Reply on RC3
James R. Christian et al.

Author comment on "Ocean biogeochemistry in the Canadian Earth System Model version 5.0.3: CanESM5 and CanESM5-CanOE" by James R. Christian et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-327-AC3, 2022

Reviewer# 3 (Anh Pham)

Summary
The submitted manuscript by Christian and colleagues presents some major developments of the marine biogeochemistry component of the Canadian Earth System Model(s) v.5, focusing on representing a prognostic iron cycle and denitrification and including flexible phytoplankton elemental ratios and interactions between multiple food chains. These improvements are described in details and results of the Canadian Earth System model version (CanESM5-CanOE), which includes this newly improved marine biogeochemistry component, are presented and compared with results from two other CanESM versions (CanESM5-CMOC and CanESM2). While CanESM5-CMOC differs from CanESM5-CanOE in its ocean biogeochemistry component, CanESM2 is the older CanESM version, having different ocean circulation. The results show that CanESM5 versions are much better than CanESM2 when compared with available observations thanks to improvement in ocean circulation. The improvements in performance of CanESM5-CanOE over CanESM5-CMOC are not as clear due to sparsity in observations and uncertainties in historical trends. However, the inclusion of prognostic schemes for ocean Fe cycling and denitrification would be more suitable to address climate change problems.

Assessment
In general, I think that this manuscript is suitable for publication in Geoscientific Model Development, serving as a documentation on the development of an important model member of Earth System Models participating in CMIP. However, I do have some comments and suggestions, which hopefully can improve the quality of the manuscript.

First, while I understand that the main purpose of this manuscript is to describe recent developments in the ocean biogeochemistry component of the CanESM and to compare performance of its different versions, having more explanations as to why there are improvements of CanESM5-CanOE over CanESM5-CMOC in some areas but not all would be helpful. In addition, given that the comparison is performed also with CanESM2, which uses different ocean circulation, I would expect more discussions on which improvements of CanESM5 over CanESM2 are due to physics and which are due to biogeochemistry.
Second, I find the naming convention throughout the manuscript is somewhat confusing since there are three model versions involved in the comparison, of which two are under the CanESM5 umbrella. Sometimes it is difficult to figure out which model version of the CanESM5 that the authors are referring to. In some places, the authors explicitly wrote CanESM5-CanOE and CanESM5- CMOC, but in others, they wrote only CanESM5 or just CanOE and CMOC. It would be better if the authors could keep the naming consistent throughout the manuscript.

We thank the reviewer for a thorough and constructive review. His concerns in the first two points overlap with those of the other reviewers, and we have addressed them in the revised MS.

Third, since the model developments focus on Fe and N cycles, I was thinking that the authors should do a more comprehensive comparison of the modeled Fe distribution with observations, taking advantage of the growing GEOTRACES data. I understand that there is no climatological Fe dataset yet, but comparison with observed Fe transects from GEOTRACES should give an indication of the model performance on ocean Fe cycling.

We have done some of this analysis but we do not believe that it adds very much to the paper in terms of process understanding. i.e., what biogeochemical processes are missing from or simplistically represented in the model? The longest available transect is GA02 in the Atlantic, and the model-data agreement along this transect is generally good, with the usual caveats about the temporal mismatch (one-time snapshot of obs vs long-term average of model). GP02 in the North Pacific shows clearly that the model does not reproduce the mid-depth maxima associated with the North Pacific oxygen minimum zone, which is already discussed in the paper and demonstrated by our other analyses. We are exploring other ways of presenting these data, and will add new plots to the Supplemental if we think the information is useful to readers, i.e., if it tells us something about model biases or model representation of specific processes that is not apparent from the existing analyses.

Fourth, while export production is an important biogeochemistry feature, using it as a metric to evaluate model performance is difficult because of the uncertainty in the observational estimates, as the authors already pointed out. Primary production/chlorophyll might be a better metric.

This issue was also raised by more than one reviewer, and the point is important. However, we believe our choice of plots and metrics is correct. Export production is not included mainly for purposes of model validation but rather, like CO2 uptake, it is included so that readers can see how our models compare to other CMIP6 models on several global metrics that are commonly used and of broad interest.

We include several observation-based metrics of phytoplankton biomass (e.g., Figures 16-18) and present them in a way that we believe helps the reader understand the important differences in the way our two biology models are formulated. Aggregate export production is important for global ocean biogeochemistry and ocean CO2 uptake; primary production is important for impacts on higher trophic levels but the same atoms can cycle faster or slower in the surface layer without any net uptake of CO2. Global spatial distribution of chlorophyll or primary production does not provide a very strong constraint on model performance due to the very strong enhancement in coastal regions that is unresolved by coarse resolution global models.

Finally, since the historical trends section forms an important part of the manuscript, I
would suggest the authors give more details on how the historical model runs are performed (i.e., which CO2 and atmospheric forcings are used...), how the results are analyzed, and why analyzing and comparing model historical trends is important.

In accordance with the comments of this and other reviewers, we have provided a more detailed explanation for the inclusion of the historical trends section. While these are standard CMIP6 experiments, we have expanded the description of the experimental setup slightly as per the reviewer's suggestion, to make sure that there is no confusion or ambiguity.

Some specific comments:

Line 27: some areas? Which areas? Please be more specific if possible.

**This is clarified in the revised MS.**

Line 30-32: Which CanESM version that shows these results?

**This is clarified in the revised MS.**

Line 127-128: Do you mean CanESM5 uses the same carbon chemistry as CanESM2?

**No. Carbon chemistry was slightly different in CanESM2, as the code was written before the current standard protocols were defined. This is clarified in the revised MS.**

Line 500: Change can not to cannot.

**Both of these are valid English. Possibly it is a difference between US and UK/Commonwealth English.**

Line 608-609: Which model version are you referring to here?

**This is clarified in the revised MS.**

Line 648-650: it might be worth to mention the difference between CanESM2 and CanESM5 in the nitrate initialization field earlier in the text. Introduction or section 2, for example.

**done**