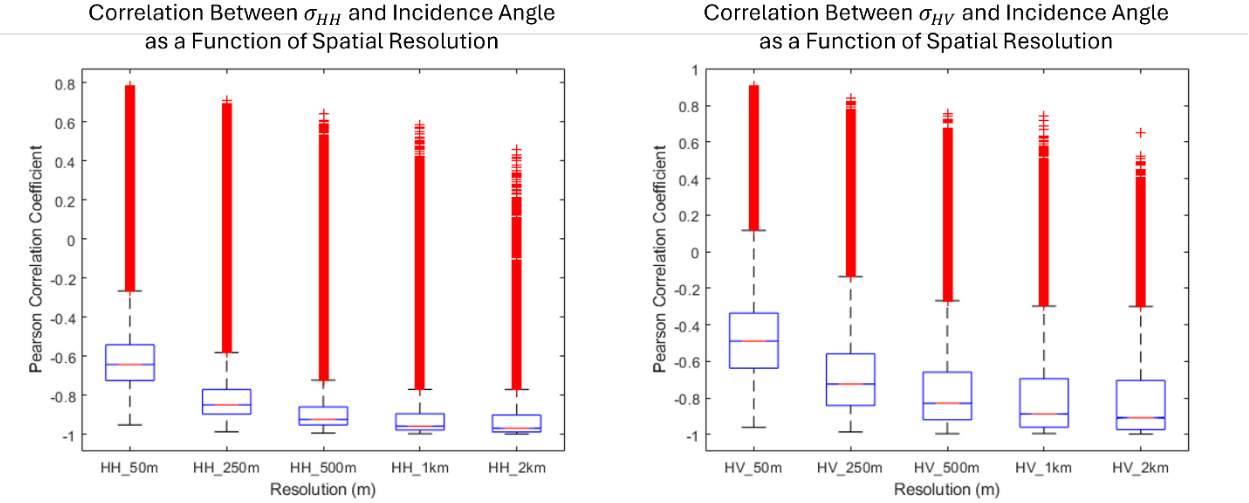
Review of "Sentinel-1 Detection of Ice Slabs on the Greenland Ice Sheet"

This paper presents an investigation of the potential for detecting ice slabs in the Greenland Ice Sheet using Sentinel-1 HH and HV C-band radar backscatter data. The paper is interesting and seems to show some promise for the method. The authors provide an appropriate degree of assessment indicating the regions of most uncertainty. I believe the paper can be published after some revisions, as follows:

Thank you for the thoughtful review!

[R1] *Why was 500 m resolution used? The authors never discuss this. Finer resolution would be of interest. Was there a reason it was not pursued?*

Since we use a backscatter threshold method for detecting ice slabs, it is critical to our results that the mosaics primarily reflect spatial variations in backscatter due to surface properties, rather than speckle, look angle, or temporal variations. Multi-looking the data to 500 m resolution gives us the necessary speckle reduction and significantly improves the linear correlation between backscatter and incidence angle as a result, which is also critical to achieving a good incidence angle correction. We conducted a sensitivity test over a region of Southwest Greenland spanning from the divide to the coast and found that 500 m resolution is a good balance between high resolution and ensuring reasonably good linear correlation between the backscatter and incidence angle. These results are shown in the figure below.



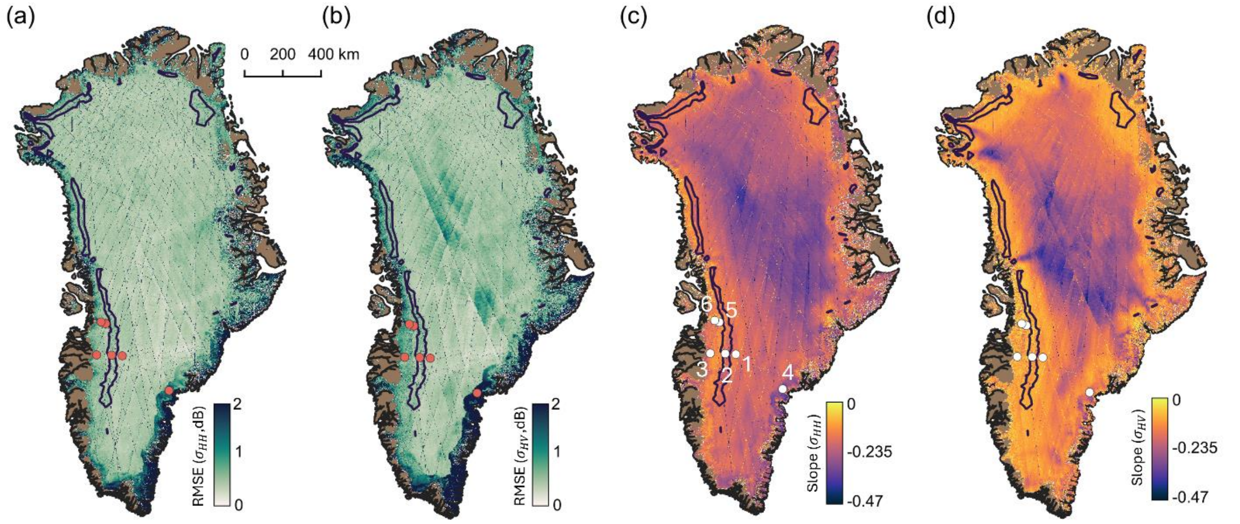
In the revised manuscript, we will discuss this point more clearly and include the above figure as supplementary information.

[R2] *The authors discuss their reasoning for using only 1 year of data in the paper's conclusions. I can see their points, but the paper would be more impactful if a multi-year study were performed. I recommend at a minimum that the authors describe their reason for using only a single year earlier in the paper.*

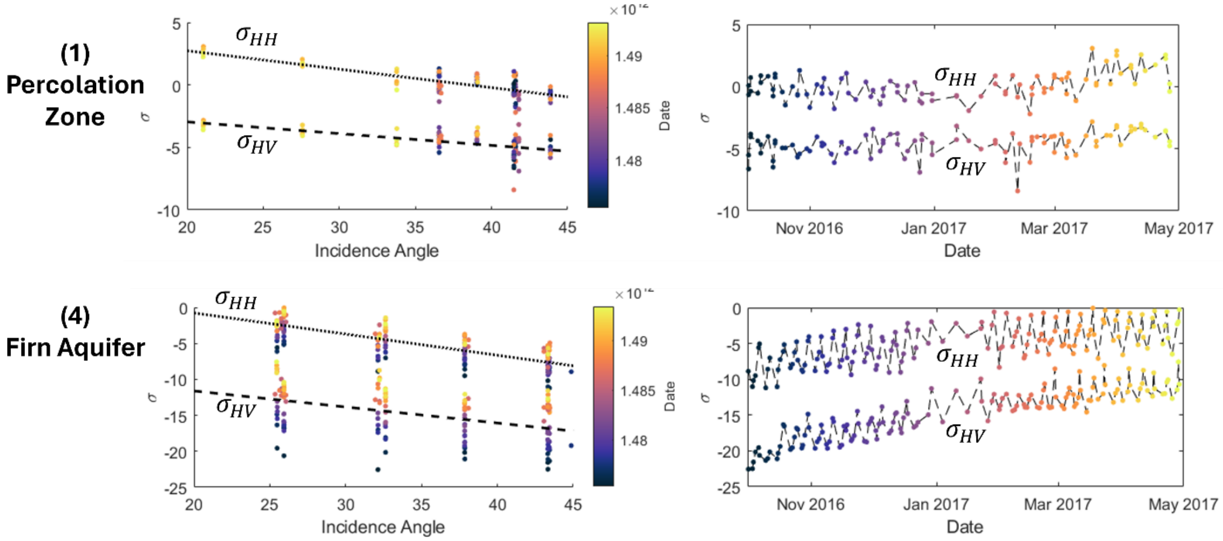
We will move this discussion to Section 3.1 when we first introduce the dataset. We do feel strongly that this is already a long paper (particularly once we add some of the sensitivity analyses introduced during review) and that it makes the most sense to conduct a follow-up study using multiple years of data once this algorithm has been established and published.

[R3] *Bottom of p. 2: should say radiometer not radar data. 2nd paragraph of p. 3: dielectric misspelled*
Corrected in text.

[R4] *Could the authors provide more information on the residual errors after the angle correction is performed, e.g. a plot or two of the data before and after angle correction? I'm assuming there is a lot of residual here due to the highly variable topography, etc. which makes such effects also a potential source of error that could be included in later discussions.*
We will add the figure below to supplementary information discussing the residuals from the linear incidence angle correction. In general, we find that when we multilook to 500 m resolution before calculating the linear regression, the residuals are quite reasonable (mostly less than 2B – see panels a and b below). We also show in a detailed sensitivity analysis for Reviewer #1 (see that response for details) that the ice slab detection results are insensitive to the number of observations per pixel as long at it exceeds ~117 observations and insensitive to the range of incidence angles if there are measurements from at least 10 different incidence angles per pixel. When using the full data set from 2016-10-01 to 2017-04-30 as originally presented in the paper, these conditions are met for our area of interest.

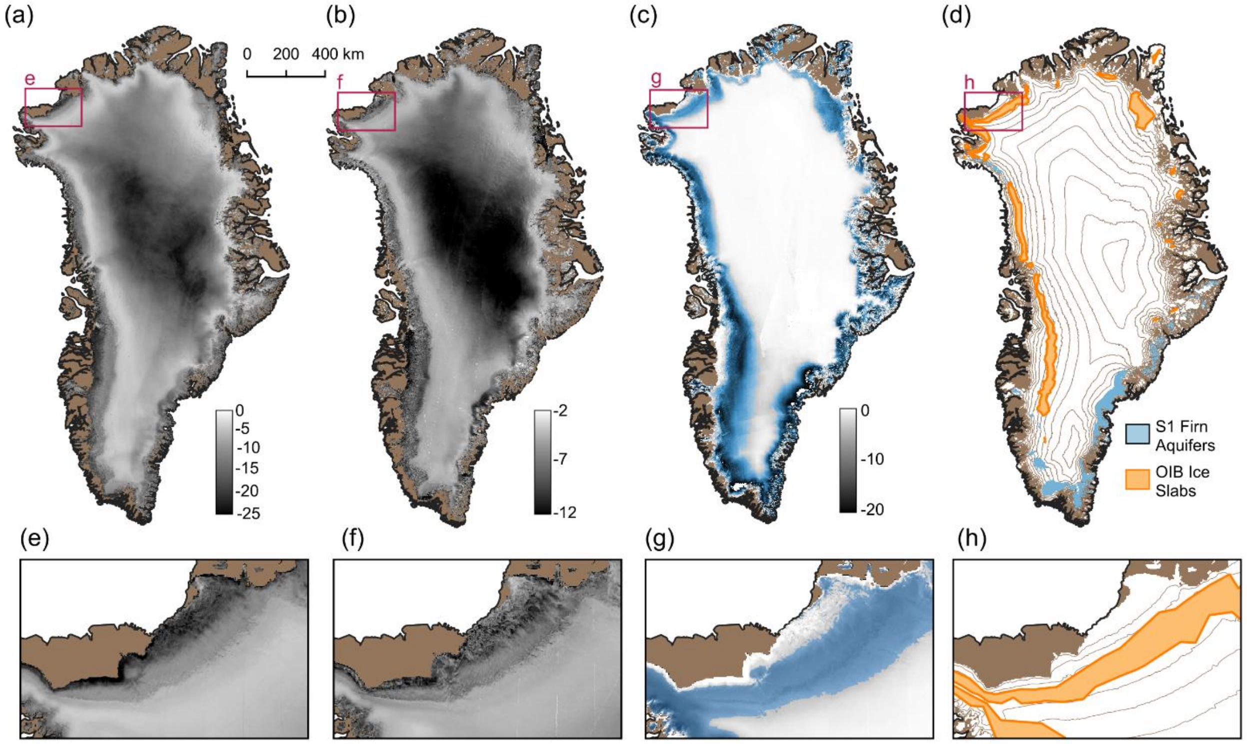

Residuals and Slope – Linear Incidence Angle Correction

The regions with the largest residuals, such as firn aquifer areas on the east coast (site #4 in the above image) have large residuals because of large temporal variations in backscatter, rather than obvious failures to follow a linear trend. The figure below shows, on the left, scatter plots of incidence angle vs. backscatter for both polarizations and on the right, the time series of backscatter. The top two panels show a firn aquifer site (#4) compared with percolation zone site (#1) to illustrate the impact of time-varying backscatter. This makes the important point that even in regions with relatively stable winter backscatter time series, our incidence angle correction serves a second purpose – it averages the backscatter in time to achieve an estimate of the mean winter backscatter that does not reflect small-scale temporal variations due to snowfall events, etc. In this way, some amount of residual expected and reasonable, since it quantifies the temporal variability in backscatter that we average out.



We will include these figures and discussion in supplementary information in the revised paper.

[R5] *Figure 2 could be more appealing if zoomed somehow in the manner of Figure 6. As is, we mostly see the interior ice sheet regions that are not of interest in this study. Perhaps rotate 90 degrees and separate into rotated images of the East and West coasts of the ice sheet?*
We think that there is still value in showing the entire mosaic, in part because it emphasizes the non-uniqueness of the ice slab radiometric signature and the need to filter out the firn aquifer and dry snow zone regions. We also attempted to rearrange this figure as suggested, but the figure quickly becomes extremely large with many panels and the need for additional locator maps to explain where each panel was located on the ice sheet. Given that the full datasets will be published with the paper, we also think that the interested reader can easily download and zoom around on the full mosaics themselves. However, as a compromise, we have now included zoom-in panels over a selected region over the North Greenland ice slabs in Figure 2 to show a clearer example of spatial variations in backscatter in our regions of interest. The revised figure is shown below:



[R6] *Was the same angle correction used for the Summer data? If so, this should also be noted as a potential (minor) issue.*

Yes, the same incidence angle correction method is applied to the summer data. We will note this explicitly in the revised text.

[R7] *The apparent alignment of the contours in Figure 3 with a "45 degree line" suggests that a classifier based on sigma_HH + alpha sigma_HV where alpha is some constant might also be successful here, rather that the separate thresholds on each quantity?*

This certainly might work! However, we think we would still need to optimize two separate quantities – the α in the summation suggested above and some β threshold for the summation image that delineates between ice slabs and not ice slabs. For this reason, it does not seem that it would significantly simplify the algorithm, so we are unlikely to pursue this method in revising this paper.

[R8] *Figure 5 appears somewhat redundant given that everything has been described in the text by this point.*

We think the visual presentation of the algorithm might still be helpful to some readers, but if the editor is concerned about the length of the paper, this will be the first thing we cut.

[R9] *I found the distinction between the "minimum", "maximum", etc. classifiers somewhat confusing. Please introduce these ideas more clearly early on in the discussion of the classifier.*

We have thoroughly revised Section 3.3 in the manuscript in response to this comment and suggestions from other reviewers. The revised text is shown below in the blue text, along with a revised Figure 4.

Section 3.3 Threshold Optimization and Uncertainty Analysis

Sections 3.1 and 3.2 describe how we form σ^0_{xpol} and σ^0_{HV} mosaics over high-melt regions where ice slab formation might be possible. To then map ice slab extent, we need to choose backscatter thresholds that can delineate regions with ice slabs from regions without ice slabs. We also wish to assess uncertainty by quantifying the range of plausible S-1 inferred ice slab extents that would be consistent with the OIB airborne ice-penetrating radar observations. We approach this problem in two steps. First, we use all available OIB ice slab detections to find the optimal backscatter thresholds that produce the best ice-sheet-wide agreement between the S-1 inferred ice slab extent and the OIB ice slab extent. By applying these optimal thresholds to the backscatter mosaics, we produce a map of the most likely ice slab extent across the ice sheet. Then, to assess uncertainty, we use a 10-fold cross validation scheme where we generate 10 new sets of thresholds, each optimized using only a small subset of the OIB data. From the results of these ten trials, we use the backscatter thresholds that produce the largest total ice slab area to define the maximum plausible ice slab extent, and the thresholds that produce smallest total ice slab area to define the minimum plausible ice slab extent. Together, this quantifies the range of plausible S-1 inferred ice slab extents that are still a good fit to the OIB observations. Below, we describe how we optimize these thresholds in detail.

3.3.1 Most Likely Ice Slab Extent

We use a training data set built from the Jullien et al. (2023) high-end estimate of ice slab extent derived from OIB flight lines surveyed in March-May 2017. (This high-end estimate corresponds to the maximum likely refrozen ice content given the observed ice-penetrating radar signal strength.) For each flight line that passes through an ice slab area, we extract the portion of the flight line that overflies the ice slabs, as well as an additional 50 km buffer that extends inland of the upper limit of the ice slabs. We discretize these lines into points every 50 m and assign each point a value of 1 if an ice slab was detected in the OIB data at that location or 0 if no ice slab was detected. These observations are then used to optimize the backscatter thresholds.

We use a brute force search to find optimal values of α and β that maximize the agreement between the upper elevation limit of the ice slabs as detected by airborne ice-penetrating radar, and the upper limit of the ice slabs as estimated by S-1. Areas where $\sigma^0_{HV} < \alpha$ and $\sigma^0_{xpol} < \beta$ are taken to be ice slabs. We then test all combinations of thresholds for -7.12 dB $< \beta <$ -2.37 dB and -13.6 dB $< \alpha <$-2.1 dB, calculate the F1 score for each combination, and choose the threshold values that give the highest F1 score. The F1 score is a measure of the accuracy of a binary classification and is calculated following Equation (1).

$$F1 = \frac{2 * true\ positive}{2 * true\ positive + false\ positive + false\ negative}$$

Figure 3 shows this optimization trade space with the optimal threshold combination shown in the white dot. We find that using both σ^0_{xpol} and σ^0_{HV} thresholds together leads to modestly better agreement with the OIB detections, compared to using only σ^0_{xpol}. When only σ^0_{xpol} is used to delineate the upper elevation limit of the ice slabs, this F1 score is 0.787, compared to an F1 score of 0.811 when both backscatter thresholds are used. When using only σ^0_{HV} to delineate the upper elevation limit of the ice slabs, the F1 score is only 0.674, so it is clear that σ_{xpol}^{0} provides additional information that significantly improves the delineation of the upper boundary.

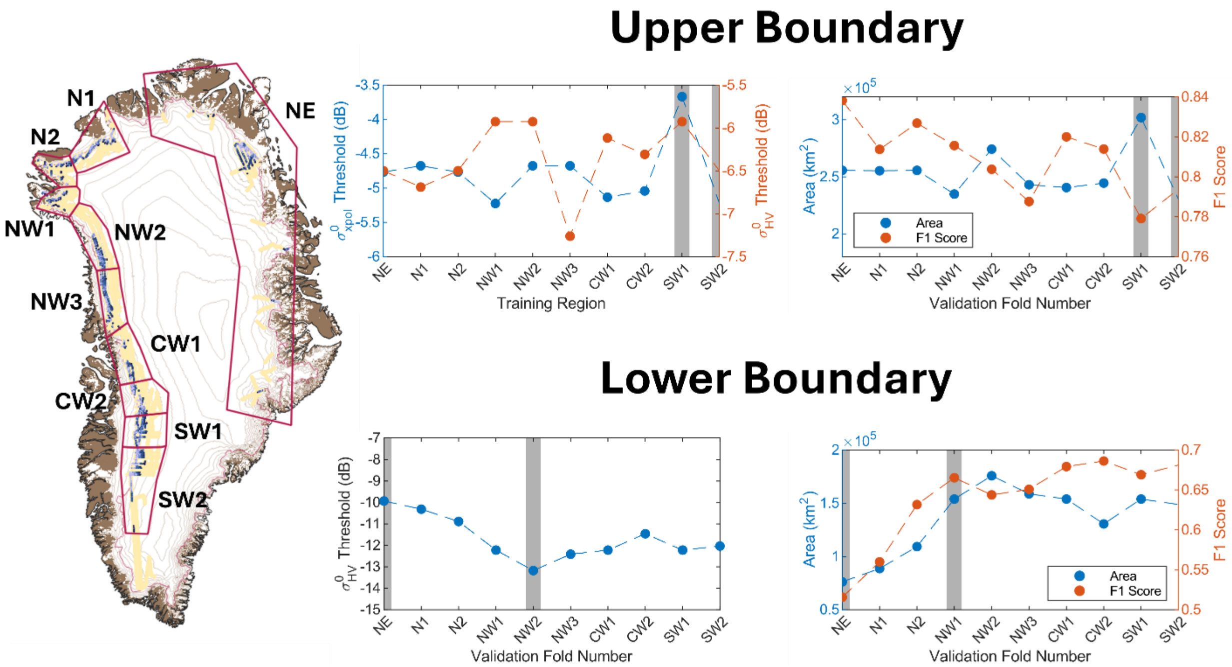
Initial analysis of the backscatter mosaics suggests that σ_{xpol}^{0} does not display any distinct change in behavior associated with the lower boundary (see Figure 1a), so we optimize a separate threshold, $\sigma_{HV}^{0} > \phi$, to delineate the lower elevation limit of the ice slabs. We optimize ϕ following the same method as described above, but using a new version of the OIB training dataset that covers the ice slab region and a 50 km buffer down-flow into the ablation zone. Altogether, the area defined by $\sigma_{xpol}^{0} < \beta$ and $\phi < \sigma_{HV}^{0} < \alpha$ is our most likely estimate of the spatial extent of ice slabs across the ice sheet.

3.3.2 Maximum and Minimum Ice Slab Extent

To quantify uncertainty in this most likely estimate of ice slab extent, we use a 10-fold cross-validation scheme. We divide our training dataset into 10 subsets, each containing OIB ice slab detections from a different region of the ice sheet (see Figure 4). For each of the ten regions, we use a brute force search to find the values of α, β, and ϕ that produce the best agreement between OIB ice slab detections and S-1 inferred ice slab extent in that region. We then apply those local thresholds to the entire ice sheet and calculate the F1 score by comparing the S-1 ice slab mapping to the ~90% of the OIB observations that were not used to choose α, β, and ϕ in that trial. As with the most likely ice slab extent, we calculate separate F1 scores for the upper and lower limits of the ice slabs. From the results of these ten trials, we use the backscatter thresholds that produce the largest total ice slab area to define the maximum plausible ice slab extent, and the thresholds that produce smallest total ice slab area to define the minimum plausible ice slab extent.

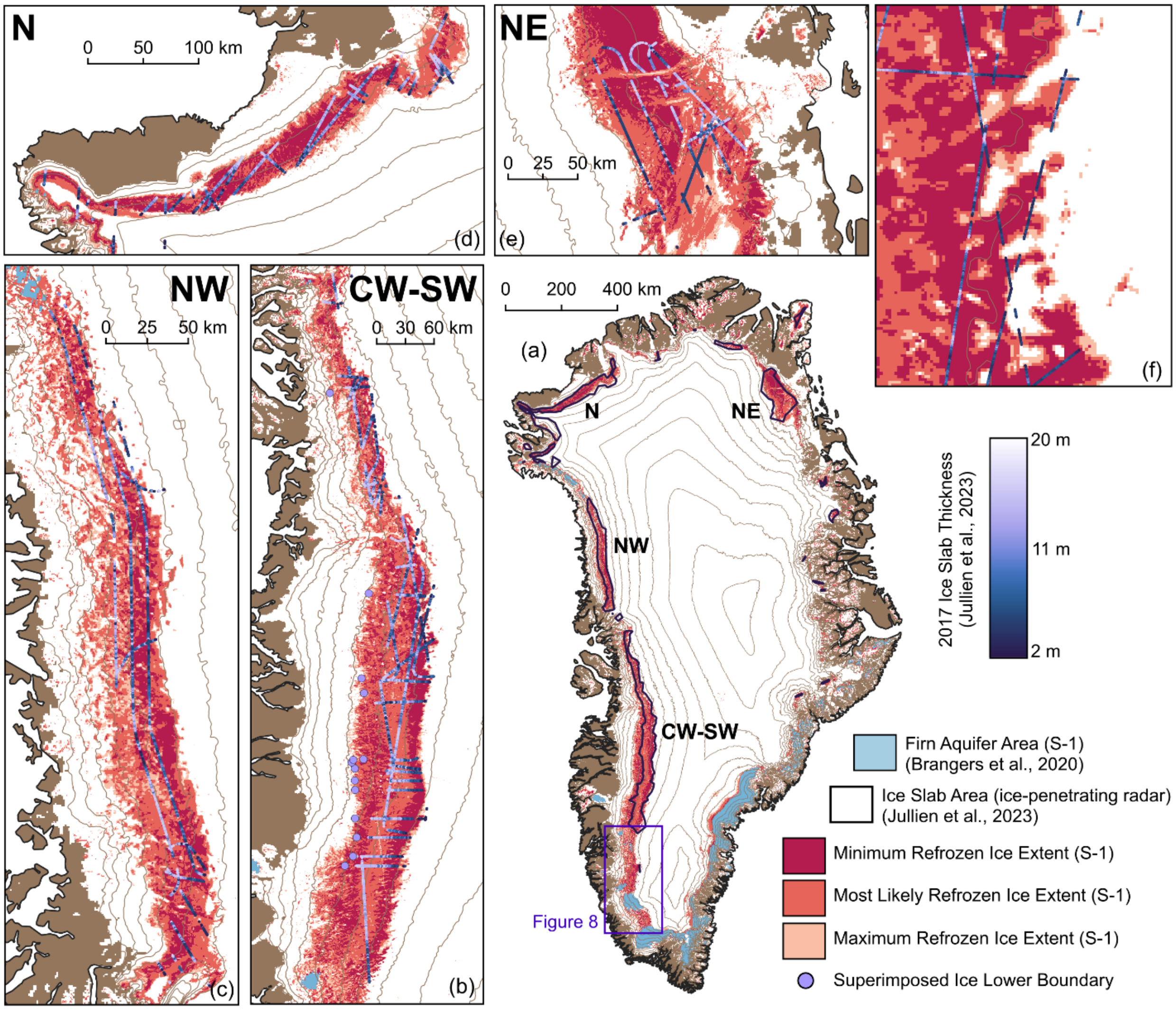
Figure 4 shows the results of this cross-validation. We find that across the 10 validation trials, F1 scores for the upper elevation limit of the ice slabs vary from 0.78-0.84, with no clear spatial trend. Since the F1 score for the most likely ice slab extent is 0.811, this suggest that values of α and β chosen based on data from one region of the ice sheet generalize well to other regions. Indeed, these thresholds vary by only $\sim \pm 1\ dB$ across all regions of the ice sheet. Therefore, the algorithm is reasonably spatially robust.

We do find a clear spatial trend in the generalizability of ϕ between regions. In particular, when ϕ is derived only using data from regions NE and N1, the resulting S-1 inferred ice slab extent in Northwest and Southwest Greenland agrees poorly with the OIB observations. However, conversely, the value of ϕ estimated using only data from the Northwest and Southwest does apply well to the North and Northeast. We suggest several reasons for this behavior. First, the North and Northeast regions have the least number of ice slab detections, so thresholds derived from data in those regions may be overfit to conditions that are not representative of larger areas. Second, we see steeper gradients in backscatter as a function of elevation in the North compared to the Northwest and Southwest. This suggests that small variations in ϕ would lead to large changes in ice slab area in the Northwest and Southwest, but small changes in ice slab area in the North and Northeast. As a result, the agreement between the OIB observations and S-1 detections is much more sensitive to errors in ϕ in the Northwest and Southwest than in the North and Northeast.

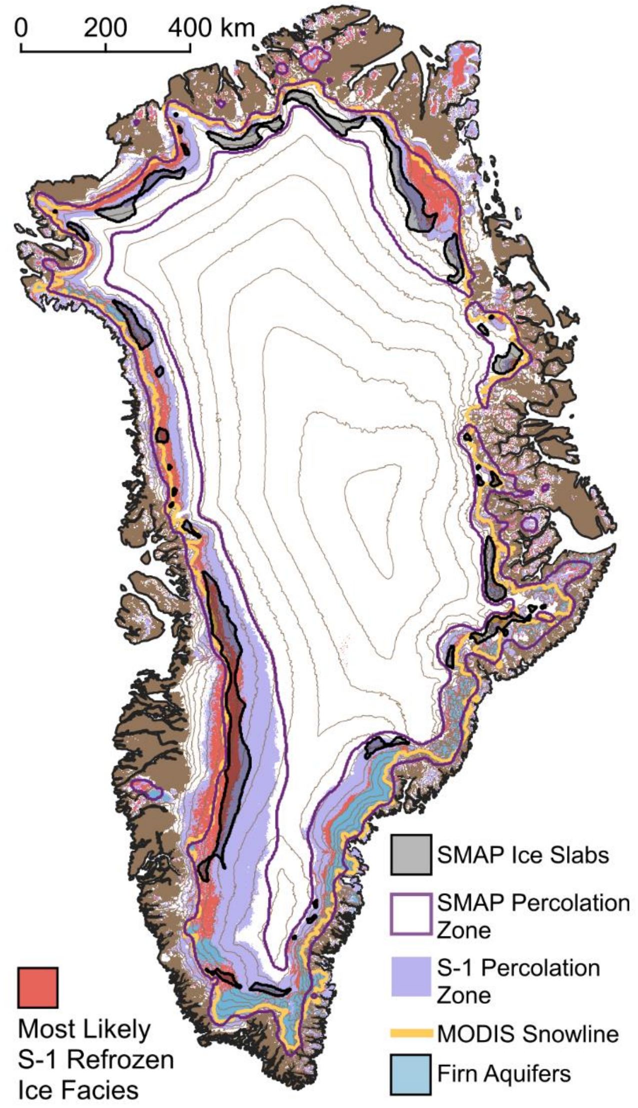

Upper Boundary
Lower Boundary
N1
N2
NE
NW1
NW2
NW3
CW1
CW2
SW1
SW2
σ^0_{xpol} Threshold (dB)
σ^0_{HV} Threshold (dB)
Training Region
Area (km^2)
F1 Score
$\times 10^5$
Area
F1 Score
Validation Fold Number
NE N1 N2 NW1 NW2 NW3 CW1 CW2 SW1 SW2

[R10] *Dashed lines in Figures 6 and 11 are hard to discern especially when the topographic lines are included. Is there a way of doing this that makes them more clear? This comparison is key so making it easy to follow is crucial.*

In both Figure 6 and 11, we have changed the dashed lines to solid lines, which we believe makes the outlines clearer and easier to interpret. The new versions of the figures are shown below:



Revised Figure 6. Dashed lines in Panel A have been changed to solid line.



Revised Figure 11. All dashed lines have been changed to solid lines.

[R11] *Some of the discussions of differences seemed a little long and overly complex. Consider trying to simplify these discussions if possible, i.e. simply state "may have more volume scattering compared to xxx" etc.*

As we revise the paper, we will work to make the discussion more concise.