Reviewer 2

The authors present an observational framework that transforms mooring time series into spatial, cross-stream distributions, which are used to estimate isopycnal eddy diffusivity from mixing lengths in the Antarctic Circumpolar Current. Overall this work is interesting and very relevant to property distributions in the Southern Ocean. My main reservation is technical and related to the assumptions of the mixing length method. An ad hoc cutoff is used, which seems inadequate or at the very least needs some more justification. A second reservation is that the length scales measured in the various datasets is not clear and these length scales will affect gradients. Lastly, there are a few comments to improve presentation, which is for the authors to consider rather than a requirement. These suggestions include: reducing subplots and plotting in density coordinates. Overall, it is an excellent start and with a little more work this manuscript will be a solid contribution.

We thank the reviewer for their thorough review and constructive feedback on our manuscript. We have made a number of changes in the data visualisation, the explanation of our methodology and the writing style that we feel improved the paper. Below we give detailed replies to all comments.

Major comments:

1. Do the assumptions required for the mixing length method hold? Are eddies doing most of the mixing? Is no variance advected from upstream? Does the background vary slowly over the mixing length?

Maybe some of the data are obtained in the mixed layer, in which case the mixing is not due to eddies. The mixed layer should be excluded from the calculations. Argo could be used to examine the advection of tracer variance in the area- maybe this is already in Cole et al (2015) or Roach et al (2018). If the mixing length is 1000 km and the section is only 200 km over which we see isopycnal shoaling, then the mixing length is probably not valid because the background is changing over 200 km. More explanation is needed to show how the data over a limited range can support much larger mixing lengths.

Unfortunately, it is not possible to verify the assumptions for the mixing length method with the data we have. Therefore, following NG11, we just proceeded with the assumptions. This is a caveat that is already in the NG11 methodology. In this article, we do not aim to address the caveats of that method, but rather take the opportunities the method presents and adjust it to work for mooring data as well, increasing its versatility. This is important to mention in the article though to manage the reader's expectations, so we added this at the beginning of Section 4 (lines 181-184).

Although we cannot verify the assumptions a priori, we can still consider whether they make sense a posteriori by looking at which of the resulting values make sense and which do not. We discuss this in Section 6.2 (lines 409-433).

Regarding the mixed layer: this should not be a problem since we only consider data at 500 m depth or deeper. As seen in e.g., Close et al. (2013), the winter mixed layer in northern Drake Passage to the north of the Subantarctic Front reaches, at most, ~500 m depth, such that its influence in our analysis should be at most marginal. This is now noted at the end of Section 4.2 (lines 251-253).

Regarding the section width: the total SR1b section is 700 km wide. The low gradient cut-off

ensures that the mixing lengths do not become greater than 700 km, which would indeed be invalid. We mention this now in lines 243-244.

2. In the same vein, about 2/3 of the data for section SR1b is below the cutoff (Fig 7a) and so it's not clear how useful these data are. Argo seems better but certain depth ranges near 1500 m and 3000 m seem unviable too (Fig 7b, c).

So is this a problem with the low gradient cutoff? Or is it a more serious problem with the method? Maybe it's not possible to convert time variability to spatial variability because the assumptions of the mixing length method do not hold here.
The main problem is likely in determining the gradients (see also Section 6.2). Not having the low gradient cutoff would lead to unphysically large values of the mixing length. The section data is more affected than the mooring data, likely because the section has fewer data points. The time-to-space conversion is not the problem, because the issue of weak gradients is also there for the section, which is a spatial dataset.

The Argo climatology is objectively mapped using certain time and length scales. How do these compare to section SR1b. The differences in their methods of making a gridded dataset will make a difference in the gradients used here. These should be explicitly mentioned somewhere,
In the end, what we use from the different products is just the phi-Y fit. The gradients are then computed just from the mooring data. The differences between the phi-Y fits are visible in Figure 2, plus the spatial resolutions of the different datasets are mentioned in Table 1.

Could some of these problems be mitigated with coarser binning in density and coarser binning in space? Then there are more data points per bin. Or would finer binning in density help in converting depth to density and vice versa?
Both points are valid. We chose to stick with the finer binning to more accurately convert density to depth and vice versa, following the value used by NG11. The bins are fine enough to have good conversion, but contain enough data points to have robust spline fits for almost every density level. This is motivated in lines 226-227.

3. Why does Fig 7 only extend south of the mooring? Are there no streamline displacements to the north?
There is some confusion here caused by a poor choice of data visualisation. The dots in these figures show the instrument depths along the mooring, but they are not shown at the location of the mooring, but rather somewhere on top of the data. We changed this so that the dots are now shown along the vertical axis.

4. To make the manuscript more concise, I think some of the comparisons to the various data/climatologies could be moved to supporting info and the number of virtually identical subplots in a figure could be reduced. Much of what is the appendices could also go in supporting info.

These various datasets strongly support the main point but are not essential to understanding it. Having these multiple additional subplots in the figures in the main text is an extra burden on a reader. A similar argument is on lines 260-266.

So maybe just chose your favorite conversion.
We agree that we can reduce the number of panels in the main text, so we have removed the panels showing the mooring data using the Argo dataset for phi-to-Y conversion from the main text; these are now only in the Appendix. The policy of OS is to include normal size figures that do not need to be included in the main text as appendices rather than supplements (https://www.ocean-science.net/submission.html#manuscriptcomposition), so we kept the figures as appendices.

Continuing in the same vein, there are places where all the various subplots are described, such as Section 5.2. The description is all valid but the main idea is maybe a bit lost in all the descriptions here and elsewhere. I believe the main point is on lines 10-13: "The mooring-derived eddy diffusivity distributions capture some key features predicted by kinematic theories: reduced diffusivities at the Sub-Antarctic Front jet and increased diffusivities at mid-depths, both in line with mean flow suppression theory." It should be clear in each sub-/section how that section contributes to this main point.
We have edited the text, particularly in Section 5, to make it more concise.

5. All of the figures are presented as lat vs depth. I wonder if you should not convert to depth and just use lat vs density. The isopycnals are tilted and so averaging in density coordinates makes more sense, as you have done except not in the figures. What do the mean profiles (averaged in the y direction) look like in density coordinates? It may aid your interpretation and more easily accommodate error bars. This is a suggestion not a requirement.
We feel that depth is a more intuitive vertical coordinate than density, and moreover we need the data as a function of depth to compare the diffusivities to those from other datasets (Figure 9). However, we have included a figure showing the mooring-derived property distributions as a function of density in the Appendix, and mentioned this at the end of Section 4.2 (line 246).

6. Typical values of diffusivity at scales of 100 km are of order 100 m^2/s from Okubo (1971). With results 10-100x that in Fig 9 from your estimates and those by others, it's worth some explanation.
The Okubo paper is a quite old study, and a lot has been learned about spatial variations of eddy diffusivity since then. For example, see the discussion in Zika et al (2020) (https://os.copernicus.org/articles/16/323/2020/os-16-323-2020.pdf), based on the work of Garrett (1983) about the scale dependence of isopycnal diffusivity: it's much smaller before filaments start interacting with one another as one reaches the mesoscale. To our knowledge, Okubo's work doesn't capture this regime change, nor does it know about suppression of diffusivity by the presence of mean flows or topography.

Estimates of eddy diffusivity in the Southern Ocean reported in more recent studies span the range of 10^2-10^3 m^2/s (with even higher values at lower latitudes). See the sources cited in the Introduction (lines 38-45).

Our method sometimes yields values as high as 10^4 m^2/s, which is unrealistically high; this matter is addressed in Section 6.2 and possible reasons are discussed. Such high values also show up in the original data of NG11 (see also major point 1).

Comments by line number:

36 - how big is the deformation radius?
We mentioned that the deformation radius becomes of order 10 km or smaller, and included an extra reference (lines 36-37).

40 - how big is the diffusivity? Later in the manuscript values are much bigger than expected for open ocean from Okubo (1971). Please explain why the mixing here is much bigger.
We mentioned that the studies cited here give diffusivities in the range of 10^2-10^3 m^2/s (line 45). See also reply to major comment 6.

65 - what is the value of gamma? how do you obtain it? It's mentioned later in the manuscript but could go here.
We kept this at the end of Section 5.3, but added an explanation of how Klocker & Abernathey (2014) obtained this value there (lines 311-313).

76 - do these assumptions hold? "the gradient does not vary significantly over the mixing length (i.e. there is a scale separation between the eddy and mean flow scales) and that the temperature fluctuations are generated locally, not transported in from regions upstream." Also the background gradient should be uniform over the mixing length.
See reply to major comment 1

110 - while applicable to the ACC and its equivalent barotropic current, maybe it is worth pointing out that this will not work in other places. Or that another variable can be used to make the cross-stream coordinate, which you mention later
This is discussed in the last paragraph of Section 6.1 (lines 384-391).

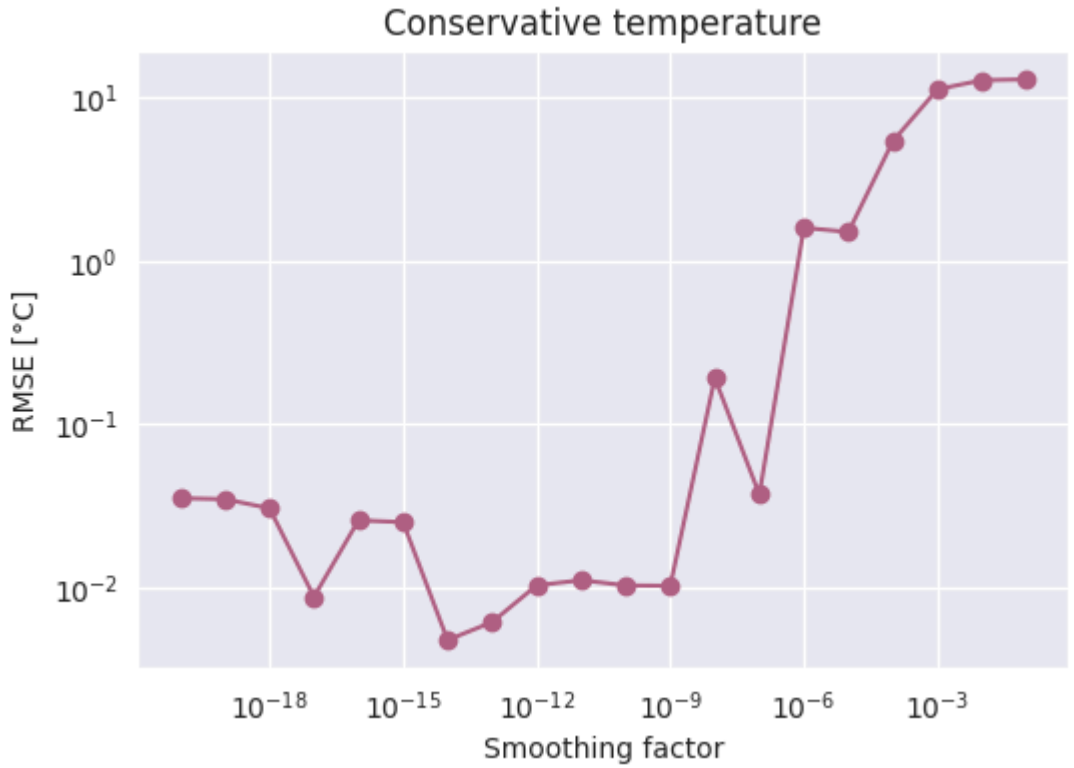
121 - some numbers here would help here as would some idea of how how variable this adjustment is
We mentioned that the suppression can alter the unsuppressed diffusivity by up to a factor 100 (Groeskamp et al., 2020) (line 126).

Fig 1 - we just need the location of C and so could omit the inset
We prefer keeping the inset because we also explain in the text that the mooring is part of an array, and motivate why we only use the middle one. Since the inset is not obscuring any important features in the figure, we think it is not a problem to keep it.

168 - Cole et al (2015) is a similar method. You should compare to that dataset too. Roach et al (2018) made a comparison and found the magnitude was different but the spatial variability was similar. Roach's diffusivity data are available online too. They have spatial

coverage and could show how representative your location is of a wider area.
We did already try to compare to both datasets, but unfortunately, they were not suitable for the comparison. Cole et al (2015) does not have any data along the SR1b section, and Roach et al (2018) only has data at the surface and at 1000 m depth, but not full profiles.

212 - smoothing factor is not explained. This is an important detail since you are converting temporal variability to spatial variability.
Good point. We chose this value because it minimises the root mean squared difference between the spline fit and Theta for the SR1b data; see figure below. We have now described how we got to this value in the text (lines 221-222).



233 - can you convert these minimum gradient to maximum distances? How do these max distances compare to length/time scales in your observations? Since your results for K are big, does this suggest a larger minimum gradient?
With this cut-off value of the gradient, we ensure that the mixing length stays below 700 km, which is the width of the SR1b section. We mentioned this in the text (lines 243-244).

241- How about a figure of U_e? Fig B4a could go in the paper or at least be referenced.
Figure 8 shows U_e.

255 - I prefer the topic sentence of a paragraph like: “the four different data sets used to transform the mooring profiles to spatial coordinates using the fit..” instead of “Fig 5 shows something” as you have it now. This is a matter of style and so up to you, Similarly line 284 and likely elsewhere in the manuscript.
Thank you for this comment; we have changed the style in various places throughout Section 5.

Fig 6b, c - These seem very similar and so one could be moved to the appendix or even supporting material.
The point of these figures is to make a comparison between the Theta distribution from the SR1b section data when all available SR1b profiles are taken into account vs when only the profiles with a phi value within the mooring data range are taken into account. They show that the Polar Front is missed if we only consider the range of phi values sampled by the mooring, but the rest looks very similar.

Fig 7- align color map with the limit. eg, Fig 7c has a lot of dark blue but only some of it is excluded. Hard to tell if the stripes extend over the whole region of drak blue. Maybe use dots instead or just blank out the “bad” data with the colormap
Good point, the stripes are not very clear. We now used dots instead.

273 - AAIW?
Fixed.

294 - “From the hydrographic properties presented in the previous sections, we can derive the mixing length from equation (3)” (Figure 7g–i).
There are numerous statements like ”This is shown in Figure 7”- just put the figure reference in brackets. It’s more concise.
We have changed this at various places in Section 5.

301 - Large mixing lengths in regions of low gradient. Does this result suggest a problem with the low gradient cutoff or even the method?
There is only a problem with the method to the extent that it doesn’t make sense to talk about mixing of mixed water (near-zero gradient = mixed water). In any case, our method produces large mixing lengths in this scenario, which is what one would expect. The low gradient cutoff ensures the mixing lengths stay below 700 km, the section width.

302 - delete “can” and add equation reference
Fixed.

307 - any thoughts on why the fronts don’t correspond to high U_e?
The frontal definitions pinpoint the axis of each frontal jet, whereas EKE tends to be higher off this axis, e.g., in the area between the SAF and PF. This is as observed by NG11, for example. We added this in the text in Section 5.3 (lines 309-310).

Fig 9 - The other methods in Fig 9c, d look like you could make a mean profile. They do not vary much by latitude. You could try making mean profiles for your results in Fig 9a, b. And then give some thought about error bars. Overall the Argo and SR1b results look pretty similar. Error bars can help determine if the differences are significant. Making the mean profiles in density coordinates is probably best. You could make a few mean profiles- one at the PF, one at the SAF, and one in the north or something similar. Your results are a bit patchy when plotted as lat vs depth. Making mean profiles provides more averaging which may help.
Making vertical profiles is not the kind of detail that this methodology is aimed at. Rather, the method can give estimates of orders of magnitude and cross-stream patterns. The cross-stream structure would be lost in mean profiles.

342 - “Differences in the isopycnal gradients.” Gradients depend on the scale they are evaluated on. The Argo section and the hydrographic section are constructed differently. It would be good to be explicit about what scales they measure. Daily mooring values in 1 m/s current are 86 km scales for example. The hydrographic section may have 30 profiles per 10 degrees and so has 30 km scales but then is averaged over many years. So the equivalent length scale is no doubt much bigger than 30 km but what is it? Argo has a grid resolution is 1 degree, but the data have been objectively mapped in space and time. Even if a model has finer resolution, numerical diffusion and averaging of the results affects the resolved spatial scales. There are several places whee high spatial or temporal resolution is noted. Instead just quote the actual number. The best comparison would be using the same spatial scale for all the data- models, hydrographic section, Argo, and mooring. If that is not feasible, then stating the actual scales they are measuring would be second best.
See reply to major comment 2

345 - “On the other hand, numerical inaccuracies and sensitivity of the spline fits may play a role.” Numerical inaccuracy is probably not the problem. Spline fits maybe are an explanation, but 3 of the 4 seem pretty close. Maybe plotting d(dyn ht)/dy might show where the conversion is going to be the most sensitive.
We believe this comment is about the phi-Y spline fits (Figure 2), but in the text the reference is to the Theta(Y) spline fits, which are used to compute the isopycnal gradients that this sentence refers to. This is indeed confusing, so we now stated explicitly that we are talking about the Theta(Y) spline fits.

412 - the gradient cutoff is indeed a problem. It cuts off a lot of data as noted earlier and it is ad hoc. I suspect the resulting mixing lengths are very sensitive to the gradient as you approach 0. Some further careful consideration/justification of this value is in order.
See reply to comment about line number 233

414 - it should be verified somehow that the assumptions for the mixing length are valid.
See reply to major comment 1

426 - don’t -> do not
Fixed.

430 - while the averaging period may matter, it may not. Eddy transport might be pretty steady averaged over a year. However, gradients calculated over 1 km or 100 km will be very different. Differing spatial scales may be affecting the various calculations and comparisons.
This may indeed be the case, that is why we have clarified our averaging scales for reproducibility in the future.