Review: "Assessment of the structure and variability of Weddell Sea water masses in distinct ocean reanalysis products" by Tiago S. Dotto, Rodrigo Kerr, Mauricio M. Mata, Marina Azaneu and Ilana Wainer

In this article the authors assess the accuracy of five global ocean reanalysis products (ECMWF, CFSR, MyOcean, ECCO2 and SODA) in representing the water mass structure of the Weddell Sea over the past few decades. The focus on the Weddell Sea is warranted by its importance for production of Antarctic Bottom Water, which ventilates the deep ocean and may substantially impact climate over timescales of hundreds to thousands of years. The Weddell Sea is a particularly challenging region for ocean reanalyses due to its small dynamical length scales and the presence of dense overflows from the Antarctic continental shelf.

The evaluation consists of a direct comparison between observed and reanalysis temperature, salinity and neutral density along the WOCE SR4 and A12 sections. All of the reanalysis products examined capture the qualitative structure of the water masses in both sections. The largest discrepancies are typically found in the Antarctic Surface Water (AASW), where the models struggle to represent the interaction between the ocean, sea ice and the atmosphere. Large discrepancies are also found in the boundary current at the Antarctic shelf break, and in the deepest Weddell Sea Bottom Water (WSBW) layer, which, for example, is largely absent in the ECCO2 model. The reanalyses exhibit multi-decadal trends in the properties of AASW, Warm Deep Water (WDW), Weddell Sea Seep Water (WSDW) and WSBW, though the various products disagree on the signs of these trends, and the magnitudes of the trends are often much smaller than the discrepancy between the observations and the reanalyses.

The comparison between the reanalysis products and observations is thorough, aided by eleven multi-panel figures comparing the mean water properties and their variability, and presenting statistical measures of the error. In general it seems that all of the reanalysis products require considerable development to improve their representation of the Weddell Sea water masses and their trends, particularly in their representation of dense water production and overflows. I am therefore not sure that I agree with the authors' conclusion that "ocean reanalysis products are a valuable option for studying the climatological states of the deep layers of the Weddell Sea" (L600–601), but this may be a matter of opinion. Either way, this article allows interested physical oceanographers and climate scientists to judge the relative strengths and weaknesses of the reanalysis products at a glance, and is therefore of practical use to the scientific community.

Comments/questions

- L129–189 and Table 1: Looking over the observational datasets, it appears that the WOCE sections used as the basis of the authors' evaluation is itself assimilated into all of the reanalysis

products except CFSR. The WOCE dataset cannot therefore be regarded an independent test of the reanalysis products, and it is unsurprising that CFSR deviates from the WOCE observations much more than any of the other products. I don't think that this invalidates the authors' approach: it is still helpful to evaluate the agreement between the reanalyses and observations. In the absence of any independent observations, it is sensible to base such an evaluation on the WOCE data because they are representative of the Weddell Sea water mass structure and have been obtained consistently over a long period. However, the authors should include some discussion of these points prominently in the manuscript.

- 178–181: I would like to see a more detailed discussion surrounding the authors' decision to exclude the later ECCO2 data. These years are presumably some of the most important, being the most recent, and if the purpose of the evaluation is to allow other scientists to judge the quality of these reanalysis, then why obfuscate anomalous behaviour in the models?

- L260–262 and L269–271: I find it difficult to see the patterns of discrepancy described for the AASW. It is easier to see distinguish the patterns of discrepancy described later for the WDW, WSDW and WSBW with the aid of Figures 7–12 and Table 4. Perhaps the authors could include the mean error in θ, S and γ^{n} in AASW layer in the text here to reassure the reader.

- L310–311: Dense water overflows are critical to the structure of the Weddell Sea, and to the global overturning circulation. How did the authors determine that the overflow of WSBW was absent in the reanalyses? Do none of these reanalysis products include some kind of dense overflow parametrization?

- L323–324: Reanalyses are judged to be accurate if their RMSE in a given property is smaller than the average, including CFSR. Given that CFSR is the only reanalysis product not to assimilate the WOCE sections, would it be fairer to judge the accuracy based on the mean RMSE *excluding* CFSR?

- L326–328: Could the authors elaborate on how the stronger currents in WOCE A12 might enhance turbulent processes and salt diffusivity? Presumably the resolution of the reanalysis products is too low to see any mesoscale turbulence, so I infer that we are talking about small-scale turbulence. Do the authors mean that large vertical shear may be generated, activating a vertical mixing parametrization that depends on the gradient Richardson number? If so, is the vertical shear really large enough at these depths for this to be a plausible explanation?

- L394–397: The authors mention that formation of a large open-ocean polynya in ECCO2 leads to anomalous behaviour beyond 2004. Is this also the cause of the anomalous behaviour in ECMWF and MyOcean?

- L37–38 and L473–475: I'm not convinced that this statement is justified. In no case is the horizontal resolution "merely" increased: simulations with different resolutions are compared between entirely different reanalysis products, and judged based on hydrographic sections from a very small part of the ocean. If horizontal resolution alone (as implied by "merely") were increased, then in fact I think there almost certainly would be an improvement in the quality of all of these reanalyses.

- L479–480: I find it questionable to say that parametrizations are as fundamental as horizontal resolution. In principle, in the limit of arbitrarily fine horizontal (and vertical) resolution, the ocean fluid dynamics would reduce to the Navier-Stokes equations, and no parametrizations would be required. So grid resolution is certainly more fundamental. I think that perhaps what the authors mean is that improvements in parametrizations may have as much impact on the reanalyses as modest increases in horizontal resolution.

Minor comments/typos

- The standard of the English prose in this manuscript is generally very good, but here and there I found sentences to read a little awkwardly. I won't provide an exhaustive list, but I suggest that the authors give the manuscript a careful proof-read.
- L39–41: I understood what was meant by this sentence, but I think it should be rephrased for clarity, particularly in the abstract.
- L106–108: I think it would be appropriate to add citations in support of this statement.
- L204–205: I didn't understand this sentence. Please rephrase.
- L209–211: I suggest a minor rephrasing to emphasize that these numbers refer to the grid spacings for the interpolation.
- L284–285: I would understand "compensate" to mean that the changes in density due to freshening and cooling exactly cancel one another, whereas I infer the authors mean that the fresh bias and cold bias have opposing effects on the density.