Rebuttal “Brief communication: Intercomparison study reveals pathways for improving the representation of sea-ice biogeochemistry in models” by Tedesco et al., 2025

Reviewer #2

General comments

This study provides an intercomparison of six 1D sea-ice biogeochemical models with a focus on the assessment of simulating spring ice algae blooms and associated nutrient variability. The main findings are that: none of the models adequately captured blooms with their default parameters; tuning improved the ice algae blooms but not the nutrient variability; and more systematic tuning strategies are suggested as a next step. I think it is a great effort to conduct an intercomparison study for sea-ice BGC models, which has not been done except for Watanabe et al. (2019). The manuscript is generally easy to follow and clearly structured. However, it would benefit from more careful proofreading to address minor editorial issues and improve overall readability.

We thank Reviewer #2 for their positive assessment of our work and for recognising the value of this sea-ice biogeochemical model intercomparison. We also appreciate the comment regarding editorial quality. In response, we will carefully proofread the revised manuscript to correct minor editorial issues and improve clarity and readability throughout.

The manuscript type is “brief communication”, so I understand that it is written briefly. However, I find it a bit too brief considering the following three points. Therefore, I recommend major revisions and provide suggestions below.

We thank the reviewer for their constructive suggestions. While we chose the Brief Communication format to highlight the core outcomes of the intercomparison in a concise manner, we agree that additional detail would enhance the clarity and utility of the study. In the revised manuscript, we will address the points raised below by expanding key sections, while keeping within the scope and length limitations of the format. Where necessary, we will provide supplementary material to ensure transparency and completeness.

Physical data. The manuscript lacks the presentation of physical data, even though the text mentions the existence of such data (e.g., L21, L137, L143). Given that physical processes drive the circulation of biogeochemical variables, it seems essential to show the comparison of physical model and observational data, such as snow thickness, ice thickness, and sea surface temperature. With these additions, the study may be able to address (or at least speculate) whether the simulated differences and biases are due to the physical processes.

We thank the reviewer for their suggestion. We fully agree that including physical data is valuable for interpreting the biogeochemical model performance. In the revised manuscript, we will add a figure comparing observed and simulated physical variables for the models that include their own physical components. We will also expand the Discussion to reflect on how differences in physical conditions may have contributed to the observed model biases and variability.

Quantitative assessment. Table 2 can be improved by incorporating quantitative findings. Currently, it is a qualitative description that is not very informative and is a bit difficult to follow; one can easily guess the qualitative changes as described in Table 2 (e.g., lower biomass was increased by lowering silica limitation). What would be informative and advance the knowledge is to report the amount of improvements by the amount of parameter adjustments.

We thank the reviewer for their comment. In the revised manuscript, we will update Table 2 to include specific parameter values before and after tuning (where available), as well as the corresponding changes in key model outputs (e.g., peak chlorophyll-a concentration, bloom timing, nutrient drawdown). We agree that this will allow readers to better assess the magnitude of improvements achieved through tuning and how these relate to parameter adjustments.

Connection to previous studies. The results and discussion section as well as the conclusions section (L162 onwards) do not appear to contain any reference to previous studies. Hence, it is unclear how this study contributes to the field. This can be achieved by incorporating discussion of the results with previous studies. Specifically, I think that the discussion can be improved by incorporating tuning strategies and intercomparison studies conducted for ocean BGC modelling (e.g., Schartau et al., 2017). Some of these are already mentioned in the manuscript (e.g., L226-234), but it would be better to link these with relevant previous studies to provide a practical direction for future studies.

Schartau et al. (2017). Reviews and syntheses: parameter identification in marine planktonic ecosystem modelling. Biogeosciences.

We thank the reviewer for their comment. We agree that linking our results more clearly to previous biogeochemical modelling and intercomparison studies would strengthen the context and relevance of our findings. In the revised manuscript, we will expand these parts to explicitly reference relevant works such as Schartau et al. (2017). At the same time, we will emphasise that, to our knowledge, this is the first intercomparison specifically focused on one-dimensional sea-ice biogeochemical models. As such, our study fills a gap in the literature and offers a novel perspective on tuning approaches, model diversity, and shared challenges within the sea-ice modelling community.

Specific comments

L21. “N-ICE2015” is too technical for the abstract. It is better to inform the region and season instead (e.g., north of Svalbard during April-June, 2015).

We thank the reviewer for their comment. We will revise the manuscript accordingly.

L22. “tuning” and “adjustments” are used together and they seem to mean the same thing, but this is unclear. I suggest replacing “without tuning, adjustments” by “using their default parameter sets, tuning”.

We thank the reviewer for their suggestion. We will revise the manuscript accordingly.

L23. It would be good to add a sentence here to explain why “adjustments improved biomass simulations but had a limited impact on nutrient representation”. (at least speculate even though the cause is unknown)

We thank the reviewer for their comment. The limited improvement in nutrient representation compared to biomass is primarily because most model groups prioritised fitting their simulations to the Chl-a observations during the tuning phase, as these data were more temporally resolved and directly linked to the main focus of the study, i.e., the ice algal bloom. In contrast, nutrient observations were limited to a single time point, which made them more difficult to constrain reliably. We will clarify this point in the revised manuscript.

L24. It may be informative to add a few words to describe what “harmonised” means here.

We thank the reviewer for their suggestion. We refer to the development of more coordinated or standardised tuning approaches across models, for example using common performance metrics or agreed-upon parameter bounds. We will revise the sentence to reflect this more explicitly in the following way:

"Variability in tuning strategies underscores key knowledge gaps and the need for further model development using more coordinated approaches such as common evaluation criteria or shared parameter ranges."

L28. Should “ice algae” be “bottom ice” instead, given the following phrase “representing the largest biomass fraction in sea ice”?

We thank the reviewer for this comment. However, we respectfully maintain the use of “ice algae” in this sentence. The term refers to the community of microalgae that inhabit the sea ice and is widely used in the literature to describe the biological component responsible for the largest biomass fraction in sea ice (e.g., Poulin et al., 2011). In contrast, “bottom ice” refers to a physical ice layer and not the biological community itself. We will retain the original wording for clarity and consistency with established terminology.

L47. “IAMIP1” should be spelled out.

We thank the reviewer for their comment. In the revised manuscript, we will spell out “IAMIP1” upon first mention as *“Ice Algae Model Intercomparison Project – Phase 1 (IAMIP1)”* to improve clarity for readers unfamiliar with the acronym.

L53. “CMIP6” should be spelled out.

We thank the reviewer for their comment. We will spell out “CMIP6” upon first mention in the revised manuscript as *“Coupled Model Intercomparison Project Phase 6 (CMIP6)”* to ensure clarity for all readers.

L67. I suggest replacing “existing” by “participating”, as the former sounds like these are all 1D models that exist.

We thank the reviewer for this suggestion. In the revised manuscript, we will replace “existing” with “participating” to clarify that the six models represent those that were available and contributed to this specific experiment.

L79. “little” or none? Horizontal advection terms are neglected.

We thank the reviewer for this comment. We agree and will include a clarification in the revised manuscript that 1D process models are typically designed to represent vertical processes only, under the assumption that horizontal advection is negligible.

L88. It would be helpful to briefly explain what “dynamic layering” means.

We thank the reviewer for this helpful suggestion. In the revised manuscript, we will briefly explain what is meant by *“dynamic layering”* in this context. Specifically, it refers to the model’s ability to adjust the thickness of vertical layers within the sea ice in response to growth and melt processes, thereby allowing for a more realistic representation of habitat structure and biogeochemical gradients.

L92. “Chemical Functional Families (CFF)” does not sound familiar in marine BGC modelling. Please use an alternative term or provide a reference.

We thank the reviewer for their comment. To improve clarity, we will revise the manuscript to use *“Plankton Functional Types (PFTs)”* instead, which more accurately describes the grouping of organisms based on shared functional traits relevant to biogeochemical cycling.

L124. Please correct the latitudinal range “83 to 83 N”.

We thank the reviewer for spotting this typo. We will correct the latitudinal range for *“83 to 80°N”* in the revised manuscript.

L126. It is more intuitive to write the range in an increasing order “80.5 and 81.8 N”.

We thank the reviewer for their suggestion. To balance clarity and scientific accuracy, we will revise the sentence to explicitly describe the southward drift from 81.8°N to 80.5°N, making both the direction and latitudinal range intuitive for readers: *“Among the four ice floes monitored during the study period, the refrozen lead data were derived from Floe 3, which was studied from mid-April to early June 2015 as it drifted southward from 81.8°N to 80.5°N.”*

L139. “Duarte et al. 2017” is one of the participating models in this study? If so, why is the performance poor? Presumably, the model was previously tuned to this study site.

We thank the reviewer for this observation. The model described in Duarte et al. (2017) is indeed one of the participating models in this study. However, it is important to note that the configuration used in our intercomparison did not retain the site-specific tuning applied in the original publication. Instead, all models were initially run using their respective baseline parameterisations, which were designed for broader applicability rather than tailored to the N-ICE2015 lead environment. The poorer performance in the default run reflects this generality and highlights the challenge of transferring model setups across sites without retuning. This underlines the importance of the harmonised tuning phase included in our intercomparison design.

L143. I do not see these physical metrics compared (sea-ice season timing, ice thickness, and snow thickness).

We thank the reviewer for their comment. As already proposed, in the revised manuscript, we will include a comparison of key physical metrics between the N-ICE2015 observations and the models that simulate their own physics. This will be presented in a new figure and discussed in both the Results and Discussion sections to help interpret the influence of physical conditions on biogeochemical model performance.

L149. “the extent of biases” will depend on the location where tuning was conducted for the default parameter sets. Hence, it would be helpful to indicate for which region each model was tuned (in Table 1 and/or the text). This will also give an indication for the “portability” of the model (Sec 2.8 of Friedrichs et al. 2007).

Friedrichs et al. (2007). Assessment of skill and portability in regional marine biogeochemical models: Role of multiple planktonic groups, JGR-Oceans.

We thank the reviewer for their suggestion. We agree that the extent of model biases in the default runs can be influenced by the region for which each model was originally tuned, and that this information is important for evaluating model portability. In the revised manuscript, we will add a column to Table 1 indicating the geographic region or study site associated with the original tuning of each model’s default parameter set.

L158. Please describe the source of the atmospheric forcing used here.

We thank the reviewer for their comment. In the revised manuscript, we will add a description of the atmospheric forcing used in the experiment. Specifically, the models were forced with atmospheric data collected directly during the N-ICE expedition. This information will be included in the Methods section for clarity and completeness.

L159. Please specify which model is the one without a thermodynamic component.

We thank the reviewer for their suggestion. In the revised manuscript, we will specify that the model without a thermodynamic component is SIMBA. Unlike the other models, it relies on prescribed physical fields rather than simulating them dynamically. This distinction will be clearly stated in the Methods section and reflected in Table 1.

L180. “Change in the initial simulation date” seems strange to be considered a tuning parameter.

We thank the reviewer for this comment. We agree that a change in the initial simulation date is not a tuning parameter in the conventional sense. We will revise the text to clarify that this change was a modelling decision made to improve alignment with observational context, rather than a formal tuning of model parameters.