Dear Stef Lhermitte,

We thank you and Reviewer 1 (R1) for the helpful reviews of this paper. We correspondingly added all of the below recommended edits in the new version of the paper.

Thank you,
Marissa Dattler, Brooke Medley, and C. Max Stevens

Abstract

- **L11:** Replace "*Snow Radiative Transfer model*" by "*Snow Microwave Radiative Transfer model (SMRT)*"
- **L11:** Replace "*Community Firn Model*" by "*Community Firn Model (CFM)*"
- **L12:** I am not sure what this sentence means (specifically, I don't understand the word "hybridize" in this context), but this could also be because I am a non-native English speaker.
  *"In the process, we also **hybridize** our method to statistical techniques …"*
- **L14:** In the rest of the manuscript, you use a hyphen in "statically based", add here as well.
- **L19:** Add comma: *"… Antarctic snow, and (b) …"*

Introduction

- **L36-37:** Consider adding a reference (e.g.: Hofer & Mätzler (1980); Mote & Anderson, 1995)
- **L57:** For people not familiar with microwave sensors, it might be good to introduce AMSR-2. For example by adding some information to this sentence: "*…microwave radiometer Advanced Microwave Scanning Radiometer 2 (AMSR-2)…*"
- **L58:** I think it is "*Automatic Weather Station (AWS)"* instead of "Antarctic Weather Station (AWS)".
- **L59:** Suggestion to replace "*Antarctic sites*" by "*AWS*"
- **L60:** "*Other techniques*" sounds a bit vague. What do you mean by this?

Data & Models

- **L63:** Could you replace this by "*thirteen AWS*". Or are the three "dry sites" not automatic weather stations?
- **L64:** analyses (instead of analysis)
- **L83-85:** Please double-check the abbreviations. Once you have introduced abbreviations (e.g., CFM, AIS), for consistency, continue using these instead of the full names.

Methodology

- **L58:** Physically-based

Results

- **L340:** What do you mean with this sentence: "*This likeness is not the case for the pattern of melt end dates*"?

Discussion

- **L402:** "*19H*" instead of "*18H*"
- **L419-429:** Very interesting paragraph. I had some similar findings in a paper where I compared different statistically-based methods for melt detection (de Roda Husman, et al., 2022).
- Really nice and well structures discussion!

Conclusion

- **L533:** I thought *thirteen* sites?

Figures

- **Figure 1:** Replace "*AIS*" by "*AWS*" in figure caption
- **Figure 2:** Consider rewriting the figure caption, with a first sentence that describes main idea figure, and summarize the rest of the text. Same holds for **figure 4, figure 7, figure 8, figure 10, figure 11, figure 12.**
- **Figure 3:** Shouldn't "*Calculate correlation length*" be replaced by "*Calculate microwave grain size*"?
- **Figure 4:** I don't see the red area in (c), is this missing?
- **Figure 7:** I would use a sequential color palette for the melt duration instead of a diverging one. The white color (around 65 days) seems to have no melt now (if you quickly look at the figure), which is a bit confusing. Same holds for **figure 12**.
- **Figure 8:** In (b) and (d), you cannot see the Picard et al. melt days if the Hybrid Method shows melt. I would suggest to use a symbol for days where both method show melt, or make Picard et al.'s symbol a bit larger, because now it seems that on days where the Hybrid Method shows melt, Picard et al's method does not.
- **Figure 9:** Consider adding the Pearson Correlation Coefficient to (b). Same holds for **figures 10-12.**
- **Figure 10:** The variation in color bar limits is somewhat confusing. Would you consider standardizing the limits to enhance clarity?

Used references

- T. L. Mote and M. R. Anderson (1995), "Variations in snowpack melt on the Greenland ice sheet based on passive-microwave measurements," Journal of Glaciology, vol. 41, no. 137, pp. 51–60.

Minor comments by editor:
- L31+51+59+... "statistical techniques" vs "statistically-based techniques" vs "statistically-based thresholding techniques" vs "statistically-based melt detection techniques" vs "statistical thresholding techniques". All these terms are used interchangeably, but it would be beneficial to use one common term throughout the manuscript. As an editor I prefer "statistical thresholding techniques" as they contain the most info + are most concise.
- L63: "We run our melt detection technique for ten AWS and three additional dry sites".
* The difference between AWS and additional sites is not very clear as the description is not very clear. How was the melt determined on the additional sites where no AWS was present? Is the difference between them that they are wet vs dry or are the AWS sites also dry/dry sites also with a AWS? Later on it becomes clear that the dry sites are used separately for validation of the temperature + grain size, but this is not clear when describing the sites
* Why do you mention three dry sites when it seems only Dome C is actually used in the manuscript? The others are shown in the manuscript but never referred to so it remains unclear what their role in the paper is.
- L94 "merge": How was this merging done? Based on which criterium? How were the characteristics of the merged layers determined?
- Figure 5e + Figure 6 + L311-316: It would be beneficial to add the melt fraction (in %) at the right hand side of the figure so the reader can directly see the percentages without having to look them up in Table 1.
- L365 "2018). The Pearson correlation coefficient between these two grain size variables is 0.88 ($p < 0.01$)." Refer to Fig. 9b + add other common statistics like RMSE, slope etc. Moreover, I think it would be beneficial to add the regression line and 1:1 line to Fig. 9b.
- Figure 7+12: some panels use a diverging color palette whereas the data are not diverging. This may lead to misinterpretation etc. Therefore, check "The misuse of colour in science communication" https://www.nature.com/articles/s41467-020-19160-7 and specifically Fig.5+6 for better choice of color palettes.
- L377: "we can see". Refrain from subjective phrasing
- L379-380 "On Dronning Maud Land there is moderate durations of melt and moderate microwave grain sizes, and on the Antarctic Peninsula there is higher durations of melt and higher microwave grain sizes" -> there are + longer durations + larger grain sizes
- L404 + L442"computational efficiency" I think it is important to clarify for the reader the computational impact. Indicating how long does it take to run the hybrid model on what type of computer setting, would help the reader to better understand the computational impact.
- Supplement: what is the role of the supplement in the paper (?) as it seems nowhere in the paper there is any reference to the supplementary material.