Response to review from Tate Meehan

Within, "Mapping the vertical heterogeneity of Greenland's firn from 2011-2019 using airborne radar and laser altimetry" the authors devise a radar surface echo retrieval for shallow firn facies classification. The manuscript is well developed and utilizes modeling and in-situ data to improve radar signal interpretation. Much thought is given to the electromagnetic modeling set-up regarding the presence of snow cover on ablated ice and ice layers within the percolation zone of the firn. Modeling results provide a compelling argument for the divergence of radar surface echo waveforms and the misclassification of surface picks based solely on maximum amplitude. When the surface echo retrieval is applied ice sheet wide, interesting and confirmatory information is discovered, such as the replenishment of homogeneous firn within the percolation zone after the significant 2012 melt event. Results from this radar retrieval approximately confirm facies boundaries established by reanalysis model results, but I am left wanting a bit more analysis on how closely the radar and reanalysis resemble each other in a spatiotemporal context, as well as some discussion as to how this new radar information can inform or validate firn modeling. Those points aside, this manuscript is well-written, complete from an investigative standpoint, and deserving of publication. I recommend minor revisions which I have annotated in the attached .pdf.

Thank you much for your time and effort reviewing our manuscript! We really appreciate the detailed and constructive suggestions, and we agree with all comments and have implemented them in the manuscript. We believe that these changes further improved the manuscript.

Please find answers to all comments below, using the following color code:

Blue indicates comments from the authors

Green italic indicates changes in the manuscript, with revised text indicated marked as underscored

Major Comments:

Throughout the analysis, I gained a more intimate understanding of how the VHF radar signal interacts with the near-surface, kudos to the authors for this. The efforts put forth in the modeling suggest that a given ice layer in the percolation zone must be buried greater than 8 m as to not interfere with the wave form surface echo. This is approximately 10 years' worth of snow accumulation. However, empirically it was determined and the authors state that "between 2013 and 2014, areas of heterogeneous firn in the dry-snow zone reverted back to homogeneous firn." This disagreement between the modelled information and radar retrieval is not reconciled within the manuscript discussion. Explanation for this phenomenon should be provided.

Thank you for catching this! Together with comments from the other reviewer, we realized that i) the theoretical range resolution we calculated for MCoRDS was too large (due to a wrong k_t value), and ii) the theoretical range resolution for the modeled waveform is likely larger than for the actually measured MCoRDS signal. We re-calculated the theoretical range resolution in firn for the measured MCoRDS data to 5.7 m (instead of 7.5 m), and the range resolution of the modeled signal to ~11 m (calculated using the peak withd at a 3dB drop from the maximum peak, similar to as the MCoRDS windowing factor for the range resolution is calculated). We attribute the difference between the modeled and observed signal range resolution to the use of different windowing parameters applied in the signal processing, and potentially also due to slightly different input signals.

To explain this, we added the following text to the manuscript: "*We note that the vertical range resolution of the modelled signal, expressed as the pulse width 3 dB down from the peak (similar to the calculation of the MCoRDS windowing factor, https://data.cresis.ku.edu/data/rds/rds_readme.pdf), is ~8.4 m in ice and ~11 m in firn, The larger range resolution is likely due different windowing factors applied in the pulse compression (here we use the standard RadSPy Hanning window), and the input pulse not being an exact replica of the MCoRDS pulse. However, we expect the general behaviour of the modelled signal to be representative of the MCoRDS signal, where depths of layer interfaces should be considered in a relative sense to the different theoretical range resolutions.*"

With this knowledge, we updated Figure 4 (see below) to represent the depth and thickness of ice layers in firn also as a function of the range resolution. From this, it becomes evident that ice layers affect the surface return when in the top 4 m for the observed MCoRDS signal (0.7xRange Resolution), and not the previously stated 8 m. (A similar secondary x-axis was applied to Figure 3, effects of snow depth, to express the thickness as function of range resolution).

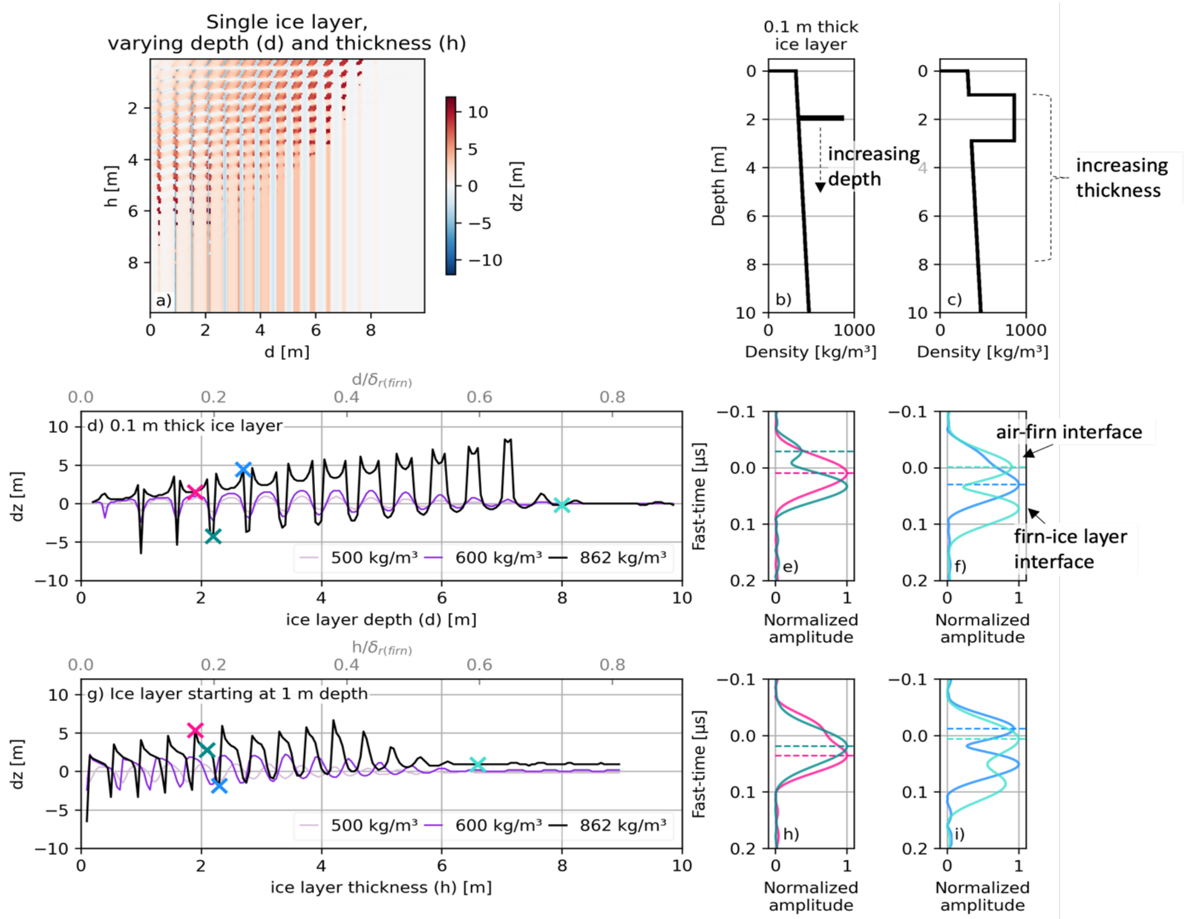

Figure 1: a) Surface peak offset (dz) for RadSPy simulated MCoRDS surface returns over firn stratigraphies consisting of a single ice layer placed at various depths and with different thicknesses. b) and c) show example model input profiles. d) dz for a 0.1 m thick ice layer, as well as layers with densities of 500 kg/m³ and 600 kg/m³ at different depths, with waveforms for selected ice layer depths (marked with crosses) shown in e and f. g) dz for an ice layer, as well as layers with densities of 500 kg/m³ and 600 kg/m³, starting at 1 m depth and with increasing layer thickness, with waveforms for selected ice layer thicknesses (marked with crosses) show in h and i. The dotted lines in the waveform plots show the picked peak identified used to calculate dz. The x-axis on the top of panels d) and g) show the ice layer depth and thickness scaled by the rang resolution δ_r of the modelled signal in firn (~11 m).

We changed the text in the results section of the manuscript accordingly:

"When an ice layer lies at depths greater than ~0.7δ_r, thus ~4 m for the MCoRDS signal in firn, the radar's surface return displays two distinct peaks…. Thus, ice layers located deeper than ~0.7δ_r have negligible impact on the surface reflection peaks, leading to dz ≈ 0."

This now matches better with observations where heterogeneous firn switches from 2013 to homogeneous firn in 2014, where burying the 2012 melt layer to 4 m depth between 2012-2014 is more reasonable. We also note that the change in *dz* only represents a "bulk" near-surface heterogeneity, and not a heterogeneity at a specific depth. For example, *dz* could have been most sensitive to already deeper layers, which then were pushed out of the sensitive depth (~4-5.7 m) in the following year. Thus, not the entire 4-5.7 m firn column needs to be replaced with homogeneous firn. To clarify this, we added the following sentence in the discussion section:

"We also note that dz represents a bulk near-surface heterogeneity over the MCoRDS surface return sensitive depth and cannot be used to identify heterogeneity/ice layers at a specific depth."

The correct identification of the "true" radar surface required an algorithm which upsamples the MCoRDS data to finer fast-time resolution. The author's choice of resampling algorithm has a significant effect on the outcome of their results. If for example a piecewise linear upsampling was applied, the original MCoRDS signal and the upsampled signal would appear identical. Using a Fourier method as you have described, introduces "Gibb's Phenomenon" which can be seen as the oscillations at the tail end of the surface reflection signal. I suspect also that such phenomena are occurring to produce the troughs in the signal seen in Figure 6d & e (i.e., the Fourier series is struggling to represent the flat discontinuity of the signal between samples 3&4 of 6d). I appreciate the approach you have taken, and it is fortuitous that the upsampled signal reasonably recreates the modelled signal. However, acknowledging the origins of this "spurious" waveform bulge, be that an effect of the resampling algorithm or your choice in sample lags, should be considered and described more thoroughly.

Thank you for this comment. We would like to note that the trough in the signal is not "spurious", but is a real signal (as suggested from the modeling), which we attribute to the interaction between the radar signal and the heterogeneous subsurface. We argue that the observed MCoRDS signal simply does no show such troughs because of the lower sampling rate, therefore not capturing the complete signal. Thus, we intentionally upsample the signal with a Fourier transform introducing the Gibbs Phenomenon, allowing us to recover the first return (i.e. getting a peak and trough before the maximum peak).

We tried to make this more clear in the manuscript by adding: "*Therefore, using an upsampling approach based on a Fourier method that introduces oscillations that recovers closely spaced waveforms is a reasonable strategy for picking the first return. However, we note that depending on the relative position of the signal peak, this phenomenon can cause some uncertainties in dz.*"

We also added a sentence in the discussion section outlining that changes in dz could occur from the location of the signal peak relative to the waveform sampling:

"While our modelling and observations generally align, some discrepancies exist. For instance, modelling predicts dz = 0 over the ablation- and dry-snow zones, yet our observations also include non-zero (and negative) values in these areas. Such deviations could stem from factors not accounted for in the model, such as the surface slope (given that the laser is nadir looking, versus the radar records the nearest return, which may result in negative dz values over sloping surfaces) surface roughness and surface anomalies like crevasses. Additionally, uncertainties in dz may arise from how the MCoRDS signal is sampled, …."

Explanations of why dz values near the summit of Greenland are significantly higher than the surrounding dry-snow zone data retrievals remains unsubstantiated. Answers to these concerns are not supported in Scanlan et al. (2023), which pertain to higher frequency radar systems and show both consistently low retrieved density and high density within the same region of Greenland's summit. Through tighter analysis, or a fleshed-out hypothesis, effort should be considered to reconcile the work of these authors.

We agree that the discussion on these anomalies could have been a bit more thorough. We changed this paragraph to include the possibility of small density contrasts from thin layers and surface roughness. Also changed the wording to make the connection to Scanlan et al (2023) - we mean to point out the similarity of these unexpected observations rather than explaining our results with the high-densities found in Scanlan et al (2023).

"Our results also reveal areas with complex surface processes, for example, localised high dz areas in the dry-snow zone, particularly near the summit of the ice sheet. It is possible that these areas consist of a high small-scale density variability in the near-surface (e.g. from wind scour or hoar formation) (Hörhold et al., 2011), causing strong enough density contrasts to affect the radar surface peak. Additionally, increased dz may be caused by increased surface roughness (e.g. from sastrugi), which introduces waveform interferences when the radar signal from facets with different heights when the signal is coherently integrated during processing. We note that this observation aligns with unexpectedly high surface densities in the summit region *reported by Scanlan et al., (2023). The persistence of this anomalous feature over several years is unclear, but may be attributed to the relatively low snowfall in the area, implying that the near-surface varies only over long timescales."*

Minor Comments: We addressed the minor comments from the PDF as follows:

L15: *"In this study, we use concurrent VHF airborne radar (MCoRDS, 195 MHz) radar and laser altimetry (ATM)…"*

L16: *"…to investigate our hypothesis that heterogeneities in firn (i.e. ice layers) cause vertical offsets in the radar surface reflection (dz).*

L18: *"… effectively delineates between vertically homogeneous and vertically heterogeneous firn profiles over a depth range of ~4 m"*

L25: Changed *survived* to *endured*

L30: Done

L32: Done

L95: We added the typical width of the surface return (using the examples from Figure 2): *"… (hereby referred to as the broader surface signal of elevated amplitudes, typically 0.1-0.3 µs wide, and can encompass multiple peaks)…"*

L125: Done

L138: Done

L139-142: We agree that this section includes some observations from the modeling results section, but we used these to help calibrate the radar peak offsets. We therefore keep this part here to help explain how we identified systematic offsets. However, we added a sentence about potential remaining timing issues in the radar measurements to the discussion section:

"Such deviations could stem from factors not accounted for in the model, such as the surface slope (given that the laser is nadir looking, versus the radar records the nearest return, which may result in negative dz values over sloping surfaces) surface roughness and surface anomalies like crevasses. Additionally, uncertainties in dz may arise from how the MCoRDS signal is sampled, and timing issues (i.e. cable delays) in the radar measurements that were not fully identified and corrected in our calibration."

L142: Done

L143: We split the sentences and change it to: *"The calibration zone started 20 km East from the western ice margin to avoid the steepest and most crevassed part of the ice, and extended to…"*

L145: Done

L162: Done

L164-165: We changed the sentence to: *"Additionally, we added a snow layer of varying thickness (up to 2 m) to the ablation zone model (three-layer model), using a density of 341 kg/m^3…"*

L256: Done

L273-275: We agree that there is a bit of methods in this paragraph, but we believe it is easier to follow if this stays in the "Effects of waveform sampling" section and therefore leave it here.

L287: addressed above

L290: This is explained in the first paragraph of this section. To help make this connection, we replace the term "*lag*" with "*varying time-zero offsets*" to be consistent.

L297: That is a good point with the absolute values, and we changed it to: *"Specifically, low absolute dz values (<1m)…"*

L375: Done

L385: Done

424: Curious as to what the correlation between dz and MAR melt is. Correlation would likely break down in the ablation zone, but it would be interesting/confirmatory data in the percolation and dry zones.

We agree that it could be interesting to do a correlation between dz and MAR melt. However, we do not include such an analysis here as we believe the manuscript contains enough information to show the spatial correlation between dz and the percolation zone, which was derived from the mean MAR melt.

L471: What is the proposed cause of higher firn densities near the summit?

We agree that the discussion on these anomalies could have been a bit more thorough. We changed this paragraph to include the possibility of thin higher-density layers or surface roughness.

"Our results also reveal areas with complex surface processes, for example, localised high dz areas in the dry-snow zone, particularly near the summit of the ice sheet. It is possible that these areas consist of a high small-scale density variability in the near-surface (e.g. from wind scour or hoar formation) (Hörhold et al., 2011)*, causing strong enough density contrasts to affect the radar surface peak. Additionally, increased dz may be caused by increased surface roughness (e.g. from sastrugi), which introduces waveform interferences when the radar signal from facets with different heights when the signal is coherently integrated during processing. We note that this observation aligns with unexpectedly high surface densities* in the summit

region *reported by Scanlan et al., (2023). The persistence of this anomalous feature over several years is unclear, but may be attributed to the relatively low snowfall in the area, implying that the near-surface varies only over long timescales."*

L498: We agree that giving a vertical depth is useful – following the modeling results, we use the depth of 0.7 x the theoretical vertical range resolution as a maximum depth, and change this in the manuscript to: "*Our results, supported by modelling and in-situ firn core analyses, demonstrate that dz serves as an effective tool to delineate between vertically homogeneous and heterogeneous firn profiles ~0-4 m depth, where temporal changes in dz over 2-5 years align well with known climatic events."*

L503: Done

L505: Done

References added:

Hörhold, M. W., Kipfstuhl, S., Wilhelms, F., Freitag, J., and Frenzel, A.: The densification of layered polar firn, Journal of Geophysical Research: Earth Surface, 116, https://doi.org/10.1029/2009JF001630, 2011.