

## Comment on **essd-2022-294**

Anonymous Referee #1

---

Referee comment on "OceanSODA-UNEXE: a multi-year gridded Amazon and Congo River outflow surface ocean carbonate system dataset" by Richard P. Sims et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2022-294-RC1>, 2022

---

### General comments:

A complete and robust dataset on the riverine carbon output and variation is definitely interesting to the ocean carbon community and the international stakeholders. According to the title, the data then tend to present would be highly appreciated by many readers and research communities. However, the manuscript itself have some major flaws, which discourage me to recommend to publish it on its current status. The major aspect concerning me including: 1) Language quality is not sufficient and need quite some effort to improve to the level of concise and straightforward academic English. 2) There is a large mismatch between the stated content of the dataset and the actual content of the dataset, specifically, the full carbonate system dataset was said generated, but actually the dataset only consists of the variables DIC and TA. 3) the description the key method , pyCO2SYS v1.7.1 software, for carbon (pCO<sub>2</sub>, fCO<sub>2</sub>) estimate is missing.

### Specific comments:

- I did not really understand the processing flow of the data generate, so DIC and TA were estimated with some optimal algorithms and pCO<sub>2</sub>, fCO<sub>2</sub>, and pH were the output from the software pyCO2SYS v1.7.1.. but under the title of "full carbonate system data set", TA and DIC are the primary output and took much of the manuscript and evaluation. However, as a reader and a researcher in the ocean carbon community, I would expect pCO<sub>2</sub> and fCO<sub>2</sub> to be the major variables in the "full carbon system dataset". Please make this one clear.
- Throughout the manuscript, the language is not concise or straight forward enough, meaning not academic, and it takes quite some efforts to understand many sentences to grasp their real meaning. And the logic in many of the paragraphs do not really flow.
- In the abstract, the author mentioned they generated a dataset of full carbonate system, but the variables they mainly present were TA and DIC. So, there should be a statement on the linkage between the variables in the full carbonate system and the TA

& DIC. Or there should be a summary on all variable consist of the dataset and their spatial and temporal resolutions.

- Line26-28, the uncertainty of the TA and DIC were expressed with absolute RMSE and bias. I suggest the percentage uncertainty should be included, i.e., how much the RMSE and bias account for the minima and minimum of the estimated value of TA and DIC.
- line 68-69, "Episodic changes in the carbonate system caused by river plumes can result in financial and biodiversity losses and are of paramount interest to local communities, businesses and policy makers (Doney et al., 2020)." , please give specific examples on what kind of financial and biodiversity loess does it cause and how it is of interest to the stakeholders.
- the first paragraph in section 2.1 is not necessary.
- line 101-102:" This is a clear weakness of comparing wRMSD values from different sources and across differing regions (Land et al., 2019)." If there is a clear weakness of wRMSD, what is the reason to use it to indicate the quality of the dataset?
- 7line 175, what is Type A uncertainty?
- 219- 226, the brief introduction to the software pyCO2SYS v1.7.1 should be included.

10, line 244-250, should be in the method section instead of results.

Technical corrections:

- line 39-41 "The inorganic carbon content of rivers is poorly constrained due to the difficulties of sampling these highly spatial and temporal variable river outflows." The logic is not correct, please revise it.
- line 64-65, " River plumes can negatively influence wild fisheries and the aquaculture industry (Mathis et al.,2015;Cattano et al., 2018) as plumes can transport low pH waters that can impact the growth and 65 life stages of many marine organisms (Cai et al., 2021) Additionally," a punctuation is missing.
- line 102-103," Following the methodology of Land, Findlay et al. (2019) we derive RMSDe from wRMSD," does not make sense., please revise it.
- line 120-125, "To be included in the algorithm evaluation, algorithms needed to be applicable within the..... chlorophyll-a." the sentence is too long to understand, please split it.