Point–by-point responses to editor and reviewers

Dear Editors and reviewers,

We thank all reviewers for their thoughtful, supportive and constructive comments.

Both reviewers concluded that our manuscript should be published after revision and we have now responded fully to all of the points raised and we will update the manuscript.

The reviewer and editor comments are in black and our responses are in red. In the file containing the revised paper, we will used track changes to show the revised text.

We look forward to hearing from you and thank you for your time.

Best wishes,

Richard Sims and co-authors

Reviewer 2

The authors reconstructed gridded carbonate system datasets by using a data matchup method based on relationships between carbonate parameters and others which had already been established in the past. While many studies have explored such relationships based on ship-based observations during these decades, the authors utilized these efforts in an effective manner. Such a study is unique and is worth being published, but there are major concerns to be clarified before publication in this journal. I'd like to encourage the authors to improve the study and to revise the manuscript for better understanding.

Thank you for your positive comments and support of our work. We have now addressed all of your comments in full and will revise our manuscript.

General comments

Oceanographic characteristics of the studied areas considered, one of the important points of the method is skill to estimate carbonate parameters of low salinity seawaters, which are complexly influenced from both river outflows and heavy precipitation along the ITCZ. On the other hand, relatively higher salinity ($S >$ approx. 34) seawaters in these regions have similar chemical properties to those in the nearest open ocean, where large scale ocean circulations dominate the seawater carbonate chemistry. According to attached supplement files, measurement data used in the matchup process were not necessarily restricted to those of low salinity seawaters. It should be emphasized that the presented method derived more appropriate TA and DIC of low salinity seawaters than others did.

Yes, this is a point that has been covered and evaluated in our previous work (within Land et al., 2019).

We will highlight this important point within our manuscript and explain its impact based predominantly on our findings in Land et al., (2019).

Moreover, secular trends of CO2 were not considered in this study, though time-series reconstructions were addressed. It is needed to show reasonable explanation about that.

Nowadays prevalent machine learning-based methods are used for carbonate system reconstructions; five of the six methods which were cited for evaluating observation-based CO2 sink in the IPCC AR6 assessment used machine learning (Canadell et al, 2021, e.g. Fig. 5.8). It should be explained carefully that this study has some limitation that novel reconstructions cannot be included and legacy of past studies only be used.

We evaluated all the algorithms equitably and as they are presented in the original literature, so we did not modify any relationships as that would be a further substantial amount of work (and was not the focus of this work). We agree that environmental conditions will have changed in the time since these algorithms were derived, e.g. due to oceanic uptake of CO_2 and increased freshwater content of the oceans. TA is a conserved quantity in the ocean so is not impacted by the uptake of CO_2 but may be lower today in the surface ocean than in the past due to freshwater inputs.

We will now include this point within the introduction of our paper.

We do not use any literature algorithms for CO_2 and instead use algorithms to derive DIC; this does mean that the large relative ~ 20 µatm increase per decade in CO_2 are much smaller for DIC as a percentage of the total inorganic carbon pool. Theoretically an additional term could be added to each DIC algorithms which accounts for the increase in oceanic DIC since each algorithm was developed. However, such a correction effectively equates to a bias in DIC. Furthermore, to fully account for secular trends, the insitu data would also need to be standardised to a reference year.

We will add a sentence to the paper to explain this point.

As part of the OCEANSODA project (from which this paper has been written), the performance of our approach was assessed against the machine learning OCEANSODA-ETHZ output from Gregor and Gruber (2021) (which is one of the methods that the reviewer refers that was used within the latest IPCC assessment). OCEANSODA-ETHZ is a state of the art machine learning approach which was also recently included in the 2021 global carbon budget estimate (Friedlingstein, Jones et al. 2022). We found that in the Amazon outflow the Gregor and Gruber (2021) TA had a wRMSD of 54.03 µmol kg^{-1} (matchups N=87), whereas the best TA uncertainty in our regionally tuned empirical outputs (given within our manuscript) gives wRMSD 34.97 µmol kg^{-1}. This is not surprising as these riverine regionally-specific empirical algorithms were trained on data for these regions whereas the machine learning approaches are trained on global data (which is increasing reduced into sub-regions during training, but it is unlikely to become riverine-outflow region specific). The machine learning approaches cannot be applied to these riverine regions due to them requiring large datasets for training.

We will add additional text into the discussion to explain why globally applied approaches, including machine learning techniques, are likely to perform poorly on a regional basis.

Specific comment

Overall

Unnatural uses of brackets “()” have to be checked.

We will go through the whole paper and we will check all uses of brackets.

P3 71

Before OceanSODA is presented, successive efforts of investigating empirical relationships between TA/pCO2/DIC and other parameters based on observations have to be mentioned here.

We will add a sentence explaining the reviewers point.

P3 L72-76

A brief explanation of OceanSODA is necessary.

We will now define the acronym in the text and briefly introduce the objective of the OceanSODA project.

P4 L103

A brief explanation of RMSDe is necessary.

Agreed. We will include a brief explanation and explain that the full details can be found in Land et al (2019).

P8 L244- Figure 1 Fig. 1 obviously shows that the four selected algorithms have the lowest RMSDe, but doesn't explain whether they are the best even in low salinity regions. It is questionable that Lee et al. 2000; 2006, which propounded global algorithms and (the latter) didn't use salinity as explanatory variables, have the best skill in low salinity Congo basin. This point should be clarified.

Whilst the Lee et. al. (2000;2006) papers provide global algorithms they also provide separate algorithms for different ocean sub-regions. We use these sub-region algorithms and not the global algorithms, and so these sub-region algorithms do use salinity as a predictor.

Within Lee et.al 2000, salinity is not specifically used as a predictor variable however, the relationships are for salinity normalised DIC, so whilst salinity it not a direct input the algorithm, the output is scaled by salinity.

We will add text to explain these points.

Fig. 4, 5, 8, 9

If DICs were successfully reconstructed, trends of increase in DIC and pCO2 and decrease in pH and Ωs would be also derived. The trends are worth being mentioned in the text to support the validity of this datasets.

This is an excellent point and we will include these in the revised manuscript.

Canadell, J. G. et al.: Global Carbon and other Biogeochemical Cycles and Feedbacks. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 673–816, https://doi.org/10.1017/9781009157896.007. 2021

Lee, K. et al.: Global relationships of total inorganic carbon with temperature and nitrate in surface seawater, Global Biogeochemical Cycles, 14, 979-994, https://doi.org/10.1029/1998GB001087, 2000.

Lee, K. et al.: Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans, Geophysical Research Letters, 33, L19605, https://doi.org/10.1029/2006GL027207, 2006.