The firn meltwater Retention Model Intercomparison Project (RetMIP): Evaluation of nine firn models at four weather station sites on the Greenland ice sheet

Baptiste Vandecrux[1,2], Ruth Mottram[3], Peter L. Langen[3], Robert S. Fausto[1], Martin Olesen[3], C. Max Stevens[4], Vincent Verjans[5], Amber Leeson[5], Stefan Ligtenberg[6], Peter Kuipers Munneke[6], Sergey Marchenko[7], Ward van Pelt[7], Colin Meyer[8], Sebastian B. Simonsen[9], Achim Heilig[10], Samira Samimi[11], Shawn Marshall[11], Horst Machguth[12], Michael MacFerrin[13], Masashi Niwano[14], Olivia Miller[15], Clifford I. Voss[16], Jason E. Box[1]

[1] Geological Survey of Denmark and Greenland, Copenhagen, Denmark.
[2] Department of Civil Engineering, Technical University of Denmark, Lyngby, Denmark.
[3] Danish Meteorological Institute, Copenhagen, Denmark
[4] Department of Earth and Space Sciences, University of Washington, WA USA
[5] Lancaster Environment Centre, Lancaster University, Lancaster, UK
[6] IMAU, Utrecht University, The Netherlands
[7] Department of Earth Sciences, Uppsala University, Uppsala, Sweden
[8] Thayer School of Engineering, Dartmouth College
[9] National Space Institute, Technical University of Denmark, Kgs. Lyngby, Denmark
[10] Department of Earth and Environmental Sciences, LMU, Munich, Germany
[11] Department of Geography, University of Calgary, Calgary, AB, Canada
[12] Department of Geosciences, University of Fribourg, Switzerland
[13] Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA
[14] Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, 305-0052 Japan
[15] U. S. Geological Survey, Utah Water Science Center, Salt Lake City, UT, USA
[16] U. S. Geological Survey, Menlo Park, CA, USA

Correspondence to: B. Vandecrux (bav@geus.dk)

Abstract. Perennial snow, or firn, covers 80% of the Greenland ice sheet and has the capacity to retain surface meltwater, influencing the mass balance of the ice sheet and its contribution to sea level rise. Multi-layer firn models are traditionally used to simulate firn processes and estimate meltwater retention. Here, we present output from nine firn models, forced by mass and energy fluxes derived from automatic weather stations at four sites that represent dry snow, percolation, ice slab and firn-aquifer areas. The model spread in firn density, temperature and water content, and the deviation from observations, increases with increasing melt, due to differences in how the models simulate meltwater infiltration. Models accounting for deep meltwater percolation overestimate percolation depth and firn temperature at the percolation and ice-slab sites but accurately simulate recharge of the firn aquifer. Models using Darcy's law and models using a bucket scheme compare favourably to observations at the percolation site, but at the ice slab sites only the Darcy models accurately simulate firn

temperature and meltwater percolation. Despite good performance at certain sites, no single model currently simulates meltwater infiltration adequately at all sites. The model spread in meltwater retention increases with surface meltwater input, reaching ±60% at KAN_U in 2012. That year models calculate that 30±24% of melt was run off which is low compared to a punctual runoff observation. We identify potential causes for the model spread and the mismatch with observation and provide recommendations for future model development and firn investigation.

1. Introduction

Responding to higher air temperatures and increased surface melt, the Greenland ice sheet has been losing mass at an accelerating rate over recent decades and is responsible for about 20% of observed global sea level rise (Van den Broeke et al. 2016, IMBIE Team 2019). Increasing temperatures have introduced melt at higher elevations where melt was previously seldom observed (Nghiem et al. 2012). In these colder, elevated areas, snow builds up into a thick layer of firn. Increased surface melt in the firn area of the Greenland ice sheet affects firn structure (Machguth et al. 2016; Mikkelsen et al. 2015), density (De La Peña et al. 2015; Vandecrux et al. 2018), air content (van Angelen et al. 2013; Vandecrux et al. 2019) and temperature (Polashenski et al. 2014; Van den Broeke et al., 2016). Changing firn characteristics affect its meltwater storage capacity; either in terms of refreezing within the firn (Pfeffer et al., 1991; Braithwaite et al., 1994) or as liquid water retained in perennial firn aquifers (e.g. Forster et al. 2014; Miège et al. 2016), therefore impacting the ice-sheet contribution to sea-level rise (Harper et al., 2012; Machguth et al. 2016; Mikkelsen et al. 2015; Van As et al. 2017). Meltwater refreezing can also form continuous ice layers that are several meters thick (MacFerrin et al. 2019). These ice slabs impede vertical meltwater percolation, enhance surface-water runoff, and lower the surface albedo, further amplifying Greenland's contribution to sea-level rise (Charalampidis et al. 2015). The evolution of firn on the Greenland ice sheet is important for two additional reasons: first, knowledge about how firn air content evolves through time is necessary for the conversion of space-borne observations of ice sheet volume change into mass change (e.g. Sørensen et al. 2011; Simonsen et al. 2013). Secondly, the depth of firn to ice transition, as well as the mobility of gases through the firn before they are trapped in bubbles within glacial ice, are necessary for the interpretation of ice cores and heavily depend on the fine coupling between the firn characteristics and surface conditions (e.g. Schwander et al., 1993).

Firn models traditionally take as input energy and mass fluxes at the surface and calculate the evolution of firn characteristics and meltwater retention at scales ranging from tens of metres to tens of kilometres. The performance of these models, when coupled to regional and global climate models, has a direct impact on the quality of ice-sheet mass-balance calculations (Fettweis et al., 2020) and sea-level change estimations (Nowicki et al., 2016). In previous work, Reijmer at al. (2012) suggested that, provided reasonable tuning, simple parameterizations of the subsurface processes calculate refreezing rates for the Greenland ice sheet in agreement with results from physically based, layered subsurface models. However, spatial patterns varied widely and evaluation against field observations remained challenging. Steger et al. (2017) and more recently Verjans et al. (2019) investigated the impact of meltwater infiltration schemes on the simulated properties of the firn in

Greenland. These studies highlighted the potential of deep-percolation schemes, for instance for the simulation of firn aquifer, but also the sensitivity of simulated infiltration to the firn structure and hydraulic properties. In these previous studies, the surface conditions were prescribed by a regional climate model. Inaccuracies in this forcing could therefore explain some of the deviation between model outputs and firn observations and prevented a full assessment of different firn model designs.

The meltwater Retention Model Intercomparison Project (RetMIP) compares results from ten firn models currently used for the Greenland ice sheet. The models are forced with consistent surface inputs of mass and energy and simulations are performed at four sites where surface conditions could be derived from automatic weather station (AWS) observations and where firn observations are available. These four sites were chosen to represent various climatic zones of the Greenland ice sheet firn area: the dry snow area, where melt is rare and temperatures are low, is represented by Summit; the percolation area, where melt occurs every summer at the surface, infiltrates in the snow and firn and refreezes entirely there, is represented by Dye-2; ice slab regions, where a thick ice layer hinders deep meltwater percolation, is represented by KAN_U; and firn aquifer regions, where infiltrated meltwater remains liquid at depth is represented by FA. At each site, we compare simulated temperature, density and the resulting meltwater infiltration patterns between models and to in situ measurements. We discuss model features that can be responsible for model spread and deviation from observations. Lastly, we evaluate how differences in simulated firn characteristics result in various simulated refreezing and runoff values at sites where melt and/or runoff occur and attempt to quantify uncertainties linked to firn models.

2. Models

The multi-layer firn models investigated here are listed in Table 1. They all have density, temperature, and liquid water content as prognostic variables and apply a framework whereby firn is divided into multiple layers for which these characteristics can be calculated. The number of layers varies in each model (Table 2) and we distinguish between two distinct types of layer management strategies: all models except DMIHH and MeyerHewitt follow a Lagrangian framework, i.e. they add new layers at the top of the model column during snowfall and these layers are advected downward as new material accumulates at the surface. DMIHH and MeyerHewitt follow a Eulerian framework in which the layers have either fixed mass or fixed volumes. During snowfall, new material is added to the first layer and an equivalent mass/volume is transferred by each layer to its underlying neighbour. At each time step, the models calculate firn density according to different densification formulations and update the layer temperature using different values of thermal conductivity (Table 2). The DMIHH, GEUS and DTU models have a fixed temperature at the bottom of their column (Dirichlet boundary condition) while other models have a fixed temperature gradient (Neuman boundary condition).

All models simulate meltwater percolation and transfer water vertically from one layer to the next according to the routines listed in Table 2. They also simulate meltwater refreezing and latent heat release. All models simulate the retention of meltwater within a layer due to capillary suction, either explicitly (MeyerHewitt and CFM model) or, for all other models,

parameterised through the use of an irreducible water content after Coléou and Lesaffre (1998). When meltwater cannot be transferred to the next layer or be retained within the layer by capillary suction, lateral runoff can occur according to model-specific rules (Table 2). The background and specifics of each model are described in greater detail in the following paragraphs.

Table 1: Models evaluated in this study.

Model code name	Developing institute	References
CFM-Cr CFM-KM	University of Washington, Lancaster University	Stevens et al. (2020), Verjans et al. (2019)
DTU	Technical University of Denmark – National Space Institute	Sørensen et al. (2011), Simonsen et al. (2013)
DMIHH	Danish Meteorological Institute	Langen et al. (2017)
GEUS	Geological Survey of Denmark and Greenland	Vandecrux et al. (2018, 2020)
IMAU-FDM	Institute for Marine and Atmospheric research Utrecht (IMAU), Utrecht University	Ligtenberg et al. (2011, 2018), Kuipers Munneke et al. (2015)
MeyerHewitt	Thayer School of Engineering, Dartmouth College	Meyer and Hewitt (2017)
UppsalaUniBucket UppsalaUniDeepPerc	Uppsala University	Van Pelt et al. (2012,2019), Marchenko et al. (2017)

2.1. CFM-Cr and CFM-KM models

The Community Firn Model (CFM) is an open-source, modular model framework designed to simulate a range of physical processes in firn (Stevens et al., 2020). The number of layers for a particular model run is fixed and determined by the accumulation rate and time-step size. New snow accumulation at each time step is added as a new layer, and a layer is removed from the bottom of the model domain. A layer-merging routine prevents the number of layers from becoming too large. CFM-Cr and CFM-KM use the Crocus (Vionnet et al., 2012) and Kuipers Munneke et al. (2015) densification schemes, respectively (Table 2). Both use the same meltwater percolation scheme: a dual-domain approach that closely follows the

implementation of the SNOWPACK snow model (Wever et al., 2016). It accounts for the duality of water flow in firn by simulating both slow matrix flow and fast, localised, preferential flow (Verjans et al., 2019). In the matrix flow domain, water percolation is prescribed by the Richards Equation; ice layers are impermeable, and runoff is allowed. In contrast, preferential flow can bypass such barriers and no runoff is simulated. Water is exchanged between both domains as a function of the firn layer properties: density, temperature and grain size. As such, when water in the matrix flow domain accumulates above an ice layer, it is progressively depleted by runoff and by transfer of water into the preferential flow domain. In the deepest firn layers, above the impermeable ice-sheet, water accumulates, and no runoff is prescribed, which allows for the build-up of firn aquifers.

2.2. DTU model

The DTU firn model was developed to derive the Greenland ice sheet mass balance from the satellite observations of ice sheet elevation change (Sørensen et al., 2011) and to describe the firn stratigraphy and annual layers in the dry-snow zone along the EGIG-line in central Greenland (Simonsen et al., 2013). The DTU model uses the densification scheme from Arthern et al. (2010) and a bucket scheme for meltwater infiltration and retention. If meltwater is conveyed to a model layer, the water is refrozen if sufficient pore space and cold content are available in the layer. Additional liquid water can be retained in a layer by capillary forces calculated after Schneider and Jansson (2004). This formulation does not allow for the formation of firn aquifers. Percolation continues until the water encounters a layer at ice density or the bottom of the model where, in both cases, it is assumed to run off. The model follows a Lagrangian scheme of advection of layers down into the firn and the model layering is defined by the time-stepping of the model.

2.3. DMIHH model

The DMIHH model was developed to provide firn subsurface details for the HIRHAM regional climate model experiments (Langen et al., 2017). DMIHH employs 32 layers within which snow, ice and liquid water fractions can vary and where each layer has a constant mass. Layer thicknesses increase with depth to increase resolution near the surface and give a full model depth of 60 m water equivalent (w.e.). Mass added at the surface (e.g., snowfall) or removed as runoff causes the scheme to advect mass downward or upward to ensure the constant w.e. layer thicknesses. In addition to the saturated and unsaturated hydraulic conductivities (Table 2), the water flow through layers containing ice follows the analytical model of Colbeck (1975) for a snowpack with discontinuous ice layers. A parameter describing the ratio between the characteristic distance between two adjacent ice lenses and the characteristic width of an ice lens was set to 1, meaning that ice lenses have a horizontal extent of half the unit area. A layer is considered impermeable if its bulk dry density exceeds 810 kg m^{-3}. Runoff is calculated from the water in excess of the irreducible saturation with a characteristic local runoff time-scale that increases as the surface slope tends to zero (Zuo and Oerlemans, 1996), with the coefficients of the time-scale parameterization from Lefebre et al. (2003). DMIHH has an initial value of 0.1 mm for the grain diameter of freshly fallen snow. The column grain

size distribution is initialized in these experiments as columns taken at the specific sites from the spinup experiments performed by Langen et al. (2017).

2.4. GEUS model

The GEUS model is based on the DMIHH model (Langen et al., 2017) and is further developed in Vandecrux et al. (2018, 2020). As in the DMIHH model, the layer's ice content decreases its hydraulic conductivity according to Colbeck (1974) but we set the geometry parameter to 0.1 as detailed in Vandecrux et al. (2018). At the end of a time step, water exceeding the irreducible water content that could not be percolated downward is assumed to run off and is removed from the layer at a rate that depends on the firn characteristics and on surface slope, according to Darcy's law. More details about this runoff scheme are provided in the Supplementary text S1.

2.5. IMAU-FDM

The IMAU-FDM model has been used in combination with the RACMO regional climate model in Greenland, Arctic Canadian ice caps, and Antarctica. Firn compaction follows a semi-empirical, temperature-based equation from Arthern (2010). The compaction rate is tuned to observations from Greenland firn cores using an accumulation-based correction factor (Kuipers Munneke et al., 2015). IMAU-FDM includes meltwater percolation following a tipping-bucket approach. Percolating meltwater is refrozen if there is space available in the layer, and if the latent heat of refreezing can be released in the layer. As opposed to other models in this study, runoff is not allowed over ice layers, but only when percolating meltwater has reached the pore close-off depth. Upon reaching that depth, runoff is instantaneous. The rationale for allowing percolation through thick ice slabs is that IMAU-FDM is mainly used to simulate firn at scales of tens to hundreds of square kilometres, and at these spatial scales, meltwater is assumed to always find a way through even the thickest of ice slabs.

2.6. MeyerHewitt model

Meyer and Hewitt (2017) present a continuum model for meltwater percolation in compacting snow and firn. The MeyerHewitt model includes heat conduction, meltwater percolation and refreezing, as well as mechanical compaction using the empirical Herron and Langway (1980) model. In the MeyerHewitt model, water percolation is described using Darcy's law, allowing for both partially and fully saturated pore space. Water is allowed to run off from the surface if the snow is fully saturated. Using an enthalpy formulation for the problem, the MeyerHewitt model is discretized using the conservative finite volume method that is fixed in the frame of the firn surface and is Eulerian, meaning that material can flow into and out of the domain.

174 **Table 2: Model characteristics.**

Model	Discretization	Meltwater routing	Hydraulic conductivity (Saturated, unsaturated)	Firn densification	Runoff calculation	Thermal conductivity
CFM-Cr	Unlimited number of layers. Lagrangian	Richards equation and dual-domain preferential flow scheme (Wever et al., 2016; Verjans et al., 2019)	Calonne et al. (2012); van Genuchten (1980) with coefficients from Daanen and Nieber (2009)	Vionnet et al. (2012)	Zuo and Oerlemans (1996)	Anderson (1976)
CFM-KM				Kuipers Munneke et al. (2015)		
DTU	Dynamically allocated, based on accumulation rates, timestep and depth range. Lagrangian	Bucket scheme	-	Sørensen et al. (2011); Simonsen et al. (2013)	Immediate runoff on top of an ice layer	Schwander et al. (1997)
GEUS	200 layers dynamically allocated, Lagrangian	Darcy's law	Calonne et al. (2012), van Genuchten (1980) with coefficient from Hirashima et al. (2010)	Vionnet et al. (2012)	Darcy flow to adjacent cell given surface slope	Calonne et al. (2011)
DMIHH	32 layers, Eulerian				Zuo and Oerlemans (1996)	Yen (1981)
IMAU-FDM	maximum of 3000 layers, Lagrangian	Bucket scheme	-	Kuipers Munneke et al. (2015)	Only at the bottom of the column	Anderson (1976)
MeyerHewitt	finite volume, Eulerian, 600 layers	Darcy's law	Carman-Kozeny (Bear, 1972); Gray (1996)	Herron and Langway (1980)	Excess surface water	Meyer and Hewitt (2017)
UppsalaUniBucket	600 layers, max 0.1 m layer thickness. Lagrangian	Bucket scheme	-	Ligtenberg et al. (2011)	Only at the bottom of the column	Sturm et al. (1997)
UppsalaUniDeepPerc		Deep percolation scheme; linear distribution down to 6 m (Marchenko et al. 2017)				

175

2.7. UppsalaUniBucket and UppsalaUniDeepPerc models

UppsalaUniBucket and UppsalaUniDeepPerc have been developed for the Norwegian Arctic (Van Pelt et al. 2012; 2019; Marchenko et al., 2017) and only differ in their representation of vertical water transport. UppsalaUniBucket simulates melt water percolation according to the tipping-bucket scheme while UppsalaUniDeepPerc uses a deep percolation scheme which mimics the effect of fast vertical transport due to preferential flow (Marchenko et al. 2017). The water transport model incorporates irreducible water storage but does not allow for standing water to accumulate on top of the impermeable ice; instead all water that reaches the base of the firn column is set to runoff instantaneously. References for the parameterizations used for gravitational settling, thermal conductivity, irreducible water storage and water percolation are given in Table 2.

3. Methods

3.1. Site selection and surface forcing

Differences between firn-model outputs and observations depend on both the model formulation, and the forcing data that are given to the model (e.g., Ligtenberg et al., 2018); i.e. any bias in forcing data propagate into the model output. To make sure we compare and evaluate the models independently of biases that may exist in forcing datasets that come from RCMs, we use meteorological fields derived from five AWS at four sites.

These sites represent a broad range of climatic conditions on the Greenland ice sheet (Table 3, Figure 1) that produce a wide variety of firn density and temperature profiles. For example, the cold and dry climate at Summit Station produces cold firn with low compaction rates representative of the "dry snow" area as defined by Benson (1962). Dye-2, located in an area with higher melt (Table 3), is representative of the "percolation area" (Benson, 1962) where meltwater generated at the surface percolates into the firn and releases latent heat when refreezing into ice lenses. At the KAN_U site, lower accumulation rates and increasing melt have led to the formation of thick ice slabs (Machguth et al., 2016; MacFerrin et al., 2019) that impede meltwater percolation below 5 m. The Firn Aquifer (FA) site in Southeast Greenland has both high surface melt and high accumulation rate, leading to the formation of a perennial body of liquid water at a depth of 12 m and below (Forster et al., 2012; Kuipers Munneke et al. 2014).

We use data from GC-Net AWS at Dye-2 and Summit (Steffen et al., 1996), from the PROMICE station at KAN_U (Ahlstrøm, et al., 2008; Charalampidis, et al., 2015) and from IMAU, Utrecht University at the Firn Aquifer site (see Supplementary Text S2 for station description). For Dye-2 in 2016 we use a AWS installed by the University of Calgary (see Supplementary Text S2 for station description). Since this station was more recently installed than the GC-Net station, this

ensures the best meteorological forcing for the models over that melting season, during which an extensive observational dataset is available for model evaluation.

Data from each AWS were quality checked and obvious sensor malfunctions were discarded. No data were discarded at FA and Dye-2_16. The resulting data gaps were filled using either nearby stations or HIRHAM5 data as in Vandecrux et al. (2018). Downward longwave radiation is not monitored by the GC-Net stations (Dye-2 and Summit) and is taken entirely from HIRHAM5 output.

Meteorological fields from the AWS (temperature, wind speed, humidity, air pressure, incoming and outgoing shortwave and longwave radiation, and snow surface height) are gap-filled with adjusted values from HIRHAM5, following Vandecrux et al. (2018), and used to calculate the surface energy balance, based on the model developed by van As et al. (2005) and applied in Vandecrux et al. (2018). This surface energy and mass balance provides, at three-hourly resolution, the three surface forcing fields that were used by all models: the surface "skin" temperature, the amount of meltwater generated at the surface, and net snow accumulation (precipitation – sublimation + deposition). Only the MeyerHewitt model required minor adaptation of these forcing fields (see Supplementary Text S1). Rain is not monitored at any site, so is not included in the mass fluxes. Tilt of the radiation sensor was not corrected for at Dye-2 and Summit stations although this correction was seen to increase the calculated melt by 35 mm w.e. yr^{-1} at Dye-2 (Vandecrux et al., 2020). The surface forcing data are illustrated in Figure S3.

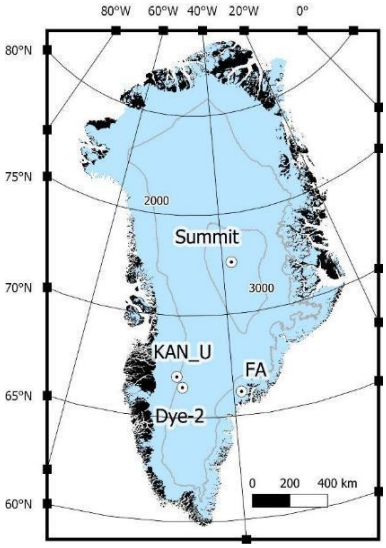

Figure 1: Map of the four study sites.

Table 3: Information about the four sites and five model runs considered in the comparison including mean annual accumulation ($\bar{b}$), mean annual air temperature ($\bar{T_a}$) and measured average deep firn temperature (T_{deep}).

Station name	Elevation (m a.s.l.)	Start date	End date	$\bar{b}$ (mm w.e.)	$\bar{T_a}$ (°C)	Surface slope (°)	T_{deep}(°C)	Initial firn density
KAN_U	1840	01-05-2012	31-12-2016	543	-12	0.5	-9@5m	Top 10 m: core_1_2012 (Machguth et al. 2016) From 10 to 60 m: Site J, 1989 (Kameda et al. 1995)
Dye-2_long	2165	01-06-1998	02-05-2015	476	-16	0.2	-15.5 @10 m	Dye-2 1998 core B (Mosley-Thompson et al. 2001)
Dye-2_16	2165	02-05-2016	28-10-2016	476	-16	0.2	-13@9m	Top 18 m: Core_10_2016 (B. Vandecrux et al. 2018) From 10 to 60 m: Dye-2 1998 core B (Mosley-Thompson, et al., 2001)
Summit	3254	02-07-2000	08-03-2015	159	-26	0	-31@10m	Top 8m: core from 1990 by Mayewski & Whitlow., (2016) From 8 to 60 m: GRIP core
Firn Aquifer (FA)	1563	12-04-2014	02-12-2014	1739	-7	0.6	0@25m	Top 8 m: FA-14 (Montgomery et al., 2018) From 8 to 60 m: FA-13 (Koenig et al. 2014)

3.2. Boundary conditions

To allow fair comparison of the various firn models, as many boundary conditions as possible were specified in common for all models. A key parameter in firn models is the density of fresh snow added at the top of the model column. Here, all models used the value of 315 kg m^{-3} from Fausto et al. (2018) which is derived from a compilation of 200 top 10 cm snow density observations from the Greenland ice sheet. The local surface slope and deep firn temperature, were prescribed according to Table 3. Initial profiles for density, temperature and liquid water content (only at FA) were provided to all models and illustrated in Supplementary Figure S4. The references for the initial density profiles are given in Table 3. Initial temperature profiles were calculated using the first reading of air temperature (as first guess of surface temperature), the first valid measurement of firn temperature, and the deep firn temperature (Table 3). The deep firn temperature was calculated as the long-term mean at the deepest firn temperature measurement available. Initial liquid water content at FA is calculated according to the observations from Koenig et al. (2014) which indicate pore saturation below 12.2 m depth.

3.3. Intercomparison and evaluation of model output

Participating models provided simulated firn density, temperature and liquid water content in 3 h time steps, interpolated to a common 10 cm grid from the surface to 20 m depth. Additionally, three-hourly vertically integrated refreezing and runoff were calculated by each model.

Three types of datasets are available at our sites for model evaluation: i) firn-temperature observations from AWS as presented by Vandecrux et al. (2020) at Summit and Dye-2, Heilig et al. (2018) at Dye-2 in 2016 , Charalampidis et al. (2015) at KAN_U and Koenig et al. (2014) at the FA station; ii) firn density profiles (Table 4); and iii) observations of meltwater infiltration depth at Dye-2 over the summer 2016 (Heilig et al., 2018).

Table 4. Firn cores used for model evaluation

	Date	Reference
Summit	05-03-2001	Dibb and Fahnestock (2004)
	29-05-2015	Vandecrux et al. (2018)
Dye-2	01-06-2011	Forster et al. (2014)
	05-05-2013	Machguth et al. (2016)
	21-05-2015	Vandecrux et al. (2018)
KAN_U	01-05-2012	Machguth et al. (2016)
	27-04-2013	
	05-05-2015	
	28-04-2016	MacFerrin et al. (2019)
	28-04-2017	

For firn density, we first compare the simulated density profiles to the firn core data at each site. We also calculate for each time step the average firn density over the 0-1 m, 1-10 m and 10-20 m depth ranges and discuss the standard deviation of these values among models and their deviation from firn core observations. For firn temperature, we compare hourly observations of firn temperature to interpolated temperature from the closest model layers and use the Mean Error (ME), Root Mean Squared Error (RMSE) and coefficient of determination (R^2) to quantify the performance of the models with respect to the observations.

4. Results

In the following, we present comparisons of firn model outputs and model deviations from observations for firn temperature, density, and liquid water content at the sites representing different firn and meltwater regimes: dry firn (Summit), the percolation zone (Dye-2), ice slabs (KAN_U), and a firn aquifer (FA site).

4.1. Dry firn site: Summit

First, we evaluate the performance of the firn models at Summit, where almost no surface melt occurs, and percolation and refreezing play a negligible role. Results for this site allow us to assess firn densification and thermal advection-diffusion processes simulated by each model without the complications of density and temperature change by refreezing of percolating meltwater.

At Summit, density evolves in a similar manner across all models: low density snow is deposited at the surface and is advected to greater depth (Figure 2a). All models except DMIHH and MeyerHewitt preserve the layering of the initial density profiles as it gets advected and generate layered firn at the surface from the forcing there. The temporal evolution of the average density for the 0-1 m depth range follows similar seasonality and slight increasing trend (Figure 3a). Over the 1-10 m and 10-20 m depth (Figure 3b,c), most models produce increasing firn density apart from IMAU-FDM in which the firn density slightly decreases. All models agree relatively well on the average density independent of the depth range, with a maximum standard deviation among models of 15 kg m^{-3} for the top 1 m average density (of 336 kg m^{-3}), 27 kg m^{-3} for the 1-10 m range (420 kg m^{-3} on average) and 23 kg m^{-3} for the 10-20 m range (542 kg m^{-3} on average) during the 15 year long simulation period (Figure 3). In comparison with a 1 m density profile from 2001 and a firn core from 2015, most models reproduce vertical variability in firn density within observation uncertainties (Figure 3de). The evaluation of the density profile reveals that IMAU-FDM underestimates firn density between 5 and 15 m depth.

Regarding firn temperature, in most models, seasonal skin temperature fluctuations drive firn temperature variability in the top few metres of the column. However, seasonal temperature fluctuations propagate much deeper in the DTU model while it is almost not visible in MeyerHewitt model (Figure 2b). This results in much lower R^2 when comparing these two models to firn temperature observation: 0.41 and 0.28 for DTU and MeyerHewitt respectively. This results from the numerical

strategy and/or thermal diffusivity used in these models. Models that have explicit formulation for deep meltwater infiltration (CFM-Cr, CFM-KM and UppsalaUniDeepPerc) have positive ME 0.6 to 0.7 ^{o}C. This is due to the simulation of short-lived deep percolation events that infiltrates the minor melt from the surface down to 5 m, and to the subsequent refreezing and latent heat release. DMIHH, GEUS, IMAU-FDM and UppsalaUniBucket provide the lowest ME compared to firn temperature observations (Figure 2c). Yet, it should be noted that IMAU-FDM calculates adequate heat diffusion while underestimating the firn density (Figure 3e). Either the firn density underestimation in IMAU-FDM is not sufficient to induce a noticeable change in thermal conductivity or the thermal conductivity and/or numerical scheme used by IMAU-FDM compensates for the underestimated density and result in adequate simulated firn temperature.

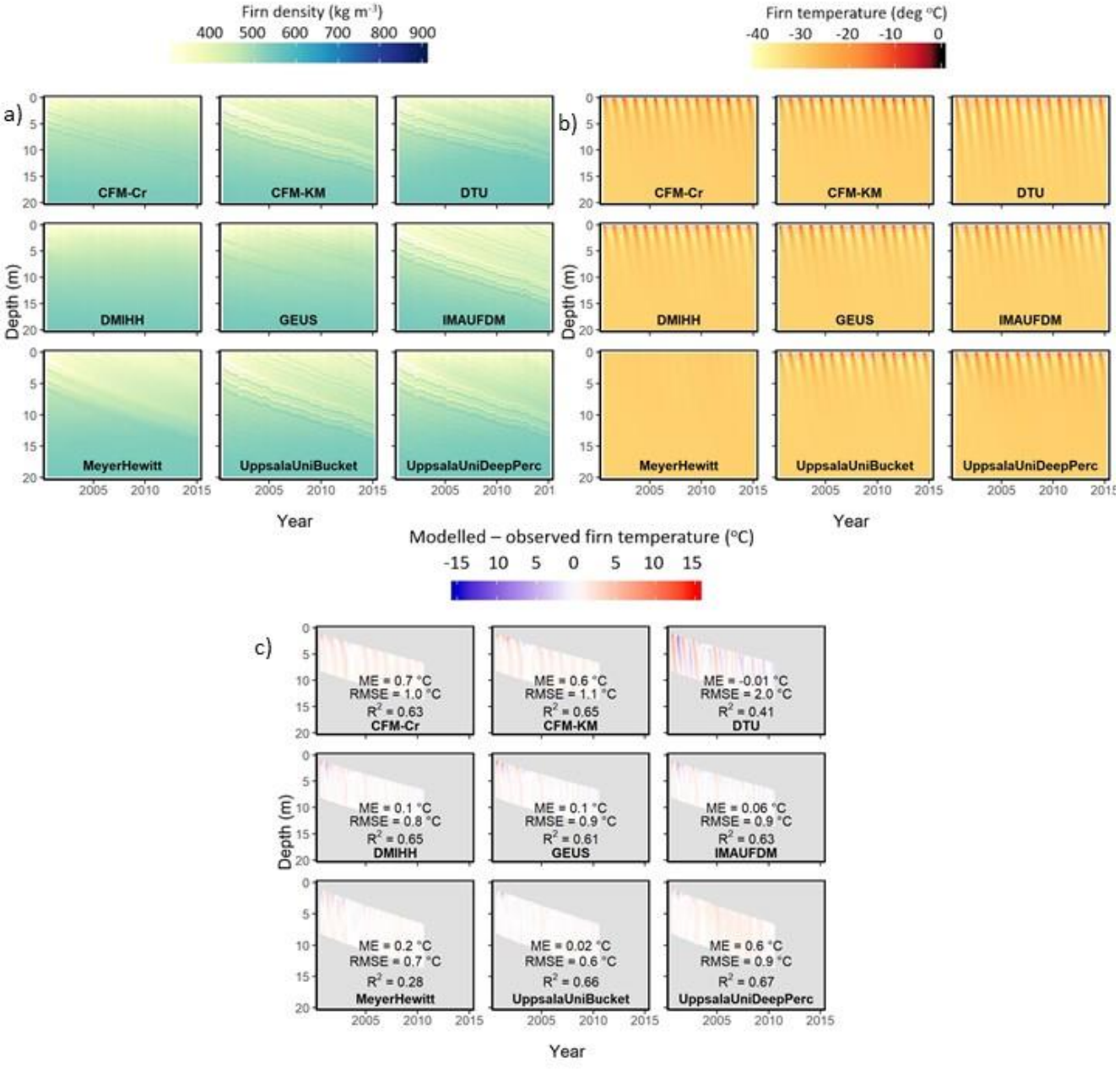

Figure 2: Simulated firn density (a), temperature (b) and deviation between simulated and observed firn temperature (c) at Summit.

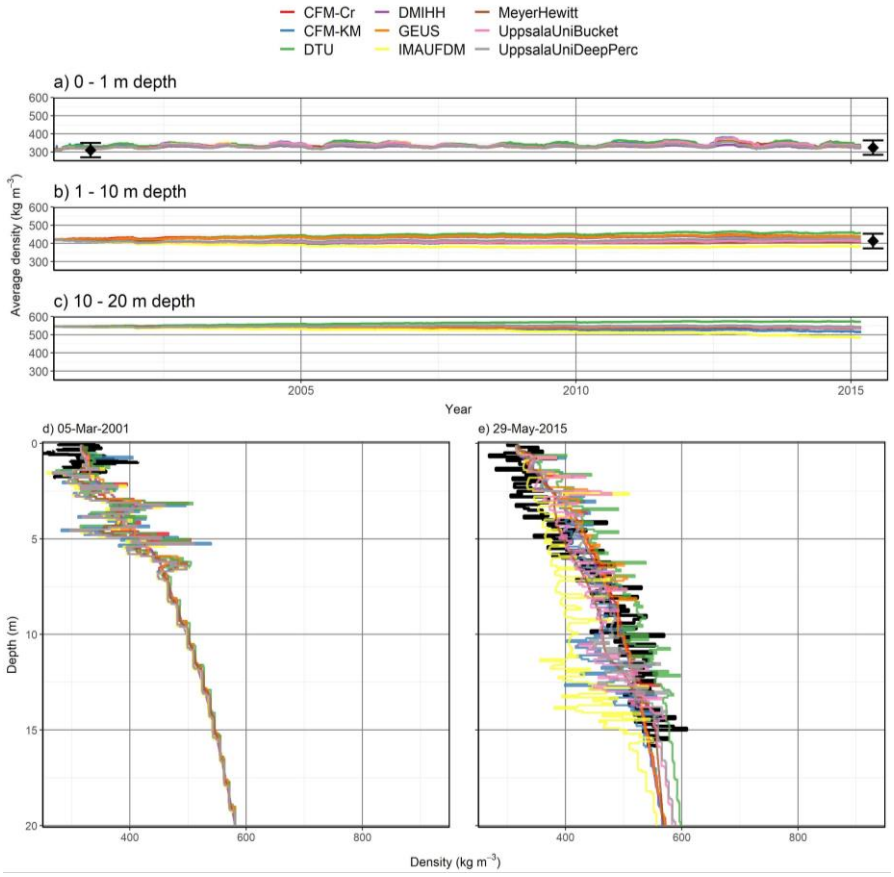

Figure 3: Modelled (coloured lines) and observed (black dots with 40 kg m-3 uncertainty bars) average firn density for the top 1 m (a), for the 1-10 m depth range (b) and 10-20 depth range (c) at Summit. Comparison of simulated and observed firn density in firn cores (d-f).

4.2. Percolation site: Dye-2

At Dye-2 surface melt occurs every summer. Consequently, refreezing of percolating meltwater has a significant effect on simulated density and temperature. The investigated models span a large spectrum of meltwater infiltration strategies (Table 2), leading to greater differences between models in firn density, temperature and liquid water content (Figure 4). Simulated meltwater percolation depth varies greatly among the models (Figure 4c). At one end of the spectrum, the DTU model only allows meltwater in the top model layer; an ice layer is built right at the start of the simulation and water is not able to penetrate ice layers in this model. At the other end, CFM-Cr and CFM-KM, which do allow meltwater to pass through ice layers and explicitly account for fast 'preferential flow', simulate percolation down to 10 m depth. In between these end-member models, UppsalaUniDeepPerc simulates percolation, up to ~5 m depth. IMAUFDM, UppsalaUniBucket, DMIHH and GEUS models give similar results and percolate water down to 1-3 m.

These differences in meltwater infiltration, when accumulated over a 17-years-long run, lead to large differences in firn density and temperature evolution across models (Figure 4). Models that include deep water infiltration (CFM-Cr, CFM-KM and UppsalaUniDeepPerc) build up a thick high-density layer at 3-10 m depth. In contrast, DTU, GEUS, IMAUFDM and UppsalaUniBucket simulate thinner, high-density, layers that form each summer at the surface and are buried in the following months and years. These sharp contrasts between low- and high-density layers are diluted in the Eulerian DMIHH and MeyerHewitt models. The long term simulated firn temperature evolution at Dye-2 (Figure 4b) and therefore the comparison of simulated temperatures to observations (Figure 4d) responds closely to the simulated meltwater infiltration each summer (Figure 4c). Models that include explicitly deep percolation (CFM-Cr, CFM-Kr, UppsalaUniDeep) also present the greatest firn warming at depth, due to refreezing and latent heat release (Figure 4b), and consequently have a positive ME ranging from 3.6 ^{o}C to 6.2 ^{o}C (Figure 4d). The DTU model does not percolate meltwater deep into the firn (Figure 4c) and consequently firn temperature evolves only due to heat diffusion, which leads to a cold bias (ME = -1.6 ^{o}C, Figure 4d). The remaining models (DMIHH, GEUS, IMAU-FDM, UppsalaUniBucket and MeyerHewitt) simulate similar inter-annual variability in meltwater infiltration and similar performance in firn temperature with a ME $< \pm 1$ ^{o}C and $R^2 > 0.5$.

The impact of these different infiltration patterns on the long term evolution of the average firn density and how simulated firn density compares to observation are presented in Figure 5. The standard deviation (model spread) of density reaches 161 kg m^{-3} in the top meter of firn and 141 kg m^{-3} for the 1-10 m layer (Figure 5). Lower deviation (29 kg m^{-3}) between 10-20 m stems from the limited time span of the simulation that does not allow the advection of the portion of firn where models disagree below 10 m depth (Figure 4 and 5).

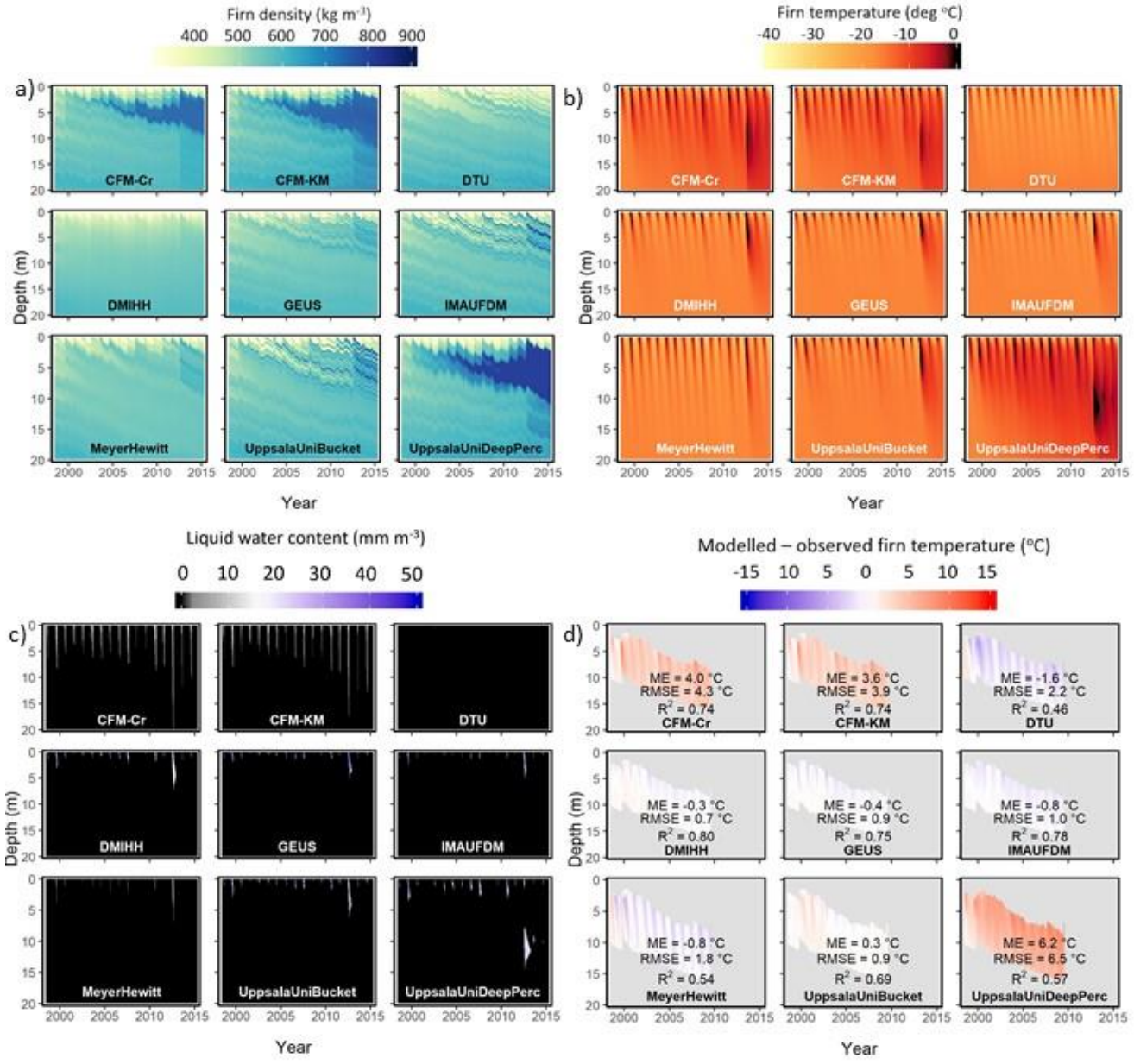

Figure 4: Simulated firn density (a), temperature (b), water content (note different y-axis) (c) and deviation between simulated and observed firn temperature (d) at Dye-2_long.

335

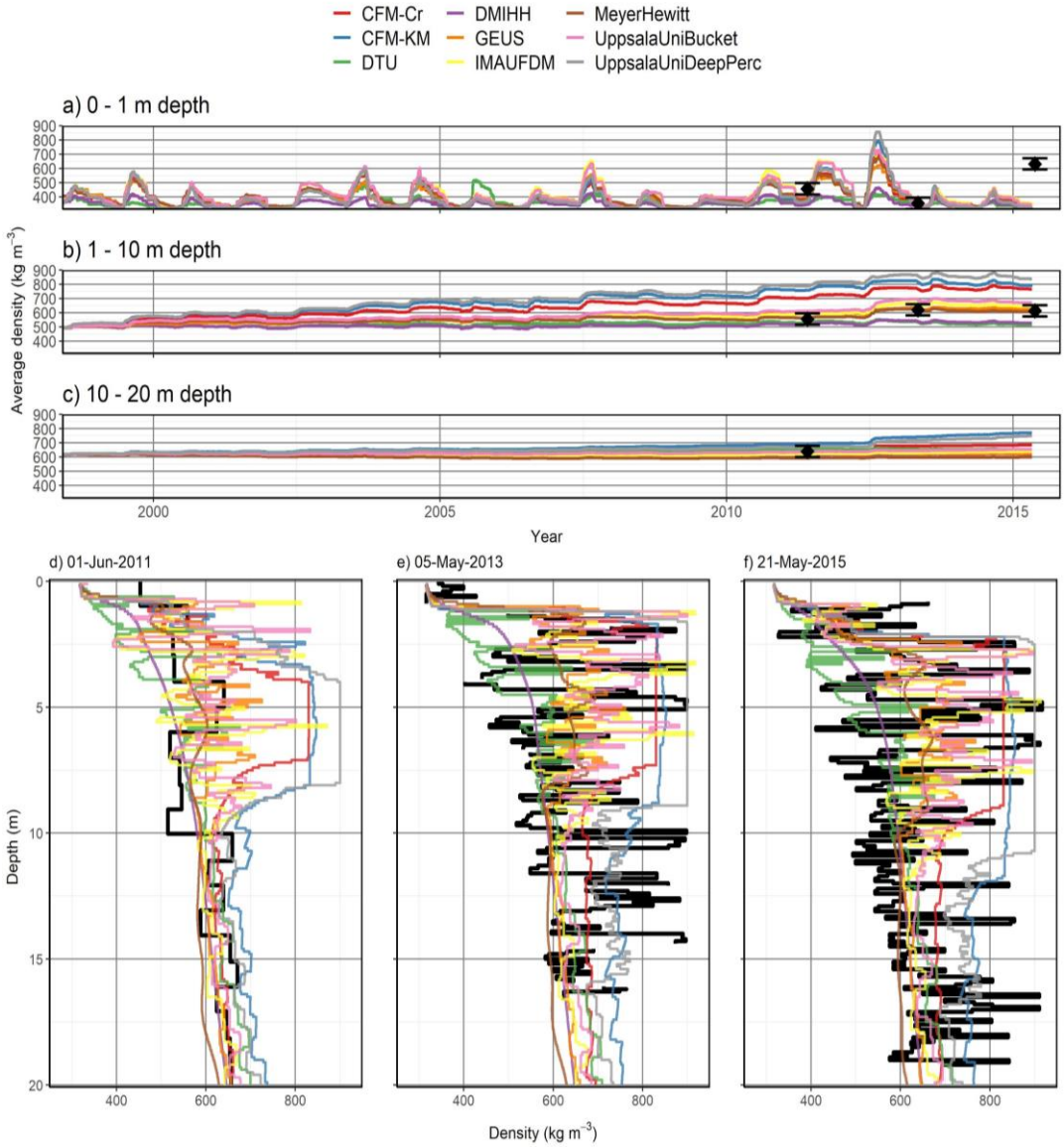

Figure 5: Modelled (coloured lines) and observed (black dots with 40 kg m-3 uncertainty bars) average firn density for the top 1 m (a), for the 1-10 m depth range (b) and 10-20 depth range (c) at Dye-2. Observed and simulated vertical variability in density at Dye-2 (d-g).

340

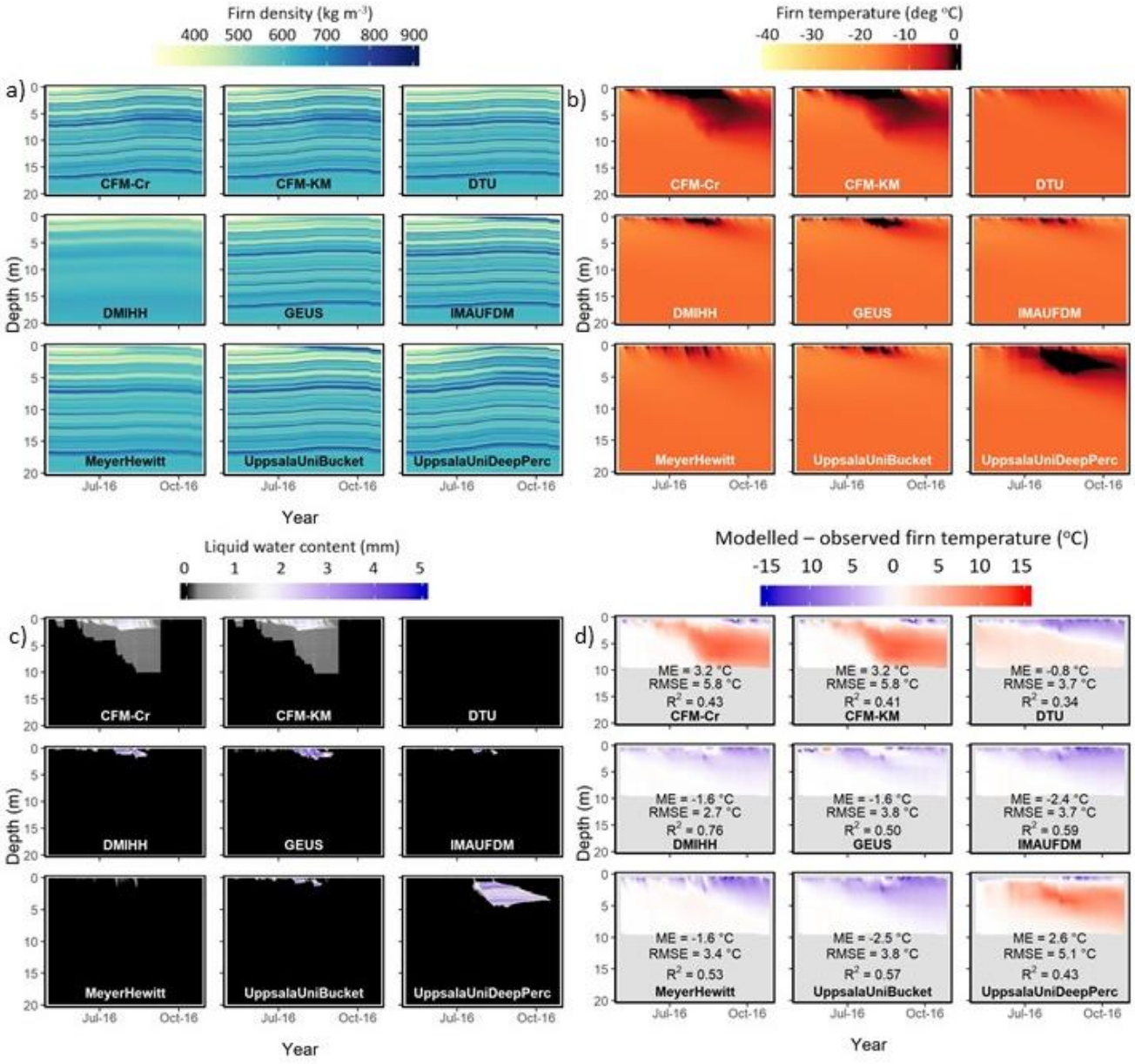

Figure 6: Simulated firn density (a), temperature (b), water content (note different y-axis) (c) and deviation between simulated and observed firn temperature (d) in Dye-2_16.

345 The use of a more recent AWS to derive the climate forcing at Dye-2 in the summer of 2016 allows to assess the firn models and their infiltration schemes in the best conditions. Over a single melt season, the meltwater infiltration and refreezing does not produce drastic changes in the simulated density profiles (Figure 6a). Yet, the meltwater is distributed at different depths and with different timing depending on the model (Figure 6c). The dual-domain approach of CFM-Cr and CFM-KM is visible

with higher liquid water content close to the surface, corresponding to the matrix flow, and low water content infiltrating down to 10 m depth in the heterogenous percolation domain. UppsalaUniDeep, which also includes deep percolation, infiltrates water down to 5 m, deeper than the models using a parametrization of Darcy's law (DMIHH, GEUS models) and bucket scheme (IMAU-FDM, UppsalaUniBucket models) which do not show liquid water below ~2 m depth (Figure 6c). As a result of these differences in meltwater infiltration, the location of the meltwater refreezing, of the latent heat release and consequently of the firn temperature differs from model to model (Figure 6b). The deep percolation models (CFM-Cr, CFM-KM and UppsalaUniDeep) have a marked positive bias (ME>2.6 ºC). The DTU model, which does not infiltrate water below the first few layers show a cold bias in the top part 5 m of the firn, where all the other models simulated meltwater infiltration. All the other models simulate colder conditions than observed with ME ranging from -2.5 ºC in UppsalaUniBucket to -1.6 ºC in the GEUS model.

UpGPR observations (Figure 7) show that the meltwater did not reach below 2.5 m depth during the 2016 melt season. The melt was concentrated around three periods of increasing intensity between May and June and a period when meltwater was continuously present in the firn between 20 July and 25 September. Compared to the upGPR, the CFM-CR and CFM-KM models substantially overestimate percolation depth (Figure 7a, red and blue lines), suggesting that, in the current configuration, these models exaggerate the effects of preferential flow, at least at this location. The DTU model does not simulate any percolation, and the MeyerHewitt model simulates the presence of meltwater in short-lived, episodic pulses rather than the continuous presence of meltwater that the upGPR observed. The other models simulate a percolation depth and temporal behaviour closer to the upGPR observations.

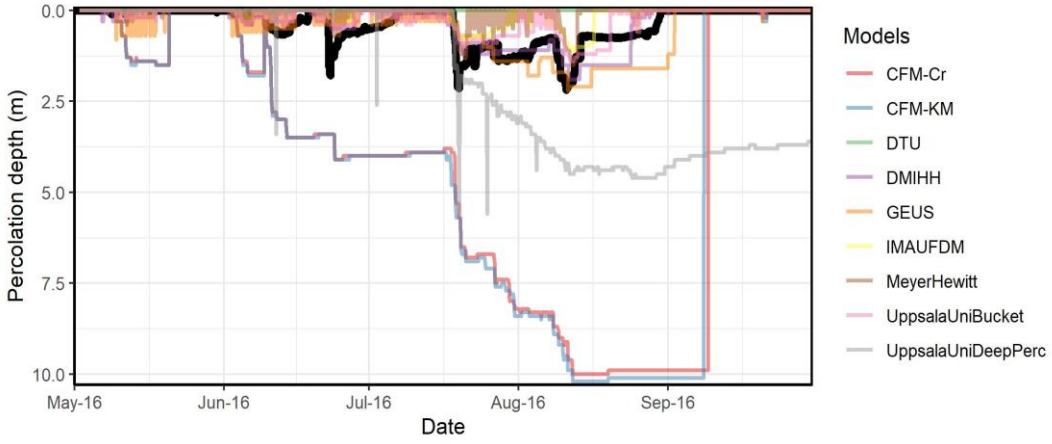

Figure 7: Comparison of the simulated (coloured lines) and upGPR-derived (black line) meltwater percolation depth at Dye-2 over the 2016 melting season. All three plots show the same period of melt-season evolution, with results grouped by models for clarity.

4.3. Ice-slab formation: KAN_U

At KAN-U, surface melt is more intense than at Dye-2. As a result, refreezing of infiltrated meltwater forms ice slabs that can be tens of centimetres to several metres thick. This site is therefore an interesting test for the firn models to see how they handle the presence of an ice-slab, and the effects of ice slabs on the vertical profiles of temperature and liquid water. Note that the firn models are initialized in spring 2012 with a pre-existing ice slab, which means that we do not assess the model capacity to form an ice slab: we only assess the effect of the ice slab on the evolution of the firn column.

The evolution of the density profile at KAN_U strongly depends on whether the model allows percolation past the ice slab (Figure 8a and 8c). The DMIHH model does not allow such percolation at all, and thus refreezing-related densification only occurs on top of the ice slab. As a consequence of no latent heat release below the ice slab in these models, the DMIHH model exhibits colder temperatures than observed (Figure 8b, 8c). Another group of models (CFM-Cr, CFM-KM, IMAUFDM, UppsalaUniBucket and UppsalaUniDeepPerc) do allow for percolation of meltwater through the ice slab, to depths of 10-15 m. As a result, the small amount of available pore space within the ice slab is used for refreezing and progressively filled (Figure 8a). Nevertheless, the sealing of the ice slab in these models does not prevent the meltwater from percolating through, and meltwater refreezing continues to occur at depth and to densify the firn there. These models overestimate deep firn temperatures compared to observations at KAN_U, presumably as a result of excess refreezing. In the MeyerHewitt and DMIHH models, the initial ice layers are gradually smoothed over time (Figure 9d-h). We relate this behavior to their Eulerian framework that implies frequent averaging of firn density and temperature when mass is added or removed from the model column. Still, they keep higher density between 5 and 10 m depth where the ice slab is. The model spread in top 1 m average density is minimal in the spring and increases in the summer (Figure 9a). The simulated average densities for 0-1, 1-10 and 10-20 m depth ranges compare well with punctual observations (Figure 9abc) but deviation between simulated and observed density profiles increase with time (Figure 9d-h). Comparison of the simulated firn temperature to hourly observations confirms that models including deep percolation (CFM-Cr, CFM-KM and UppsalaUniDeep) and bucket schemes (IMAUFDM and UppsalaUniBucket) infiltrate too much water at depth, resulting in a positive bias in temperature and a ME ranging from 1.8 to 4.6 ^{o}C. DTU and MeyerHewitt models do not show any meltwater infiltration or latent heat release at depth (Figure 8b, 8c). Consequently, they show lower firn temperature than observed with ME of -5.3 and -3.6 ^{o}C respectively. The GEUS model uses a low, but not null, permeability to ice layers and thus simulates reduced infiltration through the ice slab (Figure 8c) which leads, after this water refreezes, to firn temperatures closest to observations (ME = 0.6 ^{o}C).

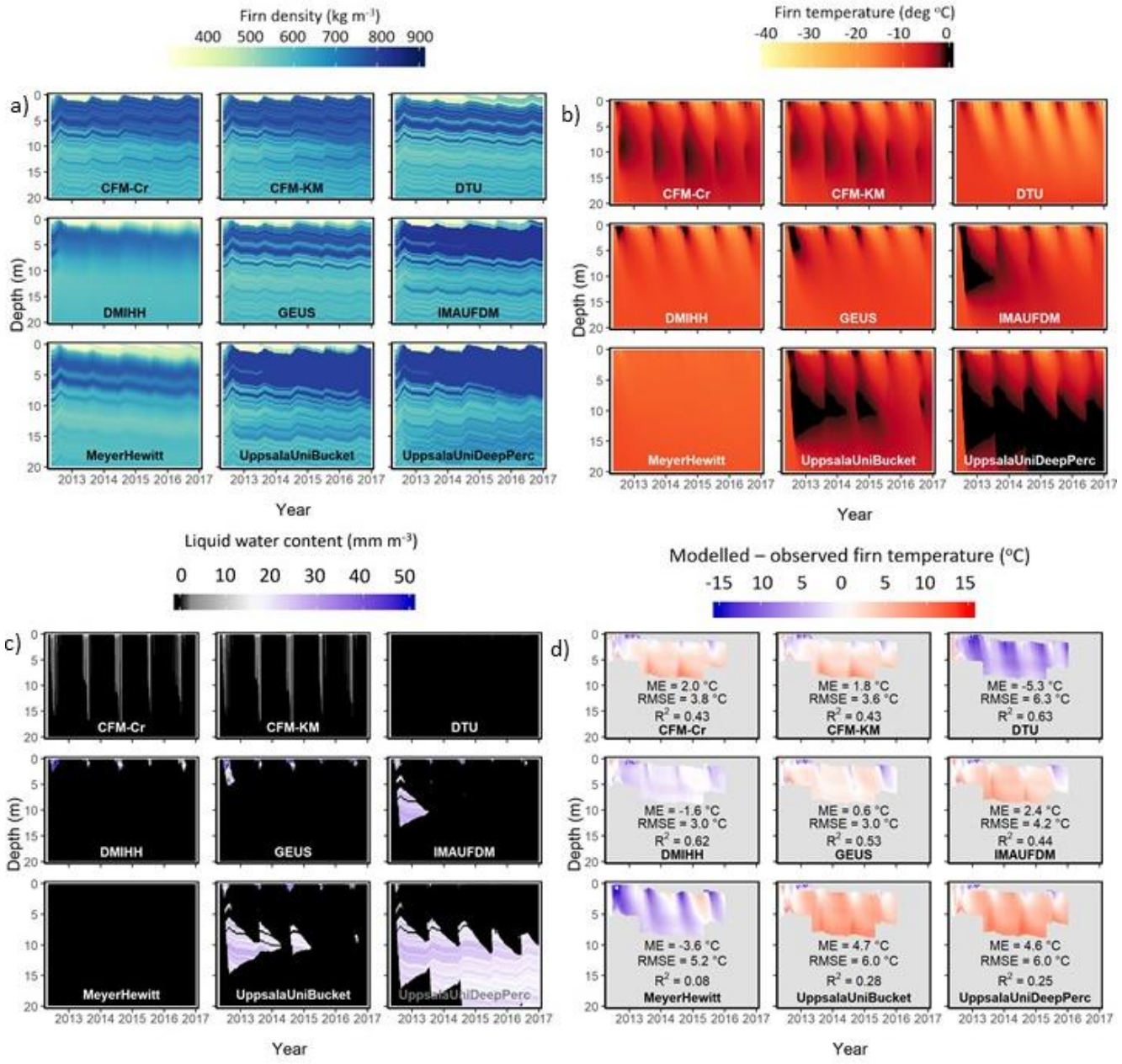

Figure 8: Simulated firn density (a), temperature (b), water content (note different y-axis) (c) and deviation between simulated and observed firn temperature (d) at KAN_U.

405

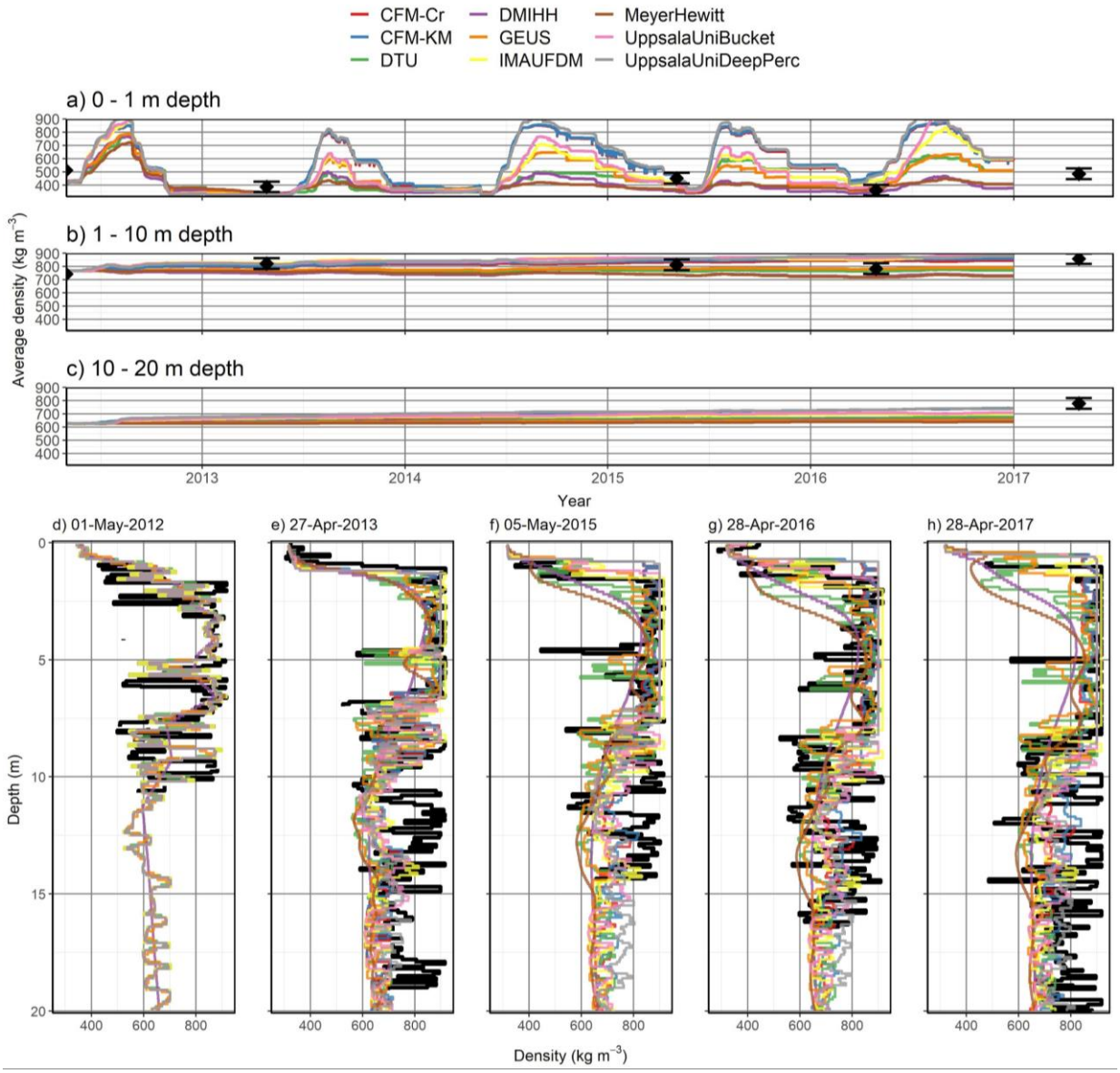

Figure 9: Modelled (coloured lines) and observed (black dots with 40 kg m-3 uncertainty bars) average firn density for the top 1 m (a), for the 1-10 m depth range (b) and 10-20 depth range (c) at KAN_U. Observed and simulated density at KAN_U (d-f).

4.4. Firn aquifers: FA site

At the firn aquifer site, both melting and snowfall are high, leading to perennial storage of liquid water within the firn pack. In terms of firn density, vertical gradients are similar among models but both the MayerHewitt and DMIHH models simulate smoother profiles (Figure 10a). This is likely due to their use of an Eulerian framework, as also seen in the results for KAN-U. Temporal evolution in density is also similar among models given the short span of the simulation. The DTU model

simulates slightly denser firn in the top few metres of the column as a result of refreezing (Figure 10a). Models which account for preferential flow in firn (both CFM models and UppsalaUniDeep) simulate meltwater infiltration to the aquifer, although with a slight difference in timing (Figure 10b, 10c). Unfortunately, the firn temperature observations do not allow us to ascertain how much water was transferred to the aquifer, but only that the whole firn column was at 0 ºC from mid-August to late September 2014, when cold surface temperature started to diffuse into the firn. These three deep percolation models overestimate shallow firn temperature in summer, and underestimate shallow firn temperature in winter, when compared to observations (Figure 10d). The DTU model simulates fast meltwater infiltration through the top 12 m and thus simulates a firn column entirely at 0 ºC (Figure 10b), in accordance with firn temperature observations (Figure 10d), but this meltwater runs off shortly after it percolates (Figure 10c). The other models simulate a firn column that is slightly too cold with ME between -0.1 and -0.6 ºC. As a result of the prescribed liquid water at depth in the initial conditions, deep firn temperature remains at melting point year-round in all models (Figure 10b), with liquid water at depth in all models except DTU (Figure 10c).

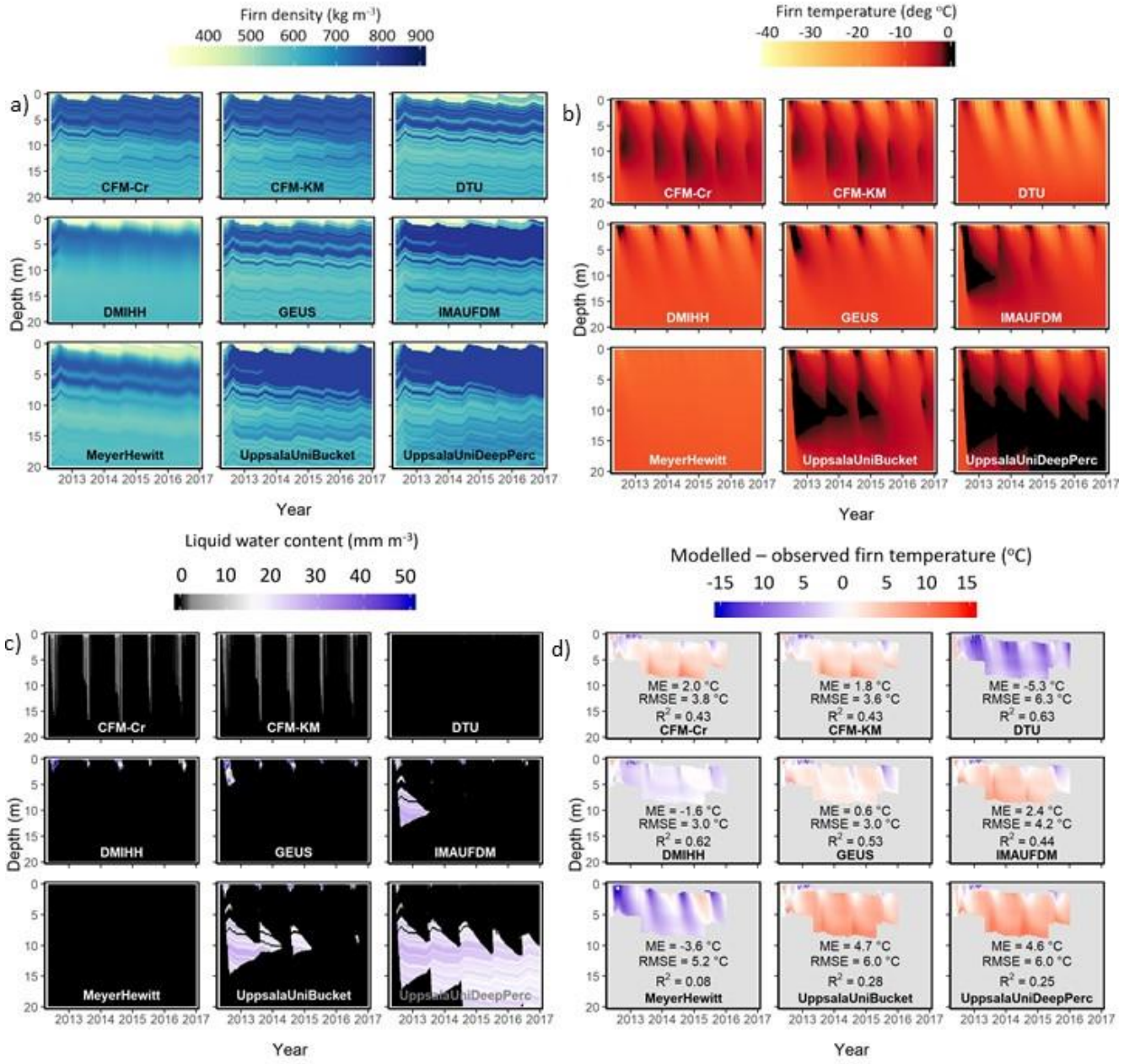

Figure 10: Simulated firn density (a), temperature (b), water content (c) and deviation between simulated and observed firn temperature (d) at FA.

5. Discussion

The variability in simulated firn density, temperature, and water content among the models, and the deviation between simulations and observations (Section 4) can be explained by the various ways physical processes are accounted for in the models. In this section we detail what can be learned from the comparison and we explore current knowledge gaps and potential improvements for firn models.

5.1. Dry firn and heat transfer

At Summit, comparisons with observations suggest that with appropriate forcing, the various densification formulations perform similarly and within observational uncertainty. The ability of firn models in the dry snow area to reproduce measured density profiles has been established from previous comparisons (Steger et al., 2017; Alexander et al., 2019), and can be attributed to the fact that most densification schemes are calibrated against firn density profiles from dry snow areas. The simulated densities at Summit show that densification schemes provide similar outputs, despite modelled temperatures spanning a wide range. Still, the ability of firn densification models to simulate firn changes in a transient climate is less certain (Lundin et al., 2017), and should remain a priority for future study. We also note that densifications schemes developed for dry firn are applied to wet-firn zones, and further research is needed to determine the validity of this assumption.

Models exhibit small but clearly discernible differences in firn temperature at Summit (Figure 2b). In our model experiments, downward advection due to accumulation was identical for all models, suggesting that this spread must be caused by the parameterization of thermal conductivity and/or the models' differing numerical schemes. Also, a suite of models exhibits colder temperatures compared with observations at Summit (DTU, DMIHH, GEUS, IMAUFDM, UppsalaUniBucket). We interpret this as an indication that heat transfer through the firn is still not accurately handled in most firn models. The heterogeneous nature of the firn, the presence of vertical ice features in the firn, the variability in surface snow density/thermal conductivity as well as firn ventilation are processes that are over-simplified or absent in the models and should be the subject of future research. Errors due to inaccurate estimates of thermal conductivity affect firn temperature, densification rates, and meltwater refreezing potential. We recommend that further work investigates potential improvements of the parameterization of thermal conductivity, either using recent studies (e.g., Calonne et al., 2019, Marchenko et al., 2019) or model calibration to observed firn temperature at dry firn locations. Other causes of model-data mismatch could be that certain processes (e.g. radiation penetration, Kuipers Munneke et al., 2009) are not provided to the models, and that uncertainty in the forcing data derived from AWS observations will propagate into the model simulations.

5.2. Meltwater percolation and refreezing

Many observational studies have demonstrated that there are two pathways for meltwater to infiltrate into the firn, namely by homogeneous wetting front, also called matrix flow, and by preferential flow through vertically extended channels (e.g. Marsh and Woo, 1984, Pfeffer and Humphrey, 1996). Some of the nine participating firn models do include both percolation regimes, and others do not. The lack of preferential flow routines has recently been described as a limitation of firn models (e.g. van As et al., 2016). Yet, little is known about how often this phenomenon occurs in the firn, how deep meltwater is transported, and which process triggers preferential flow. Here, the models that explicitly include deep percolation (CFM-Cr, CFM-KM and UppsalaDeepPerc) overestimate percolation depth and firn temperature at Summit, even though surface meltwater production at Summit is minimal. In their current configurations, the deep percolation schemes seem less adapted for areas with minor melt. Our results suggest that until the physics of preferential flow in firn are better understood, these more-complex models do not necessarily provide better results than simple bucket schemes. We recommend targeted field campaigns and lab studies to better understand preferential flow, and using those to constrain under which firn conditions and meltwater input deep percolation occurs. These steps are necessary to develop accurate deep-percolation schemes in firn models.

On the other hand, models that keep meltwater close to the surface - because they do not include any form of deep percolation - exhibit temperatures that are too cold compared with the observations at most sites (DTU, DMIHH, GEUS, IMAUFDM, UppsalaUniBucket). The cold bias could be due partly to an underestimation of thermal conductivity (section 5.1), but also due to insufficient meltwater percolation. The model evaluation at Dye-2 in 2016 indicates a reasonable percolation depth for all these models except DTU. It is conceivable that these models do simulate a reasonable percolation depth, but that the volume of percolating and refreezing meltwater is underestimated. Firn temperature observations and upGPR measurements can detect the presence of liquid water, but currently, no technique allows the vertically resolved measurement of water content. The models that use Darcy's law (CFM-Cr, CFM-KM, DMIHH, GEUS, MeyerHewitt) use different formulations for the firn permeability (Table 2) which also contribute to differences in meltwater percolation and refreezing results. Firn permeability can be related to grain size and firn density (Calonne et al., 2012). However, firn grain size and permeability observations are scarce, and these variables remain totally unconstrained in current models. Future model evaluation should include the existing data where available (e.g. Albert and Shultz, 2002) and more field observations of these grain-scale characteristics should be collected.

5.3. Ice slabs

The formation of ice slabs is a complex interplay between accumulation, densification, meltwater percolation, and refreezing (Machguth et al., 2016). Simulation of ice slabs by a firn model is therefore highly challenging, and success or failure to reproduce ice slabs depends on a number of processes that are closely linked and difficult to disentangle. Models that include

deep percolation (CFM-Cr, CFM-KM and UppsalaUniDeepPerc) grow an ice layer of several metres thickness close to the surface at Dye-2, where no such ice slabs are observed. This model behaviour can be explained by the simulation of water percolation bypassing ice layers and thus refreezing in cold underlying firn. At KAN_U, where ice slabs do exist, the DMIHH and GEUS models predict firn temperatures closest to the observations(lowest RMSE and highest R^2 for the DMIHH, lowest ME for GEUS) when compared to observations (Figure 8d). The performance of DMIHH at KAN_U can be explained by the absence of meltwater infiltration below the ice slab (Figure 8c) which agrees with recent field evidence of the ice slabs' impermeability (MacFerrin et al., 2019). In DMIHH, the blocking of percolation originates from a simple permeability criterion: if a layer's density is higher than 810 kg m^{-3}, then the layer is impermeable, and any incoming meltwater is sent to runoff. The choice of this value was based on work in Antarctica which found that firn permeability reaches zero over a range of densities centred on 810 kg m^{-3} (Gregory et al., 2014). Unfortunately, such studies remain scarce in Greenland and results do not provide a definite constraint on permeability (e.g. Albert and Schulz, 2002; Sommers et al., 2017). The DTU model uses a similar threshold density to characterize a layer's impermeability but found that 917 kg m^{-3} gave the best match with observed firn density profiles (Simonsen et al., 2013). On the contrary, the IMAU-FDM model assumes that, at the horizontal resolution on which it usually operates (1-25 km^2), ice layers can be assumed to be discontinuous and are therefore never impermeable. We note that the ice slab has a low, but not null, permeability as illustrated by rarely observed meltwater refreezing events within the ice slab (Charalampidis et al., 2016). Unfortunately, few observations are available to evaluate the effective permeability of ice slabs, both at local and regional scales and either confirm or contradict some of the assumptions made by the models. We recommend further investigation of the permeability of ice-dominated firn in relation to the firn density, the ice layer thickness and the various spatial and temporal scales at which the firn models are used.

Two models with a bucket-type percolation scheme, IMAUFDM and UppsalaUniBucket both use irreducible water content established by Coléou and Lesaffre (1998) from laboratory measurements. They consequently present similar percolation depths at KAN_U and Dye-2 (Figure 5 and 8). IMAUFDM and UppsalaUniBucket slightly underestimate percolation depth at Dye-2 in 2016 (Figure 6). This could indicate that the parametrization from Coléou and Lesaffre (1998), in combination with these firn models, overestimates irreducible water content, as suggested by Verjans et al. (2019). The current irreducible water content formulation could be complemented by observations in natural firn or adapted to the specific needs of bucket-scheme models. On one hand, meltwater routing in bucket-scheme models compare favourably to observations and to the DMIHH and GEUS models, which include more advanced meltwater routing schemes (Figure 6). On the other hand, the two bucket-scheme models both overestimate percolation in the presence of an ice slab, like at KAN_U: this is evident from a warm bias there, relative to the observations and to models that inhibit deep meltwater infiltration (Figure 8). Indeed, it was already established that they can overestimate percolation depth, and that more advanced routing schemes show slightly better performance in simulating meltwater runoff from alpine snowpacks (Wever et al. 2014). We therefore conclude that bucket schemes perform relatively well in the absence of ice slabs, but that accuracy in percolation depth could benefit from an improved representation of flow-impeding ice layers and from a slightly lower irreducible water content.

Finally, we make a note on discretization strategies of firn models. In Lagrangian models, the numerical grid follows the firn as layers get buried under accumulating snow. In Eulerian models the firn is being transferred through a fixed numerical grid. The Eulerian models, DMIHH and MeyerHewitt, smooth the firn density profile, reducing and dissipating contrasts in firn density (Figures 2, 4 and 8). This smoothing is not prevented by increased vertical resolution since MeyerHewitt has 18 times more layers than DMIHH. At KAN_U, these two models gradually lose the contrast between the layers that compose the ice slab and the firn below (Figure 8). Therefore, Eulerian models tend to represent ice slabs in terms of a depth range with increased density, rather than marked layers of ice. This limitation of Eulerian models does not prevent the DMIHH model from simulating adequately firn temperature at KAN_U (Figure 8d) and water infiltration at Dye-2 (Figure 6).

5.4. Firn aquifers

Like ice slabs, firn aquifers form in locations with a complex combination of accumulation, surface melt, percolation, and refreezing (Forster et al., 2014; Kuipers Munneke et al., 2014). Both the thermodynamic and the hydrological components of a firn model play an important role in its capacity to simulate firn aquifers.

As a general observation, aquifers are poorly represented in the firn models considered in this intercomparison, which poses the question of the suitability of the models to simulate aquifers in Greenland. For example, horizontal water flow at depth plays a crucial role in the evolution of firn aquifers (Miller et al., 2018). However, the nine models investigated here, and to our knowledge all firn models currently used to evaluate surface mass balance on the Greenland ice sheet, are one-dimensional. As such, lateral water movement in these models is governed by poorly constrained parameterizations, which are unlikely to accurately represent horizontal flow. Also, IMAUFDM and UppsalaUniBucket do not allow for the presence of water beyond the irreducible water content: after the initialization of these models, all the excess water within the aquifer is discarded as runoff instantaneously. As a result, these models are incapable of modelling actual aquifers (defined as saturated firn). Still, the regional climate model RACMO2, which includes IMAUFDM, has been used previously to map aquifers over the entire ice sheet (Forster et al., 2015). Areas where the model showed residual subsurface water (within the irreducible water content) remaining in spring was assumed to represent areas where firn aquifers might be present. Although this approach succeeded at mapping the current firn aquifer areas, the difference between what is tracked in the model and what actually happens at firn aquifer puts doubt on the current capacity of firn models to predict firn aquifer evolution in future climate. Other models show an intermediate type of behaviour: the DMIHH model runs off excess water according to the parametrization by Zuo and Oerlemans (1996). This leads to the gradual decrease of water content within the aquifer. The GEUS model incorporates a Darcy-like parametrization of the subsurface runoff, which results in faster drainage of the aquifer than the Zuo-Oerlemans parameterization. However, observations showed that excess water in the aquifer does not run off immediately but flows laterally and can remain in the aquifer for several decades (Miller et al., 2019).

Another challenging question for understanding and modelling of firn aquifers is: Where and when does the meltwater generated at the surface percolate down to the aquifer? Firn-temperature observations show that the top 20 m of firn remained at melting point during the 2014 melt season. This indicates that meltwater from the surface reached the aquifer. The firn models do not conclusively answer how and where deep percolation to the firn aquifer takes place. Given the same surface forcing and initial firn conditions, only the models with explicit deep-percolation schemes (CFM-Cr, CFM-KM and UppsalaUniDeepPerc) simulate water below 10 m depth. A simple interpretation from this result could be that the recharge of the firn aquifer has to be through heterogeneous percolation because it is the only way firn models can mimic observations. However, such a systematic infiltration through vertical channels should leave visible traces in the form of ice columns (Marsh and Woo, 1984) or show repeatedly in firn temperature observations when meltwater infiltrates into cold firn in spring (Pfeffer and Humphrey, 1996; Charalampidis et al,2016). Future field investigation should ascertain whether piping is indeed the only process infiltrating water to the aquifer. Another interpretation could be that models using a bucket scheme (DTU, IMAUFDM and UppsalaUniBucket) or Darcy's law (DMIHH, GEUS and MeyerHewitt) do not infiltrate water deep enough because of inappropriate irreducible water content or firn permeability for the firn aquifer site. The FA results reinforce our earlier suggestions that models need constraints on firn permeability, irreducible water content and occurrence of heterogeneous percolation. One last possibility could also be the misrepresentation of surface conditions: the melt calculated at the surface is subject to the biases and the uncertainties that apply to the so-called "bulk approach" used here in the energy budget calculation (Box and Steffen, 2001; Fausto et al., 2016). Although it was ensured that the calculated skin surface temperature agreed with observations available at KAN_U and FA, no direct observation of melt is available at our sites. Furthermore, the horizontal mobility of the meltwater, especially at high-melt sites such as FA, could lead to the injection at the surface of more meltwater than what is being melted. Therefore, more work is needed to quantify liquid water input at the top of the model in the firn aquifer region.

6. Towards ensemble-based uncertainty estimates for firn model outputs

Given the complexity of the firn models, it is difficult to propagate uncertainty and account for model assumptions and parameterisations. As a consequence, firn model outputs have commonly been given without uncertainty range which prevents assessing the robustness of model-based inferences. Taking inspiration fromn previous ensemble-based modelling approaches (e.g. Nowicki et al., 2016), we provide a multi-model estimation of the uncertainty that applies to any simulated value of firn temperature and density, and more importantly, to the simulated values of meltwater retention (through refreezing) and runoff.

6.1. Firn temperature and density uncertainty

We see from Figure 2 to 7 that the spread among models increases as we move from the dry snow area to the percolation area, peaking in areas with high-melt features such as ice slabs and firn aquifers. We suggest that the model spread presented here can provide a baseline for uncertainty whenever a single model is used. At Summit, representative of the dry snow area, modelled average density in the top metre of firn have a standard deviation of 13 kg m^{-3}. Hence, a two-standard deviation

(±2σ) uncertainty envelope of ±26 kg m^{-3}, or ±8%, can be used to describe the modelling uncertainty. At Dye-2, representative of the percolation area, the top 1 m average density simulated by the models have a maximum standard deviation of 145 kg m^{-3} during the 15 year-long simulation. This indicates that a substantial level of uncertainty, ±280 kg m^{-3}, or ±75%, applies to the modelled average density for the top metre. Similar uncertainty (±77%) applies to the modelled top 1 m average density at KAN_U. As for density, the models' spread in simulated firn temperature can be investigated by calculating the maximum standard deviation of firn temperature at 5 m depth among models. At Summit the ±2σ uncertainty envelope on simulated 5 m firn temperature is ±4ºC. This model uncertainty envelope is wider at Dye-2, ±14ºC because of the different meltwater infiltration depths simulated by the models. At KAN_U, the uncertainty in 5 m temperature, within the ice slab is ±10ºC. The uncertainty range increases closer to the surface and at sites or depths where meltwater infiltration may be captured differently by the models. The level of uncertainty, both for density and temperature, increases when narrowing the depth range over which averages are calculated, and conversely. This result indicates that firn models are still very variable when considering a specific depth but agree better when looking at the average firn property over a larger depth range. The uncertainty ranges provided here represent the largest deviation seen among models at each site and are therefore conservative. They can nevertheless be used as a metric for uncertainty in the absence of observational constraints or when using a single model.

6.2. Mass balance

The differences among simulated firn density, temperature, and liquid water distribution can cause them to retain and run off different amounts of meltwater and therefore affect the surface mass balance. All the models agree that all meltwater is retained at Summit. At Dye-2 and KAN_U, the inter-model average and ±2σ values can be used as a multi-model estimation of the meltwater retention, runoff and of the uncertainty on these estimates.

At Dye-2, the DTU model produces unrealistic runoff values (Figure 11c) because of the impermeability of thin ice layers blocking downward percolation and enhancing runoff. We therefore do not consider this model in our multi-model uncertainty estimation. All the other models agree that runoff is minimal compared to refreezing at Dye-2 (Figure 11abc). The bucket-scheme models (IMAUFDM and UppsalaUniBucket) retain all the meltwater generated at the surface. All the other models predict that runoff occurs regularly (Figure 11b), with a peak in 2012. Nevertheless, the uncertainty envelope applying to runoff always includes zero (Figure 11b). In years with absent or minor runoff, the annual refreezing totals reflect the inter-annual variability of surface melt and have uncertainties ranging from 3% to 24% of the average refreezing value depending on the year (Figure 11a).

At KAN_U, the impact of the ice slab on the surface mass balance is critical. The different simulated meltwater infiltration patterns (Figure 8c) lead to varying total amounts of meltwater either refrozen or runoff (Figure 11abc). The bucket schemes (IMAUFDM, UppsalaUniBucket) and UppsalaUniDeep percolate meltwater through the firn and all the meltwater refreezes

below the ice slab in these models. In all the other models, the presence of ice layers prevents or slows down meltwater infiltration, triggers ponding and lateral runoff, including in the CFM models where the preferential flow domain is unable to accommodate all the incoming water. The firn models' uncertainty on annual refreezing ranges from 28% in 2015 to 67% in 2016 relative to each year's average total refreezing. In 2012, the average runoff among models was 400 mm w.e., about 30% of the prescribed surface meltwater. For comparison, Machguth et al. (2016) calculated from firn cores that 75 ± 15% of the surface meltwater went to runoff at KAN_U in 2012. Although the observations are subject to considerable uncertainty, they indicate that most of the models underestimate the runoff at KAN_U in 2012. The spread in the model outputs leads to an uncertainty envelope which includes both zero runoff and the observed value (Figure 11f).

We do not evaluate meltwater retention and runoff at FA owing to the major limitations that we highlighted in the current handling of firn aquifers in firn models. In the percolation sites represented here by Dye-2 and KAN_U, the model spread generally increases with increasing surface melt and when more of that melts runs off. We therefore highlight the disagreement of the firn models in their simulations of the meltwater retention, refreezing, and runoff in the lower accumulation area of the ice sheet. High-melt accumulation areas should therefore be the subject of further field investigations to ascertain the actual meltwater retention there and better constrain firn models.

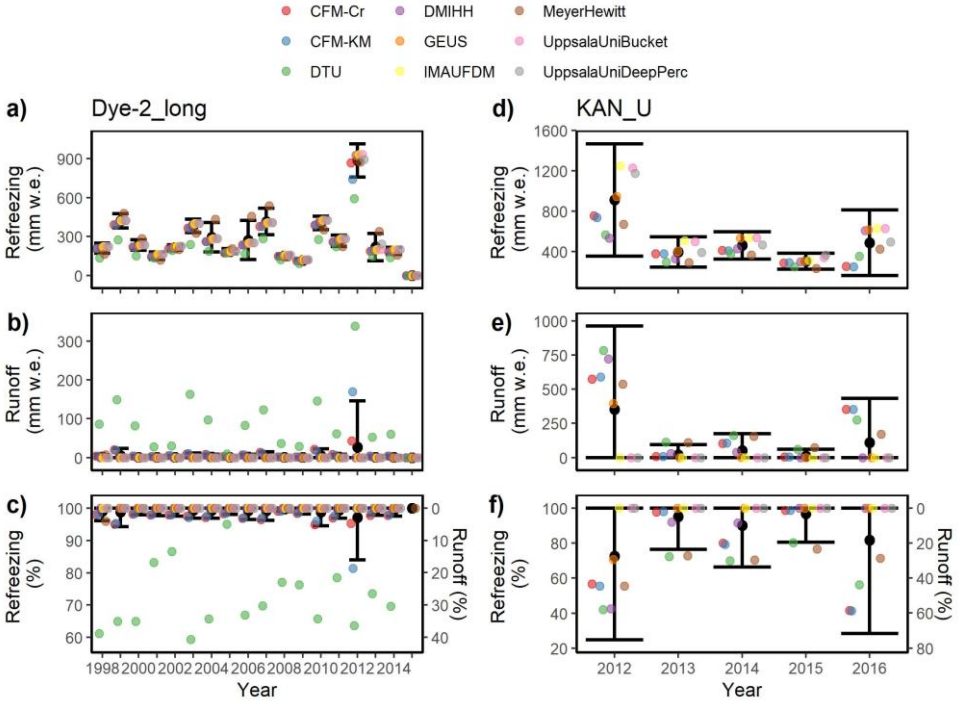

Figure 11. Yearly meltwater refreezing (a,d) or runoff (b,e), as totals (a,b,d,e) or fraction of the total meltwater input (c,f) at Dye-2 (a,b,c) and KAN_U (d,e,f). For each panel, yearly inter-model averages (black dots) and ±2σ values (error bars) are calculated from all models except the DTU model.

7. Summary remarks and perspectives

Nine state-of-the-art firn models were forced with mass and energy fluxes calculated from weather station data at four sites representative of various climate and firn zones of the Greenland ice sheet. From the intercomparison of their simulated firn temperature, density, and water content, and from evaluation against various firn observations, we identified specific routines within the models that are responsible for the models' behaviours. We later quantified uncertainties that apply to the firn model outputs and on their evaluation of meltwater retention. We identified key topics for future development of models and for the investigation of firn processes.

We identified the following disagreement among models and model-observation discrepancies. Runoff-enhancing ice slabs were formed in certain models at the Dye-2 site where they are not observed. At the KAN_U site, where models were initialized with a several-meter thick ice slab according to observations, models do not agree whether such ice layers allow meltwater infiltration or not. Models that explicitly include deep percolation allow water percolation through the ice slab, which disagrees with the relatively cold observed firn temperatures at depth. At the aquifer site, only deep-percolation models infiltrate meltwater to the aquifer. Nevertheless, all models misrepresent the aquifer either because of the inability of some models to simulate saturated conditions, the different time scales at which the excess water is sent to runoff, and the absence of horizontal subsurface water movement. At all sites, Eulerian models smooth the firn density profile and dissipate contrast in firn density even in a model with high vertical resolution. Further testing of such models should investigate how this numerical diffusion affects the modeled firn characteristics over longer runs and how runoff-enhancing ice slabs are represented in these models. Model spread and deviation between simulated and observed firn density and temperature is largest at the sites that experience more melt. Using twice the standard deviation in model outputs as an indicator of uncertainty envelope, we found that firn models can estimate the top 1 m average firn density within ±26 kg m^{-3} at a dry snow site and that uncertainty increases to ±280 kg m^{-3} at percolation sites. The similarity between modelled and observed firn density at the nearly melt-free Summit site indicates that for the top 20 m of firn, the densification equations perform similarly under dry-snow conditions given identical forcing. However, variability in simulated firn temperature at Summit indicates that heat transfer through the firn is still not handled consistently in firn models. Consequently, none of the tested models compared positively with observations at all four sites.

Differences in simulated firn characteristics in the nine models led to different amounts of meltwater being retained through refreezing or escaping the site through runoff. Models that percolate meltwater deeper (shallower) calculate higher (lower) retention through refreezing and therefore less (more) lateral runoff. The spread among models regarding annual meltwater retention is positively correlated with surface meltwater input and reaches 70% in 2012, the highest melt year at KAN_U. Still, during that year, the inter-model average runoff is 28% of the meltwater input. Therefore, assessment of model spread should be conducted at sites presenting higher fractions of meltwater running off.

These mixed results show that even the newest models need further development to perform satisfactorily under the wide range of climate and firn conditions of the Greenland ice sheet. We recommend the following topics for future investigations:

- More observations of firn permeability should be conducted both at point and regional scale. Measurements of grain size and other microstructural properties would also help to evaluate the parametrizations currently used by some of the firn models for permeability. These measurements should focus on the lower percolation area where meltwater infiltration and runoff play an important role in the surface mass balance.
- Bucket schemes, which do not calculate firn permeability, would benefit from a density-based impermeability criterion. This criterion needs to be drawn from field evidence at the scale the models operate.
- Recent work on firn thermal conductivity (e.g. Calonne et al., 2019, Marchenko et al., 2019) should also be used to improve the firn models. Furthermore, the impact of vertical ice features and firn ventilation on firn temperature is currently not included in any of the firn models. Firn temperature observations are now available to assess the model performance and should be part of the standard evaluation protocol.
- Eulerian models should be used bearing in mind that they gradually average firn characteristics. This issue does not prevent the use of such models, as long as the features that are being studied (e.g. ice slab, runoff, firn aquifer...) are being defined in ways that are compatible with the Eulerian framework.
- A major rethinking of firn models is necessary to better represent firn aquifer. In these regions, models need to allow saturated conditions and lateral subsurface water flow either explicitly with a multi-dimensional model or through an adapted parameterisation. More field observations are also needed to ascertain the surface meltwater input at these sites, whether near-surface drainage occurs and, if it does, the size of such drainage area.
- Recent efforts were made to explicitly describe heterogeneous meltwater infiltration in firn models. While they allowed better performance at the firn aquifer site, they infiltrate water too deeply and produce positive biases in firn temperature at the dry snow site and the two percolation sites. Further work is needed to understand, under various surface and firn conditions, when heterogeneous percolation occurs, how deep it should reach and how much water it should transport. Only after these questions are understood can a reliable preferential-flow be developed.
- The fresh snow density is known to have an impact on the firn model outputs but was here set to a site-invariant value derived from observations. Fresh snow density is known to vary considerably in space and time although no statistically robust parameterisation exists up to date for the Greenland ice sheet. Future measurement campaigns and modelling efforts could help to prescribe surface snow density and understand how it interacts with the densification and heat transfer scheme.

Considering the number of firn characteristics that remain to be investigated and the cost of field surveys, laboratory experiments could be highly valuable if they can address the boundary effects, the scale of the process being investigated, and provide realistic surface and firn conditions. Investigation of the points listed above will collectively improve our understanding of firn and meltwater dynamics, improve the representation of these processes in firn models, and eventually

reduce the uncertainty that applies to their output when assessing the surface mass balance of the Greenland ice sheet in past, present, and future times.

8. Data availability

The forcing data sets and all the model output is available on https://www.promice.org/PromiceDataPortal/. The code for all the plots are available on https://github.com/BaptisteVandecrux/RetMIP. The source code for the CFM model is available at https://github.com/UWGlaciology/CommunityFirnModel; the GEUS model code can be found at https://github.com/BaptisteVandecrux/SEB_Firn_model. The RetMIP protocol is available at http://retain.geus.dk/index.php/retmip/.

9. Funding

This work is part of the Retain project funded by the Danish Council for Independent Research (Grant no. 4002-00234) and the Programme for Monitoring of the Greenland Ice Sheet (www.PROMICE.dk). Achim Heilig was supported by DFG grant HE 7501/1-1, Horst Machguth acknowledges support by ERC CoG Nr. 818994 The AWS used at Dye-2 during the 2016 melt season is supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada. ArcTrain and Arctic Institute of North America (NSTP). Olivia Miller and Clifford Voss were supported by the U.S. Geological Survey. C. Max Stevens and Michael MacFerrin were supported by the National Aeronautics and Space Administration (NASA) grant NNX15AC62G.

10. Acknowledgement

We are grateful to Ian Hewitt for his insight on the MeyerHewitt model. We thank our scientific editor Xavier Fettweis as well as Samuel Morin, Kendall FitzGerald and an anonymous reviewer for comments and suggestions that significantly improved the study.

References

Ahlstrøm, A. P., Gravesen, P., Andersen, S. B., van As, D., Citterio, M., Fausto, R. S., Nielsen, S., Jepsen, H. F., Kristensen, S. S., Christensen, E. L., Stenseng, L., Forsberg, R., Hanson, S. and Petersen, D.: A new programme for monitoring the mass loss of the Greenland ice sheet, Geol. Surv. Denmark Greenl. Bull., (15), 61–64 http://www.geus.dk/DK/publications/geol-survey-dk-gl-bull/15/Documents/nr15_p61-64.pdf , 2008.

Albert, M.R. and Shultz, E.F.: Snow and firn properties and air–snow transport processes at Summit, Greenland. Atmospheric Environment, 36(15-16), pp.2789-2797, 2002.

Alexander, P. M., Tedesco, M., Koenig, L. and Fettweis, X.: Evaluating a Regional Climate Model Simulation of Greenland Ice Sheet Snow and Firn Density for Improved Surface Mass Balance Estimates, Geophys. Res. Lett., 46(21), 12073–12082, https://doi.org/10.1029/2019GL084101, 2019.

Anderson, E. A.: A point energy and mass balance model of a snow cover. [online] Available from: http://www.csa.com/partners/viewrecord.php?requester=gs&collection=ENV&recid=7611864%5Cnhttp://www.agu.org/pubs/crossref/2009/2009JD011949.shtml, 1976.

Arthern, R. J., Vaughan, D. G., Rankin, A. M., Mulvaney, R. and Thomas, E. R.: In situ measurements of Antarctic snow compaction compared with predictions of models, J. Geophys. Res. Earth Surf., 115(3), 1–12, https://doi.org/10.1029/2009JF001306, 2010.

Benson, C.S.: Stratigraphic studies in the snow and firn of the Greenland ice sheet. SIPRE Res. Rep. 70, 76–83, 1962.

Box, J. E., and Steffen, K.: Sublimation on the Greenland Ice Sheet from automated weather station observations, J. Geophys. Res., 106(D24), 33965– 33981, doi:10.1029/2001JD900219, 2001.

Box, J. E.: Greenland ice sheet mass balance reconstruction. Part II: Surface mass balance (1840-2010), J. Clim., 26(18), 6974–6989, https://doi.org/10.1175/JCLI-D-12-00518.1, 2013.

Box, J. E., Cressie, N., Bromwich, D. H., Jung, J. H., Van Den Broeke, M., Van Angelen, J. H., Forster, R. R., Miège, C., Mosley-Thompson, E., Vinther, B. and Mcconnell, J. R.: Greenland ice sheet mass balance reconstruction. Part I: Net snow accumulation (1600-2009), J. Clim., 26(11), 3919–3934, https://doi.org/10.1175/JCLI-D-12-00373.1, 2013.

Bear, J.:Dynamics of fluids in porous media, Dover, 1972.

Braithwaite, R. J., Laternser, M. and Pfeffer, W. T.: Variations of near-surface firn density in the lower accumulation area of the Greenland ice sheet, Pakitsoq, West Greenland, J. Glaciol., 40(136), 477–485, https://doi.org/10.1017/S002214300001234X, 1994.

Briggs, M. A., Walvoord, M. A., McKenzie, J. M., Voss, C. I., Day-Lewis, F. D., & Lane, J. W.: New permafrost is forming around shrinking Arctic lakes, but will it last? Geophys. Res. Lett., 1585–1592. https://doi.org/10.1002/2014GL059251, 2014.

Calonne, N., Flin, F., Morin, S., Lesaffre, B., Du Roscoat, S. R. and Geindreau, C.: Numerical and experimental investigations of the effective thermal conductivity of snow, Geophys. Res. Lett., 38(23), 1–6, https://doi.org/10.1029/2011GL049234, 2011.

Calonne, N., Geindreau, C., Flin, F., Morin, S., Lesaffre, B., Roscoat, S. R. and Charrier, P. : 3-D image-based numerical computations of snow permeability: Links to specific surface area, density, and microstructural anisotropy. Cryosphere, 6(5), 939–951, https://doi.org/ 10.5194/tc-6-939-2012, 2012.

Charalampidis, C., Van As, D., Box, J. E., Van Den Broeke, M. R., Colgan, W. T., Doyle, S. H., Hubbard, A. L., MacFerrin, M., Machguth, H. and P. Smeets, C. J. P.: Changing surface-atmosphere energy exchange and refreezing capacity of the lower accumulation area, West Greenland, Cryosphere, 9(6), 2163–2181, https://doi.org/10.5194/tc-9-2163-2015, 2015.

Colbeck, S. C.: A theory for water flow through a layered snowpack, Water Resour. Res., 11(2), 261–266, https://doi.org/10.1029/WR011i002p00261, 1975.

Coléou, C. and Lesaffre, B.: Irreducible water saturation in snow: experimental results in a cold laboratory, Ann. Glaciol., 26(2), 64–68, https://doi.org/10.3189/1998aog26-1-64-68, 1998.

Daanen, R. and Nieber, J.: Model for Coupled Liquid Water Flow and Heat Transport with Phase Change in a Snowpack, J. Cold Reg. Eng., 23, 43–68, doi:10.1061/(ASCE)0887-381X(2009)23:2(43), 2009.

Dibb, J.E. and Fahnestock, M.,: Snow accumulation, surface height change, and firn densification at Summit, Greenland: Insights from 2 years of in situ observation. *Journal of Geophysical Research: Atmospheres*, *109*(D24), 2004.

Evans, S. G., Godsey, S. E., Rushlow, C. R., & Voss, C. I.: Water tracks enhance water flow above permafrost in upland Arctic Alaska hillslopes, J. Geophys. Res.: Earth Surf., https://doi.org/10.1029/2019JF005256, 2020.

Fausto, R. S., van As, D., Box, J. E., Colgan, W., Langen, P. L., and Mottram, R. H.: The implication of nonradiative energy fluxes dominating Greenland ice sheet exceptional ablation area surface melt in 2012, Geophys. Res. Lett., 43, 2649– 2658, doi:10.1002/2016GL067720, 2016.

Fausto, R. S., Box, J. E., Vandecrux, B., van As, D., Steffen, K.,MacFerrin, M., Machguth H., Colgan W., Koenig L. S., McGrath D., Charalampidis, C., and Braithwaite, R. J.: A Snow Density Dataset for Improving Surface Boundary Conditions in Greenland Ice Sheet Firn Modeling, Front. Earth Sci., 6, 51 pp.,https://doi.org/10.3389/feart.2018.00051, 2018

Fausto, R. S., Ahlstrøm, A. P., Van As, D., Johnsen, S. J., Langen, P. L. and Steffen, K.: Improving surface boundary conditions with focus on coupling snow densification and meltwater retention in large-scale ice-sheet models of Greenland, J. Glaciol., 55(193), 869–878, https://doi.org/10.3189/002214309790152537, 2009.

Ge, S., McKenzie, J., Voss, C., & Wu, Q.: Exchange of groundwater and surface-water mediated by permafrost response to seasonal and long term air temperature variation. Geophys.l Res. Lett., 38(14), https://doi.org/10.1029/2011GL047911, 2011.

Gregory, S. A., Albert, M. R., and Baker, I.: Impact of physical properties and accumulation rate on pore close-off in layered firn. Cryosphere 8,91–105. https://doi.org/10.5194/tc-8-91-2014. 2014.

Heilig, A., Eisen, O., MacFerrin, M., Tedesco, M., and Fettweis, X.: Seasonal monitoring of melt and accumulation within the deep percolation zone of the Greenland Ice Sheet and comparison with simulations of regional climate modeling, The Cryosphere, 12, 1851–1866, https://doi.org/10.5194/tc-12-1851-2018, 2018.

Herron, M. M. and Langway, C. C.: Firn Densification: An Empirical Model, J. Glaciol., 25(93), 373–385, https://doi.org/10.3189/s0022143000015239, 1980.

Hirashima, H., Yamaguchi, S., Sato, A. and Lehning, M.: Numerical modeling of liquid water movement through layered snow based on new measurements of the water retention curve, Cold Reg. Sci. Technol., 64(2), 94–103, https://doi.org/10.1016/j.coldregions.2010.09.003, 2010.

Howat, I. M., Negrete, A. and Smith, B. E.: The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets, Cryosphere, 8(4), 1509–1518, https://doi.org/10.5194/tc-8-1509-2014, 2014.

Kameda, T., Narita, H., Shoji, H., Nishio, F., Fujii, Y. and Watanabe, O.: Melt features in ice cores from Site J, southern Greenland: some implications for summer climate since AD 1550, Ann. Glaciol., 21, 51–58, https://doi.org/10.3189/s0260305500015597, 1995.

Katsushima, T., Kumakura, T. and Takeuchi, Y.: A multiple snow layer model including a parameterization of vertical water channel process in snowpack, Cold Reg. Sci. Technol., 59(2–3), 143–151, https://doi.org/10.1016/j.coldregions.2009.09.002, 2009.

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(1), 81–85, https://doi.org/10.1002/2013GL058083, 2014.

Kuipers Munneke, P., Van den Broeke, M.R., Reijmer, C.H., Helsen, M.M., Boot, W., Schneebeli, M. and Steffen, K.,: The role of radiation penetration in the energy budget of the snowpack at Summit, Greenland. *The Cryosphere*, *3*, pp.155-165, 2009.

Kuipers Munneke, P., M. 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. Geophysical Research Letters, 41(2), pp.476-48, 2014.

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, Cryosphere, 9(6), 2009–2025, https://doi.org/10.5194/tc-9-2009-2015, 2015.

Kurylyk, B. L., MacQuarrie, K. T. B., & Voss, C. I.: Climate change impacts on the temperature and magnitude of groundwater discharge from shallow, unconfined aquifers. Water Resources. Res., 3253–3274. https://doi.org/10.1002/2013WR014588, 2014.

Lefebre, F., Gallée, H., van Ypersele, J.-P. and Greuell, W.: Modeling of snow and ice melt at ETH Camp (West Greenland): A study of surface albedo, J. Geophys. Res., 108(D8), 4231, https://doi.org/10.1029/2001JD001160, 2003

Ligtenberg, S. R. M., Helsen, M. M. and Van Den Broeke, M. R.: An improved semi-empirical model for the densification of Antarctic firn, Cryosphere, 5(4), 809–819, https://doi.org/10.5194/tc-5-809-2011, 2011.

Ligtenberg, S. R. M., Munneke, P. K., Noël, B. P. Y. and Van Den Broeke, M. R.: Brief communication: Improved simulation of the present-day Greenland firn layer (1960-2016), Cryosphere, 12(5), 1643–1649, https://doi.org/10.5194/tc-12-1643-2018, 2018.

Lundin, J. M. D., Stevens, C. M., Arthern, R., Buizert, C., Orsi, A., Ligtenberg, S. R. M., Simonsen, S. B., Cummings, E., Essery, R., Leahy, W., Harris, P., Helsen, M. M. and Waddington, E. D.: Firn Model Intercomparison Experiment (FirnMICE), J. Glaciol., 63(239), 401–422, https://doi.org/10.1017/jog.2016.114, 2017.

MacFerrin, M., Machguth, H., As, D. van, Charalampidis, C., Stevens, C. M., Heilig, A., Vandecrux, B., Langen, P. L., Mottram, R., Fettweis, X., Van Den Broeke, M. R., Pfeffer, W. T., Moussavi, M. S. and Abdalati, W.: Rapid expansion of Greenland's low-permeability ice slabs, Nature, 573(7774), 403–407, https://doi.org/10.1038/s41586-019-1550-3, 2019.

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. Chang., 6(4), 390–393, https://doi.org/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(March), https://doi.org/10.3389/feart.2017.00016, 2017.

Marchenko, S., Cheng, G., Lötstedt, P., Pohjola, V., Pettersson, R., Van Pelt, W. and Reijmer, C.: Thermal conductivity of firn at Lomonosovfonna, Svalbard, derived from subsurface temperature measurements. *Cryosphere*, *13*(7), pp.1843-1859, 2019.

Marsh, P. and Woo, M.K.,: Wetting front advance and freezing of meltwater within a snow cover: 1. Observations in the Canadian Arctic. Water Resources Research, 20(12), pp.1853-1864, 1984.

Mayewski, P. and Whitlow S.: Snow Pit Data from Greenland Summit, 1989 to 1993, NSF Arctic Data Center, https://doi.org/10.5065/D6NP22KX, 2016b.

Meyer, C. R. and Hewitt, I. J.: A continuum model for meltwater flow through compacting snow, Cryosphere, 11(6), 2799–2813, https://doi.org/10.5194/tc-11-2799-2017, 2017.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 Surf., 121(12), 2381–2398, https://doi.org/10.1002/2016JF003869, 2016.

McKenzie, J.M., Voss, C.I. and Siegel, D.I., 2007. Groundwater flow with energy transport and water–ice phase change: numerical simulations, benchmarks, and application to freezing in peat bogs. *Advances in water resources*, *30*(4), pp.966-983.

Miller, O., Solomon, D.K., Miège, C., Koenig, L., Forster, R., Schmerr, N., Ligtenberg, S.R. and Montgomery, L.: Direct evidence of meltwater flow within a firn aquifer in southeast Greenland. *Geophysical Research Letters*, *45*(1), pp.207-215, 2018.

Miller, O., Solomon, D.K., Miège, C., Koenig, L., Forster, R., Schmerr, N., Ligtenberg, S.R., Legchenko, A., Voss, C.I., Montgomery, L. and McConnell, J.R., Hydrology of a perennial firn aquifer in Southeast Greenland: an overview driven by field data. *Water Resources Research*, p.e2019WR026348, 2019.

Montgomery, L., Koenig, L. and Alexander, P.: The SUMup dataset: Compiled measurements of surface mass balance components over ice sheets and sea ice with analysis over Greenland, Earth Syst. Sci. Data, 10(4), 1959–1985, https://doi.org/10.5194/essd-10-1959-2018, 2018.

Mosley-Thompson, E., McConnell, J. R., Bales, R. C., Li, Z., Lin, P. N., Steffen, K., Thompson, L. G., Edwards, R. and Bathke, D.: Local to regional-scale variability of annual net accumulation on the Greenland ice sheet from PARCA cores, J. Geophys. Res. Atmos., 106(D24), 33839–33851, https://doi.org/10.1029/2001JD900067, 2001.

Noël, B., Van De Berg, W. J., Van Wessem, J. M., Van Meijgaard, E., Van As, Di., Lenaerts, J. T. M., Lhermitte, S., Munneke, P. K., Smeets, C. J. P. P., Van Ulft, L. H., Van De Wal, R. S. W. and Van Den Broeke, M. R.: Modelling the climate and surface mass balance of polar ice sheets using RACMO2 - Part 1: Greenland (1958-2016), Cryosphere, 12(3), 811–831, https://doi.org/10.5194/tc-12-811-2018, 2018.

Nowicki, S. M. J., Payne, A., Larour, E., Seroussi, H., Goelzer, H., Lipscomb, W., Gregory, J., Abe-Ouchi, A., and Shepherd, A.: Ice Sheet Model Intercomparison Project (ISMIP6) contribution to CMIP6, Geosci. Model Dev., 9, 4521–4545, https://doi.org/10.5194/gmd-9-4521-2016, 2016.

Pfeffer, W. T., Meier, M. F. and Illangasekare, T. H.: Retention of Greenland runoff by refreezing: implications for projected future sea level change, J. Geophys. Res., 96(C12), 22117, https://doi.org/10.1029/91jc02502, 1991.

Pfeffer, W. T. and Humphrey, N. F.: Determination of timing and location of water movement and ice-layer formation by temperature measurements in sub-freezing snow, Journal of Glaciology. Cambridge University Press, 42(141), pp. 292–304. doi: 10.3189/S0022143000004159, 1996.

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(12), 4238–4246, https://doi.org/10.1002/2014GL059806, 2014.

Reeh, N., Fisher, D. A., Koerner, R. M. and Clausen, H. B.: An empirical firn-densification model comprising ice lenses, Ann. Glaciol., 42, 101–106, https://doi.org/10.3189/172756405781812871, 2005.

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, Cryosphere, 6(4), 743–762, https://doi.org/10.5194/tc-6-743-2012, 2012.

Samimi, S. and Marshall, S. J.: Diurnal cycles of meltwater percolation, refreezing, and drainage in the supraglacial snowpack of Haig Glacier, Canadian Rocky Mountains, Front. Earth Sci., 5(February), 1–15, https://doi.org/10.3389/feart.2017.00006, 2017.

Rushlow, C. R., Sawyer, A. H., Voss, C. I., Godsey,S.E.: The influence of snow cover, air temperature, and groundwater flow on the active-layer thermal regime of Arctic hillslopes drained by water tracks. Hydrogeo. J., 10.1007/s10040-020-02166-2, 2020.

Schneider, T. and Jansson, P.: Internal accumulation in firn and its significance for the mass balance of Storglaciären, Sweden, J. Glaciol., 50(168), 25–34, https://doi.org/10.3189/172756504781830277, 2004.

Schwander, J., Barnola, J.-M., Andrié, C., Leuenberger, M., Ludin, A., Raynaud, D., and Stauffer, B.: The age of the air in the firn and the ice at Summit, Greenland, *J. Geophys. Res.*, 98(D2), 2831– 2838, doi:10.1029/92JD02383, 1993.

Schwander, J., Sowers, T., Barnola, J.M., Blunier, T., Fuchs, A. and Malaizé, B.: Age scale of the air in the Summit ice: implication for glacial–interglacial temperature change. J. Geophys. Res., 102(D16), 19 483–19 493, 1997.

Simonsen, S. B., Stenseng, L., Adalgeirsdóttir, G., Fausto, R. S., Hvidberg, C. S. and Lucas-Picher, P.: Assessing a multilayered dynamic firn-compaction model for Greenland with ASIRAS radar measurements, J. Glaciol., 59(215), 545–558, https://doi.org/10.3189/2013JoG12J158, 2013.

Steffen, C., Box, J. and Abdalati, W.: Greenland Climate Network: GC-Net., 1996.Steger, C. R., Reijmer, C. H. and Van Den Broeke, M. R.: The modelled liquid water balance of the Greenland Ice Sheet, Cryosphere, 11(6), 2507–2526, https://doi.org/10.5194/tc-11-2507-2017, 2017.

Stevens, C. M., Verjans, V., Lundin, J. M. D., Kahle, E. C., Horlings, A. N., Horlings, B. I., and Waddington, E. D.: The Community Firn Model (CFM) v1.0, Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-361, in review, 2020.

Steger, C. R., Reijmer, C. H., van den Broeke, M. R., Wever, N., Forster, R. R., Koenig, L. S., Munneke, P. K., 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(January), https://doi.org/10.3389/feart.2017.00003, 2017.

Sturm, M., Holmgren, J., König, M. and Morris, K.: The thermal conductivity of seasonal snow, J. Glaciol., 43(143), 26–41, https://doi.org/10.1017/S0022143000002781, 1997.

Sommers, A.N., Rajaram, H., Weber, E.P., MacFerrin, M.J., Colgan, W.T. and Stevens, C.M.: Inferring firn permeability from pneumatic testing: a case study on the Greenland ice sheet. Frontiers in Earth Science, 5, p.20, 2017.

Sørensen, L. S., Simonsen, S. B., Nielsen, K., Lucas-Picher, P., Spada, G., Adalgeirsdottir, G., Forsberg, R. and Hvidberg, C. S.: Mass balance of the Greenland ice sheet (2003-2008) from ICESat data - The impact of interpolation, sampling and firn density, Cryosphere, 5(1), 173–186, https://doi.org/10.5194/tc-5-173-2011, 2011.

Team, T. I.: Mass balance of the Greenland Ice Sheet from 1992 to 2018, Nature, https://doi.org/10.1038/s41586-019-1855-2, 2019.

Van Angelen, J. H., Lenaerts, J. T. M., Van Den Broeke, M. R., Fettweis, X. and Van Meijgaard, E.: Rapid loss of firn pore space accelerates 21st century Greenland mass loss, Geophys. Res. Lett., 40(10), 2109–2113, https://doi.org/10.1002/grl.50490, 2013.

Van As, D., Mikkelsen, A. B., Nielsen, M. H., Box, J. E., Liljedahl, L. C., Lindbäck, K., Pitcher, L. and Hasholt, B.: Hypsometric amplification and routing moderation of Greenland ice sheet meltwater release, Cryosphere, 11(3), 1371–1386, https://doi.org/10.5194/tc-11-1371-2017, 2017.

Van As, D., van den Broeke, M., Reijmer, C. and van de Wal, R.: The summer surface energy balance of the high Antarctic plateau, Boundary-Layer Meteorol., 115(2), 289–317, https://doi.org/10.1007/s10546-004-4631-1, 2005.

van As, D., Box, J. E. and Fausto, R. S.: Challenges of quantifying meltwater retention in snow and firn: An expert elicitation, Front. Earth Sci., 4(November), 1–5, https://doi.org/10.3389/feart.2016.00101, 2016.

Van Den Broeke, M. R., Enderlin, E. M., Howat, I. M., Kuipers Munneke, P., Noël, B. P. Y., Jan Van De Berg, W., Van Meijgaard, E. and Wouters, B.: On the recent contribution of the Greenland ice sheet to sea level change, Cryosphere, 10(5), 1933–1946, https://doi.org/10.5194/tc-10-1933-2016, 2016.

Van Genuchten, M. T.: Closed-Form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils., Soil Sci. Soc. Am. J., 44(5), 892–898, https://doi.org/10.2136/sssaj1980.03615995004400050002x, 1980.

Van Kampenhout, L., Lenaerts, J. T. M., Lipscomb, W. H., Sacks, W. J., Lawrence, D. M., Slater, A. G. and van den Broeke, M. R.: Improving the Representation of Polar Snow and Firn in the Community Earth System Model, J. Adv. Model. Earth Syst., 9(7), 2583–2600, https://doi.org/10.1002/2017MS000988, 2017.

Van Pelt, W. J. J., Oerlemans, J., Reijmer, C. H., Pohjola, V. A., Pettersson, R. and Van Angelen, J. H.: Simulating melt, runoff and refreezing on Nordenskiöldbreen, Svalbard, using a coupled snow and energy balance model, Cryosphere, 6(3), 641–659, https://doi.org/10.5194/tc-6-641-2012, 2012.

Van Pelt, W., Pohjola, V., Pettersson, R., Marchenko, S., Kohler, J., Luks, B., Ove Hagen, J., Schuler, T. V., Dunse, T., Noël, B. and Reijmer, C.: A long-term dataset of climatic mass balance, snow conditions, and runoff in Svalbard (1957-2018), Cryosphere, 13(9), 2259–2280, https://doi.org/10.5194/tc-13-2259-2019, 2019.

Vandecrux, B., Fausto, R. S., Langen, P. L., van As, D., MacFerrin, M., Colgan, W. T., Ingeman-Nielsen, T., Steffen, K., Jensen, N. S., Møller, M. T. and Box, J. E.: Drivers of Firn Density on the Greenland Ice Sheet Revealed by Weather Station Observations and Modeling, J. Geophys. Res. Earth Surf., 123(10), 2563–2576, https://doi.org/10.1029/2017JF004597, 2018.

Vandecrux, B., MacFerrin, M., MacHguth, H., Colgan, W. T., Van As, D., Heilig, A., Max Stevens, C., Charalampidis, C., Fausto, R. S., Morris, E. M., Mosley-Thompson, E., Koenig, L., Montgomery, L. N., Miège, C., Simonsen, S. B., Ingeman-Nielsen, T. and Box, J. E.: Firn data compilation reveals widespread decrease of firn air content in western Greenland, Cryosphere, 13(3), 845–859, https://doi.org/10.5194/tc-13-845-2019, 2019.

Vandecrux, B., Fausto, R. S., van As, D., Colgan, W., Langen, P. L., Haubner, K., Ingeman-Nielsen, T., Heilig, A., Stevens, C. M., MacFerrin, M., Niwano, M., Steffen, K. and Box, J. E.: Firn cold content evolution at nine sites on the Greenland ice sheet between 1998 and 2017, Journal of Glaciology, pp. 1–12, https://doi.org/0.1017/jog.2020.30, 2020.

Verjans, V., Leeson, A. A., Max Stevens, C., MacFerrin, M., Noël, B. and Van Den Broeke, M. R.: Development of physically based liquid water schemes for Greenland firn-densification models, Cryosphere, 13(7), 1819–1842, https://doi.org/10.5194/tc-13-1819-2019, 2019.

Vionnet, V., Brun, E., Morin, S., Boone, A., Faroux, S., Le Moigne, P., Martin, E. and Willemet, J. M.: The detailed snowpack scheme Crocus and its implementation in SURFEX v7.2, Geosci. Model Dev., 5(3), 773–791, https://doi.org/10.5194/gmd-5-773-2012, 2012.

Walvoord, M. A., Voss, C., Ebel, B., & Minsley, B.: Development of perennial thaw zones in boreal hillslopes enhances potential mobilization of permafrost carbon. Env. Res. Lett., 14(015003). https://doi.org/10.1088/1748-9326/aaf0cc, 2019.

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, Cryosphere, 8(1), 257–274, https://doi.org/10.5194/tc-8-257-2014, 2014.

Wever, N., Würzer, S., Fierz, C. and Lehning, M.: Simulating ice layer formation under the presence of preferential flow in layered snowpacks, Cryosphere, 10(6), 2731–2744, https://doi.org/10.5194/tc-10-2731-2016, 2016.

Yamaguchi, S., Katsushima, T., Sato, A. and Kumakura, T.: Water retention curve of snow with different grain sizes, Cold Reg. Sci. Technol., 64(2), 87–93, https://doi.org/10.1016/j.coldregions.2010.05.008, 2010.

Zuo, Z. and Oerlemans, J.: Modelling albedo and specific balance of the Greenland ice sheet: Calculations for the Søndre Stromfjord transect, J. Glaciol., 42(141), 305–316, https://doi.org/10.3189/s0022143000004160, 1996.