Author response to Referee #1

This manuscript provided a spatially informed comparison of the mechanisms which dictate differences in the relationship between rising air temperature and meltwater production over the ice shelves of Antarctica. The results represent a novel and potentially important contribution to the literature pending suitable revisions. I have included some general feedback followed by a list of specific recommendations below.

We would like to thank the referee for their comments and we address them here. In black are given the comments, in blue our response and in orange the changes we would implement in the manuscript.

General comments

The authors briefly state that RACMO may struggle to accurately represent longwave radiation and turbulent heat fluxes in the opening paragraph of section 3.1. This deserves more discussion. How does RACMO struggle to represent these variables? Are the biases in these variables uniform across antarctica or regionally dependent? How does RACMO accurately resolve surface air temperature and SMB if these important surface energy balance components are not resolved accurately in the model? What implications, if any, do these shortcomings have for the results of the study?

We plan to given more detail on the biases in section 3, including the locations of the weather stations where these biases are based on. We also will briefly discuss how the biases in SEB terms translate into melt estimates as follows (note that this text will be moved to the methods section 2.3 as suggested by Referee #2):

"RACMO(ERA5) has been extensively evaluated against weather station and mass balance observations in van Dalum et al. (2025). Using observations from automatic weather stations on the Antarctic Peninsula and in Dronning Maud Land, this study shows that version 2.4p1 performs well in simulating Antarctica's near-surface air temperature (bias of -1.40 and an RMSE of 4.38 ◦C) and shortwave radiation (bias of 8.5 and -8.8 W m−2 for downward and upward shortwave radiation, respectively), but has larger differences with observations for longwave radiative fluxes (bias of -20.4 and 11.7 W m−2 for downward and upward longwave radiation, respectively). Turbulent fluxes have small bias (-0.3 and 1.5 W m−2 for latent and sensible heat flux, respectively), but large spread (RMSE of 5.0 and 14.2 W m−2, respectively). As the biases in longwave and shortwave radiation partially offset one another, the resulting melt rates are less affected. This is reflected in the good agreement between simulated meltwater presence in the snow and satellite-based estimates (van Dalum et al., 2025)."

One of the strengths of this study is the detailed consideration in climate-driven spatial differences in SMB response between the various ice shelves. Some of the results presented in section 3.1 worked counter to this strength. For example, spatial averages in Table 1 could obscure biases that are locally relevant to a specific ice shelf. Furthermore, I did not see how this examination of biases was considered in the interpretation of rest of the results. Do these biases have any implications for the

conclusions of the study? Here again, spatial averaging makes it difficult to answer this question.

Thank you for this comment, we agree that the discussion of biases in the ESM-forced simulations would benefit from more discussion on regional differences. Therefore we have now added an extra figure with difference plots, as in Fig. 1, but for the main SEB terms (Swnet, Lwnet, LH, SH), where we can discuss in more detail the regional differences between the ESM-forced and ERA5-forced simulations. The results of Table 1 are now less central in the text and that is why it is moved to the appendix.

Note that now also the difference maps in the main text display the differences for DJF, to have more consistent focus on summer air temperatures and SEB. We have replaced the sub-figure for SMB with melt in Fig. 1 to be able to discuss the regional differences in the biases in melt.

That the differences between ESM- and ERA5-forced simulations did not impact our results is added to the discussion: "We find that the temperature sensitivity of the surface energy balance components are very similar across all simulations, indicating that the ESM-forced simulations reliably reproduce the relevant physical processes and can therefore be used to extend the temperature and melt range beyond that from the historical RACMO(ERA5) simulation."

I often found it difficult to see from where the authors were basing their claims. I believe the manuscript would benefit from more detailed explanations of how the figures support their claims. This is particularly true for the discussion of figures 4-6.

In response, we have clarified the description of Figure 4 and moved Appendix Figure C1 to the main text, as it provides more quantitative evidence for the spatial variability in the temperature-melt relationship linked to snowfall. In addition, we have improved the interpretability of Figures 5 and 6 by changing the trend lines to black and labeling selected lines. This allows for a clearer reference to the relationships shown in these figures.

Specific comments

L13: delete the s from "becomes"

Done

L74: I would suggest rewriting as "…and therefore provide a better representation of areas such as…"

Done

L79: delete "in" from "penetrate in the snowpack"

Done

L103: It can be assumed from the description that ER>0 for erosion and ER<0 for deposition, but it would not hurt to state this explicitly.

We have included this.

L109: A citation is needed for the ERA5 reanalysis dataset

We have included this now: "We use the RACMO simulation forced by ERA5 reanalysis data (Hersbach et al., 2017),.."

L111: "Future projection simulations" is a bit redundant. I would suggest "Projections spanning 2015 to 2099 were forced using..."

Thank you, we have now changed this.

L114: A short rationale for why SSP3-7.0 was chosen could be included here.

We clarify this as follows: "SSP3-7.0 was chosen within the PolarRES framework because it is considered a more plausible high-emission pathway than SSP5-8.5."

L117-120: This information would fit better in the previous paragraph, which is where you first introduce the future period simulations.

Thank you for the suggestion, we will restructure so that the first paragraph discusses the projection runs as part of PolarRES, and the second paragraph discusses the ERA5 forced simulation as reference simulation. The first pagraph then reads as follows:

"We use RACMO2.4p1 with a domain covering Antarctica and the southern tip of South America with a horizontal resolution of 11 km, forced with both reanalysis and Earth-System-Model (ESM) datasets. The simulations for this domain were performed as part of the PolarRES project, an EU Horizon 2020 funded project that uses RCMs to simulate the current and future climate of the polar regions. Projections spanning 2015 to 2099 were forced using boundary conditions from two CMIP6 ESMs: the Community Earth System Model 2 (CESM2) and the Max-Planck Institute Earth System model (MPI-ESM), under the high emission SSP3-7.0 scenario. SSP3-7.0 was chosen within the PolarRES framework because it is considered a more plausible high-emission pathway than SSP5-8.5. CESM2 and MPI-ESM were then selected from the CMIP6 ESMs using a storyline approach to represent two contrasting but plausible Antarctic climate futures: CESM2 reflects a future with extensive sea ice loss and an earlier summertime stratospheric polar vortex breakdown, while MPI-ESM captures a scenario with limited sea ice loss and a delayed polar vortex breakdown (Williams et al., 2024).

L124: Is this calculation of sea ice temperature performed within RACMO? Also, a citation is needed for this slab model.

We will clarify that the sea ice slab model is part of the ECMWF IFS model that was cited before in the text. “Sea ice temperature is calculated using the four-layer sea ice slab model from the ECMWF IFS model, which assumes a fixed maximum thickness of 1.5 m.”

L130: Can you clarify what is meant by “average climate in the historical simulation…”? As written, I would assume that RACMO(ERA5) is the historical simulation, since it is forced by an observationally constrained dataset. From what is written in the remainder of the paragraph, it seems the comparison referenced here is between RACMO(ERA5) and RACMO forced by the ESM’s representation of historical conditions. It is also not clear what the “average climate” is here. Perhaps long-term mean would be more accurate?

To clarify this we rephrase this sentence to: “Because ESM-forced simulations are not constrained by data assimilation, we evaluate the ESM-forced simulations by comparing the mean and variability of near-surface climate variables over 1985–2014 with those from RACMO(ERA5) for the same period.”

L132: Is this statement summarizing your own attempts at validating RACMO against in situ data? If so, where is this data presented? If not, a citation is needed.

This statement is based on the paper by van Dalum 2025, which is clarified as follows “RACMO(ERA5) has been extensively evaluated against weather station and mass balance observations in van Dalum et al. (2025). Using observations from stations on the Antarctic Peninsula and in Dronning Maud Land, this study shows that version 2.4p1 performs well in simulating Antarctica’s near-surface air temperature (bias of -1.40 and an RMSE of 4.38 ◦C) and shortwave …”

L138: The word “significant” is usually reserved for instances of statistical significance. Was a statistical analysis performed here? If so, what method was used and where are these results presented?

We have not done a statistical test and therefore will remove the wording ‘significant’. Instead we compare the difference with the typical year-to-year variability through comparing it with the standard deviation. We have rephrased L138 to: “Over the Antarctic continent, RACMO(MPI-ESM) temperatures differ little from RACMO(ERA5) relative to the year-to-year variability. Exceptions are Dronning Maud Land, which is warmer, and the high Antarctic Plateau, where temperatures are lower than RACMO(ERA5) by more than the inter-annual standard deviation.”

Figure 1: It is somewhat unusual in my experience to use hatching to highlight areas of small differences relative to internal variability. This is also a bit confusing in the context of the discussion, where significant differences are emphasized (L138). This is related to my previous comment, but if a statistical significance test was conducted, I think it would make more sense to highlight areas of statistical significance.

In combination with the previous comment, we now avoid use of the word 'significant' when referring to the difference between simulations, but rather discuss in terms of larger/smaller than inter-annual standard deviation. We also have changed the hatching so it indicates regions where the difference is larger than the inter-annual standard deviation. We have clarified the use of hatching in the figure in the text by adding the following: "In the figure, hatching indicates areas where the mean difference between the ESM-forced simulations and RACMO(ERA5) is larger than the ERA5 interannual standard deviation over the historical period."

Table 1: If the focus of this paper is on the Antarctic ice shelves, how relevant is a spatial average of model biases across the whole of the Antarctic Ice Sheet? It seems that spatially informed biases are of critical importance to the question at hand, and the information in Table 1 may mask some of these locally relevant biases by averaging biases of opposing sign in different regions (e.g., the strong negative precipitation bias over east Antarctica versus the strong positive bias over Dronning Maud Land in RACMO(MPI-ESM)).

We agree that spatially averaged biases can mask locally relevant differences, especially as there are regions with opposing signs in biases. We have now moved Table 1 to the appendix and instead added spatial maps of the differences in main SEB terms (SWnet, Lwnet, LH and SH) between the ESM-forced simulations and RACMO(ERA5). The text is therefore also discussing more the regional biases instead of Antarctic-wide integrated differences.

Figure 2: Panels are referred to by letter in the figure caption, but there is no lettering on the figure.

We have now included the lettering in the figure.

L185: "consistently stronger" is a bit unclear. Perhaps something like "…21st century; however, output from RACMO(CESM2) consistently shows a greater rate of warming than RACMO(MPI-ESM).

Thank you for this suggestion; we have now implemented this change.

Figure 3: The error bars in each panel can be hard to read. Would it be possible to spread them out more so as to avoid overlap?

We have now adjusted this.

L221: What are the authors relying on to make this claim about increased snowfall over cold ice shelves? This explanation makes sense from a physical standpoint, but did the authors verify an increasing trend in snowfall over these ice shelves in their RACMO simulations?

We have verified this by checking similar scatter plots between DJF air temperatures and DJF precipitation and cloud cover. The relation with cloud cover was most significant. We have rephrased the sentence to:
"This can be attributed to increased atmospheric moisture content during warm summers, leading primarily to increased cloudiness, with a smaller contribution from increased snowfall (not shown). "

L224: Why is "near" in parentheses?

This should be near-infrared instead of (near) infrared. We have changed this now.

Figure 5 caption: Caption refers to figure panels by letter, but letter labels are missing from the figure.

Figure includes subplot labels now.

Figure 5: Points and lines are color coded according to average snowfall rate. I do not see where mean snowfall rate is discussed in the context of the albedo-temperature relationship.

This is discussed around line 240 in the original MS. We have added further discussion on this with reference to example ice shelves that are now indicated in the figure as follows:
"This is evident in Fig. 6 where the strongest decreases in albedo with increasing temperature occur at ice shelves with very low snowfall rates (e.g. Nansen and Publications ice shelves), whereas some ice shelves with high snowfall rates show little to no decline, or even an increase, in albedo."

Figure 5: it is difficult to read these plots. As noted by the author in the previous paragraph, one of the more interesting pieces of information conveyed here is the slope of the albedo-temperature relationship is different among ice shelves. This is evident in the fit lines, but it is hard to distinguish the fit lines from the points. Perhaps using different color scales for the fit lines and points could help? Also, while it is not practical to label all fit lines, perhaps annotating a few lines to highlight the difference between the relatively cold and warm ice shelves could clarify things.

We will adapt the figure so that the trend lines are not color-coded but are plotted in black on top of the scatter to improve clarity of the figure. We will add labels of the ice shelves to a selection of the fit lines.

Figure 7 caption: delete "is" from last line.

Done

L324: should be spelled "satellites"

Done

L341: Might read better as “Not only does the sensible heat flux become more important … at 0 °C, but atmospheric temperatures and moisture content can also continue to rise.”

Thank you for the suggestion, we have incoorporated this.