Dear Editor,

We greatly appreciate your continued consideration of our paper. We have gone to great lengths, with many changes and additions to the revised manuscript, to address all remaining comments, questions, and suggestions from the reviewer and yourself. We feel that the revised manuscript is now substantially stronger as a result. In the following, we have repeated your comments in bold and provided responses immediately after each comment.

**Thank you for your diligent revisions. After some consideration, I am returning this manuscript to you for additional revisions. The reviewers have identified some remaining concerns.**

**First, it is important that you address the twin-experiment idea of reviewer II. Either that is completed and the results shown, or you state very clearly why it is not logical to do this.**

Following the request from Reviewer 2, we have now added figures and a discussion to Appendix B outlining an additional twin simulation experiment with the most sensitive BGC and physical model parameters. We also now include results and discussion for a sensitivity analysis of the physical parameters. These results confirm that we can optimize these additional model components along with the BGC model parameters. This analysis has also provided insights into the relative impacts of BGC and physical parameters on model results and has shown that the physical parameters generally have the largest impact on the oxygen state variable.

**Second, please address the concern of reviewer I around the definition of "physical parameters". Both bottom boundary relaxation coefficients and sinking speed of organic matter are more commonly attributed to the biogeochemical parameters, so you may consider renaming "physical parameters" or providing additional justification.**

We agree that certain parameters are difficult to define as solely "physical" or "BGC" and we have generally addressed this comment by being more explicit in the text, as well as in Table 2, about the identities of the parameters. For example, in Table 2 we no longer refer to "physical parameters" and have instead provided subheadings within the table (i.e., advection components, boundary conditions, and turbulent mixing model) to more clearly identify the different parameter types. We also now note in the text after Eq. (3) that we consider the parameters associated with the first term in Eq. (3) to be the exclusively BGC parameters that would also be present in a 0D (i.e., time-only) implementation of the model. The last two terms in Eq. (3) represent advective and turbulent mixing processes, respectively. The parameters associated with these processes, along with the boundary condition parameters, are only introduced when the BGC model is coupled to the physical model (POM1D in this case). As such, these parameters are considered to be associated with the physical model, although their values may be specific to certain BGC species, particularly in the case of the settling velocity and the boundary conditions. We hope that these clarifications address the comments about the physical versus BGC terminology.

**Third, please discuss how transferable the results are to a 3D context (where there is no bottom boundary relaxation).**

This is an important point, and we now include in the Conclusions some commentary on the limitations of our results in the present study, as well as the possibility for extending the method to more complex 3D physical scenarios. In short, although the present models may not transfer well to 3D scenarios, the method itself is completely transferrable to higher dimensions. However, the overall computational cost of the parameter estimation will increase considerably due to the higher dimensionality of the simulations.

Sincerely, Skyler Kern on behalf of the co-authors.

**Response to Reviewer**

We greatly appreciate the time taken by the reviewer to read our manuscript. We have taken into consideration and addressed all comments, questions, and suggestions from the reviewer, and we feel that the revised manuscript is now substantially stronger as a result. Changes made to the text at the request of the reviewer have been highlighted in red in the revised manuscript. In the following, reviewer comments are repeated in bold, and our responses are provided after each comment.

**Major comments:**
**The authors have addressed most of my comments and I find the manuscript to be much improved. Thank you for that! Remaining issues refer mainly to the following points:**

**(a) It would be helpful to state more prominently and already in the abstract that the nutrient content and partly also the oxygen content of the water column (controlled by the bottom boundary relaxation coefficients) and sinking of organic matter are in this specific setup attributed to the physical and not the biogeochemical model component, because this is somewhat unexpected. The authors might consider to rename "physical parameters" to avoid confusion.**

We have updated the abstract to clarify that the nutrient and oxygen contents are strongly affected by non-BGC parameters. We have also updated many references of "physical parameters" to more clearly and explicitly identify the processes to which the parameters correspond. In particular, the parameters are now referred to as the advection, boundary condition, and turbulent diffusion parameters, as shown for example in Table 2. The relaxation parameters are a boundary condition within the physical model (POM1D), so it should now be clear what part of the optimization they are included in.

We also now note in the text after Eq. (3) that we consider the parameters associated with the first term in Eq. (3) to be the exclusively BGC parameters that would also be present in a 0D (i.e., time-only) implementation of the model. The last two terms in Eq. (3) represent advection and turbulent diffusion processes, respectively. The parameters associated with these processes, along with the boundary condition parameters, are only introduced when the BGC model is coupled to the physical model (POM1D, in this case). As such, these parameters are considered to be associated with the physical model, although their values may be specific to certain BGC species, particularly in the case of the settling velocity and the boundary conditions. We hope that these clarifications address the reviewer's comments about the physical versus BGC terminology.

**(b) I would find it nice to discuss in the Conclusions that the obtained parameter optimum might not be unique and that it cannot be guaranteed that it is global. It could also be indicated that model performances on unseen observational data, not used for fitting, have not been tested and it remains an open question how the models perform and compare under new circumstances.**

We agree with both these points and have added text to the Conclusions reiterating that the present optimization methodology does not guarantee a global minimum. We also explain that the performance of the model on test data is not examined in this paper, although this could be an interesting direction for future research.

**(c) I still did not find how the model was initialized apart from the information that the initial conditions were prescribed (sorry in case I overlooked something). It would be helpful to add this information somewhere prominently in Section 3 and also clarify how the second step in the sequential optimization was initialized. I regard this information as important because, e.g., initializing with the "truth" only once might give the sequentially optimized model more time to drift away from it. I think this is not what happens but still it would be good to clarify.**

Thank you for identifying this point of confusion. A brief statement was included in Section 4 of the previous revision, but we have now moved this content to a new subsection (Section 3.5, titled

"Model Setup") and expanded the description of the model to clearly address how the state variables are initialized.

**(d) It might be nice to add to the Conclusions that the presented results may not be directly transferable to 3-dimensional models where physical observations, such as temperature and salinity, would typically be used in addition to constrain the physical model parameters - while in the presented study the physical parameters are constrained based on biogeochemical observations only (monthly physical profiles are prescribed). Also, physical parameters of interest would typically differ in a 3D environment and there is typically no bottom boundary relaxation which seems to have a large impact here. This could nicely complement the already existing discussion around line 413ff.**

We thank the reviewer for identifying these additional important points. We have added text to the Conclusions emphasizing that the present models may not transfer well to 3D scenarios, but that the methodology itself can be applied in 3D contexts, where the overall computational cost will increase considerably due to the higher dimensionality of the simulations.

**Specific comments:**

**Ln 10/11: It would be helpful to add that nutrient inventory and partly also the oxygen inventory of the water column (controlled by the bottom boundary relaxation coefficients) and sinking of organic matter are tuning parameters attributed to the physical and not the biogeochemical model component.**

Throughout the paper, we have replaced the use of the generic phrase "physical parameters" with more specific references to the processes to which each of the model components correspond. As part of this revision, the indicated sentence has been updated as suggested by the reviewer.

**Ln 12/13: This sentence does not become clear to me – I understood that the optimization starts with a global search?**

Indeed, the algorithm does use a truncated global search to find starting locations for local optimization runs whose results are compared to determine the final parameter set. We thank the reviewer for pointing out the ambiguity introduced by this sentence, and we have updated it to more clearly represent both steps in our approach to performing parameter estimations in high-dimensional spaces.

**Ln 26: It might be clearer to reformulate – it is neither obvious if and how simultaneous optimization can solve the above-mentioned issues nor if and why the optimized model "generalizes" better. I am fine with a statement that this is the aim and that the presented study might serve as a first step.**

We appreciate this suggestion and the referenced line has been updated to more clearly indicate this is an exploration of an approach to addressing the issue, instead of suggesting we have completely solved the indicated problem.

**Ln 40: It would be helpful to add here as well that bottom boundary relaxation coefficients and sinking of organic matter are tuning parameters attributed to the physical and not the biogeochemical model component and to explain the implications of these parameters for the water column budgets.**

Following the reviewer's suggestions, generic references to "physical model parameters" have been removed wherever possible and updated to more explicitly indicate the model components. We now note that the advective, boundary condition, and turbulent diffusion parameters are those associated with the physical model (POM1D).

**Ln 83: As DAKOTA is strongly highlighted, it might be nice to summarize all the information somewhere and add a general overview, alongside with information how to get access to the code (a respective subsection or appendix?). Here I would suggest to be a bit more specific on the methods rather than DAKOTA.**

In Section 2, we have attempted to provide information on the specific algorithmic choices made in our implementation of DAKOTA, and we now additionally direct the reader to the DAKOTA website (https://dakota.sandia.gov) where the code can be downloaded, as well as added a link for the reader to download our DAKOTA test case (https://doi.org/10.5281/zenodo.16700702). Further details are also provided in Kern et al. (2024), which we now point out to the reader. We hope that this additional information is helpful for any reader interested in learning more about DAKOTA.

**Ln 162: Why were the twin experiments limited to the biogeochemistry? It might be interesting if all strategies lead to the original parameter set or if this is not possible (a nice discussion point).**

We have now expanded the twin simulation experiments, as well as the sensitivity analysis, to also include physical model parameters, and the results of these additional tests have been added to Appendix B. These results have provided additional interesting insights, including that the parameters associated with vertical advection by larger eddies are generally less significant than other variables, but that these parameters also have the greatest impact on the oxygen state variable. A discussion of these results has also been added to the appendix.

**Ln 267: I find the order confusing. It would help to indicate here that "sequential" means that first physical and then biogeochemical parameters have been optimized. It might also be good to mention the rationale behind this order (as I understood, the nutrient budget is set by the so-called physical parameters before optimizing the biogeochemistry).**

We agree that the order may have been confusing, because BGC was listed first despite not being first in the sequence. We were also still using the "physical parameter" phrasing instead of the current more explicit verbiage. We have updated this line and others to be more explicit when describing the parameters, which should reduce confusion.

**Ln 305: Better "tendency" instead of "trend"?**

We agree, and this word has been updated as suggested by the reviewer.

**Ln 310/319: It would be particularly interesting to add respective information on estimated sinking speeds to this discussion (same in the conclusion).**

We agree that this could be an interesting direction for future research and have added a comment in the Conclusions to this effect. We also now note near the end of Section 4.2 that the settling velocity at HOTS is higher than at BATS. At present, the settling velocity we include in the model only applies to the detritus and is a constant value (i.e., it has no temporal or spatial variation and is not applied to other state variables).

**It seems to me that at HOTs simulated oxygen gets much worse by the biogeochemical parameter optimization (after the "physics" has been optimized), while the representation of PON strongly improves (Table 3 and Fig.). For BATs this effect less pronounced. I find this very interesting. Could the authors comment on that and elaborate on the potential reasons? Are there other parameters not included in the optimization which might change this result? Could the worsening of oxygen be prevented if sinking speed and bottom boundary relaxation coefficients were considered biogeochemical parameters? Understanding the mechanisms might help to understand how versatile the obtained results are and might raise interest of a wider audience.**

We thank the reviewer for pointing out this interesting result and for suggesting the need for more analysis and discussion. We agree that, in the context of the present estimation approach, the result makes sense because, as the reviewer points out, the oxygen results worsen while the PON results significantly improve, resulting in an overall lower model error. We now note in the text on page 18 that the worsening of the oxygen results could be ameliorated by including the sinking speed and boundary relaxation coefficients with the BGC parameters, but that the overall improvement in the sequential optimization approach would be unchanged.

**Ln 377: "trend"?**

We have now clarified the terminology in this sentence by using the phrase "annual cycle" instead of "trend".

**Ln 381: How would you treat locations with no or only few observations?**

This is an important point, and we have added text in the Conclusions noting that in situations where there are only a few observations, the present methodology could still be used. However, care would have to be taken in applying the resulting models to new scenarios, since the parameter estimation would have been based on only limited information. That is, the parameters could become over-tuned to give good agreement with the available data, at the expense of broader agreement with more observational data.

**Ln 385 Conclusions: Again, it would help to clarify what is considered as "physical parameters".**

We agree and have replaced references to "physical parameters" throughout the revised paper with more explicit references to the processes associated with each parameter.

**Ln 392: Please add "in a 1D environment"**

The sentence has been updated as suggested.

**Ln 402: It might be helpful to add that this is mainly due to the simulated amount of dissolved oxygen which worsens after the physical optimization potentially due to PON-improvements. How does this result depend on the choices of parameters to optimize?**

We have added a comment to the Conclusions noting, as the reviewer has helpfully pointed out, that the dissolved oxygen results worsen, resulting in the higher overall objective function value in the sequential case. With the substantial additional discussion of this result on page 18, we hope that the reviewer's helpful comments on this result have now been addressed.

**Ln 417: cf. ln 12/13 above**

The sentence has been updated to make it clear that the methodology employs both a global search and local optimizations.

**Ln 420: To advertise DAKOTA the code and a directly applicable test case would be very nice.**

We thank the reviewer for this suggestion and have added text in Section 2 directing the reader to the DAKOTA website (https://dakota.sandia.gov) where the code can be downloaded. We have also added a link to download our DAKOTA test case (https://doi.org/10.5281/zenodo.16700702). Further details are also provided in Kern et al. (2024), which we now point out to the reader.

**Ln 424: I am not 100% sure what is meant here.**

We updated the text on this line to more clearly state that the method should be able to identify parameter sets that produce better results relative to the other probed locations. Those best performing parameter sets are assumed to correspond to regions good for us to target and further test for identifying a minimum. We therefore use those parameter sets as the initial guess for local optimizations. The new wording should more clearly summarize this process, which is more fully explained earlier in the paper.

**Figure 2: Could you indicate in the figure how many parameters were optimized per arrow and add the abbreviations from Tab.1? It might be nice to have a similar sketch also on the physical model.**

We thank the reviewer for this suggestion, and the figure has been updated accordingly. We attempted a similar figure for the physical model, but it is substantially less complicated than the BGC model in Figure 2 and we hope that the description in the text is sufficient.

**Code availability: A link to the mentioned Zenodo-directory would be nice.**

We have provided Zenodo links to the data sets (https://doi.org/10.5281/zenodo.16696156), the modeling code (https://doi.org/10.5281/zenodo.16702246), and the DAKOTA test case (https://doi.org/10.5281/zenodo.16700702) with our initial paper submission; these links can be found under the "Assets" tab on the Preprint page for our paper (https://doi.org/10.5194/egusphere-2025-3795). We also provide these links in the paper text under "Code and data availability" on page 29.