Review of Dattler et al. 'A physics-based Antarctic melt detection technique: Combining AMSR-2, radiative transfer modeling, and firn modeling'
By Sophie de Roda Husman (S.deRodaHusman@tudelft.nl)

The manuscript introduces a novel algorithm, referred to as the Hybrid Method, designed for detecting surface melt in Antarctica. This approach leverages not only remote sensing data, as seen in conventional methods, but also integrates outputs from both the Community Firn Model and Snow Microwave Radiative Transfer Model. In this study, the threshold for discerning surface melt varies on a daily basis and is determined based on the variance of the correlation length over a specific time frame. Notably, this study also presents the correlation length as an output.

I consider this study to offer a highly valuable and innovative approach to melt detection. However, the manuscript could benefit from providing further clarity on the added value of the Hybrid Method. To strengthen its contribution, additional validation efforts should be undertaken. This might involve expanding the temporal scope for comparisons, offering less aggregated results, and broadening the spatial validation by incorporating more than two weather stations. These enhancements would contribute to a more robust evaluation of the method's effectiveness.

Strengths of this manuscript

Enhancing an algorithm with additional physical information is a noteworthy contribution to the existing literature. The authors have skillfully integrated numerous data products, and the clarity of the figures is notable. I found the flow chart in Figure 3 particularly effective in explaining the somewhat intricate methodology used in the manuscript.

Major comments

1. The **abstract of the manuscript requires refinement** as it lacks clarity regarding the research objectives, methodology, and key findings. It is crucial to distinctly highlight the primary advantage of the Hybrid Method over traditional thresholding techniques for melt detection. The assertion that "...this method is as accurate as previous statistically based thresholding techniques..." (L15) lacks persuasiveness and does not provide clear guidance on when and where to preferentially apply the Hybrid Method over established approaches.
2. This brings me to a critical point of consideration: upon reviewing the manuscript, I find myself uncertain about the **specific scenarios in which the Hybrid Method surpasses existing methodologies**. While Table 1 and Figure 8 offer valuable insights, they present aggregated data. It is crucial to discern when, within the melt season, and ideally, where in Antarctica, the Hybrid Method demonstrates superior performance, along with understanding the underlying reasons for this, as well as why alternative methods may fall short. Figure 7 could potentially provide assistance in this matter, although it hasn't been addressed in the manuscript as of now.
3. I find it crucial for the manuscript to **include an explanation of the physical interpretation of 'correlation length'**. Understanding the significance of this parameter and why it serves as an indicator of surface melt presence is

essential. A clear description of the correlation length would not only fortify the manuscript's overall strength but also enhance its readability.

4. Some of the important choices for developing the **melt detection threshold**, seemed arbitrary, examples are:
    a. $4\overline{\sigma_{p\_exp}}$ (L198, where is the 4 coming from?)
    b. +/- 7 days (L125, why 7 days?)
    c. 31-day window used to compute $4\overline{\sigma_{p\_exp}}$ (L197, why 31 days?)

   The manuscript would benefit from some further explanation of these numbers and potentially some sensitivity studies in which a variety of values is tested.
5. At present, the manuscript primarily relies on point-based comparisons for surface melt assessment involving AWS, the Hybrid Method, and other melt detection algorithms. Including **Antarctic-wide spatial maps** of the Hybrid Method, Picard's method (or other conventional melt detection algorithms), and the differences between the two regarding parameters like (1) number of melt days; (2) onset of the melt season; and (3) end of the melt season, would offer supplementary perspectives on the strengths and limitations of the Hybrid Method. Additionally, incorporating Antarctic-wide spatial maps of the newly introduced correlation length, along with snow grain size data from the Mosaic of Antarctica, would serve as valuable supplements to the manuscript.
6. The authors really challenged themselves by making a manuscript of which the **goal is two-fold**: creating a correlation length product and developing a new melt detection algorithm. I would suggest **adding some extra structure** to the manuscript (for example, adding clear sections to the result and discussion section in which both topics are presented).
7. The evaluation of the method is confined to just **two weather stations** (AWS17 and AWS18), which may be deemed **inadequate for comprehensive validation**. The manuscript mentions this selection was based on the significant overlap with AMSR-2 data, which left me somewhat puzzled. As far as I am aware, AMSR-2 data has been accessible since 2012. This prompts me to wonder why data from stations like AWS4, AWS5, AWS11, AWS14, AWS15, AWS19, and the Neumayer station are not utilized in Section 4.2.
8. In this study, both the **horizontal and vertical frequencies** of AMSR-2 are employed. I propose focusing solely on the horizontal frequency for the Hybrid Method, and the vertical frequency for the derivation of the correlation length. Now, sometimes one frequency (for example in Table 1), and sometimes both frequencies (for example in Figure 8) are used for the Hybrid Method. It remains unclear why both frequencies are employed for deriving surface melt presence and which one is the preferred choice for the Hybrid Method.

<u>Minor comments</u>

Abstract

- L10: I suggest to mention the underlying challenge here. Why does the scientific community necessitate a novel method for melt detection? It appears that you have well-founded reasons, as outlined in lines L33-L36.

Perhaps you could further emphasize that existing melt detection methods may falter in certain specific scenarios.

- L12: I suggest explicitly referring to your new method as the “Hybrid Method”, and introducing the name of the method here. For example, by changing “… to create a hybrid method …” into “… we created a novel method, referred to as the Hybrid Method, …”.
- L15: Quantifying the performance of the Hybrid Method compared to other methods (for example by using accuracy values of the Hybrid Method and other methods) would strengthen your point of introducing your novel method.
- L17: I would add the sentence about significant correlation (did you check for significance, by the way?) at the beginning of the abstract, thereby making the argument that correlation length is a good variable to use for thresholding surface melt presence.

I. Introduction

- L21: This is the first time the abbreviation ‘AIS’ is used, so please expand it to the *Antarctic Ice Sheet (AIS)* here (and not in L26).
- L34: Can you include a citation here? Also, consider referring to some studies that showed why and when certain conventional thresholding techniques fail. Maybe Johnson et al. (2020) could help here.
- L40: melt detection (instead of *melt-detection*)
- L44: I recommend providing a concise explanation of correlation length. Currently, it is only briefly mentioned in parentheses, yet it holds significant importance for the remainder of the manuscript.
- L45: Please introduce the abbreviations CFM and SMRT.
- L48: statistically-based (instead of *statistically based*)
- L49: What do you mean by ‘intermediary calculations of snow microstructure’?

II. Data & Models

- L51: What is meant by ‘AIS point’? Maybe replace it with something like ‘Automatic weather stations’
- L53: I recently also used the data of Jakobs et al. (2020) and noted there was a small typo in Table 2 with the coordinates. Not sure if you used this table, but just to be sure, I believe the coordinates of AWS18 are (66.40; -63.37) instead of (66.40; -63.73).
- L57-L60: Why are there two starting and two ending days mentioned for AWS17 and AWS18? Can the last sentence (L59-L60) be removed?
- L66: Section 2.2 would benefit from some further elaborations. Some questions I still have:
    - Which period do you use?
    - What polarizations do you use (this becomes clear in Section 3, but might be nice to add here as well)?
    - From where do you download the data product mentioned in L66?
- L65: You refer to 18.7 GHz by shortening it to 19 GHz, while throughout the rest of the manuscript, 18 GHz is consistently used. In my suggestion, I recommend adhering to 19 GHz consistently, as it aligns with conventional terminology (as seen in works like Johnson et al., 2020).

- L72: add the source of the data product Mosaic of Antarctica
- L72: What does it mean for your final melt product that CFM underestimates snow grain size (compared to Mosaic of Antarctica)?
- L82: Mention what the exact output of CFM is (snow density and temperature, right?).
- L84: Consider replacing thermal emission with *brightness temperature* and backscattering with *backscatter intensity,* for consistency.

III. Methodology

- Consider adding a section (3.1) on 'deriving surface melt presence with conventional thresholding techniques'. The melt presence derived from Picard et al., Torinesi et al., and Zwally and Fiegles are presented in the figures, but a short explanation of the methods is missing.
- L102-103: You mention that the CFM grain size profiles can be converted into correlation length with 'some uncertainty'. Can you add some additional explanation here?
- L112: What 'outside source' is meant here?
- L134: Replace 'these two polarizations' with 'horizontal and vertical polarizations'
- L139: How do you select a 'high' and 'low' correlation length value? And what is the physical meaning of a 'high' and 'low' correlation length? I can imagine that for a low correlation length, the snow properties are more heterogeneous. Therefore, I expect conditions to be more conducive to melt, as local heterogeneities and variations in temperature or impurity concentration can lead to localized melting processes.
- L144: It's a bit confusing that terms like 'many times' and 'narrower bands' aren't quantified here. While this information is provided later in the manuscript (L157-159), I would suggest revising this section to ensure a smoother flow of the narrative, without the need to constantly refer back and forth.
- L148: You might want to consider giving more intuitive names to Process 1, Process 2, and Process 3, in the text. For example, you could label them as follows: "Step 1: Calculate Correlation Length", "Step 2: Assign Melt Days / Dry Days", and "Step 3: ??", as you do in Figure 3. By the way, it appears that "Step 3" (or Process 3) is not represented in Figure 3.
- L190: What type of 'other information' is meant here?
- L191: To maintain consistency, I recommend using 'melt' and 'non-melt' days, as you have done in this line, rather than 'melt' and 'dry' days, which are used elsewhere in section 3.1 and Figure 3.
- L198: You could add 'hereafter referred to as $4\overline{\sigma_{p\_exp}}$ *or dynamic threshold*).

IV. Results

- L215-L222: These lines introduce a significant amount of supplementary information, which may divert attention from the main narrative. Could this be included in the method section rather than the results?

- L224: Wouldn't it be more appropriate to introduce SSA in the data section? And also the next line (L225), seems to be more fitting for the method section that for the results.
- L245: Consider referring to AWS melt, instead of SEB melt, in the section title.
- L258: What is meant by 'increases across each of the four austral summers shown'? I was expecting the average correlation length per melt season to progressively rise compared to the preceding season, but this trend isn't clearly evident in Figure 5b.
- L265: Also Torinesi's method yields an accuracy of 91.7%, correct? While you refer to it as the 'second best technique', from my perspective, it actually attains the lowest accuracy (tied with Zwally & Fiegles' method). Could you clarify if I'm overlooking something in this regard?
- L290: As I highlighted in the 'Major comments' section, I believe the manuscript would greatly benefit from including a comparison with additional weather stations. Comparing the Hybrid Method with other statically-based techniques may not provide substantially new insights. Given that statistically-based methods are acknowledged to have limitations in accurately detecting melt (which prompted the development of this new method, as you state in the introduction section), what meaningful insights do we gain from this comparison?
- L295: I would consistently use capitals for your new method, using 'Hybrid Method' rather than 'hybrid method'. Currently, both are employed (sometimes also in one sentence, for example in L599).
- L301: Could you explain why this is not the case for AWS15?
- L309: I would suggest to introduce MOA in the method section.
- L303-L322: This appears somewhat abrupt and disrupts the flow of the narrative. Perhaps consider creating a new subsection in the results dedicated to the evaluation of the generated correlation length product?
- L314: I'd be interested in seeing a spatial map that compares MOA's optical grain size with your correlation length product. Additionally, a scatter plot illustrating the correlation between both products would be informative.

Discussion

- General comment: I recommend using distinct titles for the topics discussed in the subsection to enhance readability.
- L325: Consider incorporating more specific statistical terminology, instead of mentioning that the methods are 'tied'? This might enhance the clarity of the comparison.
- L338: Which two sites do you mean?
- L339: Can you quantify this statement?
- L363: What do you mean by 'resulting overestimate of emissivity'?
- L379: What do you mean by 'that site'? Are you referring to a specific location or is this a general statement?
- L379: Did you test that there was a significant correlation? This might be good to add to the methods section.
- L399: In my (somewhat limited – sorry about that!) understanding of correlation length, I thought that surface melt could also result in a decrease

in correlation length. As the ice undergoes a phase change from solid to liquid, this transition can introduce irregularities in the ice structure, such as the formation of pores, cracks, or fractures filled with meltwater. I thought this process could reduce the correlation length. I assume this is incorrect, right?

Conclusion

- Looks good!

Author contributions

- Shouldn't it be MS instead of CS?

Data availability

- Really nice that the data are shared on Zenodo.
- Could you provide information about the sources from which you obtained all the data products utilized in this study?

Figures

- **Figure 1:** Consider adding a title to the legend, indicating something like 'AWS located on:' for clarity. Currently, it might be a bit confusing at first glance.
- **Figure 1:** In the caption, I think 'AIS' should be replaced by 'AWS'. Or, even better, avoid abbreviations in the figure captions, and replace by 'Antarctic Weather Stations'.
- **Figure 1:** The shades of blue and cyan appear quite similar to me. Please consider replacing one of the colors?

- **Figure 2:** Nice figure! One very small comment, maybe you could add one line with a general description at the beginning of the caption?

- **Figure 3:** Nice figure!
- **Figure 3:** Process 3 is mentioned in the text (L148), why is this step not presented in the figure?
- **Figure 3:** What are TB_1 and TB_2? Do they refer to horizontal and vertical polarizations, or the 'low' and 'high' brightness temperatures coming from the 'low' and 'high' correlation lengths?

- **Figure 4:** Again, I struggle to see the difference between blue and cyan.
- **Figure 4:** Could you clarify what is referred to as "NIR SSA" in the caption? Is this data product discussed in the text? Additionally, consider using more intuitive labels, such as indicating the 4 and 10 meter penetration depths, instead of " REFL. SP1" and "REFL. SP2".
- **Figure 4:** In subpanels (c) and (d), I cannot see the AMSR-2 line. Consider increasing the linewidth, so you see it is overlapping with the other lines.
- **Figure 4:** I would suggest not abbreviating 'BC' in the caption, as there are already numerous abbreviations in use, which can hinder readability.
- **Figure 4:** In panel (c) and (d), the legend states that **-**$4\overline{\sigma_{p\_exp}}$ is shown. Shouldn't this be **+**$4\overline{\sigma_{p\_exp}}$? And in panel (b), both + and – are shown. I got a

bit confused here, when to use – (like in panels c and d), and when to use – and + $4\overline{\sigma_{p\_exp}}$ (like in panel b)?

- **Figure 4:** Please consider merging panels (c) and (d). If they are plotted separately, the distinction isn't clear to me. Alternatively, displaying only panel (d) and (e) could effectively convey the message.
- **Figure 4:** In panel (a), 18H is depicted, while in panels (c) and (d), 18V is utilized. It is not entirely clear to me why 18H is employed for melt detection, and 18V for establishing the threshold.
- **Figure 4:** In panel (b), I find it challenging to discern the median filter. Additionally, could you clarify the purpose of using the median filter?

- **Figure 5** and **Figure 6:** I'm having difficulty comparing the melt days obtained from the Hybrid Method (panel d) with those derived from the statistically-based method (panel a). Could you consider plotting them in a single figure for easier comparison?

- **Figure 7:** This figure is not mentioned in the text, please include it in the results section.

- **Figure 9:** I recommend placing this as the initial figure, possibly in the form of a scatter plot where you plot % melt days against correlation length. This visualization would effectively demonstrate that correlation length is a reliable indicator for surface melt, thereby bolstering your argument for using this method.

Used references

- Johnson, A., Fahnestock, M., & Hock, R. (2020). Evaluation of passive microwave melt detection methods on Antarctic Peninsula ice shelves using time series of Sentinel-1 SAR. Remote Sensing of Environment, 250, 112044.