Dear Professor Middelburg

Thank you for your comments on our responses to the reviewer comments. We have re-examined the replies we submitted and on reflection can see that we did not clearly and explicitly state how we will address many of the reviewers' legitimate concerns. We appreciate that without more detailed replies and the edited manuscript it is not possible for you to determine whether we have addressed the reviewer comments appropriately. In addition to the revised manuscript that I will now upload, we elaborate on a number of our previous comments and provide additional detail on the implemented changes below.

We wish to highlight that the first reviewers concern regarding the RMSE of the model (which was stated as ~1 PSU in the original manuscript) should now be resolved. We have corrected our statistic for model performance so that it is only calculated for the part of the model domain we use to generate fCO_2, RMSE is dramatically reduced to 0.213 PSU. This makes a huge difference to the perceived accuracy of the model salinity and thus predicted fCO_2.

We now calculate a standard combined uncertainty in fCO_2 which is consistent with the International Bureau of Weights and Measures (BIPM) Guide to the expression of uncertainty in measurement (GUM) methodology. The standard combined uncertainty includes the uncertainty in the ξfCO_2 and ξS relationship (18.95 µatm), uncertainty due to ξS (8.48 µatm) and the uncertainty in the fCO_2 observations (L4 fCO2 = 6.9 µatm, from the on station comparison, Figure 2). These uncertainties add in quadrature to give a standard combined uncertainty of 21.88 µatm. When this uncertainty is shown with the data in Figure 12 we are able to demonstrate that the differences in the average model fCO_2 and fCO_2 at L4 are not due to the uncertainty of our estimate.

We would like to highlight that the salinity-based correction is effectively a bias correction on the L4-based estimate of seawater fCO_2 concentration in the near coastal zone, and results in a worsening of the agreement between L4 and Landschützer estimates.. We inevitably have much more certainty in this correction during the period where there was more frequent sampling and have clearly stated this limitation in the revised manuscript.

The agreement between the two pCO_2 systems on station is due to the respective response times of the two systems. The pCO_2 agreement is better on station at L4 because the ship remains stationary and the water mass does not change drastically, which gives both systems enough time to fully achieve equilibrium. When the ship is moving across a heterogeneous environment, the membrane system (shorter response time) responds faster and can fully equilibrate; the showerhead system (longer response time) is unable to achieve full equilibration. The membrane equilibrator reflects the changes in fCO_2 in the different water masses, but the showerhead cannot do so in such a dynamic coastal environment. This is why the agreement between the two systems is much worse in transit then when at station L4.

The three CO_2 data sources we present in Figure 12a are all at different temporal scales, which presented some problems when calculating CO_2 flux. Calculating comparable CO_2 fluxes required a number of concessions, either averaging or interpolating, and the reviewers were right to point this out. Plotting the fluxes doesn't necessarily add much to the narrative of the paper, so we have decided to simplify Figure 12 by removing the bottom panel (12b) and all associated text.

Thank you for all your time and hard work on this submission.

Yours sincerely

Richard Sims