

In the following, referee comments are in black, while our responses are in green and added material is indicated in blue.

Reviewer 1

Overall comments

The manuscript “Ocean biogeochemistry in the coupled ocean-sea ice-biogeochemistry model FESOM2.1-REcoM3” is a comprehensive analysis of a new iteration (version 2.1) of the FESOM global ocean model, coupled with the biogeochemical model REcoM3 (also an updated version). The authors set out to evaluate the model for a recent time period (1958-2021) to assess if the model can be used to evaluate climate change effects and CO₂ increases on a century timescale.

Overall, I think the authors have presented a clear and thorough manuscript that was generally well-written and easy to read. The authors presented model results that showed relatively good agreement with observations, and illustrated that the new iteration of the FESOM and REcoM models had some improvements for the model's overall performance. I believe this manuscript will be a valuable contribution. I just have some relatively minor comments, which are detailed below.

We would like to thank the referee for the overall positive assessment of our manuscript .

Specific comments

1. I thought in general the authors had a really nice and detailed Methods section. For example, the description of the FESOM model grid was very clear and I appreciated the section explicitly listing changes to REcoM. However, I found there were a few details missing from the Methods. See specific lines below:

Methods: Forcing (line 200ff): I'm curious about what river forcings you used, if any? Perhaps I missed it but it doesn't seem to be detailed in your methods. I think this could be particularly important to note since you comment at line 564-565 that future work is looking at the role of rivers for carbon and nutrient transport.

We have added this information to the text:

The freshwater supplied by rivers is a climatology and provided by Large and Yeager (2004) as part of the CORE forcing. Nutrient, carbon and alkalinity supply via river discharge is not included in the experiments described here.

Line 208ff: I think a few more details should be specified about the initialization fields; in particular, I think it would be worthwhile to explicitly state the years that the various climatologies span. Additionally, you point to GLODAP for Alkalinity and DIC initialization but don't describe what years from GLODAP were used?

We slightly changed sentences to include the time period as follows:

Initial fields for temperature and salinity were taken from the Winter statistical fields of Polar Science Center Hydrographic Climatology (PHC3, updated from Steele et al., 2001) that ingests observations from the period 1900-1994. Total alkalinity (Alk) and preindustrial dissolved inorganic carbon (DIC) were initialized from version 2 of the Global Ocean Data Analysis Project (GLODAPv2) climatology (Lauvset et al., 2016) based on data collected between 1972 and 2013. Dissolved inorganic nitrogen (DIN) and dissolved silicic acid (DSi) were started with values from the Levitus World Ocean Atlas climatology of 2013 (Garcia et al., 2014) occupied between 1955 and 2012. We used the Levitus

World Ocean Atlas climatology of 2018 for dissolved oxygen (Garcia et al., 2019a) (See Table 2) [based on data for the time span 1955-2017](#).

Line 246: I found it interesting that you only showed modelled mean fields for 2012-2021, given the long spin-up period as well as the authors stating in other locations (e.g. the abstract) that they are analyzing the period of 1958-2021. Why was this specific period chosen?

The reviewer is right that we did not justify our choice. We averaged the model results over the period of 2012–2021 to get rid of interannual variability in order to evaluate the mean state in a recent period. Additionally, we used the last decade because we wanted to keep the presented results comparable to the time span where most pCO₂ observations were taken. We kept this time period to evaluate other variables for consistency throughout the manuscript.

2. I found section 3.3.3 DIC Inventory Changes a bit hard to follow (lines 522ff). I think part of it is just some awkward wording and potentially a need for it to be a bit more fleshed out (without adding too much more text!). Additionally, the authors state that “FESOM-2.1-REcoM3 is thus one of the few ocean biogeochemistry models that falls within the range of interior ocean anthropogenic carbon accumulation”, which is again emphasized in the Conclusions. I think this should be discussed a bit more here. For example, if the numbers are readily available, I wonder if adding in estimates of C inventory from other models into Table 4 will help illustrate this point and facilitate a bit more discussion.

Thank you for pointing this out. We have reworked the text with the aim to make it more digestible. We would, however, not like to discuss the different carbon components (anthropogenic vs natural, steady state vs non-steady state) here in detail and have added references for this so that the interested reader can follow up.

We have also added the numbers for the anthropogenic DIC inventory change 1994-2007 for the other models to the text. We decided not to add them to the table, because it's not our data and it wouldn't be appropriate to give these numbers for the first time in our model description paper. We will make sure that these numbers will be published in the appendix of the GCB 2023 paper to be accessible to the scientific community.

In regards to Table 4, perhaps it's personal preference, but I think the units should be in the column header as well as in the table caption? I also think it's a bit confusing to report the total inventory in 1994 vs the increase in C inventory for the other periods under the same columns/without any delineation.

We restructured the table to distinguish the total inventory from the differences and introduced subheadings in the table. We also added the units to the column headers.

3. There are a few sentences that require more references/details about references. See specifics below:

Line 270: What other works? Maybe also include specific numbers of some sort from other studies?

References are added.

Line 277: Please reference what “known other works” you are referring to

References are added and the text is revisited.

Figure 19: You reference a “Globally integrated annual air-sea CO₂ flux from Global Ocean Biogeochemistry Models” but don’t state what these models are. My guess is that the GOBMs used are from the Global Carbon Budget 2022 (?) but as written it’s not clear and I think you’re missing the appropriate reference to 2022 Carbon Budget here. I also think that it could be useful to explicitly state in the figure caption the number of GOBMs used (10 I believe?) and where in the 2022 Carbon Budget the readers can find information about these models (i.e. reference what table they’re listed in). Potentially more details about the pCO₂-based data products could also be briefly given.

The number of GOBMs (10) and pCO₂-products (7) is now mentioned in the caption together with the reference to GCB 2022 and all references for the individual models and pCO₂-products. We have also amended the discussion on the differences between GOBMs and pCO₂-products in the text around Figure 19, referring to model biases in ventilation and carbonate chemistry as well as the uncertainties in the pCO₂-products.

Other minor comments/suggestions

Line 545: Is this a spatially averaged oxygen concentration or for a specific region? What were the values presented in Cocco et al. 2013 that you’re comparing to?

- Line 23: I’m a bit confused here... since a rate of 2.9 PgC per year was given but then the authors state it’s in 2021. It’s particularly confusing since the previous sentences refer to a 250 year trend in increasing atmospheric CO₂

We have added a sentence on the longer term uptake to link better to the previous sentence, and have then given the numbers for the recent decade 2012-2021. The unit PgC/yr is then correct for the average flux in this period.

- Line 23-24 (about terrestrial/land fluxes) is also a bit confusing

We find it important to put the ocean carbon sink in perspective not only to the fossil fuel emissions, but also to the second large carbon sink: the terrestrial biosphere. We have reformulated to link the sentences better.

- Line 55ff: I think I would suggest moving this paragraph up to right before introducing REcoM and move the introduction of REcoM down to where you introduce FESOM. That way, you’ve then fully described the problem you’re trying to solve and followed by an explanation of how you’re solving it.

Moved as suggested.

- Line 260: I’m curious if you have any theories as to why FESOM is too saline at the surface compared to climatology that could be commented on in the text?

Indeed, we speculate that the bias is caused by the imperfections in the river discharge used in the simulations. We prescribed the river runoff as the annual mean runoff from Large and Yeager (2004), which was a standard practice in CORE-type simulations. However, when combined with the relatively low surface salinity restoring, using a piston velocity of 50m/300 days, the SSS bias drifts to the values depicted in Figure 3f.

- Line 305: Any idea why the moderately high silicic acid values in the northern high latitudes is not reproduced well?

Thank you, this is a good question. We cannot say with certainty, but it is likely linked to the winter deep mixing that is too shallow. It could also be overly strong silicic acid draw-down by diatoms, which in turn could be linked to parameter choices or iron limitation that may be too weak. We have added text accordingly.

- Line 425: I'm not sure what is meant by "A more detailed description of zooplankton can increase NPP by 25%" or even how it fits in with the surrounding sentences

We have amended the text and added reference to section 3.2.2 where this is explained in some more detail. The text now reads: "A more detailed description of zooplankton results in more efficient nutrient recycling and can thus increase NPP by 25% (see also explanation in section 3.2.2, Karakus et al., 2022)."

- Figure 13: I'm confused about what the difference between the orange dots and brown line are for MAREDAT? This should be more clearly defined in the figure caption.

We have added the information to the figure caption: "orange dots for individual observations and solid brown line for the zonal mean of the observations"

- Figure 16: Panel A-B – could be nice to denote on the colourbar (or in the figure caption) what values (positive or negative) are sources/sinks of CO₂ for those unfamiliar

We have added a sentence to clarify that negative indicates a flux into the ocean.

- Line 480: Perhaps outside the scope of this study but comparison to more regional studies could be interesting. For example, the North Atlantic shelves are shown to act as sinks of CO₂ (which is consistent with other global studies) but some regional studies have shown that the Scotian Shelf, for example, acts as a source of CO₂ (see Shadwick & Thomas 2014: <https://doi.org/10.1016/j.marchem.2014.01.009>; Rutherford et al. 2021: <https://doi.org/10.5194/bg-18-6271-2021>)

Thank you for the comment. Indeed, this is interesting, but we consider this out of scope here. If discussing specifically the Scotian shelf, then we would need to discuss other regions at the same level of detail. However, we have added a sentence on the patterns in the global coastal ocean in comparison to a recent synthesis based on low- and high-resolution models and pCO₂-products (Resplandy et al.).

FESOM2.1-REcoM also generally captures the large-scale patterns of coastal CO₂ fluxes with CO₂ uptake in the mid- and high-latitudes (poleward of 25°N/S) and outgassing in the tropical coastal ocean, as described in a recent synthesis based on low- and high-resolution models and pCO₂-products (Resplandy et al., 2023).

- Lines 502ff: I think this is a super interesting result! One would definitely assume that there is a similar bias in the control and historical simulations...

Thank you! Yes, it is puzzling, isn't it?

- Line 514: What years of the model simulation are these flux estimates calculated over?

We have simplified the sentence so that this information is more easily spotted. It now reads: “After accounting for the bias in simulation B, the simulated ocean carbon sink (1990-1999) is 1.74 ± 0.11 PgC yr⁻¹ and 2.17 ± 0.13 PgC yr⁻¹ for FESOM1.4-REcoM2 and FESOM2.1-REcoM3, respectively.”

- Line 516: is this an observationally based estimate?

We refer the reviewer to the summary In Friedlingstein et al 2022: “This is based on indirect observations with seven different methodologies and their uncertainties and further use of the three of these methods that are deemed most reliable for the assessment of this quantity (Denman et al., 2007; Ciais et al., 2013). The observation-based estimates use the ocean–land CO₂ sink partitioning from observed atmospheric CO₂ and O₂/N₂ concentration trends (Manning and Keeling, 2006; Keeling and Manning, 2014), an oceanic inversion method constrained by ocean biogeochemistry data (Mikaloff Fletcher et al., 2006), and a method based on penetration timescale for chlorofluorocarbons (McNeil, 2003). The IPCC estimate of 2.2 GtC yr⁻¹ for the 1990s is consistent with a range of methods (Wanninkhof et al., 2013).”

We refrain from repeating this full information here, and just add “based on seven different methodologies”.

Minor suggested wording changes

- Line 4: I think this should be slightly changed to: “Marine biogeochemical models are a useful tool but, as any model, are a simplification and need to be continually improved.”

Rephrased as suggested.

- Line 13; “Dissolved oxygen is **also** added as a new tracer.”

Amended as suggested.

- Line 14: remove comma after 1958-2021

Amended as suggested.

- Line 21: sentence is a bit awkward

Revisited.

- Line 55: I think remove the colon and replace it with a period

Amended as suggested.

- Line 60: Use either “such as” or “e.g.” – you don’t need both

We removed “e.g.” from the text.

- Line 75: add comma after “Here”

Amended as suggested.

- Line 85: Add a comma after “FESOM1.4”

Amended as suggested.

- Line 98: suggest adding a colon after “defined” so it reads “A pair of control volumes are defined: the vector control volumes are the prisms based on elements.”

Amended as suggested.

- Line 141: maybe add some commas (and remove some of the “ands”) to make things clearer?

Amended as suggested.

- Lines 324-325: both sentences are a bit awkward. Line 324, maybe say something like “The larger magnitude of dissolved iron in the model”? In line 325, maybe remove the word “probably”

Rephrased and amended as suggested.

- Line 326-327 is also a bit awkward

Rephrased it.

- Line 384: “The more severe than expected limitation in iron” is awkward

Thanks for spotting this. We think that the sentence is subjective. Besides, the strength of iron limitation for phytoplankton is hard to compare to observations. We rephrased it.

- Line 395: awkward sentence

Thank you for spotting this. It now reads: “Generally, the NPP and chlorophyll differences to satellite-based estimates could also be linked to model deficiencies, such as coarse model resolution and associated weak upwelling, missing complexity in simulated phytoplankton classes, but also the so far unconsidered nutrient input from terrigenous sources.”

- Line 422: other **global** modelling studies?

Amended as suggested.

- Line 424: global **ocean** and the Southern Ocean

Amended as suggested.