): *", and aligning tracer sampling protocols to improve intercomparison is not always feasible."*

- Paragraphs 2.7 and 2.8 from L264: How many sediment cores were collected each time for the different tracers? Looks like three to me for luminophores from the text (but two from Fig. 5c). On the other hand, the information I cannot find out for Chl-a. Linked to this, there is no variation in the data presented in fig. 5. (line 350). Using a small 3.6 cm diameter tube in a 30 cm x 30 cm square gave you many opportunities for replication to catch the biological heterogeneity. A point you mentioned in line 543. Enhancing replication could be suggested as a solution here.
    - We sampled three luminophore cores per time-point indeed, but they could not all be analysed due to artefacts. For clarification, we added to the methods (line 287): *"The resulting sediment cores (one per plot)"*; and (line 291-292): *".. leaving two replicate cores per site per time point."*. We also clarify the number of cores for Chl-*a* (line 301): "*respectively two and three cores per site*". Furthermore, we added to the discussion (line 572-575): *"., and a higher number core replicates may therefore have increased consistency of tracers profiles for these short-term tracers."*
- Fig. 3. Distinguish the colours of the 250-500 and 500-1000 micron fractions a little better.
    - We adapted figure 3.
- L341: Even though *Scrobicularia plana* is able to bury deeper than *Cerastoderma edule*, 18 cm and 20 cm deep is extremely deep (if not impossible to reach) for those two bivalve species and more especially small individuals (as captured in 3.6 cm diam core). I suggest nuancing this explanation and perhaps also suggesting the hypothesis that you may have captured a burrow opening into which the tracers entered?
    - This is a good point, and although we measured a few luminophore pixels at these depths, the peaks in luminophores lie shallower (at the living depth of the bivalves). We mention the hypothesis on burrow infiltration in the discussion (line 510-511: "*Other bioturbators might also have contributed to non-local transport of fine tracers at both sites, either by active particle transport or through passive processes such as burrow infiltration*."). We furthermore added depths and rephrased line 367-371: *"Some luminophore particles ... around their burrows (Appendix A5, Fig. A2; Scrobicularia plana at 12–15 cm depth at the sandy site and Cerastoderma edule at 0–3 cm depth at the muddy site)."*
- L346, more replication could have helped resolve this variability issue in such a heterogeneous environment.
    - We recommend more replications in line the discussion (line 572-573)
- Paragraph 4.3 the "community part" from L469: Here, you establish some indirect connexions between some tracers distributions and the community present overall (i.e. not determined from the specific tracer cores). On the other hand, for the luminophores cores, a direct relationship with specific species found in the core is proposed. To be consistent with my previous comment on the role of bivalves in tracers transfer in one luminophore core, I suggest having the sentence "...as evidenced by luminophores that had moved rapidly to the living depth of this bivalve (Fig. A2)." presented as a hypothesis rather than a fact.
    - To rephrase this as a hypothesis, we modified this sentence to "*as suggested by luminophores..*".

- L578: "...the sand fraction is not intensively reworked". This could be different under the influence of other benthic bioturbating communities.
    - *We added: "is not intensively reworked by the present bioturbating community"*