Response to Emanuele Santi (Referee 2) on egusphere-2023-1556

First, we would like to thank the Referee for reviewing and commenting on the manuscript, which will improve the quality of the manuscript. Please find the item-by-item reply below, with the original comments in *italics* and the responses in blue. All the suggested changes will be implemented in the revised text that will be uploaded.

The subject of this manuscript is of definite interest for the scientific community. Introduction correctly frames this study in the existing literature, language is clear, and thread deploys smoothly. Innovation with respect to other studies should be however better pointed out and description should be improved in some respects, as well as the presentation of the results. Beside this, the paper suffers from some lacks in the microwave background and I'm suspecting two conceptual issues: the first deals with the attempt to retrieve the density for the 4 cm top layer, which should be quite transparent at the considered MW frequencies in dry conditions. The second concern is about merging direct satellite measurements and derived indices in the RF inputs: based on the information theory, the indices should not bring any additional information independent of the Tb from which they have been computed, so, also based on my experience, these indices should negligibly affect the results.

We thank the referee for the constructive review and suggestions. We will polish the introduction and the presentation of the results in the revised manuscript. Regarding the penetration of the MW frequencies, while we agree that theoretically a depth of 0.1—2 m should be a more reasonable choice for both 19 GHz and 37 GHz, as we cited in Line 94 of the manuscript (Surdyk, 2002; Brucker et al., 2010), our study is also based on the assumption that the frequency ratios should reflect near-surface (0—2 cm) density, as in Champollion et al. (2013) and Leduc-Leballeur et al. (2017). Therefore, as a compromise between the theoretical penetration depth and the aforementioned applications, we will switch to a depth of 0.12 m to perform the experiments in the revised manuscript. But certainly, as the Referee also pointed out in the detailed comments, the relatively reasonable results can be obtained based on indirect correlation of the top layer density with deeper layers which indeed influence the adopted frequency more. This will also be included in our discussion.

Regarding the validity of using the derived indices, our study was motivated by Tran et al. (2007) and Champollion et al. (2013). Tran et al. (2007) combined a derived Tb ratio with Tb values to cluster snow facies in both Greenland and Antarctica, and Champollion et al. (2013) could associate frequency ratios to near-surface grain size and density at Dome C, Antarctica to a certain extent. In both studies, the validity of using such ratios exists to an extent that should be interesting to discuss, hence we included them.

Moreover, since we use Random Forest regression, we do not agree that the indices cannot bring additional information or performance. While the principle of information theory indeed suggests that indices derived from the original should not introduce additional information, it's important to consider the context in which certain techniques, such as random forest regression, operate. Random forest regression is a powerful ensemble learning method that harnesses the collective strength of multiple decision trees. In the case of random forest regression, the combination of diverse decision trees allows for the detection and extraction of intricate patterns and relationships within the data that may not be readily apparent in the original dataset. Each tree contributes its unique perspective, and the ensemble's output is often more robust and accurate than that of an individual tree. Therefore, although the indices derived from the original data may seem, from an information theory standpoint, to contain similar information, the strength of random forest lies in its ability to uncover latent, complex patterns that might not be explicitly present in the raw data. This enables the model to provide more nuanced and accurate predictions, surpassing the limitations of a single decision tree.

However, since our study aims to assess a method and discuss the validity of the parameters, we assume that a sensitivity analysis of using different combinations of parameters could be added, where we use as input:

- All parameters as what we are using now
- Only absolute Tb and sigma0
- Only absolute Tb and sigma0, and derived ratios (as also pointed out by Referee 1)

Detailed comments:

Introduction.

- *The introduction contains a review of the state of the art more than enough to frame this paper. I would only suggest clarifying the aspects related to different spatial resolution, coverage and revisiting when mentioning active and passive MW.*
  This will be clarified in the revised manuscript.

Section 2

Section 2.1.

- *Equation 1 and 2 are properly referred to the original publications, however a short sentence about the physical principles behind would be useful for the reader.*
  This will be added in the revised manuscript.
- *Line 94-95. The dramatic change in emission mechanism due to the presence of liquid water within the ice sheet might be commented, although this point is mentioned later in section 3.2. Same applies to the scattering in section 2.2.*
  This will be added in the revised manuscript.

Section 2.2.

- *the linear correction for local incidence (LIA) sounds me a bit odd. LIA should be already accounted for when computing NRCS to extract the backscattering (σ°). In any case the backscattering dependence on LIA is not linear at all. Finally, as far as I understand from pag. 5 line 125, at the end you did not use data corrected with eq. 3. Could you further clarify?*
  Perhaps we did not specify the parameters properly. Equation 3 does not describe how we processed the data, but which kind of dataset we used. The same equation can be found in Lindsley and Long (2010), Eq. 3 on page 3. What we are using is the σ° normalized to the reference angle (40°), referred to as A in this equation. The A

products we are using are already available via Brigham Young University (BYU) Microwave Earth Remote Sensing (MERS) laboratory platform and are directly used in our study. However, since A as a single letter could be misleading, we called it σ° again in the following texts of the manuscript, which is more familiar to the common knowledge. We will clarify that in the revised manuscript.

- *The spatial and temporal co-registration between ASCAT and SSMIS should be better described, this could lead to error and artifacts depending on the processing you applied. At the end, how many co-located Tb and σ° you obtained? It is an important information for better understanding the RF implementation, although something is addressed later.*
  Eventually, we obtained 19,027 valid pixels within the Antarctic ice sheet range (Table 1 of this document). We admit that with a linear interpolation, artifacts occur at the edge of the images. However, we filtered them out using the coastline from Depoorter et al. (2013). What falls within the range of the Antarctic ice sheet should be reliable.

Section 2.3

- *line 135 – 138. As stated in the general comments, the attempt to retrieve density at 4 cm raises a conceptual issue. The top 4 cm layer should be almost transparent not only at C-band but also at Ka band in case of dry firn. I'm wondering if you are obtaining results based on indirect correlation of the top layer density with deeper layers to which microwaves are instead sensitive. No wonders if RF achieves successful retrievals: machine learning can exploit almost any kind of input/output relation, but the risk of finding out something based on apparent relationships is always around the corner. If used as "black boxes", ML could potentially relate newborns in China and weather in USA, but which is the utility? I believe a robust physical justification is needed.*
  We agree that it is likely that we have obtained results based on indirect correlation of the top layer density with deeper layers. Please refer to the major comments.
- *Line 138 – 140. The sentence is unclear to me, could you rephrase please. Where was density at 1m depth used later?*
  It was shown in Fig. 4 of the manuscript to prove that melt events have a prolonged impact on deeper snow densities, hence our clustering step to separate dry and melted pixels was quite reliable. However, we agree that overall it does not have added values to the following analyses, hence will remove it in the revised manuscript.

Section 3

Section 3.2.

- *Is Tb Ratio the same of eq. 1? If so, no need to introduce it again with reference.*
  Yes. This will be changed in the revised manuscript.
- *In my understanding, volume decorrelation was not introduced before. The cited work by Rizzoli is using X band SAR, it is not clear if this finding is also valid for radiometric measurements (scattering and emission are complimentary each other)*
  This application was introduced in Line 59 as previous studies. We repeatedly mentioned it here to point out the difference between our clustering and the previous

studies. However, X-band SAR indeed does not have anything to do with our method, hence we will revise this paragraph.

- *Line 195 – 199. The normalization by firn temperature is embedded in both parameters you defined in eq. 1 and 2. Which is therefore the reason for removing the average seasonal Tb signal? And which the one for doing the same with backscattering that is almost insensitive to temperature? Moreover, machine learning techniques as RF can cope with redundant, noisy, and biased data, so dealing with timeseries of measurements or their anomalies should not change much the results. Finally, there is also a concern in merging Tb with their ratios that is commented below.*
  This step (everything in Section 3.2) serves to separate dry pixels from pixels that suffered from melt, therefore Tb ratios and RF are not used here. While backscattering is almost insensitive to temperature, it is very sensitive to melt events and subsequent melt layers, hence the anomalies should be a good indicator of melt pixels. We agree that this motivation has not been clarified in the manuscript, hence we will improve it.
- *Lines 201 – 211. The clustering algorithm should be better explained maybe with a supporting figure/diagram. I don't believe a reader unfamiliar with Ward algorithm can understand this section.*
  This will be improved in the revised manuscript.

Section 3.3

- *Lines 231 – 246. With "sample" do you refer to the set of temporally and spatially coregistered SSMIS and ASCAT measurements for the given pixel? In my understanding, for both subsets I and II you selected randomly 100 pixels from the 7 clusters over Antarctica described in section 3.2 (that is spatial, 25 km resolution each pixel) and you considered the timeseries of satellite measurements (that is temporal, approx. 1 set of SSMIS + ASCAT measurements per pixel day per 10 years). At the end you should have used 365300 sets for training and the same data amount for testing. In other words, you considered about 125000 Km2 for training and testing and applied the trained RF on the remaining ≃14000000 Km2 of Antarctic surface, which is notable. Maybe some more information could be provided…*
  We indeed refer to "the set of temporally and spatially coregistered SSMIS and ASCAT measurements for the given pixel". This will be clarified in the revised manuscript. But for subsets I and II we selected 100 pixels from the 4 dry clusters instead of the 7 clusters. We should also clarify a mistake in the original manuscript that since we use the 10-day resolution IMAU-FDM, the total time slots should be 366 instead of 3653. Therefore, we used approximately 0.6% of the total data for training. We will use 10% of the total data for training (as subset I) in the revised manuscript, as is slightly improves the result (Fig. 1 of this document) and is theoretically more reasonable than using 100 pixels (~0.6% of the data).

Table 1. Statistics of pixels per cluster and pixels used for RF estimating in the original manuscript.

Cluster	Number of pixels	Number of training pixels
Firn 1	4540	26

Firn 2	7360	42
Firn 3	3465	20
Firn 4	2284	12
Firn 5	429	0
Firn 6	325	0
Firn 7	624	0
Total	19027	100

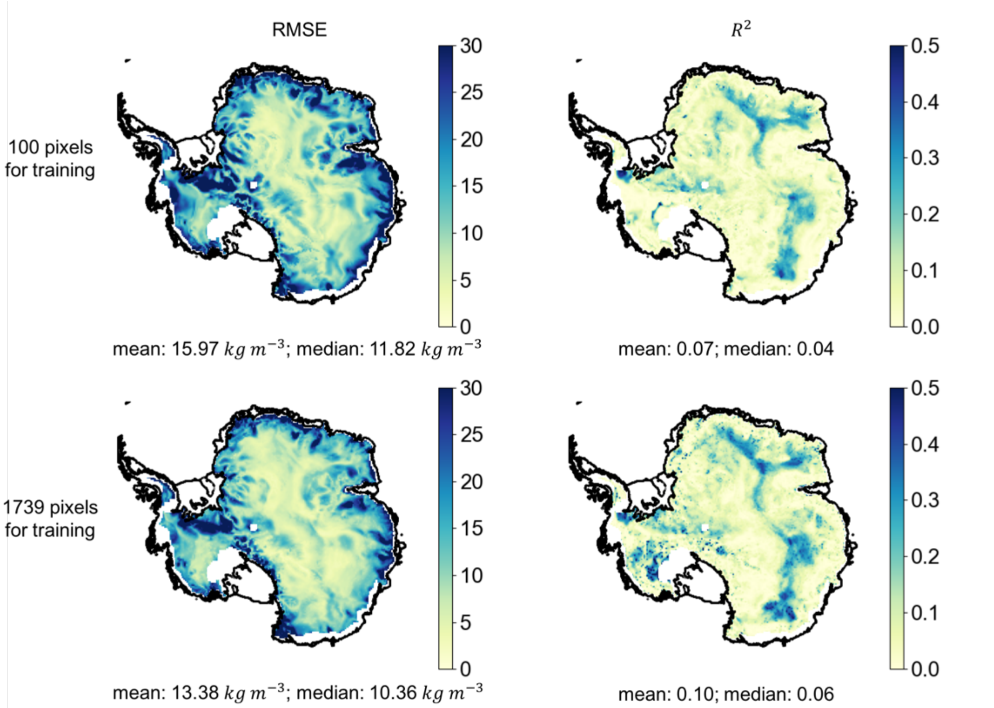

Figure 1. Comparison between using 100 pixels and 10% of the data for training.

- *Equation 4. The proposed input combination raises another concern: from the information theory, the Tb ratios do not bring to the RF additional information independent of the Tb from which they have been computed, therefore (this is also my personal experience) the results should not be affected by these inputs (or conversely by Tb if you use the ratios). Clarification is needed.*
  Please refer to the major comments, where we argue why derived indices can effectively add information when used in Random Forest regression as these are based on decision trees and derived indices can play an important role there (as they might be important in different phases of the decision tree)
- *Line 258. Gini importance should be better referred and briefly commented. Which is the difference with e.g., predictor importance proposed by Breiman?*
  This will be improved in the revised manuscript. Regarding the difference between the Gini importance and the permutation (Breiman) importance (Fig. 2), we notice that using the permutation importance, the ranking of the original horizontal channels goes down. We will use both in the revised manuscript, or switch to the permutation importance.

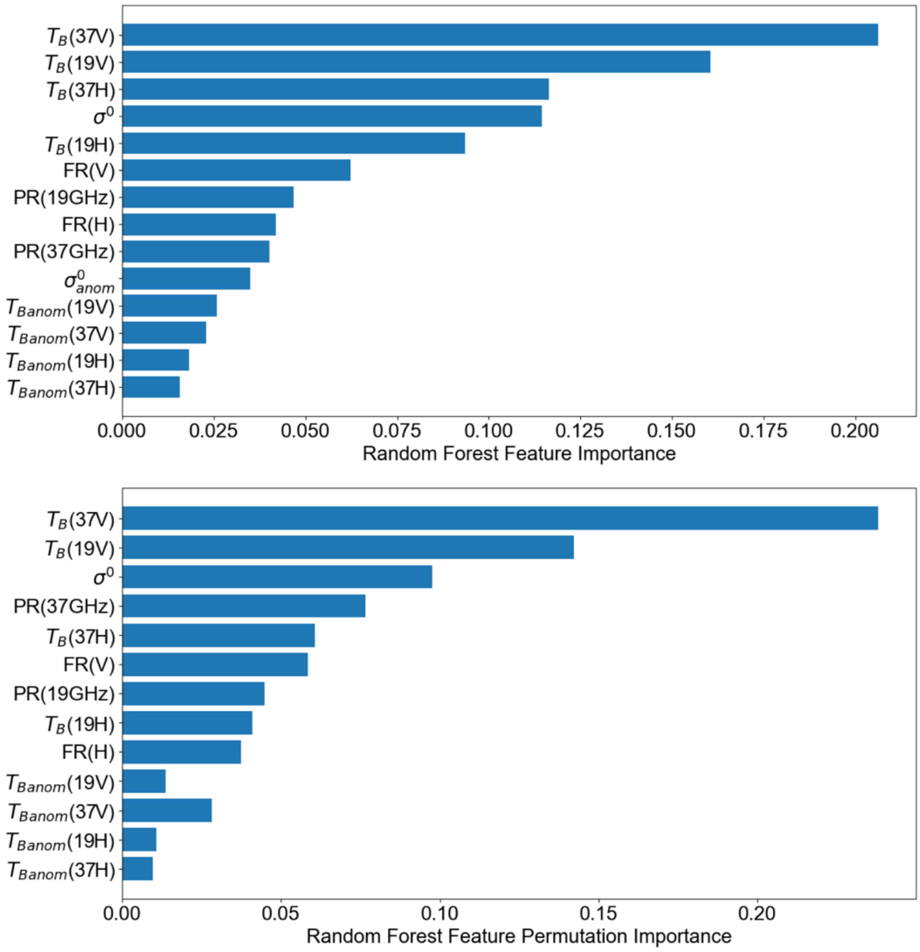

Figure 2. Comparison between upper: Gini importance, and lower: permutation (Breiman) importance.

Section 4

- *Section 4.1. Following the comment above, this is the core of my concerns: the scarce correlation with density at 4 cm could be depending on the microwaves' scarce sensitivity to such shallow depth. Also, the reverse correlation along the coasts should be depending on melting not entirely removed that occurs more frequently than in the central part of Antarctica. Again, the physics behind should be analysed.*
  We agree. A deeper snow density (12 cm) will be assessed, and the potential melting will be added in the discussion.
- *Figure 1. Although referred to in section 4.1, I find this figure poorly informative. My suggestion is to remove or replace with something more meaningful.*
  We would like to keep it to give the reader an overview of the parameters we used, including their spatial patterns.
- *Figure 2. Did you evaluate the correlation with density at 1 m? At the end which was the role of this parameter in your study?*
  We did not evaluate the correlation with density at 1 m. Ideally we intended to show that our analysis should serve for multiple depths, however this was not well addressed in the manuscript. Please note that in the revised manuscript, we opt for

assessing the densities at 12 cm depth instead of 4 cm, therefore the descriptions will be revised accordingly.

- *Figure 4. The plots in the figure are quite small and difficult to read. I would suggest revising.*
  This will be improved in the revised manuscript.
- *Figure 5 left: the scatterplot should refer to the test results (i.e. those obtained on subset II), not to the training results (Subset I). Usually, retrieval scatterplots show the estimated vs. target, not vice-versa. The plot or caption should also cite the statistics and total data amount. Finally, the R value seems even worse than the one of direct correlation with Tb Ku and Ka in figure 2 for most of the pixels. Isn't it? Which is the explanation?*
  It was a mistake in the captions. We indeed used Subset II for this analysis. This figure will be improved in the revised manuscript where we will also swap the axes.

  However, here the R^2 value (ranging between 0 and 1) refers to the linarity, i.e. if we fit a line to the estimated vs. target scatter plot, how the goodness of fit is. This is not the same indicator as the correlation coefficient (ranging between -1 and 1) in Fig. 2.
- *Figure 6 with doubled colorbar is difficult to interpret (especially figure 6d). I would suggest revising.*
  This will be changed into different markers, as also pointed out by Referee 1.
- *Figure 7 why do not also add the Correlation/Determination coefficient maps as those in figure 2? In my view this is more informative than e.g. the 10- years averaged maps of figure 6.*
  We have shown R^2 in Fig. 7 of the manuscript which is the coefficient of determination between FDM and RF, so we believe it has been an informative indicator already. An averaged map in our opinion shows that our method works reasonably well spatially (in contrast to the performance temporally, as shown in the sections afterwards). However, we agree that the metrics of our assessments should be clarified.

Reference

Brucker, L., Picard, G., and Fily, M.: Snow grain-size profiles deduced from microwave snow emissivities in Antarctica, Journal of Glaciol ogy, 56, 514–526, https://doi.org/10.3189/002214310792447806, 2010.

Champollion, N., Picard, G., Arnaud, L., Lefebvre, E., and Fily, M.: Hoar crystal development and disappearance at Dome C, Antarctica: observation by near-infrared photography and passive microwave satellite, The Cryosphere, 7, 1247–1262, https://doi.org/10.5194/tc-7-1247-2013, 2013.

Depoorter, M. A., Bamber, J. L., Griggs, J., Lenaerts, J. T. M., Ligtenberg, S. R. M., van den Broeke, M. R., and Moholdt, G.: Synthesized grounding line and ice shelf mask for Antarctica, https://doi.org/10.1594/PANGAEA.819151, supplement to: Depoorter, MA et al. (2013): Calving fluxes and basal melt rates of Antarctic ice shelves. Nature, 502, 89-92, https://doi.org/10.1038/nature12567, 2013.

Leduc-Leballeur, M., Picard, G., Macelloni, G., Arnaud, L., Brogioni, M., Mialon, A., and Kerr, Y.: Influence of snow surface properties on L-band brightness temperature at Dome C, Antarctica, Remote Sensing of Environment, 199, 427–436, https://doi.org/https://doi.org/10.1016/j.rse.2017.07.035, 2017.

Lindsley, R. D. and Long, D. G.: Standard BYU ASCAT Land/Ice Image Products, Tech. rep., Brigham Young University Microwave Earth Remote Sensing (MERS) Laboratory, https://www.scp.byu.edu/docs/pdf/MERS1002.pdf, [Access date: Oct. 12, 2023], 2010.

Surdyk, S.: Using microwave brightness temperature to detect short-term surface air temperature changes in Antarctica: An analytical approach, Remote Sensing of Environment, 80, 256–271, https://doi.org/10.1016/s0034-4257(01)00308-x, 2002.

Tran, N., Remy, F., Feng, H., and Femenias, P.: Snow Facies Over Ice Sheets Derived From Envisat Active and Passive Observations, IEEE Transactions on Geoscience and Remote Sensing, 46, 3694–3708, https://doi.org/10.1109/tgrs.2008.2000818, 2008.