

# Answers to the reviewers' comments

We would like to thank the reviewers for their positive feedback and detailed comments. We have made minor changes to the manuscript based on their comments.

We have also corrected the data in Figure 8f because during the 1<sup>st</sup> revision, a confusion occurred as to which regional profile belongs into which figure panel. This does not change the results described in the text, and the data in this panel now matches with the initially submitted version again.

## Report #1

General comments:

The authors have made considerable efforts to improve the manuscript and to address my previous comments. The resulting manuscript is a much better read and includes a thorough model analysis and evaluation. At this point, I have only a few minor comments and one suggested addition to the introduction or discussion.

Thank you for your kind feedback.

A point that has not been mentioned in the manuscript and perhaps a nice counterpoint to the idea that physical DA just disrupts a carefully calibrated coupled model is that the physical DA often reveals errors in the BGC parametrization. I am referring to statements such as "The question therefore arises to what extent an ecosystem model coupled to a data-assimilated physical model also represents a more realistic biogeochemistry..." (l 88). Studies, such as Löptien and Dietze (2019; DOI: 10.5194/bg-16-1865-2019), demonstrate that compensating for physical model errors through biogeochemical parameter estimation can lead to issues in forecasting climate-relevant metrics, despite reducing errors during the estimation period. So, if the physical DA has a large negative impact on biogeochemical estimates, this may point to problems in the biogeochemical model that may influence model forecasts even without DA. For example, if a physical ocean model underestimates coastal upwelling, maybe due to its coarse resolution, the biogeochemical may have an elevated maximum growth rate for phytoplankton to compensate for the reduced nutrient supply. If a climate forecast increases the nutrient supply to the euphotic zone, perhaps via increased wind stress, the response of the biogeochemical model may be unrealistically high. Physical data assimilation could help to reveal issues like this one, though solving the issue would subsequently require a parameter estimation experiment. Here, I am not suggesting any revision, but the authors could include this point in the manuscript.

We have included this point in the Introduction and Discussion (what was added in red):

Line 85: 'Furthermore, any coupled ecosystem model is adapted to its associated physical model with its strengths and weaknesses through carefully selected parameter values and a spin-up to near-equilibrium. Accordingly, the modeled carbon cycle may be ~~react-very~~ sensitive to deviations from the physical state that is typical for this model (Kriest et al., 2020; Spring et al., 2021). Potentially, this leads to biases in the carbon cycle through physics DA. Such effects highlight where physical model errors are compensated for by BGC parameters, and thereby DA may reveal critical areas for potentially unrealistic BGC model behavior in projections in a changing climate (Löptien and Dietze, 2019).'

Line 625: 'For example, surface chlorophyll (Fig. 11f) and pCO<sub>2</sub> (Fig. 9f) in the central Greenland Sea deteriorate in response to improvements of SST (Fig. 1c), SSS (Fig. 1f) and sea-ice concentration (not shown). This could indicate that the BGC parametrization compensates for flaws in the free running physical model in this region. The parameter mismatch might cause difficulties in modeling the change of BGC variables under the ongoing loss of Arctic sea ice (Chen et al., 2016).'

#### Specific Comments:

L 7: "The main effects": I would suggest adding "of the assimilation", the next sentence could then start with "Its", dropping the "of assimilation" there.

L 46: Why the "However" here? Isn't this another example where undersampling leads to issues?

L 55: "few years": I would suggest using "short time periods".

L 58: "Data assimilation also provides...": I think it is a bit more accurate to state that "Data assimilation can also be utilized to provide...".

L 69: "the uptake of atmospheric CO<sub>2</sub> depends ultimately on the modeled physical carbon transport": What does "depends ultimately" mean here? Sinking of particles and perhaps even vertical migration of zooplankton can lead to carbon export. I would suggest using "in large parts" or similar. Also, why use "modeled" in this sentence?

L 196: I think it is more intuitive to use % for both fractions here. Why write "at each step" here, these are surely average fractions?

L 208: Subscript e appears to be missing for "rand".

L 221: "The ensemble standard deviation of the local instantaneous air-sea CO<sub>2</sub> fluxes that results from the perturbation": I would suggest starting a new paragraph here.

L 240: I would suggest making it a bit more explicit that subscripts will denote the hemispheres.

Eq 5: What about the units here? There is still a °C in the exp.

L 245: For a consistent naming scheme, I would suggest adding the NA subscript to the "Newfoundland Basin+" region. I at first didn't think it was referring to a region and wasn't sure what to make of the +.

L 338: I had to read it twice to notice that surface values are stabilized while the DA continues to correct subsurface values. I would suggest using "subsurface" instead of "3D" in this sentence to make that more explicit.

Fig 6 and 8: If this is not too difficult, I would encourage the authors to add region labels to panel a in both figures. I found myself jumping between the text, Fig 8 (to look up the result) and Fig 7 (to look up the region name again).

L 623: "The major effects of physics DA on BGC variables seem to be related to changes of SST and are largely uniform over the full period of DA": While I know what is meant here, this statement could be misunderstood by readers. Changes in pCO<sub>2</sub> are mostly affected by changes in DIC and alkalinity (which is clearly stated a few lines later). I would suggest rephrasing, so that readers don't get the impression surface pCO<sub>2</sub> is mostly directly modified through changes in SST.

Thanks for the suggestions. We have implemented each of the above points, either omitting the words that were rather causing confusion or making minor additions where they were asked for (see Track Changes document).

L 219: "The strongest inflation ( $\rho = 0.95$ ) is applied during the first two weeks of the DA process.": Is this when the DA increments are strongest because state estimates are furthest from the observations?

That's right. We now added this as an explanation.

L 115: "surface-layer width": Is this the same as height?

That's also right, and we have changed the term to "layer thickness", as it's also called elsewhere in text.

L 173: Are the in-situ observations removed using a similar threshold as for the SST data?

No, we added "without excluding observations".

L 78: "data assimilation of ocean physics into a global ocean biogeochemistry model": Though data assimilation experts will know what is meant, the wording is a bit unclear. I would suggest rephrasing "data assimilation of physical observations into a coupled physical-biogeochemistry global ocean model".

Thanks, we changed this to: 'data assimilation of physical observations into a global ocean general circulation model coupled to a biogeochemistry model'.

L 103: What distinguishes diffusion and mixing here?

Repetitive and we have now deleted "mixing".

L 241: The descriptions that define the regions STSS\_SO+ just says "the area where the assimilation leads to a more negative air-sea flux". I assume this is based on the difference between ASML and FREE. If so, please mention this and introduce  $\Delta\text{FCO}_2$  at that point already

and not 3 lines below it. If not, please improve the description. Overall, I would suggest moving the description and definition of  $\Delta\text{FCO}_2$  towards the top, as it is important for defining the regions used in the study.

That's right, and the paragraph now reads:

To study the effect of DA on the  $\text{CO}_2$  flux, we define regions where the time-mean air-sea  $\text{CO}_2$  flux difference ASML–FREE ( $\Delta\text{FCO}_2$ ) is pronounced, based on the biome definition of Fay and McKinley (2014). Originally, these are, going polewards from the subtropics in each hemisphere, the Subtropical Seasonally Stratified Biome (STSS), the Subpolar Seasonally Stratified Biome (SPSS) and the Sea-Ice Biome (ICE). In the Southern Ocean (denoted by subscript  $_{\text{SO}}$ ) within the  $\text{STSS}_{\text{SO}}$ , we differentiate between the area where  $\Delta\text{FCO}_2$  is positive (the assimilation leads to a flux change directed out of the ocean) referred to as region 'STSS $_{\text{SO}}$ +' and the area where  $\Delta\text{FCO}_2$  is negative referred to as region 'STSS $_{\text{SO}}$ –'. All Southern Ocean regions are outlined in Fig. 5a.

L 297: "On a global average, the SST in FREE is 0.14°C colder than the observations (...) In total, the global mean absolute difference of SST to the observations is reduced from 0.59°C to 0.32°C." I know there are many lines between this and re-reading it I see that one is based on absolute values, but I would suggest further adding a comparable number here. As a reader I was waiting to see by how much the 0.14°C bias was improved.

We have changed the order of these sentences, and added: "Thereby, the global mean model-observation difference is reduced from -0.14°C to -0.12°C, and from 0.59°C to 0.32°C in absolute terms."

## Report #2

General comments:

The authors have done an excellent job of responding to the reviewer comments, and the paper is improved as a result.

We thank you for the positive feedback.

Specific comments:

My only extremely minor comment on reading through is:

Line 349-350: "deep water formation events characterised by a mixed-layer depth of more than 1000 m or 500 m occur less frequently in ASML (not shown)." Is there a significance of "more than 1000 m or 500 m", or could this just read "more than 500 m"?

We changed this to 1000m as it's used e.g. here:

<https://os.copernicus.org/articles/13/609/2017/os-13-609-2017.html>