Reviewer Rik Wanninkhof NOAA/AOML

This is a well written manuscript thoughtfully describes and rectifies inconsistencies between several global ocean air-sea CO2 flux products. Notably, differences in ocean surface area depending if marginal and coastal seas are included; and providing an adjustment to quadratic gas transfer parameterizations to match the global bomb 14C constraint. This "harmonization" leads to a better agreement between the different products. As a last step the authors use an ensemble approach to determine a global mean air sea CO2 flux of 1.92 +-0.35 Pg C/yr where the uncertainty is a 2 sigma (95 %) confidence interval and is based on agreement of the 18-member ensemble using on 6 interpolated and area normalized surface water pCO2 products and three different windspeed products with normalized gas transfer velocities. The uncertainty (+-0.17) (1-sigma) basically reflects the differences in interpolation techniques of the same dataset (SOCAT). Using this uncertainty for purposes as stated "help to identify missing fluxes as we strive to close the global carbon budget" is not appropriate. While beyond the scope of ESSD, it looks like the authors are trying "to sneak a number" into the peer reviewed literature.

The manuscript appears a mixture of a data description as expected for ESSD and commentary/interpretation which is beyond the scope of this journal. As listed on the home page of ESSD: "for the publication of articles on original research data (sets),. The editors encourage submissions on original data or data collections which are of sufficient quality and have the potential to contribute to these aims. Any interpretation of data is outside the scope of regular articles."

Response: Many thanks for the overall positive assessment of our study and the helpful comments. We have edited the manuscript to be inline with the scope of the ESSD journal, specifically focusing on the novel aspect of the computational tools provided by the pySeaFlux package. While we do still include the mean air-sea CO2 flux estimate from this group of products in the manuscript text we have removed it from the abstract. We have also focused on the uncertainty quantification at each step in the pySeaFlux package, including uncertainties inherent to the observations themselves, and elaborate on such in the text. We have added an additional paragraph to the Area-filling discussion (Section 2.1) that specifically reports estimates of uncertainty for this method. Additionally, we have added a section to the Results and Discussion section that outlines various types of uncertainty: the intrinsic uncertainty, introduced uncertainty, and the spread of the resulting ensemble of products.

R1: Why isn't the "Landschutzer, P., Laruelle, G., Roobaert, A., and Regnier, P.: A combined global ocean pCO2 climatology combining open ocean and coastal areas (NCEI Accession 0209633)," included as one of the interpolated products? This seems to be one of the most complete sets wrt area. Also, citing the paper in ESSD rather than the data product might be more appropriate.

Response: The Landschützer et al 2020 climatology is not included as an interpolated product because it is limited in its temporal coverage- it is a 12 month climatology rather than a time-evolving product. We have edited the manuscript to reference both the ESSD manuscript and the data product itself.

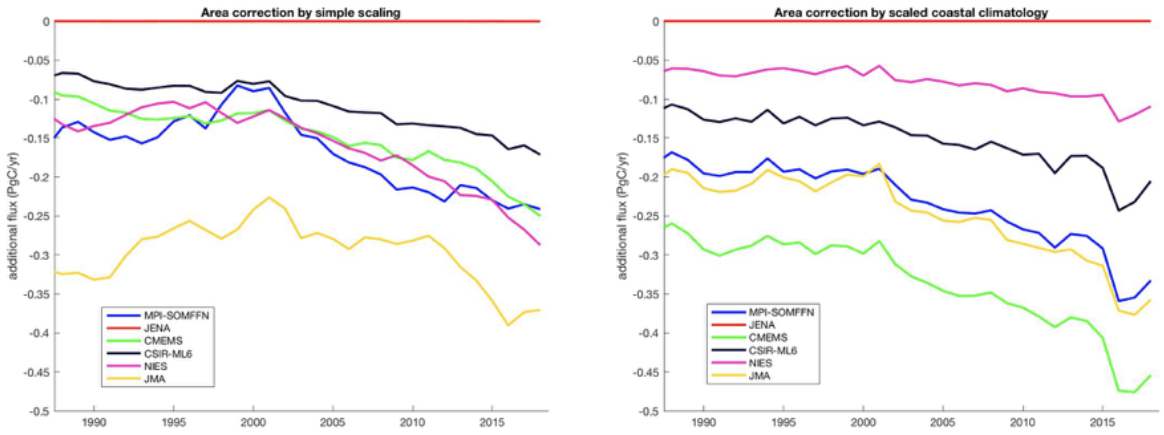
R1: The authors refer to the "Seaflux product and "Seaflux package" but the full description of either is lacking in this paper. The Gregor and Fay, 2021 referenced is the dataset without a complete description. What exactly is the Seaflux package? I would encourage the authors to focus on describing this product and tools in this manuscript.

Response: Thank you for this comment and suggestion. To clarify this, we have decided to call the Python package pySeaFlux, which contains all the code to calculate air-sea CO2 flux from pCO2. The SeaFlux product (the subject of this manuscript) is the ensemble of the 6 pCO2 products with area-filling to produce global coverage. This product will therefore be evolving as more products become available.

R1: It is mentioned that the area normalization has been previously applied in models and products in a rudimentary fashion. How different are the global fluxes using simple extrapolation methods compared to the approach used here? Eyeballing the results it appears that scaling global fluxes to a consistent area, and has been done it the past seems to work reasonably well.

Response: The additional flux for each product resulting from the area-filling method proposed here in the pySeaFlux package is on the same scale and magnitude as other simpler methods. Below is an annual time series of the additional flux amount calculated by the area-weighted method used in the Global Carbon Budget (a) and a similar plot showing the annual additional flux using the SeaFlux methodology (b). We have added this comparison figure to the revised manuscript in the Discussion section as we see the Global Carbon Budget interpolation as a primary potential use of a product such as SeaFlux.

For the area-weighting method, the interannual variability in this additional flux is a direct result of the IAV of the total global flux. Also, products with larger fluxes will have a larger correction inherently with this method, even if they aren't missing the largest area. For example, if two products were missing the exact same regions/gridcells, but one product had flux that was 0.3 PgC/yr larger for that year, the correction applied to the two products would be different, even though they were missing the same area. This assumes that the interannual variability of pCO2 in missing areas would be the same as for the rest of the product.
Another consideration is that simple area-weighting does not take sea ice cover into account, which is important given that the high latitudes are often the region lacking coverage. By first filling the product pCO2 maps with full spatial coverage and then calculating the flux, you account for this ice fraction.



(a) (b)

R1: Figure 1A is not clear. Is "the changing fraction of area covered by observations" essentially seasonal changes in ice coverage? If so, perhaps include the different expressions for gas exchange in partial ice overage (e.g. Takahashi 2009), and different ice products in the analysis.

Response: We thank the reviewer for this comment as we had edited Figure 1 and neglected to update the figure caption sufficiently. These maps show the fraction of total months with coverage for each gridcell. Blue areas with a fraction of 1 represent regions that have coverage for all months of the product (here, 1988-2018). Yellow areas show where the product has no coverage for any months of the time series. We have amended the Figure 1 caption as such: Figure 1: Maps showing the fraction of months (1988-2018) with coverage available for each of the six pCO2 data products used in this study. Blue regions represent full temporal coverage of pCO2 in the product while yellow areas show regions with no reported pCO2 values for any month of the time series.

Minor issues:
R1: The gas transfer velocity is listed as piston velocity and exchange coefficient: be consistent

Response: Thank you for this comment. We have revised the manuscript and maintained consistency using the terminology "gas transfer velocity" and "coefficient of gas transfer" and removed piston velocity.

R1: I don't think that "improved" in the title is appropriate. The title for the SeaFlux product seem better as title for this paper: "SeaFlux data set: Air-sea CO2 fluxes for surface pCO2 data products using a standardised approach"

Response: Thank you for this suggestion. We have removed the word "improved" from the title and revised the title to: "SeaFlux: harmonization of air-sea CO2 fluxes from surface pCO2 data products using a standardised approach"

R1: Tables and figures are good but lines are difficult to read (for those with color impaired eyesight)

Response: Thank you for this comment. We have altered the colors on the figures to improve readability by those with color impairment by using the "colorblind" color scheme available from the Seaborn python visualization library (as shown in the figure below).