Review

The study uses a fully coupled Earth System Model to investigate different ways of quantifying the efficiency of ocean alkalinity enhancement (OAE) for regional OAE deployment scenarios along the European coastlines. This includes both changes in the internal DIC and alkalinity inventories following OAE, but also cumulative additional air-sea carbon fluxes. The main focus of the paper is to examine the role of internal variability in the oceans' carbonate system which is in different states between the OAE and baseline scenarios and complicates the signal-to-noise ratio of the OAE signal. The authors try different spatial filtering approaches to discuss the potential best way of calculating an informative OAE efficiency and highlight pros and cons of each of them. They find that especially the El-Nino southern oscillation has a strong impact on the global ocean DIC inventory and its variability is stronger than the signal in global DIC of small scale coastal OAE deployments, which are however still likely unrealistic, highlighting the need for other metrics. Approaches to follow the plume of added alkalinity and only calculate changes of DIC, alkalinity or air-sea carbon fluxes within the plume allow for a clearer extraction of the OAE signal. But different methods are still prone to different aspects affecting the final efficiency value and leading to strong uncertainties, which are described in the study.

The paper addresses a relevant issue within the calculation of OAE efficiency in global fully coupled Earth system models with interactive carbon cycling. The effort to define a way to best define the OAE efficiency in these settings is very important for the community and the scope of the paper fits to Biogeosciences. I would like to thank the authors for performing these simulations and analysis in order to advance the scientific understanding of OAE in an important aspect! However, I have some major remarks regarding the submitted manuscript, which are in detail listed below, and need to be addressed before I recommend the manuscript for publication. I therefore suggest major revisions, since the manuscript has much room for improvement, but should ideally be published and can really contribute to the science on OAE.

General comments:

Threshold-based budgets

The connected grid cells approach is pretty cool and quite promising! But I do not really understand why the threshold based plume characterization is only applied at the surface ocean cells. Is the grid adjacency approach not possible adding the cells above and below? Surface only leaves out a lot of ocean volume with both additional alkalinity and DIC, which could pass the same thresholds and influence the efficiency. You say this in the discussion in L279. Why have you only looked at the surface? Even in surface alkalinity addition, it basically adds alkalinity to the mixed layer which is larger than the 6m of the surface cell in your model. And since mixing can also bring Δ DIC and Δ ALK to even deeper layers, I think this approach is potentially missing large parts of the signal. I would recommend including the full 3D field for the threshold approach for Δ DIC. And if not, you should give compelling reasons why not to include the 3D field, e.g. comparisons to observational methods to be potentially used for OAE quantifications, if those would point to surface only observations and hence surface only modelling.

Another point I think worth discussing is how this approach would work for a pulse of OAE and not a continuous application. For a pulse addition or a termination of OAE at some point, the whole alkalinity and DIC would dilute over very large areas and at some point, no cells would pass the threshold any more, basically meaning that efficiency can not be defined any more. I think this is relevant to discuss for the applicability of this approach and highlights an advantage of the flux based approach.

Influence of ENSO

I am having trouble with the description about the influence of El-Nino. Regarding El-Nino, the described effect on the ocean carbon sink is not what the community currently thinks and also not what the mentioned reference says. You say: "during El Niño events, higher sea surface temperatures lead to enhanced outgassing, while colder sea surface temperatures during La Niña lead to a stronger CO_2 uptake (*Sun et al.*, 2025)." However, it is the opposite and in El-Nino, the effect on upwelling seems to be the dominant driver on ocean carbon uptake and not the solubility response. During El-Nino, warmer surface waters lead to enhanced stratification and reduced upwelling of deep carbon rich waters, therefore reduced outgassing, vice versa during La-Nina. *Sun et al.* (2025), but also another already mentioned reference of *Müller et al.* (2025) describe this process. Your data and plots also support this conclusion, which is the opposite of the text, this needs to be revised.

But more generally, the abstract and also the paper (e.g. L247) to me reads in the way, that the OAE deployment and its effect on the climate is the driver of an ENSO phase shift, when I believe it is just another state of internal variability between both simulations, but there is no climate pathway from OAE to ENSO through changed atmospheric CO_2 concentrations. I would recommend to remove this discussion mostly and just say, that ENSO is the main driver of the global ocean carbon sink (Sun 2025, Müller 2025) and that different internal states of the climate system and hence ocean DIC variability create too much noise to detect the OAE response against. If the authors find in their results that OAE indeed impacts ENSO through physical processes, this needs to be better supported and would be an incredible finding on OAE dynamics.

Influence of Biological processes

You describe that the differences between Δ ALK and ALK_{add} are caused by OAE-induced changes in biogeochemical processes, especially nitrate and phosphate cycling through microbial activity. I totally agree that these cycles can influence the alkalinity inventory. However, you later describe that denitrification and nitrogen fixation occur far away from the deployment site and OAE additions are rather small. This makes me wonder, how OAE leads to these changes or if it is again just a state of different internal variability between both simulations, that causes these offsets. Additionally, is microbial activity represented in your model? Is this the process that leads to the mentioned partitioning between phosphorus pools in the ocean? This needs to be better described and explained, one correlation is not enough for me to explain the connected processes in the model and their response to OAE. You later describe the effect of high biological activity in upwelling regions, like of the Iberian peninsula. I am wondering, if you see a strong effect of OAE on biological activity, or if it is again just the state of internal variability between the simulations that adds uncertainty in the OAE quantification. I think this needs to be clarified.

I am also missing a discussion about calcification and how potential changes in calcite and aragonite production impact the internal inventory of alkalinity and DIC. It has been shown that there exists a negative feedback on OAE since calcification is generally enhanced by OAE, but then removes more alkalinity from the system (*Lehmann and Bach*, 2025). I think calcification is one of the main drivers of the alkalinity response to OAE and this needs to be taken into perspective and compared to the effect of nitrogen cycling, when such statements would like to be made. Again, if it is just different variabilities between both simulations, that is enough to argue why a global change in alkalinity inventory might not be the best metric to calculate efficiency for regional and smaller scale OAE deployments.

Clarity of selected results and their interpretation

I think the performed simulations offer a lot more to interpret and discuss than what is currently done in the manuscript. While reading the results, I was left with some questions on why the results are this way, but this was not resolved neither in the results nor in the discussion section. Specific points are discussed in the minor comments below, I wanted to highlight this as a general remark though, as it comes up throughout the manuscript. This would of course need a lot more analysis and likely text, which can be too much for this manuscript. But I think it can be worth it for some of the comments to improve the quality of this manuscript and I want to urge the authors to go through the manuscript again and think about what they say or not say, how much effort it would be to discuss such points and if it is interesting enough to do so.

Other comments and suggestions:

Figure captions, legends and labels could be improved and more detailed to improve clarity and readability of understanding the figures without the text. This also applies for the supplementary figures.

L 18 - 19: Better to write out carbon dioxide removal instead of CO_2 to introduce the abbreviation. And you can choose to do that in L18 or L19.

L 25: I think the statement that OAE has “promising potential” needs some citations or further explanations why that is.

L 30: I think it would make sense to distinguish between the theoretical maximum potential of OAE following the buffer capacity of the ocean (described by Renforth and Henderson 2017) and the observed changes in oceanic DIC for an alkalinity addition (η=Δ DIC/ALK$_{add}$).

L 72: I would recommend shifting the introduction of the background scenario to later in the paragraph where you then discuss the current emission levels (L82). And then I was wondering why you decided for SSP3-7.0, when you then say that climate policies are more likely to follow SSP2 and also later in the manuscript discuss the applicability of SSP3 (L325). I know that sometimes you run a simulation and have to use it and cant afford to run everything again. Maybe try to further discuss the choice of scenario and its implications both in the methods and discussion.

L 89: Tiny thing, but better be even more precise and say that the BASE simulation follows the same emission pathway and are also run in emission-driven mode.

L 113: I was wondering why you look only at annual data. I know that adding the seasonal cycle adds a lot of complexity and more noise to the system and there are good reasons to use annual means. I think however, that you need to discuss this choice and its impacts on the efficiency metrics somewhere in the manuscript.

L 121-122: I was super confused at first why the CO_2 flux is lower under alkalinity addition and then realizing that it’s just based on the way the flux is defined in your model, it is opposite in mine. Best note this here and say that the air-sea flux is defined upward and more negative means more ocean carbon uptake.

L 140: It would help to quickly explain why DIC and ALK show these patterns and why they are so different between the two? There is a lot of research on this and it would help the reader to get a short explanation supported by a reference, one feels a bit left hanging waiting for the explanation.

L 146-148: This is not precise enough as ocean internal biogeochemical responses can still lead to a difference between Δ ALK and ALK_{add}, even when different climate components don't react to each other. You show and discuss this later in the paper, I then would not make this statement here.

Figure 3: “GD & DK” in legend, but other plots say “GER & DK”

L 171-174: You describe the scenarios with small alkalinity additions, but what about larger alkalinity additions? What is the ratio between inventory variability to the alkalinity addition, also in comparison to the smaller additions? This information is missing here and the reader is curious about the differences between scenarios and if so, why.

L 182-183: Same here, you say how much efficiencies range between scenarios, but there is no explanation why they differ. I also do not think there is enough discussion about this in the discussion section. I am wondering about the effect of different local ocean dynamics and/or biogeochemical processes for each of the deployment sites and the following efficiencies. This maybe also informs about the applicability of the different introduced efficiency metrics depending on the deployment site. I think the opportunity of having these different regions is not used enough in this study.

Figure 4: I would note again that coastal oceans refers to the deployment region; mixed “left” and “right” axis in the caption, they do not fit the figure; And at least for the EU and UK sites, Δ ALK seems to be far enough from zero to not impact the efficiency calculation, why omitting these as well? I think their evolution in the starting phase of OAE can be very interesting.

L 191: You state that efficiencies can go up, please add a small explanation and also discuss that it goes down in other regions and why it is different there.

Table 1: you say mean CO_2 uptake efficiency, but the flux efficiency is the final value in 2047 and not a temporal mean, right? Highlight this in the caption. Also applies for Table 2

L 204: I am generally missing a discussion about the choice of thresholds vs the scale of alkalinity addition. I feel like your methods and results could allow to state how high the addition would have to be to be able to detect it against natural variability in both alkalinity and DIC. This would be very informative for OAE signal detection in ESM simulations.

L 209: You state results for the 50 μmol/kg threshold, but not for the other thresholds. First, why is the distribution different between the fluxes and the threshold approach and how does this differ by the choice of the threshold?

L 220: You state that the uptake efficiency is higher for most simulations, but if I understand the results and table correctly, it is way lower for the ICE and NE&BE regions. Why is there such a difference in sites, I think this is good to discuss and also relates to my earlier point of including local ocean dynamics and features into the explanations of varying efficiencies between the regions.

Figure 5: This refers to the internal variability of surface alkalinity and its ranges around the mean, right? Better be more precise here.

L 227: The text says 0.167, but the figure says 0.162, if I understood it correctly. Which one is correct? And why does the standard deviation increase here?

L 227: “in one respect” makes one think there is another aspect to it or a “but”, but you dont describe something else. Best remove “in one respect”.

L 232: Do you mean the reduction in atmospheric CO_2 by OAE with "atmospheric CO_2 removal"? You dont discuss or show this until this point in the paper and only bring it up later in the manuscript (L312), I therefore struggle with mentioning that you "introduced it in this study". Maybe just stick to oceanic carbon uptake at this point.

L 238: missing "in": "internal variability IN the baseline simulation"

L 242: This reads a bit like OAE leads to changes in water temperature and hence changed carbon fluxes, even though you later also say that you can not find a significant response of OAE onto the atmospheric CO_2 concentration, complicating the statement. I think the main oceanic feedback is the one you describe in the following sentence and which has been shown also in other papers (*Acksen et al.*, 2026; *Grosselindemann et al.*, 2026). I would focus on this point and just remove the "e.g., changes in water temperature (*Müller et al.*, 2025)." statement, since the connection is not that clear to me, or describe it further.

L 244: what does "vlg." mean? Maybe I don't know the abbreviation myself :D

L 251-253: I do not get this sentence at all, it seems some grammatical structure has gone missing and/or two sentences were connected?

L 265-267: I am wondering, how the issues with this criterion would be different for larger deployment areas and how to include export pathways. There will always be export out of the domain if you do not look at global changes in DIC and ALK. Maybe marginal seas or enclosed basins with large residence times make this possible, but most coastal areas in the world would see strong export. And regarding the consideration of export pathways, this seems to me to be exactly the plume tracking mechanisms you introduce in the study. Is there a benefit of defining a fixed domain to calculate efficiency in instead of tracking a plume? Maybe one wouldn't have to define thresholds and can easier quantify the background variability within the fixed domain, but this still comes with issues. I think this is worth discussing.

L 274-275: Either variabilities in plural or "has" instead of "have"

L 298: I am wondering how strong the effects of biological processes on the background carbon flux variability are in comparison to physical ocean processes like mixing. Maybe it makes sense to briefly discuss this. This would affect the monitoring necessary in such dynamic regions and guide if one should focus on phyiscal or biological processes.

L 306-308: It makes sense to me that the threshold approach neglects both Δ DIC and Δ ALK in cells which do not pass the threshold, but what is the threshold dependence here? It is not really clear to me, why using Δ DIC and Δ ALK instead of fluxes shows a lower threshold dependence. It would be good to elaborate here.

L 315-317: This very likely depends on the deployment scale of OAE, very large deployments could potentially be visible on those timescales.

L 325-327: I think you need a reference saying that OAE only makes sense in low emission scenarios. Other people would argue it allows us to emit more when doing large scale CDR, which is of course not what we want.

L 329-331: The word "current" is a bit misleading here, just remove it. And be precise in the end, that effects are only visible "after many decades of large scale OAE". For small OAE now, effects would still not be visible after many decades.

L 336-339: Combine these sentences in a way that says studies have either used non-interactive

CO_2 simulations or have added very large amounts. The first sentence alone is a strong statement and there are studies who have used them in interactive CO_2 settings (*Schwinger et al.*, 2024; *Grosselindemann et al.*, 2026), even though likely only using global Δ DIC inventories or fluxes and not regional.

L 337: “were” instead of “where”

L 366-368: Maybe highlight here that the integrated carbon fluxes do not include carbon loss since you filter it by cells that only show significant gains.

L 373: Thanks for making the code and data available to the reviewers and saying that you will upload it for publication. Short reminder to then update this sentence here to include the link to the cloud containing both code and data.

References

Acksen, S., W. Koeve, M. Pahlow, C. J. Somes, and A. Oschlies (2026), Influence of Deep-Sea Carbonate Sediments on the Long-Term Durability of Carbon Storage From Ocean Alkalinity Enhancement, *Global Biogeochemical Cycles*, *40*(8), e2026GB009,203, doi:10.1029/2026GB009203.

Grosselindemann, H., F. A. Burger, and T. L. Frölicher (2026), The efficiency and ocean acidification mitigation potential of ocean alkalinity enhancement on multi-centennial timescales, *Biogeosciences*, *23*(10), 3299–3321, doi:10.5194/bg-23-3299-2026.

Lehmann, N., and L. T. Bach (2025), Global carbonate chemistry gradients reveal a negative feedback on ocean alkalinity enhancement, *Nature Geoscience*, pp. 1–7, doi:10.1038/s41561-025-01644-0.

Müller, J. D., N. Gruber, A. Schneuwly, D. C. E. Bakker, M. Gehlen, L. Gregor, J. Hauck, P. Landschützer, and G. A. McKinley (2025), Unexpected decline in the ocean carbon sink under record-high sea surface temperatures in 2023, *Nature Climate Change*, *15*(9), 978–985, doi:10.1038/s41558-025-02380-4.

Schwinger, J., T. Bourgeois, and W. Rickels (2024), On the emission-path dependency of the efficiency of ocean alkalinity enhancement, *Environmental Research Letters*, *19*(7), 074,067, doi:10.1088/1748-9326/ad5a27.

Sun, C., E. Liao, and X. Zhu (2025), Asymmetrical Ocean Carbon Responses in the Tropical Pacific Ocean to La Niña and El Niño, *Geophysical Research Letters*, *52*(4), e2024GL112,039, doi:10.1029/2024GL112039.