Review #5

We sincerely appreciate the editor's thorough and constructive review of our manuscript. We have carefully considered each remark and recommendation. Implementing the revisions has significantly enhanced the quality and clarity of our work. Below, we provide point-by-point responses to the editor's comments, using the following color-coding for clarity:

- **Editor's remarks**: Black (original text)
- **Our responses**: Red
- **Text modifications**: Blue (tracked changes in the revised manuscript)

L3: For greater clarity, please specify the actual years this period represents.

We added: (1995 – 2019).

L10: It would be helpful to include a brief comment on whether phosphorus retention (the 12% average) was found to be stable or also varied with changing riverine loads.

We added: In contrast, phosphorus retention is independent on loads.

L15: Please specify under what conditions such achievement is unlikely. e.g., "...unlikely under the current management targets"?

We followed the editor's advice and added: under the current management targets

L68: The phrasing "improve the input" is ambiguous in this context. Do you mean "... improve the representation of organic matter inputs from rivers in the model"? Please clarify.

We followed the editor's advice and changed the sentence to: They also argue that river-specific organic matter retention factors in coastal waters would improve the representation of organic matter inputs from rivers in models.

L71: "a spatially high resolved…" -> "a high-spatial-resolution"

We followed the advice of the editor and changed the sentence accordingly.

L87: "mode" -> "model"

We corrected the typo.

L101: "light climate" -> "light field"?

We changed light climate into light field.

L143: The single-sentence paragraph is redundant and can be deleted. The structure of the section is already logically implied by the subsection titles.

We removed the paragraph.

L145: “reasonable” -> “reasonably”

We corrected reasonable to reasonably.

L145-155: Please cite the specific figures where the data (e.g., bottom oxygen, mass transport values, stratification) are presented.

We added references to figures of oxygen, mass transports, and stratification in the appendix.

L156: Please briefly specify the key variables assessed and provide a short summary of the overall conclusion from the model performance evaluation in the appendix. This will give readers context before they navigate to the supplementary material.

We restructured this subsection and added missing information. The text reads now:

Appendix A presents an expanded model performance analysis. We evaluate time series and climatology of hydrodynamic (temperature, salinity, stratification, and mass transport), and biogeochemical (nutrients, chlorophyll-a, and bottom oxygen) parameters.

The model reasonably reproduces the climatology as well as the interannual variability of nutrients, temperature, salinity at stations KHM and C (for station locations, see Fig. 1). For bottom oxygen (Figs. A1 and A2), the model predicts lower values than in the observations. Reasons are that (i) measurements are not as close to the bottom as the model data, (ii) the measurement platform (vessel) itself disturbs the stratification, and (iii) at station C commercial ship traffic induces strong vertical mixing which is not part of the model.

The mass transport through the Dziwna channel (for location, see Fig. 1) is elevated compared to known values (Fig. A6). The reason is the truncated Dziwna channel in the model reducing the hydraulic resistance and facilitating an enhanced discharge at the expense of a lower discharge through the Swina channel.

The stratification of the water column at station KHM (for location, see Fig. 1) compares well with observed stratification (Fig. A8). Stratification establishes as events, such as those in summer, while during winter the water is well mixed. Stratification results in oxygen deficiencies, which yield phosphate liberation from the sediment. This process is not directly observed but is indirectly indicated by elevated phosphorus concentrations in summer (Fig. A7).

Our evaluation confirms that the model reasonably represents the Oder Lagoon ecosystem’s key dynamics, establishing it as a reliable tool for experimental analysis and system property investigation.

L160: Please provide the specific oxygen concentration thresholds used to define oxygen-depleted conditions and anoxia. e.g., < 2 mg/L or 1 mg/L?

Depleted conditions refer in our text to a zero oxygen concentration. We added (zero oxygen concentration) to the text.

L184-188: The statement that “a significant relationship between the retention rate and phosphorus loads does not exist” appears to be contradicted by Fig. 4b. The time series in Fig. 4b consistently shows higher relative P retention under the control run compared to the 50% load reduction, suggesting a load-dependent response. This conclusion is more accurately supported by Fig. 5d. The scatter plot in Fig. 5d shows no significant correlation when all data points are considered, which aligns with the stated lack of a significant relationship. Please clarify.

The editor is right. Both figures appear contradicting. We checked again the data base for these two figures and did not found any mistake. Indeed, we get a positive correlation between phosphorus load and relative retention as figure 4b suggests. However, this correlation is weak and the p-value of a Spearman rank test is 0.68 which confirms H0 that no correlation exists. The same test for nitrogen retention gives a p-value of zero and rejects H0. We noted this in the revised manuscript: A Spearman rank correlation test yields a non-significant regression coefficient ($p = 0.68$). This result fails to reject the null hypothesis of no correlation, thereby indicating that the apparent differences in the Fig. 4b are not statistically significant at conventional confidence levels.

L187-188: Please provide the result of this t-test (e.g., p-value).

See our comments to the previous editor’s remark. In the revised version, we used a Spearman rank test to account for non-Gaussian data. However, the result is the same.

L190: Please provide statistical evidence to support the qualitative claim that the correlation for P “appears less pronounced” than for N in Fig. 6.

We changed the paragraph and provided the statistical indicators: Our analysis reveals a further relationship between nitrogen and phosphorus retention efficiency and the lagoon's water residence time (Fig 6). The relationship for phosphorus is less pronounced but statistically robust ($p(N) = 0.0004$, $p(P) = 0.0017$). In general, we can state that the longer water remains in the lagoon, the higher the relative retention of nutrients in the lagoon.

L193: Please clarify how the 40% and 12% retention estimates were derived. Are they the mean of the annual retention efficiencies calculated over the entire 25-year simulation period?

We changed the sentence for more clarity: The mean of the annual retention over the 25 years simulation period is approximately 40% for nitrogen and 12% for phosphorus.

L193-194: As commented earlier, the conclusion that "the phosphorus retention capacity remains largely independent of load variations" is inconsistent with Fig. 4b. The argument should be based on the lack of a significant correlative relationship in the scattered data of Fig. 5d. Please clarify.

We changed the sentence to: Nitrogen retention capacity increases with reduced lagoon nutrient loads, while phosphorus retention remains largely load-independent, statistically confirmed by analysis of Fig. 5d data.

See also our response to remark L184-188.

L246-247: Please cite the specific panel of the figure.

We specified the panel number. Fig. 7f.

Figures 5-7: For the regression analysis presented in Figs. 5-7, please report the associated p-values to assess whether the correlations are statistically significant.

We introduced a sentence how we evaluated the significance of correlations presented in the manuscript at the end of chapter 2. This should prevent replications throughout the text:

The observed statistical relationships were evaluated using Spearman rank correlation tests, with significance assessed via p-values. The null hypothesis (H0) posits no correlation between the examined variables. If the calculated p-value falls below the conventional significance threshold (alpha = 0.05), we reject the null hypothesis, thereby providing statistical evidence for a significant relationship between the variables.

We present the p-values for correlations in Fig. 5 in the text; see our response to editor's remark "L184-188".

We added to the caption of Fig. 6: Both correlations are highly significant ($p < 0.002$), indicating strong non-random associations between variables.

We added to the caption of Fig. 7: Panels b-e show highly significant correlations ($p < 0.002$), while panel f demonstrates no significant relationship ($p = 0.35$).

L257: Suggest rephrasing the subsection title as "4.1 Model performance and limitations"

We followed the editor's advice.

L279-285: The statements in this paragraph should clearly distinguish between findings derived from the model and those from observations or literature.

All presented findings in this paragraph represent model results unless observations are explicitly noted. We have added clarifications to distinguish between simulated and empirical data.

We start this paragraph now: Model simulations show that near-bottom oxygen deficiency in the Oder Lagoon occurs as widespread, episodic events. The most affected …

L285: Please delete the citation of Fig. 1. Fig. 1 shows the location of Oder River but can't see the high nutrient loads. Citing Fig. 1 is misleading here.

We deleted the reference.

L290-291: The claim that the newly introduced limnic phytoplankton group was "necessary to achieve realistic biomass concentrations" requires substantiation. Please clarify: 1) where are the results for this group's abundance and its impact on biomass presented (e.g., which figure)? 2) what evidence supports the conclusion of "necessity"? Was this based on a model sensitivity experiment that compared simulations with and without this functional group? 3) please moderate the language to reflect the specific context of this study. Does this finding imply that this group is universally necessary for all model setups in this region, or was it a required component within the specific configuration of this model?

We have introduced a table showing the fraction of the model phytoplankton groups in the Oder Lagoon. Furthermore, we clarified the raised questions by rewriting the whole paragraph. We think, the role of the limnic phytoplankton group is now much better described. We also emphasis the fact, that our findings could be useful for model configurations of the whole Baltic Sea. The revised text:

The newly introduced phytoplankton functional group (limnic phytoplankton) is by far the most abundant model phytoplankton group (Tab. 2). This new group was necessary to achieve realistic biomass concentrations in the lagoon. The limnic group's adaptation to low-light, CDOM-rich conditions enables realistic phytoplankton biomass simulation in the model. In environments outside the lagoon, where salinity levels are substantially higher, the limnic phytoplankton group becomes effectively absent from the community composition. This exclusion results from growth limitations imposed by elevated salinity conditions, which exceed the group's threshold. This mechanism, in combination with the other three groups, allows us to apply the biogeochemical model ERGOM in coastal waters of the Baltic Sea as well as in the open Baltic Sea without parameter tuning. This is especially important when the model is set up for the entire Baltic Sea at a high spatial resolution, for example, 2~km or finer, where lagoons are partly resolved. In this case, the ecosystem model provides reasonable results in coastal waters, lagoons, and the open Baltic Sea. We applied this Baltic Sea model to create open boundary conditions for the Oder Lagoon model, ensuring proper domain connectivity with the larger Baltic Sea model. Thus, simulations with a coarse-grained model deliver nearly seamless data for the open boundaries of the local model setups.

L292-293: The statement that the model can be applied in "coastal waters as well as in the open Baltic Sea" is somewhat vague. Please specify if "coastal waters" refers specifically to the Baltic Sea's coastal areas or to coastal waters in a global, general sense.

We clarified this imprecise formulation in the revised paragraph provided in our response to the previous editor’s comment.

L294: Please express the spatial resolution in kilometers instead of nautical miles.

Changed into 2 km.

L300: The reference to Appendix A is too general. Please provide a brief summary regarding

which aspects of model performance are analyzed in the appendix and state the overall conclusion of that evaluation to give the reader necessary context.

We have given the validation summary and reference to Appendix A already in chapter 3.1. Thus, we decided to delete this sentence because we think it is superfluous.

L305-312: Please clarify the methodological basis of the three cited studies. E.g., observational or modelling.

We added information on the used methods; budget calculations based on observations and observed removal rates.

L310-311: Please cite the specific figure to support this statement.

We refer to Table 1 which show the numbers.

L330: Please specify the exact panel (e.g., Fig. 5b) and report the corresponding correlation coefficient and p-value.

We corrected the sentence: A significant correlation ($p < 0.001$) could be established for this relationship (Fig. 5b).

L337: Please specify the spatial resolution that is considered necessary (e.g., "on the order of 1 km or finer") to resolve the key coastal filter processes discussed in this study.

We added (on the order of 1~km or finer) to the sentence.

L357-359: The concluding statement that "For phosphorus, we did not find a similar dependence" on water residence time appears to be inconsistent with the visual data presented in Fig. 6b, which suggests a strong positive correlation. This claim cannot be definitively supported without a statistical test. Please clarify.

We appreciate the editor's careful review. As correctly indicated by the p-values in Fig. 6, phosphorus retention shows a significant relationship with residence time, and we have revised the text accordingly to reflect this finding.

Additionally, both nitrogen and phosphorus retention show significant dependence on water residence time in the lagoon system.