Manuscript: Substantial inter-model variation in OAE efficiency between the CESM2/MARBL and ECCO-Darwin ocean biogeochemistry models

Review round 2 comments and replies.

We thank both reviewers for the thorough review and all the constructive comments !

Our responses in blue below:

Reviewer #1

I thank the authors for the revisions and for addressing the reviewer comments. The responses to the reviewer comments are clear, and the revisions made to the manuscript have improved the overall quality and clarity of the manuscript.
I only noted a minor typo on legend of Figure 6 (dolid instead of solid) that should be corrected.

Fixed. Thank you for spotting that.

Reviewer #2

Overall, I think the paper is improved since I last read. I continue to think it is a valuable and timely contribution to the ocean biogeochemistry and mcdr literature and should be published. Mostly, my comments have been addressed very nicely, but the revised manuscript and response to review has not adequately responded at a few points. So, I recommend another round of minor revisions. I don't necessarily need to review it again.

My comments below are comment-by-comment regarding the response to reviews and then line-by-line in the revised manuscript:

--Table 1. “We added a more extensive discussion of the features of the Darwin biology model, its dependence on the carbonate state and processes not modeled here” Maybe I missed something, but I did not see it.

It appears that during the final editing process we ended up summarizing these features in the table with references to the relevant papers and this discussion was shortened more than intended. With respect to the biological ablation on ECCO-Darwin we’ve added back some more detailed description on the relative dependence of biological growth rates, DIC and Alk, and circulation, as follows:

“The ECCO-Darwin ecology includes five phytoplankton function types (diatoms, other large eukaryotes, Synechococcus, and low- and high-light adapted Prochlorococcus) and two zooplankton types with different preferential grazing behavior. (Brix et al., 2015). The

biological rates are driven by light, temperature, and macro/micronutrients (nitrogen, phosphorus, iron, and silica) but do not explicitly depend on DIC and Alk. Conversely, the circulation does not depend on the tracers."

L150 of the revised manuscript: Are the biological rate equations explicitly dependent on DIC and Alk concentrations in MARBL and Darwin?

In Darwin, DIC and Alk influence the system thermodynamically (through carbonate chemistry) but they do not directly appear as inputs in the biological rate equations themselves.

--"we believe this negative result is worth publishing" Ok, but you need to include the context to interpret these results (i.e., actually make changes to the manuscript following the previous comment above).

We added the following text to Section 2.5:

"While biological processes modeled in ECCO-Darwin consume or produce CO_2 and/or Alkalinity (through carbonate shell creation or dissolution), the growth rates are not explicitly coupled to DIC or Alk. We therefore reasoned that the majority of the biogeochemical model should have very little, if any, influence on the CO_2 uptake curves, provided the background state of the DIC gradients (which biology helps establish) are present at the start of the simulation. If true, a significant fraction of computational cost could be saved in future simulations.

To test this hypothesis we created an ablated version of the ECCO-Darwin in whic…"

--"it is not necessary to simulate the processes that maintain them for short-medium length simulations during which the OAE pulses equilibrate"
Related to L155-156 in revised manuscript: "(plus an unperturbed reference run)"

I agree, but you need to emphasize more at the beginning that it is necessary to run an extra OAE counterfactual or reference simulation branch as well as an OAE perturbation branch off of each baseline simulation with biology. In principle, this needs to be done for each virtual OAE deployment time. Running a new physics simulation has significant cost, and the reference simulations can become costly if different times are being considered requiring multiple reference simulations as well.

Yes, it is true that for each chosen pulse injection time a biology-free reference simulation has to be run. However the savings in compute cost when not simulating the biology tracers is so significant, that the extra reference run is easily worth it if more than one OAE perturbation branch is going to be simulated from that time point.
For example, in the OAE atlas work (Zhou et al., 2024), many thousand different OAE simulations are started, at only 4 different starting times, which would have necessitated only 4 additional reference runs where biology tracers aren't advected. The overall compute savings would have been substantial.

We added the sentences “We note that for each deployment time a no-biology reference simulation must also be branched off the main simulation. However, if many OAE perturbations are being tested at different locations (as in Zhou et al.,2024) the computational cost savings can be substantial overall.”

I think the computationally cheapest (but perhaps not easiest) thing to do would be to run the OAE dynamics with “offline” physical and biological state since there is no feedback between the DIC and Alk perturbations of OAE and the biological or physical state variables in ECCO-Darwin or MARBL or the feedback is very weak.

Related: In Fig S3, I understand these are small deviations, but aren’t the deviations large compared to the OAE anomaly, thus you do need to run 2 experiments without biology: one with OAE and one without OAE to form the altered counterfactual/reference.

Yes, that’s correct.

In Fig 12, I think you should mention that the reference/counterfactual simulation differs between the two for clarity.

We have added “The η(t) were also calculated against a separate reference simulation, likewise without biological processes.”

--“The partial derivative … in equation 2 is also unitless, since it’s concentration divided by concentration” and later on “The partial derivatives … = eta_max are both unitless quantities.

I don’t quite follow. Are you saying “mols of anything” is the same unit, regardless of the constituent under consideration? I think the text should be more explicit about distinguishing the different constituents involved. I guess you are implicitly using mols C/kg-seawater for both numerator and denominator or just mols/kg or mols-eq/kg of Alk? There are a lot of different entities here, e.g. DIC includes carbonate, bicarbonate.

Are you including carbonic acid with [CO2]?

Yes, this is very commonly done as the concentration of actual carbonic acid is so tiny (<0.3% of CO2(aq)) that for all intents and purposes we can treat CO2 and H2CO3 together as one species.
Some authors denote the combined concentration as [CO2*], [CO2] or [H2CO3*].
Here we follow the convention used by Zeebe and Wolf-Gladrow where they use the symbol [CO2] = [CO2(aq)] + [H2CO3]
However we added a reminder for the reader that [CO2] represents the sum of the two individual chemical species:

“(Note that throughout this paper, the [CO_2] includes both dissolved CO_2 and undissociated carbonic acid H_2CO_3 following the convention used by Zeebe et al., 2001)”

Specifically, what are the units of [HCO3-]+2[CO32-] in the numerator of the second term?

All concentrations ([HCO3-] and [CO32-], [DIC] etc) in this paper have units of mol/L, therefore the units of the sum [HCO3-]+2[CO32-] are also in mol/L. In the literature, alkalinity is sometimes referred to as mol equivalents per L, but mathematically it is strictly defined here as a sum of concentrations [Alk] = [HCO3-]+2[CO32-]+[OH-]-[H+]+....
and therefore also, technically speaking, has units of mol/L, even though there is no simple chemical species called "alkalinity". These definitions are given in the Supplement.
Even though all these terms refer to different chemical species or represent sums of concentrations they all share the same units. A ratio of DIC to Alk is therefore unitless even though the denominator and numerator refer to different sums of different species. The same applies to a partial derivative or the Revelle factor.

--"section 3.8…much easier to follow. The presentation is now a single figure" and L610 in the revised: "Overall, the largest contributor was found to be the mixed layer depth…" and Fig. S2:
Section 3.8 is easier to follow now, but I have a hard time quantitatively evaluating the importance of MLD differences between models vs other differences between models based on the figures in the main manuscript, since Fig 13 doesn't include MLD effects.

It would be great to put the deBoyer Montegut MLD from Argo floats as black lines on Fig S2 for comparison https://mixedlayer.ucsd.edu (they have the same deBoyer Montegut density threshold 0.03 kg/m3 in the gridded monthly dataset – see Holte et al. 2017 GRL).

Great suggestion! We've added the ARGO data as suggested in Figure S2 as well as prose: "A direct comparison of the seasonal MLD for both models, for each location, is shown in Figure S2, together with MLD data from ARGO floats (Holte et al., 2017). In general we observe that the MLDs in ECCO-Darwin agree more closely with the ARGO data than the CESM2/MARBL model."

--L1 Abstract: "dissolved inorganic carbon (DIC) deficit" I think using pCO2 is better here, and more consistent with the Introduction.
Changed to "Induction of a CO_2 partial pressure (pCO_2) deficit in…"

--L25: "temperatures" -> "temperature changes"
Thank you for spotting this. We did not mean to suggest that IPCC is intending to stabilize the global average temperature at 2°C.

--L74, L85 or elsewhere regarding interannual variability: Please give some rationale for the choices of different years and some insight into whether those years might be representative of particular climate mode or ocean dynamical differences. Or at least discuss whether the choices are sufficient to represent the typical interannual variability.

The year 1999 was chosen to match the year used by the Zhou et al. study, as indicated (L74). Likewise the years 2000, 2003, 2006, 2009, and 2012 were chosen to match those from Yankovsky et al. (L85). Neither of these were free choices therefore.
The only choice we made here was to add the year 1992 to supplement the 15yr simulation in 1999. It was not chosen for particular climatological reasons but simply because it

represents the earliest year available for ECCO-Darwin data assimilation and therefore offers the potential for the longest continuous simulation without resorting to looping of the forcing. We ended up not doing that but wanted to leave that door open for future studies. We added an explanatory sentence to this section ("The year 1992 was not chosen for any particular climatological reasons, but rather, being the earliest year in ECCO-Darwin's data assimilative period, allows for potentially the longest continuous simulation.")

--L91: I think this is the better reference: Tsujino, H., Urakawa, S., Nakano, H., Small, R.J., Kim, W.M., Yeager, S.G., Danabasoglu, G., Suzuki, T., Bamber, J.L., Bentsen, M. and Böning, C.W., 2018. JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do). Ocean Modelling, 130, pp.79-139

As suggested, we added this reference.

--Around L100: Mention explicitly here that this version of ECCO LLC270 differs from the more commonly used production version of ECCO at LLC90. Most ECCO users would not be using LLC270, and non-ECCO users might not be aware of the distinctions. These distinctions are important, because LLC270 is not necessarily better than the LLC90 in all ways. And the physical solution of LLC270 is not as extensively studied and vetted by the community.

Thank you for this comment. We added: "It should be noted that this configuration uses the ECCO LLC270 grid, which is a higher-resolution variant distinct from the more commonly used LLC90 production version (Forget et al., 2015)."

--L465 "experimental data" This is confusing. Maybe "observations" or "measurements"? This makes me think of controlled experiments in a lab. Making observations at sea might yield experimental data, but it might also just fill in gaps.

Yes, absolutely. We changed this to "observational data"

--L 588-590: I disagree. I do not think you can say "The most" We really need a better understanding of the processes and parameterization why models are different and how models compare to observations. You start to unpack that here, and it is helpful. Adding more models is not necessarily going to help with this.

We agree with the reviewer's point — absolutely the most important limitation overall is our limited understanding of the processes and the need for further observations. The sentence here was specifically about limitations to our conclusions about model variance, which is a narrower scope and to which the major limitation is the fact we only compared two of them. Incorporating the reviewer's suggestion and clarifying both points we have rephrased this as follows:

"A major limitation to our conclusions is of course the fact that we were only able to compare two models; thus a large-scale OAE inter-model comparison is sorely needed to gain more insights into the model variance. Overall, however, much more observational data will be required to make progress on model accuracy."