Answers to the reviewers' comments

We would like to thank the editor and both reviewers for their feedback and many helpful suggestions to improve the text, for pointing precisely to small errors, and for their contributions to the discussion.

In response to the reviewers' comments, we have made the following major changes:

- We reproduced the last year of our simulations to provide additional output on the uncertainty represented by the ensemble
- We saved additional diagnostics (also for the last year of the simulation) on the sources and sinks of DIC and alkalinity through biological processes and advective and diffusive transport. These allow us to better understand and quantify where biological or physical processes affect pCO_2, and where biological and physical processes compensate
- We expanded the introduction
- We restructured the method section and added explanations for technical terms
- We restructured the result section and worked on a more precise wording, also including the new information based on the additional output

Otherwise, we updated the order of authors. The author's contributions are still the same.

In the following, we respond to the reviewers' comments point-by-point. Note that when text was modified and/or new text included, we use extracts from the LaTeX differences template. This highlights deleted text in crossed red text and new text in blue, underlined).

RC 01

1 **Reviewer's comment:**

The manuscript presents an application of an ensemble-based physical data assimilation technique to a global biogeochemical ocean model, with a focus on the effect of physical data assimilation on climate-relevant carbon estimates. The manuscript is mostly well written and offers some valuable insights on the effects of physical DA, but the text could be improved in places and several aspects of the DA experiments should be examined further.

One aspect that is becoming more important in modeling studies but is seemingly ignored in the current version of the manuscript is the reporting of model uncertainty -- even though ensembles are used to generate the results. The authors mention ranges of estimates when reporting results from other studies. However, in their own analysis, the focus is solely on the ensemble mean, without examining the full model ensemble or reporting any uncertainty estimates. It would be beneficial to explore ensemble-based ranges of estimates and compare them to the improvements brought about by data assimilation. This could lead to interesting questions, such as the extent to which data assimilation constrains estimates and whether the estimates improve in areas where they are more constrained. Additionally, figures like Fig. 4 and the seasonal difference plots could be enhanced by including uncertainty estimates, such as the ensemble standard deviation or the interquartile range.

Answer:

Thank you for the suggestion. Indeed, a reduction in the uncertainty of the CO_2 flux estimate would be a very relevant result in addition to an improved estimate of the mean CO_2 flux.

However, the standard deviation in the Kalman filter methodology does not directly translate into an uncertainty estimate. Here, the ensemble standard deviation (STD) of the variables affected during the assimilation step (T, S, SSH, u, v) is reduced. In ASML, most of the reduction in ensemble spread occurs over the course of the first year. After that, the STD remains stable, precisely because we tune our ensemble perturbation and ensemble inflation in such a way that the STD of temperature is maintained after the initial phase (Figure R1; yellow and green lines).

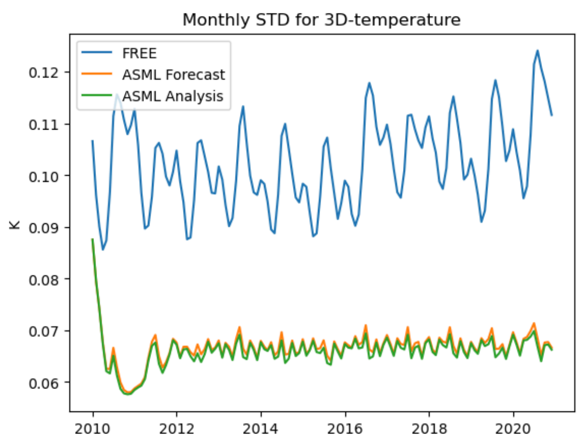


Figure R1: Ensemble standard deviation for 3D-temperature. Note: No volume-weighting applied for the global mean (includes empty cells).

It is thus expected that the ensemble standard deviation of CO_2 flux decreases as well in ASML, but this is a result of the model and not part of the tuning. Indeed, we find that the STD for the local CO_2 fluxes in ASML is reduced to about 75-80% of the STD in FREE after the first year of assimilation (see example in Figure R2; however, this data is not area-weighted).

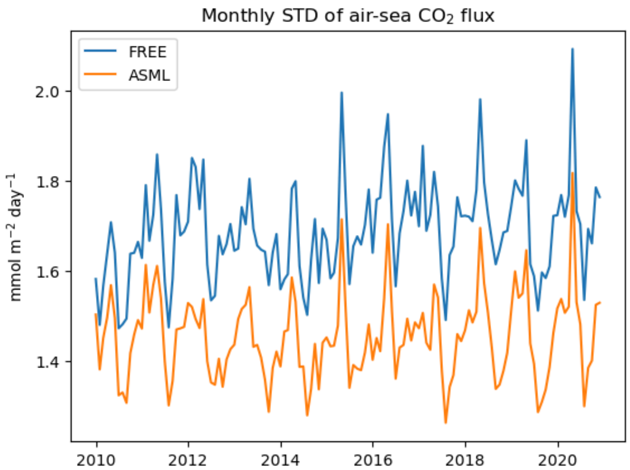


Figure R2: Ensemble standard deviation for the CO2 flux. Note: No area-weighting applied for the global mean.

We have added information on the uncertainty estimates in the revised manuscript (sub-subsection 2.3.1, subsection 3.2 and section 4). Rerunning the simulations was required for additional ensemble member output, and to save computing, we did this only for the year 2020.

In the manuscripts, this reads:

2.3.2 Assimilation method and implementation

(Line 220 in manuscript)

> the forgetting factor is set to either 0.99 or 1.0 depending on the ensemble standard deviation of temperature. The ensemble standard deviation of the local instantaneous air-sea CO_2 fluxes that results from the perturbation of physical fields is larger than that of the global CO_2 flux, with a mean standard deviation of $0.32\,\mathrm{mmol\,m^{-2}\,day^{-1}}$ for monthly means of local fluxes compared to a standard deviation of $0.0068\,\mathrm{mmol\,m^{-2}\,day^{-1}}$ ($0.01\,\mathrm{Pg\,C\,yr^{-1}}$) for the annual global flux in FREE in the year 2020. The largest ensemble standard deviation is generated in the Southern Ocean, the North Atlantic and the North Pacific

> (map in Fig. A1a), which corresponds to regions of high uncertainty in existing CO_2 flux estimates (Pérez et al., 2024; Hauck et al., 2023a; Mayot et al., 2024). However, the modelled standard deviation should not be understood as the true uncertainty of the model, but as a value dependent on tuning (Evensen, 2003).

3.2 Effect of DA on global CO2 flux

> $0.08\,\mathrm{Pg\,C\,yr^{-1}}$ in ASML (not significantly different according to F-test). Through DA, the ensemble standard deviation of the global CO_2 flux in 2020 is reduced from $1.0\times10^{-2}\,\mathrm{Pg\,C\,yr^{-1}}$ in FREE to $0.7\times10^{-2}\,\mathrm{Pg\,C\,yr^{-1}}$ in ASML.

4. Discussion

ing changes in these variables (Fig. 9, Fig. 10 and Fig. 11f). The uncertainty represented by the ensemble is reduced by the DA, which has the most obvious effect on the directly assimilated fields (SST in Fig. 6d and e and density in Fig. 8f). The ensemble standard deviation of the CO_2 flux, where it is large in FREE, is constrained by the DA to globally more uniform and smaller values (Fig. 5c-f, Fig. 7c-f and Fig. A1). Only in the North Pacific, the standard deviation of CO_2 fluxes is equally high in ASML and FREE, precisely in a region that also presents a challenge for pCO_2 products (compare Fig. A1 and Mayot et al., 2024, Figure 5a). In the rest of the ocean, the reduced uncertainty represented by the ensemble does not necessarily

coincide with improved agreement with BGC observations.

The respective figures (and captions) have been updated to show the range of ensemble members through semi-transparent shading.

Figure 5:

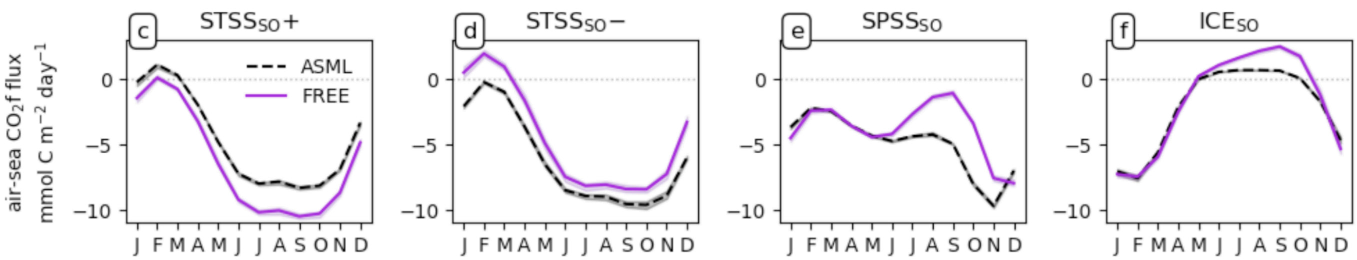
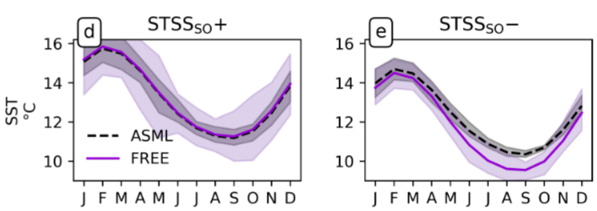


Figure 6:



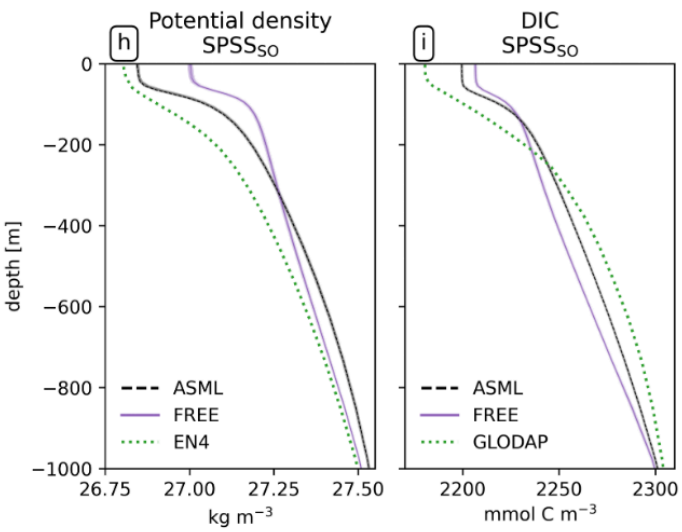


Figure 7:

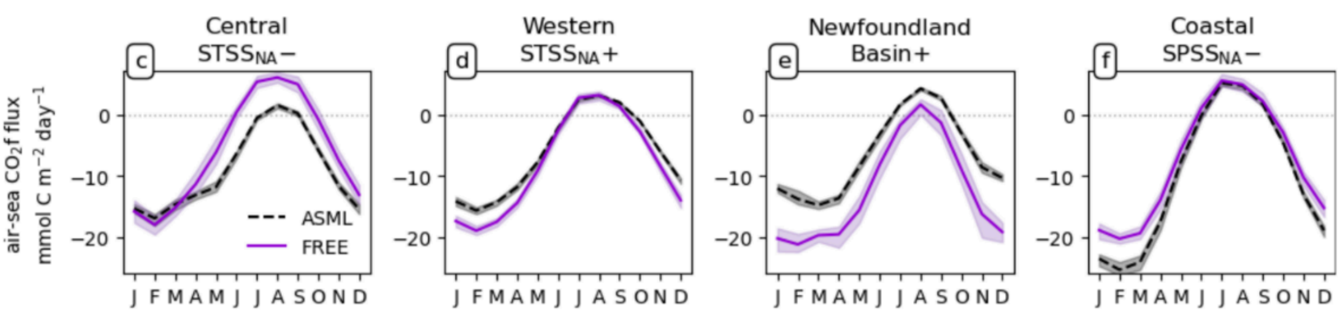
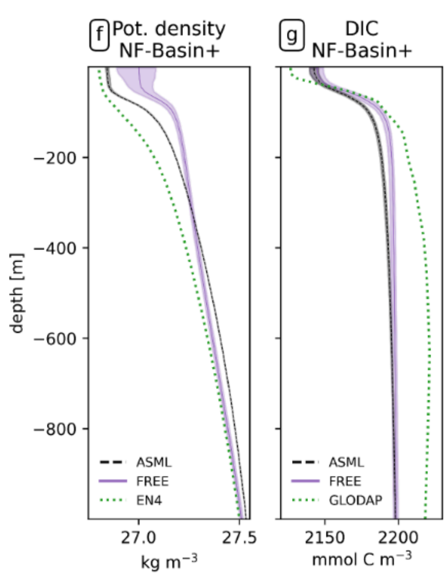
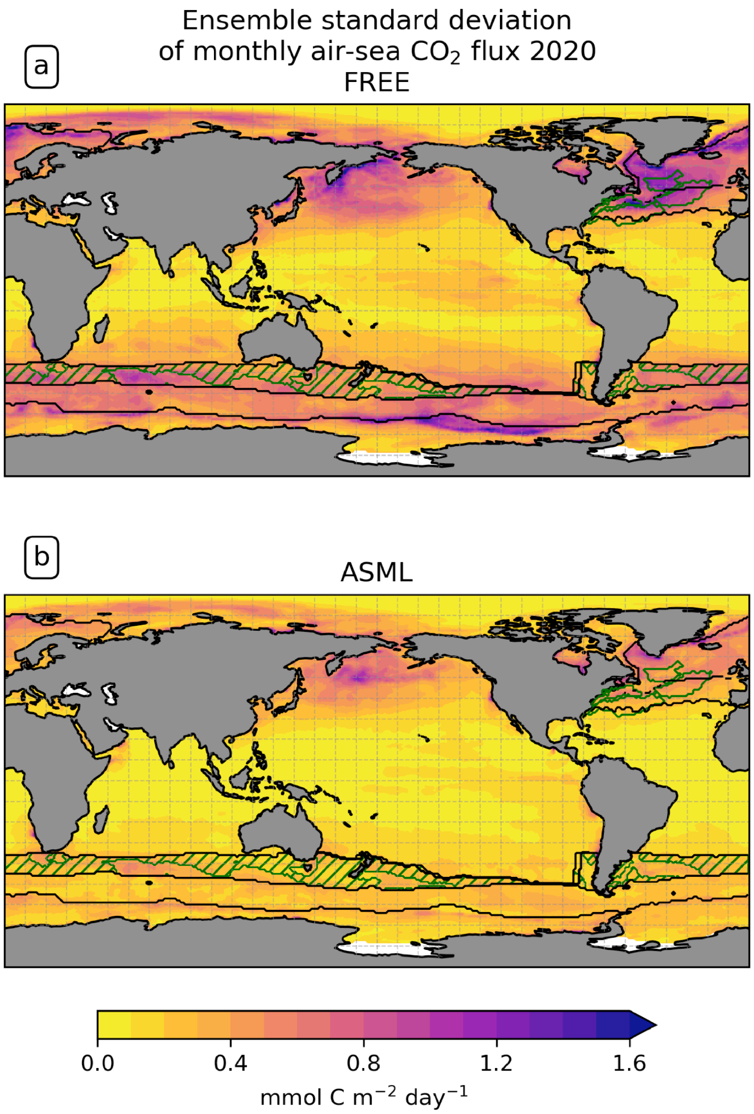


Figure 8:



We have not marked the range of ensemble members in Figure 4 because, for area-integrated fluxes globally and zonally, the uncertainty is so small that it cannot be seen.

We provide Appendix Figure A1:



2 **Reviewer's comment:**

The manuscript emphasizes carbon storage through physical transport, i.e. "upwelling and subduction of DIC, as well as the physical transport of other biogeochemical tracers" (l 60). However, the role of biological carbon fixation and sinking of particulate organic matter seems underexplored. Given that the model includes both slow and fast sinking detritus variables, a more comprehensive examination of these processes would be valuable. Here, it would help to clarify whether the biological carbon export at 200m (l 379 and following) is primarily due to sinking or physical transport. A closer examination or clearer description of the effects of the DA on the biological drivers of carbon export would help to improve the manuscript.

Answer:

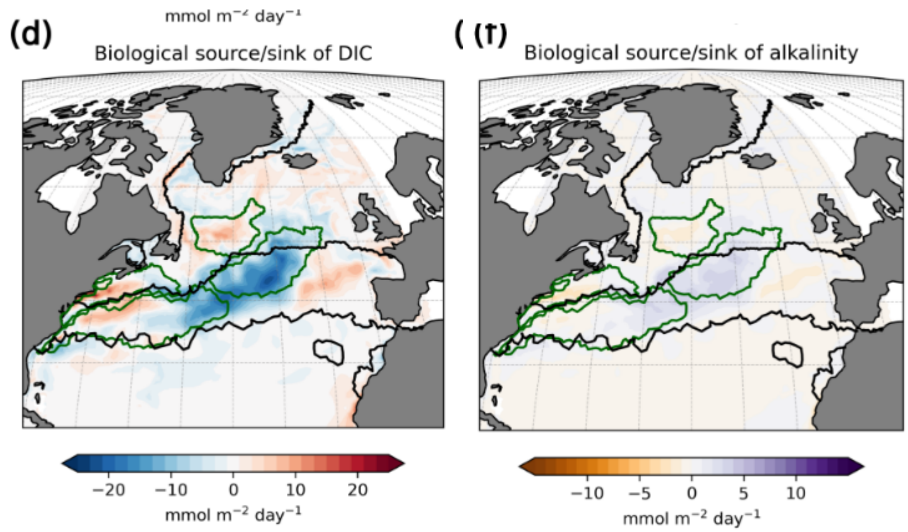
We would like to note that we're most interested in anthropogenic CO_2 uptake, which is primarily physically driven (e.g., Gruber et al., 2023 https://doi.org/10.1038/s43017-022-00381-x). Yet, on a regional scale, changes in the biological carbon sink contribute to the overall carbon balance and thus may have noticeable effects on the regional net CO_2 fluxes. A much closer examination of the biological carbon pump would be interesting, but is beyond the scope of this paper.

In response to the reviewer's comment, in the revised manuscript, we have analyzed additional output for the year 2020, as indicated below, namely:

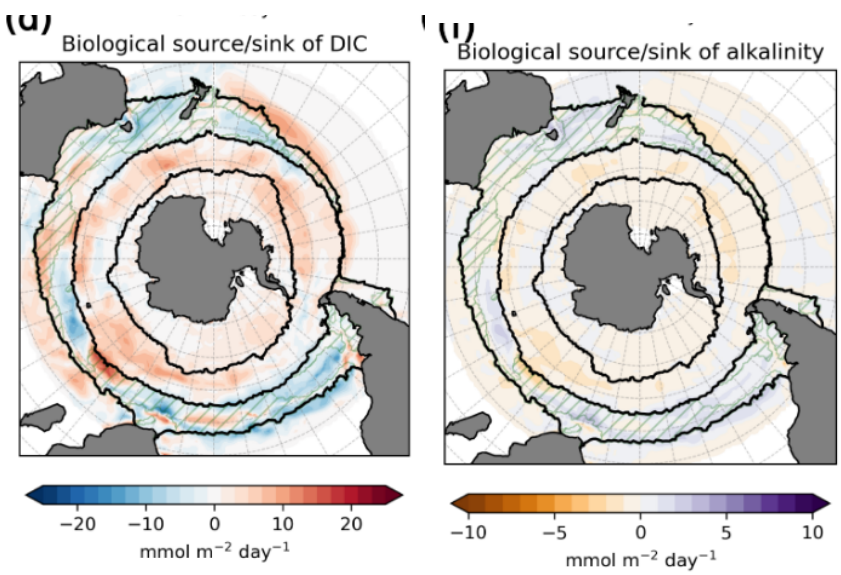
biological net sources or sinks of DIC and alkalinity through combined biological processes:

- For DIC, the net biological term is the sum of photosynthesis, respiration, remineralization of dissolved organic carbon, and formation and dissolution of calcite (Gürses et al., 2023, equation A6).
- For alkalinity, the net biological term is the sum of nitrogen assimilation and remineralization, and formation and dissolution of calcite (Gürses et al., 2023, equation A7).

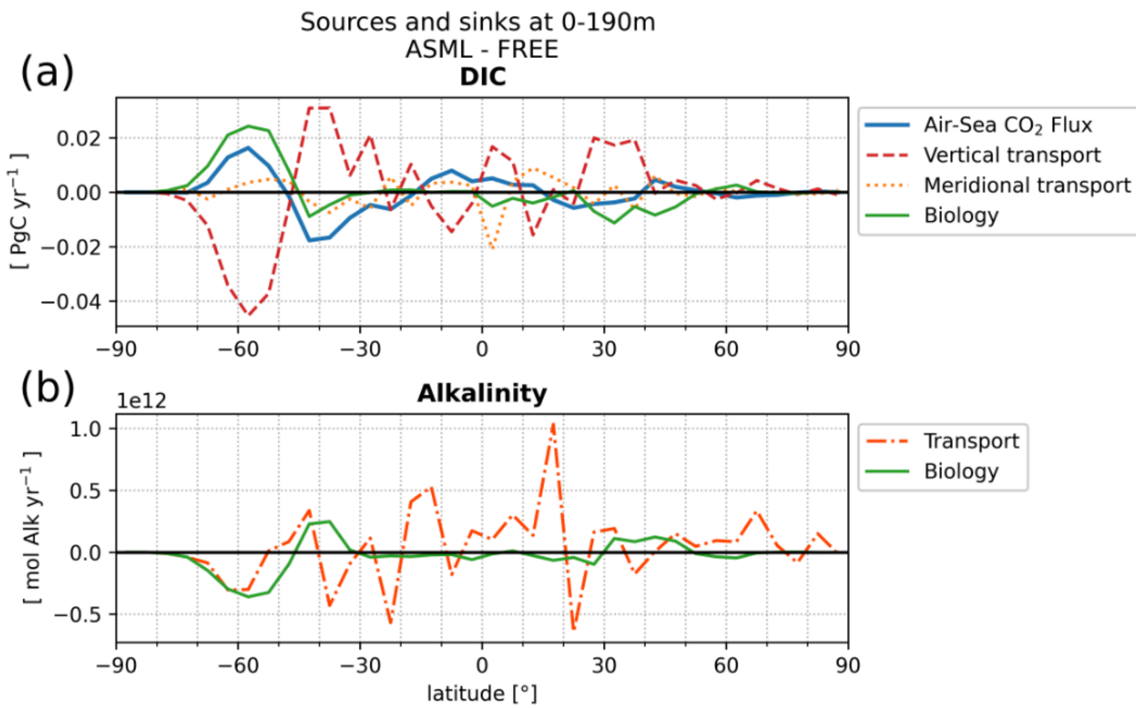
For these, differences ASML-FREE (integrated over 0-190m) are shown in Appendix Figure A15:



Appendix Figure A12:



Appendix Figure A10 (green line in a and b):



In the depth range 0-190m, the biological source/sink term for DIC is negative (-7.5 PgC yr^{-1} globally in FREE for the year 2020). It describes the net transformation of DIC into organic carbon, and therefore only contains the part of biologically fixed carbon that is not remineralised within this depth range again. Thus, while a small amount of this term might add to an increase or decrease of biomass at the same depth range on annual time-scales, most of it is transported to below 190m depth through sinking of detritus (-5.3 PgC yr^{-1}; the gravitational pump; Boyd et al., 2019: https://doi.org/10.1038/s41586-019-1098-2), and some of it is transported to below 190m depth through advection and diffusion of organic material (-1.4 PgC yr^{-1}).

Wherever we have found that DA has a considerable effect on the biological source/sink term in a certain region, we have indicated this in the manuscript (see Track Changes document). This reads:

Section 3.2 Effect of DA on regional CO_2 fluxes and their drivers

~~the southern part of the ICE biome, near the Antarctic continent, the effect of the~~ fragmented area of the $STSS_{SO}$+, different factors contribute to regional changes of the surface DIC and alkalinity budget in ASML (sources minus sinks of DIC and alkalinity in Fig. A11). Depending on location, an increased upward transport of DIC through mixing, an increase of DIC through a reduced biological sink of DIC in spring, or a decrease of alkalinity through changes in horizontal and vertical advection dominates. The seasonality of the effect of DA on the ~~CO_2 flux is small.~~

(Line 533 ff)

~~to~~ Where there is lower DIC in the ~~STSS- region~~ $STSS_{SO}-$ in ASML (~~not shown). The seasonality of the effect of DA on~~

~~summer. In contrast, in the STSS- region, the~~ Fig. 6b), this can mostly be explained by an increased biological sink of DIC, with the addition of sharply defined local changes in horizontal advection of DIC and alkalinity (Fig. A12). Additionally, seasonal

…

DIC observations, albeit some differences to GLODAP still exist. Besides the fact that the differences in stratification and ~~boundary layer~~ boundary-layer depth affect the vertical DIC profile, they also imply less available surface nutrients in ASML. Probably due to ~~that, in the SPSS,~~ a combination of lower nutrient availability and colder surface temperature, ASML features lower NPP, lower chlorophyll concentrations and a lower phytoplankton biomass in the $SPSS_{SO}$ (not shown). Thereby, the modeled biogeochemical cycle adjusts to the lower transport of nutrients to the surface by transferring less organic material to depth, ultimately acting to compensate about 60% of the difference in physical transport of DIC (Fig. A13a) and adding to the reduction in surface alkalinity (Fig. A13b). Within the $SPSS_{SO}$ (roughly south of 50 °S), differences between FREE and ASML in terms of the temperature effect on pCO_2, vertical transport of DIC and alkalinity and biological sources and sinks are larger than at any other latitude (Fig. A13).

…

than in the climatology, leading to less agreement in the Central $STSS_{NA}-$ (Fig. A6). Likely facilitated by more available nutrients through deeper mixing in winter and spring, ASML features a higher biological sink of DIC above 190 m (Fig. A15 d), more biological carbon export ~~at 200~~through sinking of detritus at 190 m, ~~surface chlorophyll and~~ more column integrated phytoplankton biomass and surface chlorophyll in spring, which is ~~shown~~ illustrated by the example of surface chlorophyll difference between ASML and FREE in Fig. 8e. ~~This is facilitated by more available nutrients through winter mixing in the surface boundary layer, which is deeper in ASML (Fig. 8d).~~ In combination, the higher ~~biological export of carbon~~ alkalinity associated with NAC transport and the higher ~~alkalinity in ASML~~ biological sink of DIC result in lowered surface pCO_2 and higher oceanic uptake.

…

~~and surface chlorophyll are reduced (surface chlorophyll in Fig. 8e)~~Furthermore, ASML represents less surface chlorophyll in the Newfoundland Basin+ (Fig. 8e) as a result of a redistribution of biomass from the surface to 50-400 m depth due to spring mixing (not shown). The downward mixing of biomass results in an increase of the biological sink of DIC above 50 m likely due to more primary production near the surface, ~~probably due to shallower mixing. In the Western Boundary STSS, the increased CO_2 uptake and reduced pCO_2~~ and a decrease of the biological sink at 50-400 m likely due to more remineralization at this depth. However, the differences in the biological sink of DIC are compensated by mixing of DIC (profiles not shown). Overall, differences of the regional DIC profile to the observational GLODAP climatology slightly increase (Fig. 8g).

4 Discussion:

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 (Section 3.3). Surface chlorophyll changes follow SST changes (~~Fig. 11 and Fig. 1), as the~~ Figs. 1 and 11). The modeled phytoplankton growth is temperature-dependent (Gürses et al., 2023). Furthermore, indirect temperature effects on plankton dynamics due to stratification and mixing changes contribute, albeit those can have heterogeneous effects and the correlation of chlorophyll and boundary-layer depth is less clear (not shown). The changes of surface DIC and alkalinity

…

above 200 m and higher DIC between 200-600 m depth. In regions of substantial DA effects on vertical transport of DIC, as for example in the Central $STSS_{NA}$ – or in the $SPSS_{SO}$ (Section 3.2), the modelled biogeochemical cycles adjust dynamically to the altered vertical transport. The resulting changes in biological sources and sinks of DIC compensate for 20-70% of the changes in vertical transport of DIC (Fig. A10a). In addition to changes in stability and mixing, the assimilation affects the

3 **Reviewer's comment:**

The assimilation of physical observations that only directly updates the physical variables can lead to "shocks" in the biogeochemical variables. It would be valuable to know if the authors observed any negative effects of daily physical updates on the biogeochemical state, such as unexpected phytoplankton blooms (for example, caused by a deepening of the mixed layer transporting nutrients, formerly below the mixed layer, to the surface).

Answer:

We are not aware of any such shocks. This might relate to our overall finding that the modeled carbon fluxes and other inspected variables such as chlorophyll-a, NPP and plankton biomass act almost surprisingly indifferent to substantial differences in the model physics. The most rapid assimilation-induced changes take place in the first few months after the start of the assimilation, yet there was no noticeable shock.

4 **Reviewer's comment:**

Several aspects of the model setup and data assimilation process could benefit from further explanation or discussion. For instance, the restoration of surface salinity towards climatology may interfere with the assimilation of salinity data. It would be informative to know if the authors have experimented with switching off the nudging when or where salinity data is being assimilated, and how well the salinity climatology aligns with the assimilated data.

Answer:

The main effects of SSS assimilation and salinity restoring are to reduce the simulated SSS globally. In addition, there are certain regions of model bias, such as the Amazon river inflow area and the North Atlantic Current, where both methods are consistent with each other. While there are gaps in the SSS-CCI data near the poles, the salinity restoring towards climatology is with global coverage. Experiments with and without salinity restoring show that without it, in FREE, sea surface salinity drifts by approximately +0.05 psu during the first year after switching it off. In ASML, the difference between switching salinity restoring off or on is smaller (less than 0.01 psu globally), because the assimilation compensates for the lack of restoring. In ASML, global SSS is reduced by approximately 0.15 or 0.2 psu compared to FREE, respectively, after one year, which shows that the assimilation has a stronger effect than the restoring. The best agreement with SSS-CCI observations is achieved when assimilation and salinity restoring are used simultaneously.

In summary, we added the following to the manuscript, 3.1 Effect of DA on ocean physics:

> The assimilation also improves the agreement with the assimilated SSS observations. Additional experiments with and without salinity restoring towards climatology show that the best agreement with the SSS-CCI observations is achieved by simultaneously using assimilation and restoring. A benefit of the additional use of restoring is the global coverage of the SSS climatology. FREE shows a global SSS bias (0.49 psu, Fig. 1d). The assimilation leads to a global surface freshening (Fig. 1e).

5 **Reviewer's comment:**

Similarly, the exclusion of temperature observations from the DA when the model-observation difference exceeds 2.4°C could use a better explanation, as this seems to hinder assimilation where it might be most needed.

Answer:

By excluding these observations, the aim is to prevent strong and sudden corrections from making the model unstable, especially in the initial phase. Instead, a 'gentler' correction is made by assimilating neighboring points. Because we use a gap-filled SST observational product, observations are continuously available in the neighboring domains. We have added some text to reflect this to the manuscript, on SST assimilation:

> than the nominal resolution of the model grid. An observation error standard deviation of 0.8°C is prescribed for the DA following Nerger et al. (2020). Observations are excluded in the DA process if the difference between the model and observation exceeds three times the observation error standard deviation, thus 2.4°C, and at grid points with sea ice in the model, as in Tang et al. (2020) and Mu et al. (2022). This exclusion keeps the model stable despite large differences between model and observations at these sites, in particular as water temperature and salinity develop differently under sea ice than under the influence of the atmosphere (Tang et al., 2020). Instead, a 'gentler' correction is made by assimilating neighboring points. After the initial phase, about 7% of SST observations are excluded because of the 2.4°C-threshold. Nevertheless, the data assimilation still has a strong effect in areas where these large model-observation discrepancies are typically found (North Atlantic, Japan and Southern Ocean).

6 **Reviewer's comment:**

To improve readability, particularly for readers less familiar with data assimilation techniques and carbon modeling, brief explanations of key concepts and modeling choices would be beneficial. These would include descriptions of the term used to perturb atmospheric forcing, the role of ensemble inflation, and the rationale behind the choice of γ_DIC and γ_Alk in Equations 4 and 5 (see also my specific comments below). Currently, the manuscript often uses references to other studies to motivate implementation details, and an additional sentence here and there could help the reader to better understand these details without having to go through other papers.

In places, the structure of the manuscript can be improved to enhance clarity and flow. Sections 4.2 and 4.3 are quite lengthy and could be subdivided based on location (Southern Ocean, Atlantic) and the different data products used in the comparisons. Section 3, which contains results from the two ensemble simulations, could be merged with Section 4 to create a more cohesive results section.

Answer:

Thank you for the suggestion. We have rearranged the sections and section titles accordingly. The structure of the revised manuscript is now as follows:

To add structure to Sections 3.2.1. and 3.2.2, we use bold font to state which region is described in the following paragraph, e.g.:

STSS$_{SO}$ In the northernmost biome of the Southern Ocean, the subtropical seasonally stratified biome (STSS~~, outlined in Fig. 5a~~$_{SO}$), the mean oceanic CO_2 uptake is comparably high (Fig. 5a). ~~Here, the~~ The uptake is largest in austral winter

7 **Reviewer's comment:**

Overall the figures look very good and are helpful, I only have a minor suggestion here: it might be more informative to report ASML-OBS instead of ASML-FREE in Figures 1-3. This would provide a clearer picture of the model error following data assimilation. Also, some of the figures, such as Figure 7, have lots of whitespace that could be reduced.

Answer:

Indeed, for temperature and salinity, ASML–OBS provides a clear picture of the model error after data assimilation (see SST, Figure R6).

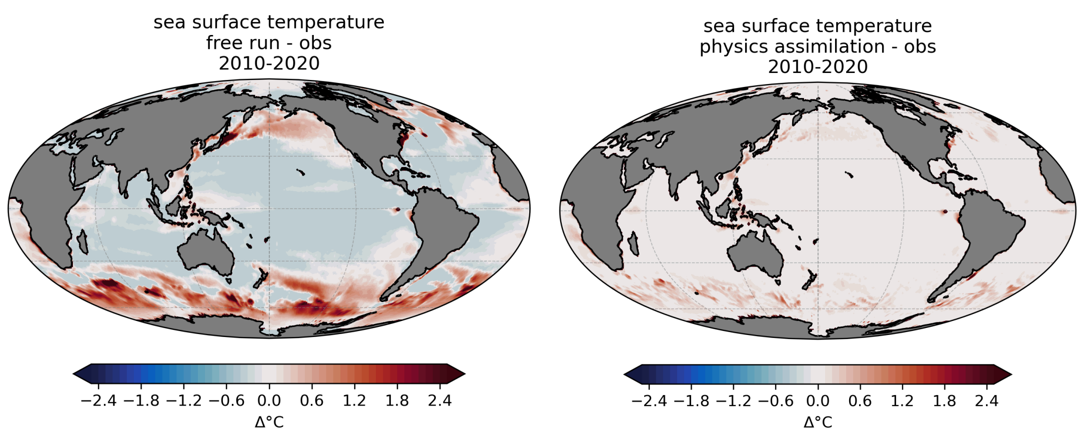


Figure R6: FREE-OBS and AMSL-OBS for SST, useful to illustrate the model error before and after assimilation

However, for the biogeochemical variables, FREE–OBS and ASML–OBS are visually too similar to recognize the differences (see chlorophyll, Figure R7).

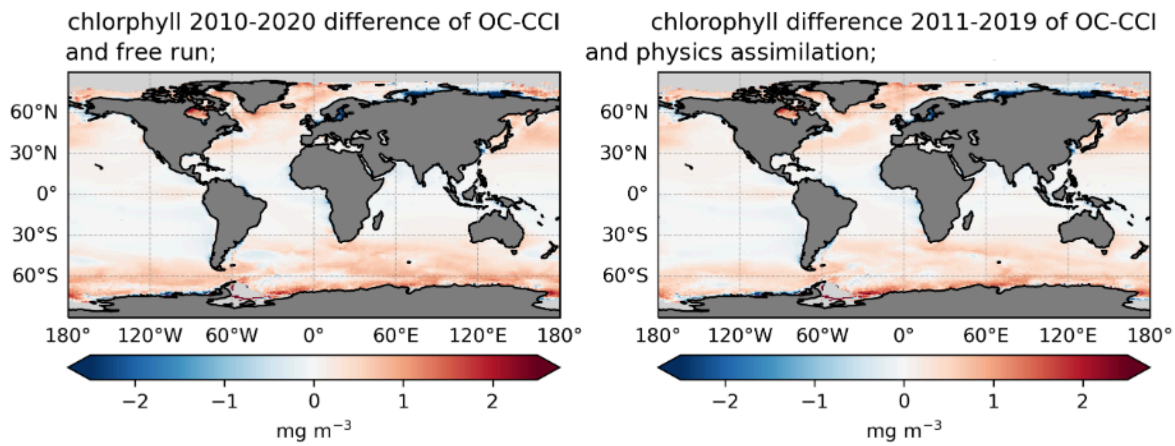


Figure R7: FREE-OBS and AMSL-OBS for chlorophyll, the effect of the assimilation is almost invisible

Therefore, we have chosen to show ASML–FREE because it allows us to visualize comparatively small changes in the biogeochemical variables. Showing ASML–FREE for all variables throughout the manuscript allows one to recognize correlations between the effects of DA on different variables.

8 **Reviewer's comment:**

L 8: "the mean CO2 uptake increases by 0.18 Pg C yr−1": Add "regionally" here to make it explicit that this increase is not a resulting global estimate.

Answer:

Thank you for the detailed comments here and below. Implemented here:

Ocean during winter. South of $50\,^{\circ}$S, winter CO_2 outgassing is reduced and thus the ~~mean~~ regional CO_2 uptake increases by $0.18\,\mathrm{Pg\,C\,yr^{-1}}$ through the assimilation. Other particularly strong regional effects on the air-sea CO_2 flux are located in the

9 **Reviewer's comment:**

L 40: "the model mean": It would be helpful to the reader to add a few words about the kind of models that were considered here.

Answer:

Done here:

> atmospheric oxygen data and atmospheric inversions (Friedlingstein et al., 2023). For the years 2010-2020, pCO_2 products included in the Global Carbon Project suggest a mean oceanic sink of $3.0 \pm 0.4\,\mathrm{Pg\,C\,yr^{-1}}$, while the model mean mean of Global Carbon Project GOBMs is $2.5 \pm 0.4\,\mathrm{Pg\,C\,yr^{-1}}$, with trends of (data provided by Friedlingstein et al., 2023). Trends over the same time period are $0.7\,\mathrm{Pg\,C\,yr^{-1}\,dec^{-1}}$ and $0.3\,\mathrm{Pg\,C\,yr^{-1}\,dec^{-1}}$, respectively(data provided by Friedlingstein et al.,

10 Reviewer's comment:

L 65: "DIC" was used before the abbreviation is introduced here (l 59). The earlier sentence actually makes a quite similar point about subduction of DIC and also mentions upwelling, perhaps this could be made more concise.

Answer:

Rearranged to merge the two sentences that make similar points into one sentence, introducing DIC at its first use, now reads:

> While previous studies indicate that the available BGC observations, when assimilated in isolation, are too sparse to constrain the modeled carbon cycle (Verdy and Mazloff, 2017)(Verdy and Mazloff, 2017; Spring et al., 2021), the assimilation of physical variables is expected to have a significant indirect effect on the modeled CO_2 uptake because upwelling and subduction of DIC, as well as the physical transport of other biogeochemical tracers, will be affected air-sea CO_2 fluxes (Bernardello et al., 2024). This is because the uptake of atmospheric CO_2 depends ultimately on the modeled physical carbon transport between the surface, the mixed layer and the deep ocean in the form of dissolved inorganic carbon (DIC) through mixing, upwelling and subduction (Doney et al., 2004). According to current knowledge, ocean physics is the dominant driver

11 Reviewer's comment:

L 65: "It was shown that assimilating ocean physics at the initial state of a model simulation has a stronger and more positive impact on the modeled carbon cycle than assimilating the BGC initial state": Is this due to the lack of BGC observations mentioned earlier, the importance of physical processes for carbon export, or a large physical model error that cannot be decreased through BGC DA?

Answer:

Fransner et al., 2020 relate the strong and positive effect of assimilating ocean physics to the strong control ocean physics exerts on the biogeochemical variability on interannual to decadal time scales (rather than low availability of BGC observations or strong physical model errors). Thus, we have added:

According to current knowledge, ocean physics is the dominant driver of interannual variability of the global air-sea CO_2 flux and also responsible for stagnation and acceleration of the CO_2 uptake on decadal scales (Doney et al., 2009; Keppler and Lands

. Related to the strong control that physics exert on the interannual variability of air-sea ~~interface across the mixed layer into the deep ocean in the form of dissolved inorganic carbon (DIC) (Davila et al., 2022). It was shown~~ CO_2 fluxes, it was shown in one idealized study that assimilating ocean physics at the initial state of a model simulation has a stronger and more positive impact on the modeled carbon cycle on interannual time-scales than assimilating the BGC initial state (Fransner et al., 2020).

12 **Reviewer's comment:**

The next sentence brings up the question of which processes are most important. Maybe a few candidates could be named and briefly discussed here before going into the details of the DA algorithm.

Answer:

Naming and discussing a few candidates here:

physics DA. The question therefore arises ~~as~~ to what extent an ecosystem model coupled to a data-assimilated physical model also represents a more realistic biogeochemistry~~. We will present cases where physics data-assimilation leads to worse and better agreement with BGC observations.~~ , and which mechanisms are responsible for the response of the CO_2 flux in physics DA approaches. One possible driver is the physical transport of DIC and alkalinity because velocities and diffusivity are changed by the DA, affecting in particular the upwelling of carbon-rich waters and subduction, which is important to capture the ocean storage of anthropogenic carbon (Davila et al., 2022). Furthermore, physics DA may change pCO_2 directly through its temperature-dependence, an effect emphasized by Verdy and Mazloff (2017). Additionally, the modelled biological pump might be altered, for example through the temperature-dependency of phytoplankton growth or through effects of stratification on nutrient availability.

13 **Reviewer's comment:**

L 70: "continuously assimilating ocean-physics for eleven years": A bit more detail could be useful here as well: What does assimilating ocean physics entail, what observations are being used for the DA here?

Answer:

More details added to the introduction:

of ocean physics a prerequisite for a realistic simulation of the contemporary CO_2 flux. We here use ensemble-based data assimilation of ocean physics into a global ocean biogeochemistry model aiming to improve the modeled air-sea CO_2 flux for the years 2010-2020. For ~~assimilation we use an ensemble Kalman filter variant (Nerger et al., 2012). With thisapproach, we describe the impact of continuously assimilating ocean-physics~~ this, we continuously assimilate temperature and salinity observations from remote-sensing at the surface and from in-situ profile measurements for eleven years ~~on the model's air-sea~~

~~the thermally, DIC- and alkalinity induced components and changes in mixing, and lateral and vertical transport~~and update the modelled temperature, salinity, horizontal velocities and sea surface height, using an ensemble Kalman filter variant (Nerger et al., 2012).

14 Reviewer's comment:

L 89: "The model allows for a variable mesh resolution": What is a typical coarse and fine resolution used in the model grid?

Answer:

We have now moved Section "Simulation set-up" up here, clarifying:

2.2 Simulation set-up

The model setup for both simulations closely follows Gürses et al. (2023). The mesh resolution is nominally 1 degree, ranging between 120 km and 20 km with enhanced resolution in the equatorial belt and north of 50 °N (126858 surface nodes). It has 47 vertical layers with thickness ranging from 5 m at the surface to 250 m in the deep ocean, as described by Scholz et al. (2019, COR

15 Reviewer's comment:

L 93: A salinity flux of 0.1m/day? Please describe this better.

Answer:

Thanks for asking, in fact, this number was a typo. We corrected the number and added Eq. (1) to clarify:

. The surface salinity (SSS) is restored towards ~~climatology and scaled by a fictional flux at the ocean surface of 0.1m/day~~the World Ocean Atlas climatology through a fictional surface flux with $v_{SSS} = 50\,\mathrm{m}/300\,\mathrm{days}$ according to equation 1 and as in Gürses et al. (2023):

$$(SSS_{clim} - SSS_{model}) * v_{SSS} * (h_{surf})^{-1} \quad (1)$$

with surface-layer width h_{surf}. A detailed description of FESOM2.1 ~~is provided by Danilov et al. (2017)~~ and a model assess-

For the example of a salinity bias of 0.5 psu and with the surface-layer width being around 5m (more or less depending on sea surface height etc.), this would yield a correction of approx. 0.016 psu per day.

16 **Reviewer's comment:**

L 96: "DIC" is introduced again, a quick search shows 7 introductions of "DIC", also counting captions.

Answer:

Thanks, we only kept the introduction of "DIC" once in the Introduction, and once more in the Conclusion.

17 **Reviewer's comment:**

L 117: "observations are weighted by distance": This is not a precise statement that could confuse some readers, express more clearly that the ensemble estimated correlation between a model grid point and an observation is down-weighted using a distance-based metric. Is vertical localization applied as well?

Answer:

The localization acts in the horizontal only. We have phrased more precisely:

> With localization of the LESTKF, ~~observations are weighted by distance , thereby avoiding the model being~~ the observation error is increased for an increasing horizontal distance between an observation and a model grid point, which weighs down the influence of a more distant observation. This avoids that the model is influenced by observations at distant locations through spurious ensemble estimated correlations. We use a localization radius of 200 km and choose a 5th-order polynomial

18 **Reviewer's comment:**

Eq. L 124: It would be useful to add equation numbers to all equations, even those that are not referenced in the text, so that they can be more easily referenced in other texts, such as this one.

Why does a larger ensemble amplify rand? It does not seem that intuitive to have larger perturbations in a larger ensemble.

Answer:

We added equation numbers to all equations.

The incomplete definition of 'rand' in the initial manuscript has led to an obvious misunderstanding: In fact, there are no larger perturbations in a larger ensemble. The factor (N_ens-1) compensates that the values of 'rand',

defined as elements of a stochastic matrix which sum up to 1, become smaller with increasing ensemble size because the matrix becomes larger. In detail, the values for rand are generated by Second-Order Exact Sampling from a trajectory of atmospheric forcing fields, a method introduced by Pham et al., see e.g.:
https://doi.org/10.1175/1520-0493(2001)129<1194:SMFSDA>2.0.CO;2
and briefly explained here:
https://pdaf.awi.de/trac/wiki/EnsembleGeneration

To clarify, in the updated manuscript we added a few sentences on the generation of the initial perturbation, and we have now redefined the stochastic element (still called ‘rand’), so that it already includes the factor (N_ens-1) that initially caused confusion.

To maintain ensemble spread, we apply a perturbed atmospheric forcing with an autoregressive perturbation ~~($\mathrm{perturb}_{\mathrm{n}}$)~~ ($\mathrm{perturb}_{\mathrm{e,n}}$) at every model time step (n) to each ensemble member (e), with:

~~$\mathrm{perturb}_{\mathrm{n+1}} = (1-\mathrm{arc}) * \mathrm{perturb}_{\mathrm{n}} + \mathrm{arc} * s * (N_{\mathrm{ens}}-1) * \mathrm{rand}$~~

$$\mathrm{perturb}_{\mathrm{e,n+1}} = (1-\mathrm{arc}) * \mathrm{perturb}_{\mathrm{e,n}} + \mathrm{arc} * s * \mathrm{rand}_{\mathrm{e}} \quad (2)$$

where rand is a stochastic element~~that is based on a covariance matrix derived~~, again generated by second-order exact sampling from a 72-days-long ~~period~~ trajectory of atmospheric forcing ~~; the~~ fields that captures patterns of day-to-day atmospheric variability. The autoregression coefficient (arc) ~~is~~ can be used to tune how quickly the perturbation changes and is set to the inverse number of model steps per day~~; and~~. s is a scaling factor for each perturbed atmospheric forcing field. For spe-

19 **Reviewer’s comment:**

L 153: "model values are computed as the average of the grid points of the triangle enclosing the observation because the number of observations is fewer than model grid points": Averaging is required to interpolate the model solution at the observation locations, why is this dependent on the number of observations?

Answer:

Thanks for pointing out how this can lead to confusion. In fact, we simply meant:

1. If observations are spatially highly resolved, they are interpolated to the model grid (as for SST and SSS).
2. If observations are available only at a few points, it is the other way round and the model solution is interpolated to the observation locations (as for the profile data).

Because this was unnecessarily confusing, we have left it out. The text now reads:

The assimilated temperature and salinity profiles are taken from the EN.4.2.2 data set (Good et al., 2013). The EN4 dataset contains quality-controlled profiles from various in-situ ocean profiling instruments. To assimilate the profiles, the observations are assigned to the respective model layers (depth range) in the vertical. In the horizontal, the model values are computed as the average of the grid points of the triangle enclosing the observation. The observation error standard deviation is set to 0.8 °C for temperature and to 0.5 psu for salinity, as in Tang et al. (2020).

20 **Reviewer's comment:**

L 157: This information about the model grid is missing from Section 2.1 where the model grid is described for the first time. It would also be useful to describe the atmospheric forcing before describing the perturbation to it (Section 2.2.1).

Answer:

Rearranged to:

2.1 Model FESOM-REcoM
2.2 Simulation set-up *(here, we describe the grid and atmospheric forcing)*
2.3 Data Assimilation
2.3.1 Assimilated observations
2.3.2 Assimilation method and implementation *(here, we describe the perturbation to the atmospheric forcing)*

21 **Reviewer's comment:**

L 171 "the river flux adjustment (...) is applied to the pCO2 products. ...": It is not entirely clear what this means, the focus here is just the CO2 flux associated with the oceans, I presume? The next sentence provides some more information but it seems to imply that the RECCAP2 CO2 flux is not being used for comparison, when previous sentences stated that it was. Some clearer language would be useful here.

Our model and other GOBMs do not account for the natural river flux, which is (simplified):

1. rivers carry organic carbon into the ocean
2. as a consequence, carbon, once remineralized, outgasses from the ocean into the atmosphere
3. fixation of atmospheric CO_2 by terrestrial and freshwater ecosystems, and export via rivers (→ 1.)

The river flux adjustment (https://www.nature.com/articles/s41586-021-04339-9) serves to make GOBM estimates of the air-sea CO_2 flux comparable with other estimates, which, in contrast, do account for the river flux.

To clarify, we have rephrased:

> ~~To assess the model results we focus on the ensemble mean.~~ We present CO_2 flux estimates for the period 2010-2020, that are compared to the 'Regional Carbon Cycle Assessment and Processes 2' (RECCAP2) global air-sea CO_2 flux estimates (DeVries et al., 2023). ~~For the comparison of the global air-sea CO_2 flux in our simulations with the~~ The RECCAP2 ~~CO_2 flux estimates, the~~ pCO_2 products account for oceanic outgassing of river carbon into the atmosphere. To make them comparable with our estimate stemming from a model without river carbon input, we apply a river flux adjustment (Friedlingstein et al., 2023; Regnier et al., 2022) ~~is applied to the~~ to the RECCAP2 pCO_2 products. Thus, we quantify the anthropogenic perturbation of the ocean carbon sink ~~without rivers~~ (as S_{OCEAN} in the Global Carbon Budget Friedlingstein et al., 2023; Hauck et al., 2020), and not the contemporary net air-sea CO_2 flux with outgassing of river carbon (as in the original RECCAP2 pCO_2 products).

22 **Reviewer's comment:**

L 183: Should the US East Coast be considered subpolar, are all regions characterized by seasonal stratification, or does SPSS stand for something different here? A alternative choice of region names may be suitable and would avoid confusion with the region names in the Southern Ocean.

Answer:

According to the definition of Fay and McKinley, the STSS, SPSS and ICE biomes exist analogously in both hemispheres (https://doi.org/10.5194/essd-6-273-2014). Therefore, there is an SPSS and STSS biome in the Southern Ocean and in the North Atlantic, of which we discuss only specific parts (e.g. the Coastal SPSS).

The Fay and McKinley biomes are used widely in the ocean carbon cycle community (see e.g. RECCAP papers, https://reccap2-ocean.github.io/publications/).

To avoid confusion with the regions names in the North Atlantic (NA) and Southern Ocean (SO), we have added subscripts to the names, e.g. $STSS_{SO}$+ and Coastal $SPSS_{NA}$–. The "+" and "–" symbols denote the sign of the effect by which each region is defined. In the revised manuscript, this reads:

> To study the effect of DA on the CO_2 flux, we define regions where the effect is pronounced and where different mechanisms are active~~. In the Southern Ocean, we use~~ , based on the biomes defined by Fay and McKinley (2014). These are, ~~from North to South~~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)~~(see Fig. 5). Within the STSS~~. In the Southern Ocean ($_{SO}$), within the $STSS_{SO}$, we differentiate between the area where the assimilation leads to a more positive air-sea CO_2 flux (positive: out of the ocean), referred to as $STSS_{SO}+$ and the area where the assimilation leads to a more negative air-sea flux, the ~~STSS-~~ $STSS_{SO}-$ (Fig. 5a and b). In the North Atlantic ($_{NA}$), we consider four coherent regions within the $STSS_{NA}$ and $SPSS_{NA}$, defined by the time-mean difference of the air-sea CO_2 fluxes in ASML and FREE (ΔF_{CO_2}~~, Fig. 7a and b~~). The Central $STSS_{NA}-$ and Western $STSS_{NA}+$ are located in the central North Atlantic $STSS_{NA}$ biome and are confined by $\Delta F_{CO_2} < -1\,\mathrm{mmol\,C\,day^{-1}\,m^{-2}}$ and $\Delta F_{CO_2} > 1\,\mathrm{mmol\,C\,day^{1}\,m^{-2}}$, respectively (see Fig. 7b). The Newfoundland Basin~~and East Coast SPSS~~+ and Coastal $SPSS_{NA}-$ are part of the $SPSS_{NA}$. The former is located east of Newfoundland and south of Greenland, and is confined by $\Delta F_{CO_2} > 3\,\mathrm{mmol\,C\,day^{-1}\,m^{-2}}$; and the latter is located off the North American coast and confined by $\Delta F_{CO_2} < -1\,\mathrm{mmol\,C\,day^{-1}\,m^{-2}}$. The Central $STSS_{NA}-$ and Western $STSS_{NA}+$ lie on the warm side of the ~~NAC~~North Atlantic Current (NAC), and the Newfoundland Basin~~and West Coast SPSS~~+ and Coastal $SPSS_{NA}-$ lie on the cold side of the NAC, which is evident from the modeled surface velocity field (Fig. A2a).

23 Reviewer’s comment:

L 185: Please explain "NAC".

Answer:

Defined in Line 307:

> ~~NAC~~North Atlantic Current (NAC),

24 Reviewer’s comment:

Eq 1 and 2: Is there an easy to communicate motivation for the choice of γ_DIC and γ_Alk ?

Answer:

We describe the motivation here:

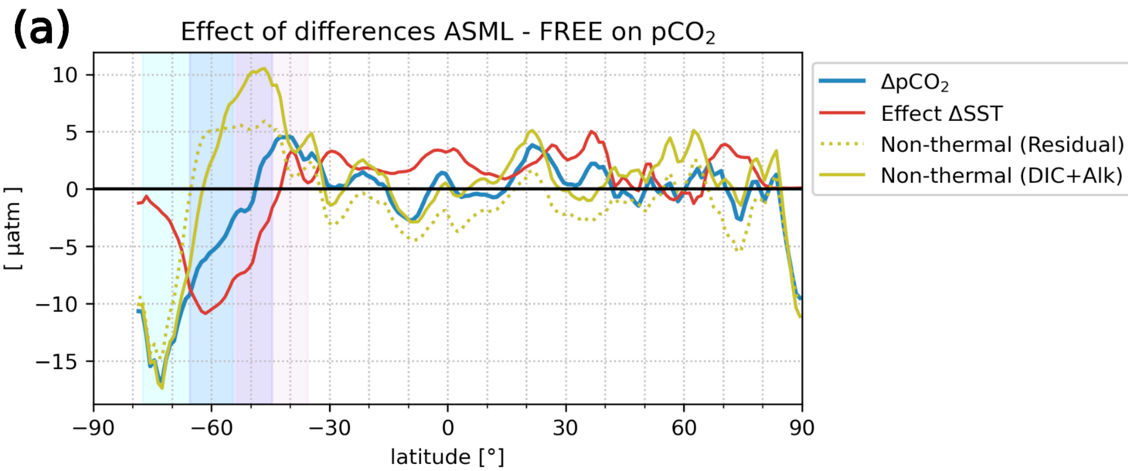
~~and surface chlorophyll~~physical and biogeochemical fields. In order to assess the ~~drivers of~~ dynamic DA effects on surface pCO_2, it is useful to distinguish between different variables that constitute the change in pCO_2. Oceanic pCO_2 varies mainly with temperature, DIC and alkalinity. Thus, we decompose changes in pCO_2 ~~are decomposed after the simulation~~ into their contributions from changes in SST~~(SST)~~, surface DIC ~~(DIC) and~~ and surface alkalinity (Alk)~~following the linear~~. For that, we apply the following approximations of Sarmiento and Gruber (2006) and Takahashi et al. (1993):

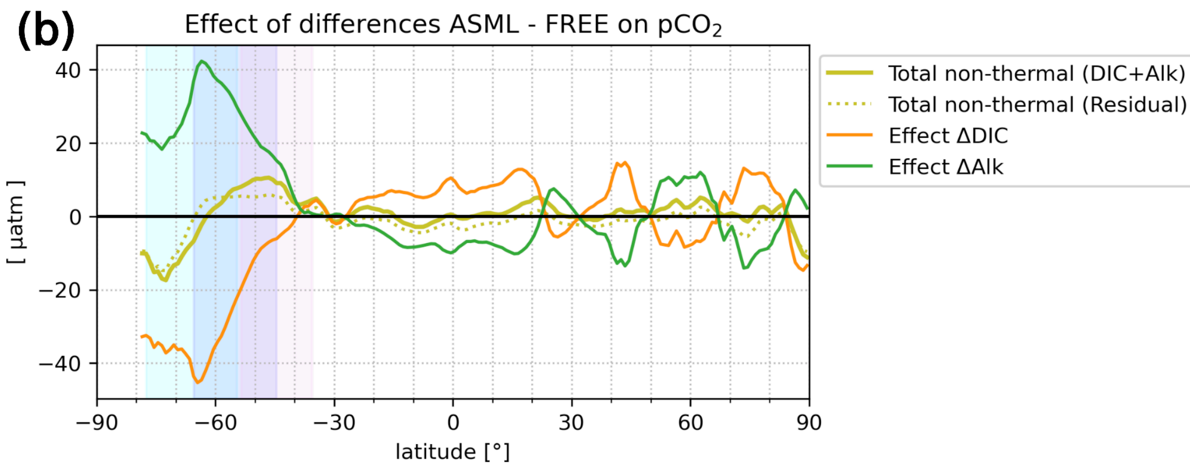
[equations]

the computation. The sensitivities γ_{DIC} and γ_{Alk} describe how pCO_2 varies with changes in one variable while keeping all other variables constant. For the sensitivities, we use an approximation derived from seawater carbonate chemistry following Sarmiento and Gruber (2006):

[equations]

In the appendix, we illustrate that the net pCO_2 difference (ASML – FREE; blue line in panel a) can approximately be explained by the sum of these three terms. Figure A9:





Here, the non-thermal effect is calculated, firstly, as the sum of alkalinity and DIC effects, and secondly as the residual (i.e. "net ΔpCO_2 minus thermal").

25 **Reviewer's comment:**

Eq 1, 2 and 3: Previously Delta denoted the difference between ASML and FREE, is this still the case here? If so, are the regular terms (e.g. DIC in Eq 1 or the terms in γ_DIC) from the FREE experiment? This should be mentioned in the description.

Answer:

Yes, delta is the difference between ASML and FREE and the regular terms are calculated from the average of the two simulations - this has been added here (Line 321):

~~with~~ Here, differences between ASML and FREE are denoted by Δ; else, the average of ASML and FREE is used for the computation. The sensitivities γ_{DIC} and γ_{Alk} describe how pCO_2 varies with changes in one variable while keeping the other

26 Reviewer's comment:

L 220: Why not mention EN4-OA earlier when the other data products are introduced?

Answer:

Makes sense, we have rearranged this. Firstly, all observational products that are assimilated are introduced in Section 2.3.1. Secondly, all observations used for validation are introduced in Section 2.4, here:

To evaluate the impact of ~~the DA on the modeled~~ DA on ocean physics, we compare the simulated SST and SSS to the assimilated observations (Section 2.3.1). For temperature and salinity at depth, we use the EN4-OA product (Good et al., 2013, updated to v . EN4-OA is an objective analysis ingesting the assimilated EN4 profile data, interpolated to global coverage on 42 depth levels. Furthermore, we compare the sea-ice concentration with remote sensing observations from OSI-SAF 2010-2020 (EUMETSAT, 2022) , the mixed-layer depthin the year 2020 with the profile-observation based climatology of de Boyer Montégut et al. (2004, updated version and the horizontal near-surface velocities 2010-2020 with the drifter-based climatology of Laurindo et al. (2017).

To evaluate the impact of the DA on biogeochemistry, we compare model outputs with ~~independent~~ observational datasets of surface pCO_2, DIC, alkalinity and surface chlorophyll. ~~For each observation type (OBS), we define the improvement as:~~

~~$$\text{improvement}_{OBS} = |\text{FREE} - \text{OBS}| - |\text{ASML} - \text{OBS}|$$~~

To evaluate surface pCO_2, we use observations from the Surface Ocean CO_2 Atlas (SOCAT Version 2023, Bakker et al., 2023, 2016), which are provided as a monthly gridded and quality-controlled compilation.

To assess DIC and alkalinity, we compare the modeled surface fields to the GLODAPv2.2023 bottle data (Lauvset et al., 2024b). At depth, we compare the model output to the GLODAPv2 DIC and alkalinity climatology (Lauvset et al., 2016), which is based on observations from the period 1972-2013 and normalized to 2002.

To evaluate global surface chlorophyll, we use observations from ESA-CCI, which is a multi-sensor satellite ocean-color chlorophyll-a dataset with monthly global coverage (Sathyendranath et al., 2021). In addition, for the Southern Ocean, we use

the mean of three satellite products (Johnson et al., 2013) that were processed with more suitable algorithms for southern high latitudes. For each observation type (OBS), we define the improvement as:

$$\text{improvement}_{OBS} = |\text{FREE} - \text{OBS}| - |\text{ASML} - \text{OBS}| \quad (9)$$

27 Reviewer's comment:

L 250: "at greater depth than 500 m, where the model's subsurface temperature": The "subsurface" can be deleted here.

Answer:

Thanks, Line 392:

> at greater depth than 500 m, where the model's ~~subsurface~~ temperature is colder

28 Reviewer's comment:

L 266: Please explain what a 15%-line is.

Answer:

See Lines 408-410:

> Sea ice reacts dynamically to the changed ocean physical state. In the Southern Ocean, FREE is characterized by a lower sea-ice concentration compared to OSI-SAF observations. The ~~maximum extentof~~ sea-ice extent, here defined as the area where the sea-ice ~~in September~~ concentration is more than 15%, reaches a maximum in September. The maximum extent is smaller in FREE than OSI-SAF, which is demonstrated by the 15%-~~lines~~ -line surrounding that area for FREE and OSI-SAF (Fig. 3a;

29 Reviewer's comment:

L 301: "In the more northern part of the STSS, which we call the STSS+, the CO2 uptake is reduced ...": The text here could be considered misleading because STSS+ is not defined as the northern part of the STSS, but as the part of the STSS with a positive CO2 flux difference. I would prefer a change in formulation that avoids this ambiguity, for example: "The part of the STSS characterized by a positive CO2 flux difference between ASML and FREE, which we call the STSS+ and in which the CO2 uptake is reduced, forms an outer (northern) ring around the STSS region." The same comment applies to STSS+ a few lines below.

Thank you for the suggested wording. We have used it:

The part of the $STSS_{SO}$ characterized by a positive CO_2 flux difference between ASML and FREE (positive difference: reduced uptake through assimilation), which we call the $STSS_{SO}$+, roughly forms an outer northerly ring around the $STSS_{SO}$ biome (hatched area in Fig. 5a and b).

and spring (June to November, Fig. 5c and d). ~~In the more northern~~ The part of the STSS~~, which we call the STSS+, the CO_2 uptake is reduced through the assimilation, demonstrated~~ $_{SO}$ characterized by a positive CO_2 flux difference between ASML and FREE ~~in this area (Fig. 5b). The reduction is greatest in winter and spring, which is shown through a positive flux differencebetween ASML and FREE from July to October (Fig. 5g). In contrast, in the more southern part of the STSS,~~ (positive difference: reduced uptake through assimilation), which we call the ~~STSS-, the assimilation increases the oceanic CO_2 uptake (Fig. 5~~$STSS_{SO}+$, roughly forms an outer northerly ring around the $STSS_{SO}$ biome (hatched area in Fig. 5a and b). The

30 **Reviewer's comment:**

L 373: "the effect of the DA is towards increased uptake of CO2 during boreal summer and autumn in ASML (Fig. 6g). This prevents summer outgassing": The increased summer uptake prevents summer outgassing, isn't this just describing the same effect? I would suggest rewording this sentence.

Answer:

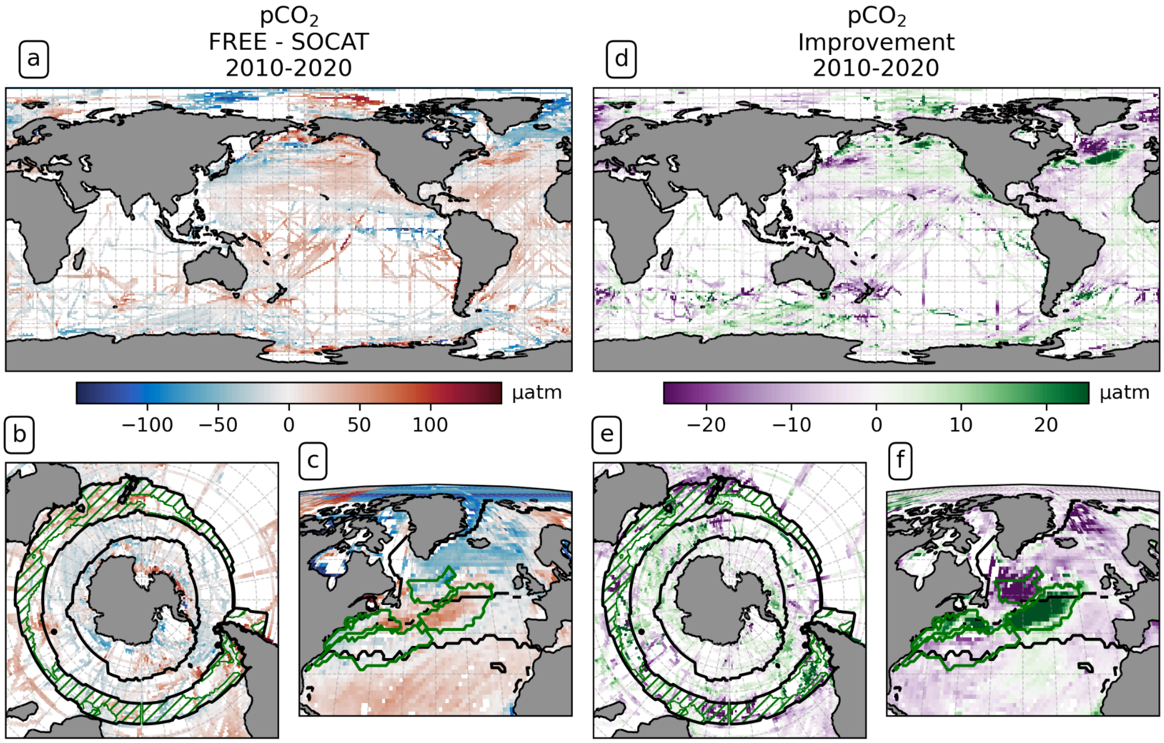
Reworded to emphasize the seasonal difference between uptake and outgassing:

Central $STSS_{NA}-$ In the Central $STSS_{NA}-$, the effect of the DA is ~~towards increased uptake of CO_2 during boreal summer and autumn in ASML~~ overall towards a more negative flux of CO_2 from May to November (Fig. 7g). ~~This prevents summer outgassing , which occurs from June to October in FREE~~ Thus, spring and autumn CO_2 uptake are increased and summer outgassing is prevented in ASML (Fig. 7c). The reason for ~~the higher uptake and~~ decreased surface pCO_2 is higher

31 **Reviewer's comment:**

L 411: "(difference of FREE and SOCAT in (Fig. 9a); difference of ASML and SOCAT not shown)": The figure label claims that ASML - SOCAT is shown.

Thank you for noting this. Figure data and labels have been updated to show FREE - SOCAT, as indicated in the text. Figure 9:



RC 02

32 **Reviewer's comment:**

The manuscript describes a study assimilating temperature and salinity observations into a global physics-biogeochemistry ocean model, with the aim of improving the modelled air-sea CO_2 flux. The assimilation brought the model temperature and salinity closer to the assimilated observations, and had a mixed impact on the carbon variables and wider biogeochemistry. The global mean change was small, but could be regionally significant, with the mechanisms explored.

The experiments are well conceived, and the manuscript generally well written and well presented. I just have some comments where aspects could use clarifying or expanding on.

L51: "Data assimilation (DA) has been employed …" This paragraph doesn't need to be comprehensive, but could be modified and expanded a little to more fully represent the available literature. Valsala and Maksyutov (2010, https://doi.org/10.1111/j.1600-0889.2010.00495.x) ran a global assimilation for 1996-2004; not multidecadal but almost as long as the present study. The paragraph states "In each of these studies, an Adjoint or Green's Function DA approach is used", but the Gerber et al. (2009) study referenced used an EnKF – another non-adjoint/Green's function example is While et al. (2012, https://doi.org/10.1029/2010JC006815) who used a sequential analysis correction scheme to assimilate pCO2. The paragraph opens by talking about "DA studies of the air-sea CO2 flux" in general terms, only semi-clarifying later that it's focussing on studies which directly assimilated pCO2 data. There have also been other studies which, like the present one, looked at the impact of assimilating other variables on the air-sea CO2 flux, e.g. Ciavatta et al. (2016; https://doi.org/10.1002/2015JC011496) and other papers from that group, and Ford and Barciela (2017, https://doi.org/10.1016/j.rse.2017.03.040).

Thank you for the references to the literature, great! We acknowledge these. The expanded paragraph reads:

> *Data assimilation (DA) can be employed to address the emerging discrepancies between pCO_2-products and models (Carroll et al., 2020). Several studies assimilating ocean surface pCO_2 have focused on specific regions (e.g., a baseline state of air-sea CO_2 fluxes in the Southern Ocean; Verdy and Mazloff, 2017), few years (e.g., optimized biogeochemical initial fields for the period 2009-2011 in Brix et al., 2015) or the climatological mean state (e.g., corrections of large-scale pCO_2 model biases in While et al., 2012). These studies capture well the assimilated pCO_2 observations, while obeying physical laws and biogeochemical (BGC) equations. Data assimilation also provides a better understanding of various components of the ocean carbon cycle, such as the transport of anthropogenic CO_2 in the ocean (e.g., a reconstruction of anthropogenic carbon storage since 1770 in Gerber et al., 2009), regional and interannual variability of the air-sea CO_2 flux (e.g., global reanalysis in Ford and Barciela, 2017; Carroll et al., 2020; Valsala and Maksyutov, 2010), the biological carbon pump (e.g., carbon export at a nutrient-rich and nutrient-poor site and estimation of BGC parameters related to air-sea CO_2 fluxes in Sursham, 2018; Hemmings et al., 2008) and specific ecosystems (e.g., the North West European Shelf ecosystem in Ciavatta et al., 2016, 2018). So far, however, there is no data assimilation product that provides a long-term, annually updated estimate of global ocean CO_2 uptake.*

33 **Reviewer's comment:**

L65: "It was shown that assimilating ocean physics at the initial state of a model simulation has a stronger and more positive impact on the modeled carbon cycle than assimilating the BGC initial state (Fransner et al., 2020)." In no way diminishing the motivation for this current study – which is undoubtedly important for the reasons stated in Fransner et al. (2020) and others – it could be clarified that this was a single model study and may or may not hold in general. The relative importance of physics vs biogeochemistry initialisation on different variables and time scales remains an open question – see e.g. the discussion in Section 4.4 of Lebehot et al. (2019, https://doi.org/10.1029/2019GB006186) and indeed the ultimate conclusions of this current manuscript.

Answer:

Thank you for providing the literature, which we have included:

~~the deep ocean in the form of dissolved inorganic carbon (DIC) (Davila et al., 2022). It was shown~~ CO_2 fluxes, it was shown in one idealized study that assimilating ocean physics at the initial state of a model simulation has a stronger and more positive impact on the modeled carbon cycle on interannual time-scales than assimilating the BGC initial state (Fransner et al., 2020). ~~Therefore the question arises which processes are most important when altered physics change CO_2 fluxes in DA approaches.~~ ~~We~~ However, the relative importance of uncertainties in physical and biogeochemical fields generally remains an open research question (e.g. Séférian et al., 2014; Li et al., 2016; Lebehot et al., 2019). Therefore, we here use ensemble-based data assimila-

34 **Reviewer's comment:**

L67: "Therefore the question arises which processes are most important when altered physics change CO2 fluxes in DA approaches." I think I understand the meaning of this sentence, but it could be reworded for clarity.

Answer:

Reworded for clarity to:

physics DA. The question therefore arises ~~as~~ to what extent an ecosystem model coupled to a data-assimilated physical model also represents a more realistic biogeochemistry~~. We will present cases where physics data-assimilation leads to worse and better agreement with BGC observations.~~ , and which mechanisms drive the response of the CO_2 flux in physics DA approaches. One possible driver is the physical transport of DIC and alkalinity because velocities and diffusivity are changed

35 **Reviewer's comment:**

L68: "to improve" – a better wording could be "to aim to improve"?

Answer:

Included in Line 90:

aiming to improve the modeled air-sea CO_2 flux

36 **Reviewer's comment:**

L75-79: The issues discussed by Park et al. (2018) and others, mentioned later in the manuscript, could be introduced at this point.

Answer:

We describe these issues now in the Introduction, instead of later in the manuscript:

An accurate representation of ocean physics is a prerequisite, but not necessarily sufficient for a realistic simulation of the CO_2 flux. Coupled ecosystem models are adapted to the associated physics Several difficulties are associated with physics DA into GOBMs. A common issue is erroneous equatorial upwelling leading to unrealistically high biological productivity in the tropics (Park et al., 2018; Gasparin et al., 2021; Raghukumar et al., 2015). Furthermore, any coupled ecosystem model is adapted to its associated physical model with its strengths and weaknesses through carefully selected parameter values . Furthermore, the natural carbon cycle in models is tuned to an equilibrium for the physical model state at pre-industrial conditions without DA, and it was shown that the and a spin-up to near-equilibrium. Accordingly, the modeled carbon cycle may react very sensitive to deviations from this physical state , leading 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 data assimilation (Spring et al physics DA. The question therefore arises as to what extent an ecosystem model coupled to a data-assimilated physical

37 Reviewer's comment:

L103: "Alkalinity is restored by a fictional surface flux of 10m/yr." Is there a reference for this, or was it introduced in this study?

Answer:

We follow the set-up of Gurses et al. (2023). This alkalinity restoring has been used by Hauck et al. (2013) and Schourup-Kristensen (2014) as well.

Gurses: doi.org/10.5194/gmd-16-4883-2023
Hauck: doi.org/10.1002/2013GB004600
Schourup-Kristensen: doi.org/10.5194/gmd-7-2769-2014

Citations added in Line 147:

pute pCO_2 and air-sea CO_2 flux, employing the gas-exchange parameterization of Wanninkhof (2014). Alkalinity is restored by a fictional surface flux of 10m/yr $10\,\mathrm{m\,yr^{-1}}$ (as in Hauck et al., 2013; Schourup-Kristensen et al., 2014; Gürses et al., 2023)

38 Reviewer's comment:

L121: "After each assimilation step, corrections are applied to the analysis state to ensure the consistency of model physics." Can you give an indication of whether these corrections need to be applied regularly or just occasionally?

Answer:

This has been clarified here:

temperature, salinity, horizontal velocities and sea surface height. After each assimilation step, corrections are applied to the analysis state to ensure the consistency of model physics: Salinity is set to a minimum value of zero and temperature to a minimum value of $-2\,°C$, if ~~necessary. The~~ the value is otherwise below. The increment of sea surface height (SSH) ~~update~~ is limited to two standard deviations of the ensemble. While in the simulation the correction was necessary for about 10% of SSH updates and 0.01 ‰ of temperature values at each step, the correction of salinity was never required. The analysis step is

39 **Reviewer's comment:**

L148: How is the weekly-resolution SSS used in the daily assimilation?

Answer:

SSS data is provided daily. To clarify, see Line 186:

The assimilated SSS data is taken from the European Space Agency (ESA) Sea Surface Salinity Climate Change Initiative (CCI) v03.21 data set (Boutin et al., 2021). ESA-CCI contains daily data at a spatial resolution of 50 km, albeit not capturing temporal variability below weekly. The ESA-CCI observations are averaged to the FESOM2.1 model grid. We prescribe a

The daily sampling of data resolving weekly variability is described in Boutin (2021): doi.org/10.1029/2021JC017676

It is not necessary that the observations capture the day-to-day variability, as the data assimilation has a comparatively slow effect: For example, it takes several months of assimilation to achieve the maximum feasible correction of a large-scale model bias.

40 **Reviewer's comment:**

L153: "model values are computed as the average of the grid points of the triangle enclosing" – what's done in the vertical?

Answer:

See Line 191:

The assimilated temperature and salinity profiles are taken from the EN.4.2.2 data set (Good et al., 2013). The EN4 dataset contains quality-controlled profiles from various in-situ ocean profiling instruments. To assimilate the profiles, the observations are assigned to the respective model layers (depth range) in the vertical. In the horizontal, the model values are computed as

41 **Reviewer's comment:**

L171: "For the comparison …" – this paragraph would benefit from a clearer explanation of what adjustments have been made to what products and why, including the model estimates from this study (which presumably have no river carbon inputs?).

Answer:

As both reviewers have asked for a clearer explanation, please see our answer to Reviewer's comment 21.

42 **Reviewer's comment:**

L206: "we define the improvement as" – I'm in two minds whether calling the statistic "improvement" is good as it's clear and intuitive, or if it should be more objective and phrased as "reduction in mean absolute difference" or something equally dry. On balance I'm happy how it is, given it's clearly defined, but will keep this comment here for completeness. It can be a little odd when positive and negative improvement gets discussed (e.g. L254, L258).

Answer:

The term 'improvement' was used before (see e.g. Losa et al., 2012: https://doi.org/10.1016/j.jmarsys.2012.07.008, with positive and negative improvements in Figure 1 and 2).

43 **Reviewer's comment:**

L220: "EN4-OA" – this is a reasonable product to use for comparison, but my understanding is that it includes no observations beyond the assimilated data, just interpolation between data points. So calling it "partly-independent" or "non-assimilated" (L244) may be misleading. Furthermore, it could have been introduced in the previous section.

Answer:

Thanks, we have adjusted the wording, saying that EN4-OA is an objective analysis ingesting the assimilated EN4 profile data. We have also changed the text structure so that all comparison datasets are described in one place. Please see our answer to Reviewer's comment 26.

44 **Reviewer's comment:**

L228: "in particularly" – in particular

Answer:

Thanks, Line 366:

FREE shows regional SST biases in ~~particularly~~ particular near strong currents or in eddy-rich regions,

45 **Reviewer's comment:**

L240: "particularly much" – "particularly"

Thanks, Line 381:

particularly ~~much~~ in the North Atlantic Central $STSS_{NA}$ –

46 **Reviewer's comment:**

L241: "Albeit negative side effects of temperature assimilation" – how is it judged that the temperature assimilation is responsible?

Answer:

We know from experiments during the test phase, assimilating only one variable at the time for a shorter period. Line 382:

Southern Ocean $STSS_{SO}$ (Fig. 1f). ~~Albeit negative side effects~~ Tests with the assimilation of temperature alone show negative side-effects of temperature assimilation on SSS in some locations (not shown). In the final set-up with combined assimilation, negative effects on SSS are found in 9% of the observed area~~, the global~~. Globally, the mean absolute difference is reduced

47 **Reviewer's comment:**

Fig. 1 and others: My instinct would be to plot ASML – OBS rather than ASML – FREE. However, I've argued about this with coauthors on papers before, and appreciate others strongly feel ASML – OBS is the better choice. So I'm merely flagging it as something to consider, I can see the argument both ways.

Answer:

We have chosen ASML - FREE because it allows us to visualize comparatively small changes in some of the biogeochemical variables. Please see our answer to Reviewer's comment 7.

48 **Reviewer's comment:**

L275: "see Appendix Text A1 for further discussion". Appendix Text A1 is a single short paragraph, I don't understand why it's in an appendix. It would be better in the main manuscript, either here or in the Discussion section.

Answer:

This paragraph has been expanded and is now included in the main manuscript (Section 3.1 Effect of DA on ocean physics):

> locities~~(see Appendix Text A1 for further discussion). The~~ . Throughout the assimilation period, spurious, spatially limited and often deep overturning structures emerge, evolve through several months or years, and disappear in the tropical Indian, Pacific and Atlantic basin (not shown). Thereby, the surface overturning cell sometimes breaks apart where it should extend over the equator, exposing the bottom cell to the surface (Fig. A8b). Transport in the North Atlantic at 26.5 °N, an indicator for the strength of the Atlantic Meridional Overturning Circulation, is between 8-9 Sv in FREE. In ASML, during the first two years of assimilation, transport at 26.5 °N decreases to below 3 Sv and, during the following years, recovers to 7-8 Sv (2016-2020). One possible cause is the effect of data assimilation on the eddy parameterisation (Gent and Mcwilliams, 1990). The parameterised eddy activity is relevant for the dynamics in the deep ocean, and corrupting it may have a negative impact on the large-scale oceanic circulation, as described in Sidorenko (2004, Chapter 5.5 onwards) for a previous version of the ocean model FESOM.

49 **Reviewer's comment:**

L276: "Thus, it can be assumed that the velocities in the upper part of the ocean are also well represented." I don't think you can make this assumption, certainly not for vertical velocities. See e.g. Raghukumar et al. (2015, https://doi.org/10.1016/j.pocean.2015.01.004) and Gasparin et al. (2021, https://doi.org/10.1016/j.ocemod.2021.101768). The data assimilation will continually update the observed variables to better match the observations, without necessarily leading to improvements in non-observed variables such as velocities – although of course that's the aim. The current study certainly doesn't seem to have the issues with vertical velocities the above studies do, but without providing assessment of the wider circulation there's no guarantee it's improved.

Answer:

We agree that there is no guarantee that it's improved and have therefore rephrased: "This can be interpreted *as an indication* that the velocities in the upper part of the ocean are also well represented."

This indication becomes more reliable, though, through additional evaluation of horizontal surface velocities and mixed-layer depth:

~~productivity in the tropics (Park et al., 2018). In FESOM-REcoM, the~~ The boundary-layer depth and mixed-layer depth are mostly reduced through DA. In particular, 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). This improves the agreement with the profile-observation based mixed-layer climatology of de Boyer Montégut et al. (2004), reducing the mean absolute difference to the climatology from 27 m to 19 m (comparison of mixer-layer depth in Fig. A6). In addition, the absolute difference of near-surface horizontal

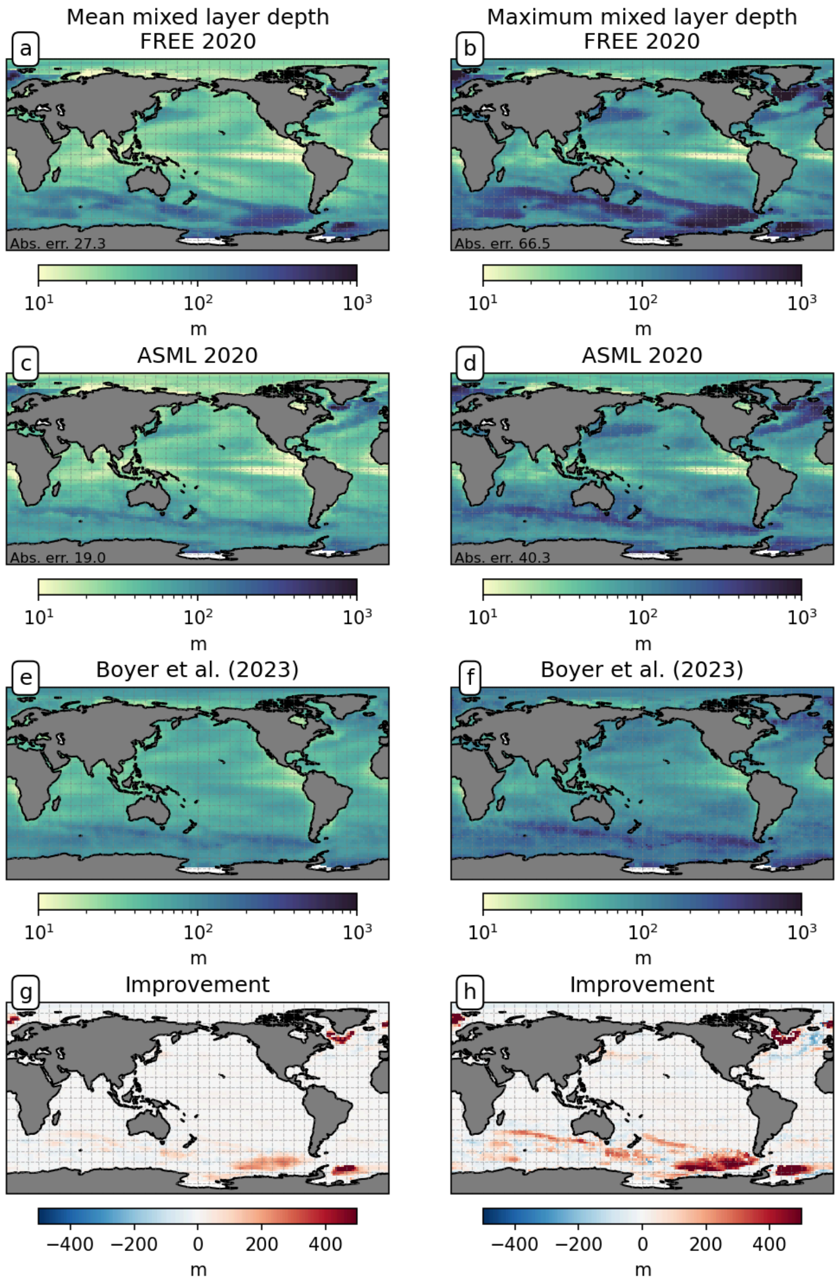
velocities to the drifter-observation based climatology of Laurindo et al. (2017) is reduced by about 10% through DA (comparison of surface velocities in Fig. A7). The biological productivity near the equator is stable in ASML and FREE, indicating that FESOM-REcoM does not suffer from the erroneous upwelling known from previous DA studies (Park et al., 2018). The

…

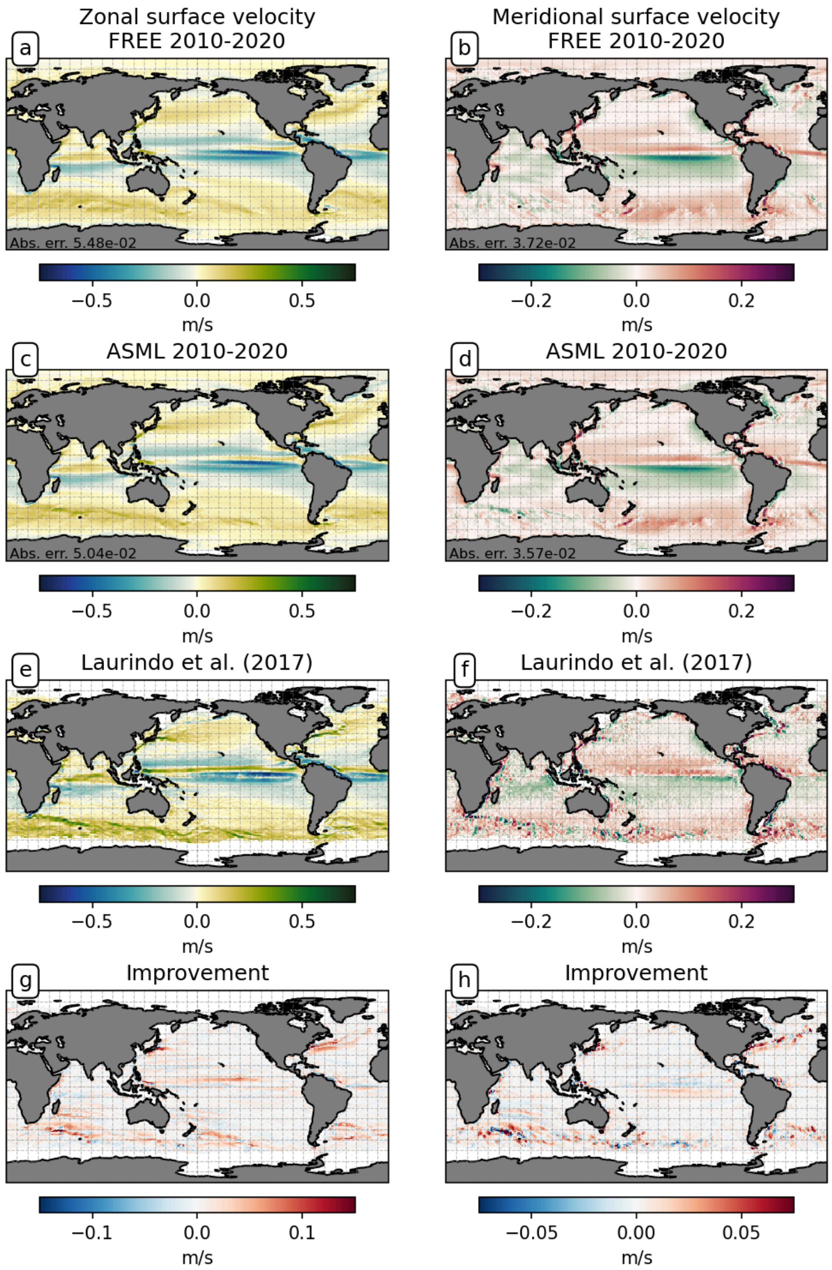
In summary, the ASML temperature and salinity fields ~~at and near the surface in ASML~~ from the surface to several hundred meters below, and mixed-layer depth are in good agreement with ~~the observations. Thus, it can be assumed that the velocities~~ observations, and the agreement of horizontal near-surface velocities with observations is improved. This can be interpreted as an indication that the velocity field in the upper part of the ocean ~~are~~ is also well represented. ~~Therefore~~Although the spurious effects on deep ocean circulation should be further addressed in future work, we are confident that the DA provides an improved physical state in the upper ocean, which serves as an improved basis to estimate the air-sea CO_2 flux~~, although the spurious effects on deep ocean circulation should be further addressed in future work~~.

We show the comparison of mixed-layer depth and horizontal velocities in the Appendix.

Mixed layer in Figure A6



Horizontal surface velocities in Figure A7



50 **Reviewer's comment:**

L280: "4 Results" – Section 3, "Effect of DA on ocean physics" is also results. Perhaps Section 4 should be "Effect of DA on ocean biogeochemistry".

Answer:

We have adjusted the section titles based on your suggestion. For the structure of the revised manuscript with all sections and subsections, please see the table of contents in our answer to Reviewer's comment 6.

51 **Reviewer's comment:**

L282: "The ocean absorbs 2.78 Pg C dec^{-1}" – is this the correct unit? From Fig. 4a, it looks to be absorbing 2.78 Pg C yr^{-1} on average over the decade.

Answer:

Thank you for having taken a closer look. Indeed, this was a typo and is now fixed in this and several other places, e.g.:

> The ocean absorbs ~~2.78 Pg C dec^{-1}~~ 2.78 Pg C yr^{-1} in ASML and ~~2.83 Pg C dec^{-1}~~ 2.83 Pg C yr^{-1} in FREE during 2010-2020 (Fig. 4b), thus the assimilation decreases the global mean oceanic CO_2 uptake by ~~0.05 Pg C dec^{-1}~~ 0.05 Pg C yr^{-1}.

52 **Reviewer's comment:**

L290: "air-sea CO2 flux (negative: into the ocean)" – if negative's into the ocean shouldn't it be "sea-air CO2 flux"?

Answer:

While the direction of air-sea CO_2 flux is not uniformly defined in the literature, the term 'air-sea' is commonly used for both for some reason, see e.g. Global Carbon Budget (Friedlingstein et al., 2023): 'air-sea flux' is positive into the ocean; and Roobaert et al. (2023): 'air-sea exchange' is negative into the ocean.

By defining outgassing as positive, the direction of CO_2 flux corresponds to the pCO_2 effect: Higher oceanic pCO_2 values result in a more positive flux.

53 **Reviewer's comment:**

L301: While STSS+ is broadly the northern bit and STSS- southern, it's a bit more nuanced than that and that should be reflected in the text.

Answer:

We have rephrased this (giving credits to the other reviewer's suggestions). Line 479:

> $STSS_{SO}$+, roughly forms an outer northerly ring around the $STSS_{SO}$ biome (hatched area in Fig. 5a and b).

and Line 527:

SO characterized by a negative CO_2 flux difference between ASML and FREE, which we call the $STSS_{SO}-$, is a fragmented region and roughly consists of segments of an inner southerly ring (non-hatched area in Fig. 5a and b). ~~In addition, reduced~~

54 **Reviewer's comment:**

Fig. 5: Add to the caption that the lines in a and b denote the regions, and the hashing (striping?) denotes STSS+.

Answer:

Added to the captions of figures 5 and 7:

Figure 5. Effect of data assimilation on Southern Ocean CO_2 flux and its seasonality averaged over the period 2010-2020. Negative numbers indicate a flux into the ocean. Additionally, lines in a and b denote the regions, and the green hatching denotes the $STSS_{SO}+$. (a) Map of

55 **Reviewer's comment:**

L462: "a pCO2-independent proxy for primary production" – I'm not sure "pCO2-independent" is needed here, I don't quite understand what's meant.

Answer:

We agree that it is not needed here. Line 726:

The representation of chlorophyll by the model is of interest as a ~~pCO_2-independent~~ proxy for primary production.

Originally, we meant to point out that there is no direct relationship of chlorophyll and pCO_2 through the carbonate chemistry of seawater - unlike for all other variables (T, S, DIC and Alk) that are included in the observation comparisons.

56 **Reviewer's comment:**

L480: "as the modelled phytoplankton growth is temperature-dependent" – how sure are you the change is due to the direct temperature dependence rather than the indirect influence of stratification and mixing changes?

Answer:

We cannot separate these effects and have therefore rephrased the text:

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 (Section 3.3). Surface chlorophyll changes follow SST changes (~~Fig. 11 and Fig. 1), as the~~ Figs. 1 and 11). The modeled phytoplankton growth is temperature-dependent (Gürses et al., 2023). Furthermore, indirect temperature effects on plankton dynamics due to stratification and mixing changes contribute, albeit those can have heterogeneous effects and the correlation of chlorophyll and boundary-layer depth is less clear (not shown). The changes of surface DIC and alkalinity

As the link between sea surface temperature and mixing is not straight-forward, the temperature-dependence of growth is a more likely candidate to explain the similar spatial patterns of SST and chlorophyll changes (Figure R8).

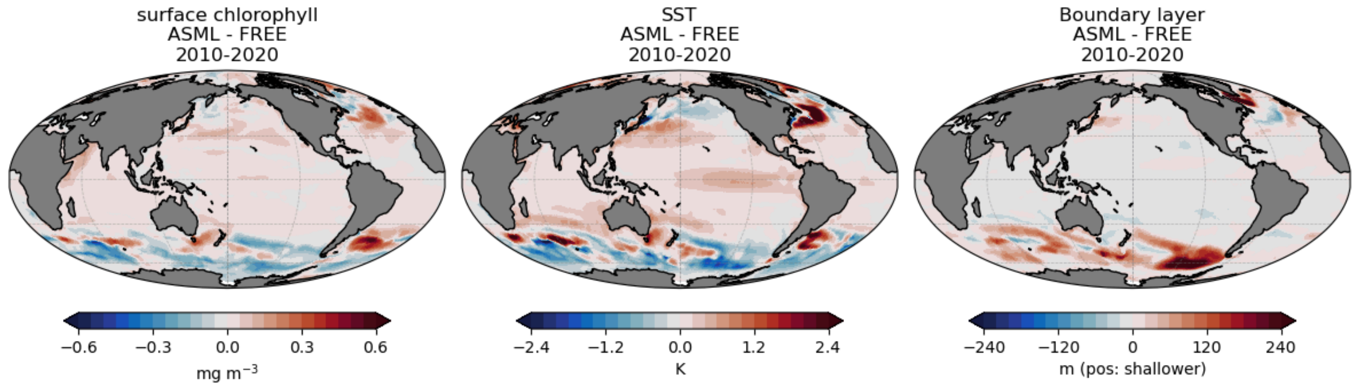


Figure R8: Spatial patterns of the difference ASML-FREE for surface chlorophyll, SST and boundary layer depth.

57 **Reviewer's comment:**

L515: "There are two other data assimilating BGC model approaches" – there are many other data assimilating BGC model approaches! Perhaps a more accurate phrasing might be: "We compare here to two other data assimilating BGC model approaches …"

Thank you for the rephrasing suggestion, we used it (Line 833):

~~There are~~ We compare here to two other data assimilating BGC model approaches, namely ECCO-Darwin (global; Carroll et al., 2020) ~~)~~ and B-SOSE, which is restricted to the Southern Ocean (Verdy and Mazloff, 2017). Both approaches use

58 **Reviewer's comment:**

L524: "suggesting that a flawed representation of ocean physics as an argument for the models underestimating the CO2 flux trend is unlikely" – I broadly agree, though it may depend on how well the wider circulation is represented.

L559: “suggests that the physical processes are already well represented in FREE” – again I broadly agree, but there may still be pertinent limitations, especially depending on the time and space scale.

Answer:

We agree with the reviewer that there are limitations. The revised Discussion no longer contains these statements (at least not verbatim). Furthermore, because “the free running model already represents temperature and salinity *rather well*” is a subjective assessment, we have also reworded the abstract:

> over the period 2010-2020 to study the effect on the air-sea CO_2 flux and other biogeochemical variables. ~~While the free running model already represents~~ The assimilation nearly halves the model-observation differences in sea surface temperature and salinity~~rather well, the assimilation further improves it and hence influences the~~ , with modest effects on the modeled ecosystem and CO_2 fluxes. The ~~assimilation has mainly regional~~ main effects on the air-sea CO_2 flux ~~, with the~~ occur on small scales in highly dynamic regions, which pose challenges to ocean models. The largest imprint of assimilation is in the Southern

59 **Reviewer’s comment:**

L565: “the adjustment of the ocean’s carbon cycle to changes in the circulation” – true, though it’s also possible that this might itself introduce biases in the carbon chemistry. See e.g. Lebehot et al. (2019, https://doi.org/10.1029/2019GB006186).

Answer:

We acknowledge that changes in the circulation may lead to imbalances of the ocean’s carbon cycle, in particular during the adjustment phase, which may however take hundreds of years. The corresponding paragraph now reads:

> than changing the global mean SST, which differs by only 0.02 °C between FREE and ASML. DA-induced differences in vertical transport of DIC are comparably large south of 50 °S, but approximately 95% of them are balanced globally by opposing changes in vertical transport further north (vertical transport of DIC in Fig. A13a). In particular, the effect of DA on subduction of DIC through vertical advection into the ocean’s deeper layers (not shown), which is the rate-limiting step on oceanic uptake of anthropogenic CO_2 emissions (DeVries, 2022), appears small, which may be due to an insufficient amount of deep observations. Besides, experiments on longer time scales might be necessary to generate a visible effect of deep circulation changes on the ocean’s carbon cycle (Cao et al., 2009), which could however lead to imbalances in the CO_2 flux (Lebehot et al., 2019; Kriest et al., 2020; Primeau and Deleersnijder, 2009). Another possible reason why the DA effect on the