Review of "Quantifying the feedback between Antarctic meltwater release and subsurface Southern Ocean warming" by Lambert et al.

In this manuscript, Lambert and co-authors analyze the potential impact of meltwater-induced ocean warming feedbacks on Antarctic ice sheet melting and the resulting sea-level contributions. I really enjoyed reading this paper and was impressed that they have managed to distill a very complicated analysis down to a relatively easy-to-understand story. They use a novel method involving Linear Response Functions derived from ocean temperature anomalies in perturbation experiments with EC-Earth3. By applying these functions across a suite of ESMs, they evaluate how different ocean warming scenarios, both with and without the feedback effect, may influence basal melt rates (and hence sea level rise). This approach provides valuable insights into the interactions between oceanic and cryospheric processes, enhancing our understanding of how these systems influence each other. The manuscript is well-structured, clearly written, and supported by high-quality figures. The analysis is comprehensive, the results are presented clearly, and the conclusions appear robust. The topic is of significant importance, as it addresses critical uncertainties in projecting future Antarctic ice-ocean interactions. It is, therefore, fitting for the audience Earth System Dynamics and will make a valuable contribution to the literature. Additionally, the findings are timely and relevant as the scientific community prepares for ISMIP7 and CMIP7. I have a few comments and concerns that I would like the authors to address prior to publication. My main concern is the chosen horizontal extent for ocean temperature extraction and the need for more discussion on the uncertainty of basing the ORFs on a single ESM.

General comments

1. **ORFs based on a single ESM (EC-Earth3)**
   As the authors note in the discussion, a significant uncertainty and limitation of this study is that the ORFs are derived solely from the EC-Earth3 model's response to meltwater input. The main assumption here is that other ESMs would exhibit a similar response. I understand that this is the only way to do this study without running experiments with many models (something covered by the SOFIA initiative), and I completely agree with your approach. I do not suggest changing this. However, I would appreciate a more extensive discussion on this topic. Specifically, it is likely that different models exhibit varying magnitudes of subsurface warming, as highlighted by Swart et al. (2023) and the recent work by Chen et al. (2023, https://doi.org/10.1029/2023GL106492). Furthermore, the assertion that most studies consistently show subsurface warming may be somewhat overstated. For instance, while you reference Thomas et al. (2023) and their findings of regional cooling, you could also include Beadling et al. (2022, https://doi.org/10.1029/2021JC017608), who demonstrated cooling around Western Antarctica in response to meltwater input in the GFDL-CM4 model. These studies suggest that a universal warming response is not guaranteed. If the ORFs were constructed using models like HadGEM3 or GFDL-CM4, the outcomes would likely be different. For example, Beadling et al. (2022) propose a mechanism in high-resolution GFDL-CM4 where additional meltwater isolates the West Antarctic shelf from offshore warm waters, a process that may not be well-represented in EC-Earth3. I believe the uncertainty stemming from this limitation is under-communicated, particularly in the results and conclusion sections. I recommend emphasizing earlier in the manuscript (perhaps also in the abstract) that the results are contingent on the assumption that all ESMs respond similarly to EC-Earth3, and that this

assumption could significantly affect the findings. I also suggest adding "in EC-Earth3" to the following sentence in the abstract: "Increased meltwater release from five individual Antarctic ice sheet regions is found to unambiguously warm the subsurface Southern Ocean at centennial time scales"

2. **Choice of horizontal and vertical boundaries for ocean temperature extraction**
   I am concerned about the uncertainties associated with the large horizontal extent and limited vertical extent when extracting temperature data from the 3D fields. Specifically, I question the inclusion of water masses located far offshore and hence spatial averaging over an area where know there are large gradients in T/S. Additionally, the quadratic melt rate parameterization (Eq.1) is for near the ice front and not far field (e.g. see discussion in Hattermann et al. (2010, https://link.springer.com/article/10.1007/s00382-009-0643-3). Would it not be more appropriate to restrict the analysis to water masses that are directly on the shelf, particularly those south of the 1000m isobath, as suggested by Barthel et al. (2020, https://doi.org/10.5194/tc-14-855-2020)? Including offshore anomalies may risk incorporating temperature changes that do not influence the shelf directly, which is supported by findings from Beadling et al. (2022). Have you conducted a sensitivity analysis to determine whether the inclusion of these offshore waters significantly affects the results? Additionally, the narrow vertical extent makes the findings particularly susceptible to vertical biases within the models, as indicated by the results presented in Figure 4. The depth of the temperature anomalies varies from model to model, which could introduce further uncertainty. Have you explored how different the results might be if a wider depth range on the shelf were included? These depth estimates are realistic and based on reality, but in the model world, they might not capture the right water masses. I acknowledge that redoing the experiments and analysis is too much, and I do not request this, also I am well aware that this is nicely consistent with Levermann et al. (2020), but especially the choice of horizontal regions is suboptimal and inconsistent with similar studies.

3. **The peninsula ocean region includes both east and west**
   Given the distinct dynamics and differing water masses between the western and eastern parts of the Antarctic Peninsula, I question whether a weighted average between the two regions is meaningful or realistic from an ocean perspective. I understand that the IRF is just one function for the peninsula which makes sense from an ice perspective, but the oceanic feedback is likely very different on either side. I suggest separating these and either include both the western and eastern parts of the peninsula separately or just the eastern part. This could be done by using the same IRF, but different ORFs.

4. **Quadratic relation between basal melt rates and thermal forcing**
   In this paper, you chose a relatively simple way of converting thermal forcing to basal melt rates, which I appreciate given the already large uncertainty. However, you mention that "basal melt representation can be improved significantly by taking into account a more appropriate extrapolation into ice shelf cavities (Jourdain et al., 2020) and a more sophisticated calculation of basal melting (e.g., Lambert et al., 2023)". Can you explain why you did not use either of these two methods? Also, could you speculate on how these different methods might affect your

results? This sentence warrants some more discussion. Additionally, I am left with uncertainty about whether I should trust the calibrated or fixed basal melt parameter more. Using the fixed parameter gives a large difference between feedback and no feedback, whereas the calibrated parameter gives only a tiny (5%) difference. Could you be more clear on what the take-home message should be for the reader? Is it correct to assume that the calibrated parameter is more realistic and that I, therefore, can conclude that the inclusion of the oceanic feedback does not matter that much for the future projections? Or have I misunderstood this?
Additionally, I am curious to see the historical period (Figure 8) also for fixed non-callibrated gamma. This comparison is interesting and warrants some more discussion.

Specific comments

Line 4	Here, you have the opportunity to explain some more context for non-specialists. I miss the "why" in the abstract. Why is this important? Why should we care about this feedback that you are aiming to quantify? By highlighting that this feedback is currently ignored in most simulations, you underscore the novelty and necessity of this work. It is clear in the introduction but would be useful to highlight this more in the abstract as well.
Line 27	You could add here that it decreases deep convection
Line 41	"on average" a positive one
Line 78	How representative is EC-Earth3 compared to other ESMs in simulating the mean state of the Southern Ocean? While Figure 3 provides a basic comparison with temperature data from reanalysis, it would be beneficial to include more references or a detailed analysis that evaluates how well EC-Earth3 reproduces key aspects of Southern Ocean hydrography. How do its biases compare with those of other models? Especially important for this study is how well it captures the vertical structure and the Antarctic slope current in comparison to observations. Including an additional figure and some discussion to demonstrate that EC-Earth3 is not an outlier among ESMs would strengthen the case for basing the ORFs on this model.
Line 121	"from either one of the five source regions." This was a little unclear to me at first. You run 5 experiments, where you just change one region in each experiment and do nothing with the other regions. Can you make this more clear and also explain earlier on why you do not just change all regions in one single experiment?
Line 125	(and throughout the paper): For easier comparison with other studies, please consider including a comparison of the Sverdrup equivalent to "Gt/yr" in the methods section. (at least when describing the pertubution magnitudes (see Swart et al., 2023).
Line 141	Can you include a table of the ESMs in the Appendix?
Line 151	Why constant and why -1.7? At the depths you have chosen the freezing temp should be much lower. Not? Could you refine Tf for each region by calculating the freezing temperature based on the mean salinity from the EC-Earth piControl at the depths

where you compute the melt rate? This adjustment would likely make your ORF estimates more realistic.

Line 142 “For each ESM we use the piControl simulation to bias correct long-term ocean temperature trends”. Please expand. This step is not clear. Why?

Line 252 The reanalysis data referenced here and utilized in Figure 3 should be described in greater detail within the methods section. It would be beneficial to clarify how these reanalysis products are combined and whether they have undergone any evaluation to assess their reliability. Why have these reanalysis products been chosen over observational climatologies (e.g. Jourdain et al., 2020)?

Line 276 This is a very interesting result and can be highlighted. Please expand a bit more on this.

Line 280 “coloured lines in Fig. 5.” Please specify we have to look at the rightmost column.

Line 315 This is a nice example, and illustrated the methodology very well. It was absolutely needed fo me to understand it, so thanks for this. However, why did you choose CISM-NCA? This is unclear. Can you explain this choice or is it random?

Line 332 I understand the limitation with the fixed basal melt parameter. However, Figure 7 is still a key result, and I think it deserves some more text. For example, the differences between scenarios depicted in this figure warrant a more detailed description. If Fig. 7 is to be included, it deserves a bit more than two sentences in the result section.

Line 385 I am not entirely convinced by the assertion that the general quantification of the meltwater–warming feedback is robust. While it may hold for EC-Earth3, your own discussion suggests that if the ORFs were based on a different model, the feedback could vary significantly. Therefore, it is unclear how this conclusion can be generalized. I acknowledge that this estimate is likely the best possible given the presented analysis, but I would suggest modifying this statement to reflect the associated uncertainty more accurately.

Line 395 The last paragraph of the discussion is very good and very important. I suggest moving this to the conclusion instead.

Line 402 The improvement/difference in the freshwater balance from the original CMIP6 version of EC-Earth3 and the new CTRL with the modified masks and routing is a significant result. The difference from the new distribution is substantial, and I believe it warrants a sentence or two in the conclusions. It is important for other modeling centers as you have shown that the way freshwater is added can have almost equally much to say as to how much freshwater is added. This result should be highlighted more. (Perhaps also in the abstract if you have space).

Figure 3 Caption: state clearly in the first sentence that is for EC-Earth3 (helps the reader). e.g. “Control time series of subsurface ocean temperatures from EC-Earth3.” Similar for Figures 4 and 5.

Figure 3 Caption: repeat here that the CTRL has the adjusted/improved freshwater balance.

Figure 4	It would greatly enhance the manuscript to include an additional figure, similar to the lower row in Figure 4 (or add rows to this figure) displaying the temperature anomalies to the 200 and 400 GT/yr perturbation experiment in EC-Earth. This will also provide valuable insight or confirm whether the choice of vertical/horizontal extent of temperature anomalies is good for EC-Earth3.
Figure 4	I recommend adjusting the x-axis scales for each region to optimize data visibility. It is more important to ensure that the data is clearly represented than to maintain a uniform latitudinal extent across all plots. For example, the data for the Ross region is currently difficult to discern, and refining the scale could make the patterns more apparent and interpretable.
Figure 5	Caption: please add "Note different y-axis scales".
Figure 6	Panels are very small and hard to read. Consider whether a different y-axis may be appropriate, especially for the second and third rows. (could, for example, have them all similar with the exception of the last column) Also, make sure that this Figure fills the full width of the page. Currently, the difference between feedback and no feedback is very hard to see.
Figure 8	Just EC-Earth? And add that this is for all regions combined?
Figure 9	These are the same for all regions, right? Perhaps repeat that in the caption.
Figure 10	It would be nice to include the median from Figure 7, the fixed basal melt parameter, on this figure for comparison.
Figure A1	This figure is useful for the community, but I would suggest a different y-axis for each region. The main point of this figure (that the response is the same for all experiments) still remains clear. The same is true for Figure B1.