

Dear reviewers,

Thank you very much for the constructive and detailed comments and suggestions to the manuscript. Please find below our responses to the individual points. To facilitate readability, responses are in blue and modifications in the manuscript in red.

Anonymous Referee #1

The manuscript addresses the important problem of determining the different contributors to the liquid water balance applied to Greenland. The study highlights the amount of water in the near-surface, using the subsurface scheme from SNOWPACK, which is actually mobilized and available for runoff. SNOWPACK is forced with the regional climate model RACMO2.3, while the manuscript makes a systematically and quantitative comparison of the impact of the different components in the liquid water balance on the Greenland Ice Sheet. The study shows and discusses the significance of changes in some of the critical model parameters to the overall spatial distribution of modelled water retention. However, it fails to sufficiently discuss the implication of the improved water balance on the surface mass balance.

Overall, this is a decent piece of work, but the manuscript has room for some substantial improvements.

Major points:

1. Highlight differences between Steger et al. (2017) and this study. There seems to be a lot of overlap. Fx fig 4 in Steger et al. (2017) looks almost identical to fig. 5 (e) of this study. It should also be specified if the SNOWPACK model versions and simulations are identical. Also, the firn aquifer description and discussion is very similar.

Figure 5 (e) is indeed almost identical to Fig. 4 in Steger et al. (2017) but we think it is useful to present it again to give a complete picture of the spatial patterns of the most relevant LWB components.

The model version of SNOWPACK was added to the manuscript (Sect. 2.2) and the end of this section was extended to state that the SNOWPACK runs are identical:

A more detailed description of the model setup and the applied spin-up procedure is stated in Steger et al. (2017), where the same SNOWPACK run was used.

Throughout the manuscript we removed redundant statements, particularly in the firn aquifer section and in the conclusions. However, we kept e.g. the discussion about the vertical extent of firn aquifers in view of new available observational data (Montgomery et al., 2017).

2. It is stated on p. 6, line 10-11: "At higher elevations in western Greenland SNOWPACK does simulate a pronounced warming of the firn but there are no in-situ observations available to constrain the magnitude of these changes." The authors should have a look in the extensive GC-net archive of in-situ subsurface temperatures to validate simulated temperatures.

We thank the reviewer for pointing out this data archive. We looked into subsurface temperature data from relevant stations (DYE-2, Crawford Point 1 & 2 and GITS) but the data of (at least) the recent decade is either missing or of very poor quality (large data gaps and/or unphysical high-frequency fluctuations). This is likely caused by the steady progression of sensor deterioration (personal communication with Konrad Steffen). It is therefore not possible to validate the recent simulated firn warming in the lower accumulation zone of the western GrIS with this data set. We added the following sentences to the manuscript to make this clear:

Other snow/firn temperature records are available from the Greenland Climate Network (GC-Net; Steffen and Box 2001) and for the western percolation zone (Humphrey et al., 2012; Charalampidis et al., 2016). [...] Unfortunately, the subsurface temperature data recorded at GC-Net stations located in this area (DYE-2, Crawford Point 1 & 2 and GITS) suffer from large data gaps and/or unphysical high-frequency fluctuations caused by sensor deterioration (K. Steffen, personal communication). The data are thus of insufficient quality to verify these changes.

3. Compare simulated refreezing with available firn cores in the literature. However, I believe, that this was done to some extent in Steger et al. (2017)? Please highlight the main outcome of this analysis. How good is the model performing?

We do not fully understand the reviewer's intention of comparing simulated refreezing with available firn cores. A quantity that is compared to observations in Steger et al. (2017) is snow/firn density, which is a combined result of compaction and refreezing. In terms of firn density, SNOWPACK indicates generally a better performance than the IMAU-FDM - particularly for comparably warmer climatic conditions.

We added to the manuscript:

Although modelled refreezing cannot directly be evaluated with observations, Steger et al. (2017) made a comprehensive assessment of modelled firn density, which is the combined result of dry compaction and refreezing. Results show a reasonable performance of SNOWPACK, but a general overestimation of densities in the percolation zone. This bias is likely the result of overestimated near-surface refreezing caused by neglecting heterogeneous water percolation, an overestimation of fresh snow density and errors in the atmospheric forcing (Steger et al., 2017).

4. Define "skin temperature".

Skin temperature is the equilibrium temperature of an infinitesimally thin layer without heat capacity, which represents the interface between the atmosphere and the ground. It is often used interchangeably with surface temperature. We stated this in the manuscript:

Skin temperature is the temperature of an infinitesimally thin layer without heat capacity, and is representative for surface temperature.

And we now use surface temperature instead of skin temperature throughout the script, because it is a more commonly used term.

5. Quantify statements whenever it is possible throughout the manuscript. For instance, statements like "...good model performance...", "...increase in surface melt...", "...indicate positive trends. . ." or "...temperature increases are highest. . ." in the Abstract should be quantified. Please have a look at the other sections in the manuscript to quantify similar statements. Please have a look at the Conclusions.

We checked all sections of the manuscript and supported statements with numerical values where applicable/possible. Some statements are rather general (particularly in the abstract) and are thus difficult to support with numerical quantities, because providing such number would require a detailed explanation about their validity (e.g. firn temperature increase → where exactly, over which temporal period, averaged over what depth, etc.)

6. The chosen spin up period seems to highly influence crucial subsurface parameters like density (fig 10). This will influence the interpretation of LWB results. Maybe use a different spin up method. Fx the year of 1960 could be used or a mean of the used period (1960-79).

Indeed, the spin-up is an important factor in our simulation that influences the modelled subsurface properties. In the absence of firn profile measurements or knowledge about the climatological conditions prior to the start of the simulation, one has to generate the firn layer with meteorological data from the simulation period (and assume that the prior climate was comparable to the one at the beginning of the simulation period). The definition (length, average, etc.) of the spin-up period is debatable:

As the reviewer suggests, one way is to repeatedly loop over the first year (in our case 1960), as done in other studies (Mottram et al., 2017; Schmidt et al., 2017). However, an issue of this approach is the assumption that the chosen year is representative for the climate prior to the simulation start. By doing this, one neglects the fact that this year may substantially deviate from the mean climate and the interannual variability of the climate is also not captured. Hence, we think that it is more appropriate to use a longer time span for the spin-up. The selected time span of 20 years is long enough to capture interannual variability and it is located in a period where the GrIS experienced a relatively stable climate. As the reviewer states, one could also compute an average over this 20 years for the spin-up. However, we are not convinced that this would improve the spin-up because on the one hand, one would lose interannual variability (which is present in "real" firn profiles), and on the other hand, one may create inconsistencies in the forcing data.

We are aware of the shortcomings of our approach (e.g. the cyclic occurrence of high density layers from years with intensive surface melt), but considering the above-mentioned facts, we do not see a distinctively better technique to perform the spin-up.

7. Several periods are used for comparing the different components in the liquid water balance. Please highlight/argue why all these periods are used.

In the model evaluation part, the chosen periods are constrained by the temporal availability of observational data. In the comparison of modelled and measured firn temperatures, we provide RACMO2.3 surface mass fluxes, which are averaged over the 20 years prior to the temperature measurements. Due to the lack of model data before 1960, we assume a constant climate and provide averaged mass fluxes over 1960–1979 for the temperature records of 1952/1955. For the majority of the LWB climatology, we use the two periods (1960–1989) and (1990–2014). This partitioning is very similar to the one applied in Van den Broeke et al. (2016) and it divides the entire simulated period in a first part with a rather constant climate and a second part with a distinct increase in melt. Additionally, we show averages over the entire period (1960–2014) when we discuss mean characteristics.

In Fig. 9, we illustrate the 2012 anomaly because this year was exceptional in terms of melt extent and amount. In the following discussion of changes in vertical firn properties and firn aquifers, we illustrate changes over the entire simulation period and thus select the years 1960 and 2014.

In Fig. 13, the first period is identical to the spin-up period of the model to illustrate the occurrence of firn aquifers in a steady-state climate. The second period was somehow arbitrary selected and is now changed to (2010–2014), which is identical to the period of firn aquifers observations by remote sensing (Miège et al., 2016).

We added the temporal availability of the ice discharge and GRACE data to Sect. 2.3 to elaborate on the selected temporal period for computing the MB (2003–2012). The reason for using the two periods (1960–1989 and 1990–2014) is briefly explained in Sect. 4.1. In Sect. 4.2, we state why we selected the year 2012 for a more detailed analysis in terms of refreezing and firn temperature change. Additionally, we changed the following sentence to explain the second period (2010–2014) used for the evaluation of firn aquifer occurrence:

To assess the influence of a transient climate, firn aquifer occurrence as a function of snowfall and liquid input has also been computed for the period 2010–2014, which is identical to firn aquifer observations by remote sensing (Miège et al., 2016).

8. You should state when the results are presented and discussed.

We now state at the end of Sect. 1 (Introduction) the beginning of the results and discussion sections by changing the following sentence:

Subsequently, we discuss the comparison of model output with remote sensing data (GRACE) and in situ measurements (firn temperatures). Section 4 contains the results of the LWB evaluation and a more detailed analysis of refreezing, runoff and changes in different firn properties.

9. Since you state in the introduction that a better LWB contributes to a better estimate of surface mass balance, how did your experiment modify the surface mass balance? It would be an important point, which is not addressed/discussed sufficiently. Fx Paragraph 4.4: The description is useful to understand figure 14 but it misses explanation on why those changes in snow/firn/ice melt are relevant for the SMB.

Please see answer to question 10 (major points).

10. In the comparison with grace what is due to the new liquid water balance? How is the liquid water balance influencing the surface mass balance?

In order to highlight the link between the LWB and the SMB, we also included the definition of the SMB in Sect. 2.1 and briefly mention the common components.

Considering the LWB, refreezing and runoff are the quantities that are most different between SNOWPACK and RACMO2.3's own snow model. In regions with high melt, SNOWPACK tends to simulate lower and more realistic snow/firn densities (Steger et al., 2017). This allows, together with the higher value of prescribed irreducible water, for more refreezing (and thus less runoff). To illustrate how this affects SMB, MB computed entirely with RACMO2.3 data is now included in Fig. 3.

The higher refreezing and lower runoff rates in SNOWPACK lead to a more positive SMB and thus to a reduced decline in cumulative MB. This improves the agreement with GRACE. We added the following text to the manuscript (Sect. 3.1) to discuss this:

A comparison between the derived cumulative MB and GRACE is provided in Fig. 3a. The MB is computed by taking the simulated SMB over the glaciated area either from RACMO2.3 or SNOWPACK. Both cumulative MBs indicate an excellent agreement with GRACE ($R^2 > 0.99$). In terms of linear trends, SNOWPACK agrees better with GRACE due to higher modelled refreezing fractions and thus lower amounts of runoff from the ice sheet. [...] The minima in the MBs occur both earlier and with higher magnitudes than in GRACE, where SNOWPACK performs slightly better due to smaller amounts of modelled runoff.

11. Basin scale GRACE comparisons with surface mass balance could improve our understanding.

As the reviewer states correctly, a basin-scale comparison would enhance our understanding of the modelled SMB. However, despite the availability of GRACE products on basin scales, we omit such a comparison because we suspect significant signal leakage in GRACE on the spatial scale of the eight basins. Hence, we are not convinced that a comparison of the trends (of GRACE and SMB – ice discharge) on a basin scale would yield meaningful additional insights.

Figures:

Fig. 1: Nice illustrative figure.

Thanks.

Fig. 2: Nice illustrative figure.

Thanks again.

Fig. 3: It does not make sense to state a correlation value in Fig 3 unless the time series have undergone a high pass filter, which allow the analysis of the variability of shorter time scales when compared to annual cycle. I.e. the annual cycle should be removed, as it will dominate the correlation. Here I would also recommend giving a seasonal R-squared then the mean of that and its standard deviation, which would give a good overview on the mean performance and on its variability.

We do agree with the reviewer's opinion that applying a high pass filter to the time series would be necessary to assess variability on shorter time scales than the annual cycle. However, the trend as well as the annual cycle are the signal components we actually want to compare between GRACE and the MB. Therefore, we do think it is meaningful to compute a direct correlation between the time series.

We did not incorporate the reviewer's suggestion of computing a seasonal R-squared as we were not able to understand the intended method entirely.

Fig 4: please add RACMO2.3 surface accum. (1994-2013) to illustrate if changes in accumulation is responsible any temperature changes.

We added RACMO2.3 surface accumulation (1994–2013) to Fig. 4 and added the following sentence to Sect. 3.2:

Along the entire transect, modelled increases in solid precipitation are spatially rather uniform and small (~ 0.02 m w.e. a^{-1}), and therefore likely less relevant for explaining changes in firn temperature.

Fig 5: I wonder if the quantities of figs. 5d-f are influenced by the spin up method. Also, the title of (e) and (f) should not be "refreezing" because, I suspect, that the figures show values of both refrozen and liquid water being retained in the firn. Fx basin 4 is where the perennial aquifers exist seems to be a lot of refreezing there.

The dependency of runoff on the spin-up method chosen is rather small for the majority of the GrIS, as runoff mostly originates from bare ice (Fig. 14). Part of the runoff from snow/firn originates from seasonal snow over bare ice in the ablation zone and is therefore not influenced by the spin-up either.

In Basins 4 and 5 however, a considerable amount of runoff originates from porous snow/firn. These areas are more sensitive to the selected spin-up method but other ill-constrained factors, such as vertical firn temperature initialisation and heat flux description at the bottom of the model domain, will outweigh these uncertainties because the accumulation rate rapidly refreshes the firn layer. Refreezing likely exhibits the strongest dependency on the chosen spin-up method, but because of the reason mentioned above firn porosity and cold content mainly depend on the recent climate and thus we do not expect substantially modified refreezing rates if the spin-up method were changed. The titles of (e) and (f) are correct and only refreezing rates are displayed, i.e. not the combined effect of refreezing and liquid water retention.

To clarify this, we added the following sentences to Sect. 2.2:

Neglecting heterogeneous percolation causes refreezing to occur mostly in the upper snowpack, where temperature and porosity are determined by the recent climate.

Fig 6: Same concern with refreezing vs. retention as in Fig. 5.

Here, also, only refreezing rates are displayed (not combined with liquid water retention in the firn).

Fig 7: Same concern with refreezing vs. retention as in Figs. 5 and 6.

Please see previous answers.

Fig 8: Please explain in the text, why do you show differences between the periods of 1960-1989 (30 years) and 1990-2014 (25 years). These results could also be influenced by the spin up period.

See answer to question 6 and 7 (major points): the main reason is to have a reference period with a relatively constant climate and a period with rapid change, both having comparable lengths.

Fig 9: This is a nice plot, as (a) shows where the firn in 2012 lost its capacity to retain water compared to the reference period. However, I suspect it should be a retention anomaly.

The plot shows only refrozen mass (not combined with liquid water retained in the firn).

Fig 10: Again, this figure clearly shows the influence of the spin up period. This is evident on the western side of the ice sheet with three highly identical subsurface features in the density.

True, please see answer to question 6 (major points).

Fig 11: Again, spin up problems?

Please see answer to question 6 (major points). It is also interesting to note that this only becomes evident close to the equilibrium line where accumulation is small.

Fig 12: Nice plot!

Thanks.

Fig 13: Nice plot!

Thanks.

Fig 14: Please explain the implication of changes in the liquid water balance on the modelled runoff pattern in more detail.

Please see answer to question 10 (major points).

Tab 1: I would like to see all trends even if they are not significant.

We now show all trends and marked the insignificant ones by an asterisk.

Specific points:

Line 27 p.3: Which bucket scheme? Please more details and references.

We added two references to the manuscript (in Sect. 2.2), in which the bucket scheme used in SNOWPACK is described.

Line 1 on p. 4: It should be mentioned if the two model setups are identical. If not, the differences should be highlighted.

Please see answer to question 1 (major points).

Lines 9-15 on p. 4: More detail is needed for this description of observational data and what is it being used for? Also, many datasets are available (remote sensing, station data, historical and contemporary SMB measurements...) for further validation of the forcing data and of the model output. This part could be improved, which would make the conclusions more solid.

The RACMO2.3 output, used as forcing data in this study, was already extensively evaluated in the study by Noël et al. (2015). We have now made this more explicit in the manuscript by adding the following sentence to Sect. 2.2:

The capability of RACMO2.3 to accurately simulate present-day surface climate on the GrIS was illustrated in an extensive evaluation by Noël et al. (2015).

SNOWPACK's performance in terms of snow/firn density and firn aquifer extent was assessed in the Steger et al. (2017). In this manuscript, we additionally evaluated the spatially integrated SMB of SNOWPACK (forced with RACMO2.3) and we compared observed and modelled firn temperatures (where measurements are available and a comparison is meaningful). We agree with the reviewer that further evaluations of SNOWPACK would be interesting, particularly for subsurface processes as refreezing, but we think that we exploited all presently available/suitable data sources.

Line 6 on p. 5: How is the tundra hydrology dealt with?

We kindly refer the reviewer to the first paragraph of Sect. 3.1 where the RACMO2.3 tundra snow model is described, and to the last paragraph of this section where we list relevant features of the tundra hydrology. These are currently not considered in our model framework.

Line 24 on p. 5: It is not only the subsurface temperatures that may be bias but also the density profile.

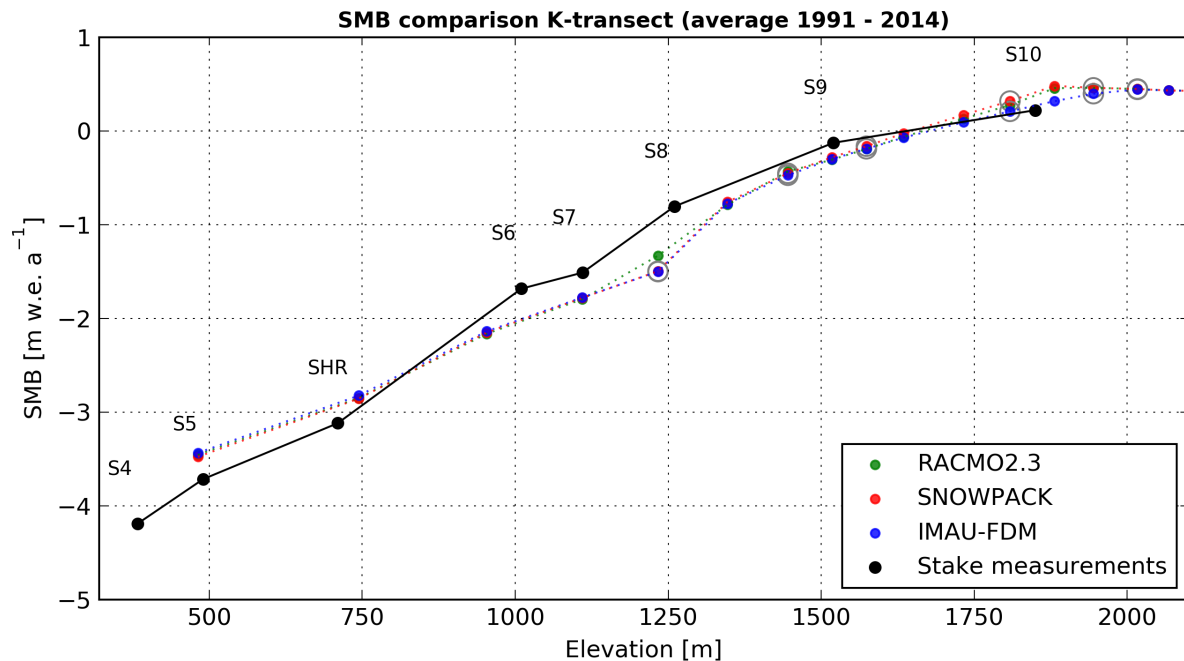
Good point; we included this consideration by changing the sentence to:

The bias for the second period is more difficult to explain in the absence of continuous firn temperature measurements and firn density records.

Line 4 to 7 p. 6: Here and later at KAN-U you mention the overestimation of bare ice zone. A quantification of the spatial extend of this bias would be useful (comparison with remotely sensed bare ice areas?). It would go hand in hand with the many observed SMB available in western Greenland (K transect EGIG line): how does the model compare to them.

A qualitative picture of this overestimation can be found in Steger et al. (2017) (Fig. 9), where vertical firn density transects of the IMAU-FDM and SNOWPACK are compared with data from the NASA Operation IceBridge accumulation radar. The figure confirms the overestimation of the bare ice zone by both models but it does not allow for an exact quantification of the bias.

A quantification of the bare ice zone with remote sensing data (by distinguishing between different surface properties of snow/firn and ice), as probably intended by the reviewer, could be performed. However, delineating ice and snow/firn areas with such a technique could result in a substantial overestimation of the bare ice zone due to a misclassification of porous firn covered by near-surface ice (Machguth et al., 2016). We consider such an effort beyond the scope of the present paper. Therefore, we restrict our evaluation to a comparison of modelled SMB values along the K-transect:



The above figure shows SMB components of the different models compared to stake measurements along the K-transect (average 1990–2014). All models return similar SMB values, which indicates that the forcing data by RACMO2.3 dominates the modelled SMB values rather than the individual snow/firn models, at least for this transect.

The values of RACMO2.3 and SNOWPACK are almost identical and there is only one larger difference around S8, which is caused by the interpolation of SNOWPACK data from the checkerboard to the full RACMO2.3 grid (Steger et al., 2017). The IMAU-FDM has higher runoff and therefore lower SMB values around S10 compared to the other models, for reasons explained earlier. Generally, modelled SMB values are in close agreement with stake measurements between S9 and S10, where the transition of bare ice to porous firn is located. However, there is a fundamental problem of comparing stake measurements to modelled SMB values at locations with porous firn: Stake measurements only capture mass changes at the surface. The modelled SMB as presented in this paper however considers vertically integrated mass changes (→ climatic SMB). I. e. the climatic SMB at S10 could be higher than the one derived by stake measurements due to refreezing of mass deeper in the firn.

We added the following sentence to the manuscript to state the difficulty of using remote sensing data to quantify the bare ice zone extent:

Inferring the bare ice zone from remote sensing data, e.g. by using the different surface properties of snow and ice, is complicated due to formation of near-surface ice layer (Machguth et al., 2016) above porous firn.

Line 6 on p. 6: Who are “they”?

“They” refers to the models. The sentence was reformulated to remove this ambiguity:

The reason is the overestimation of the bare ice zone on the western GrIS by the IMAU-FDM and SNOWPACK; i.e. the models are incapable of simulating the subsurface warming due to a deficiency of pore space for refreezing.

Lines 10-12 on p. 6: Fx, please have a look at GC-Net data.

Please see answer to question 2 (major points).

Line 13 on p. 6: Near-surface snow density depends mostly on wind and subsurface vapour fluxes.

We agree with the reviewer that wind and subsurface vapour fluxes are relevant factors for the near-surface density. Instead of adapting our sentence, we decided to remove it because it didn't really connect this paragraph to the previous one. We added a new sentence to accomplish that:

To address the above-mentioned model bias in overestimating the bare ice zone, we briefly assessed fresh snow density, which is a rather uncertain factor in our simulation.

Line 22 on p. 6: Compaction here is mostly due to wind and vapour fluxes. Line 29 on p. 6: Please quantify this inaccuracy?

We agree with the reviewer that compaction due to wind and vapour fluxes are the most important factors for near-surface snow densification. However, these effects are already taken into account in the fresh snow density parameterisation we apply. To clarify, we wrote:

The parameterisation, which accounts for near-surface densification due to wind and vapour fluxes, clearly overestimates fresh snow density for this region.

And we removed the part about overburden-dependent compaction, because it is indeed of minor relevance in the uppermost 50 cm compared to other process that induce densification.

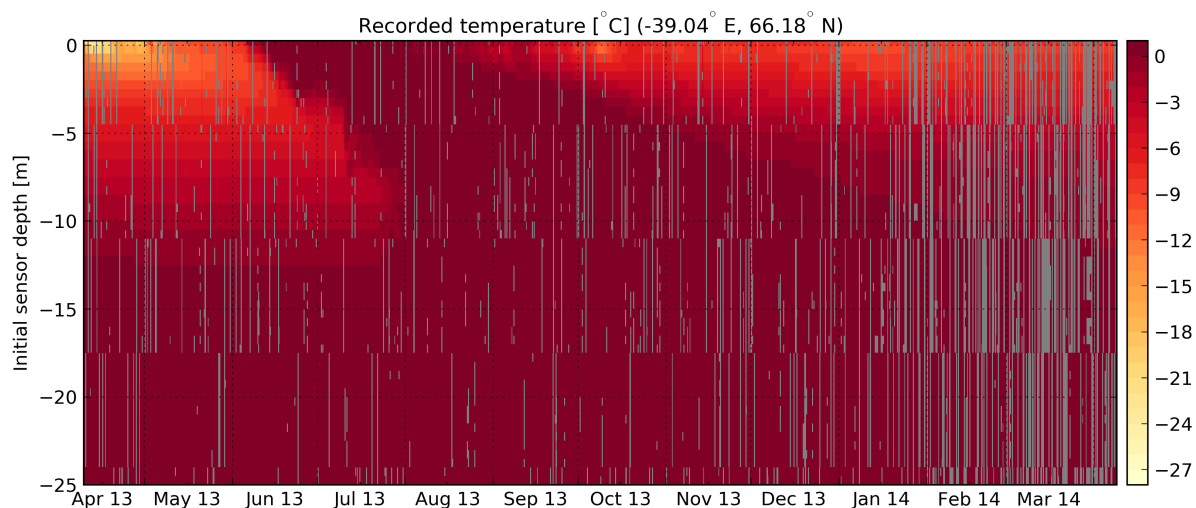
Line 19 on p. 7: Describe the results from Table 1 in more detail.

We extended the description about linear trends (1990–2014) in the manuscript (Sect. 4.1) - particularly for changes in melt and rainfall:

Changes are particularly large for Basin 5 and 6, where melt increases by $0.36 \text{ m w.e. a}^{-1}$ and $0.38 \text{ m w.e. a}^{-1}$, respectively. The dominant cause for these large changes is the comparably high increase of melt in the ablation area of the GrIS, especially in the southwest. Modelled snow melt in the ablation zone is particularly sensitive to temperature increases due to the albedo difference between snow and ice, where bare ice with a lower albedo is more rapidly exposed through accelerated melt of snow. The lowered surface albedo subsequently enhances melt of bare ice. A secondary cause is the relatively flat hypsometry of these basins, where 58 % respectively 47 % of the area is below 2000 m a.s.l. (compared to 39 % for the GrIS). Rainfall, as a further contributor to liquid input, does not exhibit a significant trend for the majority of the basins. Linear trends are comparably high for Basin 5 ($1.22 \text{ mm w.e. a}^{-2}$) and Basin 6 ($0.43 \text{ mm w.e. a}^{-2}$) but statistically insignificant. Remarkably, the northwestern Basin 8 is the only region with a significant positive trend in rainfall of $0.56 \text{ mm w.e. a}^{-2}$. This increase is not caused by a change in total precipitation but by a significant increase of the rainfall fraction in this area.

Lines 12-16 on p. 8: Should be assessed using observations

The figure below shows the temporal evolution of firn temperature between April 2013 and March 2014 for a firn aquifer location in southeast Greenland (we thank Clément Miège for providing this data).



Grey areas mark missing data and it should be noted that the vertical coordinate always refers to the initial depth of the temperature sensors (→ no correction for surface accumulation/ablation). It is evident that the vertically integrated cold content of the firn layer steadily decreases from April to June until the entire firn column is isothermal at 0°C in early August. Near-surface temperatures (\sim upmost 0.5 m) are at the melting point between mid-June and mid-August due to high incoming shortwave

radiation and/or downward directed sensible heat fluxes. Additionally, it is likely that the release of latent heat from refreezing (particularly during the night) keeps the near-surface temperature constantly at 0° C. However, to verify this, one would need temperature observations with a high temporal resolution (e.g. hourly) and with a sensor that is always in very close proximity to the surface. We are not able to assess this constant surface temperature (on a diurnal time scale) with the current observational data.

Lines 21-25 on p.8: Please explain in more detail the asymmetrical retention pattern and the consequences of this.

Corrected by changing the mentioned paragraph.

Line 10 on p. 9: Again, GC-Net firn temperatures can be used here.

Please see answer to question 2 (major points).

Lines 6-7 on p. 10: How good is this threshold? Where does it come from? Please more support for this is needed.

This threshold is derived from remote sensing of firn aquifers. It is estimated to be the approximate sensitivity of the accumulation radar to detect liquid water in the firn (Miège et al., 2016). We added the sentence:

The above-mentioned threshold for firn aquifer delineation is based on a sensitivity estimation of the NASA Operation IceBridge accumulation radar to detect liquid water in firn (Miège et al., 2016).

Referee 2 (Max Stevens)

Summary:

This paper presents results of an investigation of the liquid water balance (LWB) of the Greenland Ice Sheet using firn/snow-model (SNOWPACK). The authors' goals are to quantify the components of the LWB spatially and how those quantities have changed in recent decades; to investigate temporal and spatial patterns in refreezing and how those affect the firn; and to assess the models' ability to simulate firn aquifers. The authors force their model using climatic data from the regional climate model RACMO.

Observations of LWB components are unfortunately scarce, but the authors use available data (GRACE data, firn temperatures, firn-aquifer extents) to evaluate their model performance. The model results compare to GRACE data very well. The model does not reproduce firn temperatures as well, but the results do still compare to the data favourably.

General Comments:

The paper is written and organized well and is a topic of wide interest to the cryospheric-science community. Understanding the LWB is important, as the authors identify, because uncertainty in the runoff component of surface mass balance is a large contributor to sea-level-change estimates. Accurate model simulations are an essential contribution to this scientific issue. Additionally, the authors do a good job of discussing potential sources of model error. I recommend this manuscript for publication with minor revisions.

Thank you very much for this positive and encouraging feedback.

General points to address:

The authors mention firn "structure" numerous times. I think to many in the firn community "firn structure" refers to microstructural properties such as grain size, coordination number, etc. In this case, the authors refer specifically to firn temperature and porosity. It may be appropriate to call them by name specifically or use "firn physical properties" as the terminology.

We agree and changed this term throughout the manuscript to "firn properties".

The authors ignore any lateral flow and also any heterogeneous flow (i.e. piping). I would like a bit more discussion on how those might affect the results, or if it is even possible to know at this point. Section 2.1 asserts that pore space downhill is often filled, but can enough hydraulic head be generated to drive a significant amount of flow? Is there enough data on piping available to do an easy scale analysis of how much heat could be delivered (how deep, how fast?)

The reviewer is right that these two potentially important processes (lateral flow/heterogeneous percolation) are not considered in the current framework. The main reason for not including these processes in our study is the lack of observational data to constrain the simulation on a GrIS-wide scale. This renders it impossible to quantify the magnitude of these processes – even in a simple scale analysis. Subsequently, we listed some more detailed consideration about the two processes:

- Lateral water transport: It is likely that some fraction of the lateral runoff is retained along its path to the ocean in various parts of the glacier system (e.g. accumulation in subglacial lakes, refreezing in cold snow/firn or storage in firn aquifers). However, the limited observational data make it impossible to quantify the magnitude of this process on a GrIS-wide scale. We can only speculate that, at least in terms of near-surface water retention, the magnitude is rather small due to the mentioned seasonal upward migration of the melt area and the associated depletion of pore space and cold content in lower elevated areas.
- Heterogeneous percolation: There are different methods available to model this process, like the statistical approach of Marchenko et al. (2017) and the physical-based approach of Wever et al. (2016). However, both methods contain at least two tuning parameters that require observational data to calibrate the models. In addition, the approach by Wever et al. (2016) is computationally expensive (it demands a high vertical resolution and solving the Richards equation twice) which makes it unsuitable for distributed model runs with a large number of grid cells.

We modified Sect. 2.2 to clearly state that we neglect these two processes in our model:

We do not consider heterogeneous percolation (Wever et al., 2016; Marchenko et al., 2017) in our simulation due to an insufficient spatial coverage of observational data to calibrate such routines for the entire ice sheet and/or the too expensive computational demand. [...] Lateral flow of runoff is also not considered in our simulation.

In Sect. 3.2, we briefly discuss a location where modelled firn temperature is likely biased due to the neglect of heterogeneous percolation. The sentence about the effect of neglecting lateral flow on horizontal redistribution of mass and energy was shifted from Sect. 2.1 to Sect. 5 (and slightly modified):

This is likely a less relevant issue for horizontal near-surface redistribution of mass and energy, as surface melt typically reaches higher elevated areas later in the season, which means that lower areas are already depleted of pore space and/or cold content and thus do not provide any more storage volume for upstream runoff.

Section 3.2 and figure 4 compares model results to observations. The authors identify biases in the model and discuss lower elevation sites, but what about the higher elevation sites? The model shows a somewhat uniform temperature increase over the period, but the observations are not as spatially coherent, e.g. sites 4-050 and 4-000. Also, the model gets the lower-elevation structure correct for the modern, but does not as well for 1960. Why is this? Does your assumption that 1952/55 would be the same as 1960 break down (I do think that is a very reasonable assumption, however.)

We added the following sentences to discuss the incoherency in observed firn temperature change at higher elevations:

The spatially incoherent firn temperature change (between B 4-225 and B 4-000) in the observations is not reproduced by SNOWPACK, which simulates a uniform temperature increase of $\sim 0.3^\circ$. This incoherency in the observations may be partly explained by uncertainties in the measurements caused by errors in the sensor calibration and uncertainties in the applied correction used to retrieve 10 m firn temperature from shallower measurements (Polashenski et al., 2014).

We are uncertain to which part of the transect the reviewer is referring with “lower-elevation structure”. Discrepancies between modelled and observed temperatures could again be caused by uncertainties in the temperature measurements (as mentioned above).

Section 3.2 (end): If the different parameterization for fresh-snow density works better, why did you not just use that one? How does this uncertainty affect the results for the higher elevation sites? When you say, “An improved fresh snow density parameterization seems therefore essential to address this inaccuracy”, do you mean an entirely new parameterization is needed, or just a new-to-your-model parameterization? Does the Langen (2017) parameterization fit the criteria of an improved parameterization?

Because of computational constraints we could only test the parameterisation by Langen et al. (2017) for the mentioned regions in western and southeastern Greenland. For these areas with rather warm climate conditions, the parameterisation outperforms the one we apply from Kuipers Munneke et al. (2015). For colder conditions however, the parameterisation by Kuipers Munneke et al. (2015) is performing better and the parameterisation by Langen et al. (2017) yields too low initial densities for SNOWPACK.

A main issue in our opinion is the proper disentanglement of processes that influence near-surface snow density:

- Density of freshly fallen snow
- Density of deposited snow from snow drift → particles become smaller due to collisions/enhanced sublimation (Groot Zwaaftink et al., 2013)
- Vapour fluxes
- Melt and refreezing

With available fresh snow density parameterisations, it's often unclear which of these processes are considered in the parameterisation and which have to be explicitly modelled. It would be necessary to test SNOWPACK with different available (or newly derived) parameterisations that optimally depend on meteorological parameters, in contrast to the parameterisation by Langen et al. (2017), and compare simulated densities with snow/firn density from locations with various climate conditions to end up with a significantly improved fresh snow parameterisation. Developing such a new expression is clearly beyond the scope of this manuscript.

We simply replaced the last sentence of this paragraph to state our intention more clearly:

This model inaccuracy should be addressed in the future by testing available or newly derived fresh snow density parameterisations with SNOWPACK for various climate conditions on the GrIS.

Section 4.1: Please clarify: You say “changes in the retained liquid mass (dM_{ret}/dt) are even smaller...”. Equation 1 defines dM_{ret}/dt as the liquid water balance. Is retained liquid mass the same as the LWB? In that sentence, does the quantity in parentheses (dM_{ret}/dt) refer to ‘changes in retained liquid mass’ or to ‘retained liquid mass’? I don’t doubt the science here but it was confusing to read. If changes in LWB, defined as dM_{ret}/dt , are indeed small, then it might imply that it is not an important term in changes in SMB.

The liquid water balance, as stated in Eq. (1), represents the sum of all liquid water mass fluxes through the upper and lower boundaries of the snow/firn column and refreezing, which is an internal sink for liquid water. The residual of this balance is dM_{ret}/dt , the change of retained liquid mass in the snow/firn column with time. Horizontally integrated values of M_{ret} (e.g. over basins) are comparably small, hence values of dM_{ret}/dt are also small. M_{ret} can be substantial for model grid cells with firn aquifers (see Fig. 7 in Steger et al. 2017) but changes in the retained liquid mass (dM_{ret}/dt) are typically still small compared to components on the right-hand side of equation (1). We rephrased the following sentence in the manuscript to clarify this:

Changes in the retained liquid mass (dM_{ret}/dt) are even smaller than components on the right-hand side of Eq. (1), particularly when integrated over basins, and are thus not presented.

End of page 8, start of page 9, and Figure 8: Please clarify your language: Basin 4 shows a very large increase in firn temperature (due to refreezing I believe), but then you talk about how changes in t_{skin} are also important (but basin 4 does not show increase in t_{skin}). Adding a few sentences to clarify this would help – which phenomenon is important where?

There are two important processes that influence local firn temperature:

- Warming/cooling by heat conduction (e.g. caused by a change in surface temperature)
- Warming by release of latent heat from refreezing

In the addressed paragraph, we intended to state that for Basin 4, refreezing is clearly the dominant factor for the increase in firn temperature, whereas for Basin 2, the increase in firn temperature is rather caused by an increase in the surface temperature. We reformulated this paragraph to clarify:

In contrast to other basins, Basin 2 reveals a relatively constant firn temperature increase at lower elevations. This increase is not only caused by the rather small increase in refreezing and the associated latent heat release, but also by an enhanced vertical heat flux from the surface through an increase in surface temperature (Fig. 8d). Surface temperatures changes show a distinct spatial variability, with the largest increases occurring in the northeastern part of the ice sheet, where temperature increases by more than 1.5° C.

Specific/technical corrections:

Use of units throughout: in some places, the authors use units of $kg/m^2/a$ (e.g. Figure 5) and in others $w.e./a$ (e.g. Figure 6). It would be good to have consistency throughout. I slightly prefer $m w.e./a$ in this context because it is slightly more intuitive.

Corrected by consistently using $m w.e.$ (or $mm w.e.$ for smaller values) in all figures and in the text.

Use of vague language throughout: several instances of “it seems” or “apparently”. Just say what you mean directly. E.g. page 5 line 19: change to “it is reasonable”.

Corrected by rewriting sentences with “it seems” (Sect. 1, 3.2, 5) and “apparently” (Sect. 3.2, 4.3)

Page 1, Line 5: “good model performance” is vague; perhaps “indicate good model- observation agreement” or something along those lines

Corrected.

P1L7: “increases with” change to “increases at”

Corrected.

P1L13: be aware that upward could also mean forming at a shallower depth “migration of firn aquifers to higher elevations”

Corrected.

P1L20-21: put the e.g. section in parentheses to break up the sentence more clearly; change to “and the darkening”

Corrected.

P1L24: change to “suggest that modelled refreezing”

Corrected.

P3L15: perhaps a new paragraph at “The Greenland mass...”

Corrected.

P4L1: please clarify: do you mean densification scheme, as in how density changes with time, or the parameterization for new snow density that you use?

Thank you for pointing out this error. We indeed meant the fresh snow density parameterisation and not snow densification. We change the sentence to:

The enhanced near-surface snow compaction due to strong winds, which is implemented in SNOWPACK for Antarctic simulations (Groot Zwaaftink et al., 2013), is switched off, because the applied fresh snow density parameterisation already accounts for this effect.

P4L9: get rid of word Furthermore.

Corrected.

P4L19: it is a bit unclear what you mean by “indirectly”. “over” is probably not the best word here.

We rephrased this sentence:

Due to the lack of direct refreezing observations, we assess the model's performance in terms of refreezing indirectly by comparing the spatially integrated SMB and local snow/firn temperatures to observations.

P4L25: start new paragraph at “A comparison”, rather than at line 28.

Corrected.

P5L3: do you mean delaying runoff by 18 days in the model? Please clarify.

We rephrased this sentence:

Van Angelen et al. (2014) demonstrated that the monthly error between detrended modelled SMB and GRACE on a GrIS-wide scale could be minimised by delaying simulated runoff by 18 days.

P5L7: mean seasonal amplitude in what?

Clarified by changing the sentence to:

The mean seasonal amplitude of the detrended modelled MB derives by ~30 % from winter accumulation and summer melting of seasonal snow over the tundra (Fig. 3b).

P5L7: change to “A too-early modelled snow...”

Corrected.

P5L13-14: remove semi-colon and e.g., change to “. . . et al. 2015), or accumulating...”

Corrected by rewriting the sentence to:

Finally, runoff may also be retained in the hydrological system of the tundra by refreezing in soil, ponding on frozen ground (Johansson et al., 2015), accumulating in surface lakes (Mielko and Woo, 2006) and storage in terrestrial aquifers.

P5L20-21: change to “Figure 4 shows that SNOWPACK...”

Corrected.

P6L5: I think you have not defined IMAU-FDM acronym prior to this use.

Corrected.

P6L8: The sentence starting “Due to a different...” is a bit awkward to read – try to rephrase to be active voice.

Corrected by rewriting the sentence to:

Compared to IMAU-FDM, the overestimation of this zone is less pronounced in SNOWPACK owing to a different densification scheme, which is more accurate for relatively warm conditions (Steger et al., 2017).

P7L12: 47% refreezes – how has this changed since 1960?

The reviewer is kindly referred to Fig. 6 and Table 1. Figure 6 (panel for GrIS) shows that the refreezing fraction is rather constant for the first period (1960 – 1989) with some interannual variability. The fraction decreases during the second period (1990 – 2014) according to the linear trend stated in Table 1.

P7L17: list the four most relevant components in parentheses.

Corrected.

P7L20: “Remarkably” – why is this remarkable? Would you have expected other regions to also have more rain? Is there more precipitation in total in that region, or higher rainfall as a percentage of total precipitation – how does that compare to other areas?

We expected significant increases in rainfall also in other regions, particularly in the more southerly located basins. We checked time series and linear trends (for 1990–2014) for rainfall, snowfall, total precipitation and rainfall fractions for the GrIS and the eight basins. The only significant trends that were found are from Basin 8, where both rainfall and rainfall fraction exhibit a positive significant trend. We included this finding also in the manuscript:

This increase is not caused by a change in total precipitation but by a significant increase in the rainfall fraction in this area.

P7L27: change “does” to “do”

Corrected.

P7L31: change “one” to “trend”

The sentence was modified.

P8L6: Is melt climate? I might suggest that temperature is climate, and melt is a result (but maybe I am wrong or nit-picking or both)

We agree with the reviewer’s comment that melt is not really a climatic factor but rather the result of the climatic factor temperature. Hence, we change the sentence as follows:

Refreezing is a process that strongly depends on local climate, i.e. particularly on surface temperature which is the main driver for melt, and therewith on seasonality and elevation (Fig. 7).

P8L13: do you mean “the seasonal decrease in refreezing at higher elevations is caused by . . .”?

We merged this sentence with the previous one to clarify our statement:

Therefore, refreezing at lower elevations persists throughout the melt season but with lower rates than in spring due to a gradual decrease in the firn cold content.

P8L23-26: Are these sentences about hysteresis needed? They seem distracting and not relevant to me, but if they are relevant, include some discussion about how it affects your results and conclusions.

We decided to rewrite this paragraph and remove the term hysteresis, because it would require too much additional explanation.

P8L29: get rid of word “even”

Corrected.

P8L30: change to “causes melt and refreezing...”

Corrected.

P9L18: change “higher” to “larger”

Corrected.

P9L18 and Figure 10: perhaps adding a panel to figure 10 to show pore space across the transect?

We agree that pore space would be an interesting additional parameter to show. However, because it can be inferred from firn density and to keep the figure size reasonable, we decided not to include this additional parameter.

P10L10: get rid of word “apparently”

Corrected.

P10L16: get rid of apparently; say something like, “Instead, in these regions liquid water drains into...”

Corrected.

P10L20: Change to: “Comparing the modelled depth of the firn-aquifer top to observations is difficult...”

Corrected.

P10L23: change to “unsaturated wet layer”

Corrected.

P10L26: unclear what you mean by advection of cold interior ice here – please expand on the physics of what is going on.

What we meant is the following: We initialise firn temperatures for our simulation with a vertically constant value equal to the RACMO2.3 skin temperature (1960 – 1979 average). Next, we correct this temperature for latent heat release by refreezing with a parameterisation (Steger et al. 2017). As a result, firn temperatures are mostly initialised with temperatures close to 0°C at firn aquifer locations. With the assumption of a zero-heat flux at the base of the model domain, no heat sinks are present and modelled deeper firn temperatures remain at 0°C. The base of modelled aquifers is then determined by firn compaction (liquid water is “squeezed” out if there is no pore space left). Recent observations indicate however that refreezing conditions often prevail at the base of aquifers (and that

this effect thus limits the vertical extent of aquifers). Apparently, there is a “cold reservoir” beneath aquifers that induces a downward directed heat flux. This colder ice is advected laterally from the interior of the ice sheet, because it originates from higher elevations.

The sentence was however removed and we generally modified this section due to overlaps with Steger et al. (2017). Wherever necessary, we refer to this publication in the manuscript.

P10L28: would initializing the model with lower temperatures be a physical thing do to? Is there any reason to believe that firn temperatures were lower in 1960 (assuming that is the start of the simulation)?

Initialising deeper firn temperatures with lower values and applying a downward directed heat flux at the lower boundary of the model seems reasonable in view of recent observations that reveal refreezing conditions at the base of firn aquifer, see the explanation above. However, a more detailed knowledge of the thermodynamic conditions deeper in the firn is required to apply such modifications.

P10L32: Specify the time period over which that expansion is occurring.

We changed this sentence to clearly state the different time periods for modelled and observed firn aquifer expansion:

Figure 11 shows an expansion of the firn aquifer to higher elevations during the simulation period (1960–2014). This trend is in line with observations (2010–2016), which indicate an inland expansion of aquifers in this region (Miège et al., 2016; Montgomery et al., 2017).

Evaluations of radar measurements for an even longer period (1993 – 2016) confirm the inland expansion of firn aquifers – at least for the Helheim region (Miège et al., in preparation).

P11L4-5: are these trends or just changes?

We changed this sentence and the one before to clarify our statements:

In Basins 4 and 5, the mean surface elevation at which firn aquifers are modelled rises by ~200 m during the simulation period (1960 – 2014). Upward migration of firn aquifers is also apparent in other basins, where Basin 3 and 6 reveal a smaller change of 125 and 90 m, respectively, and Basin 8 a larger elevation increase of 215 m.

P11L11: hyphen in ice-sheet margin

Sentence was removed due to overlap with Steger et al. (2017).

P11L12: s in aquifers; change “i.e.” to “specifically”

Corrected by adding an “s” to “aquifer”. However, we think that “i.e.” fits better in this context than “specifically” – hence we did not amend this expression.

P11L14: no comma after simulation

Corrected.

P11L16: get rid of word “the” before aquifer formation

Corrected.

P11L26: get rid of word “intuitively”

Corrected.

P13, end: also mention piping perhaps

As Anonymous Referee #1 correctly stated, there was some thematic overlap between this manuscript and Steger et al. (2017). We removed this overlap and are thus not mentioning heterogeneous percolation (piping) in the conclusion but refer to Steger et al. (2017), where this process is already suggested in the conclusions as a potential model improvement.

Figure 3 caption: change to “shaded areas illustrate”

The shaded areas (interannual variability) were removed from the figure because they are not discussed in the main text.

Figure 5: It might be nice to have basin numbers labelled on this figure to avoid flipping to figure 2. Why is refreezing the only parameter you show fraction for? You could, for each LWB component, show the value (as you do) with the fraction of total in parentheses. So, for example, basin 1 in panel e would have the value 14 (30%). That way the reader could see the fraction for all of the components. Also on this figure: The outlines of the basins are tough to see (make them darker and make the ELA a dashed dark line?). The colour scales make it tough to see what is going on near the margins – not sure how to fix that.

Basin numbers are now provided in panel (a). As the reviewer suggests, it would indeed be possible to show other parameters than refreezing as fractions (e.g. melt as a fraction of the total liquid water input). However, as we are focusing on liquid water retention in snow/firn, we think that showing fractions for other values than refreezing would distract from the focus.

To enhance visibility of the basin's boundaries, we increased the line width and selected a darker grey for the lines. We agree with the reviewer that the colour scale makes it difficult to recognise patterns near the margins. This issue is caused by the fact that most plotted parameters exhibit the highest gradients close to the margins. The use of a non-linear colour scale would improve readability but it would also distort the presented values. Therefore, we decided to keep the current colour scale.

Figure 6: the legend could be larger at the bottom, especially with the thickness of the lines

Corrected (Fig. 14 was improved likewise).

Figure 8: Perhaps move the x-labels to the bottom so that the numbers and labels are next to one another. It is a bit confusing – is this showing the change in the average values for 1960-1989 subtracted from the average values for 1990 – 2014? Also on this figure: legend could be larger

We shifted the units to the bottom of the panels and increased the legend according to the reviewer's suggestion. The figure caption was also adapted to clarify the description of the presented data:

Temporal changes in refreezing, firn air content, firn temperature (averaged over 2–10 m depth) and skin temperature in 100 m elevation bins. The difference shows the 1990–2014 average minus the 1960–1989 average.

Figure 10: An additional panel showing the elevation/surface profile of the ice sheet could be useful here.

We refer the reviewer to the axis at the bottom of Fig. 10, which shows the surface elevation of the ice sheet. We agree that an additional panel with a plot of the surface elevation would enhance readability but we omit this panel to keep the vertical size of the figure in a reasonable range.

Figure 11: inset panels are quite small.

We enlarged the inset panels in (b) and (c), increased the line width and made all fonts bold to enhance readability.

Figure 13: Put a label along the colour bar: Firn Air Content (m). The dots here make this a bit tough to understand. Perhaps making the aquifer dots more contrasting colours? Or use crosses? The grey dots are tough to see.

We added a label along the colour bar as suggested by the reviewer. Additionally, we changed the colours of the dots to increase readability. We also changed the colour bar for firn air content to grey-scale, because this quantity is a rather secondary information in this figure (compared to the occurrence of firn aquifers).

Table 1: The units are difficult to understand here. Does m w.e. a^{-1} (25 a^{-1}) refer to melt, rainfall, runoff, and refreezing? Since you state in the caption that these trends are 1990 – 2014, does it not suffice to say the trend is in m w.e. per year ?

We agree with the reviewer that the units are somewhat difficult to read. Therefore, we converted the numerical values to units that are more intuitive:

- Trends in mass fluxes: $(\text{mm w.e. a}^{-1}) \text{ a}^{-1}$
- Trend in refreezing fraction: $\% \text{ a}^{-1}$

Additional references

Mottram, R., Boberg, F., Langen, P., Yang, S., Rodehacke, C., Christensen, J. H., and Madsen, M. S.: Surface mass balance of the Greenland ice sheet in the regional climate model HIRHAM5: Present state and future prospects, *Low Temperature Science*, 105:115, doi: 10.14943/lowtemsci.75.105, 2017.

Schmidt, L. S., Aðalgeirsdóttir, G., Guðmundsson, S., Langen, P. L., Pálsson, F., Mottram, R., Gascoin, S., and Björnsson, H.: The importance of accurate glacier albedo for estimates of surface mass balance on Vatnajökull: evaluating the surface energy budget in a regional climate model with automatic weather station observations, *The Cryosphere*, 11, 1665-1684, doi: 10.5194/tc-11-1665-2017, 2017.

The modelled liquid water balance of the Greenland Ice Sheet

Christian R. Steger¹, Carleen H. Reijmer¹, and Michiel R. van den Broeke¹

¹Institute for Marine and Atmospheric Research Utrecht (IMAU), Utrecht University, Utrecht, the Netherlands

Correspondence to: Christian R. Steger (C.R.Steger@uu.nl)

Abstract. Recent studies indicate that the surface mass balance will dominate the Greenland Ice Sheet's (GrIS) contribution to 21st century sea level rise. Consequently, it is crucial to understand the liquid water balance (LWB) of the ice sheet and its response to increasing surface melt. We therefore analyse a firn simulation conducted with SNOWPACK for the GrIS and over the period 1960–2014 with a special focus on the LWB and refreezing. **Indirect evaluations** of the simulated refreezing climate with GRACE and firn temperature observations indicate a good **model-observation agreement**. Results of the LWB analysis reveal a spatially uniform increase in surface melt (**0.16 m w.e. a⁻¹**) during 1990–2014. As a response, refreezing and runoff also **indicate positive changes during this period (0.05 m w.e. a⁻¹ and 0.11 m w.e. a⁻¹, respectively)**, where refreezing increases **at** only half the rate of runoff, which implies that the majority of the additional liquid input runs off the ice sheet. However, this pattern is spatially variable as e.g. in the southeastern part of the GrIS, most of the additional liquid input is buffered in the firn layer due to relatively high snowfall rates. The increase in modelled refreezing leads to a decrease in firn air content and to a substantial increase in near-surface firn temperature in some regions. On the western side of the ice sheet, modelled firn temperature increases are highest in the lower accumulation zone and are primarily caused by the exceptional melt season of 2012. On the eastern side, simulated firn temperature increases more gradually and with an associated **migration of perennial firn aquifers to higher elevations**.

1 Introduction

The mass balance (MB) of the Greenland Ice Sheet (GrIS) has been negative since the early 1990s (Van den Broeke et al., 2016). Besides increased ice discharge through the acceleration of marine-terminating outlet glaciers, the ice sheet is losing mass through increased surface melt and associated meltwater runoff. The latter process has recently become the dominant contributor to mass loss from the ice sheet (Enderlin et al., 2014). The increase in meltwater runoff and associated decrease of the surface mass balance (SMB) is attributed to processes on various spatial and temporal scales (**e.g. the polar amplification (Bekryaev et al., 2010) and the darkening of the GrIS (Tedesco et al., 2016)**) and is further promoted by the hypsometry of the ice sheet (Mikkelsen et al., 2016; Van As et al., 2017). An accurate quantification of the liquid water balance (LWB) of the ice sheet is important, as it determines how much of the liquid input at the surface ultimately reaches the ocean and contributes to sea level rise. A key parameter of the LWB is meltwater storage in the firn (Rennermalm et al., 2013a) by refreezing and liquid water retention. Previous studies suggest that **modelled** refreezing strongly depends on the model formulation (Reijmer et al., 2012; Steger et al., 2017) and that it exhibits the largest inter-model variation of all SMB components (Vernon et al., 2013).

Besides the instantaneous effect of retaining liquid water, refreezing also co-determines the future potential of firn to absorb melt, as it reduces the porosity of the firn (Noël et al., 2017) and releases large amounts of latent heat (Humphrey et al., 2012; Cox et al., 2015), which decreases the firn's cold content.

The hydrology of the GrIS is a complex system, which involves various ill-constrained processes (Fig. 1). At the surface, liquid input is determined by rainfall, evaporation/condensation and melt. In areas where the ice sheet is covered by snow and/or firn, liquid water is able to percolate vertically. These snow/firn layers may act as a buffer for runoff if liquid water either refreezes (Harper et al., 2012) or remains in its liquid state in **perennial** firn aquifers (Forster et al., 2014). Such aquifers typically form at locations with relatively high amounts of snow accumulation (Kuipers Munneke et al., 2014) and are thus particularly abundant along the southeastern and northwestern margins of the ice sheet (Forster et al., 2014). A recent study (Poinar et al., 2017) revealed that some aquifers likely drain into crevasses. To what degree the water refreezes there or reaches the bed of the ice sheet remains largely unknown. Along the southwestern and northeastern margins of the ice sheet, firn aquifers are less abundant. In these areas, percolating water typically refreezes in the firn or runs off over the ice surface. A study by Machguth et al. (2016) suggests that horizontal ice layers could inhibit vertical percolation and render underlying pore space inaccessible for liquid water. The water would hence be forced to flow laterally above such obstacles - either as surface runoff or within the firn.

In the bare ice zone, hydrological processes are better understood: Liquid water flows along surface rivers and may accumulate in supra-glacial lakes (Arnold et al., 2014) or enter the subglacial system via moulins or crevasses. The amount of water stored in supra-glacial lakes is thereby rather small compared to the magnitude of supra-glacial river fluxes, which drain liquid water efficiently from the surface (Smith 2015). Liquid water flowing into moulins or crevasses enters the en- and subglacial (Lewis and Smith, 2009; Lindbäck et al., 2015) hydrological system of the ice sheet. Here, water may refreeze, accumulate in subglacial lakes or flow along channels to the margins of the ice sheet. The relevance of en- and subglacial water storage is currently rather uncertain. Rennermalm et al. (2013b) suggests that for a watershed in southwestern Greenland, up to 54 % of meltwater may be retained during one season. **It is however possible** that this residual is partly caused by uncertainties in e.g. watershed delineation (Rennermalm et al., 2013b) and inter-basin piracy (Lindbäck et al., 2015). A more recent study for a similar catchment yielded little evidence for meltwater storage in en- and subglacial environments (Van As et al., 2017). In summary, the hydrology of the GrIS represents a complex system of pathways that transport meltwater from the surface of the ice sheet to the ocean (Chu, 2014).

In this study, we quantify the components of the LWB from the GrIS surface to the firn-ice-transition, using a state-of-the-art snow/firn model. The upper boundary conditions for the model are provided by the regional atmospheric climate model RACMO2.3 (Noël et al., 2015). Potential en- and subglacial liquid water retention is not considered as we only model the upper part of the ice sheet. The primary goal is to quantify the spatial magnitude of the different LWB components and assess how these mass fluxes evolved over the last decades. Additionally, we evaluate the spatial and seasonal occurrence of refreezing and the impact of this process on **firn density and temperature**. Furthermore, we analyse how the horizontal extent of firn aquifers, which act as perennial storage for liquid water, evolves with time. The following section provides a brief description of the model and the observational data used in this study. Subsequently, we **discuss the comparison of model output with** remote

sensing data (GRACE) and in situ measurements (firn temperatures). Section 4 contains the results of the LWB evaluation and a more detailed analysis of refreezing, runoff and changes in different firn properties.

2 Definitions, model and data

2.1 Definitions

- 5 In this study, we investigate the LWB of the upper part of the ice sheet, namely the snow/firn layer. This layer ranges from the surface down to the firn–ice–transition. If percolating water reaches the bottom of this domain, it is considered to leave the ice sheet as runoff. Potential en- and subglacial storage of liquid water are thus not accounted for. The LWB of the firn layer is defined as

$$\frac{dM_{ret}}{dt} = RA - EV + ME - RF - RU, \quad (1)$$

- 10 where M_{ret} is the retained liquid mass, RA , EV and ME are surface mass fluxes of rainfall, evaporation and meltwater respectively, RF is internal refreezing and RU is runoff at the bottom of the model domain. In this study, the term evaporation refers to phase changes of water from liquid to gaseous (evaporation) and vice versa (condensation). The SMB used in this study equals the climatic mass balance (Cogley et al., 2011), i.e. it includes subsurface processes of liquid water retention and refreezing, and is defined as

$$15 \quad SMB = RA + SN - EV - SU + SD - RU, \quad (2)$$

where SN is snowfall, SU sublimation (and resublimation) and SD deposition or erosion by snow drift. The SMB is linked to the LWB through the components rainfall, evaporation and runoff.

The Greenland mass balance (MB) derived to validate the modelled SMB with GRACE data is defined as

$$MB = SMB_{GrIS} + SMB_{PIC} - D + \frac{dM_{ts}}{dt}, \quad (3)$$

- 20 where SMB_{GrIS} and SMB_{PIC} are the SMB of the glaciated area (GrIS and peripheral ice caps/glaciers), D is ice discharge across the grounding line from marine-terminating glaciers and M_{ts} is the tundra snow mass.

2.2 Model data

- Snow/firn on the GrIS and the peripheral ice caps/glaciers is modelled with SNOWPACK (version 3.30), a state-of-the-art snow model. SNOWPACK was recently applied in different studies (Groot Zwaaftink et al., 2013; Van Tricht et al., 2016; Steger et al., 2017) to simulate snow and firn in polar regions. The model contains an overburden-dependent densification scheme and simulates the evolution of different microstructural snow properties, which are linked to thermal and mechanical snow quantities (Bartelt and Lehning, 2002; Lehning et al., 2002b, a). We run SNOWPACK on an 11 km horizontal grid and with the same ice mask (Fig. 2) as used in the regional atmospheric climate model RACMO2.3 (Noël et al., 2015). At

the snow–atmosphere interface, SNOWPACK is forced with mass fluxes (precipitation, evaporation/sublimation, snow drift and surface melt) and with skin temperature from RACMO2.3. Skin temperature is the temperature of an infinitesimally thin layer without heat capacity, and is representative for surface temperature. The capability of RACMO2.3 to accurately simulate present-day surface climate on the GrIS was illustrated in an extensive evaluation by Noël et al. (2015). Vertical water percolation is simulated with a bucket scheme (Bartelt and Lehning, 2002; Wever et al., 2014) and the irreducible water content follows the formulation of Coléou and Lesaffre (1998). We do not consider heterogeneous percolation (Wever et al., 2016; Marchenko et al., 2017) in our simulation due to an insufficient spatial coverage of observational data to calibrate such routines for the entire ice sheet and/or the too expensive computational demand. Neglecting heterogeneous percolation causes refreezing to occur mostly in the upper snowpack, where temperature and porosity are determined by the recent climate. Lateral flow of runoff is also not considered in our simulation. Fresh snow density is prescribed with an empirical parameterisation that depends on mean annual surface temperature (Kuipers Munneke et al., 2015). The enhanced near-surface snow compaction due to strong winds, which is implemented in SNOWPACK for Antarctic simulations (Groot Zwaaftink et al., 2013), is switched off, because the applied fresh snow density parameterisation already accounts for this effect. A more detailed description of the model setup and the applied spin-up procedure is stated in Steger et al. (2017), where the same SNOWPACK run was used.

2.3 Observational data

To derive a MB for Greenland, we use ice discharge data from Enderlin et al. (2014) and a GRACE gravity field solution for Greenland (Groh and Horwath, 2016). The ice discharge data contain annual estimates of ice discharge from 178 marine-terminating glaciers wider than 1 km and are available for the period 2002–2012. Following Van den Broeke et al. (2016), we neglect seasonal variations in ice discharge and assume that all intra-annual variation in the MB is induced by components of the SMB or by tundra snow. The GRACE data we apply are based on the monthly GRACE solution ITSG-Grace2016 (Mayer-Gürr et al., 2016) and is available between mid-2002 and mid-2016. We computed the MB for the overlapping period 2003–2012 where data are available from all sources throughout the year.

We use firn temperatures that were recorded along a 2700 km transect in northwest Greenland (Fig. 2), referred to as the NW GrIS transect, to evaluate our simulation. Shallow borehole temperature measurements were conducted at 14 sites between 1952–1955 (Benson, 1962) and repeated in 2013 (Polashenski et al., 2014). The former measurements were taken at a range of 3 to 16.75 m depth (predominantly at 8 m) and were corrected for seasonal influences to obtain an intercomparable, mean annual 10 m temperature. The measurements in 2013 were recorded at a depth between 5–12 m (mainly at 8.5–12 m) and were corrected with the same methodology (Polashenski et al., 2014).

3 Model evaluations

Although modelled refreezing cannot directly be evaluated with observations, Steger et al. (2017) made a comprehensive assessment of modelled firn density, which is the combined result of dry compaction and refreezing. Results show a reasonable performance of SNOWPACK, but a general overestimation of densities in the percolation zone. This bias is likely the result of

overestimated near-surface refreezing caused by neglecting heterogeneous water percolation, an overestimation of fresh snow density and errors in the atmospheric forcing (Steger et al., 2017). In this study, we use additional observations to evaluate the ability of SNOWPACK forced by RACMO2.3 to simulate the LWB and particularly the refreezing climate of the GrIS. Due to the lack of direct refreezing observations, we assess the model's performance in terms of refreezing indirectly by comparing the modelled spatially integrated SMB and local snow/firn temperatures to observations.

3.1 Model evaluation using GRACE

Due to the large footprint of GRACE, the signal also contains mass variations from Greenland's peripheral ice caps and glaciers and from tundra hydrology; primarily from seasonal snow cover. These signals are thus included in Greenland's MB as explained in Sect. 2.1. Tundra snow cover is not simulated by SNOWPACK but the signal is taken from RACMO2.3 output. In RACMO2.3, seasonal snow is simulated with a single-layer model that does not allow for refreezing and liquid water retention in the snow (Van den Broeke et al., 2016). All surface melt is hence immediately transferred to runoff.

A comparison between the derived cumulative MB and GRACE is provided in Fig. 3a. The MB is computed by taking the simulated SMB over the glaciated area either from RACMO2.3 or SNOWPACK. Both cumulative MBs indicate an excellent agreement with GRACE ($R^2 > 0.99$). In terms of linear trends, SNOWPACK agrees better with GRACE due to higher modelled refreezing fractions and thus lower amounts of runoff from the ice sheet. The detrended mean seasonal cycles (Fig. 3b) indicate a good agreement in winter and spring, when changes in cumulative SMB are mainly caused by accumulation of solid precipitation on the glaciated area and the tundra. From May on, the derived MBs show an earlier and steeper decrease compared to the GRACE signal. The minima in the MBs occur both earlier and with higher magnitudes than in GRACE, where SNOWPACK performs slightly better due to smaller amounts of modelled runoff. These findings are consistent with earlier studies (Van Angelen et al., 2014; Alexander et al., 2016), in which the average seasonal cycle of the MB and GRACE were compared. A likely contributor to this mismatch is the neglect of the time it takes meltwater runoff to reach the ocean. Van Angelen et al. (2014) demonstrated that the monthly error between detrended modelled SMB and GRACE on a GrIS-wide scale could be minimised by delaying simulated runoff by 18 days. A study for a catchment in southwestern Greenland revealed that transit times up to 10 days are required to align the modelled surface runoff and observed river hydrograph optimally (Van As et al., 2017).

Another uncertainty arises from modelled tundra snow cover and tundra hydrology. The mean seasonal amplitude of the detrended modelled MBs derives by $\sim 30\%$ from winter accumulation and summer melting of seasonal snow over the tundra (Fig. 3b). A too-early snow ablation in the tundra could hence also contribute to the bias between MBs and GRACE. This assumption is supported by a comparison of the simulated snow cover fraction (SCF) with MODIS/Terra Snow Cover data (Hall and Riggs, 2016), which revealed a too early decrease in modelled SCF in most basins (not shown). Potential causes for this bias are the neglect of refreezing and liquid water retention in the relatively simple RACMO2.3 snow model and the poor representation of tundra topography at a horizontal resolution of 11 km. Additionally, heterogeneous snow distribution on a subgrid scale could also contribute to the bias (Aas et al., 2017). Finally, runoff may also be retained in the hydrological system

of the tundra by refreezing in soil, ponding on frozen ground (Johansson et al., 2015), accumulating in surface lakes (Mielko and Woo, 2006) and storage in terrestrial aquifers. All these processes are currently not represented in our model framework.

3.2 Model evaluation with firn temperature measurements

ERA40 reanalysis data, which forces SNOWPACK via RACMO2.3, is available from 1958 onwards. SNOWPACK output for the years 1952–1955, when the first firn temperature data set along the NW GrIS transect was collected, is thus not available. However, it is reasonable to assume only small changes in firn temperature between 1952–1955 and the start of the SNOWPACK simulation. We therefore compare the 1952–1955 observations to modelled firn temperatures from 1960. Figure 4 shows that SNOWPACK forced by RACMO2.3 slightly overestimates firn temperatures in the higher elevated part of the transect for both periods. For the first period, this bias may be partly caused by the spin-up procedure of the model, where the model is looped over the reference period (1960–1979) to generate the initial firn profile (Steger et al., 2017). This means that surface temperature evolutions before this reference period are not considered. The bias for the second period is more difficult to explain in the absence of continuous firn temperature measurements and firn density records. The spatially incoherent firn temperature change (between B 4-225 and B 4-000) in the observations is not reproduced by SNOWPACK, which simulates a uniform temperature increase of $\sim 0.3^\circ$. This incoherency in the observations may be partly explained by uncertainties in the measurements caused by errors in the sensor calibration and uncertainties in the applied correction used to retrieve 10 m firn temperature from shallower measurements (Polashenski et al., 2014).

Between locations B 2-175 and B 2-070, there is a ~ 1.6 – 2.7° C warming in the observations between 1952–1955 and 2013, likely caused by latent heat release due to refreezing. This temperature increase is larger than the modelled, spatially rather uniform warming of $\sim 0.5^\circ$ C. Possible explanations for this bias are the underestimation of meltwater production at the surface or a too shallow refreezing depth, which enables the released heat to be conducted upwards to the surface and escape to the atmosphere through emission of longwave radiation. In SNOWPACK, percolating water is not allowed to pass unhindered through layers with refreezing capacity, where in reality, liquid water may move to greater depth through heterogeneous percolation (Humphrey et al., 2012; Marchenko et al., 2017). At site B 1-010, SNOWPACK simulates a local maximum in firn warming, in agreement with observations. Here, RACMO2.3 simulates a doubling of the liquid water input between the two periods considered. However, the magnitude of warming in SNOWPACK is somewhat smaller (4.1° vs. 5.7° C), which may again be linked to the neglect of heterogeneous percolation. Along the entire transect, modelled increases in solid precipitation are spatially rather uniform and small (~ 0.02 m w.e. a^{-1}), and therefore likely less relevant for explaining changes in firn temperature.

Other snow/firn temperature records are available from the Greenland Climate Network (GC-Net; Steffen and Box 2001) and for the western percolation zone (Humphrey et al., 2012; Charalampidis et al., 2016). The latter two data sets, which cover the periods 2007–2009 and 2009–2013, also indicate substantial warming of the upper ~ 10 m firn caused by latent heat release from refreezing. Firn simulations with the Institute for Marine and Atmospheric Research Utrecht Firn Densification Model (IMAU-FDM; Kuipers Munneke et al. 2015) and SNOWPACK (Steger et al., 2017), forced by RACMO2.3, do not reproduce the strong warming observed at these locations. The reason is the overestimation of the bare ice zone on the western GrIS by

the IMAU-FDM and SNOWPACK; i.e. the models are incapable of simulating the subsurface warming due to a deficiency of pore space for refreezing. Compared to IMAU-FDM, the overestimation of this zone is less pronounced in SNOWPACK owing to a different densification scheme, which is more accurate for relatively warm conditions (Steger et al., 2017). Inferring the bare ice zone from remote sensing data, e.g. by using the different surface properties of snow and ice, is complicated due to formation of near-surface ice layer (Machguth et al., 2016) above porous firn. At higher elevations in western Greenland, SNOWPACK does simulate a pronounced warming of the firn layer. Unfortunately, the subsurface temperature data recorded at GC-Net stations located in this area (DYE-2, Crawford Point 1 & 2 and GITS) suffer from large data gaps and/or unphysical high-frequency fluctuations caused by sensor deterioration (K. Steffen, personal communication). The data are thus of insufficient quality to verify these changes.

To address the above-mentioned model bias in overestimating the bare ice zone, we briefly assessed fresh snow density, which is a rather uncertain factor in our simulation. The empirical relation (Kuipers Munneke et al., 2015) we use to obtain this quantity was derived with samples from the dry snow zone and is subsequently extrapolated to lower elevation on the ice sheet. Snow/firn density profiles from a transect on the western GrIS (Harper et al., 2012) allow a comparison between observed and modelled near-surface densities: Averaging over the upper 50 cm and all samples yields a value of $\sim 345 \text{ kg m}^{-3}$ for April (i.e. before the onset of seasonal surface melt). For these locations, our fresh snow density parameterisation returns a mean density of $\sim 405 \text{ kg m}^{-3}$. The parameterisation, which accounts for near-surface densification due to wind and vapour fluxes, clearly overestimates fresh snow density for this region. A comparison of our fresh snow density parameterisation with near-surface snow density samples obtained on the northern GrIS and for spring (Koenig et al., 2016) supports the assumption that the applied parameterisation yields too high densities for comparably warm climate conditions. To test SNOWPACK's sensitivity to initial snow densities, an experiment with a lower, spatially uniform fresh snow density of 320 kg m^{-3} was carried out for the western GrIS transect. The selected initial density is comparable to what the recently published parameterisation of Langen et al. (2017) yields for this transect. With this model setting, the mismatch between the observed and modelled bare ice zone extent (and thus the firn warming) was reduced for this specific region. This model inaccuracy should be addressed in future by testing available or newly derived fresh snow density parameterisations with SNOWPACK for various climate conditions on the GrIS.

4 Climatology of the liquid water balance

The evaluations of mass changes and firn temperatures with observations presented in the previous sections inspire sufficient confidence to use SNOWPACK firn data for a description of the LWB of the GrIS. First, we discuss the mean fields and temporal evolution of the LWB components during the simulation period (1960–2014). Subsequently, refreezing, one of the key components of the balance, and its dependency and influence on firn density and temperature is discussed in more detail. Finally, we analyse the temporal evolution of perennial firn aquifer extent and the partitioning of runoff from ice and snow/firn.

4.1 The liquid water balance

Figure 5 shows the temporally averaged (1960–2014) LWB components for the GrIS and the peripheral ice caps and glaciers. Mean fluxes of rainfall and evaporation are typically at least one order of magnitude smaller than melt, runoff and refreezing. Changes in the retained liquid mass (dM_{ret}/dt) are even smaller than components on the right-hand side of Eq. (1), particularly when integrated over basins, and are thus not presented. Rainfall rates are particularly significant along the southern margin of the ice sheet and in the western ablation zone. For the northeastern part of the GrIS, the contribution of rainfall to the LWB is small and liquid water input at the surface is dominated by melt. The highest melt rates on the GrIS occur along the western ablation zone with a maximum of 129 Gt a^{-1} for Basin 6. The mean spatial runoff pattern is comparable to the one of melt but attenuated by the buffering effect of refreezing. Runoff also peaks in Basin 6 with a value of 85 Gt a^{-1} , which accounts for a third of the total GrIS runoff. Averaged over the entire ice sheet, SNOWPACK simulates that almost half (47 %) of the liquid water input at the surface refreezes in snow or firn. This fraction has a high spatial variability and is relatively low for the northeastern basins and for Basin 6, where precipitation is low and bare ice extent relatively large. As a result, refreezing rates in these regions peak more inland in the lower accumulation zone just above the equilibrium line. Refreezing in the ablation zone is, in terms of absolute liquid water retention, only relevant on intra-annual scales. The highest overall refreezing fractions, up to 75 %, are modelled along the wet southeastern margin of the ice sheet (Basin 4).

Time series of the four most relevant LWB components (melt, runoff, refreezing and rainfall) for the eight basins show no distinctive trends for the first half of the simulation period (1960–1989), but do exhibit large interannual variability, particularly for surface melt (Fig. 6). For the second half (1990–2014) however, there is a statistically significant increase in melt in all basins (Table 1). Changes are particularly large for Basin 5 and 6, where melt increases by $0.36 \text{ m w.e. a}^{-1}$ and $0.38 \text{ m w.e. a}^{-1}$, respectively. The dominant cause for these large changes is the comparably high increase of melt in the ablation area of the GrIS, especially in the southwest. Modelled snow melt in the ablation zone is particularly sensitive to temperature increases due to the albedo difference between snow and ice, where bare ice with a lower albedo is more rapidly exposed through accelerated melt of snow. The lowered surface albedo subsequently enhances melt of bare ice. A secondary cause is the relatively flat hypsometry of these basins, where 58 % respectively 47 % of the area is below 2000 m a.s.l. (compared to 39 % for the GrIS).

Rainfall, as a further contributor to liquid input, does not exhibit a significant trend for the majority of the basins. Linear trends are comparably high for Basin 5 ($1.22 \text{ mm w.e. a}^{-2}$) and Basin 6 ($0.43 \text{ mm w.e. a}^{-2}$) but statistically insignificant. Remarkably, the northwestern Basin 8 is the only region with a significant positive trend in rainfall of $0.56 \text{ mm w.e. a}^{-2}$. This increase is not caused by a change in total precipitation but by a significant increase of the rainfall fraction in this area. For all basins, melt rates peak in 2012 when the GrIS experienced unprecedented surface melt both in spatial extent (Nghiem et al., 2012) and magnitude. The exposure of relatively high-elevated regions with cold and porous firn to surface melt is the main reason that refreezing also peaks in all basins during this year. In response to the positive trends in melt, runoff also exhibits a significant increase in all basins between 1990 and 2014 (Table 1). The increase in runoff accounts in most basins for more than half of the increase in melt ($\sim 55\text{--}80\%$), i.e. most of the additional melt is not stored in the firn layer but is running off the ice sheet. As for melt, the southwestern Basins 5 and 6 show the strongest increase per area. An exception is Basin 4, where

runoff increases with only $\sim 30\%$ the rate of melt. In terms of refreezing and refreezing fraction, the response of the basins to increased surface melt is spatially less uniform: The majority of the basins do not indicate a significant trend in refreezing. This means that e.g. for Basins 1 and 2, most of the additional melt is not absorbed in the firn but runs off, similar to what happens to northern ice caps not connected to the main ice sheet (Noël et al., 2017). Basin 4, which has the highest overall mean refreezing fraction (75 %), is an exception. Refreezing in this basin shows a distinctive positive trend. This is linked to the high amounts of solid precipitation in this basin, which provide enough pore space to absorb the increase in surface melt. Refreezing is also significantly increasing in the northwestern Basins 7 and 8 but with a lower trend than runoff. Significant trends in the refreezing fraction are only apparent in Basin 1 and particularly in Basin 8, where the fraction decreases by $\sim 16\%$ in 25 years. For the entire GrIS, melt, runoff and refreezing indicate significant positive trends between 1990 and 2014. The increase in runoff is roughly twice the one in refreezing, which leads to a significant decrease in the GrIS-integrated refreezing fraction of $\sim 9\%$ over the 25 years (Table 1). The different responses of the eight basins to increasing surface melt are related to refreezing, which in turn is linked to firn porosity and temperature. This will be discussed in more detail in the following section.

4.2 Refreezing and latent heat release

Refreezing is a process that strongly depends on local climate, i.e. particularly on surface temperature which is the main driver for melt, and therewith on seasonality and elevation (Fig. 7). At the beginning of the melt season, modelled refreezing primarily occurs in the lower parts of the ice sheet, where the melt onset is earliest and meltwater percolates into the cold winter snow layer. For Basin 3–7, low-level refreezing peaks in (late) May while for the northern Basins 1, 2 and 8, the maximum occurs in mid-June. During the course of the melt season, the lower regions are gradually depleted of pore space or cold content and the area of peak refreezing moves upward. For the majority of the basins (e.g., Basin 1, 2 and 7), the availability of pore space is the limiting factor for refreezing at lower elevations (Fig. 7). Particularly for Basin 4 and 5 however, this is not the case. Therefore, refreezing at lower elevations persists throughout the melt season but with lower rates than in spring due to a gradual decrease in the firn cold content. However, even if the entire firn column has become temperate, modelled refreezing persists due to the diurnal temperature cycle, which periodically refreshes the near-surface cold content during night. This underlines the importance of using atmospheric forcing data that resolve variations on subdaily time scales. Peak refreezing rates pass the equilibrium line altitude in June (western Basins 6–8) or July (northern Basins 1 and 2). For Basins 3–5, it is not possible to define a mean equilibrium line altitude, because these basins have a very narrow ablation zone and at 11 km resolution, many model grid cells close to sea level have a positive SMB due to high accumulation rates. In July, peak refreezing moves beyond the runoff line in all basins. Especially Basins 4 and 5 reveal a percolation zone that stretches over a relatively large vertical extent, with substantial refreezing as high as 500–750 m above the modelled runoff line. A further notable feature of refreezing is its seasonal asymmetry in comparison to melt (Fig. 7). Melt peaks in July in all basins and the seasonal increase and decrease are rather symmetric around this maximum. Refreezing on the other hand peaks at the beginning of the melt season in all basins and at all elevations. As mentioned above, this is mainly caused by the decrease of either pore space or cold content during the melt season. A similar feature was found by Cullather et al. (2016) in the regional climate model MAR, when seasonal

runoff is plotted as a function of melt area. Runoff was thereby found to be higher in the second half of the melt season. As Cullather et al. (2016) states, this is an important finding for refreezing or runoff parameterisations that do not take seasonality into account.

Refreezing rates increase in all basins during 1990–2014 (Fig. 6), but not always at a significant level (Table 1). Generally, increases in refreezing are restricted to elevations above 1000 m a.s.l. in all regions (Fig. 8a). The peak of this increase is around 1500 m a.s.l. for the northern Basins 1, 2 and 8 and at higher elevations for the more southerly located regions. This increase is primarily caused by a gradual expansion of the melt area to higher elevations, which causes melt and refreezing to occur in formerly dry snow/firn. The increase in refreezing induces both a corresponding decrease in the modelled firn air content (Fig. 8b) and an increase in firn temperature (up to 4° C; Fig. 8c) due to latent heat release. Particularly for Basins 4 and 5, firn air content decreases also in areas below 1000 m a.s.l.. This reduction is not related to changes in refreezing, but rather caused by increases of melt and the subsequent transformation of formerly porous firn to bare ice. In contrast to other basins, Basin 2 reveals a relatively constant firn temperature increase at lower elevations. This increase is not only caused by the rather small increase in refreezing and the associated latent heat release, but also by an enhanced vertical heat flux from the surface through an increase in surface temperature (Fig. 8d). Surface temperatures changes show a distinct spatial variability, with the largest increases occurring in the northeastern part of the ice sheet, where temperature increases by more than 1.5° C.

The exceptional melt season of 2012 has an even stronger influence on firn temperatures according to our model simulation: Fig. 9a shows the corresponding refreezing anomaly for this year and Fig. 9b the resulting increase in firn temperature. Almost the entire ice sheet experienced exceptional refreezing rates above the equilibrium line, particularly in the southern area of the GrIS where refreezing anomalies up to +0.8 m w.e. a⁻¹ are modelled. The increase in firn temperature largely reflects the refreezing anomaly, with the strongest warming (6° C and higher) being modelled in the southwestern percolation zone of the ice sheet. Unfortunately, no observational data are available to confirm the pronounced warming. The closest available record is from the KAN_U automatic weather station of the Greenland Analogue Project (GAP) and the Programme for Monitoring the Greenland Ice Sheet (PROMICE), which is located in the lower accumulation zone of Basin 6. There, firn temperature increased by approximately 4.7° C during 2012 (Charalampidis et al., 2016). To discuss changes in the vertical firn properties over the simulation period in more detail, we present cross-sections of firn density, temperature and volumetric water content along a southern GrIS transect (Fig. 2) for the beginning of the simulation period (April 1960, Fig. 10) and as relative changes for the end (April 2014, Fig. 11). In 1960, the transition from bare ice to porous firn is modelled around station KAN_U on the western side of the ice sheet. On the eastern side, no bare ice zone has formed due to the high accumulation rates in this region. Increasing accumulation rates from west to east induce the downward bending of high-density layers east of the ice sheet divide (Fig. 10a). The larger amount of pore space on the eastern side permits larger refreezing fractions, which leads, through release of latent heat, to temperate firn condition close to the margin of the ice sheet and to the formation of a perennial firn aquifer (Fig. 10c).

During the 55 years of the simulation, the firn layer along this transect experienced some distinctive changes: near-surface density increased both on the eastern and western side of the ice sheet (Fig. 11a) with a shift of the transition between bare ice and porous firn to higher elevations on the western side. For 2012, SNOWPACK simulates a bare ice profile for KAN_U

while a retrieved density profile for this year revealed layers with porous firn (Charalampidis et al., 2016). A potential cause for this overestimation in firn density is discussed in Sect. 3.2. Modelled firn temperature increased substantially on both sides of the GrIS but with different patterns (Fig. 11b). In the west, 15 m firn temperature in the percolation zone is relatively stable until 2010 and abruptly increases afterwards by $\sim 10^\circ \text{C}$, particularly due to the exceptional melt season of 2012. Refreezing during this year induces a substantial warming of the firn down to a depth of 40 m. At the eastern side, the initial 15m firn temperature is higher by $\sim 5^\circ \text{C}$ and the simulated warming of the firn is more gradual. A major reason for the less pronounced and shallower firn temperature increase on the eastern side of the ice sheet is the temporal distribution of liquid input in 2012 (inset panel Fig. 11c). On the eastern side, liquid input at the surface is rather evenly distributed throughout the melt season. On the western side, there are several distinctive peaks with liquid input up to $\sim 40 \text{ mm w.e. day}^{-1}$. These high fluxes, together with the fact that percolating water is able to bypass layers without pore space in our model, cause the relatively deep maximum in firn warming. On the eastern side, the gradual increase in firn temperature allowed the firn aquifer to expand further inland (Fig. 11c), a process that is discussed in more detail in the following section.

4.3 Perennial Firn aquifer

In accordance with Steger et al. (2017), we classify any firn with a vertically integrated liquid water content of more than 200 kg m^{-2} in April as perennial aquifer, irrespective of water saturation. This is necessary because our model is not able to simulate saturated conditions in the used configuration due to the neglect of impermeable layers. Introducing saturated conditions in SNOWPACK would require a definition of the pore space fraction available for liquid water storage. This quantity is rather uncertain and is assumed to be in the range of 40 % (Jansson et al., 2003) to 100 % (Koenig et al., 2014). The above-mentioned threshold for firn aquifer delineation is based on a sensitivity estimation of the NASA Operation IceBridge accumulation radar to detect liquid water in firn (Miège et al., 2016).

The eastern part of the southern GrIS transect crosses the region where perennial firn aquifers were discovered in 2011 (Forster et al., 2014) and mapped in 2015/2016 (Montgomery et al., 2017). The grey shaded area in Fig. 11 indicates the horizontal extent of these mapped aquifers. The combination of RACMO2.3 and SNOWPACK underestimates the the upper limit of the firn aquifer's horizontal extent by approximately 100 m in elevation if one assumes only small changes in aquifer extent between 2014 and 2016. A brief sensitivity test of SNOWPACK with a lower fresh snow density, as described in Sect. 3.2, yields a firn aquifer that reaches higher elevations and thus reduces the mismatch. The reason for this improvement is that the lower near-surface firn density reduces the conductive heat loss of the aquifer to the atmosphere in winter. The lower limit of observed firn aquifers in this area is around 1520 m a.s.l., which coincides with crevasses in the ice stream. It has recently been demonstrated that firn aquifers do not exist in such regions because liquid water drains into crevasses, where water either refreezes or enters the subglacial drainage system (Poinar et al., 2017). This feature is not included in our model framework, which is why SNOWPACK models the presence of aquifers at lower elevations. Note that smaller firn aquifers have been mapped downstream of the crevasse fields (Poinar et al., 2017).

Observations of the vertical extent of firn aquifers in this region return an average depth of 16.2 m for the water table and 27.7 m for the aquifer base (Montgomery et al., 2017). Comparing the depth of the modelled firn-aquifer top to observations is

difficult, because observations derived from radar measurements return the depth of the water table and not the transition from dry to wet (but unsaturated) firn. The depth of the water table may roughly be in line with our simulations (Fig. 10 and 11), if an unsaturated wet layer between the aquifer top and the dry firn is assumed. The depth of the modelled firn aquifer base is clearly overestimated (>40 m). This mismatch is likely related to a positive temperature bias of the deeper firn in our simulation (Steger et al., 2017). Observations indicate that refreezing conditions typically prevail at the base of aquifers (Montgomery et al., 2017) and thus suggest to initialise deeper firn with lower temperatures and the application of a of downward-directed heat flux at the bottom of the model domain (Steger et al., 2017). Due to the present setting of SNOWPACK, which does not allow for saturated condition, the observed mean liquid water content of 16 % (Montgomery et al., 2017) is higher than modelled values. Figure 11 shows an expansion of the firn aquifer to higher elevations during the simulation period (1960–2014). This trend is in line with observations (2010–2016), which indicate an inland expansion of aquifers in this area (Miège et al., 2016; Montgomery et al., 2017). Aquifer expansion is also apparent for other regions, where significant firn aquifer areas are modelled by SNOWPACK (Fig. 12). The highest fractions of firn aquifer area are simulated in the southeastern Basins 4 and 5. In both basins, firn aquifers considerably expanded inland with time, particularly in Basin 4. This expansion to higher areas is partially compensated by a decrease of aquifer area at lower elevations, where porous firn is transformed to bare ice by increasing melt amounts (Fig. 8b). In Basins 4 and 5, the mean surface elevation at which firn aquifers are modelled rises by ~ 200 m during the simulation period (1960–2014). Upward migration of firn aquifers is also apparent in other basins, where Basin 3 and 6 reveal a smaller change of 125 and 90 m, respectively, and Basin 8 a larger elevation increase of 215 m. The modelled aquifer extent of $\sim 59,000$ km² for the entire GrIS, average over 2010–2014, is substantially larger than an estimate based on remote sensing data of 21,900 km² for the same period (Miège et al., 2016). Potential causes for this overestimation are discussed in Steger et al. (2017).

The formation of perennial firn aquifers requires specific conditions, i.e. comparably high melt rates during the summer season and high annual accumulation rates (Kuipers Munneke et al., 2014). The dependence of firn aquifers on these parameters is also apparent in our simulation when modelled GrIS grid cells are plotted as a function of snowfall and liquid input (Fig. 13). The occurrence of firn aquifers is thereby restricted to a rather well separated space, which supports the hypothesis that snowfall and liquid input are the principal predictors for aquifer formation. The period of 1960–1979 has been selected for this analysis because it is identical to the spin-up period of our simulation. The relation is thus computed for steady-state climate without any long-term trends in the forcing. To assess the influence of a transient climate, firn aquifer occurrence as a function of snowfall and liquid input has also been computed for the period 2010–2014, which is identical to firn aquifer observations by remote sensing (Miège et al., 2016). The zone of modelled aquifers shifts to a region with a higher ratio of liquid input to snowfall. This shift is likely caused by the changing climate conditions, where the spatial firn aquifer extent has not yet equilibrated to the new forcing.

4.4 Runoff partitioning

Runoff from the ice sheet can either originate from melting of bare ice in the ablation zone, in which case runoff is assumed instantaneous, or from melting of snow/firn in the ablation or accumulation zone, in which case meltwater can be retained or

refrozen. Partitioning runoff in these two classes yields insights in basin characteristics and indicates shifts in the accumulation and ablation area extent. Basins with high fractions of snow/firn runoff exhibit **likely** a higher uncertainty in runoff estimates due to potential storage of liquid water at the **source** location or along the routing path, the latter of which is not explicitly modelled. **To distinguish runoff from ice and snow/firn, we apply a threshold for firn air content of 0.02 m.**

- 5 Basins 1 and 2 show relatively similar characteristics in terms of runoff partitioning (Fig. 14). In these dry northern regions with relatively wide ablation zones, runoff from snow/firn melt is at least an order of magnitude smaller than runoff from ice melt. Both basins reveal a strong positive trend in runoff from ice between 1990 and 2014 with an increase of 30.3 Gt a⁻¹ and 15.6 Gt a⁻¹ over the 25 years. The eastern Basin 3 has a higher runoff fraction from snow/firn than the northern basins. Runoff from snow/firn is increasing in the later period (**2.4 Gt a⁻¹**), although not on a statistically significant level. A very different
- 10 picture emerges from Basin 4, which has a very narrow ablation zone. In this basin, approximately 87 % of runoff originates from snow/firn. Still, there is a **considerable increase in runoff from ice during the second half of the simulation period (2.5 Gt a⁻¹)**, which is caused by the decrease of pore space (Fig. 8b) and the gradual increase of the ablation zone. As a result, the snow/firn runoff fraction in this basin exhibits a significant negative trend (**-9 %**). Basin 5 reveals a similar pattern, but here runoff from ice and snow/firn are comparable in magnitude, particularly towards the end of the simulated period. This basin
- 15 also **reveals** the highest interannual variability in snow/firn runoff fraction, which is related to the high interannual variance of winter (**Oct. - Mar.**) snowfall in this region (**$\sigma = 0.13$ m w.e.**). Variance in **winter** snowfall is also high in Basin 4 (**$\sigma = 0.14$ m w.e.**), but the sensitivity of the snow/firn runoff fraction on winter snowfall is lower due to a smaller ratio of ablation to accumulation area. The westerly Basins 6–8 exhibit comparable runoff partitionings: all three basins are dominated by runoff from bare ice in the ablation zone and all these fluxes reveal a statistically significant positive trend in the second half of the
- 20 simulated period. These trends are particularly strong in Basins 6 and 8, where runoff from ice increases by 54.6 Gt a⁻¹ and 31.5 Gt a⁻¹ over 1990–2014. In the more northerly Basins 7 and 8, there is a small but still significant increase in runoff from snow/firn. For the entire ice sheet, both runoff originating from ice and snow/firn increase at significant rates over the period 1990–2014 by 171.7 Gt a⁻¹ and 25.6 Gt a⁻¹, respectively. The runoff fraction from snow/firn decreased over this time by 6 %.

5 Conclusions

- 25 In this study, we analysed a SNOWPACK simulation carried out for the glaciated area of Greenland and for the period 1960–2014 with a focus on the liquid water balance (LWB) of the firn layer. The model was forced by output from the regional atmospheric climate model RACMO2.3 at the upper boundary. **A comparison of the cumulative MB, derived with modelled SMB values and ice discharge data from observations, indicates an excellent agreement ($R^2 > 0.99$) with GRACE. The linear trend in cumulative MB improves when the SMB of Greenland's glaciated area is simulated by SNOWPACK instead of**
- 30 **RACMO2.3 due to higher refreezing rates in SNOWPACK and thus reduced runoff from the ice sheet.** However, the detrended mean seasonal cycles of these signals reveal significant discrepancies during the melt season. This mismatch can likely be attributed to neglecting runoff transit times and inaccuracies in the modelled tundra (snow) hydrology. The model also agrees well with observed changes in firn temperature along a 2700 km transect in northwestern Greenland and with firn aquifer

occurrence in the southeast. A direct comparison with temperature records from the western percolation zone of the ice sheet is not possible due to an overestimated bare ice zone extent in the model. Among other potential causes, such as climate biases in RACMO2.3, this mismatch is at least partly related to a bias in the fresh snow density parameterisation.

Temporally averaged LWB components over the simulation period (1960–2014) reveal that the balance is dominated by melt,

5 runoff and refreezing in all basins. Modelled changes in retained liquid mass, evaporation and rainfall are typically at least one order of magnitude smaller, even for the more southerly basins. SNOWPACK simulates a mean refreezing fraction of 47 % averaged over the entire ice sheet. This quantity reveals a high spatial variability and is smallest for the northern GrIS (30 %) and largest in the southeast (75 %), where snowfall rates are highest. During the first half of the simulation period (1960–1989), there are no distinctive trends in the components of modelled LWB but this changes for the second half (1990–
10 2014), when surface melt fluxes significantly increase in all basins. **These increases are reflected in runoff, particularly in the southwestern area of the ice sheet where runoff increases by 0.31 m w.e. a⁻¹.** Simulated trends in runoff generally exceed those in refreezing, which implies that the majority of the additional liquid water input runs off and thus contributes to sea level rise. The only exception is Basin 4 in the southeast, where most of the additional liquid input (**~70 %**) is buffered in the firn. The simulated increase in refreezing, which is linked to the gradual expansion of the melt area in all basins, impacts **firn properties**
15 by decreasing firn air content and increasing firn temperature. The exceptional melt in 2012 **particularly** causes a substantial warming of the firn, with a peak in the western percolation zone where modelled firn temperatures averaged over 2–10 m depth locally **increases** by more than 6° C. SNOWPACK also simulates a migration of the perennial firn aquifer area to higher elevations, which is, at least for an area in southwestern Greenland, in line with observations. Partitioning runoff according to its source (melting ice or snow/firn) shows that runoff from ice dominates on the ice sheet scale (78 %), with the highest runoff
20 fractions (87 %) from snow/firn modelled in the southeast of the GrIS. Thus, **this basin** likely exhibits the highest uncertainty in runoff estimates due to possible retention of runoff in snow/firn at the place of origin or along the routing path.

**The evaluation of our SNOWPACK simulation with various in situ and remote sensing observational data revealed several model inaccuracies, which are discussed in Steger et al. (2017). The current study emphasises the uncertainties in the applied fresh snow density parameterisation and the thermodynamic conditions beneath firn aquifers. The positive bias in the applied
25 fresh snow density parameterisation for comparably warm climate conditions may be addressed by a comprehensive sensitivity test of SNOWPACK with different fresh snow density parameterisation for various climatic conditions. A particular focus should be placed on the disentanglement of processes influencing near-surface density (e.g. decrease of snow particle size during wind drift, vapour fluxes) and the statistical or physical representation of these processes in the parameterisation or the snow/firn model. The uncertainties in the simulated thermodynamic conditions beneath firn aquifers may be constrained with
30 the increasing availability of in situ observations. Finally, our study reveals lateral routing of runoff as an additional relevant process that is not considered in our model. This is likely a less relevant issue for horizontal near-surface redistribution of mass and energy, as surface melt typically reaches higher elevated areas later in the season, which means that lower areas are already depleted of pore space and/or cold content and thus do not provide any more storage volume for upstream runoff. Neglecting this process complicates however comparisons of modelled SMB and GRACE on seasonal time scales. This shortcoming will
35 be addressed in a next step by coupling SNOWPACK to an offline routing scheme for the GrIS.**

Data availability. All modelled SNOWPACK data presented in this study are available on request from the authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. Christian R. Steger, Carleen H. Reijmer and Michiel R. van den Broeke acknowledge financial support from the Netherlands Polar Programme (NPP) of the Netherlands Institute for Scientific Research (NWO) and the Netherlands Earth System Science Centre (NESSC). ECMWF at Reading (UK) is acknowledged for use of the Cray supercomputing system. Graphics were made using Python Matplotlib (version 2.0.0) and Affinity Designer (version 1.5.5).

References

- Aas, K. S., Gislås, K., Westermann, S., and Berntsen, T. K.: A Tiling Approach to Represent Subgrid Snow Variability in Coupled Land Surface–Atmosphere Models, *J. Hydrometeorol.*, 18, 49–63, doi:10.1175/JHM-D-16-0026.1, 2017.
- Alexander, P. M., Tedesco, M., Schlegel, N.-J., Luthcke, S. B., Fettweis, X., and Larour, E.: Greenland Ice Sheet seasonal and spatial mass variability from model simulations and GRACE (2003–2012), *The Cryosphere*, 10, 1259–1277, doi:10.5194/tc-10-1259-2016, 2016.
- Arnold, N. S., Banwell, A. F., and Willis, I. C.: High-resolution modelling of the seasonal evolution of surface water storage on the Greenland Ice Sheet, *The Cryosphere*, 8, 1149–1160, doi:10.5194/tc-8-1149-2014, 2014.
- Bartelt, P. and Lehning, M.: A physical SNOWPACK model for the Swiss avalanche warning: Part I: numerical model, *Cold Reg. Sci. Technol.*, 35, 123–145, doi:10.1016/S0165-232X(02)00074-5, 2002.
- 10 Bekryaev, R. V., Polyakov, I. V., and Alexeev, V. A.: Role of Polar Amplification in Long-Term Surface Air Temperature Variations and Modern Arctic Warming, *J. Climate*, 23, 3888–3906, doi:10.1175/2010JCLI3297.1, 2010.
- Benson, C. S.: Stratigraphic studies in the snow and firn of the Greenland ice sheet, Research Report 70, Snow, Ice and Permafrost Research Establishment, 1962.
- Charalampidis, C., As, D. V., Colgan, W. T., Fausto, R. S., Macferrin, M., and Machguth, H.: Thermal tracing of retained meltwater in the lower accumulation area of the Southwestern Greenland ice sheet, *Ann. Glaciol.*, 57, 1–10, doi:10.1017/aog.2016.2, 2016.
- 15 Chu, V. W.: Greenland ice sheet hydrology, *Prog. Phys. Geog.*, 38, 19–54, doi:10.1177/0309133313507075, 2014.
- Cogley, J., Hock, R., Rasmussen, L., Arendt, A., Bauder, A., Braithwaite, R., Jansson, P., Kaser, G., Möller, M., Nicholson, L., and Zemp, M.: Glossary of Glacier Mass Balance and Related Terms, IHP-VII Technical Documents in Hydrology No. 86, IACS Contribution No. 2, UNESCO-IHP, Paris, 2011.
- 20 Coléou, C. and Lesaffre, B.: Irreducible water saturation in snow: experimental results in a cold laboratory, *Ann. Glaciol.*, 26, 64–68, 1998.
- Cox, C., Humphrey, N., and Harper, J.: Quantifying meltwater refreezing along a transect of sites on the Greenland ice sheet, *The Cryosphere*, 9, 691–701, doi:10.5194/tc-9-691-2015, 2015.
- Cullather, R. I., Nowicki, S. M. J., Zhao, B., and Koenig, L. S.: A Characterization of Greenland Ice Sheet Surface Melt and Runoff in Contemporary Reanalyses and a Regional Climate Model, *Front. Earth Sci.*, 4, 10, doi:10.3389/feart.2016.00010, 2016.
- 25 Enderlin, E. M., Howat, I. M., Jeong, S., Noh, M.-J., van Angelen, J. H., and van den Broeke, M. R.: An improved mass budget for the Greenland ice sheet, *Geophys. Res. Lett.*, 41, 866–872, doi:10.1002/2013GL059010, 2014.
- Forster, R. R., Box, J. E., van den Broeke, M. R., Miège, C., Burgess, E. W., van Angelen, J. H., Lenaerts, J. T. M., Koenig, L. S., Paden, J., Lewis, C., Gogineni, S. P., Leuschen, C., and McConnell, J. R.: Extensive liquid meltwater storage in firn within the Greenland ice sheet, *Nat. Geosci.*, 7, 95–98, doi:10.1038/ngeo2043, 2014.
- 30 Groh, A. and Horwath, M.: The method of tailored sensitivity kernels for GRACE mass change estimates, EGU General Assembly, Vienna, Austria, 17–22 April, EGU2016-12065, 2016.
- Groot Zwaartink, C. D., Cagnati, A., Crepaz, A., Fierz, C., Macelloni, G., Valt, M., and Lehning, M.: Event-driven deposition of snow on the Antarctic Plateau: analyzing field measurements with SNOWPACK, *The Cryosphere*, 7, 333–347, doi:10.5194/tc-7-333-2013, 2013.
- Hall, D. K. and Riggs, G. A.: MODIS/Terra Snow Cover Daily L3 Global 0.05Deg CMG, Version 6, doi:10.5067/MODIS/MOD10C1.006, 2016.
- 35 Harper, J., Humphrey, N., Pfeffer, W. T., Brown, J., and Fettweis, X.: Greenland ice-sheet contribution to sea-level rise buffered by meltwater storage in firn, *Nature*, 491, 240–243, doi:10.1038/nature11566, 2012.

- Humphrey, N. F., Harper, J. T., and Pfeffer, W. T.: Thermal tracking of meltwater retention in Greenland's accumulation area, *J. Geophys. Res.-Earth*, 117, F01 010, doi:10.1029/2011JF002083, 2012.
- Jansson, P., Hock, R., and Schneider, T.: The concept of glacier storage: a review, *J. Hydrol.*, 282, 116 – 129, doi:10.1016/S0022-1694(03)00258-0, 2003.
- 5 Johansson, E., Gustafsson, L.-G., Berglund, S., Lindborg, T., Selroos, J.-O., Liljedahl, L. C., and Destouni, G.: Data evaluation and numerical modeling of hydrological interactions between active layer, lake and talik in a permafrost catchment, Western Greenland, *J. Hydrol.*, 527, 688 – 703, doi:10.1016/j.jhydrol.2015.05.026, 2015.
- Koenig, L. S., Miège, C., Forster, R. R., and Brucker, L.: Initial in situ measurements of perennial meltwater storage in the Greenland firn aquifer, *Geophys. Res. Lett.*, 41, 81–85, doi:10.1002/2013GL058083, 2014.
- 10 Koenig, L. S., Ivanoff, A., Alexander, P. M., MacGregor, J. A., Fettweis, X., Panzer, B., Paden, J. D., Forster, R. R., Das, I., McConnell, J. R., Tedesco, M., Leuschen, C., and Gogineni, P.: Annual Greenland accumulation rates (2009–2012) from airborne snow radar, *The Cryosphere*, 10, 1739–1752, doi:10.5194/tc-10-1739-2016, 2016.
- Kuipers Munneke, P., Ligtenberg, S. R., van den Broeke, M. R., van Angelen, J. H., and Forster, R. R.: Explaining the presence of perennial liquid water bodies in the firn of the Greenland Ice Sheet, *Geophys. Res. Lett.*, 41, 476–483, doi:10.1002/2013GL058389, 2014.
- 15 Kuipers Munneke, P., Ligtenberg, S. R. M., Noël, B. P. Y., Howat, I. M., Box, J. E., Mosley-Thompson, E., McConnell, J. R., Steffen, K., Harper, J. T., Das, S. B., and van den Broeke, M. R.: Elevation change of the Greenland Ice Sheet due to surface mass balance and firn processes, 1960–2014, *The Cryosphere*, 9, 2009–2025, doi:10.5194/tc-9-2009-2015, 2015.
- Langen, P. L., Fausto, R. S., Vandecrux, B., Mottram, R. H., and Box, J. E.: Liquid Water Flow and Retention on the Greenland Ice Sheet in the Regional Climate Model HIRHAM5: Local and Large-Scale Impacts, *Front. Earth Sci.*, 4, 110, doi:10.3389/feart.2016.00110, 2017.
- 20 Lehning, M., Bartelt, P., Brown, B., and Fierz, C.: A physical SNOWPACK model for the Swiss avalanche warning: Part III: meteorological forcing, thin layer formation and evaluation, *Cold Reg. Sci. Technol.*, 35, 169 – 184, doi:10.1016/S0165-232X(02)00072-1, 2002a.
- Lehning, M., Bartelt, P., Brown, B., Fierz, C., and Satyawali, P.: A physical SNOWPACK model for the Swiss avalanche warning: Part II. Snow microstructure, *Cold Reg. Sci. Technol.*, 35, 147 – 167, doi:10.1016/S0165-232X(02)00073-3, 2002b.
- Lewis, S. M. and Smith, L. C.: Hydrologic drainage of the Greenland Ice Sheet, *Hydrol. Process.*, 23, 2004–2011, doi:10.1002/hyp.7343, 25 2009.
- Lindbäck, K., Pettersson, R., Hubbard, A. L., Doyle, S. H., van As, D., Mikkelsen, A. B., and Fitzpatrick, A. A.: Subglacial water drainage, storage, and piracy beneath the Greenland ice sheet, *Geophys. Res. Lett.*, 42, 7606–7614, doi:10.1002/2015GL065393, 2015GL065393, 2015.
- Machguth, H., MacFerrin, M., van As, D., Box, J. E., Charalampidis, C., Colgan, W., Fausto, R. S., Meijer, H. A. J., Mosley-Thompson, E., and van de Wal, R. S. W.: Greenland meltwater storage in firn limited by near-surface ice formation, *Nat. Clim. Change*, 6, 390–393, 30 doi:10.1038/nclimate2899, 2016.
- Marchenko, S., van Pelt, W. J. J., Claremar, B., Pohjola, V., Pettersson, R., Machguth, H., and Reijmer, C.: Parameterizing Deep Water Percolation Improves Subsurface Temperature Simulations by a Multilayer Firn Model, *Front. Earth Sci.*, 5, 16, doi:10.3389/feart.2017.00016, 2017.
- 35 Mayer-Gürr, T., Behzadpour, S., Ellmer, M., Kvas, A., Klinger, B., and Zehentner, N.: ITSG-Grace2016 - Monthly and Daily Gravity Field Solutions from GRACE, doi:10.5880/icgem.2016.007, 2016.

- Miège, C., Forster, R. R., Brucker, L., Koenig, L. S., Solomon, D. K., Paden, J. D., Box, J. E., Burgess, E. W., Miller, J. Z., McNerney, L., Brautigam, N., Fausto, R. S., and Gogineni, S.: Spatial extent and temporal variability of Greenland firn aquifers detected by ground and airborne radars, *J. Geophys. Res.-Earth*, 121, 2381–2398, doi:10.1002/2016JF003869, 2016JF003869, 2016.
- Mielko, C. and Woo, M.-k.: Snowmelt runoff processes in a headwater lake and its catchment, subarctic Canadian Shield, *Hydrol. Process.*, 20, 987–1000, doi:10.1002/hyp.6117, 2006.
- Mikkelsen, A. B., Hubbard, A., MacFerrin, M., Box, J. E., Doyle, S. H., Fitzpatrick, A., Hasholt, B., Bailey, H. L., Lindbäck, K., and Pettersson, R.: Extraordinary runoff from the Greenland ice sheet in 2012 amplified by hypsometry and depleted firn retention, *The Cryosphere*, 10, 1147–1159, doi:10.5194/tc-10-1147-2016, 2016.
- Montgomery, L. N., Schmerr, N., Burdick, S., Forster, R. R., Koenig, L., Legchenko, A., Ligtenberg, S., Miège, C., Miller, O. L., and Solomon, D. K.: Investigation of Firn Aquifer Structure in Southeastern Greenland Using Active Source Seismology, *Front. Earth Sci.*, 5, 10, doi:10.3389/feart.2017.00010, 2017.
- Nghiem, S. V., Hall, D. K., Mote, T. L., Tedesco, M., Albert, M. R., Keegan, K., Shuman, C. A., DiGirolamo, N. E., and Neumann, G.: The extreme melt across the Greenland ice sheet in 2012, *Geophys. Res. Lett.*, 39, L20 502, doi:10.1029/2012GL053611, 2012.
- Noël, B., van de Berg, W. J., van Meijgaard, E., Kuipers Munneke, P., van de Wal, R. S. W., and van den Broeke, M. R.: Evaluation of the updated regional climate model RACMO2.3: summer snowfall impact on the Greenland Ice Sheet, *The Cryosphere*, 9, 1831–1844, doi:10.5194/tc-9-1831-2015, 2015.
- Noël, B., van de Berg, W. J., Lhermitte, S., Wouters, B., Machguth, H., Howat, I., Citterio, M., Moholdt, G., Lenaerts, J. T. M., and van den Broeke, M. R.: A tipping point in refreezing accelerates mass loss of Greenland's glaciers and ice caps, *Nat. Commun.*, 8, 14730, doi:10.1038/ncomms14730, 2017.
- Poinar, K., Joughin, I., Lilien, D., Brucker, L., Kehrl, L., and Nowicki, S.: Drainage of Southeast Greenland firn aquifer water through crevasses to the bed, *Front. Earth Sci.*, 5, 5, doi:10.3389/feart.2017.00005, 2017.
- Polashenski, C., Courville, Z., Benson, C., Wagner, A., Chen, J., Wong, G., Hawley, R., and Hall, D.: Observations of pronounced Greenland ice sheet firn warming and implications for runoff production, *Geophys. Res. Lett.*, 41, 4238–4246, doi:10.1002/2014GL059806, 2014GL059806, 2014.
- Reijmer, C. H., van den Broeke, M. R., Fettweis, X., Ettema, J., and Stap, L. B.: Refreezing on the Greenland ice sheet: a comparison of parameterizations, *The Cryosphere*, 6, 743–762, doi:10.5194/tc-6-743-2012, 2012.
- Rennermalm, A. K., Moustafa, S. E., Mioduszewski, J., Chu, V. W., Forster, R. R., Hagedorn, B., Harper, J. T., Mote, T. L., Robinson, D. A., Shuman, C. A., Smith, L. C., and Tedesco, M.: Understanding Greenland ice sheet hydrology using an integrated multi-scale approach, *Environ. Res. Lett.*, 8, 015 017, doi:10.1088/1748-9326/8/1/015017, 2013a.
- Rennermalm, A. K., Smith, L. C., Chu, V. W., Box, J. E., Forster, R. R., Van den Broeke, M. R., Van As, D., and Moustafa, S. E.: Evidence of meltwater retention within the Greenland ice sheet, *The Cryosphere*, 7, 1433–1445, doi:10.5194/tc-7-1433-2013, 2013b.
- Steffen, K. and Box, J.: Surface climatology of the Greenland Ice Sheet: Greenland Climate Network 1995–1999, *Journal of Geophysical Research: Atmospheres*, 106, 33 951–33 964, doi:10.1029/2001JD900161, <http://dx.doi.org/10.1029/2001JD900161>, 2001.
- Steger, C. R., Reijmer, C. H., van den Broeke, M. R., Wever, N., Forster, R. R., Koenig, L. S., Kuipers Munneke, P., Lehning, M., Lhermitte, S., Ligtenberg, S. R. M., Miège, C., and Noël, B. P. Y.: Firn Meltwater Retention on the Greenland Ice Sheet: A Model Comparison, *Front. Earth Sci.*, 5, 3, doi:10.3389/feart.2017.00003, 2017.
- Tedesco, M., Doherty, S., Fettweis, X., Alexander, P., Jeyaratnam, J., and Stroeve, J.: The darkening of the Greenland ice sheet: trends, drivers, and projections (1981–2100), *The Cryosphere*, 10, 477–496, doi:10.5194/tc-10-477-2016, 2016.

- Van Angelen, J. H., van den Broeke, M. R., Wouters, B., and Lenaerts, J. T. M.: Contemporary (1960–2012) Evolution of the Climate and Surface Mass Balance of the Greenland Ice Sheet, *Surv. Geophys.*, 35, 1155–1174, doi:10.1007/s10712-013-9261-z, 2014.
- Van As, D., Bech Mikkelsen, A., Holtegaard Nielsen, M., Box, J., Claesson Liljedahl, L., Lindbäck, K., Pitcher, L., and Hasholt, B.: Hypsometric amplification and routing moderation of Greenland ice sheet meltwater release, *The Cryosphere Discuss.*, 2017, 1–30, doi:10.5194/tc-2016-285, 2017.
- 5 Van den Broeke, M. R., Enderlin, E. M., Howat, I. M., Kuipers Munneke, P., Noël, B. P. Y., van de Berg, W. J., van Meijgaard, E., and Wouters, B.: On the recent contribution of the Greenland ice sheet to sea level change, *The Cryosphere*, 10, 1933–1946, doi:10.5194/tc-10-1933-2016, 2016.
- Van Tricht, K., Lhermitte, S., Lenaerts, J. T. M., Gorodetskaya, I. V., L’Ecuyer, T. S., Noël, B., van den Broeke, M. R., Turner, D. D., and
10 van Lipzig, N. P. M.: Clouds enhance Greenland ice sheet meltwater runoff, *Nat. Commun.*, 7, 10 266, doi:10.1038/ncomms10266, 2016.
- Vernon, C. L., Bamber, J. L., Box, J. E., van den Broeke, M. R., Fettweis, X., Hanna, E., and Huybrechts, P.: Surface mass balance model intercomparison for the Greenland ice sheet, *The Cryosphere*, 7, 599–614, doi:10.5194/tc-7-599-2013, 2013.
- Wever, N., Fierz, C., Mitterer, C., Hirashima, H., and Lehning, M.: Solving Richards Equation for snow improves snowpack meltwater runoff estimations in detailed multi-layer snowpack model, *The Cryosphere*, 8, 257–274, doi:10.5194/tc-8-257-2014, 2014.
- 15 Wever, N., Würzer, S., Fierz, C., and Lehning, M.: Simulating ice layer formation under the presence of preferential flow in layered snow-packs, *The Cryosphere*, 10, 2731–2744, doi:10.5194/tc-10-2731-2016, 2016.
- Zwally, H. J., Giovinetto, M. B., Beckley, M. A., and Saba, J. L.: Antarctic and Greenland Drainage Systems, https://icesat4.gsfc.nasa.gov/cryo_data/ant_grn_drainage_systems.php, 2012.

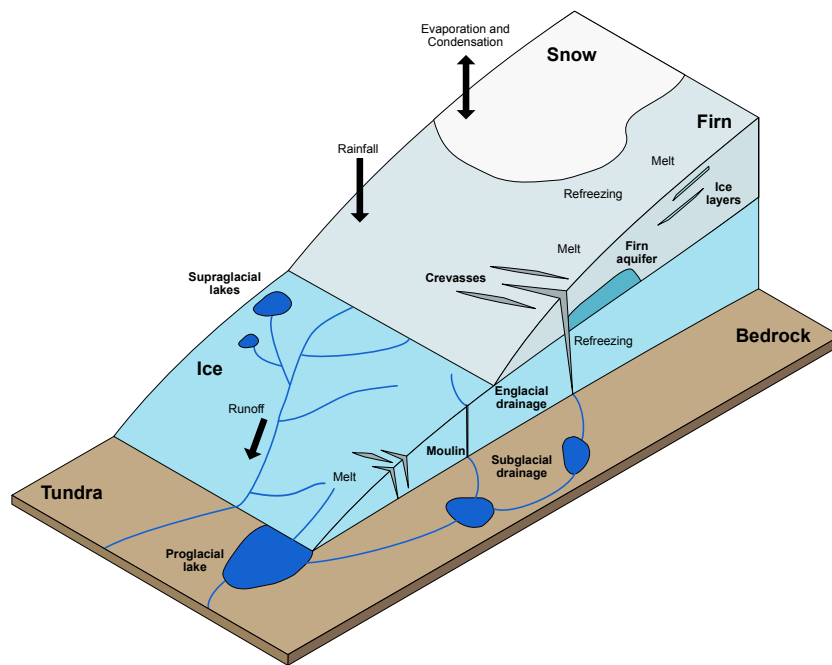


Figure 1. GrIS hydrology with the most relevant features and liquid water balance (LWB) components.

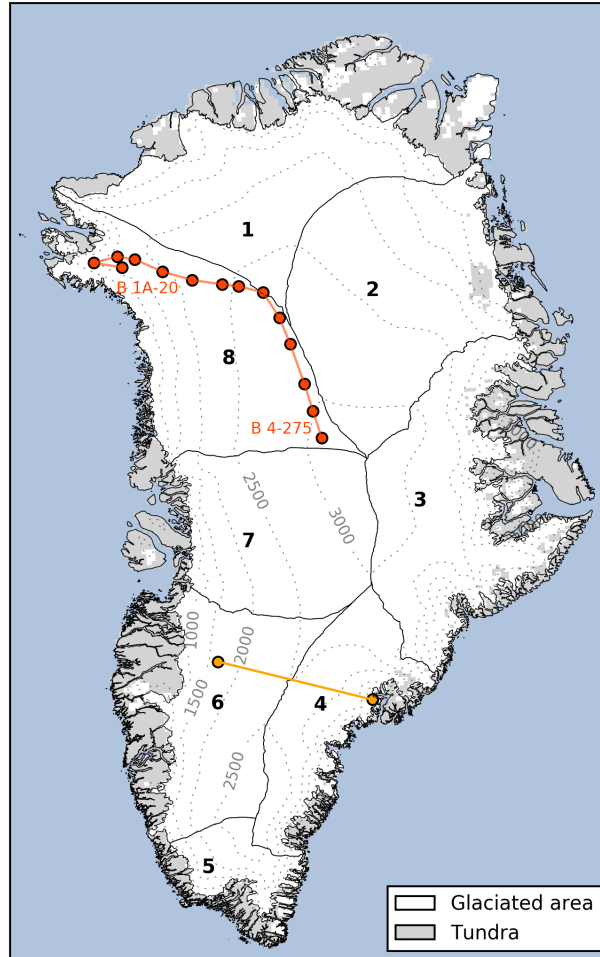


Figure 2. Map of Greenland with RACMO2.3 topography (500 m elevation contours as dashed lines) and land surface mask. Thin solid lines delineate eight drainage basins according to Zwally et al. (2012) and the connected circles indicate the locations of firn temperature measurements (red) and the defined southern GrIS transect (orange).

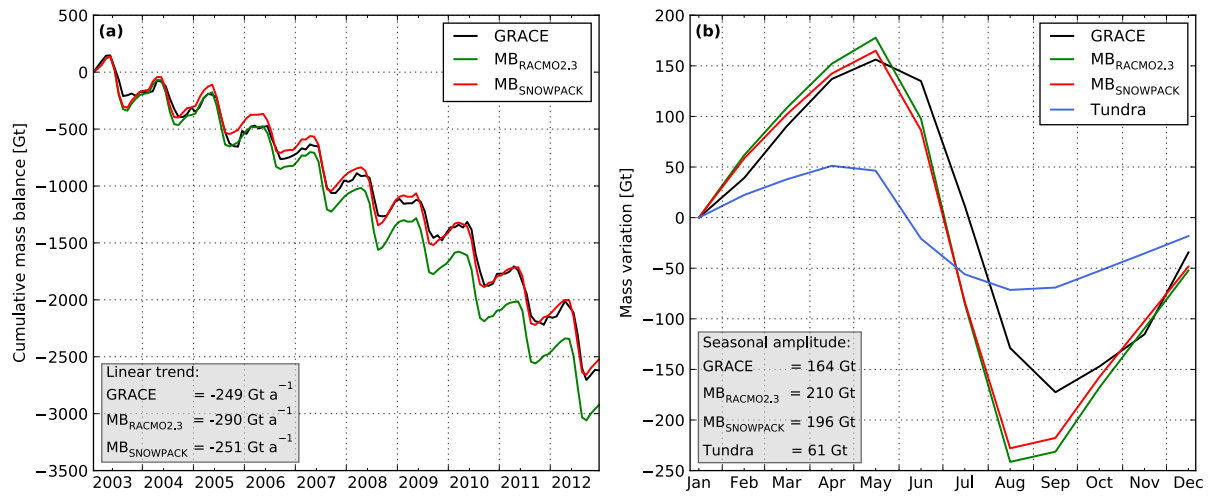


Figure 3. GRACE time series and cumulative MB (SMB of glaciated area from RACMO2.3 or SNOWPACK) (a) and detrended seasonal means of these series (b). The detrended seasonal mean of the tundra snow cover, simulated by RACMO2.3, is also shown. The inset in (a) provides linear trends and the inset in (b) seasonal amplitudes of the time series.

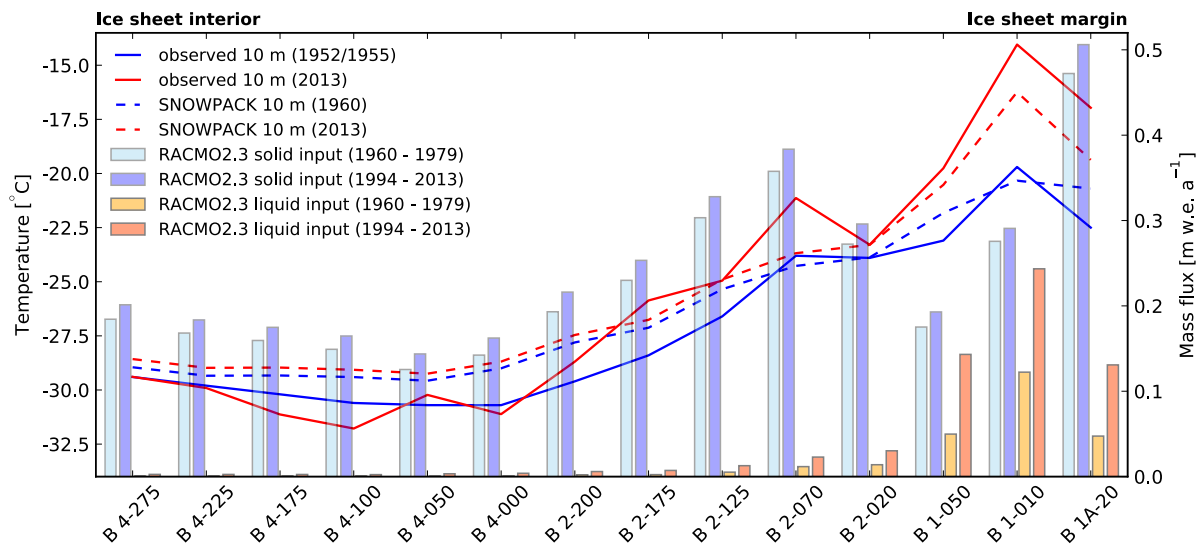


Figure 4. Observed and modelled firn temperatures (at 10 m depth) along the NW GrIS transect (Fig. 2). Bars represent RACMO2.3 solid (snowfall, sublimation and snow drift) and liquid (rainfall and snowmelt) surface inputs.

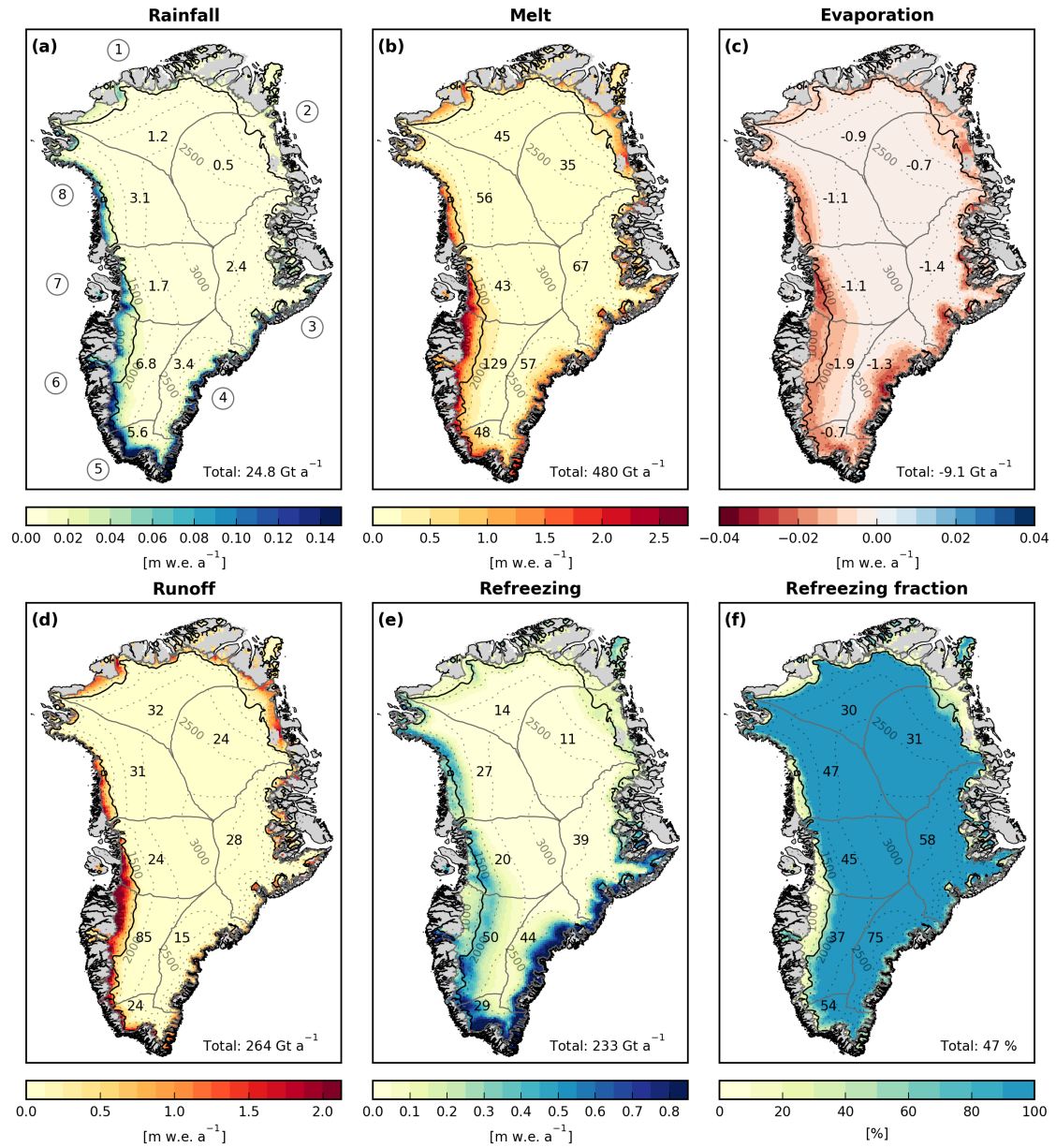


Figure 5. Components of the liquid water balance (LWB) for the glaciated area of Greenland averaged over 1960–2014 (a - e). Panel (f) shows refreezing as a fraction of liquid input (rainfall, melt and evaporation). Numbers represent basin-integrated values (excluding peripheral ice caps and glaciers) and the value in the lower right denotes the sum/average for the GrIS. The solid black line marks the mean position of the equilibrium line.

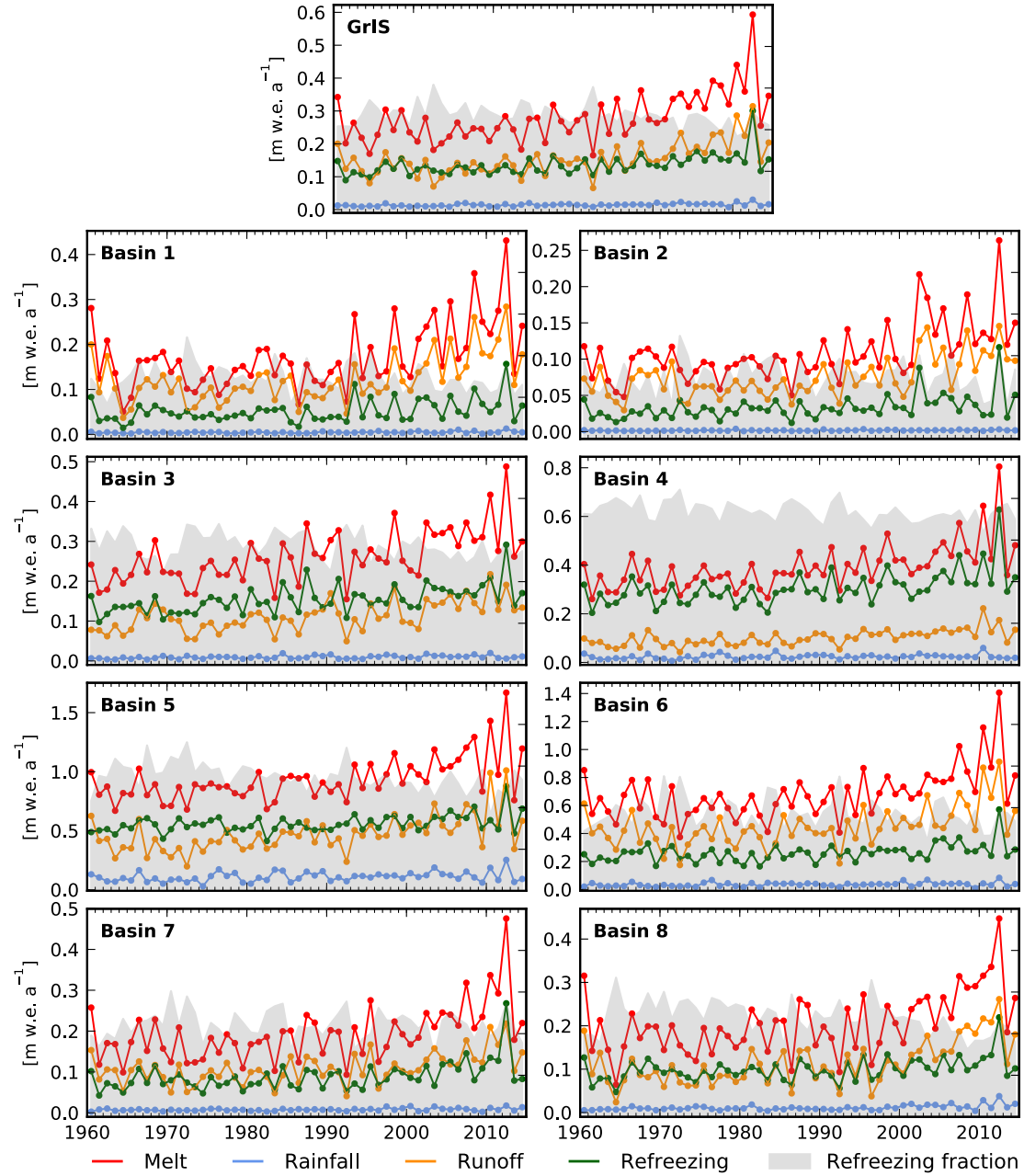


Figure 6. Time series of the liquid water balance (LWB) components for the GrIS (top) and the eight basins. Note the different vertical scales. Refreezing fractions in grey represent values between 0 and 100 %.

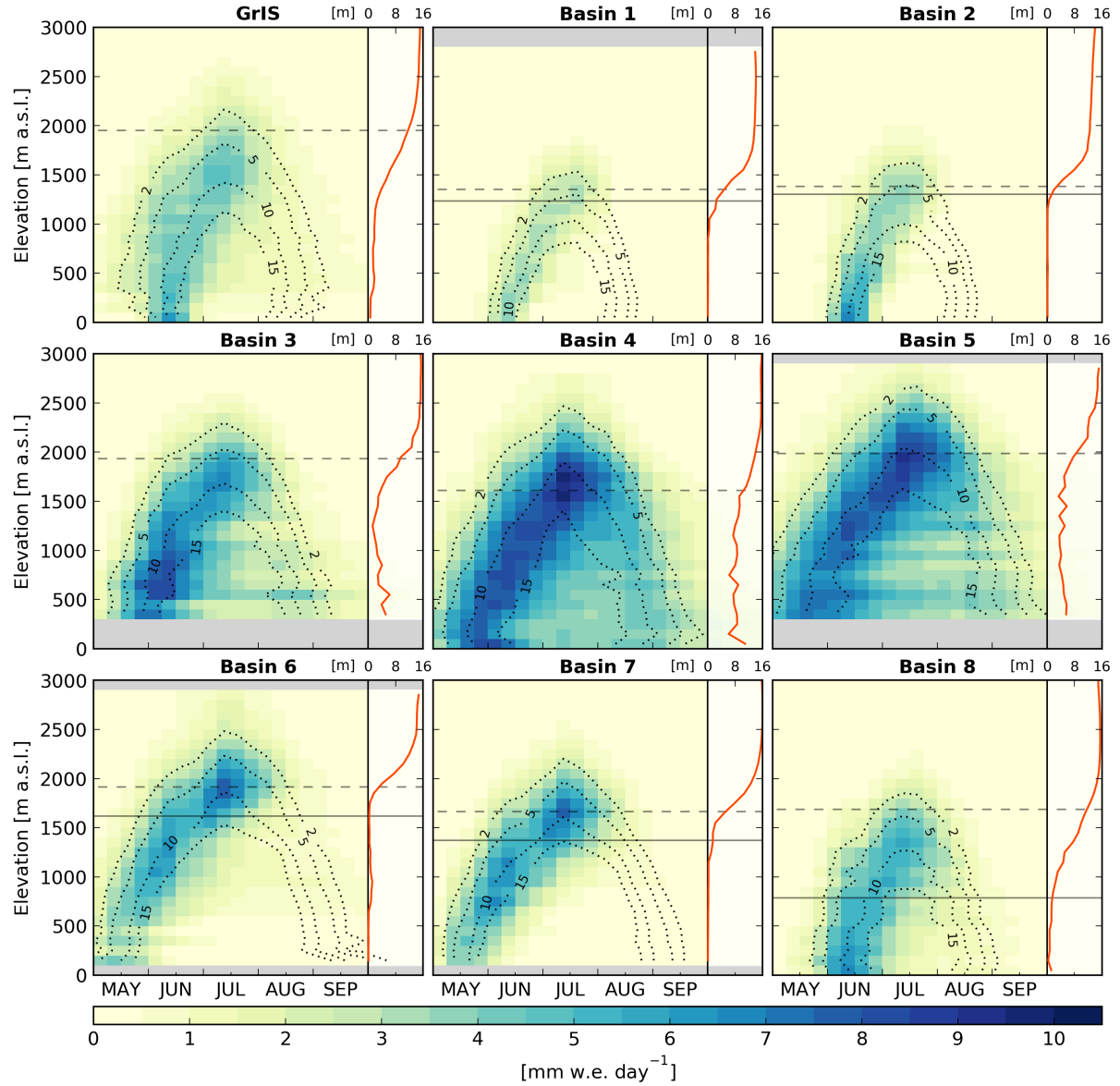


Figure 7. Mean 1960–2014 refreezing as a function of season and elevation. Each cell represents a 7.5 day period and a 100 m elevation bin. Surface melt aggregated with the same method is shown as dashed contour lines and the mean equilibrium line altitude and the elevation of the runoff line are indicated as solid and dashed lines, respectively. The red line displays the elevation bin-averaged firm air content of the upper 40 m.

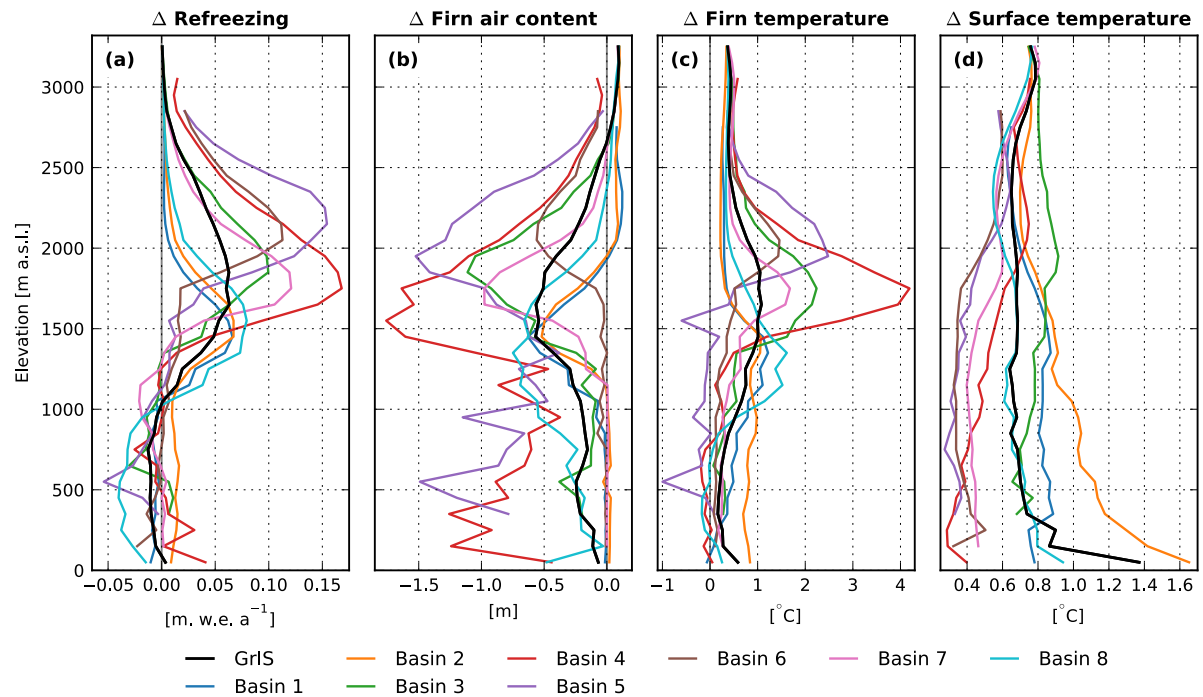


Figure 8. Temporal changes in refreezing, firn air content, firn temperature (averaged over 2–10 m depth) and surface temperature in 100 m elevation bins. The differences show the 1990–2014 average minus the 1960–1989 average.

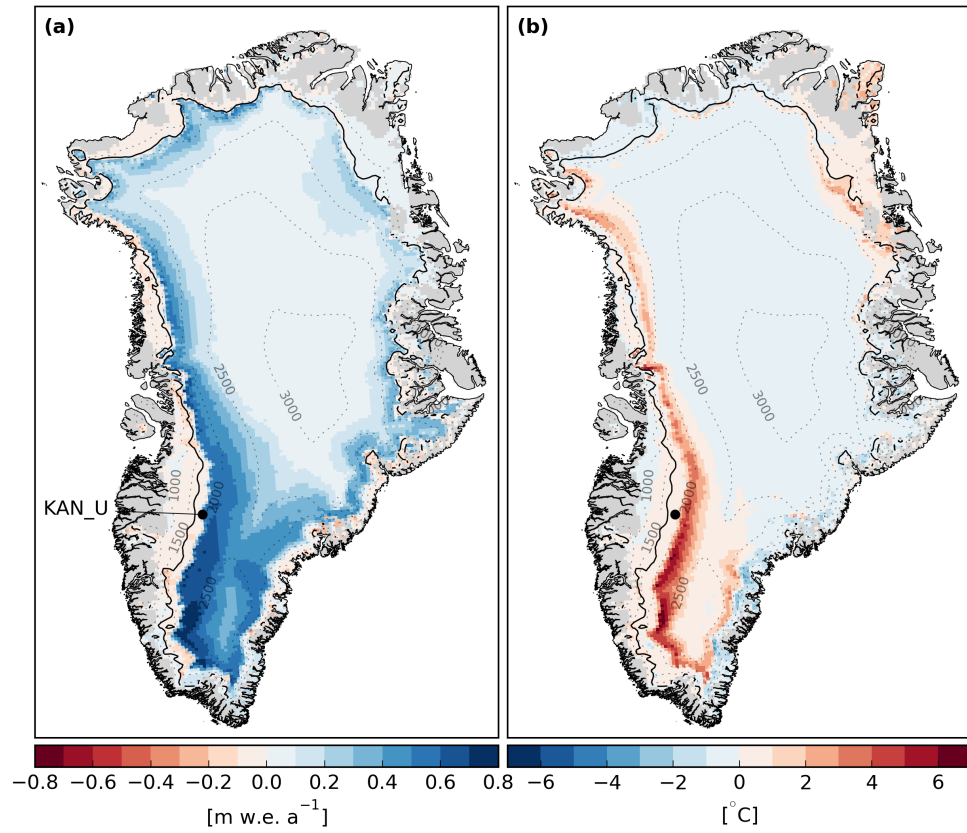


Figure 9. Refreezing anomaly of 2012 with reference period 1990–2014 (a) and firn temperature (average over 2–10 m depth) difference between 2011 and 2013 (b). The black line indicates the position of the equilibrium line for the reference period and the black dot the location of station KAN_U.

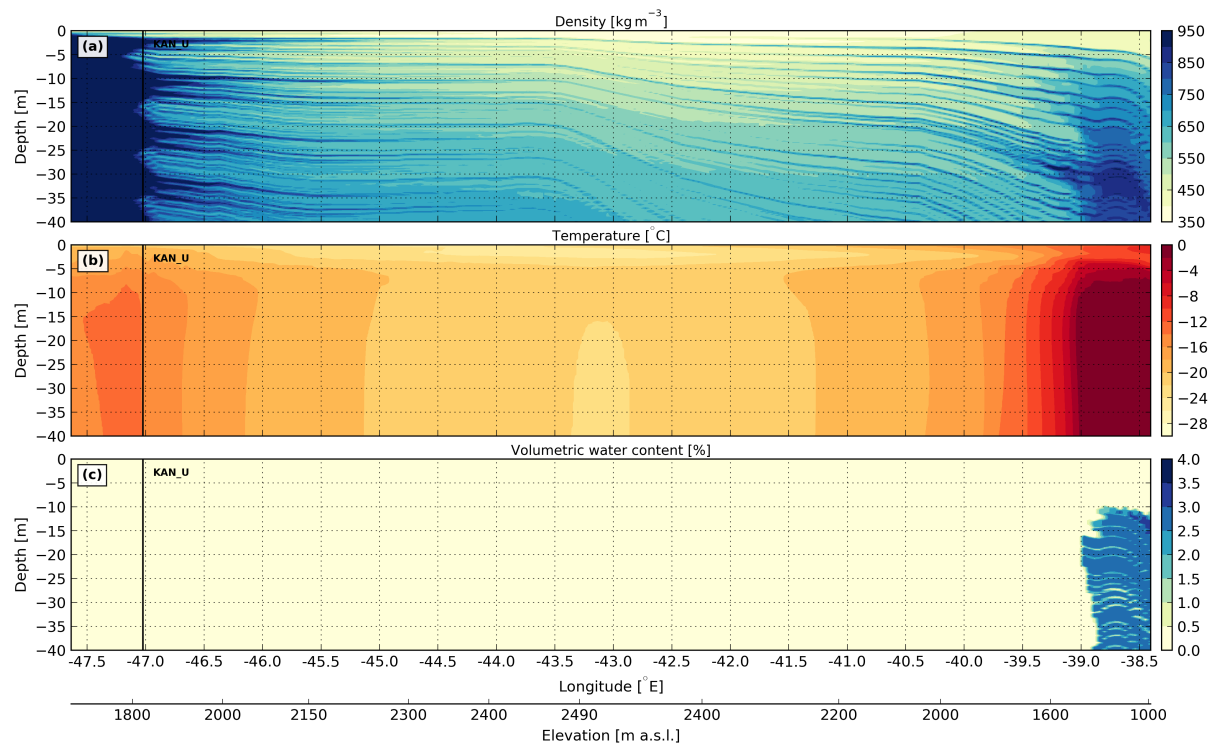


Figure 10. Modelled firn properties of the upper 40 m along the southern GrIS transect in April 1960. The vertical black line marks the location of station KAN_U.

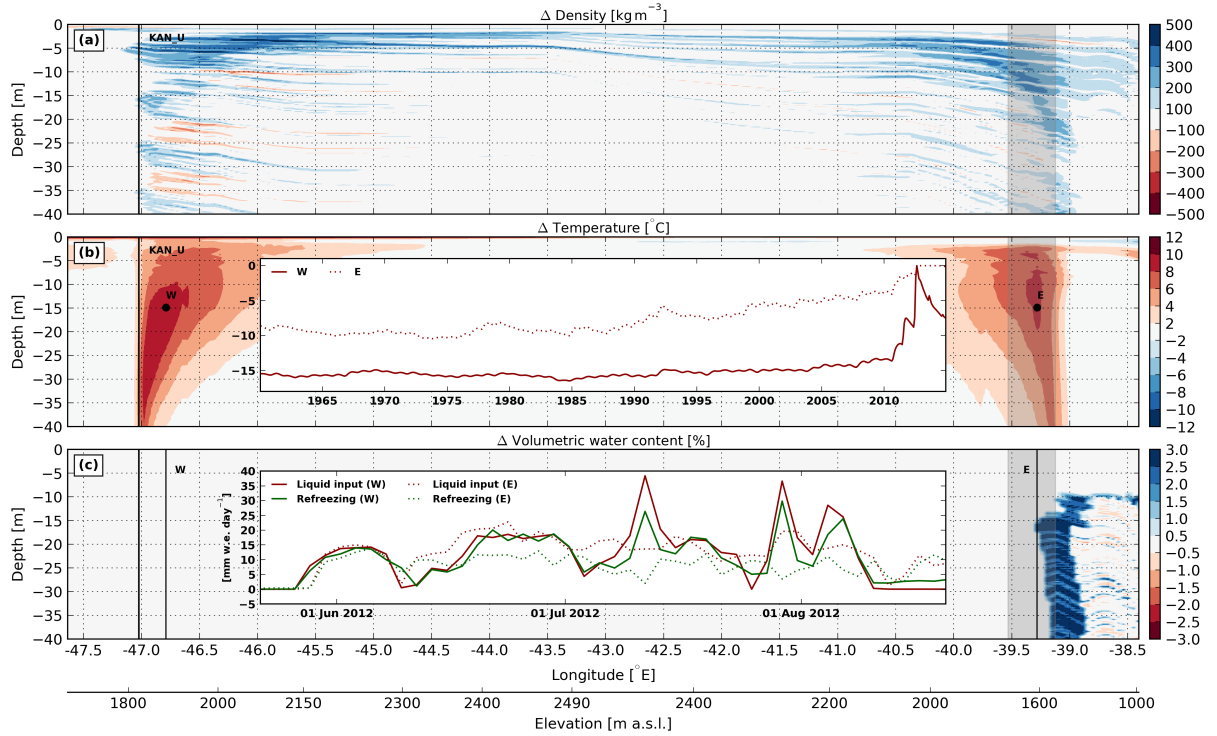


Figure 11. Modelled firn property changes of the upper 40 m along the southern GrIS transect between April 1960 and April 2014. The vertical black line marks the location of station KAN_U and the grey shaded area indicates the horizontal extent of observed firn aquifers in 2016 (Montgomery et al., 2017). The inset panel in (b) shows the temporal evolution of firn temperature at the two indicated locations. The inset panel in (c) shows daily mass fluxes of liquid input and refreezing in the summer 2012 for these locations.

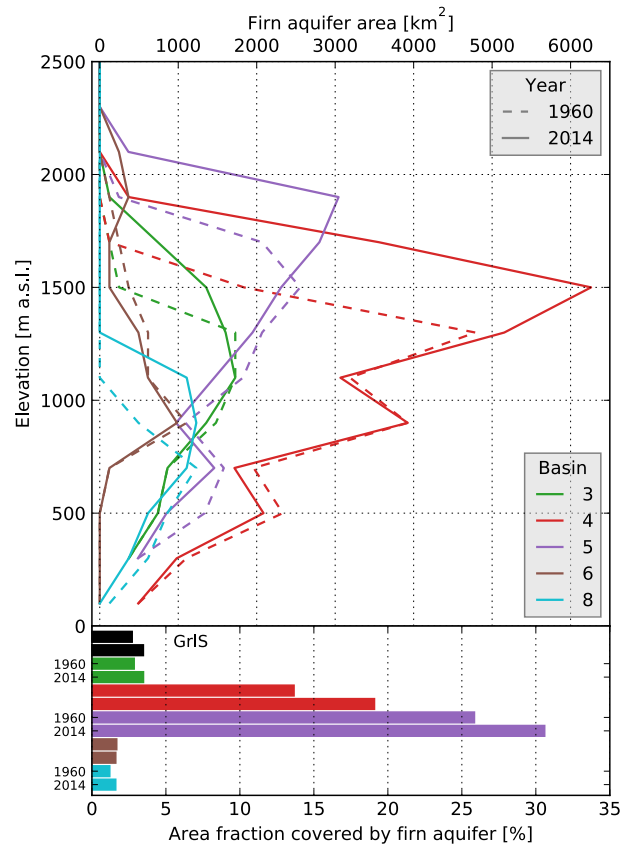


Figure 12. Elevation-dependent distribution of firn aquifer areas for the 5 basins where significant aquifers are modelled. Firn aquifer areas are delineated with a liquid water threshold of 200 kg m^{-2} and are aggregated in 200 m elevation bins. In the lower part of the figure, firn aquifer extent is shown as a fraction of the total basin area for the years 1960 and 2014.

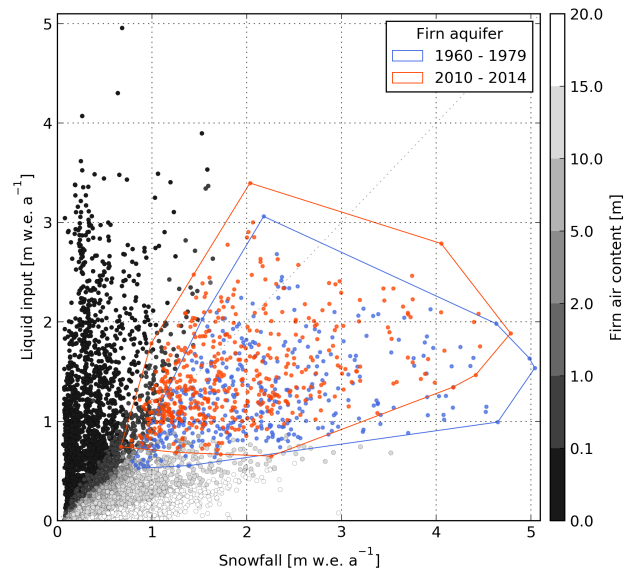


Figure 13. Model grid cells with seasonal dry firn as a function of snowfall and liquid input for the period 1960–1979. The colour map shows the firn air content of the upper 40 m for these points. Grid cells with perennial firn aquifer are delineated with a threshold of 200 kg m^{-2} of liquid water and are shown for the period 1960–1979 (blue dots) and 2010–2014 (orange dots).

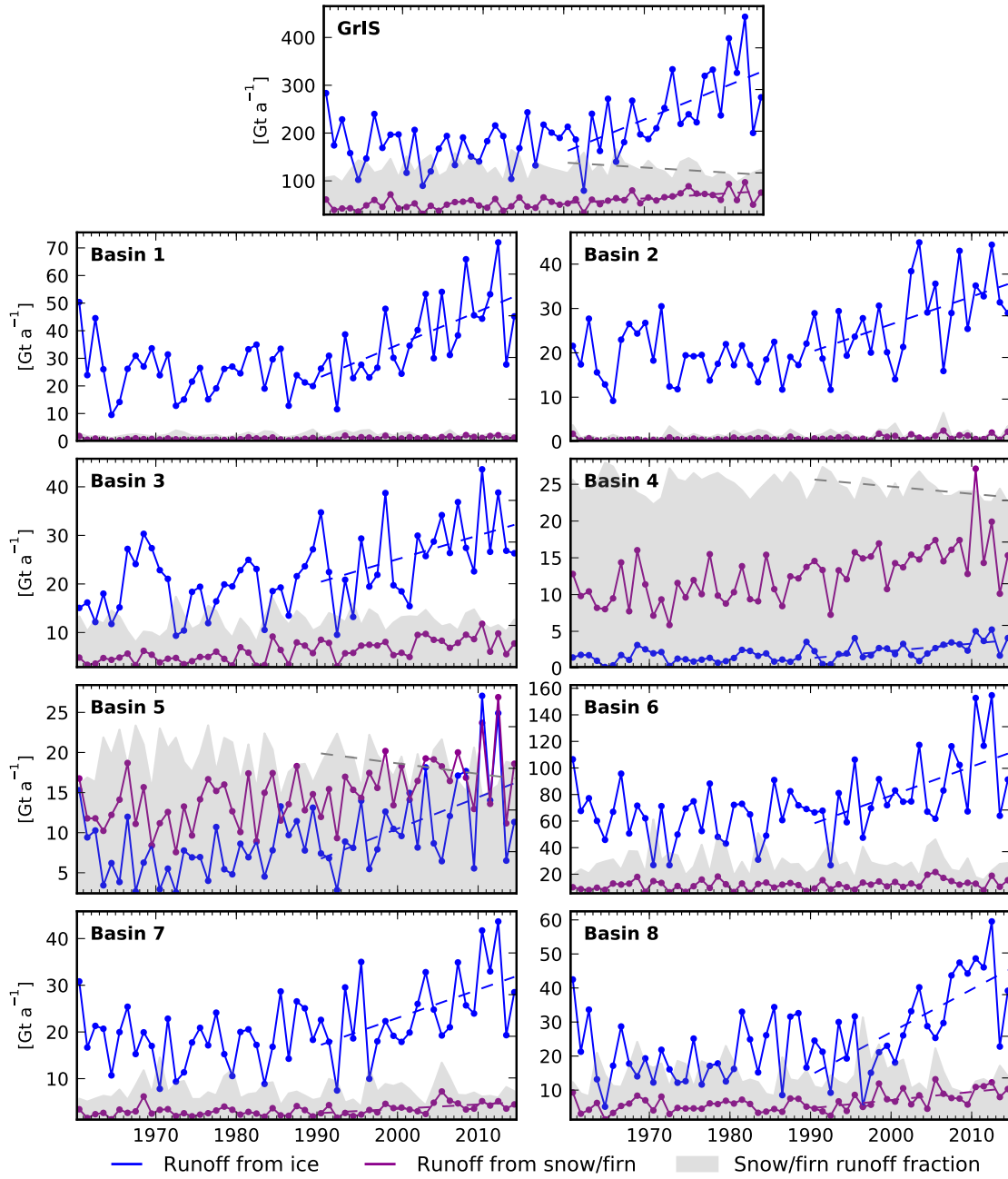


Figure 14. Time series of runoff from ice and snow/firn for the GrIS (top) and the eight basins. The grey shaded area shows runoff from snow/firn as a fraction of total runoff, with values between 0 and 100 %. Dashed lines indicate statistically significant trends (using a significance level of 0.05) between 1990 and 2014.

Table 1. Linear trends (1990–2014) in components of the liquid water balance (LWB) for the GrIS and the eight basins. Statistically insignificant trends (using a significance level of 0.05) are marked by an asterisk.

	Melt	Rainfall	Runoff	Refreezing	Refreezing fraction
			<i>mm w.e. a⁻²</i>		<i>% a⁻¹</i>
GrIS	6.50	0.22*	4.59	2.09	-0.34
Basin 1	5.79	0.08*	4.75	1.09*	-0.33
Basin 2	2.81	0.01*	2.05	0.73*	-0.22*
Basin 3	3.99	0.09*	2.29	1.74*	-0.23*
Basin 4	6.70	0.09*	1.99	4.67	-0.05*
Basin 5	14.22	1.22*	11.55	3.88*	-0.31*
Basin 6	15.37	0.43*	12.28	3.48*	-0.33*
Basin 7	5.69	0.11*	3.14	2.59	-0.13*
Basin 8	6.77	0.56	5.46	1.85	-0.63