Earth Syst. Sci. Data Discuss., referee comment RC2
https://doi.org/10.5194/essd-2021-16-RC2, 2021

Comment on essd-2021-16

Timothy DeVries (Referee)

Referee comment on "SeaFlux: harmonization of air–sea CO_2 fluxes from surface pCO_2 data products using a standardized approach" by Amanda R. Fay et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-16-RC2, 2021

Review of "Harmonization of global surface ocean pCO2 mapped products and their flux calculations; an improved estimate of the ocean carbon sink"

The authors use a consistent set of input data to estimate the air-sea flux of CO2 from six different seawater pCO2 interpolation products for the period 1988-2018. They then fill in the missing areas in these data products using a scaled estimate of the air-sea CO2 flux from a recent data-based interpolation. The authors make two claims about the results:
(i) that this methodology provides a consistent method for computing global air-sea CO2 fluxes from seawater pCO2 products, and that using this consistent methodology allows for improved intercomparison between the air-sea fluxes computed by difference seawater pCO2 products, and
(ii) that this methodology leads to improved estimates of the global air-sea CO2 flus as well as reducing the uncertainty in the global air-sea CO2 flux.

The first claim is certainly correct, and it will be very useful to have a common set of input parameters that modelers can use to determine the air-sea CO2 flux from their seawater pCO2 products. This will allow better intercomparability of results across different seawater pCO2 products. The input data and methodology are very clearly described in the paper, and seawater pCO2 modelers will find this to be a useful reference. On the strength of this aspect of the study, this is a useful study and should be published.

However, the second claim is misleading and requires substantial revisions to the paper. By using a uniform set of input parameters and assumptions, as well as a single product to fill in spatial gaps, the authors reduce the ensemble spread among the six different pCO2 products. However, this is not at all the same as reducing the actual uncertainty. The true uncertainty should reflect the uncertainty in the input parameters and assumptions used to compute the air-sea CO2 flux from seawater pCO2 products, and should also reflect the uncertainty of the air-sea CO2 flux in regions that are not covered by some products. Therefore, rather than stating that their methodology provides an "improved estimate" of the global air-sea CO2 flux with a reduced uncertainty, the authors should remove any such statements and instead explicitly discuss how their methodology artificially reduces the uncertainty in the global air-sea CO2 flux. This needs to be explicitly caveated, otherwise the community will cling to the numbers reported here as a "best estimate" of the global air-sea CO2 flux, which is not what it should be intended as. I understand the desire to create a consensus estimate of the global air-sea CO2 flux, but this consensus will emerge when multiple independent methods yield the same answer, not when one approach is uniformly applied.

In addition to this concern about how the results are presented, I have a general concern with how the missing areas are filled in each product. Filling in the missing areas with the estimates of one single product (the MPI-ULB-SOMFFN) is problematic, as this implicitly assumes zero uncertainty in the gap-filled areas. Rather, these areas are precisely where the uncertainty is the largest, and indeed the method used here depends on several assumptions which have their own uncertainties (which of course is not reported in the uncertainty because it does not contribute to the ensemble spread). The best way to fill in these missing regions is to extend the individual methods to global coverage, as this would provide a better estimate of uncertainty in these regions. In general, I am concerned the using the MPI-UMB-SOMFFN product to fill in these gaps will become entrenched, and reduce the motivation for extending the various methods to global coverage. The authors should explicitly warn against this, and discuss the necessity of having multiple independent estimates of the air-sea CO2 flux in these regions (high latitudes, shelves/seas) in order to arrive at a good estimate of the uncertainty and an improved estimate of the global air-sea CO2 flux. In the meantime, the authors might want to consider using at least one other estimate of the air-sea CO2 flux in these missing regions — the Jena-MLS has global coverage and so could be used there as well.

Below I list some specific points in the manuscript where revisions should be applied to address these more general concerns that I listed above. I also identify some areas where clarification is required.

Line 24-25: It is true that the details of how these calculations are done varies greatly among methods, but it shouldn't be said that this unnecessarily enhances the uncertainties. Rather, the differences among methods reflects (only partly) the actual, true, large uncertainty that exists when converting sparse measurements of seawater pCO2 to global estimates of air-sea CO2 flux.

Line 25-26: Applying a uniform approach to all the different methods yields a lower ensemble spread, but it does not reduce the actual uncertainty. Filling in the gaps in products with a single product of course yields a lower spread (because a single estimate is applied to all the missing areas), which artificially lowers the "uncertainty". Likewise, applying a uniform gas transfer velocity and gas exchange formulation to all products artificially lowers the true "uncertainty".

Line 31: Again, these "methodological inconsistencies" reflect our imperfect state of knowledge. Different groups chose different wind products, gas exchange parameters etc. Sure, one can repeat the calculations with a uniform set of parameters and get a lower spread, but one cannot argue then that this reflects a lowering of the true uncertainty.

Line 34: Regarding "appropriately scaled". This scaling factor has significant uncertainty. Also, it depends on the assumed relationship between wind speed and gas transfer velocity (i.e. linear, quadratic, cubic, or some mixture). Again, one can use a single number but it does not make the uncertainties go away.

Line 56-57: the "lack of systematic approach" is due to real uncertainties in how to do this calculation, not because some folks are doing it wrong

Line 57-58: These differences don't "introduce uncertainty", they capture real uncertainty

Line 60: Please clarify what you mean by "meaningfully compared". They can be compared regardless of whether or not these adjustments are made, one just has to attribute the differences correctly

Line 79-80: Yes this is a consistent approach, which should be highlighted is useful for intercomparisons AMONG the different seawater pCO2 products, so that the differences between products can be attributed to differences in their underlying seawater pCO2 estimates. However, I would argue that a less consistent approach (i.e. one that better accounts for uncertainties) would be more appropriate for intercomparisons ACROSS products, e.g. for comparison with the biogeochemical models used in the Global Carbon Budget. If the SeaFlux products is used for cross-model intercomparison it will underestimate the true uncertainty.

Line 97: It should be noted that there's no "problem" with any of the choices in Table A1, they are all reasonable assumptions. Notice they also all use the quadratic wind speed dependence, which is not necessarily "correct" (at least for all wind speeds)

Line 116: It is important to say more precisely what you mean by "coastal". I think here you mean the continental shelf waters. So does not include the littoral zone, estuaries, tidal wetlands, seagrass/mangroves/kelp forests etc.

General point about resolution: I don't think any of the products at 1 degree resolution capture "coastal" areas. So they are not missing coastal areas because of a lack of input data, but simply due to their coarse spatial resolution.

Another point: If the Jena-MLS does not require any filling, why not just fill the other products with the Jena-MLS product, instead of filling them with the MPI-ULB-SOMFFN product? Or at least, one could use both as an estimate of uncertainty.

Line 129-130: This is confusing. The way that the regions are defined shouldn't matter in terms of calculating a global air-sea CO2 flux. Do you mean differences in the coverage or lack thereof in different regions?

Line 135: What is the spatial resolution of the MPI-ULB-SOMFFN product?

Table 1: What resolution grid is used to calculate the area? Please also define the land/sea mask

Table 1: It would be good to report the actual global mean pCO2 before and after the filling is applied

Equation (2): The authors should plot this scaling factor out over time for the different models.

Paragraph starting line 155: This application of the scaling factor could be problematic for several reasons. First, the scaling factor is derived for the open ocean, but is applied in the coastal ocean (i.e. shelves and seas). As the authors mentioned, the processes driving the pCO2 in the shelf/sea regions are distinct from those in the open ocean, so this scaling factor might not be appropriate. It would be better to use a scaling based on results from a high-resolution carbon cycle model that includes shelf/sea regions. Second, this implicitly assumes no interannual variability in the shelf/sea carbon sink (leading to too-low uncertainties). This is problematic because the shelf/sea regions are highly impacted by human activity such as fishing/trawling/farming etc. So I think there are very large unreported uncertainties that arise due to this scaling factor.

As asked before: Why not just fill the other products with the Jena-MLS product? Then this scaling would not be needed.

Lines 166-168: Not true, the Jena-MLS product covers those regions.

Line 175: The area-based scaling could be called a “reasonable first order approximation” as well.

Line 197: Please clarify what is meant by “second moment of the average”? The average is the first moment.

Line 198ff: The quadratic dependence is most often used, but the actual relationship could vary from less than linear (Krakauer et al., 2006) to almost certainly cubic at high wind speeds due to bubble-mediated transfers (Stanley et al., 2009), so this is a significant source of uncertainty.

Line 212: “piston velocity” is used here, while elsewhere it is “gas transfer velocity”

Line 217-218: This was not clear to me. What is meant by “a probability distribution of wind speeds is used to optimize the gas transfer coefficient”? Also, it should not be stated that the rate of bomb 14C invasion is observed — rather, it is inferred from an estimate of the bomb 14C inventory in the ocean (which has a significant uncertainty), and also requires the intermediary of an ocean circulation model (another source of uncertainty).

Lines 219-220. This sentence is also unclear. How would you scale a gas transfer coefficient to a bomb-14C inventory? Maybe what the authors mean is that the estimated bomb 14C inventory has been used to infer a global average estimate of the gas transfer velocity, and that different methods have used gas transfer velocities that may not be exactly the same as ones that have been inferred in previous studies (e.g. Sweeney et al or Naegler et al).

Lines 220-221: This is not clear again. What do the authors mean that “the range of bomb-14C estimates is within the range of uncertainty from the associated studies”? It sounds like the authors mean that they used the gas transfer coefficients and wind speed parameterizations from the various studies listed in Table A1 and then calculated the bomb-14C inventory in the ocean, and compared that to the estimated value of the bomb 14C inventory from Naegler (2009). But I doubt that is the case because the bomb spike goes back to 1955 and it requires a model for the seawater bomb 14C as well (i.e. a circulation model). So maybe what they mean is that the globally-averaged gas transfer velocity used by the different studies listed in Table A1 is within the range of uncertainty on that particular parameter that has been deduced from studies that have inferred the globally averaged gas transfer velocity using an estimate of the bomb radiocarbon inventory and a circulation model (e.g. the estimates of Naegler, 2009).

Equation (4): Here the “a” parameter is not the same one as used in equation (3), because you have introduced the (1-ice) term in this equation. If you say that Kw (equation 3) includes the (1-ice) in its definition, then equation (4) would be divided by (1-ice), and you would eliminate the (1-ice) from equation (1). So you need to say whether you are finding a value of “a” that yields a value of Kw as defined by equation (3) that is 16.5 cm/hr on average, or whether you are finding a value of “a” that yields a value of Kw*(1-ice) that is 16.5 cm/hr on average. Also, notice the additional uncertainty that this choice introduces.

Lines 231-232: What do you mean “even with the same bomb-14C observations the scaled coefficient (a) can have a 40% range?” This sentence is puzzling because the coefficient “a” is completely independent of the bomb-14C observations. Do you mean that the value of “a” that would be inferred by an inverse model that tried to match the bomb-14C inventory in the ocean would depend on the wind speed product used by that model, such that using different wind speed products could result in optimal values of “a” that are as much as 40% different from one another?

Line 232-233: No, one cannot reduce the uncertainty in the global fluxes in this way (unless the original value used falls outside the uncertainty bounds of the average gas transfer velocity). One can reduce the ensemble spread by specifying that each model use the same globally-averaged gas transfer velocity. This is not the same as reducing the uncertainty. The true uncertainty would take into account the uncertainty in the value of "a" (which is ~20%, see Wanninkhof, 2014) as well as the uncertainty in the form of the gas transfer velocity parameterization itself (e.g. quadratic vs. cubic).

Line 240: The authors say "our results show" and then cite the study or Roobaert et al. (2018). Please clarify which study shows this, the present one or that of Roobaert et al.

Line 248-249: What do you mean by "small, but not insignificant"? Can you state the magnitude of the impact?

Line 261: Taking a global mean is straightforward regardless if you account for spatial coverage differences or not. So I think the authors mean something other than "straightforward".

Line 265-267: Please state the spread due to wind product without scaling, vs. with scaling

Line 269-270: Yes, the SeaFlux allows "a more accurate comparison of fluxes" within pCO2 products, but it does NOT lead to "increased confidence"

Figure 5: It would be useful to show the individual models before and after the corrections are applied. Why does the NIES-FFN appear to change so much relative to the others in the later period? Is it due to the gas exchange scaling or infilling?

Lines 275-276: Actually, one should have lower confidence in the uncertainties and the ensemble mean, because the uncertainties have been understated. Also, the phrase "higher confidence in the uncertainties" sounds a bit strange because it implies you are reducing the uncertainty in the uncertainty.

Line 278-280: "working towards consensus on other issues" implies that there is a consensus on the issues addressed here (gas exchange parameterization, shelf/sea CO2 flux). However, there is not a consensus on this, the authors have simply picked one reasonable approach to these issues and applied it uniformly to different seawater pCO2 products. A consensus will emerge when multiple independent methods yield the same answer, not when one approach is uniformly applied.

Line 318: The authors should explicitly discuss/restate here how this uncertainty estimate is too low. It would be unfortunate if the community mistook this as a consensus estimate or "best estimate" of the mean and/or its uncertainty.

Line 324: What do the authors mean by this "may reduce the current carbon budget imbalance"? This needs to be spelled out in more detail.

Line 329-330: Since others will be able to apply these "standardizations" to their datasets, I think this should come with a warning. These standardizations should be used for applications where there is going to be intercomparison with other seawater pCO2 products. But it should not be used as a method to arrive at a "best estimate" of the air-sea CO2 flux. What the community really needs is independent methods of estimating Fnet in order to derive a robust estimate of the mean and uncertainty. If everyone uses the same methods, assumptions, and datasets, we will never know what the true answer is.

Sincerely,

Tim DeVries