

Reviewer 3 (Steve Rackley):

I feel honoured to be asked to review the work of this distinguished group, and I hope that some at least of the following comments and suggestions—coming as they do from a non-oceanographer—will be helpful.

Response: We appreciate the careful and thorough review and are grateful for the kind words.

1. Line 51: I would argue that MRV is primarily a *deployment* rather than a *research* challenge, and that “skillful and fit-for-purpose models” will be **essential** in meeting this challenge, rather than merely “valuable”.

Response: We agree that MRV is primarily a deployment challenge but feel that the question of how MRV can be done most reliably and effectively is an ongoing research challenge. We happily replace “valuable” with “essential.” The sentence now reads (new text in bold):

*“Skillful and fit-for-purpose models will **be essential** ~~prove valuable~~ for addressing many OAE research questions including the MRV challenge, assessment of environmental impacts, and interpretation of natural analogs.”*

2. Line 66: It would be helpful to expand on what is “sufficient”, or how one would go about establishing that.

Response: What is sufficient depends on the application and research question to be addressed. For details on methods to establish that, we refer to section 4.2 about OSSEs. We have revised the sentence as follows:

*“Applications of realistic models rely on them being skillful and accurate, requiring that they include parameterizations of the relevant processes, and that they are constrained by observations that contain sufficient meaningful information (**what is sufficient depends on the application and research question**).”*

3. Line 77: Is the conflation of “scenarios” and “counterfactuals” standard nomenclature in this field? I am used to a different definition of scenarios, as not being limited to counterfactual cases. Where do simple sensitivities—essential in uncertainty analysis—fit into these four general types?

Response: We agree that counterfactuals are just a subset of scenario simulations and no, there is no standard nomenclature yet. The field is just emerging. We have changed this text as follows:

*“Scenarios, or counterfactual simulations, are unconstrained hindcasts or forecasts where one or more aspects of the model is systematically perturbed to assess the effect of the perturbation, for example, in paired simulations with and without OAE, one **would be the realistic case and the other a scenario (also referred to as counterfactual in this specific case)**. These can be used to explore even very unlikely situations, which is often required in comprehensive uncertainty and risk assessment.”*

4. Line 84-85: “especially when considering” seems superfluous, as successful model implementation will be even more of a challenge for eventual OAE deployment. Also, modelling of OAE field trials involving small alkalinity additions may still involve large spatial and temporal scales, depending on local circulation and given the long gas exchange timescale.

Response: Agree. We have removed that phrase.

5. Line 115: Suggest to replace “mCDR” with “CDR”, since terrestrial methods will also affect Earth system feedbacks.

Response: Agree. Changed mCDR to CDR.

6. Line 176: Might be useful to mention Lagrangian methods as also being relevant to modelling the physical dynamics of particles.

Response: Agree. We have added a sentence mentioning that this as a useful approach for this problem (new text in bold):

*“The least understood but potentially dominant source of uncertainty pertains to the representation of the microscale biological, chemical, and physical dynamics of particles, which is an active area of experimental and observational investigation (Subhas et al. 2022, Fuhr et al. 2022, Hartmann et al. 2023). **While the explicit multiphase modeling of the particles themselves is computationally costly, an approach wherein the parametrized evolution of inertia-less Lagrangian particles are simulated may provide a fruitful middle ground, providing a mechanism to realistically determine the alkalinity release field associated with the advection, mixing, sinking and dissolution of reactive mineral particles.**”*

7. Lines 246: “significant” is superfluous here, as even small-scale field trials are subject to the slow air-sea gas exchange. See note 4.

Response: Agree. We have removed “significant.”

8. Line 248: I think it is incorrect to state that transportation distance is not relevant in the case of alkalinity “added to seawater that is oversaturated in CO₂”. Such waters can also be transported large distances before equilibration (same gas exchange timescale applies!), and the seawater conditions at the point of

equilibration (not at the point of alkalinity addition) are what determine OAE efficiency.

Response: Agree that the same time scales for gas exchange apply in the undersaturated as in the oversaturated case. We have removed the first part of this sentence. It now reads as follows:

~~"Unless equilibrated before the addition or added to seawater that is oversaturated in CO₂, Alkalinity-enhanced waters can be transported far away from injection sites before equilibration is complete (He and Tyka 2023)"~~

9. Line 295: "wastewater treatment plants" could also be mentioned in lines 290-293 in relation to reactive mineral addition (*Planetary's* approach).

Response: Agree. Added "or through coastal outfalls (e.g., from wastewater-treatment or power plants)."

10. Line 313: "precludes" is perhaps too strong. Regional models may be sufficient if they fully contain the area of ocean circulation prior to subduction.

Response: Agree. We again removed "significant" in front of OAE deployments and removed the sentence that started with "This precludes..."

11. Lines 364-367: I do not think that **local** variations in relation to carbonate chemistry equilibrium coefficients impact eventual OAE efficiency. The relevant equilibrium coefficients are those at the point of air-sea gas equilibration or, more strictly, at the point of eventual subduction of the DIC enhanced waters, in the case that there are significant changes in seawater conditions between equilibration and subduction.

Response: This may be true for OAE deployment, which is why we use weak language here and say that "it may be important." We feel this is a relevant research topic.

12. Line 383: Challenging indeed, but perhaps not required for deployment modelling since, as suggested in line 542, there are likely reduced-complexity approaches that will be sufficient.

Response: Agree that this may not be required for routine deployment, but routine deployment should only proceed if it is found to be safe for ecosystems. Establishing the latter requires research and more complicated models.

13. Line 463: I guess this should be "and decreasing pH"?

Response: Yes, of course. Now corrected to say that CaCO₃ solubility increases with decreasing pH.

14. Line 496: Does this really stand out as “critical” among the many areas needing improvement?

Response: Agree. Removed the word “critical.”

15. Lines 535-536: The concern regarding prescribed atmospheric pCO₂ is really only relevant in the case of conceptual global studies. For actual OAE deployments, hindcast modelling with the actual pCO₂ history will be required for the quantification of removals.

Response: Agree. We don’t see any contradiction with what is written in the manuscript.

16. Lines 575-577: Re. “ultra-high-resolution modeling tools” if might be of interest to note (as a “personal communication”?) that *Planetary* is developing a Lagrangian particle tracking approach to determine the “alkalinity release field” generated by the advection, sinking, and dissolution of reactive mineral particles with a generic size distribution, which will then be imposed on a regional scale Eulerian model.

Response: We’re reluctant to endorse specific approaches with attribution given the general nature of section 2.3 “Model development needs for OAE research,” but agree that it is useful to mention Lagrangian methods and have done so now. Please see response to point 6 above.

17. Lines 580-581: “The mean state ...”. See points 4 and 11 above.

Response: Please see responses above.

18. Lines 650-652: “different points ...” seems a bit too indeterminate, given that some points in space and time are far more relevant than others! (points 4 and 11 again).

Response: Please see responses above.

19. Line 910: While the pCO₂-pH pair results in higher parameter uncertainties, I don’t think it is correct to characterize this as a “high” uncertainty in the overall context. (See e.g. CarbonPlan’s Verification Framework, which characterizes carbonate system uncertainty as Low (1 to 5% impact))

<https://carbonplan.org/research/cdr-verification/ocean-alkalinity-enhancement-mineral-Component-2;Mineral-dissolution>).

Response: We're making a different point here, namely that some pairs of the 6 carbonate system parameters provide more certain, or uncertain, estimates than others. As stated in the manuscript, unfortunately the two most amenable to autonomous observation are a very unfavorable combination (i.e. deriving DIC, alkalinity from these comes with relatively large errors). A sensor for alkalinity would significantly improve the situation.

20. Line 929: I think ensemble-based methodologies, in general, provide a useful framework for uncertainty analysis, whether incorporating DA or not.

Response: Agree. We see no contradiction to what is written.

Typographical corrections

1. Line 67: An array "is", rather than "are". Alternatively, "Many methods ... are ..."

Response: Changed to "Methods ... are ..."

2. Line 104: Replace "." by "," after "2.2".

Response: For some reason this was fine in the Word doc we uploaded but not in the version that was posted.

3. Line 314: Suggest to replace "alone and" with "alone, which".

Response: Sentence has been removed.

4. Line 414: Replace "104" with "10".

Response: Again, this was fine in the Word doc we uploaded but not in the version that was posted.

5. Figure 4: Bottom right text, "desirable".

Response: Done. Really appreciate the Reviewer's attention to detail.

6. Line 614: "pCO₂" has been used multiple times already and does not need to be defined here.

Response: Agree. Definition removed.

7. Line 713: Replace “properties. But” with “properties, but”

Response: Done.

8. Line 906: Replace “an” with “and”.

Response: Done

9. Line 932: Replace “to made” with “to be made”.

Response: Done