Dear Editors and Reviewers for ESSD:

We thank both reviewers, Dr. Doney and Dr. Tanhua, for their helpful and supportive reviews and constructive suggestions. We have made edits in response to their reviews, which we believe have strengthened the manuscript. We have copied the review text in italicized text below and placed our responses to the reviews in bold text alongside the associated comments. Red text calls attention to new additions to the manuscript.

We have also made several additional edits to the document:

Most significantly, we welcomed Daniel Sandborn as a coauthor after he provided a detailed review of the manuscript and our code that turned up an error in how the TTD shape parameters were being implemented in TRACEv1. This error has been fixed and all numbers and figures have been recomputed and recreated. These fixes have a small impact on the calculations, and many of the presented numbers are unchanged at the presented level of precision. However, you will notice that some numbers have been updated at the ~2nd significant digit. These changes are always within uncertainties of the original estimates. A new data product was uploaded to Zenodo, computed with the updated routines, and the Zenodo pointers now refer to the updated version. The code on GitHub code has also been updated. We also added some clarifying text associated with the error that was made to the text associated with the calculation of transit time distributions.

Coauthor Schwinger has also created a new Zenodo repository with the simulated CFC and SF_6 fields that we use in this study, and we have added a citation for this data product to the data availability statements.

Finally, we had one additional round of coauthor review that turned up several areas where we felt minor adjustments to the phrasing could improve the clarity of the manuscript.

Thank you again for your time editing and reviewing this submission. Sincerely,

-Brendan and coauthors

The manuscript presents a data-based method for estimating ocean anthropogenic carbon dioxide concentrations. The method can be used to estimate as a function of location, temperature and salinity the geographical patterns for both historical and projected future atmosphere CO2 levels. The "tracer-based rapid anthropogenic carbon estimation" (TRACE) method uses recent compilations of ocean transient tracer field measurements *(e.g., SF6 and CFCs) to compute water parcel age distributions (transient time distribtions or TTDs) as a function of location and water mass properties. Anthropogenic CO2 perturbations to the dissolved inorganic carbon field are then calculated from the age distributions using atmosphere CO2 time histories and assumptions about air-sea dis-equilibrium. Neural network methods are used to approximate the required intermediate steps and final anthropogenic CO2 product across the full range of geographic locations and water mass conditions. The method is tested against various model and observation-based anthropogenic CO2 estimates. The manuscript is through and well written, and it includes sufficient depth on the methods, assumptions, and uncertainties. The TRACE method builds on a couple of decades of research connecting ocean transient tracers to anthropogenic carbon and will likely be beneficial to a range of ocean scientists and other users.*

We thank Dr. Doney for these supportive statements.

Overall the manuscript is a solid scientific contribution, and I only have two issues that I recommend the authors consider addressing. In both cases, I am not suggesting that the authors modify the methodology, simply that they include a little more detail on caveats that may be useful for some readers.

In Line 153 in "Section 3.2 How TRACEv1 work" the text states: "1. uses a neural network to estimate an age distribution for seawater from a user-specified location, T, and S, and returns the mean age if this is a desired output;"

It would be helpful to add some text discussing the rationale for the variables chosen for the neural network. In particular, for many potential users location and depth would be a more straight forward variable set than location, T, and S. The need for T and S are clear for CO2 solubility but not all potential users have knowledge of T and S distributions. Presumably the neural networks are connecting T and S to density to approximate depths of isopycnal surfaces, recognizing that TTDs reflect advection and mixing along isopycnal surfaces as the primary path for the introduction of anthropogenic CO2 into the ocean interior, at least in the thermocline.

Yes, this was our thinking as well. We also had a simple pragmatic reason for choosing these two parameters – they are measured by Argo floats and regularly compiled into time varying climatologies (which will open the door to more applications going forward). We added text that highlights some of our thinking.

L179: While the ESPER_NN routines utilize many combinations of possible predictors, only *S* and *T* are chosen for the TRACE neural networks because they are among the most frequently available predictor measurements and because they collectively represent the density structure of the ocean. Advection and diffusion along density layers in the ocean is the dominant mechanism by which C_{anth} enters the ocean interior, and variations in density, both spatially and temporally, are therefore expected to correlate with the interior ocean distribution of C_{anth}.

On a related issue, T and S fields in the ocean have already changed with time and will evolve further with future climate change; warming of the thermocline (and shoaling of isopycnal surfaces) is already well documented, altering ocean ventilation rates, and these trends will continue in the future, particular under strong climate change projections. The model uses time-invariant T and S fields, thus neglecting changes in ocean properties (including CO2 solubility) and ventilation rates. To incorporate these secondary effects, users could consult climate model projections that include time-evolving ventilation patterns, transient tracer and the projected anthropogenic CO2 fields. There is a caveat in the manuscript to this effect on line 430:

"An important caveat is that these findings do not consider the impacts of changes in heat and freshwater content, circulation, or changes in the ocean's biological pump, and only reflect the impact expected from changing atmospheric xCO2 and ocean buffer capacity." and then again on Line 454: "It presumes fixed circulation and is unable to resolve most timescales and modes of Canth variability." However, it would be beneficial to explicitly include this caveat in the Abstract so readers no up front potential limitations (or alternatively, areas for future enhancement).

We agree and added this final sentence to the abstract:

L34: A notable limitation of this and similar projections made with TRACEv1 is that ongoing and potential future warming and changing circulation patterns with climate change are not captured by the method.

Review of Tracer-based Rapid Anthropogenic Carbon Estimation (TRACE) by Carter et al.

This study use presents a tool that has been developed to estimate the storage of anthropogenic carbon in the interior ocean using observational and model data. It is a valuable contribution to the science on the ocean carbon uptake. The question – how much of the observed DIC in the interior ocean is in excess to what was there before the industrial revolution – is a difficult one, and the authors identify (at least) 3 main hurdles.

The manuscript present 3 products: 1) The TRACEv1 code, 2) uncertainty estimates, and 3) a global data product.

The paper is well written, timely, and of potential high impact, and deserves to be published. I do however see a few areas where I would like the authors to improve on the manuscript.

We thank Dr. Tanhua for the supportive comments.

It is unclear how the variable alpha (that reduce the TTD shape to a single term) is calculated – please provide this information in the manuscript.

We have added some details to the relevant text and now specify the section of the manuscript in which the relationship between alpha and the terms in equation 5 is specified.

L328: The parameters are optimized using a bounded minimum "search" function ("fminsearchbnd" in MATLAB) with an initial value of *α*=1, an upper bound of *α*=1000, and a lower bound of *α*=0.001. This function uses a Nelder-Mead simplex algorithm (Lagarias et al., 2006) with iterative variations of the α term by 5 % to minimize a cost function. For each iteration of this solver, the *j*=3 (i.e., CFC-11, CFC-12, and SF_6) transient tracer constraints and the *A* are first calculated as described in Eqn. 4. The cost function that is minimized for this solver (ε^2) is the sum of the squared normalized errors of the three partial pressures and *A*, or:

$$\varepsilon^2 = \sum_{j=1}^{3}\left(\frac{pX_{\text{meas}}^{j}-pX_{\text{calc}}^{j}}{pX_{ATM_{2020}}^{j}}\right)^2 + \left(\frac{A_{\text{OCIM}}-A_{\text{calc}}}{A_{\text{Max}}}\right)^2 \quad (5)$$

The callback to Equation 4 refers to:

L236: This *α* is used to identify the ages associated with this IG probability distribution, where the age values assigned to the 500 *f*(*x*) values equal [1:500]**α* years. ...

L247: Once the age distribution is known, the atmospheric CO_2 record is convoluted into the age distribution as follows (Hall et al., 2002): First, the atmospheric record or projection is interpolated to obtain values for the year of the desired estimate minus the ages in the distribution. Then for each of the (up to) 1000 fractions of the water mass, the fraction-weighted mean ages ($\overline{A}$) and concentrations ($\overline{[X]}$) can be computed as fraction-weighted sums. E.g. for gas *X* with atmospheric concentration [*X*] summed over the *i*=1:1000 years prior to the estimate of interest this would be computed as:

$$\overline{[X]} = \sum_{i=1}^{1000} f_i [X]_i \quad (4)$$

Line 60: This sentence is difficult to understand. Please rephrase.

We have rewritten this into two shorter sentences:

L65: There are several qualities that are desirable for C_{anth} estimation strategies. Foremost among these is accuracy, but we contend that it is also helpful for estimation approaches to be (1) accessible; (2) computationally efficient; and (3) able to return estimates for the past, present or future.