This manuscript used a high-resolution coastal model to study the retention capcity of the Oder Lagoon. The topic is interesting, however, the manuscript need a major revision before publication for the following major concerns.

First, as a model study, the manuscript lacks sufficient details regarding model description and needs clarification: 1) How are model parameters determined? Please provide justification or references for parameter values. 2) In oxygen-rich environment, phosphate becomes bound to iron oxide. How does the model simulate iron oxide? Are this process and phosphate release simulated prognostically or through parameterization? 3) In this model, the light attenuation is determined by chlorophyll and CDOM. Please explain how to simulate CDOM and how well is the simulation of CDOM? 4) For riverine input, only fresh water and nutrients were described. How about other state variables, including CDOM, each functional group of phytoplankton, and etc? Since the liminic phytoplankton thrive in fresh and turbid water, I assume that the fresh water is turbid with high chlorophyll or CDOM concentration. Then, the question is, how are the riverine inputs of chlorophyll (each groups of phytoplankton) and CDOM specified, by observations or some assumptions? 5) Does meterological forcing include data of nutrient deposition? How is the nutrient deposition simulated in the model? 6) The authors state in Line 104: “the extracellular excretion of dissolved organic matter by phytoplankton results in non-Redfield carbon uptake”. This sentence is unclear to me and needs clarification. Also, I would like to ask how does the model deal with nutrient stoichiometry? Is it fixed or variable in the model? 7) Some model components are missing in the model description, including the ice model, the two-layer sediment model, and the definition of residence time. The authors answered in their response letter that they had added description of the ice model in the revised manuscript, however, it remained absent. For the residence time, please give the equation used to calculate the residence time in the manuscript.

Second, the model validation is not sufficient enough to support their conclusions. 1) The model failed to reproduce the observed season cycle of surface chlorophyll (one peak versus two peaks). Since phytoplankton growth is the primary driver of organic matter deposition and therefore is a key in N/P retention, the authors should discuss how this model bias might affect their key conclusions. Attributing the model bias to the fixed carbon-to-chlorophyll ratio is insufficient. 2) The model significantly underestimated bottom oxygen concentration. However, the authors’ justifications are unsatisfactory:

(i) mismatch in depth between model and observations – the validation should use the closest model grid cells to measurement depths for comparisons

(ii) ship-induced mixing – The authors claim in appendix that the model is more realistic and ship-induced mixing as a shortcoming of observations. I disagree with this statement. While ship disturbances occure in reality, neglecting this process is the a weakness of the model, not shortcoming of observations.

Since the oxygen controls phosphate release and denitrification, the model's failure in simulating bottom oxygen cast doubt on conclusions of this manuscript. Please improve the model performance or provide discussions on whether this bias affect conclusions. 3) There is no validation of bottom nutrients, which is an important indicator of phosphate release and nutrient retention. 4) For model validation of stratification, Figure A7 is hard to read. Please provide some metrics (e.g. RMSE, R-square) between observations and modelled results.

Third, the current version of discussion section should go to results, and some discussions looks not highly relevant. The subsection 4.2 should be shortened. In addition, some key issues are not discussed in the current mansucript. For instance, (1) as I mentioned above, How might model bias in simulating surface chlorophyll and bottom oxygen affect the nutrient retention conclusions? What sensitivity analysis can be done to test the robustness of comclusions? (2) Differences in N- and P-retention response to nutrient loads deserve explanation, including but not limited to model bias in simulating bottom oxygen, the N:P ratio of riverine input (limited nutrients), and the inherent differences between N and P cycles. Answers to this question may help clarify whether this conclusion in Odor Lagoon can be applied to other coastal systems. (3) What is the implications for larger-scale models? In the introduction part, the authors mentioned that some baltic model accounted for the filter function of nutrients by assuming bioavaliability or reduction factors. The authors should compare their finds to previous empirical approaches, and discuss the validity of the previous assumptions.

Finally, some comments and concerns raised in the initial review were not well addressed and some revisions promised in the response round were not made in the revision round. These will be listed in the detailed comments. To facilitate the evaluation of revisions, I would also suggest the authors to 1) povide locations for each revision in their response letter and 2) include revised text in the response letter.

Detailed comments:

L87: please provide the number of model grid cells or the length/width of the lagoon. This will help readers who are not familiar with this region to better understand the model resolution.

L104: Why this lead to non-Redfield ratio of nutrient uptake?

L128: How to quantify/calculate mas transport?

Section 3.1: wrong reference to Fig1

L217-228: The N-retention rate (40%) in this study is much higher than previous studies (<30%). Why? The authors attribute this difference to the fact that previous studies only account for denitrification while neglecting other source/sink terms; however, this explanation appears insufficient. Based on Figure 6, the non-denitrification processes contribute minimally to the overall nitrogen budget.

Response to reviewer #1

Comment on L199-205: the promised discussion of hypoxia were not made in the revised manuscript.

Comment on L222: Regarding "what could be transfered to other regions", this was not well addressed. The authors present several assertions yet without providing any justification or supporting evidence in the manuscript. Please also see my major comments above. In addition, the promissed discussion on "which coastal type requires additional studies" were not made.

Response to reviewer #2

Comment 5: no description of ice model in revised manuscript

Comment 9: maybe I missed it, but I can't find it in the revised manuscript

Comment 10: please add this to manuscript

Comment 12: This didn't answer the reviewer's concerns.

Comment 18: No statistical evaluation in the manuscript. The reviewer was requesting a statistic test to examine differences between the control and halved-load experiments, rather than the trends from combined experiments.

Comment 19: see major comments above

Comment 33: The response is unclear

Comment 36: see major comments above

Comment 38: It confused me. The authors kept mentioning that changes in river discharge reflect changes in nutrient load because nutrient concentrations were stable, yet at the same time argued that this relationship did not apply for long-term analyses.