The authors have replied to the comments raised during the reviews, but have not addressed thoroughly all the concerns, some of which constitute major weak points of the work. The application of the WMT framework to AABW in the region is a novel aspect of this study, is interesting and worth doing. However, I still have major concerns on some aspects of the study and its methodology, which is why I don’t recommend its acceptance and publication at this stage. I encourage the authors to do further work and resubmit at a future stage.

In particular, my main suggestion is that given that one of the main conclusions is that the reanalyses are not capturing observed variability, and that the reanalyses are not consistently or faithfully capturing the processes that drive AABW in a robust way (line 556), and that the valuable aspect of the work is the WMT transformation framework developed; why not do this analysis with a high resolution ocean model that is able to capture AABW more robustly? See for example Solodoch et al 2022 (using Kiss et al 2020 model which you mention in the conclusion). This would probably provide more insightful results because I see many issues with reanalyses throughout the manuscript, which removes relevance from your main findings listed in the abstract (volume loss in SOSE, larger interannual variability in ECCO, unphysically large variability in SODA).

Comments not addressed in the response
Below I expand further on the responses to the review that I consider did not properly address the concerns. The numbering corresponds to the comment’s number in the Author’s Response document.

3. and 9. The definition of the boundary region in this study is key to interpreting the results of the analysis (accounting for part of the volume changes according to equation 7 in the manuscript), and not enough work has been done to corroborate that the results are robust to the choice of boundary. The definition based on latitude/longitude lines seems arbitrary, based on a study that focuses on the Weddell Polynya in the vicinity of Maud Rise, rather than on other studies more relevant to the topic at hand (such as Solodoch et al 2022, that has with a description of the AABW pathways of export from the region). And the interpretation of export across these boundaries throughout the paper is also misleading. For example:
  - Line 9: in the abstract it says “we diagnose a closed form of the water mass budget for AABW that explicitly accounts for transport across the WG boundary”. The lat/lon lines chosen actually cut through the Weddell Gyre (see for example Neme et al 2021), and therefore the export does not represent export through the gyre’s boundaries. It might even be capturing recirculations within the gyre itself.
  - Line 342: “We note that though Kerr et al. (2012) obtained this value from a transect at the tip of the Antarctic Peninsula, they do attempt to capture the dominant outflow of the Weddell sea AABW.” Again, this is not a fair comparison: you are encompassing other circulation pathways and recirculations within the gyre in your export calculation
because of your boundary definition, whereas Kerr et al (2012) looked at transport across a specific hydrographic transect. You have responded to comment 7 from the review document “We think it is important to attempt to compare with observations somehow.” If comparison to observations is what you were after, given that the reanalyses are gridded products, you could have reproduced the hydrographic transect, which would have made for a better comparison against observations.

The above dot points are examples of how the selection of boundaries influence the results and conclusions of this paper. I don’t consider there is appropriate justification behind this selection.

5. This comment regarding Figures 1 and 2 has not been addressed. I understand that not all the reanalyses are available during the same period, but as your Figure 10 shows, there is significant interannual variability that is going to influence this comparison. In other words, you are not comparing apples to apples here. You have added to the text “We still compare each model to the available observation, but recognize that this introduces unknown biases in our comparison.”, but it is important to highlight that these needn’t be unknown biases. It should be easy and straightforward to do this comparison with the same time period to observe the biases, and even if afterwards you decide to keep your Figures 1 and 2 the same, at least then you have proper justification for the reviewer and the reader, as well as understanding about how your decision is impacting your results.

10. I still think that SOSE is showing that most of the transformations of your definition of AABW is accounted by surface heat fluxes, whereas in ECCO it is mostly due to surface salt fluxes (which you call brine rejection, see following comment), which are qualitatively different. In fact, ECCO shows almost no transformation due to surface cooling.

11. There is still no justification to the transformation attributed to brine rejection. The orange dashed line in your Figure 6 is showing surface salinity flux, which does not necessarily represent brine rejection. In line 346 you say that your analysis shows a negligible role from the atmosphere, but you don’t show the analysis or describe it? In your response to the reviews it says you were going to do a decomposition a la Abernathey et al 2016, but I can’t see any reference to that in the revised version.

Other comments

1. The Introduction is still not well organised: you start by (i) a description of the MOC, follow with (ii) a very brief paragraph on the Weddell Gyre, (iii) then the relevance of AABW for the global climate, (iv) then what obs. show regarding AABW export, (v) then WMT framework, and end with (vi) your research question. A clearer view of your research area, in my opinion, would be to start with (iii) the relevance of AABW, part of which is its connection to (i) the MOC. You could then follow with a description of the characteristics and processes linked to AABW production/circulation/export in your study region, which would encompass its production mediated by the (ii) Weddell Gyre, as well as (iv) the observational studies. This would provide an adequate frame and highlight that we don’t know the thermodynamic mechanisms that link variability in surface forcing to AABW export, and
describe (v) the WMT framework, which is your methodology to study your (vi) research questions (thermodynamic mechanisms).

2. This is a major comment stemming from additions to the manuscript. You have used Jackett and McDougall (1995) equation of state to define sigma2. Why have you not used the most recent equation of state (TEOS10)? There is even a python library called gws that calculates sigma2 easily.

3. You have made some changes to the Figures that have not improved them: you have used the Blue to Reds colorbar for anomalies, which is standard procedure and its perfect; but then for the standard deviation you switch from a sequential colorbar in Fig. 1 to the Blue-Red in Fig. 2 which at a glance could be misleading for the reader. And I would have also suggested to use a different colorbar for bottom temperatures, because again at a glance they could read as anomalies.

4. You use standard deviation in Figures 1 and 2 as a proxy of the (unknown) reanalyses uncertainties (line 271). This is not correct since standard deviation is a measure of variability.

5. Line 330: you say that because of volume conservation, the transport of bottom water is equal and opposite to the transport of deep water. Is this not affected by surface mass fluxes?

6. Line 374: “excess dense water is created by WMT in winter and then destroyed in summer”. This is not the case for ECCO, where there is production throughout the year (green line panel a.)

7. Line 439: the residual is your ECCO time series in Figure 10a is large, with the same order of magnitude of your time series of transport, WMT and change in volume. You say you find no explanation in this, are you sure your calculation is not flawed? This is worrying because for your interannual time series you conclude that SODA’s is flawed, SOSE is too short and therefore ECCO is the most useful, even though it presents these large errors.

8. There are still minor typos (which I won’t go into detail because it is not the focus of this review) and not-so-minor errors throughout the text that makes me think it was not proofread. These include, but are not limited to: there is a random sentence in line 247 that seems should not be there, acronyms not defined in the text (NADW line 57), one sentence paragraphs (line 121 and 170), equations that are referenced before they are defined (line 111)

References
