This is an interesting study that proves, once again, that reanalyses/state estimates cannot be used instead of observations. I particularly like the tone, including the salt puns. I have two issues that may not significantly impact the manuscript, but which I expect to take time to verify, so I recommend major revisions.

My first issue is your definition of polynyas. [Disclaimer: The following contains references to papers from my team. I am not mentioning them to pressure you into citing me. These are the studies I know of that illustrate the point I am about to make.] On several occasions in the text, you discuss how some signals look like polynya signatures. I would add that the surface salinity flux of SOSE on Fig 9 suggests this further. The issue I have is that you rule out polynyas based on sea ice concentration. Globally, and especially so over Maud Rise, there can be polynyas with a near-100% sea ice concentration... but with a very thin ice. That's why many polynya detection algorithms use a thickness threshold instead of a concentration one. A threshold of 12 cm is standard; see e.g. the latest Nakata et al. (doi: 10.1029/2020GL091353) for coastal polynyas, and Mohrmann et al. 2021 (doi:10.5194/tc-15-4281-2021) for all types of Antarctic polynyas, in models and in observations. That is particularly crucial because there was some halo / small polynya activity in the region in 2004 and 2005 (table B1 of Heuzé et al. 2021, doi: 10.5194/tc-15-3401-2021). I would therefore like to see you redo your “polynya analysis” with the sea ice thickness instead of the concentration, and if the thickness falls below the threshold or simply decreases, rewrite your discussions accordingly. An extra (supplementary?) figure to compare the sea ice in all three models would be most welcome. If they assimilate sea ice instead, do describe the data source, frequency, whether they even assimilate thickness or only concentration, etc.
My second issue is from line 365 onwards: I do not understand why you are conducting the analysis on sigma cross, the boundary between CDW and AABW, rather than on AABW itself. Especially since you say line 366 that you are using this level to study AABW. I would like to see a clearer explanation of why this particular level can be representative of AABW, and not, as I first expected, be where the signal is most dampened. I would also like to see that this choice is robust, either by providing a supplementary version of Figs 7, 8, 10 and 11 created at a denser level, or by adding “denser” lines to these figures.

Now for specific comments, in order of appearance:

- The introduction up till line 45 is quite repetitive. I would merge these paragraphs and keep only the key points, notably the very last line (first time you do not mention only air-sea interactions but also cryosphere-sea interactions).
- The paragraph line 53-59 is out of place. That should be among the very first things to write about, this overall “why should anyone who is not an oceanographer care?”. By this point, I would rather you explain why the seasonal to interannual variabilities are important to study, which processes they impact, etc.
- Line 89, for context, provide the depth range that you are looking at in this study (should also be mentioned in the introduction)
- Line 91-93: Discuss whether these limitations are problematic to study the Weddell Gyre region. Again, no pressure for citations, but Mohrmann et al. 2022 (doi:10.1029/2022GL098036) suspects that some mixing signals we see in observations in the region are the result of cabbeling and thermobaricity.
- Line 103: typo I suspect, vertical is diapycnal? Also, since you called this term \( G_h \), I would write “horizontal” instead of “lateral” in the description, to help the readers.
- Having the reanalyses introduced earlier would make section 2.1 easier to picture, to know which variables are available at which resolution, esp. vertical. Consider whether to swap sections 2 and 3.
- Line 136 onwards: your cumulative sums, bottom up or surface down?
- Equation 14: double minus = typo?
- Section 3: for all products, specify whether \( \sigma_2 \) is provided or whether you had to compute it, and if so, how.
- Line 194, ECCO: Vertical grid type? Resolution (in m/mbar)?
- Throughout the manuscript, for example lines 205 and 207, check your citation styles (citet vs citep, if using LaTeX)
- Line 2012, SOSE: Number of vertical levels? Type of vertical grid? Vertical resolution? Also, please write more clearly what daily values of 5-day averages mean: each day is the 5-day mean centred on that day?
- Line 226, SODA: Vertical grid type? Vertical resolution?
- Line 232: how many ensemble members?
- Line 245-246: I do not understand this description; it is nudged? If so, how often? Towards which variable(s)?
- Figure 1 / line 258 onwards: is the bottom at the same depth for all products? In particular, is the shelf-break shifted N-S?
- Lines 262-263: I do not understand what the standard deviation and spatial variability mean. Throughout the manuscript, use consistent terms such as “temporal standard deviation”, “horizontal standard deviation”, etc
- Figure 1: fonts are too small on the colorbars
- Figure 2: fonts are too small there as well, and adjust the caxis ranges on most panels, but especially the first column, so that we see more details
- Line 288: Help the reader by having this range highlighted with a box on Fig3.
- Line 291: you eventually give this information several pages later, but give the T-S range of CDW now or on line 288, and not just that of AABW.
- Figure 4 deserves way longer a description than this short line 294. Comment on the overestimation (?) in CDW, or on how the Gade line really sticks out, which suggests some sea ice / ice shelf interaction misrepresentation.
- Figure 3: something is off with SODA at T=0 deg C. Explain why it has a larger volume than expected, or correct if that is artificial.
- Figure 4: The caption and titles are reversed, so write more clearly “model minus WOA” (or the opposite)
- Line 318: you eventually give the densities in the different models many pages from here. You should give them here instead. Because you just finished showing how biased they are in T-S, yet here it looks like you do not account for their biases in density.
- Line 325: what does a volume gain mean in practice? Takes over other water masses higher up in the water column? Wider branch?
- Line 331: Fig 6 does not show brine rejection, only surface salinity fluxes. Could be P-E. This joins my major comment: your study of sea ice needs to be more extensive, and to be shown.
- Line 336-337: are these small values a typo? If not, which density interval am I supposed to look at right now? Because I did not notice this order of magnitude difference from ECCO.
- Line 347: I guess you do not use LaTeX after all. Please do not comment on Fig 9 before Figs 7 and 8, it is quite uncomfortable to have to go back and forth between the pages.
- Line 358: see major comment, surface cooling + freshening suggests sea ice melt (from below) to me.
- Line 420-421: have you tried contacting them? They may not have checked AABW, but maybe they’ve investigated NADW for the AMOC and that would give you clues.
- Figure 10: to help with the comparison of the variabilities, have the same width as time interval for the three panels (e.g. 1 cm per year). Ideally, align them even, so that we can directly compare the models to each other.
- Figure 11: same comment as Fig 10
- Line 541: there’s an Heuzé 2021 on CMIP6 (doi: 5194/os-17-59-2021). More models, more recent, same conclusions. As previously: no expectation of citation, just for your information.