Reviewer 1

Although the manuscript is very well written, I have substantial concerns regarding the methodology, particularly the production of the prediction maps, the estimation of OC stocks throughout the Holocene sediment column, and the approach used to demonstrate the importance of near-bottom lateral OC fluxes in shaping the spatial distribution of sedimentary OC stocks. Nevertheless, I believe that this study could make a valuable contribution to the journal following major revision. To support such a revision, I provide three major comments below, followed by several more general comments.

We thank the reviewer for their thorough review and constructive comments. We indicate below how we intend to address each comment in a revised version.

Major comments:

1. Predicted maps of OC concentrations

The authors use kriging and a deep neural network based on four predictors, all of which are related to location: latitude, longitude, water depth, and distance from shore. Although the ML approach also relies primarily on spatial information, the two methods produce markedly different spatial patterns. The authors should explore in greater detail why these differences arise. For example, does the ML model assign greater importance to water depth and distance from shore, potentially explaining the sharper spatial gradients? To what extent can the relative influence of the different predictors be assessed? I understand that deep neural networks are not inherently interpretable, but it should nevertheless be possible to investigate feature importance using appropriate methods.

This is a valid point. We plan to use permutation importance on the held-out data, plus partial-dependence and ICE curves for each predictor. One caveat applies here: water depth and distance from shore are closely correlated, both with each other and with the coordinates, and permutation importance becomes unreliable when predictors are correlated. We may therefore also give importances for groups of predictors (coordinates, water depth, distance from shore) and accumulated local effects, and compare the ranking against the feature importances of a tree-based model fitted to the same data.

Given the pronounced differences between the two approaches, I suggest adding predictions from tree-based ML models. Such models would also allow the authors to examine feature importance and use partial-dependence plots to interpret the relationships between individual predictors and predicted OC concentrations. Predictions based on alternative ML methods would help assess the robustness of the results and better characterize the uncertainty associated with the spatial predictions.

I also do not understand why the ML model uses only these four spatial variables as predictors. This limited predictor set likely weakens the model's capacity to extrapolate to areas of the Baltic Sea with sparse observations, such as the Gulf of Bothnia. It would be preferable to include predictors that are more directly and causally related to OC production, inputs, transport, and deposition. For instance, remotely sensed chlorophyll-a concentrations could provide a useful indicator of autochthonous OC production that may ultimately contribute to sedimentary OC. The manuscript itself refers to previous Baltic Sea studies using this type of information.

We have since tested this directly. Several other model classes were fitted to the same four predictors and the same data; on the held-out test set, extremely randomised trees reach an RMSE of 1.81 wt%, gradient-boosted trees 1.89 to 1.92 wt%, and Gaussian-process regression 1.92 to 2.05 wt%, against 1.80 wt% for the neural network and 1.76 wt% for a combination of the network and a tree model. These baselines will go into the revised manuscript. They show that the predicted distribution does not hinge on the particular network architecture, and the tree model supplies the feature importances and partial-dependence curves the reviewer asks for. We may also derive two predictors from the bathymetry we already use, seabed slope and a compound topographic index, and report how the choice of predictor set affects the error, so that the value of extra features is measured rather than assumed.

Furthermore, the manuscript argues that lateral OC transport is an important control on sediment accumulation. To test this hypothesis more directly, the simulated near-bottom sediment transport could be included as a predictor in an interpretable ML model. The authors could then assess its importance, for example through SHAP analysis, and its effect on predicted OC concentrations using partial-dependence plots. The sediment-thickness map used to estimate total OC stocks might also be a useful predictor.

We think that using the simulated near-bottom transport as an additional predictor of OC concentrations would not be very meaningful, since the result of the ML model was used to initialize the sediment fields of the transport model, both for OC contents and sediment composition, so feeding the results of this model back into the ML prediction would result in a kind of autoregressive feedback that would only amplify existing biases. However, we do agree that additional analysis of the transport patterns could be used to strengthen the argument that lateral transport is important for sediment and carbon accumulation, and we will add such analysis to the revision.

Other potentially relevant predictors include riverine inputs of water, sediment, and terrestrial particulate OC, perhaps derived from existing spatially explicit estimates such as Liu et al. https://www.nature.com/articles/s41561-024-01524-z. More advanced bathymetric variables that may be relevant to sediment transport and accumulation, including seabed slope and compound topographic index, could also be derived from the Baltic Sea bathymetry data.

This is a good idea and we may add such additional predictors depending on the outcome of the comparison to Parameswaran et al. (2025).

2. Calculation of total Holocene sedimentary OC stocks

The study estimates total Holocene sedimentary OC stocks by multiplying the predicted OC concentration in the surface layer by the total sediment thickness reported by Miluch et al. and by sediment bulk density. It is unclear whether the bulk-density data are representative of the entire sediment column or only of the surface layer.

More importantly, this calculation assumes that OC concentrations are constant throughout the sediment column, which is highly debatable. One could hypothesize that anthropogenic soil erosion and eutrophication have increased OC inputs and burial in recent periods, resulting in higher OC concentrations near the sediment surface than in older deposits. Conversely, post-depositional decomposition would generally be expected to reduce OC concentrations with sediment depth. The calculation therefore also appears to assume that OC decomposition within the sediment column is negligible.

The authors argue that their estimates mainly represent refractory sedimentary OC, for which decomposition may be limited. However, based on the Methods section and the description of the OC concentration data, it is unclear whether - and, if so, by which methods - the refractory OC fraction was distinguished from more reactive OC. I therefore have substantial concerns about the validity and uncertainty of the estimates of total Holocene sedimentary OC stocks and burial rates presented in the study. At a minimum, the authors should justify these assumptions, assess their sensitivity, and clearly characterize the resulting estimates as approximate if depth-resolved observations are unavailable.

We agree that the absence of depth-resolved data in our analysis can bias the estimates of Holocene OC-stocks. While we have discussed and attempted to quantify such uncertainties in 4.5 (Limitations), we agree they should be quantified in a more rigorous way. We intend to test the sensitivity of our assumptions and derive robust uncertainty estimates, e.g. using sediment OC profiles for the Baltic recently compiled by Langley et al. (2026). While we do not expect the labile OC fraction to be particularly important for the Holocene stock, we will attempt to quantify this using existing measurements and biogeochemical model results.

3. Calculation of near-bottom POC transport

The study combines the predicted map of surface-sediment OC density with an ocean-current and sediment-transport model to estimate the potential magnitude of near-bottom OC transport. It is problematic that these flux estimates cannot be evaluated against observations, although I recognize that this reflects a lack of suitable observational data rather than a shortcoming that the authors can readily resolve.

Nevertheless, I wonder whether the observed surface-sediment OC concentrations could be used to test the relevance of the simulated lateral transport more directly. I strongly recommend including variables derived from the simulated near-bottom sediment transport as predictors of surface-sediment OC concentrations. If the authors could demonstrate that these transport variables explain a substantial proportion of the spatial variation in observed OC concentrations, this would provide much stronger support for their conclusion that lateral transport drives the spatial pattern of OC burial across the Baltic Sea.

For this purpose, it may be preferable to use a more interpretable ML method, such as a random forest or another tree-based model. Such models would facilitate analyses of feature importance and partial dependence and may be more appropriate than a deep neural network for a moderately sized dataset.

We agree that more information is needed to support the validity of the transport model. As the reviewer correctly states, no data is available to directly validate such transport. As stated in a response to comment #1, we think that using the simulated near-bottom transport as an additional predictor of OC concentrations would not be very meaningful, since the result of the ML model was used to initialize the sediment fields of the transport model. Instead, we intend to analyse the transport and OC and mud redistribution pattern from our transport model in more detail to see if it matches observed areas of accumulation and in that case, it will support our argument that the simulated transport is reasonable.

General comments:

L17-18: From the abstract itself it is not fully clear what you mean with "maritime boundaries". You should make clearer that this refers to boundaries between countries, maybe by linking to the relevance of these fluxes for national carbon budgets.

We agree that this may be confusing and will attempt to clarify the phrasing in the abstract. However, we do note that we have selected and defined the terms "maritime boundary" and "maritime zone" deliberately since an EEZ boundary is not technically part of a country's national boarder, and it is not yet clearly defined which parts of a marine area (such as EEZ and territorial sea) shall count towards a country's national carbon inventory.

L67-70: Here I miss some details about the data. That is all sediment OC concentration in the top 10 cm? Or does the depth of that top layer varies between the datasets?

We will clarify that these are all near-surface samples taken mainly with grab samplers and are assumed to represent the top 10 cm.

Section 2.2: In how far have these models been validated against observations?

We will add relevant details on validation (mainly hydrodynamics; water level, currents, temperature, and salinity) performed in Kossack et al. (2023). However, we note that this validation focused on the North Sea part of the model domain, which is why we have added additional comparison to hydrodynamic measurements in the Baltic in the supplement. Nevertheless, we consider the proper representation of the North Sea as an important component since it governs the exchange of dense water with the Baltic, especially during inflow events, which are especially relevant for near-bottom flows. We may add additional validation figures for hydrodynamics to the appendix.

L131: Where do you have “deep ocean” in your study area?

The model domain encompasses part of the Northeast Atlantic. We will add a map of the entire model domain to the supplement.

L133: Are those six years representative for the period over which the top 10 cm of sediments were deposited?

There is strong evidence that the bottom circulation and inflows in the Baltic Sea have remained relatively stable throughout the last decades (Mohrholz, 2018; Jędrasik and Kowalewski, 2019). Nevertheless, our results do show interannual variability in POC flux for some sub-regions, reflecting the interannual variability in hydrodynamics. High computational demand prohibits multi-decadal simulations at this stage, but this may be necessary for longer-term flux estimates (as noted in the discussion section). The model period was selected to capture the strong inflow event of 2014/2015, though its effect is not immediately obvious in the computed POC fluxes. We may add a measure of simulated flux stability to discuss this further.

L149: Not clear how you estimated the recalcitrant fraction of the OC in the surface sediments from the data and extrapolation methods you have used.

The recalcitrant fraction has not been separated, as we assumed the contribution of the labile fraction to be relatively minor. This assumption will be qualified and discussed in the revised version. We note that there is no agreed upon definition of what constitutes “recalcitrant" in the spectrum of organic matter lability, and, to our knowledge, no dedicated experiment has attempted to separate the “labile” from the “recalcitrant” fraction in the Baltic Sea. However, we may use the results of a biogeochemical model to estimate possible proportions of this seasonally degraded fraction.

Figure 3: It would be great to have a map of the bathymetry maybe in the supplement. That could help with the interpretation of panels a) and d), as the spatial patterns here are likely driven by water depth, which was used in the ML model, but not in the Kriging based interpolation.

This is a good idea, we will show the bathymetry in the supplement in combination with the model domain.

L210: Instead of "top meter", do you mean "top 10 cm"?

Yes, this will be corrected, thank you.

Section 3.3: Have at least the water and sediment fluxes been validated against observations? It is also not clear to me if you only used the predicted map of surface sediment OC densities based on the ML model, or if you also performed these analyses with the result from the Kriging-based interpolation. It would be interesting to see in how far the results vary based on the predicted sediment OC map.

Water "fluxes" in the sense of hydrodynamics have been validated in Kossack et al. (2023) as well as using the comparison to the measured inflow event in 2014 (see supplement). However, sediment fluxes cannot be validated as this would require near-bottom SPM measurements which, to our knowledge, are not available. It is an interesting idea to repeat the simulations with the Kriging map. Since the simulations are very costly in terms of computational demand, we may simulate one year with initialization according to the Kriging solution to get a sense of this effect.

L271ff: In how far can you say that you obtain similar results? Do you only compare the average surface sediment OC concentration, or have you also compared the spatial patterns in these OC concentrations? Only if you get a similar spatial pattern, you can convince me that you get similar results with so much less predictors.

We agree with this comment, we will add a quantitative comparison of our result to that of Parameswaran et al. (2025) to see if this argument holds.

L277: But you conclude that lateral transfers have a strong impact of the distributions of surface sediment OC concentrations. If you could show that including bottom current strength improves the performance of your ML model, this would strongly support your conclusion.

We argue that using bottom currents or sediment fluxes in another ML prediction would result in autoregressive feedback, as explained above. We do agree that additional analysis of the transport patterns could be used to strengthen the argument that lateral transport is important for sediment and carbon accumulation, and we will add such analysis to the revision.

L331-333: But the estimate by Thomas et al. seems very close to your estimate.
L335-336: Ok, this means that these estimates are not comparable. You could have started with that.

This may be a misunderstanding: Thomas et al. (2010) estimate a net *export*, while we estimate a net *import* by a similar amount, so the direction of net transport is opposite. We may rephrase to clarify.

L351ff: In how far would it be necessary to distinguish terrestrial fraction from the authochthonous (produced within the marine ecosystem) fraction of the buried OC?

This will be very relevant to avoid double counting. We touch on the issue of terrestrial input in l.410-416, but may add a discussion of its relevance to accounting here as well. An estimate of total riverine OC input to the Baltic Sea may also be added.

L389-392: What about shifts over the last 100+ years following eutrophication, soil erosion, overfishing etc.? Wouldn't these anthropogenic perturbations induce substantial shifts in OC burial?

While more recent shifts may also have altered OC burial, we consider them relatively minor compared to those during the early Holocene (at least, we do not see a comparable recent shift from sediment cores). Nevertheless, we will add this aspect to the discussion and account for recent changes using depth-resolved OC content information (Langley et al., 2026). A discussion on recent human impacts based on literature (e.g., Bunke et al., 2019; Porz et al., 2023) may also be added.

L400: Again, it is not clear how you defined and quantified the stable POC. Is that already given in the observation you used?

The recalcitrant fraction has not been separated, as we assumed the contribution of the labile fraction to be relatively minor for the upper 10 cm. This assumption will be qualified in the revised version.

L410: In how far could that driver be represented by the distance to shore predictor? In how far do you think this biases your prediction maps, in particular for the near shore areas?

We meant to state that the riverine flux itself has not been included as an additional source in the transport model. However, the deposited riverine OC will be included in the ML prediction, since the measurement data does not distinguish between marine and terrestrial OC. We will rephrase to clarify this and attempt to gauge the importance of each predictor, including distance to shore.

References

Bunke, D., Leipe, T., Moros, M., Morys, C., Tauber, F., Virtasalo, J. J., Forster, S., and Arz, H. W.: Natural and Anthropogenic Sediment Mixing Processes in the South-Western Baltic Sea, Front. Mar. Sci., 6, 677, doi:10.3389/fmars.2019.00677, 2019.

Jędrasik, J. and Kowalewski, M.: Mean annual and seasonal circulation patterns and long-term variability of currents in the Baltic Sea, J. Mar. Sys., 193, 1–26, doi:10.1016/j.jmarsys.2018.12.011, 2019.

Kossack, J., Mathis, M., Daewel, U., Zhang, Y. J., and Schrum, C.: Barotropic and baroclinic tides increase primary production on the Northwest European Shelf, Front. Mar. Sci., 10, doi:10.3389/fmars.2023.1206062, 2023.

Langley, B., Burdett, H. L., Cameron, K., Juul-Pedersen, T., Rouillard, A., Slaymark, C., and Kamenos, N. A.: Hotspots of Arctic and sub-Arctic marine sediment organic carbon are dominated by the Baltic, Barents and Chukchi Seas, Comm. Earth Environ., 7, 529, doi:10.1038/s43247-026-03720-8, 2026.

Mohrholz, V.: Major Baltic Inflow Statistics – Revised, Front. Mar. Sci., 5, 384, doi:10.3389/fmars.2018.00384, 2018.

Parameswaran, N., González, E., Burwicz-Galerne, E., Braack, M., and Wallmann, K.: NN-TOC v1: global prediction of total organic carbon in marine sediments using deep neural networks, Geosci. Model Dev., 18, 2521–2544, doi:10.5194/gmd-18-2521-2025, 2025.

Porz, L., Zhang, W., and Schrum, C.: Natural and anthropogenic influences on the development of mud depocenters in the southwestern Baltic Sea, Oceanologia, 65, 182–193, doi:10.1016/j.oceano.2022.03.005, 2023.

Thomas, H., Pempkowiak, J., Wulff, F., and Nagel, K.: The Baltic Sea, in: Carbon and Nutrient Fluxes in Continental Margins: A Global Synthesis, Liu, K.-K., Atkinson, L., Quiñones, R., Talaue-McManus, L. (Eds.), Springer Berlin Heidelberg, Berlin, Heidelberg, 331–421, 2010.