

Ocean Sci. Discuss., author comment AC1
<https://doi.org/10.5194/os-2021-87-AC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Reply on RC1

David S. Trossman et al.

Author comment on "Tracer and observationally derived constraints on diapycnal diffusivities in an ocean state estimate" by David S. Trossman et al., Ocean Sci. Discuss., <https://doi.org/10.5194/os-2021-87-AC1>, 2021

Our responses are in boldface:

- State estimates like ECCO take an important place in our field. It bridges a gap between observations and numerical models. Like observations and numerical models, it has its own pros and cons. This paper particularly looks at vertical mixing resulting from the state estimate and tries to understand and improve this number. It does so by first establishing how good it is (comparing to observational products), and then trying to add the tracer oxygen, to see if this could provide extra information to constrain and improve vertical mixing estimates in ECCO. Although it seems to do so, the results are somewhat disappointing, as the authors also notice. Regardless, this is a result in its own and therefore worthy contribution to publish. Before publication, I do think there are aspects of the analyses that need to be addressed and places where clarification to be made. Hence, with some major revisions, I think this could become a good publication. Comments are in order of text and thoughts. Major comments indicated.

- **Thank you for the thoughtful review. We have performed additional analyses as well as added some textual clarification, as described below.**

- L20-25 – GM is not mixing, I think. It is advection.

• **The GM coefficient has been considered a “mixing parameter” in the literature, but you are right that it is effectively an advective term. We have changed the text to say “Ocean mixing is typically conceptualized in terms of diffusion along and across isopycnal surfaces. Subgrid-scale transport of isopycnal thickness (or bolus), which is effectively an advective contribution to tracer budgets, also must be parameterized. Ocean models often represent these unresolved processes with three parameters:...”**

•L35 – I don’t understand what is meant with the rotation derived component. . . Is this about along isopycnal rotation? Perhaps better to explain this more carefully.

•**The formulation of Katsumata (2016) separates out the parallel and perpendicular directions of the horizontal components of the eddy advection tensor to density contours on a 1000 dbar surface. The perpendicular direction is the component that corresponds to the Gent-McWilliams scheme, while the parallel direction includes the rotational eddy flux, which is part of the eddy flux not in the Gent-McWilliams scheme. The formulation of the rotational component of eddy transport in ECCO is different from that of Katsumata (2016), which is why a comparison was not performed in that paper. We have modified the text to say: “... the formulations of the perpendicular and parallel components of the eddy advection tensor relative to isopycnal surfaces are not the same in many models as in the observationally-derived Gent-McWilliams coefficient product.”**

•Major L115 – The results are compared to observations of mixing. This is good. But I think it would be worth to compare the results against the parametrization of Lavergne et al 2020. This is semi-observational estimate compares well to Whalen’s work, but covers all grids and full depth. Without data-gaps this may provide better comparison product for the model. The work is also “averaged”, like that from ECCO, and not instantaneous like that from observations. Another reason why this might be a better comparison for this effort. Of course, neither observation, nor Lavargne’s method are perfect for comparison, but having them both really adds value.

•**We now include the diapycnal diffusivities from de Lavergne et al. (2020) in our manuscript. Below about 2000 or so meters, the diapycnal diffusivities from ECCO are too small in comparison to every observationally-derived product. However, in contrast to the comparison with microstructure-inferred and the Argo-derived diapycnal diffusivities, the diapycnal diffusivities from de Lavergne et al. (2020) are smaller than the ones from ECCO in the upper 2000 or so meters.**

•Major - L120-125 - I think you should update WOA13 to WOA18, at least 5 years more data, including BGC data. This is certainly worth the upgrade, and it should be straightforward. Also, which is used, monthly means, annual means?

•We used annual means because ECCO uses a constant diapycnal diffusivity field. At the time we began this project, we also chose to use WOA13 (WOA18 wasn't available yet) because the initial conditions for oxygen concentrations in the model are derived from WOA13. Thus, the regions where there are disagreements between WOA13 and the model are regions where model drift must be occurring due to model specification/parameter errors. The adjoint sensitivities tell us how to change the model parameters so that the model won't drift as far from the initial conditions. Further, much of the extra data in the WOA18 database is at high latitudes where we don't have diapycnal diffusivity data from microstructure or Argo. If you still find this to be an important update for our manuscript, we could perform the adjoint sensitivity runs again, but we would need to get computer resources to run the additional simulations.

•L170 – for by?

•Corrected, thanks.

•Major - L200 – N2 is used from WOA13. However, WOA13 has problems of its own: to start with it is averaged on isobars and not on isopycnals. This leads to false mixing and we also see that N2 is often unstable. However, on average time scales of months, the ocean is “stable” as overturns generally have much shorter duration. Have you made N2 stable? There are simple tools for this in the TEOS-10 manual. I think this is important for your comparison.

•Yes, we did not explain this, but we use TEOS-10 to calculate N2 from WOA13 data so that it's stable. We also repeated our calculations using the N2 provided in the de Lavergne et al. (2020) data and found no qualitative difference.

•L220–Fig2a=fig2

•Corrected, thanks.

- Major - L225-230 – Equation 1 and surrounding text. The results are sensitive to prior choices made by the modeler, in the form of weights. Here a certain choice is made, but sensitivity to this choice is not studied. The sensitivity is both to the equation weighting (as done here), and to the variable weighting. The latter means that the a-priori estimate of the variable, influences the result. This should as well be considered somehow because this a-priori estimate can be wrong or vary within a range. I see no sensitivity to either of these choices. Or at the very least, a discussion on this possibility. Places to read about this are McIntosh and Rintoul 2001, but also Groeskamp et al 2014. I'm aware that this can be a pain and sometimes limited by computer power. I don't know the situation for this study. I therefore strongly recommend to do it if you can, it will strengthen the confidence in your model and the results. IF you really can't do this, then at least discuss it.

- **In regard to the choices made by the modeler (equation weighting), we performed Monte Carlo simulations to consider the uncertainties in the diapycnal diffusivity estimates as well as the uncertainties in the weights themselves for the offline estimates of the adjoint sensitivities. Our results are very close to the same when considering only the uncertainty in the diapycnal diffusivity estimates. As for sensitivity to variable weighting, because we only consider one variable at a time in the context of adjoint sensitivity experiments (not optimizations), we do not consider variable weighting as a factor, but in practice, if oxygen is used to help constrain the ECCO- estimated diapycnal diffusivities, then this is an issue to consider. We show how the first iteration's estimate of the diapycnal diffusivities is different from the final iteration's estimate because we agree with this point. We now include the following statements in the discussion section: "The ECCO-estimated k_p can also be sensitive to the a-priori estimate of k_p and we showed how one particular initial guess ($10^{-5} \text{ m}^2 \text{ s}^{-1}$ everywhere) can evolve from the first optimization iteration to the final one. ... If biogeochemical tracers are included in the misfit calculation in an optimization run, their impact on variables such as k_p would depend upon how they are weighted relative to the physical variables (e.g., temperature, salinity, and pressure)."**

- 240-245 – Below eq 2, do you here mean to fill in equation 1 into 2, so you then get a "y". You refer to "y", where there is none in eq 2.

- The "y" refers to the y in Eq 1, but we have removed all references to variables in Eq 1 after Eq 2 for clarity.**

- Section 2: You focus on analyzing ρ_{krho} . But what about all other variables. It could well

be that if you better constrain K_{ρ} , error is increased on other variables. Hence, how do you keep watch that the rest of the variables are not underperforming?

•It is likely that errors in one of the three mixing parameters bleed into each other, which makes each of their values biased. This is an issue that can potentially be resolved in a similar way that issues with atmospheric forcing fields have been handled, which is to only allow them to vary within a certain range of reanalysis products/observational estimates. This issue is important, but is a next step beyond the scope of our current study because we are motivating new optimizations of ECCO with our manuscript, not attempting them until we have reason to do so. Sidenote: in a previous version of this manuscript, we included results for the adjoint sensitivities of oxygen with respect to the diapycnal diffusivities, the Redi coefficients, and the Gent-McWilliams coefficients. They're all different, of course, and some are larger than others. The adjoint sensitivities for the Redi coefficients, for example, are the smallest, and there are many studies that have examined how strongly oxygen can respond to changing the Redi coefficients (e.g., Gnanadesikan et al., 2013). The same type of analysis we have performed with the diapycnal diffusivities was performed with the Redi coefficients, and the signs of the adjoint sensitivities agreed about 50% of the time (what you'd expect from random chance). This could be because there are imperfections with the scaling factors used in the calculation of observationally-derived Redi coefficients; in fact, if we scale down the observationally-derived Redi coefficients by a constant factor between 2 and 3, the percent volume agreement between the adjoint sensitivities goes up to 60-70%. More work should be done on observationally-derived Redi coefficients before we used them for our analysis. And the same analysis cannot be done with the Gent-McWilliams for reasons we've explained and are discussed in Katsumata (2016).

• Line 283-284 – it says, “not shown”. I would have liked to see this in the same figure, I think. This would be informative, and it sounds like you have the data anyway.

•Because the uncertainty is a factor of three and we're plotting the profiles on a depth versus log-diffusivity scale, the uncertainty looks like it tracks the profiles linearly, just offset by about one-third on the abscissa. So we didn't think it would be important to explicitly show the uncertainties. However, given your next concern, we decided to show the uncertainties in the microstructure profiles of the diapycnal diffusivities for both these average profiles and for individual campaigns to demonstrate how ECCO's diapycnal diffusivities can be within the uncertainties for some campaigns with large uncertainties and well outside of the uncertainties for other campaigns.

•Major - Section 3.1 – Observations are instantaneous profiles in a time varying field. We here are dealing with averages in a model. So I was wondering, is taking an average to compare the right thing to do? Is it not better to check individual profiles and add their error together somehow? I mean, averaging a diffusivity is always a strange thing to do. Okay, we must do something, I agree with that. But it seems like many profiles that are physically unrelated are now averaged in one big profile that therefore loses all sense of information. Is there not a possibility to compare and show some selected individual profiles and come up with a metric that addresses the error between the obs and the model, for the sum of each individual profile? I think that would be more meaningful.

• **de Lavergne et al. (2020) did this kind of comparison for the reasons you argue. We argue in the Appendix that the diapycnal diffusivities don't significantly vary over time except in a few select regions of the ocean, like in the top 1000 meters between 40-50°N in the North Atlantic Ocean. So taking an average (a geometric one, in particular) makes sense to show. However, we take your suggestion and now also show comparisons between the diapycnal diffusivities from most example microstructure campaigns and ECCO. We also show the approximate factor of three uncertainties on the microstructure profiles in these comparisons.**

• Grid lines in Fig 3. And Fig. 7 would be very helpful.

• **We're not sure what you mean because those figures are not maps, but we have edited both of these figures in any case. We hope they're presentable.**

•L295 – Do air-sea fluxes influence the results here or is this stuff too deep. What about advective processes?

• **The diapycnal diffusivities are all below the mixed layer and our comparisons are therefore for values below the mixed layer. However, over long (multi-year) time scales, the air-sea fluxes could influence the results because, as we show in the Appendix, the steric sea level budget term related to the diapycnal diffusivities are associated with the steric sea level budget term related to the air-sea fluxes. Thus, we have changed the text to say "The errors in $\kappa\rho$, ECCO could be partially compensating for errors in the vertical component of the along-isopycnal diffusivity tensor, erroneous air-sea fluxes due to inconsistencies between the sea surface and atmospheric forcing fields, and/or the presence of numerical diffusion." As for advective processes, if they influence the gradients in oxygen concentrations, then that could explain at least part of the**

disagreements in the signs of the adjoint sensitivities because the diapycnal diffusivities are not the only factor influencing oxygen concentrations. However, our results suggest that oxygen could play an overall positive role in estimating a more realistic diapycnal diffusivity field. We will need to perform an optimization to show this definitely.

- L310 - There is no table 2.2.

- **Corrected, thanks.**

- Section 3.3 discussion figure 5. How would this figure look if the differences where small? I mean, it would still be red or blue right? It only gives a sign, but does it say something about the magnitude of the mismatch? It only seems to give a direction of the mismatch. Is it not better to also say something about the magnitude of the mismatch?

- It depends on the sign of the small value (could be red or blue). If the value is small then the sign of the adjoint sensitivity shouldn't matter much. What we want the reader to see in our previous version's Figure 5 is the correspondence of signs between the different adjoint sensitivities. We white out regions where the differences are small by an uncertainty-based metric in the following figure (Figure 6 of the previous version) to quantify the percent volume over which the signs of the adjoint sensitivities agree. We show the magnitudes of the adjoint sensitivities (in Figure 7 of the previous version) when we show the scatterplots and quantify their correlations. In short, we focus on the agreement in sign in the figure in question, we indicate where the magnitudes are small in the following figure, and we focus on correlations between the magnitudes with our final figure.**

- L342 – We cannot compare white and white. Please make these locations referred to, gray in figure 6.

- Here we're referring to the white regions in Figure 6 that are not white in Figure 5.**

- Major - L360 – Oke, O2 gives extra information. But which part? Is it the vertical gradients of O2? If so, when do they give extra information compared to N2? It should be possible to argue this upfront and discuss (not do, yet) which other tracers could be of interest in future work. That is, when N2 does not have vertical gradients, but O2 does, then for that region it is worth adding O2. If both don't have gradients there, then probably some other tracer may work. A discussion on this would be helpful. Possibly a figure.

- The similar magnitude of the correlations between the vertical gradients of oxygen and the diapycnal diffusivity estimates and the correlations between the adjoint sensitivities suggests that the information that oxygen provides about the diapycnal diffusivities is related to the vertical gradients in oxygen. Our comparisons with the dissipation rates instead of the diapycnal diffusivities suggests that N2 doesn't appear to provide any information in addition to that which oxygen provides, but we did an additional calculation where we check whether this conclusion is contingent upon the vertical gradients in N2 and in oxygen. We found that the correlation between the vertical gradients in N2 and in oxygen is about 0.25, which is higher than the correlation between the diapycnal diffusivities and the vertical gradients in oxygen. This is not perfect, so we performed our comparisons between the adjoint sensitivities using only regions where the vertical gradients in N2 are relatively small and the vertical gradients in oxygen are relatively large. The correlations are about the same, again suggesting that our results are independent of the stratification. We have edited the text to say: "The spatial correlation between the annual mean vertical gradients in oxygen and the annual mean vertical gradients in stratification (N2) from the World Ocean Atlas (2013) is about 0.25, but this is only suggestive. To determine whether oxygen provides information about stratification (and through stratification, about $\kappa\rho$),..."**

- Section 4 – Could oxygen also increase the errors? It is not said that things get better. I think it is argued that the observation of O2 is accurate and thus it provides information. But for some places where there is lack of data or the values are not averaged the right way in constructing the climatology, errors could be substantial. You have regions where estimates improve, and where they get worse. This could be why. This is related to the previous comment. Again, a discussion would be helpful.

- The diapycnal diffusivities would get worse in regions where the signs of the adjoint sensitivities disagree if oxygen is the only information provided as a constraint. In practice, oxygen would be weighted as a constraint along with temperature and salinity, which may not lead to worse diapycnal diffusivities. This would take some experimentation to get the weights such that the diapycnal diffusivities are improved. All we are saying here is that the diapycnal diffusivities could improve if we include oxygen in the misfit. It is a fair point to**

suggest that the agreements between the adjoint sensitivities are highly correlated in regions where we have diapycnal diffusivity and oxygen data and/or the diapycnal diffusivities might actually have their errors increased where we do not have data. In fact, when we mask out all regions where we don't have both diapycnal diffusivity data from de Lavergne et al. (2020) and oxygen data from WOA13, the correlations between the adjoint sensitivities go up to 0.47 in the Monte Carlo simulations. We added the following sentence: "If we only consider comparing locations where we have both observationally-derived κ_p data and oxygen data, our results are qualitatively the same and the correlations increase to as much as 0.47 in the case of the de Lavergne et al. (2020) data using a Monte Carlo approach."

•L380-381 – I think with everything going on in ECCO also leading to errors compared to the real world (e.g., parameter choices), I find it is too strong a statement to simply say that data alone is insufficient. You do tone down this point a little in the sentence below, but I still think this statement as written down here is too strong.

•**Okay, we have rephrased this statement: "Some model specifications would lead to errors in κ_p , ECCO even in the presence of globally complete hydrographic observations (see Section 4.2), but we investigated whether κ_p , ECCO can benefit from new information."**

•As mentioned before, I think this might also be a place to discuss that the data product itself (WOA) has errors in there due to interpolations, and averaging techniques (e.g., horizontal instead of isopycnal averaging).

•**This is true, but we only included observations where they were taken, not interpolated/averaged (except in the vertical), in the misfit function in the simulations we performed. We have added the following sentences: "We assumed a factor of three for the observationally-derived κ_p and 2% of the oxygen concentrations. These do not account for interpolation/averaging errors that entered the data prior to our calculations, but are conservative estimates nonetheless. The observational uncertainties affect the weights given in the misfits that enter the adjoint sensitivity calculations and our Monte Carlo simulations of the correlations between the adjoint sensitivities account for the possibility that these weights are misspecified."**