

## Response to Reviewer 2

Thank you for your positive and constructive comments on our paper. Below, please see our responses (in black) to the comments (in blue). We hope that you find our manuscript to be improved and suitable for publication.

The authors propose a new method to quantitatively define the Southern Ocean fronts. The new metric, I-metric, searches boundaries of water masses in the space spanned by the principal components of the Southern Ocean hydrography. The water masses are classified by the Gaussian mixture modelling. When applied to the output of the Southern Ocean State Estimate (B-SOSE), the newly defined fronts correspond well to the traditional definitions of fronts both in monthly snapshots and in 5-year average.

I find the definition of the I-metric (equation (3)) is intuitive and natural, and am convinced that the new method has an advantage over the traditional front definitions in its ability to allow "for a wider variety of transition types between regimes" (ll.198-199). The paper is well written except for a few places where exact meanings of technical terms are not obvious (see below). I think only a minor revision is needed before acceptance.

Thank you for these comments, they were very helpful.

I.95: Doesn't B-SOSE also solve the salt conservation equation?

Yes, thank you for catching this omission. Added (line 106).

Table 1. It is stated that "a state estimate ... the changes in ... mixing parameter" (l.90). Are the mixing parameters listed in Table 1 the initial values? Or are they fixed? Or are they after adjustment?

It is true that mixing parameters are varied in some ECCO products, but in B-SOSE the parameters in Table 1 are manually adjusted and held fixed. We have changed the text to better reflect this procedure.

I.120: Explain the use of the word "normalise" in more detail. Did you calculate the spatial standard deviation at each depth at each time step in the entire region south of 30S and divide temperature with the standard deviation? Or the standard deviation includes the time dimension?

The standard deviation and the mean are calculated at each point for each variable at each depth level, over the full sample year. I have added additional sentences at this point to clarify this point. We have also added a couple of plots in a new appendix (now called appendix B) to illustrate the preprocessing steps. (line 418-420)

I.180: I believe this term "posterior probability" is used in the Bayesian probability. If this is the case, it would be helpful to explain what is the prior probability here.

We have added one line to the manuscript (at line 187) to make this point. We use the term posterior probability because it is used in other papers that use GMM, and so we use this term for consistency.

II.183, 190, etc. also equation (3): Is "maximum" right word here? The word "highest" (I.185) sounds more appropriate since all Gaussian are "maximum" (I.165) in the sense of minimum mode-data misfit.

I agree that this seems to be a more clear wording. Thank you for this suggestion. All instances have been changed to "highest" from "maximum".

I.246: By "central" and "eastern" export pathways, I interpret the former corresponds to the equatorward path around the dateline and the latter along the 120W meridian. If the "eastern" means the I-metric blob off the South America coast, more detailed explanation is needed to clarify the ambiguity.

Thank you for pointing out this ambiguity. The export pathways spatially correspond to the *low* values of the I-metric; the higher values delimit the boundary of the pathways. We have changed the text to hopefully remove this ambiguity.

II.277-278: I do not understand what is meant by "correlation coefficient between  $G_x$  and  $x$  gradient." Is this the correlation between the result of  $G_x$  operator and the result of  $d/dx$  applied to the gridded PC data?

Yes, this was a comparison between applying the  $G_x$  operator and the  $d/dx$  operator, to the same PC1 grid. We have attempted to clarify this (line 296-298)

I.283: What is "\*" operator? How is it defined? Convolution in space?

Yes, this is defined as the 2d convolution in space (line 272). I have added this to the figure caption as well to make it more obvious for the reader.

Equations (B4)(B5)(B6):  $c_n$  seems undefined.  $\mu$  in the 1st line of (B6) should be in bold face.

We have added that definition to the start of the definition, and made sure that  $\mu$  is always bold (line 432, 437, 440).

Equation (B7): Is  $k$  defined?

We have added the definition to the beginning of appendix A (line 433), thank you for pointing this out.

L.450: What is w here?

“w.” was an abbreviation for “with”. Thank you for catching this. I have replaced this with “with” (line 467).

Appendix A1.

(A2) and (A5) are identical. The content is very elementary and can be found in any textbook. I do not think this appendix is necessary. I found Appendix A2 is very informative.

Thank you, it is very helpful to hear what you found to be useful background, and what is unnecessary. We deleted appendix A1, but kept a reference to the appropriate textbook, in case students need to be guided to the right resource. We have left the second subsection A2 much as it was, and are glad that you found it helpful.