1 Reviewer 1 (Anonymous Referee 2 from revised submission)

Minor Comments

- L46-48: Thank you for pointing this out - as the reviewer says, this should have read added/passive heat and anthropogenic storage locally. This has been rewritten to correct this.
- L68-69: This is a good point. The precise definitions of excess redistributed tracers were inadvertently removed from the rewritten theory section. This precise definition has been returned to the text, which should help to clarify the statement.
- L92-96: Thank you for pointing out this ambiguity - the text here has been rewritten to clarify this point. As the reviewer correctly identifies, our approach builds on the anticorrelation noted by Williams et al. 2021 between the background fields of temperature and DIC.
- L97-99: Again, thank you for pointing out this ambiguity - the text here has been rewritten to clarify the point being made.
- L100-103: We have reconfigured the text here to address potential ambiguities and better describe the methodological approach: specifically, we describe how the decomposition simply relies on the assumption that we can identify a clear spatial relationship between natural carbon and the variable we wish to decompose.
- L129: The reviewer is correct that Δ refers to changes from 1860 to an arbitrary time t. As noted by the reviewer, it is later defined, and we have moved the definition up in the text as suggested.

Section 2.2

- a) (Equation 12 (now 13)): Thank you for spotting this, this has been fixed.
- b) As the reviewer correctly points out, we are using the arguments of Williams et al. (2021): that the background temperature and carbon fields are generallly strongly anticorrelated, and that we may assume that concentrations of natural carbon (Cnat) change only through redistribution. However, this anticorrelation varies spatially (through its soft tissue and carbonate components) and differences in the air-sea disequilibrium states of water masses when they are formed, so it is necessary to allow for this variation in order to improve estimates of local temperature (and salinity) redistribution. As suggested by the reviewer, we have added maps of κ_r^T and κ_r^S to Appendix D to illustrate their spatial variability, which we feel should help to address clarify the point raised. We have also rewritten the statement regarding the use of the word 'similarly': as the reviewer correctly points out, these parameterisations are only similar from the perspective of both encoding one of the arguments of Williams et al. 2021 each (the correlation of excess/anthropogenic carbon and excess temperature; and the anticorrelation of background/natural carbon and redistributed temperature): we have reworded this in response. Additionally, we have clarified that our use of subdecadal timescales in order to estimate $\kappa\_r\hat{T}$ and κ_r^S is in order to allow us to estimate the spatial covariability of the background fields, without it being necessary to have a decomposed temperature or salinity field at this point of the process. Regarding the appropriateness of this process over the RCP8.5 period, we have clarified that this should not be a problem due to the way in which the redistributed fields are designed and these correlations are calculated, with reference to the definitions mentioned in L68-69.

  As we have practical limitations on performing additional PAT experiments (see the point below on explicit validation), we are unable to show directly the correctness of Equation 7 (now Equation 8). However, we have instead added an appendix (E) in which we compare the estimates of total heat content change and heat content change due to redistribution in regions of the ocean which can be considered to be unventilated (determined using regions of the ocean where concentrations of anthropogenic carbon are negligible throughout the model run). Though this is of course an imperfect validation of Equation 7 (now 8), we are able to show that during the early 20th century, the high frequency (sub decadal) variability in the heat content of unventilated waters are very well captured

by the redistributed heat content of this region, and that over the 21st century, the ratio of the two asymptotes to approximately 0.8, indicating our method is reliably capturing at least 80% of the heat content change due to systematic redistribution as a result of climate change. We do note however that this is an imperfect validation due to the presence of other atmospheric forcings besides CO2, and so represents an upper limit of the uncertainty of our method: considering slightly shorter periods greatly improves the accuracy of the reconstruction, and we believe this is a bias introduced by the validation technique rather than the method.

Regarding the applicability of our technique to other tracers, the reviewer is correct to point out that we have not specifically shown that the method is applicable to other tracers. We have thus rewritten this statement to clarify that our point is merely that the assumptions and premises of our method are such that they can theoretically be applied to other tracers. The practicality of this will be investigated in further studies.

2 Reviewer 2 (Anonymous referee 3):

Minor Comments

- L14: Thank you for pointing this out, we have rephrased this and expanded context to make the statement clearer.
- L29: Thank you for pointing this out, this has been fixed. Full stops have been added to equations which require them, and we have replaced upper case with lower case theta as pointed out.

Code availability & test cases

As suggested by the reviewer, we have added a number of test cases and sample code/data to a GitHub repository (detailed and linked to at end of manuscript), following the standards of Irving et al. 2016. For ease of use and to minimise file sizes, we share a subset of the full dataset that enables a user to calculate a gamma value and decompose temperature for a single section of ocean (across the subtropical North Atlantic) for a single year, entirely reproducing our own decomposition for this region. We believe the available code should also enable users to apply our decomposition method to other simulations.

Glossary / Table of Definitions

The reviewer is correct to point out that a number of terms and concepts are introduced which may make the paper difficult to follow for a reader who is not already familiar with these terms concepts. The reviewer suggests introducing a table of definitions / glossary in order to make the paper more accessible. As we have expanded upon the definition and interpretation of terms (as noted in our response to Reviewer 1), we feel that also including these definitions in a table within the main text would become repetitive. However, we agree with the reviewer that having definitions and descriptions grouped together would help improve the accessibility of the paper, and so we have added an Appendix containing the definitions and interpretation of these terms to the paper (Appendix F).

Explicit Validation (PAT)

We agree with both reviewers when they observe that it would improve the paper were we able to directly compute the errors on our estimates of temperature redistribution. Unfortunately, these simulations were performed >5 years ago on an HPC cluster which no longer exists. Rerunning the simulations and including a passive anomalous heat tracer is thus simply not practical due to the financial, infrastructure and time investment necessary. We hope that our theoretical justification and the favourable comparison with the validated alternative carbon based methodology of Bronselaer & Zanna 2020 supports the use of this new methodology to deconvolve the components of total heat and salinity changes.