Earth Syst. Sci. Data Discuss., author comment AC3
https://doi.org/10.5194/essd-2022-71-AC3, 2022

Reply on RC3

Jiye Zeng et al.

Author comment on "A new estimate of oceanic CO_2 fluxes by machine learning reveals the impact of CO_2 trends in different methods" by Jiye Zeng et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-71-AC3, 2022

The reader has given valuable comments and kindly pointed out many misuses of terminology, grammar and spelling errors. We have corrected the errors and incorporated most advice in the revised manuscript. The followings are point-to-point response to the comments.

- The reader suggested explaining more about the choice outlined 2.1 and 2.3. We have provided python code for model creation in the supplement. Those interested in the details could go on reading related python documents and references listed in the manuscript. In the revised manual script, we added more explanations for the iteration method of rate extraction.
- Regarding the question of why not just using atmospheric xCO2 trends. It sounds simple but is questionable. First, what xCO2 trends to use? Using the annual increase rate of xCO2 in the same year or in the previous year, or the mean rate of the past two years, three years… 10 years? Second, how to prove that ocean CO2 change follows the same pattern as xCO2? The uptake of air CO2 may not the only major factor that leads to ocean CO2 increase. Changing SST could be an important factor as well, for example. For these reasons, we do not plan to use xCO2 trends directly to model flux for comparison.
- We have rewritten the discussions for comparison. First, we included the JENA product. We didn't include it initially because its spatial and temporal resolutions are different from others. The results of a comparison could depend on how a monthly product in 1x1 degree grids were derived from JENA's daily product in 2x2.5 degree grids. Seconds, we recalculated the fluxes of all products under comparison using the same method and adjusted the fluxes to the ones as if they have the same grid coverage. Third, time-series fluxes are put in the same plot for comparison. Fourth, we added a figure to summaries the mean differences.
- The reader questioned the variable of global locations of observations. As the reader pointed out "There is no way to account for or correct for the fact that some years data is dominated from the tropical pacific which could have a much different trend than that of the rest of the ocean." Our method examined the change of APPARENT GLOBAL TREND with data length and select the trend when the change become small. The implicit assumption is that the effect of unbalanced sampling becomes acceptably week with increasing data length. The method is not perfect, but it offers an alternative. If the locations were quite even geographically, it would not be necessary to use machine learning in the first place, as a simple spatial interpolation would be sufficient.
- In the revised manuscript, we emphasize that the uncertainty value in Section 3.2 is solely due to the machine learning. It is unreality to count in all uncertainty factors in

such a data paper, including uncertainties of measurements, griding, and etc.

- In the revised manuscript, we explained that a negative flux value means an efflux of carbon out the ocean and a positive flux into the ocean. Although the reader strongly suggests flipping the convention around, we think the convection makes it easier to see the trend of flux. When we made the flux map, we used the same convention and a similar colorbar palette as those in the figure 6 of GCB-2021.