

Reply to referee #2:

We highly appreciate comments raised by the referee and present a point-to-point reply for all issues raised by the referee. For an improved readability and to facilitate direct response, we sometimes subdivided comments into several paragraphs referring to similar issues

This manuscript describes the deployment of an upward-looking ground-penetrating radar (upGPR) in the Greenland firn. Using the upGPR, the authors observe percolation of water into the snowpack, and movement of water within the firn, measuring percolation depths and deriving parameters such as bulk density and volumetric water fraction. The authors compare the measurements of these percolation parameters with the outputs of a popular regional climate model, MAR, as forced by two different reanalysis datasets. In general this is an interesting manuscript that advances knowledge in the field, and is worthy of publication. The science is good with no serious flaws. The presentation is a bit lacking, primarily in the use of English. I've commented extensively (though probably not exhaustively) in the 'minor comments' section below.

We thank the referee for the evaluation and the positive assessment.

2 Major Comments

- Page 3, Lines 16-20... Seems to me like you can either use MAR to help determine how the radar does, or use the radar to determine how MAR does, but not both. At all.

This is correct and we apologize for the previous version; now changed to “To estimate the reliability of radar-derived parameters, we compare determined percolation depths with changes in temperature records and analyze monitored changes in thickness of the snow and firn column above the antennas with results of ultrasonic transducers located within a distance of less than 2 km (Steffen et al., 1996; MacFerrin et al., 2015).”

- Page 4, Lines 31-33... Not convinced that the 'compaction of the radar and the target' are actually equal; in practice the target supports may be driven down faster than the radar. Effect is likely small but will be non-zero.

Here, we respectfully disagree that this is the case given the current level of accuracy. Such an unequal compaction would most likely be gradually and would have effects on the radar derived snow/ firn height data. If the compaction of the posts holding the radar target would be faster, d_A would decrease and since d is taken as being constant L_s would gradually increase. The results should show up as systematic gradual overestimation of the radar derived L_s . This is obviously not the case in Fig. 4. However, in the manuscript, we state “to be approximately equal”.

- In general, much of the Results section appears to be more like Discussion; I have not attempted here to disentangle the discussion-like parts from the straight results, but this is a part of the paper that could use refinement.

We sincerely apologize for the sloppiness and checked carefully for appearances of discussions in the Results section. Among other parts previous P9 L4-6, P12 L5-8 as well as P14 L5-8 are now moved to the discussion part or changed to meet requirements for a results section. Please see the respective marked parts in the discussion section.

3 Minor Comments

- Abstract Line 2: 'act as meltwater' -> 'act as a meltwater'

- Abstract Line 9: 'capable to monitor quasi-continuously' -> 'capable of continuously monitoring' (though use quasi if you must)

- Page 2, Line 5: 'multiplied by a factor' -> 'has increased by a factor'

All have been changed accordingly.

- Page 2, Lines 7-8: sentence needs rewording; should start '61% of the recent mass loss is ascribed to...'

The sentence has been revised to: "van den Broeke et al. (2016) attributed 61% of the recent mass loss to a decrease in surface mass balances and only 39% to an increase in solid ice discharge."

- Page 2, Line 19: 'are exposed' -> 'have been exposed'

Changed

- Page 2, Line 23: Reference to Abdalati and Steffen 2001- this is a remote-sensing paper, not looking at in-situ data.

Changed to: "Information on melt usually just exist for the area extent of surficial melt over the GrIS (e.g., Abdalati and Steffen,2001) from remote sensing data."

- Page 2, Line 25: unless McFerrin et al is published by the time this comes to press, remove this reference.

We changed the reference to Machguth et al. (2016) and slightly modified to "massive ice lenses"

- Page 2, Line 29 'inevitable'- is this the right word? Sounds like you meant to say 'imperative'.

Changed to essential

- Page 3, Line 2: 'temporal continuous' -> 'temporally continuous'

- Page 5, Line 9: 'identification of timing' -> 'identification and timing'

Both changed accordingly

- Page 5, Line 11: reference to figure 3 before I see any reference to figure 2, the first reference I see to which is on page 6.

We switched the order of the figures.

- Page 6, Line 5: Seems 120 kg/m^3 is pretty low for Greenland snow (to me) but I don't have a reference to point to.

We dug numerous pits around Dye-2 in May 2016. For conditions with recent new snow, we found density values of 104 kg/m^3 , 52.8 kg/m^3 , 93.3 kg/m^3 and 61.8 kg/m^3 for the upper 5-10cm. The 120 kg/m^3 seems to be an appropriate estimate, especially while accounting for the progress of the season including warmer temperatures in summer. We are pretty confident that this value is not too low. We ask the reviewer to specify why the value should be higher at this particular location, e.g. by providing a reference.

- Page 8, Line 26: delete 'respectively' as it's not needed here.

Changed accordingly.

- Page 9, Line 4: I don't understand what the first part of this sentence is referencing; I don't see the change after 19 July that I think the statement is discussing.

We moved this paragraph into the discussion and modified to:

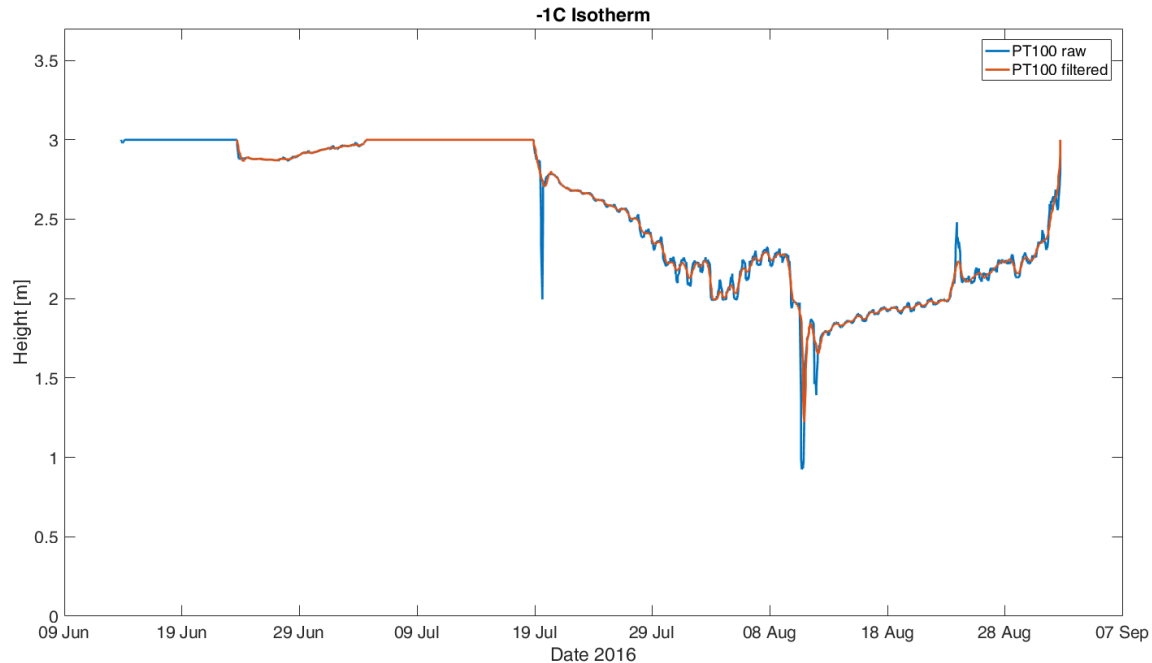
"The used wave speed model becomes incorrect when liquid water infiltrates snow and firn. Liquid water decelerates radar wave propagation significantly and, consequently, distance to reflections above the infiltration increase in measured TWT. However, snow pits and firn cores at the site can only be obtained

when the instruments are visited once a year. For data analysis only measured TWT is used and, consequently, presented heights are not relevant. Still we consider a presentation of heights, even though they are partly incorrect after certain time periods, as being more intuitive and more supportive for readability. Percolation depths are unaffected from erroneous TWT conversions since they indicate the maximum height of dry snow and firn. In contrast to the ice lens at about 2 m height, the 2015 horizon and the snow surface, all layers below the reference layer (Figure 2a and b) are basically unaffected by melt events and consequently do not show variations in TWT.”
We hope this will enhance comprehension for readers, which are not familiar with TWT domains.

- Page 9, Line 21: 'equal to' -> 'are'
Changed accordingly

- Page 9, Lines 30-31: *This description needs to be tightened up. The 'determined changes in snow and firn' are really 'extent of percolation'; the results you show in panel c are also 'changes in snow and firn'...*
Modified to: “In Figure 5, determined changes in snow and firn from upGPR (Figure 5a: extent of percolation and c: changes in SWE - brown line and volumetric liquid water content - blue line) are compared with temperature data derived from the installed thermistor string (Figure 5b).”

- Figure 5b: *Do you really believe the high-frequency variability of the isotherms (most prominent example is between 30 May and 19 Jun going from 1 to 0 m height)? I don't. These should probably be smoother curves. Filtering these data might be what you need to do before calculating isotherms. The other really major thing here is that because you discuss the -1 C isotherm so much in the manuscript, it should be delineated clearly here- a different color, or marked in some way. This way readers can see the trends you are describing on page 12 in time and depth of the -1 C isotherm.*
The referee is right that the oscillating signal of the isotherms is actually noise. Here, we are at the stated 0.25 K accuracy of the sensors. Those sensors were deployed in May 2016 and, consequently, required some time to settle and measure the environmental temperature. The noise in the signal is even much more prominent for the very first time period after installation (Fig. 5b). For the revised version, we smoothed the signal. Just for presentation, it is most likely better to present processed data, however, for analysis it is insufficient to smooth/ filter the data. Percolation depths are significantly altered by any kind of processing see figure below. Hence, data in Tab. 1 is derived from recorded raw data.



This figure demonstrates the effect of smoothed thermistor data. We used PT100 temperature sensors. The blue line displays the raw temperature signal interpolated to generate a -1°C isotherm and its location in height above the antennas. The orange line represents the same isotherm, while PT100 data has been smoothed using a quartic Savitzky-Golay filter.

- Page 11, Line 2: 'outlasted' -> 'lasted'

Changed accordingly

- Page 12, Line 9: Start a new paragraph here.

Changed accordingly.

- Page 12, line 23: *this assertion would be easier to verify if you plotted the 2015 summer surface on figure 5; that would show clearly the melt propagating below the summer 2015 surface (harder to verify quickly by looking at figures 6a and b as referenced).*

Here, we refer to figures 6a and b, which present quantification of seasonal mass transfers from snow into firn below summer surface 2015. Fig. 5 just present percolation depths. We decided to keep the reference and did not modify Fig. 5. The chosen presentation in Fig. 5a is horizontally filtered, which means that the relatively stable 2015 layer is filtered out. We believe that including a horizontal line in this figure would distract and confuse the reader, since the purpose of this figure is to demonstrate the agreement of thermistor data with radar recordings and derived parameters.

- Page 14, Lines 5-8: *Not sure why all of this is here. Why wouldn't you include the ice lenses that you observed in the pits? In that case simply report the last sentence on line 8. But it's not clear if these ice layers were measured in the pits or where they used from the radar to 'correct' the pit data? If so, not sure why it's valid to do that. Clarity of the language is required here.*

This is another paragraph, which we shifted into the discussion part. We revised it to enhance readability to: “Offsets in radar derived mass balance data (b_{15}) of about 100 kg/m^2 to manual observations can be attributed to difficulties in picking the reflection event seasonal snow above the summer horizon of 2015 in the radargram. Snow pits are usually just dug down to a remarkable crust, which is hardly penetrable with a shovel. The reflection response at this specific density gradient is masked by signal interferences with the reflection generated at the lower border of this crust, which represents the melt horizon of summer 2015. Correspondingly, including the observed ice lenses into SWE calculation of the pits results in a mass of 426 kg/m^2 for early May. This reduces the offset to values obtained from upGPR to only 2.8%.”

- Page 15, Line 2: *Rather than force the reader to look up the physical meaning of the Nash-Sutcliffe efficiency value (which one needs in order to evaluate your reported numbers), briefly define it here. Only needs a sentence.*

The sentence has been included: “NSE values of 1 indicate a perfect fit of the model with the data, while a NSE of 0 shows that the model fit is as good as simply the average value of the data. NSE for MAR NCEP1 simulations is at 0.75 and below 0 for ERA-Interim driven simulations for the whole data series.”

- Page 15, Line 8: *I think this kind of correction is probably fine- if you want to quote the statistics for it, you need to demonstrate how "removal of the strong increase" was done, and it would probably be good to show the resultant curve on figure 7.*

We agree that the description of this kind of correction was misleading. The respective sentences were modified to: “Hence, deleting only the data point of 11 May ($\Delta b_s = 35 \text{ kg/m}^2$) from analysis lead to NSE values for MAR-NCEP1 of 0.53 and MAR-ERA of 0.95.”

So, we simply removed this data point from NSE analysis. It is incorrect to plot an additional SWE curve in Fig. 7. Here, we present cumulative changes in SWE in respect to 1 May conditions. Certainly, snow accumulation happened for the 11 May. It just might not have been as large as measured by the radar due to the disturbed wind field from the shelter causing drifting snow at the position of the upGPR. However, to analyze similarity of two curves by NSE values, it is appropriate to remove single data points, if they are biased.

- Page 17, Figure 9: *I don't understand why the θ_w NCEP is illustrated the way it is. Seems like on previous figures there was a parallel track with MAR-NCEP and MAR-ERA, and this should continue in this figure.*

We decided to plot simulated liquid water content per specific layers for NCEP1 to demonstrate offsets between model runs and data. Such presentation provides more information than just a height curve with simulated $\theta_w=0$. Referee #1 explicitly asked for a more detailed description and explanation of differences in between both forcings. We consider such presentation as being supportive for this. However, we detected an error in the previous interpolation scheme of the figure plot. We changed the figure to a contour plot with contour intervals at 1 vol%. Due to the coarse depth resolution of MAR below 2 m from the surface (next data point at 3 m depth), we decided to plot the interpolated depth of the 0.1 vol% as approximate for the bottom of the wetting front to address the recommendation of the referee.

- Page 17, Line 18: *'entering' -> 'enters'*

- Page 18, Line 9: *'plain' -> 'planar'*

Changed accordingly

- Page 18, Line 16: *I'm not sure that using MAR provides adequate proof of assumption iii. Isn't MAR modeling these physical processes using assumptions of its own? An observational proof would be more*

useful here. However, I think iii is a very reasonable assumption and don't think it really needs a rigorous proof.

We added the following sentences to describe limitations of this proof: "However, MAR uses assumptions as well to estimate sublimation and evaporation. According to our knowledge, no experimental setup within the deep percolation zone of the GrIS exists to provide a more rigorous proof for assumption (iii)."

- Page 19, Line 6: Strain is by definition a dimensionless quantity, if you are using a measurement of 3.7 cm, this must be deformation, or shortening, not strain.

Changed to compaction.

- Page 19, Line 24: 'could not be' -> 'was not'

- Page 19, Line 30: 'could not' -> 'did not'

Changed accordingly