

Biogeosciences Discuss., referee comment RC1
<https://doi.org/10.5194/bg-2021-344-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on bg-2021-344

Anonymous Referee #1

Referee comment on "The sensitivity of $p\text{CO}_2$ reconstructions to sampling scales across a Southern Ocean sub-domain: a semi-idealized ocean sampling simulation approach" by Laique M. Djeutchouang et al., Biogeosciences Discuss.,
<https://doi.org/10.5194/bg-2021-344-RC1>, 2022

This paper is well-written, and some interesting conclusions can be drawn from the results. However I had trouble appreciating the paper as the conclusions were presented in what felt to be a very misleading way. Recommendations are given in the abstract and conclusions without emphasizing appropriate caveats making this paper feel biased and unscientific. I recommend major revisions to appropriately implement these caveats. The title and abstract are misleading with the conclusions not put in proper context. I would prefer if you took Southern Ocean out of the title. As it stands the paper tells me much less about the Southern Ocean than I had envisioned it would. Your decisions made in this paper are fine, but they significantly impact the results and cannot be hidden in the details. Indeed the experiment was designed to make nUSV do the best as you explain on line ~657. Here are some more specific comments (in no real particular order):

- The choice to reconstruct one year of a small box (10 degrees of lat by 20 degrees of long) and call that the Southern Ocean is quite a leap and should be presented clearly in the abstract so as to not mislead about results. This isn't the Southern Ocean and this choice has significant implication on observing system recommendations.

- As you state, Figure 2c shows this model is inconsistent with both reconstructions. Line ~195 needs to be changed. This is significant as it means the $p\text{CO}_2$ may be driven by a mechanism different from the real world. This is relevant to the objective of the study (i.e. you can predict largely from knowledge of SST) and you need to discuss this. If the correlation to SST was weaker you would need to know the biological component better and the statistics may be different. This also has significant implication on observing system recommendations.

-Is nUSV sampling realistic? How many platforms are envisioned? It seems very dense – how long would it take one saildrone to reach all those points? Shouldn't the nUSV sampling look more like the ship track? Moreover, how many nUSVs would it take to sample the full Southern Ocean with that density?

-The goal of the floats is to give the large spatial structure and this aspect of them is not valued in your experiment design. This should be mentioned when giving recommendations.

- If one were to extrapolate this to the full Southern Ocean, which is about 100 times larger than your box, what would the monetary cost to sample with the density of your experiment for each platform? For example, if it would take 100 nUSVs or 200 wavegliders is that practical?

- Even in your box you may give some information about the monetary cost of each sampling effort to give some practical perspective.

- WGs can be “driven”, but you have chosen to give them a mooring-like program. This impacts the results. You finally mention this at the end, but not in the abstract when giving recommendations

- You give the caveat “We emphasize that these experiments are intentionally made to reproduce the sampling resolution of their real-world counterparts, not necessarily their spatial resolution in practice but at least the temporal one.” In the text, but it needs to be featured more prominently (e.g. in the title)

- It appears you train with mixed layer depth (MLD) for all platforms, but surface platforms don't give MLD. How will you get that information and deal with the uncertainty from this? This wasn't discussed, or I missed it, and this is a serious caveat.

- Temporal aliasing by the floats is exacerbated by the fact you don't advect them in the model, so their quasi-Lagrangian nature is not taken advantage of. If you plotted pCO₂ in figure 9 following a water parcel (or plotted it from integrating the full material derivative of pCO₂) I assume it would be much smoother. What you show in figure 9 is likely dominated by advection and not air-sea fluxes. This discrepancy impacts your findings and should be addressed.

Related to the figure 9 comment above, it seems odd to conclude the sampling need be at least daily (line 793) and that the floats should sample 1-2 days when the model output used is daily averaged. How much signal is really being aliased? Please do a more thorough time series analysis (e.g. show power spectra density to see how much variance is truly being aliased). You call into question a discrepancy with Bushinsky et al 2019. I suggest you repeat the calculation done in that paper with the model output. Since you have the model timeseries I suggest you be quantitative with your statements.

My other major criticism is that one expects mapping estimates to never get worse when more data is added. If they are getting worse when more data is added it signifies a flaw in the mapping method and not the data sampling itself. The fact that SHIP(smr) + FLOAT(SAZ+PFZ) has a significantly higher overall RMSE than SHIP + FLOAT(PFZ) is troubling. It points to either a serious issue with your methods or that the uncertainty in the RMSEs you are discussing are very large. You need to explain what is going on here. (This also calls into question your claim of the importance of temporal sampling as it appears something else is going on as well.) Did you carry out WG(SAZ+PFZ), and if so did it also have this troubling behavior?

Minor:

- Model is 1 yr, but Southern Ocean has significant interannual variability and this should be discussed a bit more.
- Line 121 sentence is incomplete
- Supplementary Information text needs cleaning
- In SI you say what hyper-parameters you tune but you don't give the final values you used. You should give these for reproducibility and to inform the results
- Some typos in figure captions. Check these
- Time, J, is stated as a predictor in some cases and not in others. Please check for consistency.