We thank both referees for very constructive and helpful comments. Each comment of each referee is considered separately in the following. In addition, minor changes such as typos were incorporated in the MS without listing them here. In order to improve readability, comments by the respective referee are listed in italic, while responses and modifications in the MS are written regularly.

Reply to referee #1:

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

Comments to the Author
Assessment
The paper is original with good scientific quality and impact. The technical quality of the manuscript leaves some things to be desired; the text must be cleaned up, the introduction section restructured, see comments below. These are all relatively minor issues.
We appreciate the assessment by the reviewer.

Major comments
Unfortunately, the manuscript text contains multiple typo's and generally suffers from mediocre readability because of frequent unclear formulations. For instance, the introduction requires close scrutiny, several statements are unclear or inaccurate, see (non-exhaustive) listing below under minor comments. The introduction would also benefit from a more logical structure starting with the mass loss from the Greenland ice sheet and then stepwise building up to the importance of meltwater retention, its observation (especially the lack thereof) and model evaluation.
We adjusted the structure of the introduction in accordance to the suggestions made by the reviewer. To improve readability (see also comments raised by Rev#2), we carefully revised larger parts of the manuscript and checked thoroughly for typos. We sincerely apologize, if this appeared to be too sloppy before initial submission.

Adjust the number of significant digits throughout the manuscript to represent the real accuracy of the results. For instance (both on page 12), temperature probably is less significant than stated at 0.01 degree, and mass transfer and/or SWE not at 0.1 kg m-2.
We agree with the reviewer and adjusted given accuracies. However, since the accuracy of the temperature sensors is stated by the manufacturer to 0.25 K, we would like to keep this statement in the manuscript.

Section 3.5: Surely, MAR output must be available at a higher time resolution than 1 day? The model time step must be typically several minutes, so the line 10 statement that "these generate MAR output with a daily temporal resolution" appears inaccurate.
You are correct, MAR outputs are available with sub-daily temporal resolution. However, it remains debatable, whether sub-daily outputs for a 20 km x 20 km grid are valuable for comparison with sub-daily upGPR measurements for a single point. As stated in the manuscript (section 4.3), we still test MAR on its upper end of accuracy. We came to the conclusion that daily averages are adequate for such comparison. However, we changed the respective statements. "We use MAR outputs with a daily temporal resolution and two different forcings, which generate grid cells of 20 km (NCEP1) and 15 km

(ERA-Interim), respectively." In addition, we modified the respective part in the methodology to: "MAR is forced every 6 h by either NCEP1 or ERA-Interim reanalysis data. We decided to use daily outputs for comparisons."

Why does forcing MAR with different reanalysis data (NCEP and ERA-Interim) produces such large differences in surface density and refreezing characteristics, if the same snow model is used (Figs. 8 and 9)? This touches directly upon model performance and deserves to be discussed in more detail. A paragraph with discussion of possible causes for the biases found in the model is also warranted.
We included the following paragraph into the discussion section: "As stated above, predicting individual parameters of the SMB for a point location of the GrIS is beyond the scope of regional climate modeling. Here, we used two different versions of MAR with two different resolutions. This explains already a large fraction of the observed discrepancies for the analyzed parameters density and melt. Since models are usually tuned to accurately reproduce SMB data, individual parameters such as bulk density or bulk liquid water content may result in variable offsets from in-situ data for different climate forcings. In addition, the initial conditions for summer 2016 for both ERA-Interim and NCEP1 are not exactly equal, which causes the model to adjust differently for the individual parameters. Next, clouds have a large impact on the energy balance of the percolation zone of the GrIS. Due to the positive feedback of melt and albedo, small differences in the timing of melt and the amount result in significant offsets for the used forcings. However, upGPR data can help to identify misconceptions in regional climate modeling and, consequently, support further improvements in simulations of temporal changes in snow- and firnpacks."

Minor comments:
p. 1, l. 5: are observable -> have been observed (?). Also p. 2, l. 27 and 28.
Changed accordingly

p. 2, l. 4: "...average negative mass balance all over the ice sheet...": this is unclear; what you probably mean to say is that the ice sheet mass balance became persistently negative.
Changed to "… which resulted in persistent negative mass balances all over the ice sheet"…

p. 2, l. 5: multiplied -> increased
Changed accordingly

p. 2, l. 6: "...Negative annual surface mass balances over the same time period are attributed to an increase in surface melt and runoff...": inaccurate: surface mass balance integrated over the ice sheet has not yet been negative; locally it has, of course. Do you mean: Negative trends in surface mass balance?
Changed accordingly to negative trends

p. 2, l. 9: "...Since melt conditions are expected to continue...": unclear; do you mean: are expected to continue to increase?
Changed accordingly

p. 2, l. 16: "...cause a large fraction of uncertainty cause a large fraction of uncertainty...": consider replacing with: "...is a major component of the uncertainty..."
Replaced accordingly

p. 3, l. 1: "...few existing automatic weather stations ...": Nowadays there are two major AWS networkd on the GrIS: GC-Net and PROMICE.

Changed to: “Apart from several existing automatic weather stations (AWS) being unevenly distributed over the GrIS, no temporal continuous observations exist to validate results of such models.”

Table 2 does not add much information, and its contents can be absorbed in the main text.
Here, we respectfully disagree. You are correct that the content could easily be absorbed in the main text. However, this would require most likely the same space as the table. We consider having a table as much more supportive for such simple dates and numbers than several sentences describing this.

p. 12, l. 20: "...increases in accumulation ...": increases in SWE (?)
Changed to: …snow accumulation…

p. 12, l 22: how is ‘contemporary snow’ defined?
Contemporary has been deleted.