1 Response to reviewer 1

General Comments

I read this with great interest, as the author brings together a number of essential points regarding the ways models are used with estimating marine carbon cycle fluxes and budgets. The critical point here is that GOBMs play a central role in providing "estimate" of ocean carbon uptake, but they suffer from some known limitations and caveats which may result in biases. In my opinion this "letter manuscript" warrants concerted attention and further review, in my opinion it is very close to meeting the journal standard for publication.
[...]
But overall I think that this is a very valuable discussion, it's very well-reasoned and represents a dose of constructive reflection, and should also motivate some careful thinking about how to improve the way in which models are applied to estimate carbon uptake by the ocean. Overall the scientific concept of using coupled models in conjunction with GOBMs is a very constructive recommendation, and in fact at the very least this approach deserves inclusion in the GCP-type analyses of marine carbon uptake.

Response:

We thank the reviewer for their positive evaluation and the additional constructive and helpful comments, which have substantially improved our manuscript. We have taken each comment into account, provide responses to each point below, and adapted the manuscript accordingly.

Comment 1.1

I think it would be greatly beneficial if the author could consider whether ensemble simulations with ESMs could provide a means to get a more useful ESM-derived component of this story, as the ensemble-mean approach offers the real forced-trend (assuming enough members).

Response:

As suggested by the reviewer, I had indeed considered using large ensembles, such as the CESM2-LE (Rodgers et al., 2021). In the revised manuscript, I now include that consideration as suggested by the reviewer:

"Instead of using several different ESMs, one could also have used large ensembles such as the GFDL-ESM2M large ensemble (Burger et al., 2022) or the CESM2 large ensemble (Rodgers et al., 2021). The advantage of such an ensemble with 30-50 ensemble members is indeed that the forced trend can be better isolated (Li and Ilyina, 2018). However, the disadvantage is that one model can have systematic biases in all its ensemble members. CESM2, for example, appears to have a strong low bias in the Southern Ocean sea surface salinity, and hence the creation of mode and intermediate waters (Terhaar et al., 2021). The underestimation of the mode and intermediate water then leads to too little carbon uptake in the Southern Ocean, with a strong global imprint (see Figure 3 in Terhaar et al., 2022). The multi-model approach, on the other hand, has the distinct advantage that systematic biases in single models are averaged out to the best extent possible. Moreover, the use of these different models even allows to correct the multi-model estimate for biases that remain within the entire model ensemble (Terhaar et al., 2022). The disadvantage of the multi-model approach is that it has a smaller number of simulations so that the true forced trend may not be completely isolated, and some random variability may remain. However, as a spline fit is later fitted to the earth system model output, that variability is largely removed as well. Although not all internal variability might be filtered out entirely by the multi-model ensemble, I use this multi-model ensemble to avoid biases in the magnitude of the long-term carbon sink that could arise from using a large ensemble of a single model."

Li, H. and Ilyina, T.: Current and future decadal trends in the oceanic carbon uptake are dominated by internal variability, Geophys. Res. Lett., 45, 916–925, https://doi.org/10.1002/2017GL075370, 2018.

Rodgers, K. B., Lee, S.-S., Rosenbloom, N., Timmermann, A., Danabasoglu, G., Deser, C., Edwards, J., Kim, J.-E., Simpson, I. R., Stein, K., Stuecker, M. F., Yamaguchi, R., Bódai, T., Chung, E.-S., Huang, L., Kim, W. M., Lamarque, J.-F., Lombardozzi, D. L., Wieder, W. R., and Yeager, S. G.: Ubiquity of human-induced changes in climate variability, Earth Syst. Dynam., 12, 1393–1411, https://doi.org/10.5194/esd-12-1393-2021, 2021.

Terhaar, J., Frölicher, T., and Joos, F.: Southern Ocean anthropogenic carbon sink constrained by sea surface salinity, Sci. Adv., 7, 5964–5992, https://doi.org/10.1126/sciadv.abd5964, 2021.

Terhaar, J., Frölicher, T. L., and Joos, F.: Observation-constrained estimates of the global ocean carbon sink from Earth system models, Biogeosciences, 19, 4431–4457, https://doi.org/10.5194/bg-19-4431-2022, 2022.

Comment 1.2

On a related point, there is the issue of volcanoes, and the question of whether one risks double-counting something like Pinatubo by combining ESMs and GOBMs in the way described in the manuscript. Connecting this to my previous point, if I understand correctly Faye et al. (2023; GBC) ran a new ensemble with CESM1 without volcanoes, so I'm wondering if using an ensemble mean from such a set of runs and then combining this ensemble mean output with GOBM output as described in the manuscript would provide a way to avoid double-counting?

Response:

As mentioned by the reviewers, volcanoes are indeed part of the forcing sets of ESMs and GOBMs. The CESM1 ensemble (Fay et al., 2023) was indeed an excellent suggestion by the reviewer that allowed to properly assess the possible double-counting of volcanoes. For this assessment, I calculated the ensemble mean flux for the simulation with and without the eruption of Mt. Pinatubo (solid lines in Response Figure 1a). Afterwards, I calculated for each of these timeseries the long-term trend (dashed and dotted lines in Response Figure 1a) to separate each of these two timeseries into their long-term trend and short-term variability. To do so, I used the Enting spline with a cut-off period of 15 years, exactly the same way the ESM and GOBM timeseries were separated into a long-term trend and the short-term variability. The difference between the long-term trend with and without Mt. Pinatubo and the difference of the short-term variability with and without Mt. Pinatubo allow to quantify the effect of the eruption of Mt. Pinatubo on the long-term trend (dashed line in Response Figure 1b) and the short-term variability (dotted line in Response Figure 1b).

Mt. Pinatubo has a large effect on the short-term variability component of the ocean carbon sink but also a non-negligable effect on the long-term trend. As both, ESMs and GOBMs, are forced with volcanic activity (CMIP6 ESMs only until 2014), the long-term trend from ESMs includes the long-term effect of Mt. Pinatubo (and other volcanic activity) and the short-term variability from GOBMs includes the short-term variability from Mt. Pinatubo (and other volcanic activity). By adding both components, the overall effect of Mt. Pinatubo (Fay et al., 2023) are both accounted for in the composite estimate and neither the long-term effect nor the short-term variability is counted double. If, however, I had used the CESM1 ensemble simulations without Mt. Pinatubo, the long-term trend from this ensemble would not be included and the composite estimate would be off by up to 0.05 Pg C yr^{-1}. The following sentence was added for clarification to the revised manuscript:

"As GOBMs have external forcing, such as volcanoes, prescribed and ESMs have it prescribed until 2014, the end of the historical period in CMIP6, the long-term effect of volcanoes is part of the long-term ESM estimate and the short-term variability due to volcanoes is part of the short-term variability from GOBMs. As the same spline and cut-off period is used to separate long-term trend and short-term variability in GOBMs and ESMs, there is no double counting of the overall effect of volcanoes (Fay et al., 2023)."

Fay, A. R., McKinley, G. A., Lovenduski, N. S., Eddebbar, Y., Levy, M. N., Long, M. C., et al. (2023). Immediate and long-lasting impacts of the Mt. Pinatubo eruption on ocean oxygen and carbon inventories. Global Biogeochemical Cycles, 37, e2022GB007513.

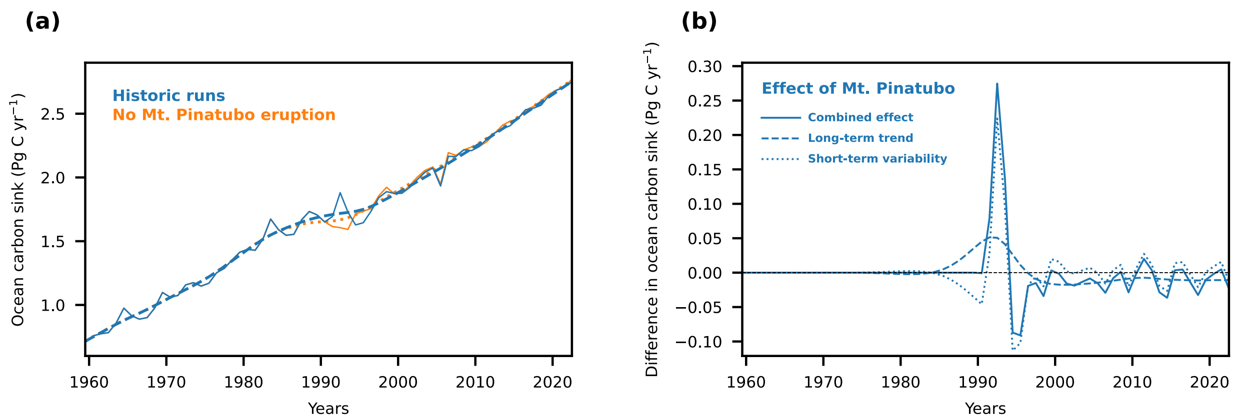

Figure 1: **The effect of Mt. Pinatubo on the long-term trend of the ocean carbon sink.** **(a)** The simulated ocean carbon sink in the CESM1 large ensembe (Fay et al., 2023) with (blue solid lines) and without (orange solid lines) the eruption of Mt. Pinatubo as well as the long term trends of both timeseries (dashed blue with Mt. Pinatubo and dotted orange line without Mt. Pinatubo) that were calculated with an Enting Spline and a cut-off period of 15 years. **(b)** The effect of Mt. Pinatubo calculated as the difference between the timeseries of the ocean carbon sink simulated by CESM1 with and without Mt. Pinatubo (solid line) as well as the long-term trend component (dashed line) calculated as the difference of the Enting splines of both ocean carbon sink estimates, and the short-term variability (dotted line) calculated as the difference of the overall effect and the long-term trend.

Comment 1.3

Another point that should be addressed is seasonality and missing mechanisms. Both GOBMs and ESMs suffer from deficiencies in representing the seasonal cycle in pCO2, and as has been pointed out by Fassbender et al. (2022; GBC) there is a rectified effect of seasonal pCO2 variations onto the mean state. To the extent that biases in the seasonal cycle of pCO2 should thereby have an impact on the rate of uptake of CO2, this cannot be remedied by a hybrid model.

Response:

Following comments 1.3 and 1.4 from the reviewer, I have added the following sentences to the revised manuscript:

"Furthermore, the composite estimate cannot remove shortcomings or uncertainties that are inherent to both GOBMs and ESMs and that were not accounted for by emergent constraints (Terhaar et al., 2022), such as the often incorrect representation of the seasonal cycle of pCO2 in both model classes (Rodgers et al., 2023; Joos et al., 2023). As it has been shown that the seasonal cycle changes in the future will affect the strength of the ocean carbon sink by 8% until 2100 under a high-emission scenario (Fassbender et al., 2022), an incorrect representation at present likely also affects the simulated ocean carbon sink by ESMs and GOBMs and hence by the composite estimate. Other processes that are also still not at all or not accurately simulated in GOBMs and ESMs and that might affect the ocean carbon sink are, for example, the ocean biological carbon pump (Doney et al., 2024; Laufkötter et al., 2015) or the land-ocean aquatic continuum (Séférian et al., 2020; Terhaar et al., 2024). Although improvements have been made in the past to account for these model weaknesses (Dinauer et al., 2022; Archibald et al., 2019; Lacroix et al., 2020; Terhaar et al., 2021a), more research is needed to improve simulated estimates of the ocean carbon sink."

Comment 1.4

With this last point (seasonality) it would be good if the author could state in a sentence or two that there are fundamental "missing processes" in current models that won't be fixed by building a hybrid product, that require further community attention.

Response:

Please see answer to comment 1.3.

Comment 1.4

A more minor point with Line 19: Didn't Ernst Maier-Reimer investigate anthropogenic carbon uptake before Sarmiento (1992)?

Response:

As suggested by the reviewer, the following reference was added to line 19:

Maier-Reimer, E., Hasselmann, K. Transport and storage of CO2 in the ocean — an inorganic ocean-circulation carbon cycle model. Climate Dynamics 2, 63–90 (1987). https://doi.org/10.1007/BF01054491