Review of Wu et al. (essd-2024-309)

General comments

Wu et al. describe a new data product that reconstructs sea surface pCO_2 in the North American Atlantic coastal ocean margin over nearly thirty years. The authors rely on the gridded Surface Ocean CO_2 Atlas (SOCAT) dataset as the baseline observations for this data product, and they reconstruct pCO_2 using a two-step random forest regression (RFR) + linear regression (LR) approach. They find that their data product (ReCAD-NAACOM-pCO_2) effectively captures coastal features and variability along the North Atlantic coastal margin. The authors report a region-wide R^2 of 0.83 and RMSE of 18.64 μatm in comparison to observations. Overall, this manuscript describes a useful product that has value for those engaged in studies of ocean acidification and air-sea CO_2 flux in the region. There are some areas, however, where more detailed explanations and thorough analyses would make this a stronger contribution.

The strategy of adjusting RFR estimates with an LR is a unique and straightforward way to mitigate possible biases in the RFR estimates. However, this aspect of the methodology could use more explanation, in particular with respect to why this correction might be needed and how it improves the product ReCAD-NAACOM-pCO_2 relative to not implementing the LR step. If, as indicated, the LR serves to "mitigate potential systematic biases in RFR-derived fCO_2 values [that] arise from spatiotemporal heterogeneities in the SOCAT observational dataset", I envision a figure like Fig. 4c before and after applying the LR would emphasize the added value of this methodological step.

I find the analysis presented in Section 3.3 to be somewhat lacking. While the similarities in large-scale climatological patterns between the raw observations and ReCAD-NAACOM-pCO_2 is encouraging, more interesting is where, when, and why the two datasets differ, and how those differences speak to the value added by the gap-filled product. In particular, I see much higher wintertime pCO_2 in the observations compared to the product in the northern region in Fig. 6. Is this result due to preferential observational coverage of high-pCO_2 areas in that season, as potentially indicated by Fig. 2d? This type of analysis I think is more interesting to readers, and more effective at communicating the utility of the new product.

The comparisons to global products detailed in Section 3.4 would benefit from some quantitative results to be presented alongside the qualitative interpretation of the annual mean climatological figures. The authors assert, for example, that compared to ULB_SOMFFN_coastal_v2 "the ReCAD-NAACOM-pCO_2 product exhibits closer values to the observations", but provide no evidence outside visual inspection of Fig. 7. Instead, by comparing (for instance) the average and RMSE of differences between the gridded SOCAT observations and corresponding values from the products within specific regions, the authors could more clearly emphasize the level of improvement provided by ReCAD-NAACOM-pCO_2.

Additional line-specific comments are provided below.

Specific comments

49: Aren't the fCO_2 data included in SOCAT from these cruises exclusively from underway measurements (not discrete)? In which case, perhaps this sentence should read "Underway measurements from these research cruises, combined with underway measurements from volunteer observing ships and buoy observations, …" or something to that effect.

Figure 1: The regional labels might be better displayed in orange rather than red. As they are now, one might understandably associate the red labels with the red 200m isobath, which can be confusing.

75–77: These two sentences say essentially the same thing and could be combined.

105: The word “enhanced” suggests a comparison for the spatial, seasonal, and decadal variability. It should be mentioned here that the capability of the product at resolving these variations is enhanced in reference to some other dataset. Global products? The gridded SOCAT observations?

109: I find the “ground-truth data” terminology to be somewhat misleading. Ground-truth suggests data that is used to evaluate some model or remote-sensing measurement, but here the data is used not only as a ground-truth but also for training the model itself. Perhaps something like “observational data”, “model-training data”, etc. might be more appropriate.

118: Sampling density also looks to be particularly low in the western Gulf of Mexico.

Figure 3: My understanding is that the “Model” (light orange box with curved sides) is the same that is applied to all satellite and reanalysis data to construct the gridded product. As such, I’d recommend some modification to this flow chart. The arrow from “Model” to “Predictive model” is confusing if those two items are indeed the same.

167–168: More explanation should be given here on exactly how the validation set is used to evaluate the model performance.

171–172: This sentence is somewhat unclear.

178: How is month treated in the model training? If you’re only using 1–12 for the months of the year, there will be an unintended effect whereby months that should be treated as similar (e.g., January vs. December) will be treated as extremely different (1 vs. 12). See Sauzède et al. (2015) or Gregor et al. (2018) for information about transforming cyclical predictors using sine and cosine functions.

206: I believe P represents the total atmospheric pressure, not the “CO_2 atmospheric pressure”.

315–324: I’m not sure this discussion is very valuable because the general features discussed here are evident in the product but also in the observations themselves. It might be more effective to discuss the seasonal cycle features in the product as they relate to the observations; what information is added by the gap-filled product?

416: It should be clarified here that this uncertainty value for the North American Pacific Coastal Ocean Margin is specific to areas within 100km of the coastline and the uncertainty provided for ReCAD-NAACOM-pCO_2 is for areas within 400km.

Technical corrections

148: Should be “arising” or “that arise”

424: Recommend changing wording here: “the performance…reduced” is somewhat awkward

References

Gregor, L., Kok, S., & Monteiro, P. M. S. (2018). Interannual drivers of the seasonal cycle of CO2 in the Southern Ocean. Biogeosciences, 15(8), 2361–2378. https://doi.org/10.5194/bg-15-2361-2018

Sauzède, R., Claustre, H., Jamet, C., Uitz, J., Ras, J., Mignot, A., & D'Ortenzio, F. (2015). Retrieving the vertical distribution of chlorophyll a concentration and phytoplankton community composition from in situ fluorescence profiles: A method based on a neural network with potential for global-scale applications. Journal of Geophysical Research: Oceans, 120(1), 451–470. https://doi.org/10.1002/2014JC010355