

Response to Referees

“TRACE-Python: Tracer-based Rapid Anthropogenic Carbon Estimation Implemented in Python (version 1.0)” (egusphere-2025-5793)

Sandborn et al. In review.

We sincerely thank Dr. Jiang and two anonymous referees for their very helpful and constructive suggestions benefiting this work. We appreciate the time and energy that went into these reviews, and hope that they have resulted in an improved work with greater value to our community.

Referee text is reproduced verbatim as indented text.

Our responses are given below each point, with line numbers and section numbers where appropriate referring to locations in the attached revised manuscript.

Response to Anonymous Referee 1

L54: This is a minor comment but since you defined “TRACE” just above you may consider using the short version here.

This change was made as suggested.

L213: Do you mean “uptake” here? (instead of “update”)

Yes. This change was made as suggested.

L258: Is this not the opposite of what is stated in the caption of Figure 6? Here it says, “Lower values of D/G were associated with higher vertical gradients ...” and in the Fig. 6 caption it says, “Higher values of D/G are associated with a higher surface-to-depth mean age gradient ...”.

You are correct: these two statements conflicted. The statement in the figure caption was changed to agree with L258 and the pattern in the figure.

L114: Add a comma after “1.3”.

This change was made as suggested.

L192: ... in Carter et al. (2025) ...

This change was made as suggested.

L195: (Table C1), which ...

This change was made as suggested.

Response to Referee Dr. Wenrui Jiang

Although the precursor paper Carter et al 2025 is available, this paper should be able to convey all the necessary information as a stand-alone paper. I personally find it very difficult to understand what the package does without referring to Carter et al 2025. A section like Section 2.2 of Carter et al 2025 would help a lot. As a reader and potential user, one would like to know what are the input and output of the tool and how it works broadly.

Thank you for this input. At your request and that of Referee no. 3, we have expanded Sections 1 and 2 to include information necessary to understand the fundamentals of the routine and its use.

As an alternative and maybe better solution to the above problem is to have a diagram or some kind of pseudo code. In a diagram, you can show what the format of the input and output are and you can also highlight the new functionalities with a different color.

Thank you for the suggestion. We agree that further explanation of the TRACE routine would aid the reader. To this end a brief numbered list identifies input and output in a more accessible format in Section 2, with much-expanded detail following.

It surprised me somewhat that the vertical interpolation and integration scheme have such a big impact on the overall results. Although I understand the choice of using a spline style interpolation to account for the stiffness of gradient and the existence of outliers, the choice still feels subjective. If you like, I suggest add an example of a vertical profile with different interpolation to justify this choice.

The choice of vertical interpolation/integration routine decreased the 2002 global Canth inventory by 7%, which was indeed surprising, if within our estimated uncertainty. While the method of 1D interpolation is somewhat arbitrary (despite our attempt to pick an interpolation least likely to introduce spurious features), we found that most of the mismatch stemmed from the integration step, rather than interpolation. The original (TRACEv1) inventories were calculated as the sum of grid cells' Canth contents instead of an explicit spatial integral. This "block average sum" introduced some bathymetry intersection errors and failed to resolve sharp spatial gradients, just as you suggest. The present work implemented a more realistic interpolation/integration routine which, while arbitrary, may be more accurate. Further, we feel that a figure of the interpolation method would distract from the thrust of the work, and the reader would better be served by consulting e.g. the Barker and McDougall paper given in our work, L136.

Somewhat related to the previous comment. The salinity in finite-volume models are grid-cell averaged quantity. The lat-lon location of a grid cell centers can also be considered the grid cell averaged

location. Instead of using neural network to predict the concentration at a point, one presumably can first calculate the averaged concentration of tracers in a volume using observations and feed that to the neural network. In this way, there will not be any ambiguity in terms of what integration approach to use.

This is an interesting suggestion, which could be particularly useful for predictor grids of higher spatial resolution or S/T observations from instruments (see also our response to you comment on L318, below). We feel this approach should be left to future work in the interest of providing an Canth inventory for the GLODAP gridded product with utility for comparison to the first TRACEv1 product and other observation-based products.

Do you have any estimate of the effect of seasonality and/or hemispheric asymmetry of CO₂? If not, a brief discussion would still be nice.

Seasonality is neglected in this analysis, as is hemispheric asymmetry in the atmospheric boundary forcing.

The latter could readily be simulated using an alternative CO₂ atmospheric trajectory representative of the latitude of a purported source water for a water parcel of interest, but assigning source waters (and contributions of various source waters) is out of the application space for an IG-TTD analysis. Other tools, such as maximum entropy age deconvolution of Green's functions, are better suited for that task. Text has been added, L86, to this effect.

We suggest that seasonality would likely have minimal effect on global ocean Canth inventories, as the solubility pump is driven by intrusion of dense (generally winter) waters into the deep ocean. The seasonality in atmospheric xCO₂ would apply equally to both the DIC calculation for the increasing xCO₂ and the reference xCO₂. When this is propagated through CO₂sys, a seasonality of 4-5 uatm would translate to 1-3% of variation in Canth. The same reasoning applies to individual Canth estimates below the winter mixed layer. Shallower waters are certainly subject to seasonality in DIC, but the relative variability in Canth is much smaller, with Green's functions for surface estimates yielding near-zero age as expected, such that the "transient equilibrium" assumption holds for TRACE as for e.g. the eMLRC* surface estimates.

There are many important issues that I think should be mentioned early on, but only got mention in the discussion, in another paper, or simply was never mentioned. These include: (1) Is xCO₂ spatially varying?

No, atmospheric CO₂ fraction is not spatially varying in TRACE; see our comments above regarding hemispheric asymmetry. Text has also been added preceding and following Eq. 3 to this effect.

(2) Can the ratio of Gamma and Delta vary in space/time?

In the real ocean, yes. In this estimation routine, no, unless the user decides to vary these parameters in a series of estimates. This was the reason this implementation of TRACE made the δ/Γ ratio user-accessible (L111). Constraining spatial (let alone temporal) variability of δ/Γ would require more data constraints than are presently globally-available from the age tracers discussed in this work, such as tritium, ^{39}Ar , etc. This should be the focus of future work predicated on expanded observational capabilities. Clearer reference has been added to the role of the spatially-invariant δ/Γ input value in Section 2.

- (3) How does preformed quantities affect C_{anth} ? (4) Is there some kind of DIC dynamics in the model, i.e. increased CO_2 decrease alkalinity and prevent future uptake of CO_2 ?

Inorganic carbon equilibria are captured in the routine by use of PyCO2SYS software to represent the response of DIC to the transient CO_2 signal as a function of estimated preformed alkalinity and nutrients. This process is discussed in the expanded Section 2. We should note that increased CO_2 dissolution does not affect total alkalinity; however, changes to buffer capacity as a result of increased C_{anth} (not temperature or salinity) are comprehended by this method via the carbonate equilibria.

- (5) For a future projection simulation, should I use the PI temperature and salinity as input or should I use the projected temperature and salinity?

The most conservative approach would be to use T/S representative of the training data of the product, i.e. the era covered by the GLODAP data. The utility of TRACE lies largely in its ability to propagate transient atmospheric CO_2 signals in the context of a physical steady state, rather than respond to T/S changes. Additional reference to the physical steady state assumption has been added to Sections 1 and 2.

- (6) Does the Green function change with time (does it assume constant diffusivity and velocity field?)

No. The Green's function, as a concise representation of tracer transport, does not change in TRACE for a given point in space as a function of time as a result of the steady state assumption of the product. Additional reference to the physical steady state assumption has been added to Sections 1 and 2.

- (7) When I opt to provide temperature, am I using a different neural network to when I opt out?

No, the same neural networks are used for all C_{anth} estimates. Information has been added, c. L73, to explain this.

L4: "user supplied coordinates, year, depth. . ." There is two issues: coordinates often include depth. Also, why not just say time instead of year? It sounds like year need to be an integer.

The indicated text has been changed to eliminate the reference to depth and change “time” to “year”.

L25-27 This sentence is quite hard to read. I also find this paragraph does not help the narrative much.

Thank you for pointing this out. The indicated text has been revised for clarity and condensed.

Equation 1: Here, I suggest talking about the functional dependency of all the parameters. Is Gamma and Delta function of space? Are you defining $t=0$ implicitly to be the time of tracer initiation? Should $G(t)$ be $G(\vec{x}, t; \vec{x}', t')$? If you can answer some of those questions here it can really help with many of my earlier confusions. Perhaps start with the most general case first and then clearly state how you simplified it. Equation 1: I believe for most readers, it is not always obvious how to get C_{anth} using G . Please write out the convolution explicitly. Also, may as well just spell out the integral to define Delta and Gamma.

The explanation of Equation 1 was significantly expanded to address these questions. $G(t)$ could be written in an explicitly vectorized form, but we opted to leave it as-is to reduce notational clutter and adhere to literature precedent, which we believe will not inhibit understanding with the now-expanded explanation. We also added the convolution of the Green’s function boundary propagator with the surface boundary as Eq. 2, which is described now in greater detail in the specific context of propagating the atmospheric transient in Sections 1 and 2.

L70-71 What is the input output format of the neural network? Without mentioning how it is trained, it is also difficult to know how it works. A diagram would help here.

The description of the inputs and outputs (bolded and italicized for additional clarity), along with a brief overview of its architecture and references to its development, were added to Section 2.

L82 and Equation 2: It is weird to call it a recursive function if $\text{pCO}_2^{\text{oce}}$ does not appear on the RHS. It just has some memory.

The work “recursive” was removed, L81.

Equation 2: the 65 in $t-65$ does not have a unit.

A unit was added to the indicated term of Eq. 2.

L87: Eq2 seems to me like a simple weighted average. To say that it accounts for both the value and the increase and decrease seems like a weird way to say that and it is slightly confusing.

We agree that the wording is confusing, and have altered it accordingly.

L101-111: This paragraph introduce too many new concepts, and many for me are very foreign. I had to find the paper for (Py)CO2SYS to understand what you mean by equilibrium parameters. Either stay at a higher level and not get in the weeds or elaborate on what each new word means.

We have elaborated on several of these concepts at the indicated location and in the following paragraphs to clarify the matter for users.

L112-119: This seems like an awfully long way of saying we should also try 0.2-1.8 apart from 1.3. May be condense the citations to fewer sentences. I also got the wrong impression that the ratio can be spatially varying here.

The indicated section was compressed to streamline the text and make it clearer that the ratio can indeed be user-modifiable to encompass spatial variability.

L126: Cite a few example of the “literature” here

Reference to a paper synthesizing many inventories was added at the indicated location. A demonstration of this application is also given in Section 4.2 for an representative inventory.

L124-128: I was under the impression that it is due to the nonlinearity of DIC chemistry that preindustrial CO2 concentration affect C_{anth}, but I later realize you are actually talking about definition of base lines, is that right? If so, it would be nice to clarify the source of ambiguity caused by the PI baseline.

Canth concentrations of a water parcel varies in this framework depending on at what baseline value (and consequently when) one starts counting. Phrasing to that effect has been added at the indicated location.

L142: Aren't all functions provided in the GitHub repo?

This line was judged to be redundant and removed.

L149: This is probably a layman question, can you say a bit more about what “tolerance for pH change” mean? Is the 1e-3 1e-4 in the log pH unit? If so, can we measure such small changes in pH in observation? I think a descent pH meter can only measure as accurate as ± 0.01 .

Thanks for pointing out this ambiguity and we've now clarified at the indicated location that the tolerance here is the lower step limit for the iterative system of equations solver utilized in CO2SYS.

L154-155: This seems like a rather small performance increase. Unless there is some reason to drastically scale up this calculation or repeat the calculation many times, I think most users don't mind wait 10 more seconds for the result. I suggest simply stating that the performance is comparable.

We agree: the performance is for most uses is comparable. This particular statistic was requested by colleagues interested in the relative suitability of TRACEv1 and TRACE-Python for rapid and repeated state estimates for operational product purposes, some of which require millions of estimates over time for global high resolution models. We opt to leave the numbers, but add the clarification that the times are comparable.

Table 1: Can you add the uncertainty to each results. I understand this table is trying to demonstrate the consistency (which is pretty impressive btw). It still looks like you have 9 significant digits.

Uncertainties were added to the check values in Table 1 as suggested.

Figure 2: I suggest moving the small 10^{-5} to the y axis label to make it more visible. Same for the other figures in the appendix. Also, just out of curiosity, where are the points that have larger deviation?

Thank you for the suggestion for the plot. This has been implemented by changing the units to pmol/kg.

The larger deviations are usually associated with points in or near marginal seas, coastlines, or areas of particularly unusual inorganic carbon chemistry, leading us to believe these points were the most affected by differences in CO2SYS iterative solver approaches, as mentioned briefly in the text. Many other deviations were stubbornly difficult to associate with a cause; a couple options being differences in machine representation of NN weights and differences in geographic interpolation among basins. An investigation thorough enough to warrant extended description in the literature is beyond the present work's scope or needs, but future investigations may address it if the need arises. Compare also our work on ESPER ports for more information and discussion.

L197-198: This comparison is very nice indeed. I suggest also list the result of the original interpolation here. I am sorry for my paranoia, but when I am reading this I thought you may have developed several vertical interpolation/integration schemes and picked the one with the best agreement. I am not accusing you of doing that, of course, but I think other reader may also have the same gut reaction. If you could make a stronger case for the Hermite splines, or say something like "the interpolation scheme we chose happens to improve the agreement", it will help with this issue.

Thank you for the feedback. The choice of interpolation scheme did indeed occur independently from comparison of previous inventories, and indeed, the weaker literature comparison with the erroneous calculation was published in the TRACEv1 manuscript. It wasn't until the discrepancy between TRACE-Python and TRACEv1 was noticed that an erroneous calculation for grid cell vertical width used for TRACEv1 was identified. The choice of interpolation schemes had a secondary impact compared to this bug fix. A disclaimer along the lines

of your suggestion was added at the indicated location.

L248: Do diapycnal diffusion and/or deep convections play no role in the age distribution? Please clarify.

Diapycnal diffusion is not explicitly represented in the IG-TTD framework. Individual deep convection events are not represented explicitly, as IG-TTD represents only average circulation and tracer transport; this doesn't mean that well-ventilated signals of e.g. LSW or NSOW are neglected, only that interannual variability is absent. This information has been added to Section 1.

L277: what is ESPER? Please provide a citation or at least provide the full name.

Thank you for drawing attention to this oversight – an explanation and citation have been provided at the indicated location.

Figure 6: provide a map of WOCE A16 or at least say which basin it is in. The panels look a bit too similar to me. Perhaps plot their difference instead?

An inset transect map was added to aid in contextualization, and the second two rows were recast into relative differences.

L291-300: Much of this should be mentioned earlier to reduce some confusions.

This makes sense. We expanded the introduction and summary of inherited method (as discussed in a previous comment) to include much of this important information.

L318: Does the model only work with GLODAP? Can a user use a different model/reanalysis?

Any reanalysis or model output product may indeed be used – these are elements of continuing work. An addition was made to L317 to that effect.

Appendix D: I am not sure what the purpose of this appendix is. Why is the 2020 CO₂ only 300 ppm?

As explained on L241, these distributions illustrate effective global ocean circulation-informed pre-industrial xCO₂ distributions for common starting points of ocean state estimates, and support the claim for the development of bimodal distributions of ventilated waters. They are left in the Appendix to avoid cluttering the manuscript, but may be of interest for future comparisons of ocean state estimates and especially model initialization.

Response to Anonymous Referee 3

I find it difficult to get an overall picture of how the TRACE method works from Section 2. There are many components in the TRACE method. But, it is difficult to see how they work together to convert various inputs to the final output.

In particular, there is no equation showing how the pCO₂ boundary condition in Eq. 2 is combined with the IG-TTD kernel in Eq. 1 to produce the point estimate of C_{anth} in Fig. 3. I suggest that the authors include a flowchart to clearly show how the various inputs go through different neural networks and PyCO₂SYS to produce the IG-TTD kernel and the boundary condition, and then the C_{anth} estimate. Such a diagram would help readers quickly understand the basics of the TRACE method without having to consult multiple references individually.

Thank you for this input. At your request and that of referee Dr. Jiang, we have expanded Sections 1 and 2 to include information necessary to understand the fundamentals of the routine and its use. Also, the convolution equation was added as Eq. 2. We feel confident that the additional information is sufficient to aid future users in understanding and deploying the method without excessive external reference.

We have opted to include a numbered list of steps rather than a flowchart, which should promote the easy reference desired from a flowchart without extraneous figures.

TRACE-Python allows users to change the shape parameter (δ/Γ) of the IG-TTD distribution from the default value of 1.3 to other values. Lines 70-72 suggest that the default IG-TTD distribution is constrained by observations of CFC-11, CFC-12 and SF₆. But it is not clear whether the IG-TTD distributions with altered shape parameters are constrained by observations as well? If not, this suggests that the C_{anth} estimates with altered shape parameter (e.g. those in Figs. 5 and 6) are not observationally constrained, hence less reliable than the default one.

The analysis of Fig. 6 suggests that altering δ/Γ also leads to a change in the mean age Γ . I am not sure how this works. In theory, one could change δ/Γ by only changing δ , without changing Γ . Is this done by the neural network that produces IG-TTD from input parameters? In addition, I find the description in lines 121-123 regarding the neural network difficult to understand. It may be useful to briefly explain how the neural network was trained in the first place.

We did retrain TRACE using the data constraints for each of the Delta/Gamma values, and the routine chooses the neural networks appropriate to the ratio

selected by the user. We feel this is a misunderstanding arising from a lack of clarity in our text. The indicated sections have been expanded to make it clearer that the optimization of the inverse-gaussian functional form of the TTD is constrained by the ratio of its moments δ/Γ , leaving a restricted solution space which may be summarized as a range of mean ages. This in turn explains the change in mean age illustrated by Fig. 6.

One could, indeed, get a change in δ/Γ by changing only δ and holding Γ constant, but the exercise may not describe real ocean circulation well, for δ and Γ tend to covary along a water mass (consider an analogous process of an inkdrop moving and spreading in a river). We opt to provide several references to the use of these moments of the IG-TTD distribution which provide a more detailed treatment of the fundamentals of TTD functional behavior than can or should be included in this work.

> Lines 212-217 It would be useful to compare the projections from TRACE-Python with those from CMIP Earth system models here.

Thank you for this suggestion. We added comparisons to the GOBM ensemble considered in the RECCAP2 analysis to sections 4.1 and 4.2. These expanded our analysis of sensitivity of observation-based Canth inventories in a very useful direction.

Line 252 “This contrasts with the findings of He et al. (2018) . . .”
Why is your finding different from He et al.? Please add a discussion to address this discrepancy.

The most significant difference between the two analyses is the incorporation of OCIM ages into our work, which likely stabilize the optimization of the IG-TTD in waters predating the transient tracers. Reference to this has been added at the indicated location.