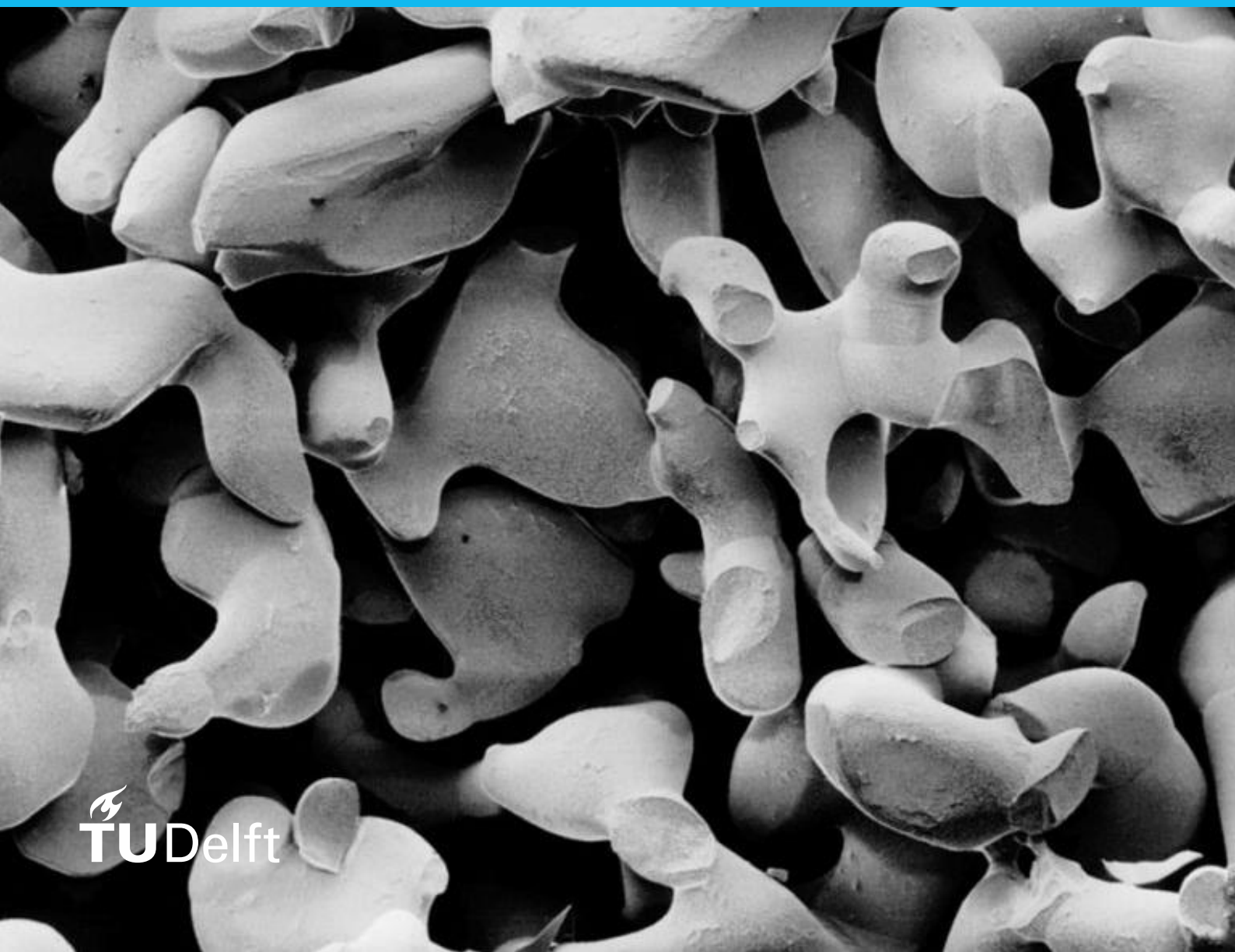
Sticky Snow

Combining Snow and Radiative Transfer Models in the Percolation Area of the Greenland Ice Sheet

D.A.S. Kreynen

MSc Thesis Report – July 2021



TU Delft

Sticky Snow

Combining Snow and Radiative Transfer Models in the Percolation Area of the Greenland Ice Sheet

by

D.A.S. Kreynen

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday July 16, 2021 at 1:00 PM.

Student number:	4176669	
Project duration:	March 2020 – July 2021	
Thesis committee:	Dr. S.L.M. Lhermitte,	TU Delft, chair, daily supervisor
	Ir. W. Li,	TU Delft, daily supervisor
	Dr. ir. D.C. Slobbe,	TU Delft, external assessor

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Abstract

Subsurface firn processes play a crucial role in ice sheet mass loss mechanisms. On Greenland surface meltwater percolates to deeper layers where porous firn retains it, directly inhibiting runoff. However, secondary effects such as the formation of impermeable ice slabs may indirectly and irreversibly accelerate runoff and with it global sea level rise. Microwave remote sensing offers opportunities to monitor these processes, but due to the simplicity of their underlying snow models retrieval methods fail over areas subject to melt and refreezing - areas where the firn's (in)ability to buffer meltwater is critical. This study presents a new forward model which given initial conditions and atmospheric forcing first solves for the firn state through a full-complexity snow model (*SNOWPACK*) and then simulates multifrequency brightness temperature (T_b) time series (using radiative transfer model *SMRT*). As part of a comprehensive sensitivity analysis three ensembles of multi-decade T_b time series (19 and 37 GHz) were modelled for the DYE-2 site in the percolation area of the Greenland Ice Sheet. Model performance based on RMSE w.r.t. independent T_b satellite observations was found to be sensitive to biases introduced in the atmospheric forcing record (with air temperature, precipitation and relative humidity controlling variance) and snow model settings (new snow grain size and albedo settings) and not to initial profile conditions. However, computed RMSEs were high (min. 17.8 K at 37V and 19.4 K at 19V) due to trends in modelled T_b consistently underestimating observed trends when taken over an accumulation season. It is shown that this can only be explained by the constant-with-time stickiness assumption used to link the snow model's microstructure representation to the sticky hard sphere model employed for the radiative transfer scheme. A seasonal stickiness signal is made evident for the conditions at DYE-2 and linked to its yearly melt-refreeze-accumulation cycle. These results demonstrate that earlier approaches to forward model microwave satellite observations based on a constant-with-time stickiness assumption (or that lack a third snow microstructure parameter altogether) are not valid for ice sheet areas prone to melt. This study is expected to be the starting point for a more sophisticated implementation that estimates a snow layer's stickiness from microstructure information already available in the snow model. If successful it would be the first of its kind and open the door to satellite-based retrieval of subsurface firn properties and processes from areas where observations are currently lacking, greatly reducing uncertainty in ice sheet mass loss and global sea level rise projections.

Cover image: micrograph of rounded snow grains, courtesy of the Agricultural Research Service USDA.

Preface

I could not have produced the work contained in this report without a couple of key people that I want to thank. Firstly, Stef and Weiran, for our numerous online meetings full of ideas, feedback, discussion and encouragement. I was always happy to log on! Stef, for inspiring the scientist in me and all the support you have given it. Weiran, for making me feel less alone in my work (I will miss your occasional Slack message at random hours of the day). Thank you to Cornelis, for your fresh pair of eyes when it mattered most. Your feedback and advice improved the final product. Thank you Stephan for taking care of the VRLAB cluster and by extension, thank you to everyone at the GRS department. I have grown and felt at home.

As an alma mater TU Delft presented me with quite the ride: somehow I found myself racing recumbent bicycles through the desert one day, constructing a major car brand's production line the next, presenting our work on space based solar power another and riding snow scooters over endless glaciers after. With this thesis finally, the "remote" was put in "remote sensing" while studying the Greenland ice sheet from space, but also the confinement of my own room. My heart goes out to all students that had to start under these conditions rather than graduate. Here's to more adventurous academic years!

One constant source of cheer have been the friends I got to meet and side with, without whom being a student would have been a boring affair. Be it over one of the numerous bachelor projects, as roommates, in the DREAM Hall, at Sjøskrenten, during the master programme or at any other point along the way: know that it was you who made the peaks so prominent and the valleys traversable. *Ingen nævnt, ingen glemt.* No one mentioned, no one forgotten. I am thankful that our paths crossed and so stoked to have you all in my life.

Last but by no means least, two people I *do* need to mention are my parents. Thank you for your endless patience and support, for providing me with all the chances I could have wished for and for our yearly trips to Berwang (which got me up close and personal with snow from a very young age). Without you none of this would have been possible – this milestone is as much your accomplishment as it is mine. Ik zie jullie graag!

D.A.S. Kreynen
Delft, July 2021

Contents

List of Figures

List of Tables

1 Introduction

1.1. Sea Level Rise and Firn

The climate is warming, glaciers and ice sheets are losing mass and global sea level is rising. The 1993-2010 contribution of glaciers and ice sheets is estimated at 1.4 mm/year, which makes it the dominant source of global mean sea level rise. Measured rates are accelerating and pose a variety of societal challenges. Sea level rise projections are used to plan adaptation and mitigation efforts but come with great uncertainty (IPCC, 2014).

"Tipping elements" of global climate change form one such source of uncertainty. When a tipping element reaches a critical threshold rapid change becomes irreversible. Both the Greenland and Antarctica ice sheets are considered such elements: on Greenland there is feedback between increased melt water runoff and polar amplification (Robinson et al., 2012), while on Antarctica uncontrolled solid ice discharge poses a danger. Although an Antarctic ice shelf collapse alone could lead to 1 m extra sea level rise by 2100 (DeConto and Pollard, 2016), the thresholds and timing of these mechanisms are uncertain forcing the IPCC to exclude these effects from its projections (IPCC, 2019). An important source of uncertainty here is the understanding and model representation of subsurface firn processes in our changing climate (Machguth et al., 2016).

Seasonal snow and perennial firn conditions play a crucial role in mass loss mechanisms. Directly, their properties affect melt rates by controlling how much incoming solar radiation is absorbed. Additionally, due to latent heat release the density and thermodynamic structure of glaciers and ice sheets are heavily affected by refreezing processes (Colgan et al., 2015, Lüthi et al., 2015, Phillips et al., 2010). When melt occurs at the surface liquid water percolates to deeper layers where porous firn retains it - in liquid or refrozen form - inhibiting runoff (fig. 1.1). Initially a buffer against mass loss (Pfeffer et al., 1991), secondary effects after refreezing further affect the net mass balance (van Pelt et al., 2016). When enough water refreezes near the surface low-permeability ice slabs can be formed which prevent additional meltwater from reaching retention capacity in deeper layers, effectively halting the buffering effect (MacFerrin et al., 2019). With runoff being the prime mass loss mechanism for Greenland (Fettweis et al., 2017, Mottram et al., 2019, van den Broeke et al., 2016) but only slightly more than half of the surface melt actually running off (Steger et al., 2017), the significance of this buffering mechanism is not to be belittled. It is estimated that by 2100 runoff over Greenland's expanding ice slabs will have doubled compared to a situation without ice slabs. This equates to an additional 17 to 74 mm of sea level rise given a high-emission scenario (MacFerrin et al., 2019). On Antarctica subsurface buffering capacity controls the expansion of meltwater streams and lakes which threaten ice shelf stability, potentially triggering disintegration, accelerating surrounding glaciers and solid ice discharge leading to another case of rapid sea level rise (Bell et al., 2017, Kuipers Munneke et al., 2014, Lenaerts et al., 2017, Pollard et al., 2015). All this illustrates that it is essential to monitor firn processes to anticipate future sea level rise.

1.2. Assessing the State of Firn

It is difficult to make in situ observations of subsurface firn properties. Traditionally this is done by drilling and measuring the properties of retrieved firn core sections: a labour intensive and expensive

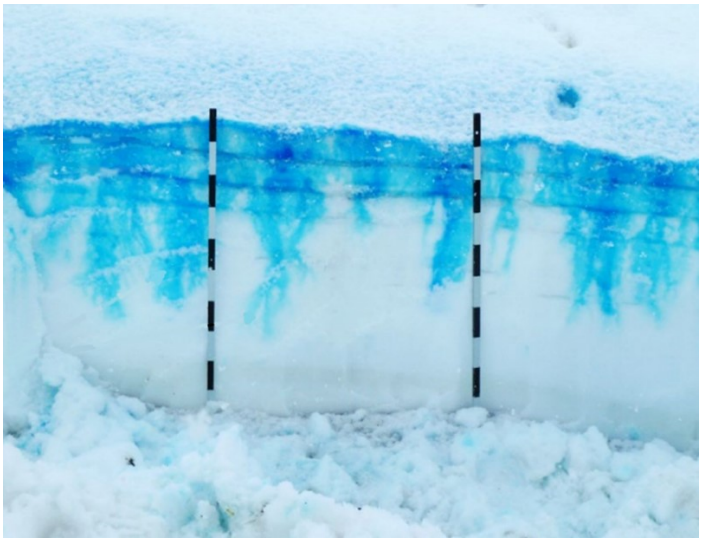


Figure 1.1: *Pecrolation and retention of surface "meltwater" visualized by applying 0.5 mg/L of erioglaucine as blue dye to the snow surface. Picture taken after 45 min of percolation, preferential flow paths and ice layers clearly visible. Source: Lazzaro et al. (2015).*

task. Even with recent advances in pneumatic measurements and ground penetrating radar, due to our ice sheets' inaccessibility and challenging conditions efforts generally focus on specific areas that are revisited on an annual basis at best, which fails to fully capture spatio-temporal variations of firn conditions (van As et al., 2016). This means that our current understanding is largely dependent on a combination of microwave remote sensing (Fettweis et al., 2011) and snow models (Lenaerts et al., 2017, Obleitner and Lehning, 2004, Reijmer et al., 2012, van Pelt et al., 2016).

Remote sensing in the microwave spectrum offers possibilities to study snow properties not only at the surface but also several meters below it. This is true because microwave radiation is able to significantly penetrate the snowpack and interact with it while doing so (fig. 1.2). A snowpack's temperature, density, liquid water content and grain size profile or the presence of ice lenses continuously affects its emissivity and scattering characteristics at different frequencies and polarizations. As a result both active and passive microwave instruments can detect (sub)surface melt and radar (terrestrial, airborne or from a satellite platform) can be used to map snow pack layering (Forster et al., 2014). Another case is the use of multi-frequency passive microwave (brightness temperature, T_b) satellite series observations to constrain and validate snow models that in combination with radiative transfer models are used to assess subsurface firn properties. Using this technique Brucker et al. (2010) have demonstrated that observed emissivities in Antarctic dry-snow zones can only be explained through an increasing grain size with depth. Similarly, ratios between horizontally and vertically polarized T_b have been used as indicator of near-surface density changes (Brucker et al., 2011). However, the use of these techniques is limited by the simplicity of the underlying snow models. E.g. Brucker et al. (2010) assumed a simple analytical relationship for the grain size profile based on a near surface grain radius and vertical gradient. This works for dry-snow zones where grain growth follows an Arrhenius-type relationship but will fail in areas with melt where grain growth is dominated by melt-refreezing cycles. Likewise, Picard et al. (2009) used a physically based snow dynamic and emission model to forward model T_b time series over Antarctica and found the model to be inadequate in the melt zones. The model was relatively simple and, for example, assumed a vertically constant and spatially uniform snow density. This proved sufficient in dry snow zones but failed where melt events complicated the snowpack structure and its temporal variations - areas where the firn's (in)ability to buffer melt water is pivotal.

To retrieve firn properties from these areas more complex, physical snow models are required. An example of a "full-complexity" snow model is *SNOWPACK* (Bartelt and Lehning, 2002, Lehning et al., 2002a,b). Given atmospheric forcing it solves the coupled energy and mass balance of the snowpack to simulate how its properties evolve over time (fig. 1.3). Originally developed for seasonal snow and in operational use as avalanche forecasting tool it is skilful in modelling stratigraphy and the formation of thin layers such as surface hoar and melt-freeze crusts, which are known to affect microwave scat-

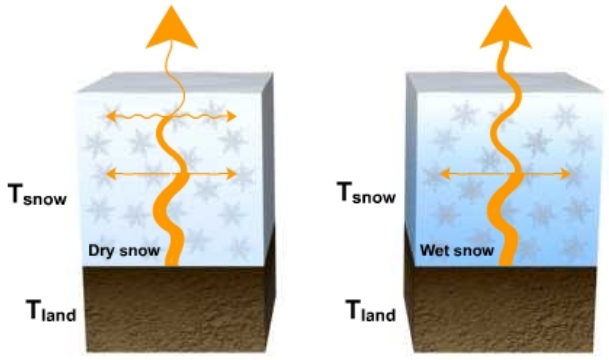


Figure 1.2: *Example illustration of how microwave radiation propagates through a snowpack, which scatters away upwelling radiation. When a snowpack starts melting it also increases its emissivity, resulting in an increase of observed radiation at the sensor. Source: The COMET Program.*

tering and emission properties (Montpetit et al., 2013). *SNOWPACK* has also been applied to polar and perennial firn conditions where thanks to its advanced hydrological, metamorphosis and feedback processes it can adequately deal with melt and refreezing (Steger et al., 2017, Wever et al., 2014). Another way in which this model sets itself apart is through its advanced snow microstructure representation. Four parameters are used to describe the complex texture of snow: grain size, bond size, dendricity and sphericity. Throughout a simulation these parameters evolve per snow layer according to rate equations, which are functions of the local environmental conditions that were based on both theory and empirical relations (Lehning et al., 2002a). The resulting microstructure is used to determine global properties of the snowpack, be it optical (e.g. albedo), mechanical (e.g. strength) or physical (e.g. thermal conductivity) properties. This is promising within the context of radiative transfer modelling because the snow microstructure also determines the scattering and emission properties of a snowpack in the microwave spectrum.

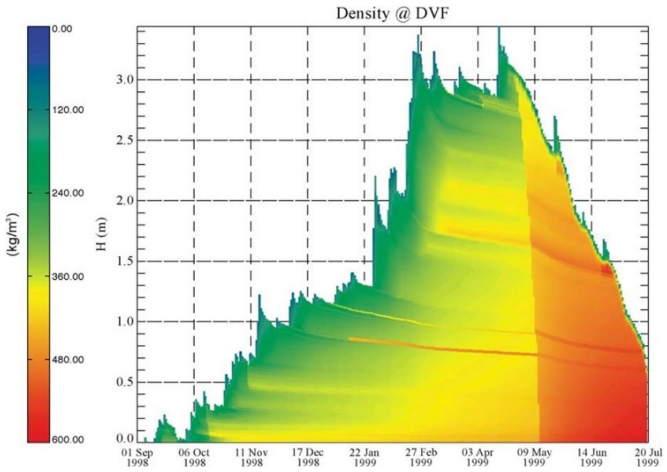


Figure 1.3: *Density profile evolution of a seasonal snowpack as modelled by SNOWPACK. Layering (horizontal streaks of higher density) and onset of spring melt (step increase in density at the beginning of May) clearly visible. For this seasonal example only atmospheric forcing is required as input for the model - no initial snowpack. Source: Bartelt and Lehning (2002).*

1.3. Snow Model Uncertainty

However, given this complexity the use of *SNOWPACK* comes with its own uncertainties: in input data (errors in initial snow pack or atmospheric forcing record), model settings and solution (errors in input parameters, different possible model configurations and the discretizations/algorithms used to compute output), and as a result in output (the precision and accuracy with which *SNOWPACK* can reproduce snowpack properties). Because it can be applied to a wide variety of conditions the number of model settings and parameterizations available in *SNOWPACK* is naturally expected to be

high. As a process-based model its different configurations depend on input parameters that have physical meaning but are in practice difficult to infer from measurements, for example the grain size of fresh snow or local aerodynamic roughness length. In other applications observations of these parameters or of initial snowpack conditions might be missing altogether. Also the atmospheric forcing record contains errors, which can be significant when variables have to be interpolated from sparsely distributed weather stations or come from climate model output. The usefulness of any model depends on the accuracy and reliability of its results. In order to quantify uncertainty in output and to focus measurement and modelling efforts on areas where they have the greatest impact it is important to understand how these different sources of uncertainty interact and propagate through the model.

Snow models have been the subject of studies assessing the relationship between model structure and model performance (Magnusson et al., 2015) and how forcing error propagates through different models (Raleigh et al., 2015). The work discussed here applies to the setting of hydrological modelling of seasonal snowpacks. Magnusson et al. (2015) compared the performance and behaviour of three different model types with varying complexity (including *SNOWPACK*) by generating an ensemble of model outputs and comparing it to snow mass and runoff observations at two sites in the European Alps. Although no clear relation between model complexity and performance could be identified, it was demonstrated that an appropriate combination of process representations is required to achieve high model performance. Raleigh et al. (2015) employed a global sensitivity analysis to investigate how different error types (biases/random), probability distributions and magnitudes introduced in the forcing record affect performance of a physical snow model. The authors found a high sensitivity to biases in the forcing (less so for random errors) and that model output uncertainty due to forcings can be comparable to or larger than model uncertainty due to model structure.

Günther et al. (2019) recognised that earlier efforts focused on either the impact of different model configurations or forcing errors, but not their possible interactions or the robustness of results given model parameter uncertainty. They undertook a global variance-based sensitivity analysis of their own, investigating the impact of these uncertainty sources simultaneously. Performance was based on a medium-complexity snow model's ability to simulate snow water equivalent observations at a monitoring station in the Austrian Alps. Within this context it was found that snow models are most sensitive to input data errors, then to model structure and last to model parameters. Significant interaction effects were discovered and the authors conclude that for a representative assessment these must be included in future analyses. The Sobol' sensitivity analysis framework proved an excellent tool to come to these conclusions: it can quantify the sensitivity of model performance to changes in input data and settings, include interaction effects and allow for grouping of input factors. The latter is crucial because it means that both scalar and non-scalar factors can be considered simultaneously, and grouping reduces the dimensionality of the problem and thus computational costs. Inspired by Günther et al. (2019) in this study a similar analysis is performed for perennial firn conditions, using a combined snow and radiative transfer model and satellite T_b observations to evaluate model performance.

1.4. Radiative Transfer Modelling of Snow

Radiative transfer models based on the understanding of how snow and electromagnetic waves interact can, given snowpack properties, be used to forward model a satellite signal. In case of passive microwave observations radiometers carried on board satellites measure thermal radiation emitted by the snowpack, which is expressed as T_b. T_b is the product of snow temperature and emissivity, with emissivity a function of snowpack density, liquid water content, grain size and microstructure. In order to be able to apply a radiative transfer scheme the snowpack is assumed to be made up out of a number of discrete layers whose properties come either from in situ snow measurements (e.g. Brucker et al., 2011) or snow models (e.g. Wiesmann et al., 2000). Additionally, different models have different approaches to the snow microstructure representation which determines a layer's electromagnetic properties required to solve the radiative transfer equation. Two theories are of particular interest to this study. Firstly, the Improved Born approximation (IBA, Mätzler, 1998) which expresses the scattering coefficient in terms of the Fourier transform of the two point correlation function. In the Microwave Emission Model of Layered Snowpacks (MEMLS, Wiesmann and Mätzler, 1999) this is simplified further by assuming an exponential function form for the correlation function and as a result only a single parameter is needed as microstructure model input (exponential correlation length - no grain size). The second theory of interest is the Dense Media Radiative Transfer Theory (DMRT, Picard et al., 2013)

where microstructure is based on approximating snow grains by spheres. However, it was found that the sphere diameter ("snow grain size") and volume fraction ("snow density") do not fully characterise the snow medium and that an extra parameter is required to do so: the stickiness parameter, resulting in the sticky hard sphere (SHS) microstructure model. The stickiness parameter is used to change the relative position of snow grains (the spheres) without changing their diameter or volume fraction and can be regarded as controlling the amount of snow grain clustering, with clusters showing different scattering properties than single well separated grains. Recent advances have proven the electromagnetic approaches of IBA/MEMLS and DMRT to be equivalent (Löwe and Picard, 2015) and by reformulating their microstructural models in a common framework, the Microwave Radiative Transfer model (*SMRT*, Picard et al., 2018) now allows the IBA electromagnetic theory to be used with five different microstructure models, including SHS. This setup was chosen for this study as it allows for using snow modelled grain sizes directly as input for the radiative transfer modelling chain.

1.5. Proposed Firn Retrieval Workflow

A novel workflow to retrieve information on the processes occurring in the firn layer is proposed here. An integrated modelling-remote sensing approach it combines a snow model with a radiative transfer model to simulate ensembles of satellite observations. Fig. 1.4, from left to right: given an area and time frame of interest, multiple snow model output realisations (time series of firn properties) are generated by forcing *SNOWPACK* with a range of atmospheric input variables that account for the uncertainties in climate model output (accounting for uncertainties in model forcing). At the same time, different model configurations using the settings/parameterizations available in *SNOWPACK* are used (accounting for uncertainties in model settings and parameters). The resulting ensemble of possible but uncertain time series of snow and firn properties serves as input for a radiative transfer model (here *SMRT*) which outputs the ensemble of modelled satellite observation time series. In a final step the last ensemble is compared to actual satellite observations and ensemble members that are not in agreement are excluded, leaving a reduced set of candidates. While this will likely not lead to a single best ensemble member (i.e. a unique solution), if the modelling chain proves successful it will certainly allow for quantification of the uncertainty in firn time series by analysing the spread in properties of the reduced set of candidates. Additionally, by looking at which model inputs and settings lead to realistic modelled satellite observations atmospheric forcing and *SNOWPACK* model uncertainties can be reduced. When for example a certain model configuration is found to consistently underperform it seems reasonable to assume it inappropriate for the conditions and exclude it from future efforts - effectively reducing model uncertainty.

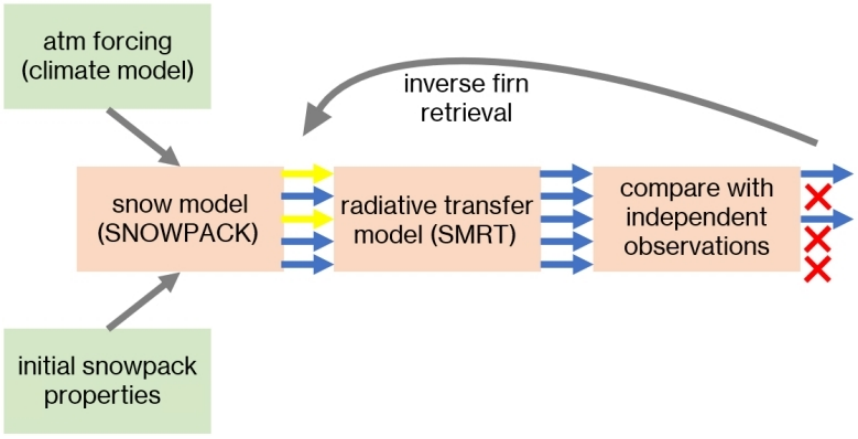


Figure 1.4: *Proposed workflow that uses an integrated modelling-remote sensing approach to assess subsurface processes occurring in the firn layer. An ensemble of firn time series is generated accounting for model uncertainties, translated into satellite observed series and compared with independent observations to make a subset of possible candidates.*

1.6. Research Objectives

In a first step towards inverse firn retrieval, the work presented in this report concerns the forward modelling aspect of the proposed workflow of fig. 1.4. The main objective reads:

> **Main Objective**: demonstrate the feasibility of using snow model *SNOWPACK* in combination with radiative transfer model *SMRT* to forward model ensembles of multi-frequency T_b time series, for comparison with independent satellite observations.

With a particular interest in ice sheet percolation areas, where the buffering mechanism of firn is both fragile and crucial but where earlier retrieval methods fall short. It aims to answer the following research questions:

> **RQ1** - What is the impact of the different *SNOWPACK* input variables and model settings on modelled T_b time series?
>
> **RQ2** - How do *SNOWPACK-SMRT* modelled T_b time series compare to independent satellite observations?
>
> **RQ3** - What is the added value of using T_b time series when constraining *SNOWPACK* ensembles?

Answering these questions will prove essential in the development of a combined snow and radiative transfer model to evaluate subsurface firn processes using microwave remote sensing observations, currently being worked on at TU Delft's Geoscience and Remote Sensing department.

1.7. Research Approach

This study presents a combined forward model which given initial conditions and atmospheric forcing first solves for the snow and firn state (through snow model *SNOWPACK*) and then simulates multi-frequency T_b observations (using radiative transfer model *SMRT*). As part of a comprehensive sensitivity analysis three ensembles of multi-decade T_b time series at 19 and 37 GHz and V and H polarizations are modelled for a study site in the percolation area of the Greenland Ice Sheet. Model performance is assessed by comparing *SNOWPACK-SMRT* modelled T_b to satellite observations and by comparing the *SNOWPACK* modelled density profiles directly to in situ profiles from firn cores. The impact of errors in atmospheric forcing variables, initial snowpack conditions and *SNOWPACK* model settings as well as their interaction effects are evaluated within the Sobol' sensitivity analysis framework, for model performance based on both T_b and density. Differences in sensitivity and discrepancies between modelled and observed T_b series are discussed, the latter linked to snow microstructure assumptions made while combining *SNOWPACK* with *SMRT*. Lastly the forward modelling approach presented here is compared with earlier efforts: while measured profiles from dry snow zones and physically modelled seasonal snowpacks have been used to model microwave emission time series before, several aspects are new. This includes the ensemble approach, the application of a model with full-complexity snow component to the percolation area of an ice sheet and modelling of multi-decade time series.

2

Models, Methods, and Data

2.1. Study Site and Data

The analysis was performed at DYE-2 (also known as Camp Raven), a site located on the Greenland Ice Sheet at 66.48°N, 46.28°W (fig.2.1). The local surface is elevated 2165 m above sea level and has a slope of 0.2°. Mean air temperature over the simulation period (June 1998 to December 2017) was measured at -17.4°C and mean annual precipitation at 424 mm water equivalent. This site was selected for two main reasons. Firstly, located in the so-called percolation area of the ice sheet, DYE-2 experiences yearly melt seasons during and after which surface melt water percolates through the firn layer. Here it can refreeze forming ice lenses and releasing latent heat - the processes at the basis of ice slab formation. From field campaigns we know that thick impermeable ice slabs are as of now not present at DYE-2 but they have been observed at nearby stations at lower elevations (MacFerrin et al., 2019). Secondly, DYE-2 is the site of a Cold War era military installation and is a well established landmark on the ice sheet. It has been and continues to be frequently revisited and as a result there is ample in-situ data available to serve as model input and for calibration/validation purposes.

2.1.1. Atmospheric Forcing

Atmospheric forcing was taken from Vandecrux et al. (2020) and is based on the GC-Net Automatic Weather Station (AWS) record at DYE-2. Vandecrux et al. (2020) carried out the following steps:

1. Outliers were rejected according to Vandecrux et al. (2018).
2. Gaps were filled by interpolation, through adjusting data from nearby KAN_U AWS (Charalampidis et al., 2015) or by using RACMO2.3p2 climate model output (Noël et al., 2018).
3. Gaps in upward shortwave radiation were filled using a MODIS daily albedo product by Box et al. (2017).
4. Shortwave radiation was corrected for AWS tilt according to Wang et al. (2016).
5. Downward longwave radiation (not measured) was taken entirely from RACMO2.3p2 model output.
6. Measurements of surface height were converted to hourly snow accumulation (from here onward “precipitation”), using a site-specific calibration factor such that the derived winter accumulation matches in-situ snow pit observations. For more information on this routine, see Vandecrux et al. (2018).

The result is a gap-free hourly meteorological data set that runs from 1 June 1998 to 31 December 2017. An overview of the variables that were used to force our model are presented in table 2.1.

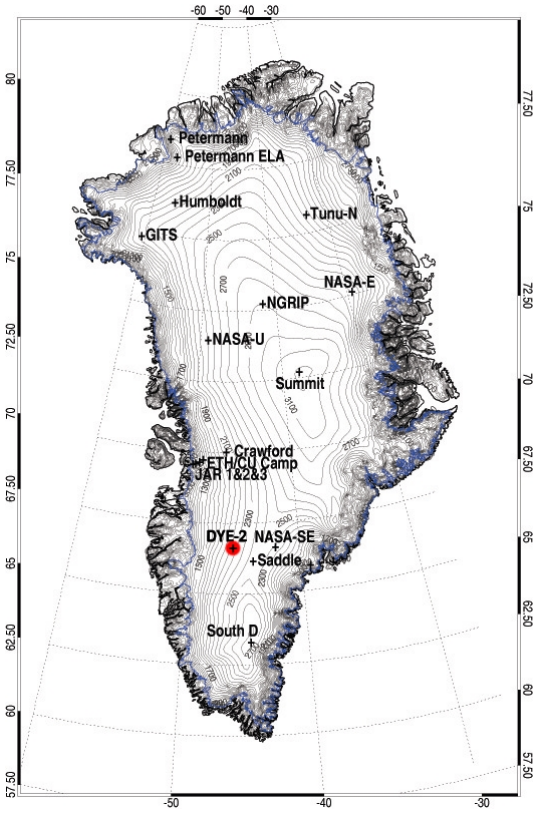


Figure 2.1: *Map of Greenland showing the GC-Net Automatic Weather Station locations, with the DYE-2 study site within the ice sheet's percolation area highlighted in red. Source: modified from the Cooperative Institute for Research in Environmental Science at UCB (Steffen et al., 1996).*

Table 2.1: *Atmospheric forcing Variables that were taken from the meteorological record at DYE-2, as presented in Vandecrux et al. (2020).*

Variable	Symbol	Units	Comment
time	t	[days]	since 1900-1-1 0:0:0 UTC
air temperature	T_a	[K]	at 2 m
relative humidity	RH	[%]	at 2 m
wind speed	U	[m/s]	at 10 m
incoming shortwave radiation	Q_{si}	[W/m^2]	-
outgoing shortwave radiation	Q_{so}	[W/m^2]	-
incoming longwave radiation	Q_{li}	[W/m^2]	-
precipitation	P	[m/hr]	water equivalent

2.1.2. Snow Density Profiles

Snow density at the beginning of the simulation period comes from a Program for Arctic Regional Climate Assessment (PARCA) core drilled at DYE-2 in 1998, with a 10 cm resolution and total depth of 60 m (Mosley-Thompson et al., 2001). Repeat observations for comparison with snow modelled densities come from the SUMup dataset (Montgomery et al., 2018) which was filtered on observations from the DYE-2 area corresponding to our simulation period. Five firn core records were found, all from 2013 onwards (table 2.2).

2.1.3. Satellite Observed T_b

Independent satellite observations of T_b at 37 and 19.35 GHz come from the Special Sensor Microwave Imager (SSM/I) and the Special Sensor Microwave Imager/Sounder (SSMIS) sensors carried on board a series of Defense Meteorological Satellite Program (DMSP) satellites. Here data from the F13 (operational from 1995 until end of 2008) and F17 (operational from end of 2006 onwards) platforms was

Table 2.2: *Firn core density observations used for model evaluation, retrieved from the SUMup dataset (Montgomery et al., 2018). Note that cores 2013A and 2013B were drilled on the same day and henceforth only the 2013 core is referred to, whose density profile was computed as the mean of cores 2013A and 2013B.*

Core	Date	Depth range [m]	Reference
2013A	05 May 2013	0.00 - 16.64	Machguth et al. (2016)
2013B	05 May 2013	0.00 - 16.45	Machguth et al. (2016)
2015	21 May 2015	0.81 - 19.27	MacFerrin et al. (tion)
2016	06 May 2016	0.00 - 17.37	MacFerrin et al. (tion)
2017	11 May 2017	0.00 - 22.97	MacFerrin et al. (tion)

used, allowing for daily observations over our entire simulation period. Daily images were downloaded for both vertically and horizontally polarized channels (Maslanik and Stroeve, 2018), which had spatial resolutions of 25 km. The values of the single pixel containing the study site were extracted and for days were both F13 and F17 data was available the mean value was taken to arrive at time series of daily T_b observations for four channels: 37V, 37H, 19V and 19H.

2.2. Models

2.2.1. Snow Model

SNOWPACK was chosen as snow model for this study. It is a one-dimensional, physical model driven by atmospheric forcing (Bartelt and Lehning, 2002) and implemented in C++. It is skilful in modelling the formation of thin layers (Lehning et al., 2002a), has a proven track record of dealing with melt and refreezing in polar environments (Steger et al., 2017, Wever et al., 2014) and sets itself apart from other models by its advanced snow microstructure and metamorphism representation (Lehning et al., 2002b), which makes it an excellent candidate for the combined snow and radiative transfer model.

For initial snow conditions the 1998 observed density profile was supplemented by assumed grain size and temperature profiles from (Vandecrux et al., 2020), resulting in the same set up as the “DYE-2_long” experiment presented by those authors and ensuring comparability of results. However, given the advanced microstructure representation *SNOWPACK* required some additional parameters. In case of an undisturbed model run, the bond radius (rb, [mm]) was assumed as 0.5 times the grain size, dendricity (dd, [-], between 0 and 1) was set to 0.2 and sphericity (sp, [-], between 0 and 1) to 0.8. These variables were assumed constant with depth and any other parameters or markers were set to 0. To speed up computational times the resolution of the initial profile was adjusted with depth. For the top 10 m the original resolution of 10 cm was preserved, between 10 and 20 m this was reduced to 20 cm, between 20 and 30 m to 50 cm and below 30 m to 100 cm. This reduced the total number of initial snow layers from 600 to 200 while maintaining fine snow layer resolution at the top of the profile where strong gradients in properties are expected due to interaction with the atmosphere. Below the final depth of 60 m 10 insulating ice layers were added to prevent heat escaping from the bottom of the domain. These had a thickness of 10 m each and were set to the study site’s mean air temperature.

Table 2.3 contains an overview of the *SNOWPACK* model settings considered in this study. Over the years different groups have experimented with their own implementations of *SNOWPACK* and as a result many options are available in the model. Most of the processes can be represented by a number of different parameterizations, other options can be turned on or off, increasing or decreasing the model’s complexity or changing its working altogether. An extended list of used *SNOWPACK* settings as well as an explanation on the non-standard RACMO albedo parameterization can be found in appendix A. For a comprehensive overview of the model’s workings the reader is kindly referred to Bartelt and Lehning (2002), Lehning et al. (2002a,b).

Table 2.3: *SNOWPACK model settings and parameters considered in this study. Appendix A contains a list of the other model settings (that were kept fixed) as well as an explanation on the RACMO albedo parameterization.*

Available setting	Options considerd	Comment
SW_MODE	BOTH, INCOMING	measured or SNOWPACK albedo
SNOW_ALBEDO	FIXED, PARAMETERIZED	SNOWPACK albedo: fixed value or parameterized
ALBEDO_FIXEDVALUE	range of values	[-], only when SNOW_ALBEDO is set to FIXED
ALBEDO_PARAMETERIZATION	LEHNING_0-2, SCHMUCKI_1-2, “RACMO”	6 options total, only when SNOW_ALBEDO is set to PARAM.
ALBEDO_AGING	TRUE, FALSE	snow age taken into account for albedo, only some param.
SW_ABSORPTION_SCHEME	MULTI_BAND, SINGLE_BAND	single or multi band shortwave radiation absorption scheme
HN_DENSITY	FIXED, PARAMETERIZED	sixed or parameterized new snow density
HN_FIXEDVALUE	range of values	[kg/m3], only when HN_DENSITY is set to FIXED
HN_PARAMETERIZATION	LEHNING_NEW-OLD, ZWARTE, BELLAIRE, PAHAUT	5 options in total, only when HN_DENSITY is set to PARAM.
ATMOSPHERIC_STABILITY	RICHARDS, NEUTRAL, MO_MICHLMAYR	atmospheric stability correction scheme
ROUGHNESS_LENGTH	range of values	[m], aerodynamic roughness length
NEW_SNOW_GRAIN_SIZE	range of values	[mm], grain size of new snow (diameter)
ENABLE_VAPOUR_TRANSPORT	TRUE, FALSE	whether mass transport by vapour flow is enabled

2.2.2. Radiative Transfer Model

SMRT (Picard et al., 2018) was used to translate *SNOWPACK* output into multi-frequency T_b observations. It uses the discrete ordinate and eigenvalue method to solve the radiative transfer equation, or how radiative energy propagates through the multilayered medium. In our passive case this energy comes from thermal emission of the snow itself (opposed to also from the sky in case of active mode/radar). Implemented in Python, *SMRT* differentiates itself from other models by offering a high degree of flexibility in switching between different electromagnetic theories and snow microstructure representations which together determine the exact electromagnetic behaviour of snow. In this study IBA electromagnetic theory was used in combination with SHS microstructure representation.

This IBA-SHS configuration was chosen for a couple of reasons. One, of the two electromagnetic theories of interest - being DMRT (Picard et al., 2013, Tsang et al., 2007) and IBA (Mätzler, 1998) - IBA proved to produce results most reliably. This is related to the way in which DMRT is currently implemented in *SMRT* (using a short-range approximation which requires small grain size compared to wavelength) and the large grain sizes encountered at DYE-2. That being said, exceptionally large grain sizes still had to be capped when using IBA but not to the same extent as with DMRT. Two, in *SMRT* IBA is not limited to a particular microstructure model but DMRT is only compatible with SHS. The IBA-SHS configuration therefore provided a good starting point as potential follow up work can change the microstructure representation independently from the electromagnetic theory and vice versa.

The SHS representation had an extra advantage in that it uses two parameters to characterize snow microstructure: grain size and a stickiness parameter, while the exponential representation that IBA is traditionally associated with (as implemented in MEMLS by Wiesmann and Mätzler, 1999) only uses a correlation length. It might seem advantageous to have one less parameter, but this has been found inadequate to fully characterize the microstructure of snow (Mätzler, 2002, Picard et al., 2018). Additionally, when using SHS *SNOWPACK* grain size could be used directly as input for *SMRT*. The stickiness parameter is then used to change the relative position of snow grains without changing their diameter or volume fraction, effectively controlling the clustering of snow grains. This has an effect on radiative transfer because sticky or clustered snow shows different scattering behaviour compared to uniformly distributed snow grains. Stickiness is not explicitly present in *SNOWPACK* and was assumed vertically and temporally constant.

2.2.3. Combined Model

Snow model *SNOWPACK* and radiative transfer model *SMRT* were linked together to form the “combined model” as implemented on the VRLAB cluster of the GRS department (fig. 2.2). From left to right: initial snow and firn conditions and atmospheric forcing serve as input for *SNOWPACK*, that depending on the chosen settings models how the profile of snow and firn properties evolves over time. This intermediate output (daily snow and firn properties) is saved on the server. In the “layer merging” step this is prepared as input for *SMRT*, that translates it to time series of modelled T_b - the final output of the combined model. This comprises the forward modelling chain of the proposed firn retrieval workflow of fig. 1.4.

Practically, the layer merging step is made up of custom Python functions that start off with reading the *SNOWPACK* output data, to turn it into daily profiles of snow density, temperature, liquid water content and grain radius suitable for use as input with *SMRT*. This involves the following steps:

1. The 10 insulating ice layers were removed again from the bottom the record.
2. All variables were converted to have units appropriate for *SMRT*: layer thickness [m], snow temperature [K], density [kg/m^3], volumetric ratio of liquid water [-] and grain radius [m].
3. The total depth of the profiles was limited to the first 4 m for simulations at 37.00 GHz and to 15 m for 19.35 GHz and a semi-infinite final layer was added to the bottom of the record. This was done to limit computational times. The 4 and 15 m values were chosen such that increasing the profile depths any further did not change the modelled T_b values anymore.
4. The number of snow layers was limited to 80 for simulations at 37.00 GHz and to 200 at 19.35 GHz, again to limit computational times. When more layers were present the thinnest layer was merged with the layer below it and the physical properties of the merged layer were determined as the thickness-weighted average of the properties of the individual layers. This was continued until a maximum of 80 or 200 layers were left. The maximum number of layers was determined such that increasing these numbers did not significantly change modelled T_b values anymore.
5. Snow densities were capped at 916 kg/m^3 to avoid erroneous volume fractions greater than unity in *SMRT*, which would lead to an error. *SNOWPACK* densities max out at the generally accepted pure ice density of 917 kg/m^3 and the effect of the capping was assumed negligible.
6. Grain radii were capped at 1.5 mm (equivalent to a grain size of 3.0 mm) to avoid exceptions in *SMRT*. Contrary to the density capping when large grain sizes are present in the *SNOWPACK* output capping these might significantly affect modelled T_b, but this was preferred over gaps in the output time series. Details on the amount of capping can be found in appendix B.
7. A constant (with depth and time) stickiness parameter was set (section 2.5). This is a snow variable not explicitly present in the *SNOWPACK* output but required to run *SMRT* with the SHS microstructure model.

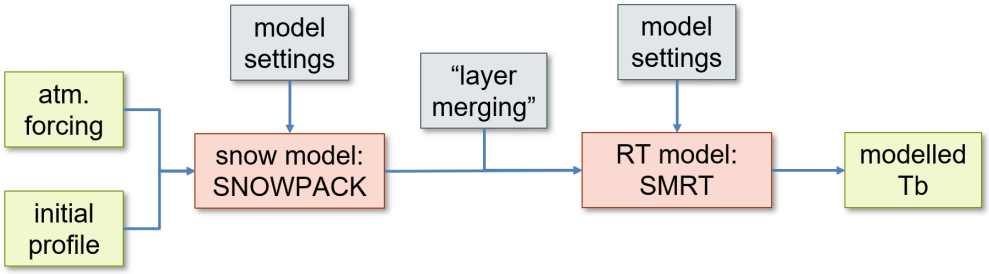


Figure 2.2: *Schematic of the combined snow (SNOWPACK) and radiative transfer (SMRT) model at the basis of this study. Left: model input (initial snow and firn properties and a record of atmospheric conditions), right: model output (modelled T_b time series).*

2.3. Sensitivity Analysis Framework

2.3.1. Sobol' Global Variance-based Method

The Sobol' sensitivity analysis framework is a global variance-based method, where the sensitivity of the model's performance to a certain change in input or model setting is quantified using sensitivity indices (Sobol, 1993). These indices range from zero to one and represent the fraction of output variance explained, i.e. the first-order sensitivity index S_i is defined as the variance in model prediction (Y) from a change in one particular parameter (X_i). Or, in the form of an equation:

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)} \tag{2.1}$$

where V is the variance operator, E the exception operator and $X_{\sim i}$ stands for all parameters except X_i. Contrary to the first-order sensitivity index, the total-order sensitivity index S_{Ti} also includes the

interaction effects of all other parameters:

$$S_{Ti} = 1 - \frac{V[E(Y|X_{\sim i})]}{V(Y)} \tag{2.2}$$

note that as interaction effects are attributed to all parameters involved the sum of all total-order sensitivity indices can exceed one. Previous studies dealing with complex snow models have revealed considerable interaction effects between different uncertainty sources (Günther et al., 2019), hence these effects need to be taken into account in order to arrive at a representative sensitivity analysis. This was an important consideration when deciding on the method presented here.

In practice three sampling matrices are constructed to be able to to evaluate equations 2.1 and 2.2 for a single parameter: A, B and $A_B^{(i)}$, where A and B are of size $N * k$ (number of samples * number of parameters). Each row corresponds to one model realisation, with its elements representing the state of the considered model parameters (the different columns). Typically the latter are picked quasi-randomly in order to optimally sample the input space. $A_B^{(i)}$ is then made up out of a combination of A and B by taking A but replacing the i^{th} column by the i^{th} column of B (Saltelli et al., 2010). The first and total-order sensitivity indices are then determined as (Günther et al., 2019, Jansen, 1999, Saltelli et al., 2010):

$$S_i = 1 - \frac{\frac{1}{2N}\sum_{j=1}^{N}(f(B)_j - f(A_B^{(i)})_j)^2}{V(Y)} \tag{2.3}$$

$$S_{Ti} = \frac{\frac{1}{2N}\sum_{j=1}^{N}(f(A)_j - f(A_B^{(i)})_j)^2}{V(Y)} \tag{2.4}$$

Where f denotes the model performance operator, e.g. $f(A)_j$ is the model skill when setting the model parameters according to the j^{th} row of sampling matrix A. Important to stress is that within the context of the presented sensitivity analysis one dimensional model performance is the final output of the system. This means that this study estimated the sensitivity to different input parameters and model settings not on modelled T_b directly, but on the ability of the combined snow-radiative transfer model to simulate realistic T_b time series according to defined performance measures.

2.3.2. Uncertainty Sources and Grouping

Parameters were grouped into uncertainty classes to limit computational costs and ease interpretation. Given the total number of input parameters and model settings and that simultaneous assessment of S_i and S_{Ti} requires $N(k+2)$ model runs it was computationally infeasible to estimate these indices on an individual level. Additionally, when computing total-order sensitivity indices these would include interaction effects of possibly all other factors making the results exceedingly difficult to interpret. The solution was to group parameters into wider uncertainty classes following the approach of the General Probabilistic Framework (Baroni and Tarantola, 2012).

Three main classes presented themselves naturally: atmospheric forcing, initial snow conditions and *SNOWPACK* model settings. These three distinct groups of uncertainty served as the basis for three different experiments in which the precise definition of these classes was varied (fig. 2.3). For each identified class n independent realisations were generated that reflect the uncertainty inherent to that class. Rather than the state of a single parameter the elements of sampling matrices A and B contained integer numbers picked quasi-randomly from the range 0 to n, each corresponding to one of the generated uncertainty class realisations.

Note that the model settings available in *SMRT* were kept fixed throughout the sensitivity analysis and considered beyond the scope of this study. There are two main reasons for this: one, the extra layer of model uncertainty originating from radiative transfer modelling would make results difficult to interpret, as it would be challenging to discern between effects from the snow model vs. the radiative transfer model. Two, model configurations available in *SMRT* that would give reliable results for the conditions at DYE-2 without heavily altering the *SNOWPACK* output were found to be limited. The decision was made to stick to a single promising configuration, hence the focus of the sensitivity analysis is on the

SNOWPACK component of the combined model. This is in line with the research objectives (notably RQ1).

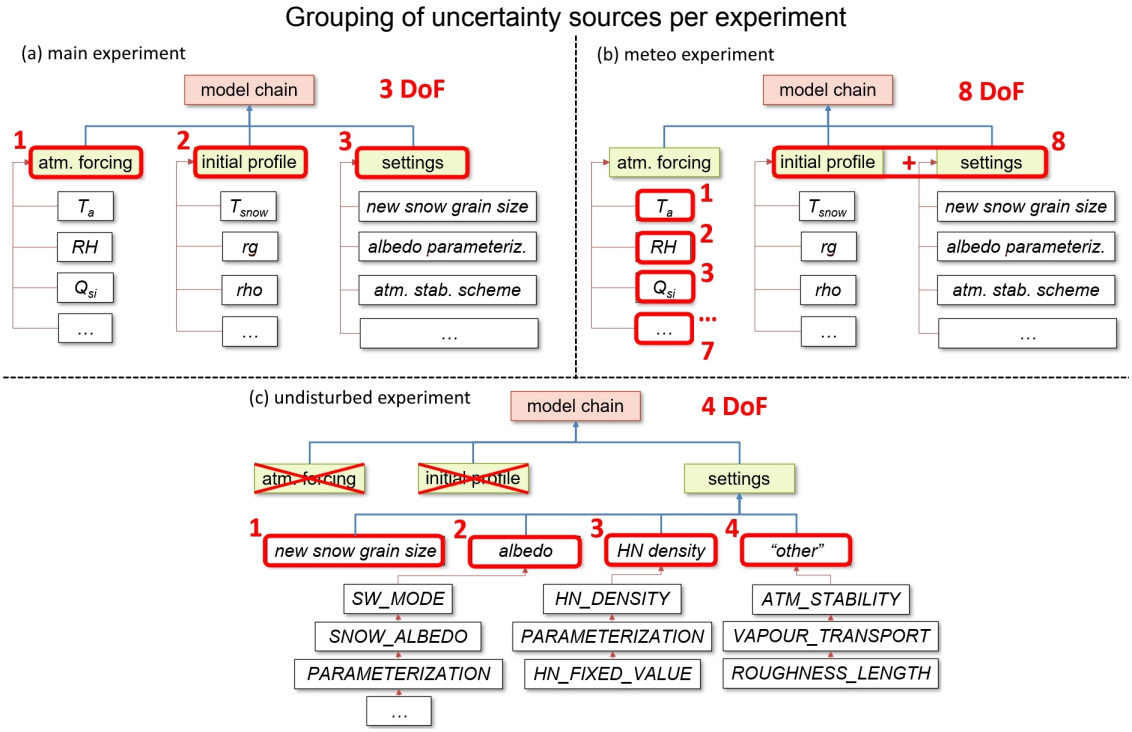


Figure 2.3: *Schematic showing the different degrees of freedom for the main (**a**), meteo (**b**) and undisturbed (**c**) experiments. The meteo experiment considered the atmospheric forcing variables individually rather than as a group and combined the initial profile and SNOWPACK model settings as a single degree of freedom. For the undisturbed experiment uncertainty in atmospheric forcing and initial profile were excluded from the analysis and only model settings were varied, grouped into four degrees of freedom.*

2.4. Description of Experiments

2.4.1. Main Experiment

The first experiment, dubbed the main experiment, has three degrees of freedom: atmospheric forcing, initial profile and *SNOWPACK* settings, with each of these three uncertainty classes consisting of a collection of individual parameters (fig. 2.3a). The goal of the main experiment was to investigate the sensitivity of model performance to biases in the atmospheric forcing and initial profile parameters, changing of the *SNOWPACK* settings and their interaction effects.

Disturbance Distributions

Uncertainty distributions were defined for every individual parameter and were then used to generate a collection of n inependent realisations per uncertainty class. Table 2.4 shows an overview of the assumed distrubutions for the atmospheric forcing variables. The notion was that they represent a realistic range of forcing errors that could be present in the input data, allowing for assessment of the impact that an uncertain variable has on model performance. It has to be acknowledged that deciding on these distributions involves a high degree of arbitrarity and decisions made here will have affected the results of the sensitivity analysis. Wherever possible earlier studies were followed to ensure comparability of results, where needed new assumptions were introduced.

Uncertainty distributions for the second degree of freedom (the initial snow profile properties) required new assumptions (table 2.5), with earlier studies dealing with seasonal snowpacks (no initial profile) and simpler snow models (less parameters). Apart from the profile variables themselves also the temperature of the insulating ice layers at the bottom of the record were disturbed (with the default being mean air temperature) and the profile resolution was varied (default resolution, fine = double

resolution or coarse = half resolution) to include possible effects that different initial layer thicknesses might have on model performance.

SNOWPACK model settings and parameters formed the third and last main experiment degree of freedom. An overview of the considered settings is shown in table 2.3. Most are discrete options and thus no distributions had to be assumed, with the exception of the new snow grain size, fixed albedo value, new snow fixed density and aerodynamic roughness length (table 2.6). Please note that some settings are incompatible: for example, when setting SW_MODE to ”BOTH” *SNOWPACK* is using measured albedo from the atmospheric forcing record and ALBEDO_PARAMETERIZATION or ALBEDO_FIXEDVALUE will go unused. Similarly, when specifying HN_DENSITY as “FIXED”, the new snow fixed density value under HN_FIXEDVALUE will be used and the HN_PARAMETERIZATION key becomes obsolete.

Table 2.4: *Distributions used to sample atmospheric forcing errors from. Note: rather than disturbing the measured outgoing shortwave radiation, “measured” albedo ($\alpha = Q_{so}/Q_{si}$) was disturbed. Symbols correspond to those in table 2.1.*

Variable	Units	Distribution	Parameters	1st, 99th perc.	Type	Reference
T_a	[K]	normal	μ = 0, σ = 1.3	-3, +3	additive	Raleigh et al. (2015)
RH	[%]	normal	μ = 0, σ = 10.7	-25, +25	additive	Raleigh et al. (2015)
U	[m/s]	normal	μ = 0, σ = 1.3	-3, +3	additive	Raleigh et al. (2015)
Q_{si}	[W/m²]	normal	μ = 0, σ = 43	-100, +100	additive	Raleigh et al. (2015)
α	[-]	normal	μ = 1, σ = 0.05	0.88, 1.12	multiplicative	own assumption
Q_{li}	[W/m²]	normal	μ = 0, σ = 10.7	-25, +25	additive	Raleigh et al. (2015)
P	[m/hr]	normal	μ = 1, σ = 0.107	0.75, 1.25	multiplicative	Raleigh et al. (2015)

Table 2.5: *Distributions used to represent uncertainty in the initial profile. Note that temperature, density and grain size profiles were disturbed, while bond radius to grain radius ratio (rb/rg), dendricity (dd) and sphericity (sp) were simply set and assumed constant with depth (marked “absolute” under type).*

Variable	Units	Distribution	Parameters	1st, 99th perc.	Type	Reference
T_{snow}	[K]	normal	μ = 0, σ = 2.5	-5.8, +5.8	additive	own assumption
ρ	[kg/m³]	normal	μ = 0, σ = 40	-93, +93	additive	based on Fausto et al. (2018)
rg	[mm]	normal	μ = 1, σ = 0.3	-0.7, +0.7	multiplicative	own assumption
rb/rg	[-]	uniform	min = 0.1, max = 0.9	-	absolute	own assumption
dd	[-]	uniform	min = 0, max = 0.5	-	absolute	own assumption
sp	[-]	uniform	min = 0.5, max = 1.0	-	absolute	own assumption
T_{insu}	[K]	normal	μ = 0, σ - 2.5	-5.8, +5.8	additive	own assumption
resolution	[-]	uniform	fine, default or coarse	-	absolute	own assumption

Table 2.6: *Distributions used to represent uncertainty in SNOWPACK model parameters (not measured).*

Variable	Units	Distribution	Parameters	1st, 99th percentile	Reference
ALBEDO_FIXEDVALUE	[-]	normal* *capped	μ = 0.8, σ = 0.1* *max = 0.99	0.57, 0.99	own assumption, based on Riihelä et al. (2019)
HN_FIXEDVAUE	[kg/m³]	normal	μ = 320, σ = 44	218, 422	own assumption, based on Fausto et al. (2018), Vandecrux et al. (2018)
NEW_SNOW_GRAIN_SIZE	[mm]	normal	μ = 0.25, σ = 0.1	0.017, 0.483	own assumption
ROUGHNESS_LENGTH	[m]	custom	-	-	refer to appendix A

Uncertainty Class Realisations

For each of the three uncertainty classes an ensemble of realisations was generated. This was done quasi-randomly using Sobol sequences, with the number of realisations determined as $n = N(k + 2)$ with N set to 100 and k the number of parameters in that class. Conveniently, the grouping of two of the three classes aligned with the input files required for *SNOWPACK*: the atmospheric forcing parameters are captured in **.smet* files and the initial profile in **.sno* files. With seven and eight variables apiece (tables 2.4 and 2.5), 900 **.smet* and 1000 **.sno* files were created (each containing one uncertainty

class realisation) and assumed to represent the uncertainty inherent to the atmospheric forcing and initial profile classes, respectively. Each file was assigned an ID, and it was this ensemble of files (or uncertainty class realisations) that was eventually sampled to generate model realisation.

Sampling realisations for the third class was less trivial because it consisted of a mix of continuous parameters (table 2.6) and discrete settings (table 2.3), some of which were mutually exclusive. The continuous settings (four in total) were combined and sampled quasi-randomly as before, resulting in 600 realisations. For the discrete settings, all combinations were considered and incompatible configurations were subsequently excluded. Because this heavily favoured configurations using parameterized as opposed to measured albedo, the configurations using measured albedo were repeated until this was balanced out. Model runs were sampled from the virtual ensemble of all possible combinations of the 600 continuous setting realisations and the albedo-balanced collection of discrete settings.

Taking Samples

The uncertainty class realisations were sampled to fill matrices A and B from section 2.3, which was again done in a quasi-random manner using Sobol sequences. In the case of the main experiment matrices A and B each contained three columns, one for each degree of freedom, and N (the number of rows) was kept at 250 to keep the number of model runs within computational feasibility. Columns of A and B were used to create matrices $A_B^{(1)}$, $A_B^{(2)}$ and $A_B^{(3)}$ required for simultaneous assessment of S_i and S_{Ti} which resulted in a total of $N(k+2)$ = 1250 model runs for the main experiment, henceforth referred to as the main experiment ensemble.

2.4.2. Meteo Experiment

The hypothesis for the main experiment was that model performance would be particularly sensitive to the atmospheric forcing uncertainty class and the “meteo experiment” was designed to identify exactly which atmospheric forcing parameters are important. As such it contains eight degrees of freedom: the seven atmospheric forcing variables individually vs. all other uncertainties from the main experiment (initial profile biases and *SNOWPACK* settings taken together in a single degree of freedom, fig. 2.3b). On an individual level the distributions (or options) from which realisations were drawn remain unchanged from the main experiment (tables 2.4, 2.5, 2.6 and 2.3). However, this time realisations were generated for the seven atmospheric forcing variables individually rather than as a group and the earlier initial profile realisations were lumped together with the model setting realisations. Sampling matrices A and B contain eight columns (one for each degree of freedom), which in turn were combined into eight $A_B^{(i)}$ matrices. N was kept at 250, for a total of $N(k+2)$ = 2500 model runs known as the meteo experiment ensemble.

2.4.3. Undisturbed Experiment

While the model settings and parameters available in *SNOWPACK* were varied the atmospheric forcing and initial profile were kept fixed for the “undisturbed experiment” (undisturbed - no biases were introduced). The goal of the settings experiment was to gain more insight into exactly which model settings are important for model performance and possibly which configurations perform well by eliminating some of the interaction effects that come with simultaneously disturbing the atmospheric and initial profile records. Considered model settings were subdivided in four degrees of freedom (fig. 2.3c):

1. Grain size:
   NEW_SNOW_GRAIN_SIZE only.

2. Albedo:
   SW_MODE, SNOW_ALBEDO, ALBEDO_PARAMETERIZATION, ALBEDO_AGING, SW_ABSORPTION_SCHEME, ALBEDO_FIXEDVALUE.

3. Density:
   HN_DENSITY, HN_PARAMETERIZATION, HN_FIXEDVALUE.

4. Other:
   ATMOSPHERIC_STABILITY, ENABLE_VAPOUR_TRANSPORT, ROUGHNESS_LENGTH.

Compared to the main and meteo experiments, the undisturbed experiment only used the model settings distributions and options from tables 2.3 and 2.6. For each of the four uncertainty classes quasi-random realisations were generated in a similar fashion as for earlier experiments and subsequently sampled using A and B matrices, this time containing four columns each. With an additional four $A_B^{(i)}$ matrices and N kept at 250, the ensemble contains $N(k+2)$ = 1500 model runs: the settings experiment ensemble.

Table 2.7: *Summary of the three experiment ensembles.*

Name	Size	DoF	Classes	Remark
Main exp.	1250	3	atm. forcing, initial profile, model settings	-
Meteo exp.	2500	8	atm. forcing variables individually, all other	37 GHz only
Undisturbed exp.	1500	4	grain size, albedo, density, other settings	undisturbed

2.5. Setting the Stickiness Parameter

The stickiness parameter was the last snow parameter required to run the radiative transfer scheme using the SHS microstructure representation and was effectively treated as a calibration parameter for the combined snow and radiative transfer model. While most snow layer properties can be taken directly from the *SNOWPACK* output and serve as *SMRT* input (temperature, density, liquid water content and grain size) this is not the case for the stickiness parameter which is not modelled by *SNOWPACK*. Instead it was assumed constant with snow depth and throughout time. Given that until recently methods to objectively estimate the stickiness parameter of natural snow were unknown (Löwe and Picard, 2015) this is not unlike the approach followed by earlier studies using the SHS microstructure model (Picard et al., 2013, Tsang et al., 2007). The value of the stickiness parameter was determined using an iterative procedure:

1. A guess for the stickiness parameter was made and the undisturbed experiment's snow model output was translated into T_b time series using this stickiness parameter.
2. The "median modelled T_b time series" was constructed by taking the ensemble's median modelled T_b value at every available data point of the time series.
3. The error between the mean of the satellite observed T_b time series and the mean of the median modelled T_b series was computed.
4. A new guess for the stickiness parameter was made: if the modelled ensemble overestimated T_b, the stickiness parameter was lowered (corresponding to less sticky/less clustered snow conditions). Vice versa, if the modelled ensemble underestimated T_b the stickiness parameter was increased.

The procedure was continued until the stickiness parameter was found that minimised the error between the mean of the median modelled T_b time series and the satellite observed signal. This was done twice: once for the 37V channel (found stickiness parameter of 0.25 for 37 GHz) and once for 19V (stickiness parameter of 0.10 for 19 GHz, but only because it could not be lowered further). The same stickiness values were then used to translate the main and meteo experiment's snow model output into modelled T_b time series. So while this study presents results using a constant stickiness parameter irrespective of the considered experiment, the assumed stickiness did change depending on the considered T_b frequency. The reasoning behind this approach is that T_b at different frequencies is sensitive to snow properties from different depth ranges and that older snow at deeper depths is expected to show different microstructural properties when compared to fresh snow that has not yet experienced any melt-refreeze cycles. Important to note is that given the size of the ensembles it was computationally intensive to translate the *SNOWPACK* output into T_b time series. Therefore the iterative procedure outlined above was run on a subset of datapoints: while it did use each and every ensemble member of the undisturbed experiment, to speed up run times only one in ten dates of every time series was modelled and compared with observations. The full time series were only modelled once an appropriate stickiness parameter was found.

Another approach to the unknown stickiness parameter would be to treat it as a free parameter that is allowed to evolve over the course of the simulation period to then solve for it using the independent satellite observations. Although in theory it would be possible to model a range of T_b time series corresponding to a range of stickiness parameter values (in order to invert for it using the observations) the computational cost to do this for a whole experiment's ensemble far exceeded the computational resources available for this project. Additionally, if the stickiness was treated as a free parameter on a per ensemble member basis it would compensate for any differences present in the snow model output to maximally match every ensemble member's time series to the observations, undermining the idea behind the sensitivity analysis. To get around these two problems and still gain insight into a possible seasonal stickiness signal one ensemble member's snow model output was used as *SMRT* input for a broad range of stickiness parameter values and the output of these model runs was used to invert for the stickiness required to match the satellite observed signal. This was done for the 37V channel using the main experiment's ensemble member that showed minimum offset between the means of the observed and modelled T_b signal (when using the ensemble's "standard" stickiness parameter of 0.25).

2.6. Quality Indicators

SMOWPACK model performance was evaluated directly by comparing modelled snow density profiles to in-situ firn core observations and the *SNOWPACK-SMRT* combined model by comparing modelled T_b time series to the satellite observed time series. Quality indicators include:

- Comparison of mean modelled snow and firn density taken over 0 - 1, 1 - 4 and 4 - 15 m depth ranges compared to observations from firn cores.
- RMSE of modelled snow and firn density profiles for dates with firn core observations (depth range of 0 - 15 m and 10 cm resolution). Within the SA framework an ensemble member's "density profile RMSE" is defined as the mean RMSE of the four RMSEs computed for the four dates with in situ observations.
- RMSE of modelled T_b time series compared to satellite observations (daily data points over June 1998 to December 2017). These were evaluated per modelled channel (37V, 37H, 19V, 19H) and used as input for the SA framework.
- Comparison of the slopes of linear segments fitted to the modelled/observed T_b signal when taken over an entire accumulation season (October up to and including May).

For visual inspection ensemble spread of mean density evolution, modelled density profiles, T_b time series and fitted accumulation season slopes were plotted together with observations and per experiment ensemble. The median modelled series/profile/slope (constructed by taking the median modelled value at every date/depth) was compared with observations and the 50 and 95% range in modelled quantities is reported. Density is an interesting variable to quantify snow model performance because it can and is regularly unambiguously measured in the field (it was therefore the only in-situ validation data considered in this study). Additionally, density gives a good indication of available pore space and thus potential for meltwater retention in firn. Given a certain profile, densities can vary from just around 100 kg/m^3 for fresh snow to the pure glacial ice density of 917 kg/m^3. Under the assumption that no liquid water is present the volumetric void fraction is simply unity minus the ratio of modelled to pure ice density. In terms of modelled T_b time series, in addition to time series RMSE the fitted slopes were included in the analysis to see if modelled T_b trends match well with observations when taken over an entire accumulation season. The snow properties in the top layers of the snowpack are expected to vary considerably over the course of the accumulation season at DYE-2, as the melt season leaves an old, dense snowpack to be gradually buried under new layers of fresh snow.

Within the SA framework estimates for the sensitivity indices focus on density profile and T_b time series RMSE, for the latter RMSE was computed over the whole simulation period (June 1998 to December 2017) which gives one estimate for every considered sensitivity index based for the sample size of 250. To gain insight into the uncertainty associated with these estimates a test was performed where the time series were split up in yearly chunks running from June 1st to May 31st, which increased the number of time series 19 fold allowing for 19 estimates per index (once for every year with T_b data)

without reducing the sample size or using data more than once. This does however come at a cost: individual time series contain less data points than before and ensemble members are now correlated as they are based on chunks coming from the same continuous snow state time series. The spread in estimates was visualized by means of boxplots showing median estimates and their interquartile range and used to confirm whether the original estimates (using the full time series) based on are robust in the face of a limited sample size.

2.7. Comment on Runtimes

For each of the experiments, input data errors were introduced, model configurations were set and the combined snow and radiative transfer model was run. This resulted in a total of 5250 model runs or a total simulation period of just over 100.000 years. The *SNOWPACK* component of the combined model took two weeks to complete using 64 CPUs on a modest cluster and resulted in 1.7 TB of daily snow model output data. In terms of *SMRT*, both the 37 GHz (all three experiments) and 19 GHz (the main and undisturbed experiments) observations each took another two weeks to model. Interesting to note for future efforts is that for the *SMRT* component memory bandwidth seemed to be the bottle neck, not CPU computational power. Results are presented per experiment.

3

Results

3.1. Main Experiment

3.1.1. Snow Densities

Fig. 3.1 shows a large spread in mean snow densities of the top 1, 1 to 4 and 4 to 15 metres of the main experiment *SNOWPACK* output. For the top 1 m for example, the mean 95% range is equal to 447 kg/m^3 taken over the entire simulation period. 22% of data points show a 95% range of at least 600 kg/m^3 meaning that on those days conditions in the top 1 m vary from light snow to ice-like depending on which ensemble member is considered. This is especially true for the ablation and very beginning of the accumulation season with 72% of those days occurring in the months of June through September. As surface melt is known to have a big impact on snow density this is likely a result of differences in the amount of melt experienced by ensemble members, depending on the disturbances introduced in the atmospheric record (e.g. positive bias in air temperature means more melt days) or applied model settings (e.g. lower albedo means more energy available for melt). Later in the accumulation season the mean modelled densities for the 1st metre and their 95% range shrink, a result of fresh snow accumulating on top of the snowpack. Of the 24% of days with a 95% range of less than 300 kg/m^3, 76% fall in the months of December through May. The 2002-2003 accumulation season forms an exception: a consistently high range in 1st metre mean modelled density is maintained throughout. This is a result of exceptionally little precipitation over that season (measured 0.133 m water equivalent from October 2002 to May 2003 compared an average of 0.309 m for the other accumulation seasons), meaning that mean densities at the end of the 2002 melt season stay largely preserved until the 2003 melt season comes along. Seasonal effects are muted when looking at deeper depth ranges (fig. 3.1b and 3.1c) which is to be expected as deeper layers are buffered from the seasonal atmosphere by the snow layers above. At these depths a trend for firn densification becomes obvious, the median mean density for 4 - 15 m is equal to 563 kg/m^3 at the beginning of the simulation (June 1998) and grows to 684 kg/m3 by the end (December 2017). Spread however remains great with mean 95% ranges of 525 and 491 kg/m^3 for the 1 - 4 and 4 - 15 m of *SNOWPACK* output, respectively. After the start of the simulation period the spread in mean modelled density can be seen to increase steadily. E.g. for 1 - 4 m the mean 95% range for the first three years is equal to 321 kg/m^3, compared to 561 kg/m^3 for the remainder of the simulation period. We can assume that ensemble members diverge due to the different biases introduced in the atmospheric forcing record and/or model configuration.

The ensemble of mean modelled densities agrees reasonably well with the available firn core observations (confirmed visually from fig. 3.1). The spread in modelled densities and limited number of observations makes it difficult to draw decisive conclusions but the median time series of modelled snow densities tends to underestimate density for the 1st metre (on average by 80 kg/m^3, three observations) and overestimate for deeper depth ranges (on average by 50 kg/m^3 for 1 - 4 m and 35 kg/m^3 for 4 - 15 m, four observations). It would be easy to blame this on shortcomings in *SNOWPACK* or model input but could also be the result of a non-linear response to disturbances in the atmospheric forcing. I.e. in case of air temperature the absolute increase in snow density caused by a positive 1 K bias might be larger than the absolute decrease caused by a negative 1 K bias. Because the firn core observations come from the end of the accumulation season (all from May, before the onset of melt)

such an effect would be most pronounced at deeper depths, below last season's accumulation.

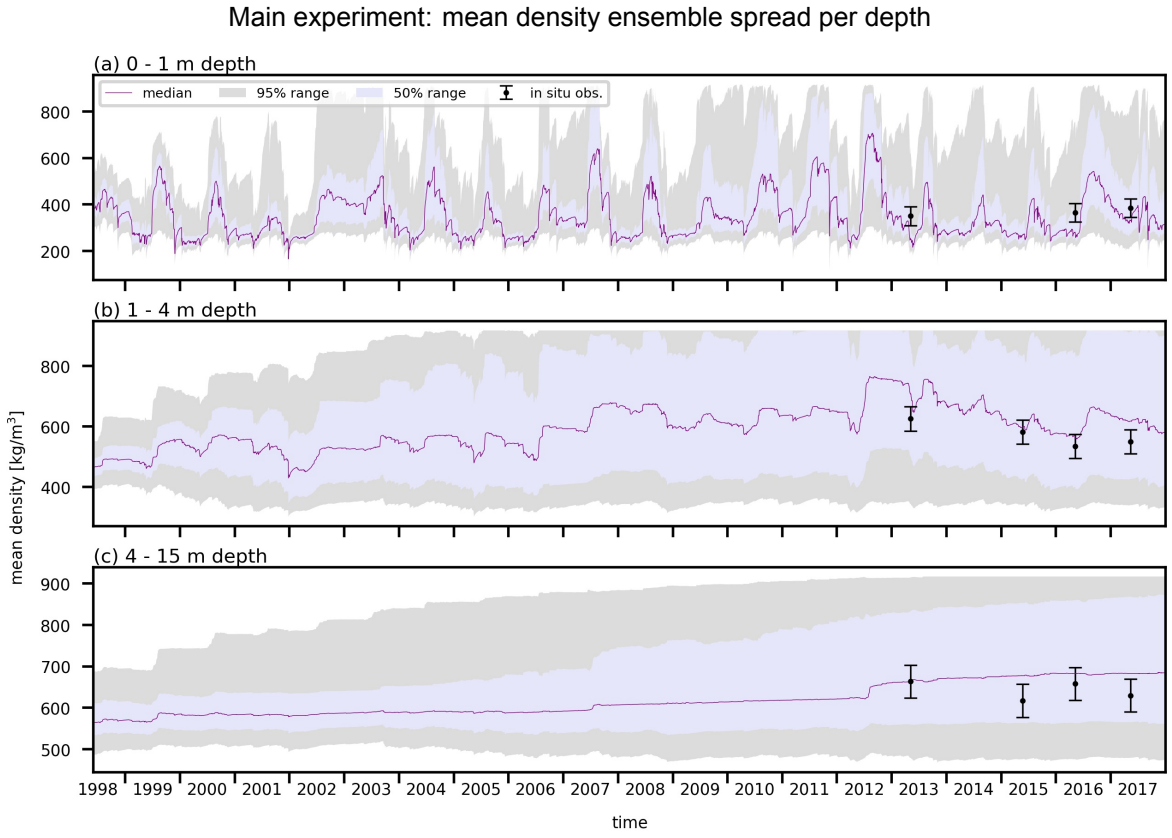


Figure 3.1: *Ensemble spread of mean modelled densities for 0 - 1 m (**a**), 1 - 4 m (**b**) and 4 - 15 m (**c**) depth of main experiment SNOWPACK output shows large spread (an average 95% range of 447, 525 and 491 kg/m^3 respectively). Seasonal effects clearly visible at shallow depths: modelled densities and spread are greatest at the end of the ablation season. In-situ observations come from firn cores and include 40 kg/m^3 error bars. The median modelled time series were constructed by taking the ensemble's median value at every data point.*

Comparison of modelled density profiles with firn core observations confirms the large spread in modelled densities (fig. 3.2). This is most pronounced in the first few metres below the surface (mean 95% range of 595 kg/m^3 for 1 - 4 m on 05.05.2013) after which with increasing depth the spread decreases steadily (mean 95% range of 448 kg/m^3 for 4 - 15 m on 05.05.2013). Unlike the observations which show clear layering the median profiles appear smooth. As can be seen from the best modelled profiles (in yellow), on an individual level some ensemble members do show clear layering albeit not necessarily as pronounced as in the observations. Given the fine resolution of 10 cm slight mismatches in the depths at which certain stratigraphy manifests itself will naturally result in a smooth median profile. This also affects the performance metric used to determine which ensemble member to plot in yellow (min. RMSE), which is suspected of preferring an accurate smooth density profile over a precisely stratified one with a depth mismatch compared to the observations. Taken over 0 - 15 m minimum RMSEs of 104, 123, 119 and 105 kg/m^3 were found for the 2013, 2015, 2016 and 2017 profiles respectively, all corresponding to different ensemble members. Qualitatively the median modelled density profiles follow the general trends of the observations well, e.g. in 2013 the step increase in density at a depth of ca. 1 m (likely the result of the extreme 2012 melt season) is captured well (fig. 3.2a). For 2015 and 2017 the median modelled profiles visibly overestimate density, especially at depth. Taken over 0 - 15 m the mean error between the median modelled and observed profiles is equal to 7.72, 45.0, 15.6 and 46.5 kg/m^3 for the 2013, 2015, 2016 and 2017 profiles respectively.

Remarkable is that a significant portion of ensemble members develop ice-like densities at shallow depths, in 2013 at least 25% of profiles show a density of 917 kg/m^3 between depths of 1 and 7 m. In subsequent years this depth range grows, in the spring of 2017 it extends to well below 9 m depth. In fact, further analysis has shown that on 05.05.2013 552 out of 1250 members contain an ice slab of at

least 1 m thick and this number grows to 578 out of 1250 on 11.05.2017. It is clear from the observations that no such ice slabs were present at DYE-2. A detailed investigation into ice slab development is outside the scope of this study but would be interesting for follow-up work as impermeable ice slab formation was identified as a tipping mechanism for Greenland Ice Sheet mass loss (MacFerrin et al., 2019).

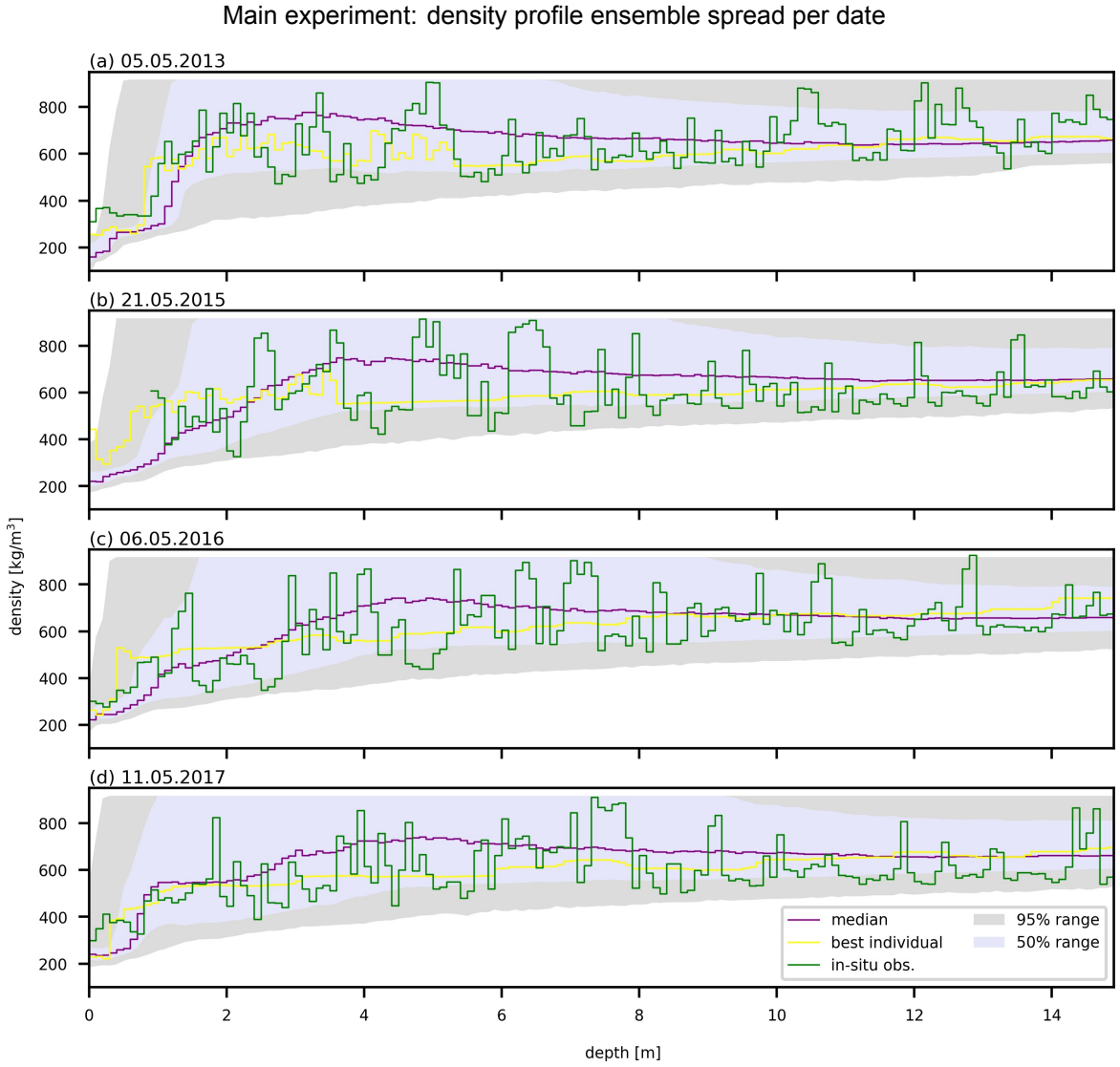


Figure 3.2: *Comparison of modelled density profile ensemble spread with an actual firn core observations in May 2013 (**a**), 2015 (**b**), 2016 (**c**) and 2017 (**d**) confirms large spread in modelled densities in the top 15 m of SNOWPACK output. The median modelled profiles were constructed by taking the ensemble's median density value at every depth and follows the general trend of the observed profile well, but does not show layering of individual ensemble members. Best ensemble members in terms of min. RMSE plotted in yellow (RMSE of 104, 123, 119 and 105 kg/m^3 respectively). All modelled and observed profiles were converted to have a depth resolution of 10 cm.*

3.1.2. T_b Time Series

Major events of the 37 GHz satellite observed signal are captured well in the median series of the modelled ensemble, but when taken over a whole accumulation season the trend of the modelled ensemble does not match with observations (fig. 3.3). With a mean 95% range of 100.6 K for the 37V channel (96.1 K for 37H) modelled T_b ensemble spread is substantial. Visually the timing of modelled melt events (indicated by the major T_b spikes in fig. 3.3a and 3.3b) match up well with observations. The same can be said for the timing and amplitude of smaller accumulation season events occurring on the scale of a couple of days or weeks. When looking at entire accumulation seasons however,

a striking effect becomes apparent: where observed T_b tends to start low after the melt season and then steadily increases until the next, modelled time series show an opposite trend (high to low over an accumulation season). This is true for both the V and H channels and will henceforth be referred to as the "opposite slope effect". Some years show greater modelled T_b ensemble spread than others, for example the 2002-2003 accumulation season (for whole of 2002 and 2003 mean 95% range at 37V is equal to 108.9 K). This is consistent with the ensemble spread for modelled snow densities (fig. 3.2a). At 37V the absolute error between the mean of the observed signal and the mean of the median modelled series is only 0.06 K, but recall that this is by design as the stickiness parameter was set to minimize this error at V polarisation. For 37H this error grows to 9.80 K, with the median modelled series overestimating T_b. This difference is too great to be compensated by a realistic topography-induced change in incidence angle and must be an effect of the modelling chain. RMSE of the median modelled series vs. observations is 20.0 K at 37V and 20.6 K at 37H. On an individual basis ensemble member's RMSE ranges from 17.8 to 62.6 K for 37V and from 19.9 to 56.9 K for 37H. Comparison of the three best performing (min. RMSE) ensemble members with observations at 37V (fig. 3.4) confirms that on an individual basis modelled time series are capable of accurately reproducing small T_b events (time scale: couple of days to weeks) but that even for these good performers the opposite slope effect is clearly visible over most accumulation seasons.

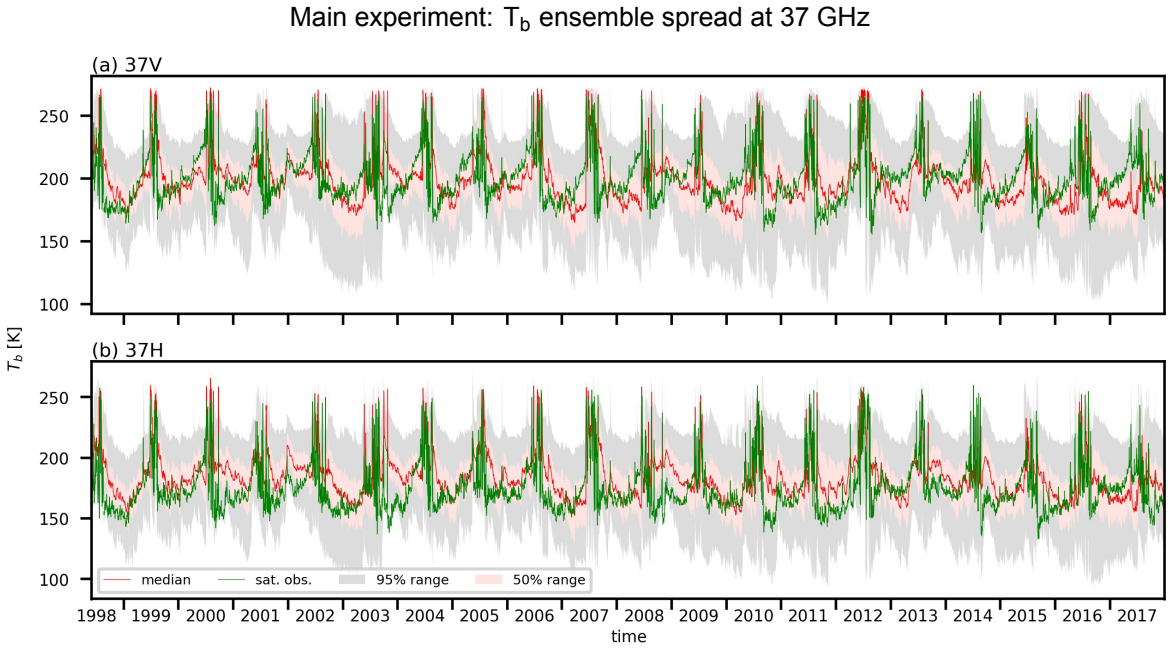


Figure 3.3: *Comparison of 37 GHz modelled T_b ensemble spread with independent satellite observations at V (**a**) and H (**b**) polarisations show that timing of melt events and smaller variations in observed T_b are captured well, but taken over a whole accumulation season the modelled series show an opposite trend ("opposite slope effect"). Spread in modelled T_b is great (mean 95% range of 100.6 K at 37V, 96.1 K at 37H). The median modelled series were constructed by taking the ensemble's median T_b value at every data point (data points are daily).*

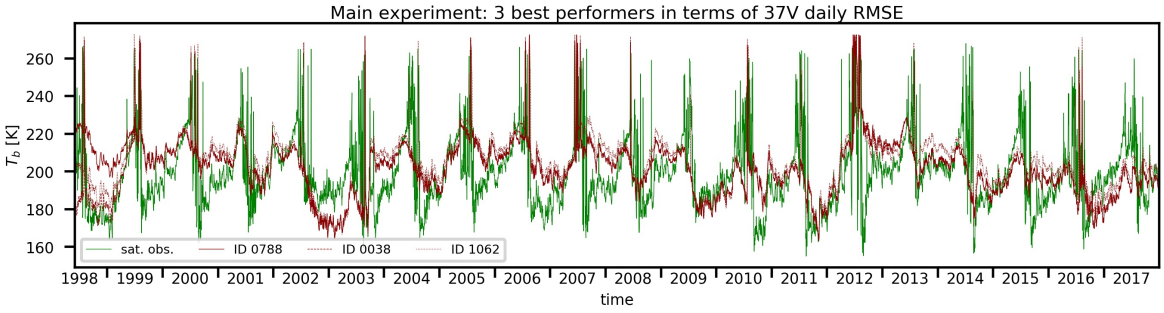


Figure 3.4: *Comparison of the three best performing (min. RMSE) main experiment ensemble members with observations at 37V. Small T_b events are captured well in individual time series but the opposite slope effect is clearly visible over most accumulation seasons. RMSE equal to 17.8, 18.3 and 18.5 K, respectively.*

While at 19 GHz the timing of melt events is captured well in the median series of the modelled ensemble, smaller high frequency variations in observed T_b are poorly represented (fig. 3.5). When they are the amplitudes of the observed variations appear muted. The opposite slope effect manifests itself again: after a melt event the observed T_b signal drops virtually instantly and the modelled series fail to replicated this behaviour. Instead the median modelled series comes down gradually over the course of the entire accumulation season. Modelled T_b ensemble spread remains substantial with a mean 95% range of 103.3 and 100.6 K at V and H channels respectively. The median modelled series visibly overestimates T_b at both 19V and 19H: the absolute error between the mean of the observed signal and the mean of the median modelled series is 16.2 K for 19V (compared to 0.06 K at 37V) and this grows to 30.9 K for 19H. The routine to set the stickiness tries to minimize the error for the vertically polarized channel, but the stickiness parameter could not be lowered below 0.10 before running into exceptions in *SMRT* - resulting in the offset between modelled and observed T_b. Not surprisingly computed RMSEs are greater for 19 GHz than for 37 GHz. RMSE between median modelled series and observations is 24.9 K for 19V and 37.1 K for 19H. On an individual basis ensemble member's RMSE range from 19.3 to 68.3 K for 19V and 21.8 to 80.3 K for 19H.

The mismatch in slopes fitted to the main experiment's modelled T_b and observed time series is obvious from fig. 3.6. The slopes fitted to the observed accumulation season's T_b series show an upward trend irrespective of the considered year or T_b frequency (no exceptions). Slopes fitted to the modelled series on the other hand show either a downward trend (all median slopes for 19V in fig. 3.6b, half of the median slopes for 37V in fig. 3.6a) or underestimate the upward trend (remaining half of the median slopes for 37V). Some years show stronger disagreement between modelled and observed slopes than others. At 37V, the 2003-2004, 2006-2007 and 2008-2009 seasons show the strongest disagreement with angles between the observed and median fitted slopes beyond perpendicular. Closer inspection revealed that for some years the fitted slopes were affected by a particularly early onset of melt (in the month of May, which was assumed accumulation season) making the slopes fitted through the observed and modelled series falsely appear similar, in particular for 2011-2012.

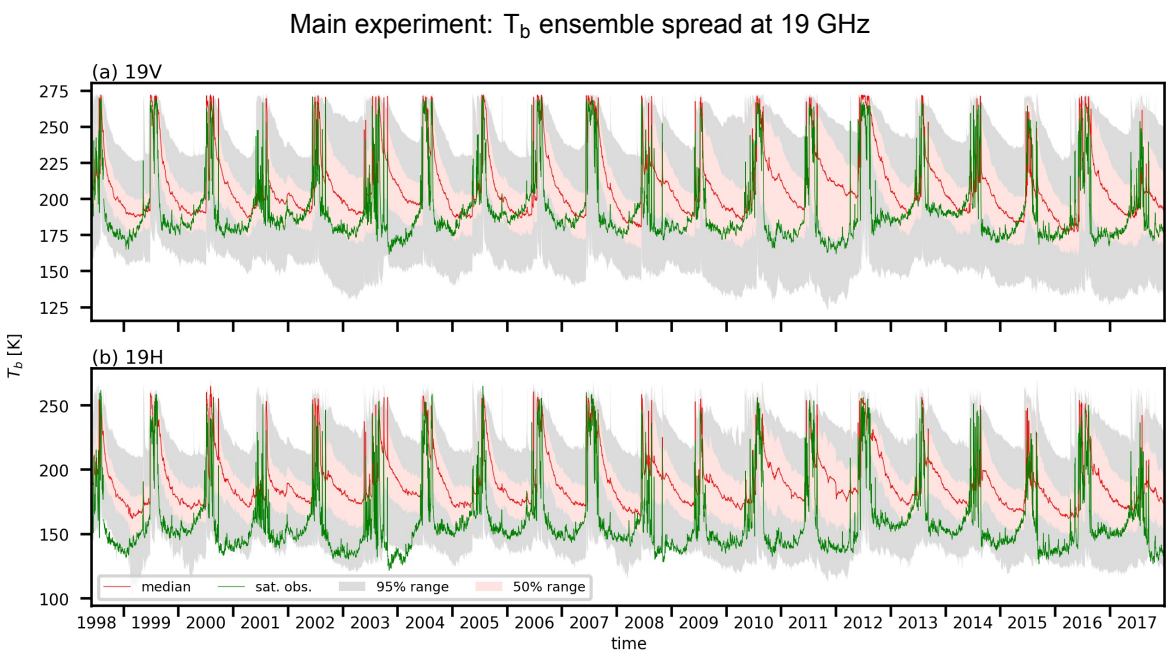


Figure 3.5: *19 GHz modelled T_b ensemble appears smooth compared to observations at V (**a**) and H (**b**) polarisations and fails to reproduce the observed drop in T_b immediately after the melt season (opposite slope effect). Modelled ensemble spread is great (mean 95% range of 103.3 at 19V, 100.6 K at 19H) and the median modelled series overestimates observed T_b (difference in mean of 16.2 K at 19V, 30.9 K at 19H). The median modelled series was constructed by taking the ensemble's median T_b value at every data point (data points are daily).*

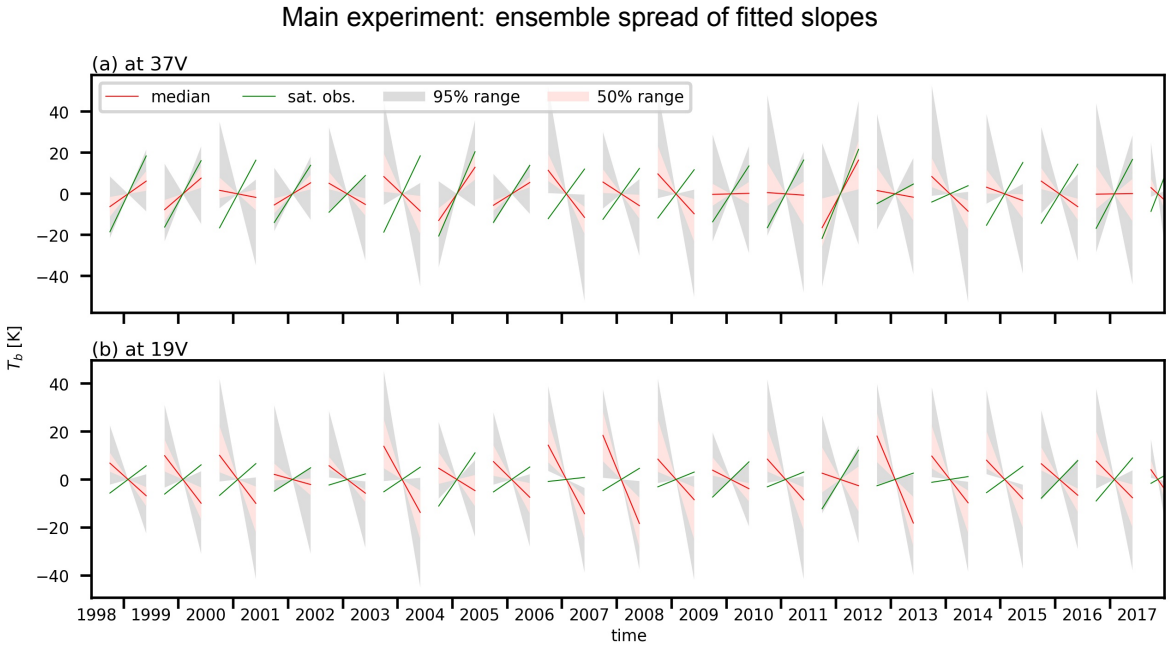


Figure 3.6: *The ensemble spread of linear segments fitted to yearly T_b signals (at 37V (**a**) and 19V (**b**)) illustrate the mismatch between observed (in green) and modelled (red) slopes when taken over an entire accumulation season. Accumulation seasons were assumed to run from October up to and including May. Means were removed on an individual basis to plot all segments at the same T_b level. Original segments fitted to the observed signal are shown in appendix C.*

3.1.3. Sensitivity Indices

Model performance is sensitive to changes in atmospheric forcing, model settings and their interaction but not to changes in the initial profile. This is true for the combined model when quantifying model performance as RMSE of the modelled vs. observed time series at 37V (fig. 3.7a) and for the snow model on its own when using modelled and observed density profiles to compute RMSE (fig 3.7a). In case of 37V RMSE 40% of the main experiment's variance in combined model performance is explained by biases introduced in the atmospheric forcing, 33% by changing *SNOWPACK* model settings and parameters. The remaining variance in model performance (27%) comes from interaction effects between these two, indicated by the difference in first and total-order sensitivity indices (fig. 3.7a). Important to keep in mind is that these are statistical estimates based on an ensemble with limited size: sensitivity indices presented here come with their own uncertainty. Low sensitivity to the initial profile can be explained by the fact that T_b is most sensitive to the top layers of the snowpack and the long simulation period, initial properties of the top layers are quickly "forgotten" due to interaction with the atmosphere. In terms of modelled densities available firn core observations all come from the latter part of the simulation period (2013 onwards) and it appears that by then also deeper layers are no longer significantly affected by initial conditions, at least in terms of density. Model performance based on T_b is more sensitive to changes in model settings compared to performance based on density, possibly a result of the fact that model settings include the new snow grain size parameter which has a big effect on the top layer's snow microstructure and thus scattering behaviour but less on modelled densities, especially at depth.

Results presented here focus on the 37V channel because unlike the other channels the modelled T_b ensemble at 37V shows no offset with respect to observations (fig. 3.3a). That said, it is important to keep the spread in modelled T_b and opposite slope effect in mind. The performance metric at 37V is therefore likely sensitive to variables that have a strong effect on the mean of modelled T_b series (i.e. variables that move the entire time series up or down) or the modelled accumulation season T_b slope behaviour (i.e. variables that cause more or less opposite slope effect), and only to a lesser extent to variables that effect the model's ability to accurately model smaller or shorter scale T_b fluctuations.

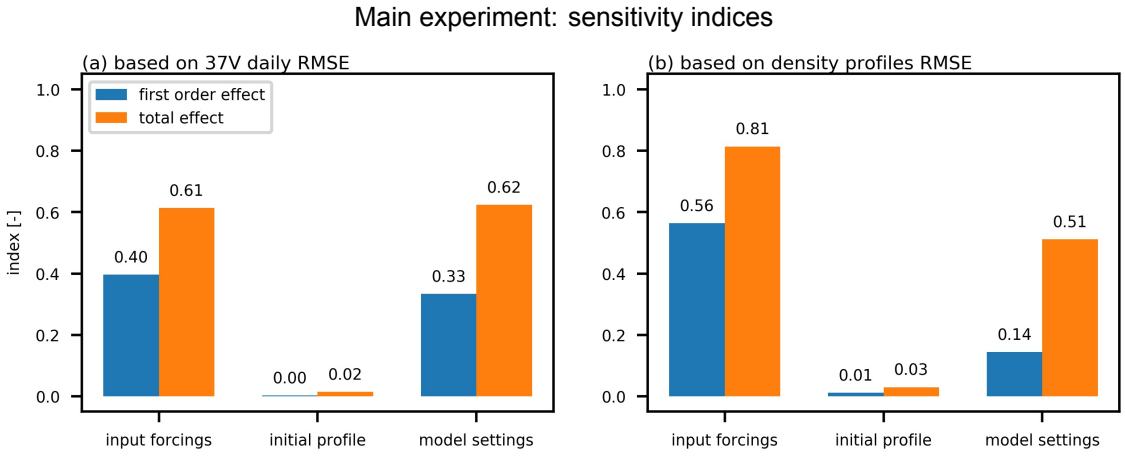


Figure 3.7: *Estimates for the main experiment first and total order sensitivity indices show that model performance is sensitive to biases introduced in the atmospheric forcing, varying model settings and their interaction effects but not to uncertainty in the initial profile. When moving from model performance based on 37V RMSE (**a**) to performance based on density profile RMSE (**b**), input forcings become comparatively more important. Estimates are based on a sample size of 250.*

3.2. Meteo Experiment

Biases and model settings of the meteo experiment come from the same distributions as the main experiment and as a result the ensemble spread of modelled snow densities and T_b time series are similar. The figures are shown in appendix D, the findings are in line with those of the main experiment. Only the sensitivity indices are presented here.

3.2.1. Sensitivity Indices

At 37V combined model performance is sensitive to biases introduced in the air temperature, precipitation and relative humidity record (and in that order, fig. 3.8a). 21% of the variance in model performance is explained by air temperature disturbances, 11% by precipitation and 9 % by relative humidity. When including interaction effects their order changes to air temperature, relative humidity and precipitation, although the differences are small for the last two (19 and 17%) in the face of sensitivity analysis estimation uncertainty. The negative sensitivity indices for in and outgoing shortwave radiation must be erroneous and are another reminder of uncertainties in S_i estimation, especially when the indices are small. The "all other" degree of freedom includes the main experiment's initial profile and model settings uncertainties combined and explains 43% of the variance in model performance at 37V (excl. interaction effects). Virtually all of this must come from the model settings (from fig. 3.7: influence of initial profile uncertainty negligible). Compared to T_b at 37 GHz performance based on modelled density profiles is more sensitive to air temperature and sensitivity to relative humidity, precipitation as well as initial profile and model settings (combined) is reduced (fig. 3.8b). Also interaction effects become more important when considering densities. These findings are in line with the results of the main experiment.

It makes sense that air temperature has a large influence on model performance as it controls the amount of melt(days) which is known to have a drastic effect on both snow microstructure (and thus grain size) and snow density. As the amount of precipitation is equivalent to snow fall, it controls how quickly older snow with higher density and bigger grain sizes is buried under layers of light, small grained snow. As a result it has a big influence on modelled T_b which is sensitive to both grain size and density of the top snowpack layers. The sizeable contribution of relative humidity disturbances to variance in model performance might come as more of a surprise, but does make sense as it controls snowpack ablation processes (evaporation and sublimation) and latent heat flux, physical processes that are incorporated in *SNOWPACK* (Lehning et al., 2002a). *SNOWPACK* is also skilful in the modelling of surface hoar formation, for which it uses latent heat flux and wind speed as input. This interaction might explain the significant wind speed total-order sensitivity index of 6% compared to its first order effect (1%, fig. 3.8a).

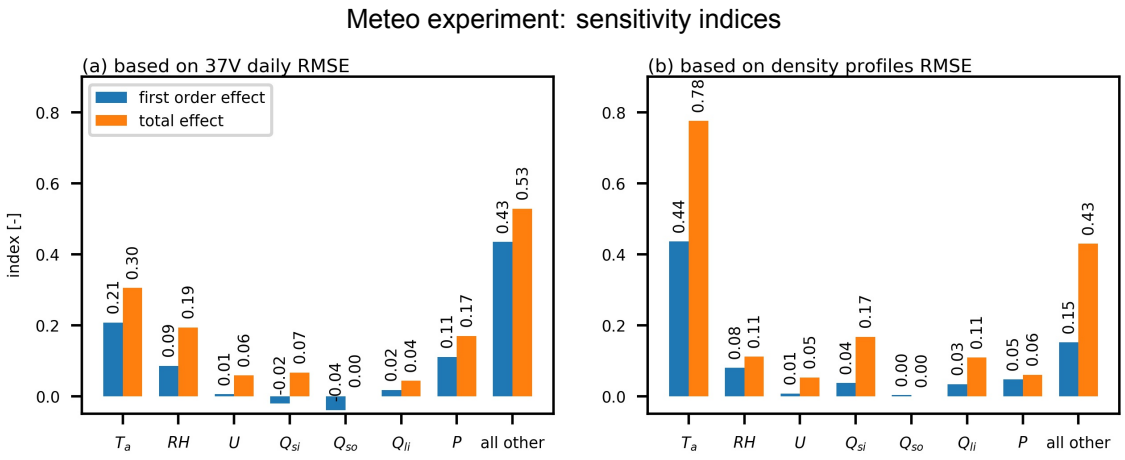


Figure 3.8: *Estimates for the meteo experiment first and total order sensitivity indices show that at 37V (**a**) model performance is most sensitive to biases introduced in air temperature, precipitation and relative humidity, respectively. When basing model performance on density profiles (**b**) air temperature and interaction effects become comparatively more important. Please note: the reported Q_{so} S_{Ti} of 0.0 was found to be erroneous (section 3.4). Estimates are based on a sample size of 250.*

3.3. Undisturbed Experiment

3.3.1. Snow Densities

Fig. 3.9 reveals considerably less spread in mean modelled densities for the undisturbed experiment compared to the main experiment ensemble (fig. 3.1), the result of no longer disturbing the initial profile or atmospheric forcing record. For the top 1 m of *SNOWPACK* output the mean 95% range is reduced from 447 kg/m^3 for the main experiment to only 241 kg/m^3 for the undisturbed case. However, spread between ensemble members remains large for the ablation seasons (especially later in the simulation

period) and for accumulation seasons with little precipitation (e.g. 2002-2003) where 95% ranges can exceed 400 kg/m^3. The steady increase in the spread of modelled densities over the course of the simulation period (fig. 3.9b and 3.9c) reveals that ensemble members still diverge even when forced by the same atmospheric record and only the model settings and parameters are varied. E.g. for 1 - 4 m the mean 95% range for the first three years is equal to 170 kg/m^3 compared to 452 kg/m^3 for the remainder of the simulation period. At depth there is still a trend of densification, albeit not as pronounced as for the main experiment. The 4 - 15 mean density is equal to 548 kg/m^3 at the beginning of the simulation and the median mean density grows to 638 kg/m^3 by the end (684 kg/m^3 for the main experiment). The median time series of modelled snow densities no longer overestimate density at deeper depth ranges (average error of only +21 kg/m^3 for 1 - 4 m and -16 kg/m^3 for 1 - 15 m, four observations) which confirms that the overestimation observed for the main experiment was indeed a result of non-linear responses to input forcing disturbances.

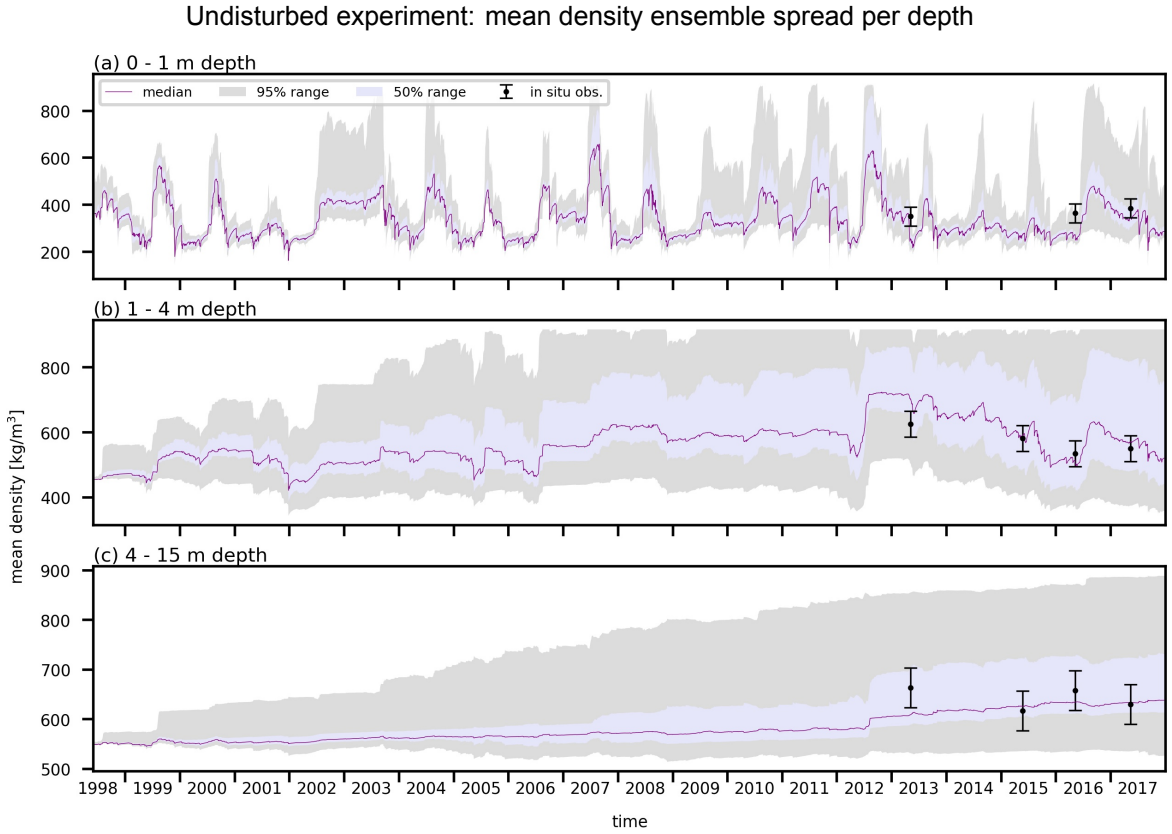


Figure 3.9: *Enemble spread of mean modelled densities for 0 - 1 m (**a**), 1 - 4 m (**b**) and 4 - 15 m (**c**) depth of undisturbed experiment output shows a reduced spread compared to the disturbed experiments (an average 95% range of 241, 409 and 224 kg/m^3 respectively). At shallow depths the spread in modelled densities remains large at the end of the ablation season. In situ observations come from firn cores and include 40 kg/m^3 error bars. The median modelled time series were constructed by taking the ensemble's median value at every data point.*

When comparing modelled density profiles of the undisturbed experiment (fig. 3.10) to those of the main experiment (fig. 3.2) findings are in line with the discussion above: there is considerably less spread in modelled snow densities (true over the entire depth range) and the median profiles no longer visibly overestimate densities at depth. For the 2013 profile (fig. 3.10a) the mean 95% range between 1 - 4 m is equal to 448 kg/m^3 (595 kg/m^3 for the main experiment) and between 4 - 15 m this shrinks to 358 kg/m^3 (448 kg/m^3). Median profiles still appear smooth compared to the stratified observations but the best individual profiles (min. RMSE, plotted in yellow) do show realistic layering in the first 8 m below the surface. Min. RMSEs of 115, 115, 134 and 95 kg/m^3 were found for the four profile dates respectively (all corresponding to different ensemble members). A portion of the ensemble members develop ice slabs at shallow depths, but the depth range in which they manifest and the number of ensemble members with ice slabs is significantly reduced from the main experiment. In 2013 at least

25% of profiles develop densities of 917 kg/m^3 between depths of 1 and 4 m, compared to between 1 and 7 m for the main experiment. On 05.05.2013 22.6% of ensemble members contain a continuous ice slab of at least 1 m thick, which grows to 28.3% on 11.05.2017 (compare to 44.2 and 46.2% for the main experiment).

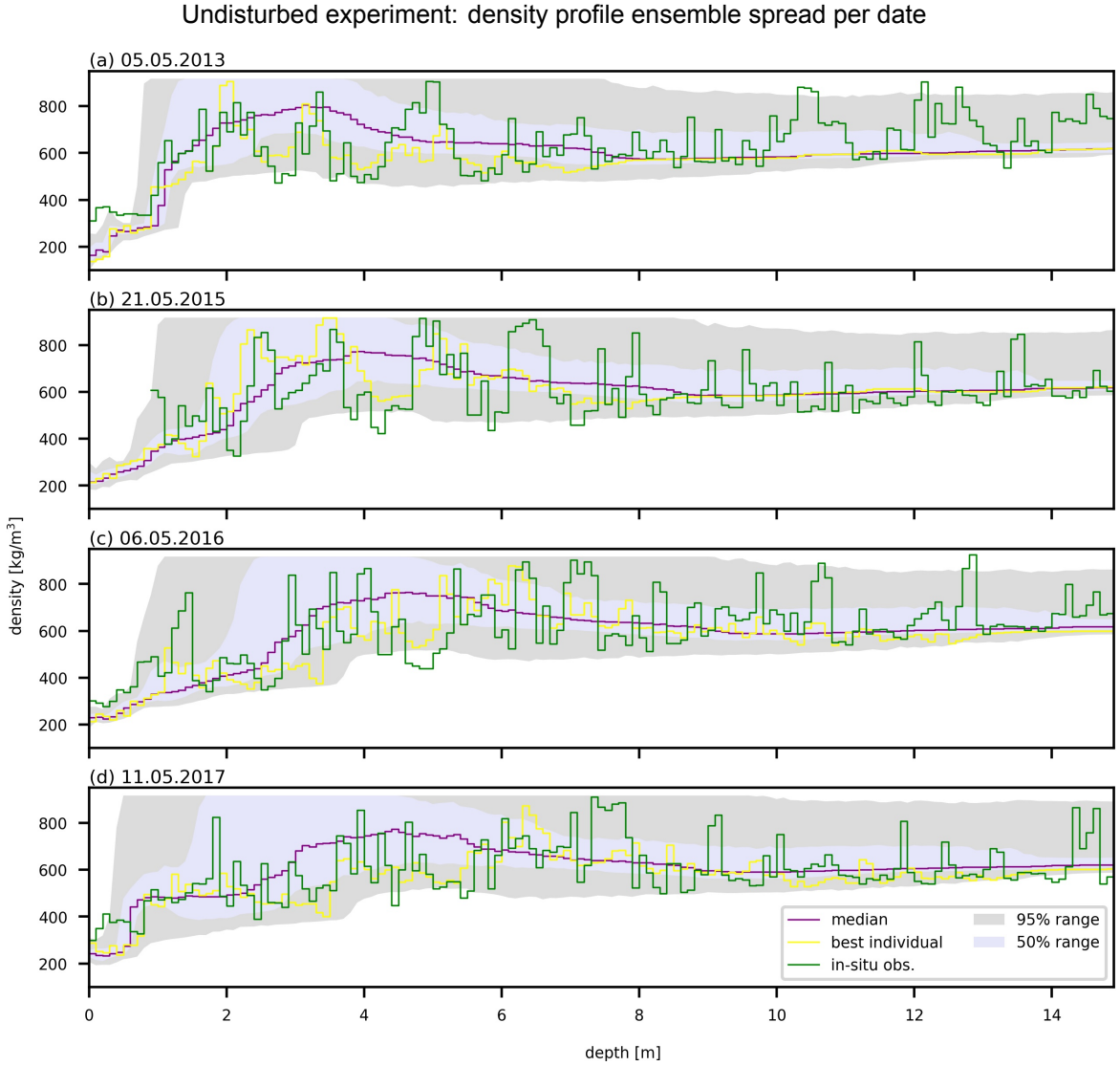


Figure 3.10: *Comparison of modelled density profile ensemble spread with an actual firn core observations in May 2013 (**a**), 2015 (**b**), 2016 (**c**) and 2017 (**d**) confirms reduced spread in modelled densities in the top 15 m of SNOWPACK output for the undisturbed experiment. Best ensemble members in terms of min. RMSE plotted in yellow (RMSE of 115, 115, 134 and 95 kg/m^3 respectively) show realistic stratigraphy until depths of 8 m. The median profiles were constructed by taking the ensemble's median density value at every depth. All modelled and observed profiles were converted to have a depth resolution of 10 cm.*

3.3.2. T_b Time Series

As expected visualisation of the spread in undisturbed modelled T_b (fig. 3.11) reveals a reduction in mean 95% range when compared to the main experiment: 74.3 and 71.6 K at 37V and 19V respectively (compared to 100.6 and 103.3 K for the main experiment). The remaining spread is the result of changing the *SNOWPACK* settings only. Other findings are in line with the main experiment's ensemble of modelled T_b observations. Worth noting is the opposite slope effect which reappears just as strong (fig. 3.12) and can therefore not be the result of disturbing the atmospheric forcing record and instead must be inherent to the modelling chain. RMSE between the median modelled vs. observations is equal o 21.6 and 23.8 K at 37V and 19V, respectively. On an individual basis ensemble member's RMSE range from 19.4 to 59.0 and 20.4 to 62.0 K at 37V and 19V. Figures and statistics for 37H and

19H are shown in appendix E.

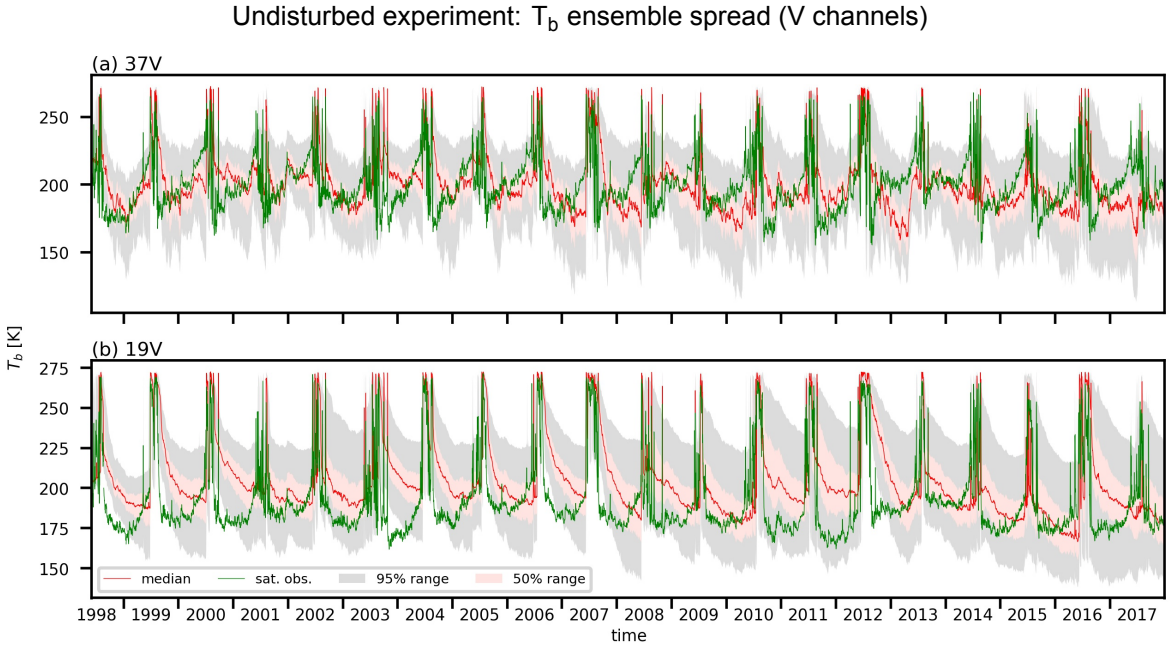


Figure 3.11: *The undisturbed experiment's 37V (**a**) and 19V (**b**) ensemble spread is reduced when compared to the main experiment but remains considerable with a mean 95% range of 74.6 (37V) and 71.6 (19V) K. The opposite slope effect still appears strong in the modelled ensemble. The median modelled series were constructed by taking the ensemble's median T_b value at every data point (data points are daily).*

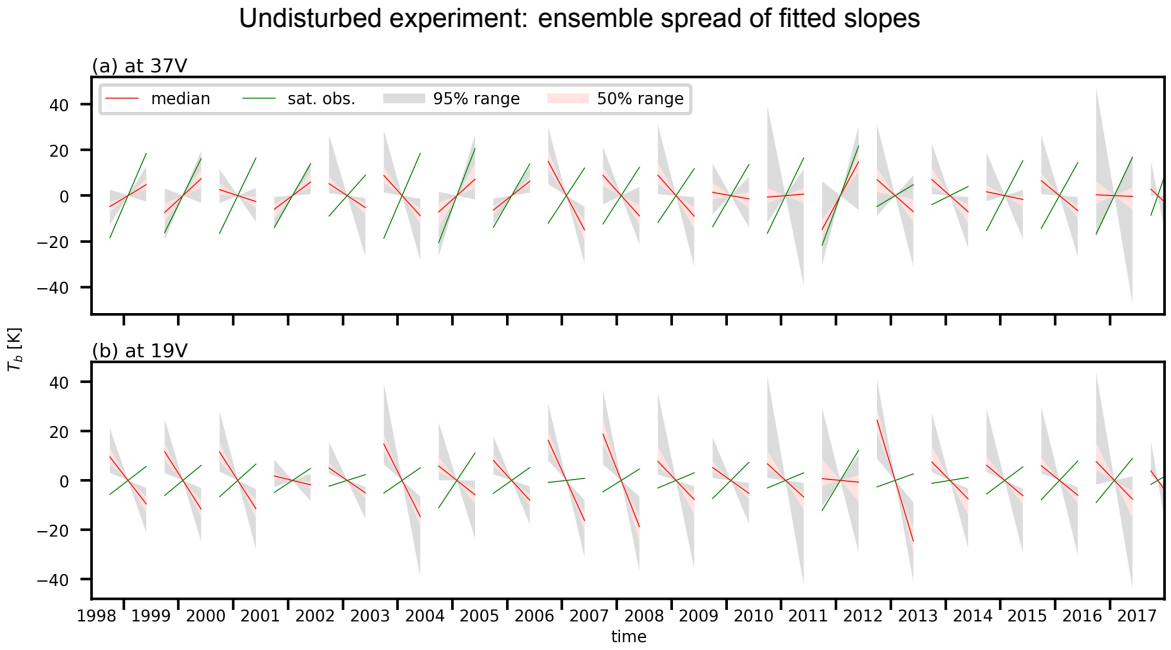


Figure 3.12: *The ensemble spread of linear segments fitted to yearly T_b signals illustrate the mismatch between observed (in green) and modelled (red) slopes when taken over an entire accumulation season. Accumulation seasons were assumed to run from October up to and including May. Means were removed on an individual basis to plot all segments at the same T_b level.*

3.3.3. Sensitivity Indices

Model performance based on 37V RMSE is most sensitive to the new snow grain size and albedo settings, while variance in model performance based on snow density profiles is dominated solely by albedo settings (fig. 3.13a and 3.13c). It makes sense that the new snow grain size (45% of first order variance) and albedo (23% of variance) settings are important at 37V, as logically the new snow grain size has a fundamental effect on the grain sizes encountered in the top layers of the snowpack - the layers which affect modelled T_b. Albedo determines how much energy gets absorbed by the snowpack which in turn is available for melt and melt is known to have drastic effects on both snowpack density and grain size. For density the new snow grain size is not important, instead density profile RMSE is dominated almost solely by the settings that govern *SNOWPACK* albedo (86% excl. interaction effects). This can at least in part be explained by the different depth ranges considered to compute model performance based on 37V T_b on the one hand and firn density profiles on the other: 37V T_b only uses the top 4 m of *SNOWPACK* output and is most sensitive to the very top snowpack layers, while the density profiles cover a depth range of 0 - 15 m. It is expected that the new snow grain size setting has a big effect on grain sizes encountered in the top layers (and thus the *SNOWPACK* physical properties of these layers), while grain sizes in deeper layers are mainly controlled by melt-refreeze cycles (in turn affected by how much energy is absorbed by the snowpack through *SNOWPACK*'s albedo settings). Indeed, when comparing sensitivity indices based on higher frequency 37 GHz to lower frequency 19 GHz model performance (fig. 3.13a and 3.13b) one can see that albedo settings become comparatively more important at lower frequencies (from 0.23 to 0.44% first order effect for 37V and 19V RMSE respectively). 19 GHz T_b is sensitive to snow properties of deeper layers as lower frequency microwave radiation is able to propagate further through the snowpack.

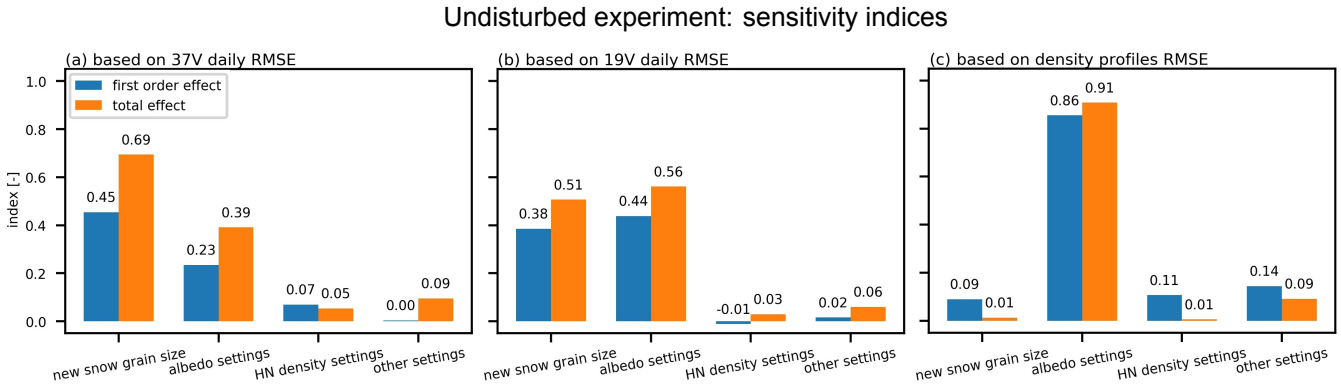


Figure 3.13: *Estimates for the undisturbed experiment first and total order sensitivity indices show that at 37V (**a**) model performance (RMSE) is most sensitive to new snow grain size then to settings related to albedo. At 19V (**b**) albedo settings become more important for model performance than new snow grain size, based on firn density profiles finally performance is dominated by albedo settings (86% of first order variance explained). Estimates are based on a sample size of 250.*

3.4. Uncertainty in Sensitivity Index Estimates

Estimates for sensitivity indices based on yearly 37V time series show that estimates vary depending on which year is considered (fig. 3.14) and give insight into the uncertainty associated with the earlier reported estimates. Although splitting 1998-2017 time series up in 19 yearly chunks means that no data is used more than once, estimates are still correlated: the 1999-2000 time series will begin with the snow and firn state where 1998-1999 left off, and so on. However, given the low sensitivity to initial profile disturbances it seems reasonable to assume they are not in order to get an idea of sensitivity index estimate uncertainty. While results vary depending on which year is considered (boxplots report the median and interquartile ranges of yearly estimates), the general order and relative magnitudes appear reasonably well established. In order to limit the variance in estimates experiments with a sample size greater than 250 are recommended (here limited due to computational costs). Please not that sensitivity indices are not normalized and results of the undisturbed experiment (fig. 3.14c) should not be compared with those of the main and meteo experiments (fig. 3.14a and 3.14b). Also worth noting are the estimates for the outgoing shortwave radiation total order sensitivity index which were all found to be equal to zero. Closer inspection revealed that for a small part of the meteo experiment

ensemble members the outgoing shortwave radiation disturbance was not introduced correctly and this result (S_{Ti} for Q_{so} = 0) should be disregarded, also in fig. 3.8.

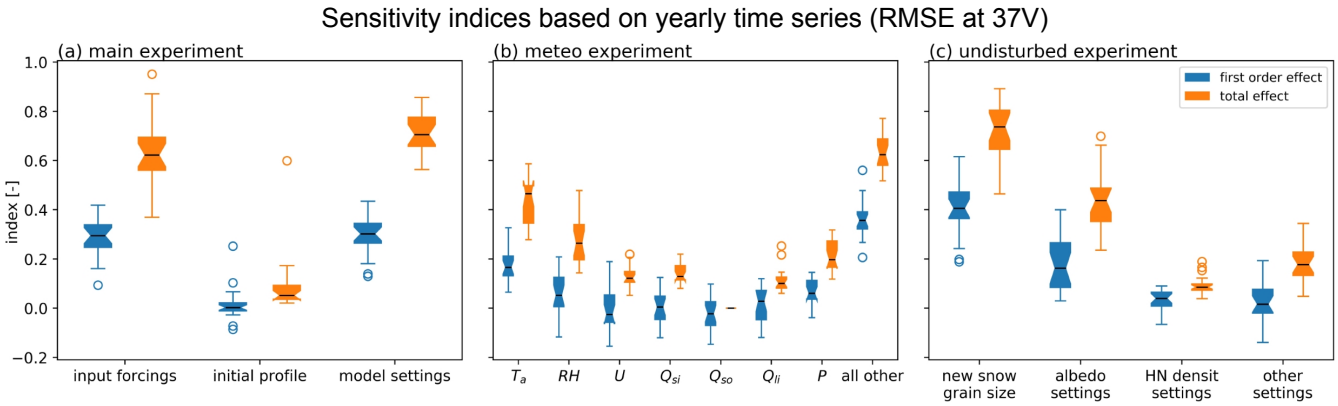


Figure 3.14: *Sensitivity index estimates vary depending on which yearly chunk of the original time series is considered (here for RMSE between modelled and observed time series at 37V for the main (**a**), meteo (**b**) and undisturbed (**c**) experiments) but their relative order of magnitude appears well established. Note: indices are not normalized and only the first order sensitivity indices of the main and meteo experiments are comparable (others should be compared within their respective experiment only). The Q_{so} total order sensitivity index was found to be erroneous and should be disregarded.*

3.5. Effect of Individual Parameters on T_b

Comparison of the 100 most positively and 100 most negatively disturbed meteo experiment ensemble members show that the air temperature disturbance has an impact on the opposite slope effect: warmer ensemble members show visibly more opposite slope compared to colder ones that fail to adequately replicate melt events (fig. 3.15a). When making the same comparison in terms of precipitation disturbance (fig. 3.15b) it becomes obvious that ensemble members with more snowfall show higher modelled T_b values. This makes sense as more snowfall means that old, large-grained snow layers (low T_b) are more effectively covered by fresh small-grained snow layers (high T_b). Lower precipitation ensemble members also show greater spread in modelled T_b (as the ensemble spread in top snowpack properties at the end of the ablation season is preserved rather than buried under fresh snow), but regarding the opposite slope effect no differences can be observed when comparing high to low precipitation time series. When considering parameters with lower sensitivity indices such as relative humidity (3.15c, others not shown) differences in modelled time series become too nuanced to make meaningful conclusions based on an ensemble where other parameters are disturbed simultaneously. New snow grain size was identified as the most important *SNOWPACK* setting for model performance at 37V and fig. 3.15d confirms a huge T_b offset between the 100 meteo experiment ensemble members with the greatest new snow grain size and the 100 members with the smallest. A greater new snow grain size generally means a colder modelled T_b time series that is more sensitive to individual precipitation events (instead T_b signals from ensemble members with a small new snow grain size appear much smoother).

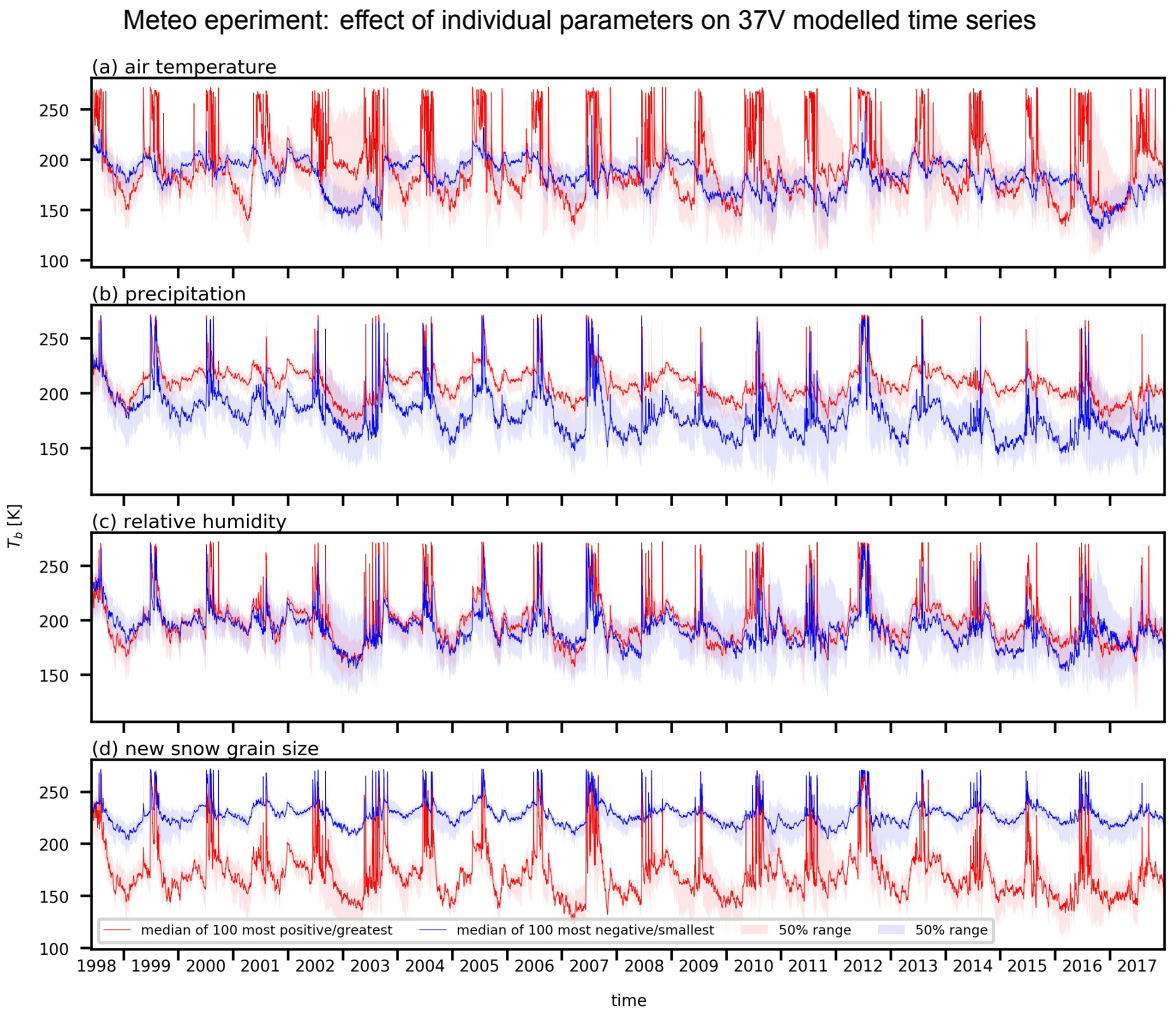


Figure 3.15: *Median time series and 50% range of the 100 most positive and 100 most negative disturbed meteo experiment ensemble members for air temperature (**a**), precipitation (**b**) and relative humidity (**c**). Colder (negative air temperature disturbance) ensemble members show less opposite slope effect but fail to replicate melt events. Great influence of the SNOWPACK new snow grain size setting visualized by comparing greatest new snow grain size ensemble members with smallest (**d**).*

Analysis of the undisturbed experiment ensemble's model performance reveals a clear relationship between an ensemble member's new snow grain size or albedo setting and RMSE between modelled and observed T_b at 37V (fig. 3.16). New snow grain size appears optimal at just over 0.2 mm, but this comes with a disclaimer: this is for a stickiness parameter of 0.25. This stickiness parameter was assumed 0.25 to minimise the offset between the modelled ensemble and 37 GHz observations. Given the great offsetting effect of the new snow grain size (fig. 3.15d) it is safe to assume that if a different input distribution was assumed for this *SNOWPACK* setting (e.g. favouring larger new snow grains, lowering the ensemble's T_b) a different stickiness parameter would have been selected and the optimal new snow grain size would shift accordingly to compensate. The albedo setting is the only discrete snow model setting for which a clear relationship with 37V RMSE can be observed: ensemble members using a parameterized albedo scheme (median RMSE of 24.7 K) perform better when compared to members using constant (26.1 K) or measured (28.7 K) albedo. It is remarkable that a simple solution like a fixed albedo value performs better than using albedo based on outgoing shortwave radiation as measured by the AWS. The conlusion is that the measurements are not accurate, which is not hard to imagine: these instruments experience harsh conditions and little maintenance, measure localised albedo which might not be representative for the albedo when taken over the general area at DYE-2 and have to deal with very low solar incidence angles due to the high geographical latitude. All of the considered albedo parameterizations showed similar performance - no winners or losers could be identified.

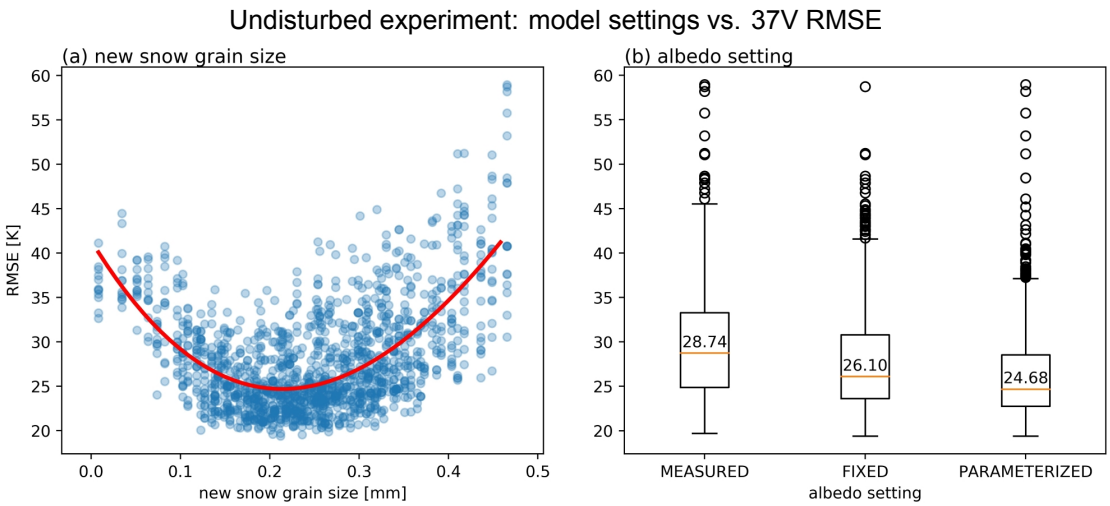


Figure 3.16: *3rd degree polynomial fitted to the new snow grain size value vs. RMSE at 37V (**a**) shows an optimum new snow grain size of just over 0.2 mm (undisturbed experiment at a stickiness of 0.25). In (**b**): using measured albedo from the atmospheric forcing record (SW_MODE = BOTH) generally leads to poorer model performance when compared to using a SNOWPACK albedo paremeterization or even a simple fixed albedo value (SW_MODE = INCOMING).*

3.6. Inverting for the Stickiness Parameter

Fig. 3.17a shows the effect that varying the stickiness parameter has on the modelled T_b series for a single ensemble member of the main experiment. Time series appear offset from each other, with a lower stickiness parameter (higher *stickiness*) corresponding to lower T_b. For one season signals appear closer to each other (data points 1500-2000), this corresponds to the 2002-2003 accumulation season with exceptionally little precipitation (high average densities in the upper snowpack layers). When inverting for the stickiness that best matches the satellite signal (fig. 3.17b) a seasonal cycle becomes apparent: after every melt season the required stickiness parameter starts low (highly sticky snow) to then gradually grow over the course of the accumulation season (evolving towards a non-sticky situation). It is speculated that this is related to fresh snow accumulating on top of the old, melt affected snow layers. During actual melt episodes a very high stickiness parameter seems to be suggested, although closer inspection reveals that at these times the modelled time series using variable stickiness appear very close to each other resulting in a noisy signal for the required stickiness parameter. In contrast, over most accumulation seasons the signal looks smooth and relatively free of noise (2002-2003 being an an exception again). By jumping from one time series to another according to the stickiness found in fig. 3.17b it is possible to construct a composite time series that closely matches the observed signal nearly everywhere and shows no opposite slope effect (fig. 3.17c). One should of course keep in mind that in this situation the stickiness is compensating for any other shortcomings in order to maximally match the observed signal, but given that the required stickiness parameter signal looks so smooth for most years the match between composite and satellite signal is still remarkable.

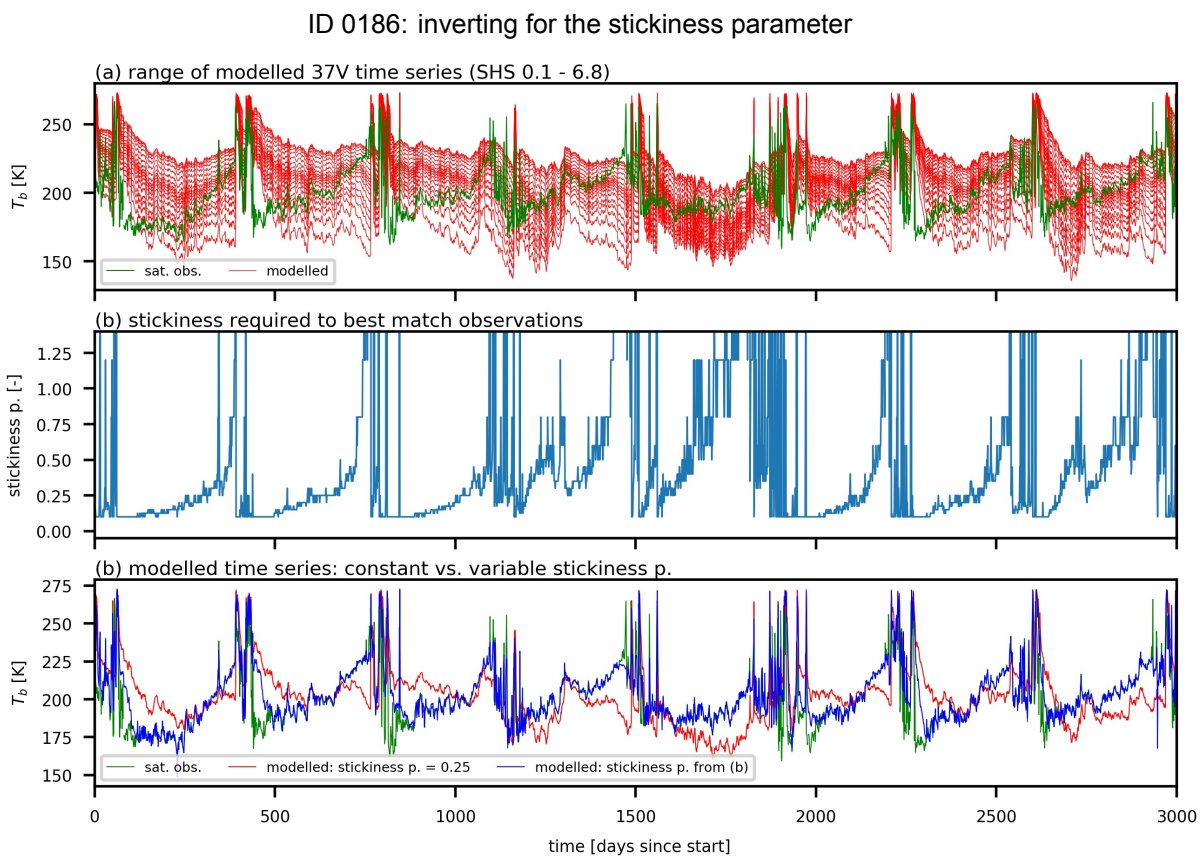


Figure 3.17: *Main experiment's ensemble member ID 0186 modelled for a range of stickiness parameter values (**a**, lowest stickiness parameter corresponds to lowest T_b time series and vice versa) and the stickiness parameter that leads to the best match with the observed signal at 37V (**b**) indicates a seasonal stickiness cycle. Over the course of an accumulation season the stickiness parameter needs to gradually increase (corresponding to a decrease in stickiness) in order to match observations. In **c**: discrete combination of the "best" time series using the stickiness from **b**. The composite time series is mostly able to match the observed signal, showing no opposite slope effect.*

4

Discussion

4.1. Origin of the Opposite Slope Effect

The most striking observation coming from this study is the opposite slope effect: the trends in modelled T_b time series consistently underestimate those of observations when taken over an accumulation season, revealing inadequacies in the modelling chain. What is the cause of this effect? Here it is argued that the constant-with-time stickiness assumption is the culprit. To discuss this it is helpful to break down the T_b parameter. On the most basic level T_b is equal to the product of a snowpack's actual temperature and its emissivity, with emissivity representing the effectiveness with which the snowpack emits thermal radiation. Temperature speaks for itself, emissivity is a function of a snowpack's density, liquid water content and (within the SHS microstructure model) grain size as well as the stickiness parameter.

Possible seasonal biases in modelled snow density, temperature or liquid water content are the easiest to disregard as main drivers behind the opposite slope effect. Modelled snow density profiles were compared with in situ observations from firn cores and general trends were found to agree reasonably well. As atmospheric temperatures are cold over the accumulation season no liquid water is expected to be present at those times, definitely not in the upper layers of the snowpack where heat is quickly lost to the atmosphere. In terms of temperature I argue that the discrepancy in modelled and observed T_b slopes is too great to be caused by biases in modelled snow temperatures alone. When using their mean values to align modelled and observed T_b series on a per accumulation season basis (as was effectively done for the fitted slopes in e.g. fig. 3.12) the mismatch in modelled and observed T_b can easily exceed 10 K at the beginning and end of the accumulation season (and often much more). If this were to originate from biases in snow temperature it is impossible to imagine that *SNOWPACK* would be capable of modelling melt events with any form of accuracy. From comparing the timing of melt events in the modelled and observed T_b time series (easily recognisable by the liquid water induced step increase in T_b) it is known that *SNOWPACK* did model the timing with reasonable accuracy, giving confidence in the snow temperature modelled for the top snowpack layers.

It is difficult to say how realistic the modelled grain size evolution is for the conditions at DYE-2. Over the course of the simulation period grain sizes grow large in *SNOWPACK* which models seemingly unbridled grain growth, in particular at depth (fig. B.1 and B.2). By the end of the undisturbed experiment it is common to encounter grain radii of 2 mm for the top 4 m and 3.5 mm for the top 15 m of snow model output (sometimes greater) and at that point grain growth shows no signs of stopping. *SNOWPACK* was originally developed for seasonal snowpacks and grain growth was validated using laboratory data of up to 40 days (Lehning et al., 2002b) but it is unclear how realistic its growth mechanisms are for perennial firn conditions. I was unable to find suitable in situ measurements of grain size profiles from the area at DYE-2 and even if they were to come available comparison is predicted to be difficult. Given the complex nature of natural snow grain size measurements are quantified by a range of dimensions which are often subjective or the link to a straightforward grain radius representation (as in *SNOWPACK*) is not clearly defined (Linow et al., 2012), hampering quantitative comparison. It is however well known that on the ice sheets grain size generally increases with depth (Brucker et al., 2010, Linow et al., 2012) and there is anecdotal evidence that snow grains do indeed grow very large

at DYE-2 (Max Stevens, personal communication, fig. 4.1). Grain size is known to have a great effect on snow emissivity with larger grains corresponding to lower T_b values but even if grain size at depth is overestimated this mechanism fails to explain an opposite slope effect: the impact of older large grained snow is expected to be greatest at the beginning of the accumulation season, before old layers are buried under fresh accumulation. If the older grain sizes are overestimated this would lower T_b observations at the beginning of the accumulation season - counteracting any opposite slope effect.

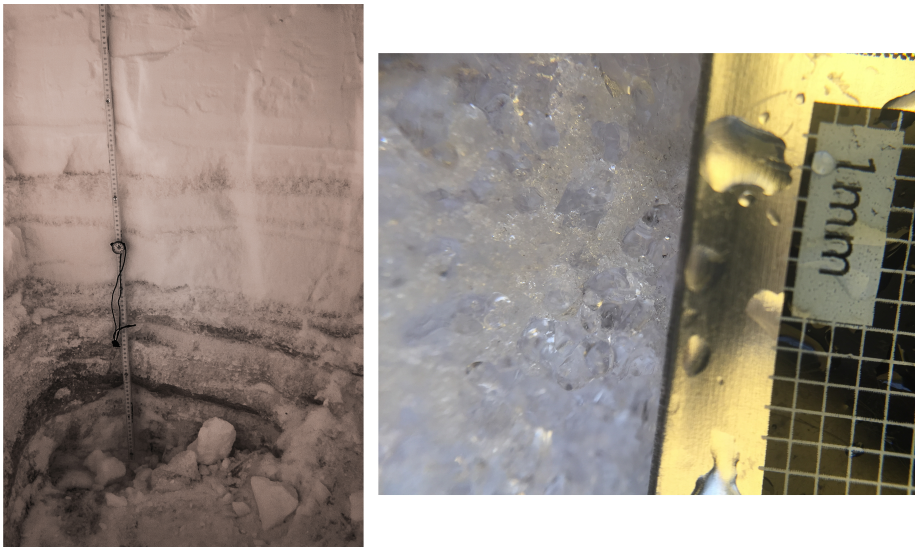


Figure 4.1: *Pictures from a snowpit at DYE-2 showing an increase in grain size with depth and stratigraphy (near-infrared photograph, left) as well as a close up of large, clustered grain sizes living side by side with smaller structures (right). Photo's correspond to the end of the simulation period (spring of 2016 or 2017). Source: Max Stevens, University of Maryland (not published).*

Harder to disregard is the impact that capping of large *SNOWPACK* grain sizes has had on the modelled T_b time series. The largest grain sizes caused the *SMRT* code to fail (likely because they lead to unrealistically large scattering coefficients compared to the absorption coefficient (Picard et al., 2018)) which caused problems especially in the latter half of the simulation period. To keep scatterers small compared to the wavelength and avoid gaps in time series *SNOWPACK* grain radii were capped at 1.5 mm before being used as input for *SMRT*. It is difficult to say exactly what impact this capping on a per-layer basis has had on the characteristics of modelled T_b time series. It is conceivable that capping would artificially lower modelled T_b more so at the beginning of the accumulation season (when old large grained snow near the surface might have to be capped) than at the end (largest grains obscured by layers of fresh snow), possibly favouring an opposite slope effect. However, in the first few years of model output virtually no capping was required (e.g. appendix B for the undisturbed ensemble). As the opposite slope manifests itself right from the start of the modelled series it is concluded that capping of large grain sizes was not the main driver behind the opposite slope effect.

Instead it is argued that the constant-with-time stickiness parameter assumption is is the main cause of the opposite slope effect. Within the SHS microstructure model the stickiness parameter effectively controls the relative position of snow grains without changing their radius or volume fraction (Löwe and Picard, 2015). A high stickiness parameter means well separated, non-clustered snow grains ("not sticky"), a low stickiness parameter means a highly clustered snow medium ("sticky"). This is relevant within the context of radiative transfer because clustered snow grains will scatter microwave radiation more effectively compared to well separated grains of the same size, impacting emissivity and thus the T_b signal. As the difference in T_b caused by a certain change in stickiness is greater for larger grain sizes (fig. 4.2) T_b at DYE-2 (with its large grain sizes, modelled and confirmed in situ) is suspected of being particularly sensitive to this effect. Under natural circumstances melt (as well as pressure from snowpack weight) causes sintering of snow particles and old snow that has experienced one or more melt-refreeze cycles is expected to appear more clustered (sticky, low stickiness parameter) than freshly fallen snow with well separated grains (not sticky, high stickiness parameter) (Colbeck, 1997). The proposed theory behind the opposite slope effect is illustrated using the grain size vs.

stickiness plane of fig. 4.3 and goes as follows: at the beginning of the accumulation season the top layers of a natural snowpack consist of large grained highly clustered particles left behind after the most recent melt-refreeze cycle - this corresponds to very low emissivities. As the season progresses fresh snow with well separated, small grain sizes accumulates on top of these old layers which makes the mean properties of the top snowpack layers gradually shift towards the small grained, low stickiness situation corresponding to very high emissivities (double effect: decrease in grain size and increase in stickiness parameter). As a result, observed T_b time series show a decisively positive trend. In contrast, in this study grain size evolution was modelled by *SNOWPACK* but the stickiness (not present in *SNOWPACK*) was assumed constant. Only a single effect (grain size evolution) is captured and the increase in emissivity over the course of the accumulation season is underestimated, giving rise to the mismatch in observed and modelled T_b trends. The results presented in fig. 3.17 attest to this story, showing a seasonal cycle in the stickiness parameter required to match the observed T_b signal and that over the course of an accumulation season the required stickiness decreases (so increase in stickiness *parameter*).

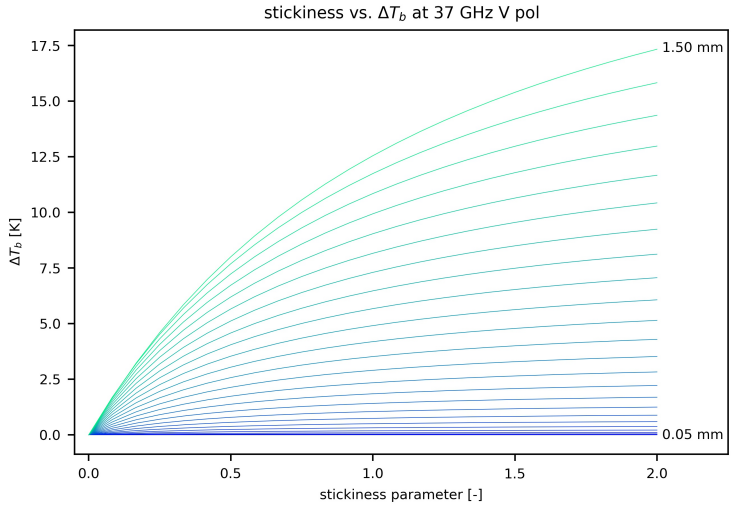


Figure 4.2: *Change in modelled 37 GHz V pol T_b as function of stickiness parameter, for grain radii ranging from 0.05 to 1.50 mm. Stickiness effect is greatest for large grain sizes, increase in stickiness parameter corresponds to decrease in stickiness. Density and temperature constant at 700 kg/m^3 and 260 K.*

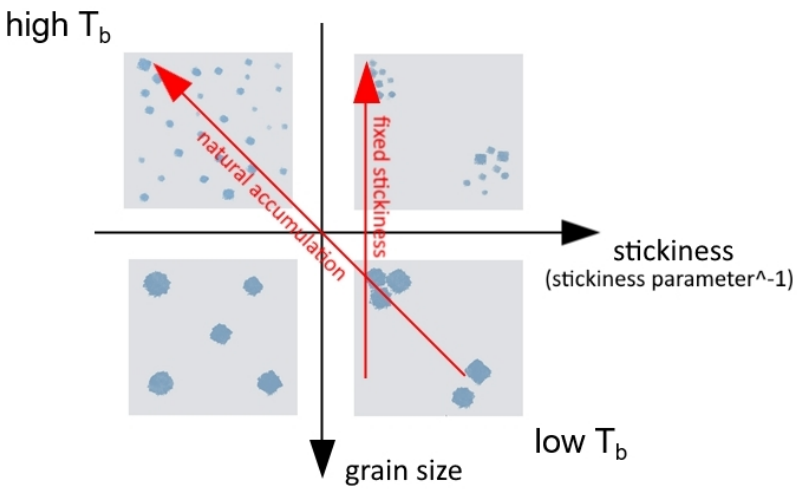


Figure 4.3: *Figure illustrating how stickiness and grain size in the top few metres of snowpack evolve over the course of an accumulation season, in case of natural circumstances and in case of the constant stickiness parameter assumption. In this figure highest emissivities are found in the top left corner (non-clustered small grains) and lowest in the bottom right (large clustered grains).*

The conclusion to this section is that to use *SNOWPACK* in combintation with *SMRT* and SHS microstructure to model microwave satellite observations from ice sheet areas prone to melt, the stickiness needs to be taken into account explicitly and treated as a variable snow parameter. Even though *SNOWPACK* is a full-complexity snow model, the combined model presented here fails to capture seasonal microstructure evolution essential for modelling T_b observations from these areas.

4.2. Comparison with Earlier Efforts

4.2.1. Stickiness Assumption

Setting appropriate microstructure parameters for microwave radiative transfer modelling from in situ measurements or snow model output is known as a tricky problem and as a result nearly every study uses a different approach. While snow density ("ice volume fraction") and snow specific surface area (m^2/kg, directly relatable to sphere radius) can be measured or predicted by models with little or no ambiguity, studies have shown that these two parameters are not sufficient to fully characterize the snow microstructure for microwave propagation modelling and within the context of the SHS microstructure representation high sensitivity to the third parameter, the stickiness parameter, has been demonstrated (Picard et al., 2018 and fig. 4.2). Stickiness values for natural snow are as of now not known because direct measurements are lacking and quantitative relationships between grain types observed in the field and a corresponding stickiness parameter have not yet been established (Löwe and Picard, 2015). While studying microwave emission from dry Alpine snowpacks Liang et al. (2008) reverted to a fixed stickiness parameter of 0.10 "because it yields 2.8 for the frequency dependence of the extinction coefficient which corresponds to the experimental values (Hallikainen et al., 1987)" but also acknowledged that there were no studies indicating how it varies with depth and that other values can be used if warranted by new information. Picard et al. (2018) noted that simply setting the stickiness to 0.10 does not fully capture the power dependence as it is also influenced by grain size and density. A series of other studies used an infinitely large stickiness parameter, i.e. non-sticky spheres in combination with a grain size scaling factor (Brucker et al., 2010, Dupont et al., 2014, Picard et al., 2014a, Roy et al., 2013). The grain size scaling factor was determined empirically by fitting model results to observations. Over-fitting was prevented by applying a single scaling factor per study, determined by fitting observations at different frequencies, polarizations and incidence angles. Picard et al. (2018) showed that the scaling factors found (2.1 to 3.5) are near-equivalent to stickiness parameter values of 0.10 to 0.13 in combination with non-scaled grain sizes. This is at the lower end of stickiness parameter values estimated by (Löwe and Picard, 2015) based on microtomography images of a variety of snow types, who suggested stickiness parameter values of 0.1 to 0.2 (although stickiness parameter values for some individual samples exceeded 0.4).

None of the earlier studies considered time series spanning melt seasons under perennial firn conditions and as far as I am aware this study is the first to demonstrate the need for a variable stickiness with time. The important implication is that under these conditions earlier approaches based on a fixed stickiness parameter or constant grain size scaling (or that lack a third parameter altogether) are not valid when modelling time series spanning multiple seasons. Regarding the constant-with-time stickiness assumption, for this study a fixed stickiness parameter of 0.25 was required at 37 GHz to best match the mean of the modelled ensemble with observations which is a slightly higher value than suggested by earlier studies. Here *SNOWPACK* modelled grain sizes were used as input for *SMRT* directly but *SNOWPACK* also provides optical equivalent grain sizes in its output files, which are smaller. It is speculated that if the optical equivalent grain sizes were used the mean modelled T_b time series would be greater and a lower stickiness parameter would have been selected, potentially in line with earlier reported values.

4.2.2. Modelling of T_b Observations over Ice Sheets

Two earlier studies dealing with the forward modelling of T_b observations from ice sheets are shortly discussed here. Brucker et al. (2011) modelled two year long time series at 18.7 and 36.5 GHz at Dome C in Antarctica's dry snow zone, driven by in situ measurements of snow temperature, density and grain size. Grain size profiles were established by relating near-infrared reflectence of a snow pit wall to specific surface area from stereological measurements based on three different relationships and kept constant throughout the simulation. Using multi-layered DMRT with non-sticky spheres to forward model T_b observations, T_b time series were overestimated for all the considered relationships

(RMSE of ca. 28 K) and grain size scaling was applied to match the satellite observations, after which results improved significantly (RMSE of ca. 1.5 K) although T_b in summer was still overestimated. In the conclusion the authors recommend to use a stickiness factor to consider the various shapes of snow grains. Picard et al. (2009) on the other hand used a physically based snow dynamic and emission model (SDEM) driven by atmospheric forcing (ERA-40 re-analysis) to forward model T_b time series at 19 and 37 GHz over the whole of the Antarctica Ice Sheet over several years. Model parameters were optimized for every 50 by 50 km pixel. The underlying snow model was relatively simple and, for example, assumed a vertically constant and spatially uniform snow density. This proved to perform well in dry snow zones (RMSE between 1.4 and 3.0 K) but failed where melt events complicated the snowpack structure and its temporal variations. Several studies have implemented more sophisticated snow models for seasonal snowpacks (e.g. Kontu et al., 2017, Wiesmann et al., 2000) but due to that context none dealt with multi-year time series where and old melt-affected snowpack is covered by fresh snow over the course of a new accumulation season. Several aspects of the study presented here are new. This includes, the application of a combined full-complexity snow and radiative transfer model to the percolation area of an ice sheet, the ensemble approach of the SA and modelling of continuous multi-decade time series.

4.3. Linking SNOWPACK to Sticky Hard Spheres

In light of the eventual goal of inverting observed microwave observations for the underlying firn properties (fig. 1.4), treating the stickiness parameter as a free variable on a per ensemble member basis is undesired. Solving for the stickiness by matching modelled time series to the observations would defeat the original aim of looking at which ensemble members correspond well to observations and which do not. At the same time it was shown that a model based on a constant stickiness fails to capture essential processes taking place in the snow and firn column. The main recommendation coming from this thesis is to instead link the snow microstructure representation of *SNOWPACK* to the SHS model such that stickiness can be taken directly from the snow model output, or be estimated from it.

In *SNOWPACK* four primary parameters are used to describe the complex texture of snow: grain size, bond size, dendricity and sphericity which are allowed to evolve per snow layer according to rate equations (functions of the local environmental conditions based on both theory and empirical relations, Lehning et al., 2002b). Of particular importance is the bond size, expressed by the bond radius *rb* of the necks that connect snow grains. Within *SNOWPACK* the bond size is allowed to grow not only through metamorphism but also pressure sintering and is the single most important parameter for determining snow viscosity and thermal conductivity in *SNOWPACK*. The SHS microstructure model in *SMRT* on the other hand is defined by grain size and the stickiness parameter and has its roots in molecular fluid theory (Baxter, 1968). The spherical particles interact through hard-core repulsion and surface adhesion and in the equilibrium state of these attractive and repulsive forces microstructures with interesting structural properties arise, with the stickiness parameter (inversely proportional to the contact adhesion) controlling clustering of the spheres. Both microstructure models are a severely simplified representation of reality (required to be able to apply theory), only mimicking the highly complex and variable texture of real snow. Yet at the same time, both models have been used in physical models capable of producing realistic output. This suggests that at least within their respective contexts the models are capable of compensating for microstructure nuances not explicitly represented by additional parameters. Assuming that grain sizes of both models are interchangeable, could there be a relationship between bond size and stickiness?

Not based on any theory, fig. 4.4 indicates there could be, with $1/rb^2$ (mean rb of the top 1 m of *SNOWPACK* output in this case) suggesting correlation with the evolution of the required stickiness parameter of fig. 3.17. Finding a robust methodology of estimating stickiness from parameters readily available in *SNOWPACK* output is left for future work but is considered a promising route to resolve the opposite slope effect. It would not only allow to set a snow model estimated stickiness per date (variable stickiness with time), but also per snow layer (variable stickiness with depth). Such an approach has not been attempted in the literature before and a successful model would be the first of its kind. Alternatively it has been suggested that the SHS representation could be implemented in models as *SNOWPACK* in lieu of their current microstructure implementations (Löwe and Picard, 2015) but considering how integral its microstructure is to *SNOWPACK*'s model physics this is expected to be a very involved undertaking.

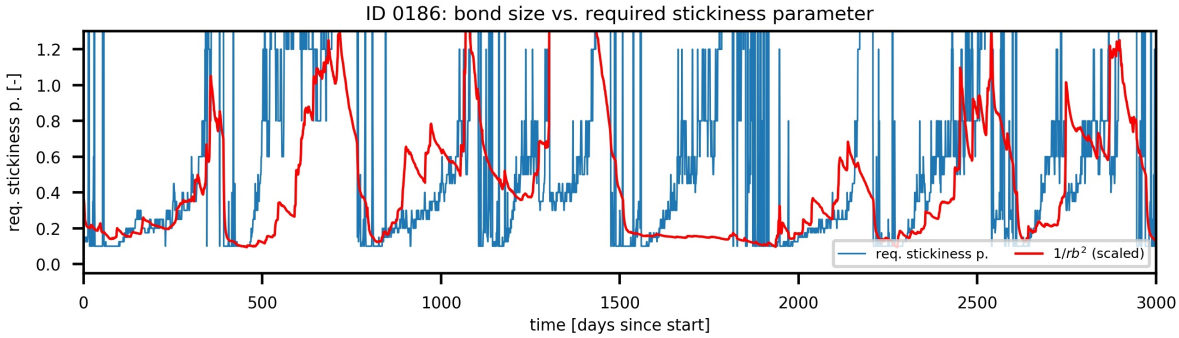


Figure 4.4: $1/rb^2$ *(scaled) plotted with the required stickiness parameter of main experiment's ID 0186 (from fig. 3.17) suggests that bond size, a variable readily available from the SNOWPACK output, might eventually be used to estimate a snow layer's stickiness parameter before input in SMRT (future work).* rb *corresponds to the mean bond radius taken over the upper 1 m of SNOWPACK output. Low stickiness/high bond radius season between datapoints 1500 and 2000 corresponds to the 2002-2003 accunulation season with exceptionally little precipitation (relatively dense and old snowpack).*

4.4. Other Limitations and Recommendations

It is acknowledged that the results presented here and their interpretation are subject to some additional limitations:

- Lack of atmosphere in radiative transfer model: the impact of the atmosphere on T_b observations is expected to be limited at 37 and 19 GHz and given the high RMSE's of modelled time series the lack of an atmosphere was not a limiting factor. *SMRT* allows for the inclusion of a simple atmospheric layer which would be an interesting addition to the model once the opposite slope effect is dealt with.
- Low sample size for SA: the sample size of 250 is relatively small for a Monte Carlo style SA for which sample sizes an order of magnitude larger are common. Computational power was a bottle neck here and for future work it might be interesting to model shorter or sparser time series or to upgrade to a more powerful cluster.
- Model output vs. satellite observation time of day: daily *SNOWPACK* output was collected at the end of the afternoon in an attempt to match the overpass time of DMSP satellite platforms F13 and F19, but satellite swaths overlap considerably near the poles and the T_b value extracted from Maslanik and Stroeve (2018) is made up of an average of multiple observations from the same day. T_b can be sensitive to time of day because of the snowpack's daily temperature evolution although daily temperature fluctuations at high latitudes are expected to be limited due to low solar incidence angles.
- Spatial resolution of imager vs. spatial variations in firn properties: firn properties are known to show lateral heterogeneity, originally created by uneven deposition of snow due to the local wind field and thereafter subsurface conditions continuously evolve where an important contribution to local spatial variations comes from the flow and refreezing of meltwater (Bell et al., 2008, Sturm and Benson, 2004). It has been shown that measured T_b can fluctuate with local properties on the meter-scale even in dry snow zones (Picard et al., 2014b). The satellite observations used here however have a spatial resolution of 25 km and are affected by properties as integrated over the general area rather than the local DYE-2 situation. The influence of lateral variations in firn properties on T_b observations needs investigating.
- Only one location: this study only considered DYE-2 and the question is how representative the results are for the percolation area in general and by extension for other areas of the Greenland Ice Sheet. For example, is less opposite slope modelled for areas in the dry snow zone? Forcing data presented by Vandecrux et al. (2020) means that it would be relatively straighforward to expand the analysis to areas marked by ice slabs, firn aquifers or the dry snow zone.
- Not all *SNOWPACK* model configurations included: over the years different groups have implemented a range of *SNOWPACK* model options and parameterizations and to limit the number

of possible configurations not all available implementations were considered here. One model option of particular interest for future efforts is the water transport scheme based on the Richards Equation as implemented by Wever et al. (2014). Here it was excluded (only the standard bucket scheme was used) because of unreliable runtimes, a consequence of the variable time step implementation.

- Only one *SMRT* configuration: *SMRT* facilitates easy experimentation with different microstructure representations and electromagnetic theories. For the situation at DYE-2 however large grain sizes posed a problem, limiting the options which produced results reliably. Nevertheless, the current configuration (IBA-SHS) allows for changing the electromagnetic theory or microstructure representation independently from the other and further experimentation with this is recommended, as well as the modelling of observations at additional frequencies. Ideally the theories available in *SMRT* get expanded upon so that in the future large scatterers (compared to the wavelength) no longer pose a problem.

4.5. Added Value of T_b Time Series

Estimated sensitivity indices showed that model performance based on T_b time series can be sensitive to different input parameters and model settings compared to performance based on in situ measurements. Compared to density profiles (the only in situ measurement considered in this study) this was most evident from the sensitivity indices of the undisturbed experiment where model performance at 37V was mostly controlled by the new snow grain size (45% of variance explained) while the density profile metric was dominated by *SNOWPACK*'s albedo settings (86%). This suggests that when the goal is to maximally constrain an ensemble of possible but uncertain *SNOWPACK* realisations and both density profile and T_b observations are available, one would benefit from considering both. Because of the sensitivity to different parameters the subset of likely candidates is expected to be different depending on the validation metric considered, but true candidates can of course only come from the overlap of the two subsets leading to a further reduction.

Additionally and importantly, as discussed in the introduction to this report in situ measurements are not nearly available everywhere on the ice sheets and where they are measurements tend to have poor temporal resolution and often a bias towards being sampled at the end of the accumulation season (all the density profiles considered here came from May, for example). The true value of the microwave remote sensing observations lies in the fact that they are available over the entire ice sheet and on a daily basis, year-round. The cost is that while in situ measurements can be very explicit the microwave observations are hard to interpret, given the coarse resolution and layers of uncertainty added by modelling (as again demonstrated here). In order to fully open the door to satellite based retrieval of subsurface firn properties and processes from areas where observations are currently lacking, work on a combined snow and radiative transfer model should be continued.

5

Conclusion

In a first step towards inverse firn retrieval from microwave remote sensing observations this study presented a new forward model which given initial conditions and atmospheric forcing first solves for the firn state through full-complexity snow model *SNOWPACK* and then simulates multifrequency T_b time series using radiative transfer model *SMRT*. As part of a comprehensive sensitivity analysis three ensembles of multi-decade T_b time series (19 and 37 GHz) were modelled for the DYE-2 site in the percolation area of the Greenland Ice Sheet and compared to independent satellite observations.

> **RQ1** - What is the impact of the different *SNOWPACK* input variables and model settings on modelled T_b time series?

Model performance (RMSE w.r.t. daily satellite observations at 37V) was sensitive to errors introduced in the atmospheric forcing record (with air temperature, precipitation and relative humidity directly controlling 21, 11 and 9% of the variance) and model settings (33%) but not initial firn conditions. In an undisturbed experiment new snow grain size and albedo settings were shown to be the most important *SNOWPACK* settings, directly explaining 45 and 23% of variance in model performance at 37V. At 19V however the albedo settings (44%) surpassed the new snow grain size (38%), explained by 19 and 37 GHz observations being sensitive to different snowpack depth ranges. Variables showing the highest sensitivity indices were found to have a big impact on the mean modelled T_b value or modelled T_b trends over the accumulation seasons. *SNOWPACK* parameterized albedo schemes generally performed better than albedo as measured by an automatic weather station.

> **RQ2** - How do *SNOWPACK-SMRT* modelled T_b time series compare to independent satellite observations?

The most striking observation was that modelled T_b consistently underestimated observed trends when taken over the course of an accumulation season, dubbed the “opposite slope effect”. It was shown that this can only be explained by the constant-with-time stickiness assumption used to link *SNOWPACK*’s microstructure representation to the SHS model used with *SMRT*, as a seasonal stickiness signal was made evident for the conditions at DYE-2 and linked to its yearly melt-refreeze-accumulation cycle. As a result RMSEs were high (min. 17.8 K at 37V and 19.4 K at 19V) although individual T_b events were well captured in the 37 GHz modelled series. 19 GHz ensembles lacked such detail, appearing smooth next to and overestimating the satellite signal (because of a lower stickiness parameter limit of 0.1) but still capturing the timing of melt events.

> **RQ3** - What is the added value of using T_b time series when constraining *SNOWPACK* ensembles?

Sensitivity indices of model performance based on in situ density profile observations were shown to differ to those based on T_b (with T_b generally more sensitive to microstructure development in the upper snowpack layers and density to the net energy absorbed) suggesting additional constraining power when both in situ and T_b observations are available. However, in situ observations are scarce

and the real power from microwave satellite observations comes from the availability of daily and year-round data giving a synoptic and historical overview of both ice sheets. Coarse spatial resolution vs. spatial variations in firn properties as well as the layers of uncertainty added by modelling aggravate interpretation.

Future Work

The link between the opposite slope effect and a seasonal stickiness signal at DYE-2 implies that earlier approaches to forward modelling microwave satellite observations based on a constant-with-time stickiness parameter (implicitly or explicitly) are not valid for ice sheet areas prone to melt-refreeze cycles. This study is expected to be the starting point for a more sophisticated implementation of the combined model that estimates a snow layer's stickiness from microstructure information already present in the snow model, and *SNOWPACK*'s bond size was presented as an interesting place to start. If successful such a model would be the first of its kind and open the door to satellite-based retrieval of subsurface firn properties and processes from areas where observations are currently lacking, greatly reducing uncertainty in ice sheet mass loss and global sea level rise projections.

A

SNOWPACK Settings

A.1. SNOWPACK Configuration

For this study *SNOWPACK* version 3.50 was used in “default” variant, compiled in conjunction with Libsnowpack 3.50 and MeteoIO 2.80. Fig. A.1 shows a *SNOWPACK* produced *-.ini* file containing assumed model settings. Worth noting is that THRESH_RAIN was set to 350 K to force solid precipitation (the only precipitation measured in the atmospheric forcing record).

Example -.ini file containing SNOWPACK configuration

```
[GENERAL]
BUFF_BEFORE = 1.5
BUFF_CHUNK_SIZE = 370

[GENERATORS]
TSG::CST::VALUE = 255.77
TSG::GENERATORS = CST

[INPUT]
COORDSYS = LOCAL
ISWR_IS_NET = false
METEO = SMET
METEOPATH = /net/labdata/dylan/Snowpack/*
NUMBER_OF_SOLUTES = 0
SNOW = SMET
SNOWFILE1 = DYE-2_1998_*
SNOWPATH = /net/labdata/dylan/Snowpack/*
SOLUTE_NAMES = NITRATE
STATION1 = DYE-2_*
TIME_ZONE = 0.0

[INTERPOLATIONS1D]
PSUM::ACCUMULATE::PERIOD = 900
PSUM::RESAMPLE = accumulate

[OUTPUT]
AGGREGATE_PRF = false
AGGREGATE_PRO = false
AVGSUM_TIME_SERIES = true
BACKUP_DAYS_BETWEEN = 100000
CLASSIFY_PROFILE = false
COORDSYS = LOCAL
CUMSUM_MASS = false
EXPERIMENT = *
FIRST_BACKUP = 100000
HARDNESS_IN_NEWTON = false
HAZARD_STEPS_BETWEEN = 2
METEO = SMET
METEOPATH = /net/labdata/dylan/Snowpack/*
OUT_CANOPY = false
OUT_CANPOY = false
OUT_HAZ = false
OUT_HEAT = false
OUT_LOAD = false
OUT_LW = false
OUT_MASS = false
OUT_METEO = false
OUT_SOILEB = false
OUT_STAB = false
OUT_SW = false
OUT_T = false
PRECIP_RATES = true
PROF_DAYS_BETWEEN = 1
PROF_FORMAT = PRO
PROF_START = 0
PROF_WRITE = true
SNOW = SMET
SNOW_WRITE = true
TIME_ZONE = 0.0
TS_DAYS_BETWEEN = 1
TS_FORMAT = MET
TS_START = 0
TS_WRITE = false
WRITE_PROCESSED_METEO = false

[SNOWPACK]
ATMOSPHERIC_STABILITY = *
CALCULATION_STEP_LENGTH = 15.0
CANOPY = false
CHANGE_BC = false
ENFORCE_MEASURED_SNOW_HEIGHTS = false
FORCING = ATMOS
HEIGHT_OF_METEO_VALUES = 2
HEIGHT_OF_WIND_VALUE = 10
MEAS_TSS = false
ROUGHNESS_LENGTH = *
SNP_SOIL = false
SW_MODE = *

[SNOWPACKADVANCED]
ADJUST_HEIGHT_OF_METEO_VALUES = false
ADJUST_HEIGHT_OF_WIND_VALUE = false
ADVECTIVE_HEAT = false
ALBEDO_AGING = *
ALBEDO_AVERAGE_SCHMUCKI = ALL_DATA
ALBEDO_FIXEDVALUE = *
ALBEDO_PARAMETERIZATION = *
ALLOW_ADAPTIVE_TIMESTEPPING = true
ALPINE3D = false
ALPINE3D_PTS = false
AVG_METHOD_HYDRAULIC_CONDUCTIVITY = ARITHMETICMEAN
CANOPY_HEAT_MASS = true
CANOPY_TRANSMISSION = true
COMBINE_ELEMENTS = true
DETECT_GRASS = false
ENABLE_VAPOUR_TRANSPORT = *
FORCE_RH_WATER = true
FORCE_SW_MODE = false
FORESTFLOOR_ALB = true
HARDNESS_PARAMETERIZATION = MONTI
HEAT_BEGIN = 0.0
HEAT_END = 0.0
HEIGHT_NEW_ELEM = 0.005
HN_DENSITY = *
HN_DENSITY_FIXEDVALUE = *
HN_DENSITY_PARAMETERIZATION = *
HOAR_DENSITY_BURIED = 125.
HOAR_DENSITY_SURF = 100.
HOAR_MIN_SIZE_BURIED = 2.
HOAR_MIN_SIZE_SURF = 0.5
HOAR_THRESH_RH = 0.97
HOAR_THRESH_TA = 1.2
HOAR_THRESH_VW = 3.5
JAM = false
LB_COND_WATERFLUX = FREEDRAINAGE
MASS_BALANCE = false
MAX_NUMBER_MEAS_TEMPERATURES = 5
MEAS_INCOMING_LONGWAVE = false
METAMORPHISM_MODEL = DEFAULT
MINIMUM_L_ELEMENT = 0.0025
MIN_DEPTH_SUBSURF = 0.07
MULTI_LAYER_SK38 = false
NEW_SNOW_GRAIN_SIZE = *
NUMBER_FIXED_RATES = 0
NUMBER_SLOPES = 1
PERP_TO_SLOPE = false
PLASTIC = false
PREVAILING_WIND_DIR = 0.
REDUCE_N_ELEMENTS = false
RESEARCH = true
SALTATION_MODEL = SORENSEN
SNOW_ALBEDO = *
SNOW_EROSION = false
SNOW_REDISTRIBUTION = false
SOIL_FLUX = false
SSI_IS_RTA = false
STRENGTH_MODEL = DEFAULT
SW_ABSORPTION_SCHEME = *
TEMP_INDEX_DEGREE_DAY = 0.
TEMP_INDEX_SWR_FACTOR = 0.
THRESH_DTEMP_AIR_SNOW = 3.0
THRESH_RAIN = 350
THRESH_RH = 0.5
TWO_LAYER_CANOPY = true
T_CRAZY_MAX = 340.
T_CRAZY_MIN = 100
VARIANT = DEFAULT
VISCOSITY_MODEL = DEFAULT
WATERTRANSPORTMODEL_SNOW = BUCKET
WATERTRANSPORTMODEL_SOIL = BUCKET
WATER_LAYER = false
WIND_SCALING_FACTOR = 1.0
```

Figure A.1: *An example -.ini file illustrates the assumed SNOWPACK configuration. The settings that were varied throughout the experiments (table 2.3) have been marked with an asterisk, leaving the settings that were kept fixed throughout the study.*

A.2. RACMO Albedo Parameterization

The “RACMO” albedo parameterization does not come standard with *SNOWPACK* 3.50 and was added to the program before compiling it. It is based on the broadband parameterization proposed by Gardner and Sharp (2010) and implemented by Kuipers Munneke et al. (2011) in regional climate model RACMO2.1, albeit in a simplified form: for this study cloud optical depth was assumed 0 and soot contamination 0.2 ppm throughout.

A.3. Aerodynamic Roughness Length

Unlike the other continuous model settings (table 2.6), uncertainty in aerodynamic roughness length was not represented directly through a normal distribution but included additional steps. First, samples were taken from the normal distribution around 0 with standard deviation 75 and the absolute value of all samples was taken. For samples greater than 1 the aerodynamic roughness length was taken as 10e-5*sample, while for samples smaller than 1 it was assumed 10e-5/sample. The resulting ensemble of assumed roughness lengths is visualised in fig. A.2 for the main experiment. Roughness lengths for the other two experiments were drawn from the same distribution.

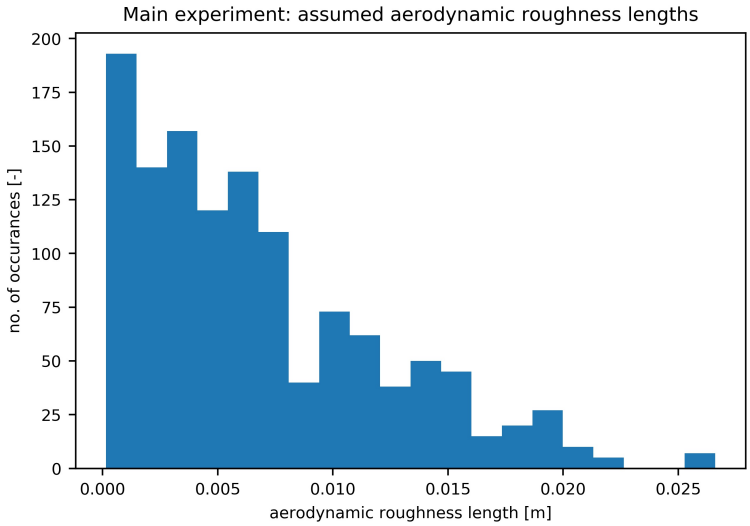


Figure A.2: *Histogram showing the distribution of aerodynamic roughness lengths used as SNOWPACK input for the main experiment. Values for the meteo and undisturbed experiments were drawn from the same distribution.*

B

Layer Merging Routine

Output from *SNOWPACK* was processed before being used as input for *SMRT*: the number of layers was limited to speed up runtimes and grain sizes had to be capped to avoid exceptions. The figures here illustrate the extent to which the number of layers had to be limited or grain radii capped in case of the undisturbed experiment.

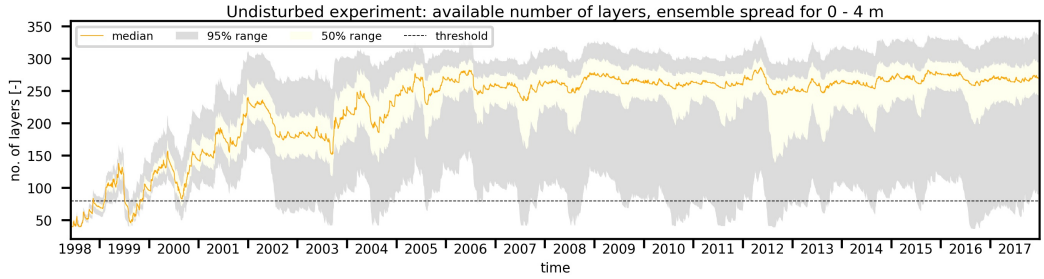


Figure B.1: *Main experiment: ensemble spread of the number of layers present in the top 4 m of SNOWPACK output (relevant for 37 GHz simulations). Number of layers for input in SMRT was limited at 80 layers, excess layers were merged with adjacent layers.*

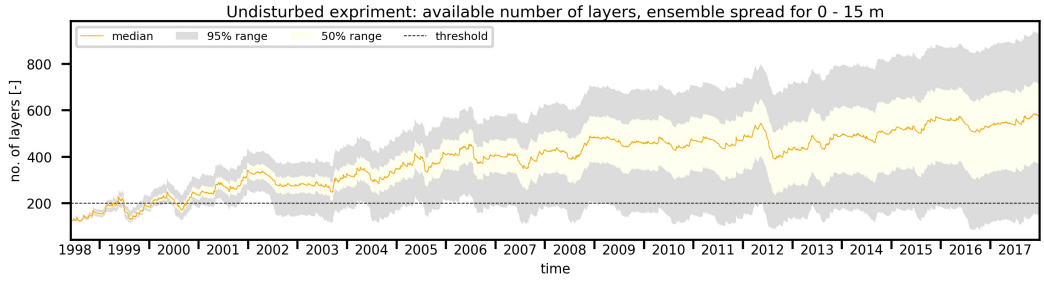


Figure B.2: *Undisturbed experiment: ensemble spread of the number of layers present in the top 4 m of SNOWPACK output (relevant for 37 GHz simulations). Number of layers for input in SMRT was limited at 200 layers, excess layers were merged with adjacent layers.*

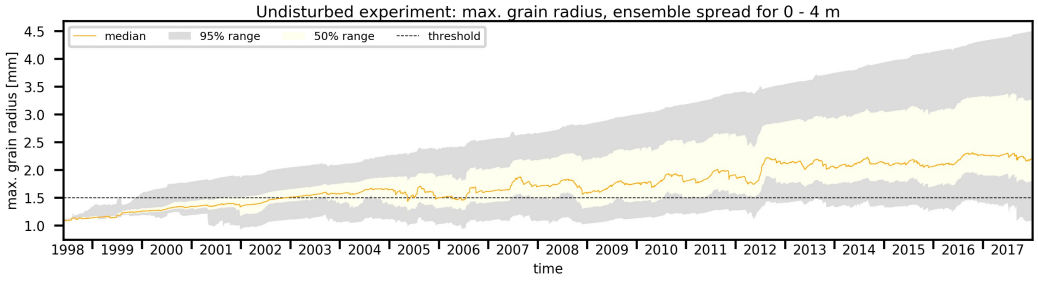


Figure B.3: *Undisturbed experiment: ensemble spread of the maximum grain radius present in the top 4 m of SNOWPACK output (relevant for 37 GHz simulations). Grain radius for input in SMRT was capped at 1.50 mm.*

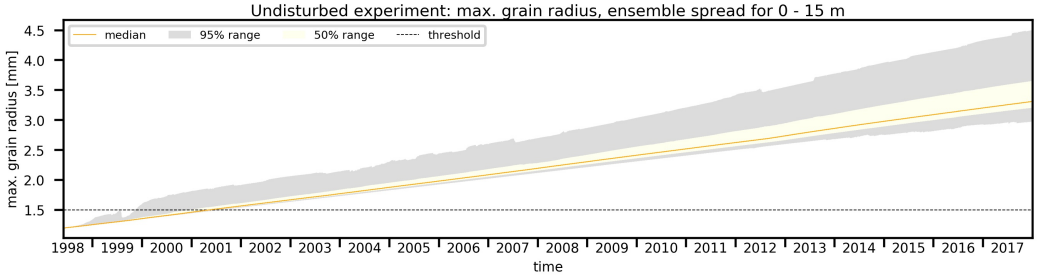


Figure B.4: *Undisturbed experiment: ensemble spread of the maximum grain radius present in the top 15 m of SNOWPACK output (relevant for 19 GHz simulations). Grain radius for input in SMRT was capped at 1.50 mm.*

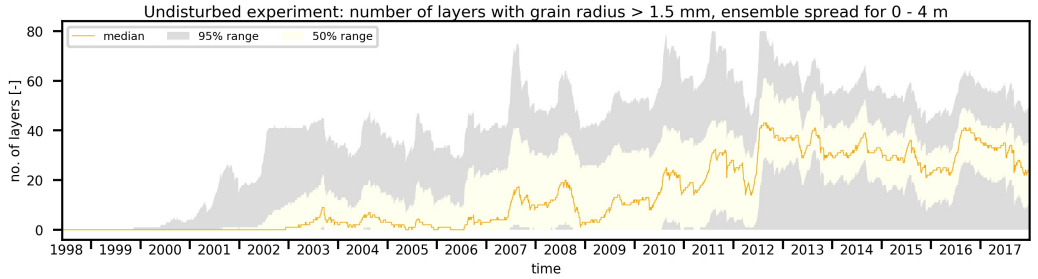


Figure B.5: *Undisturbed experiment: ensemble spread of the number of layers present in the top 4 m of SNOWPACK output (relevant for 37 GHz simulations) for which grain radius exceeded 1.50 mm. Grain radius for input in SMRT was capped at 1.50 mm.*

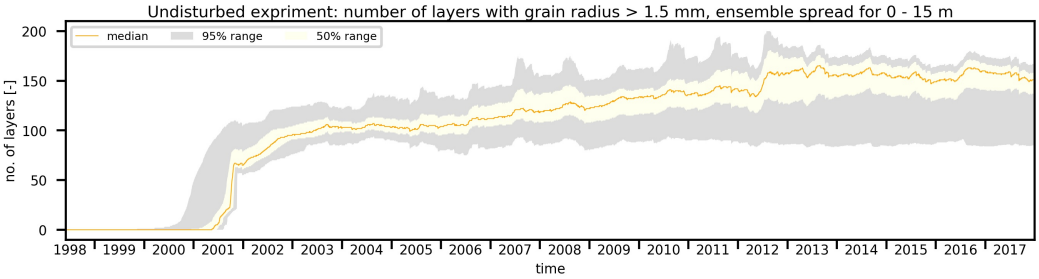


Figure B.6: *Undisturbed experiment: ensemble spread of the number of layers present in the top 15 m of SNOWPACK output (relevant for 19 GHz simulations) for which grain radius exceeded 1.50 mm. Grain radius for input in SMRT was capped at 1.50 mm.*

C

Fitting of Slopes

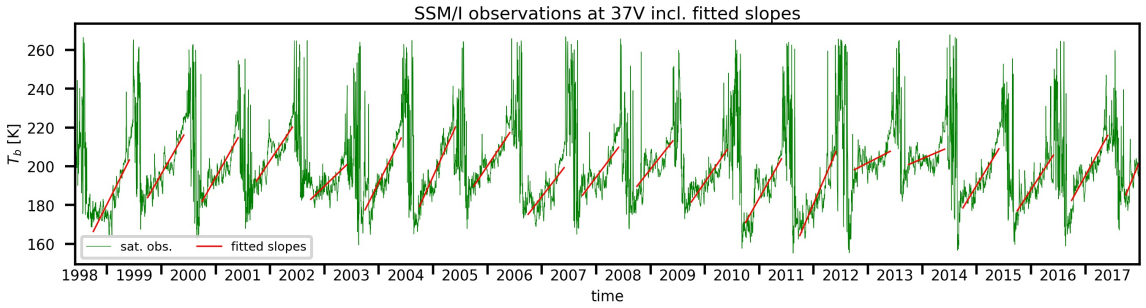


Figure C.1: *Linear segments (red) fitted to the observed T_b signal at 37V, per accumulation season (assumed to run from October up to and including May). Slopes are upward without exception.*

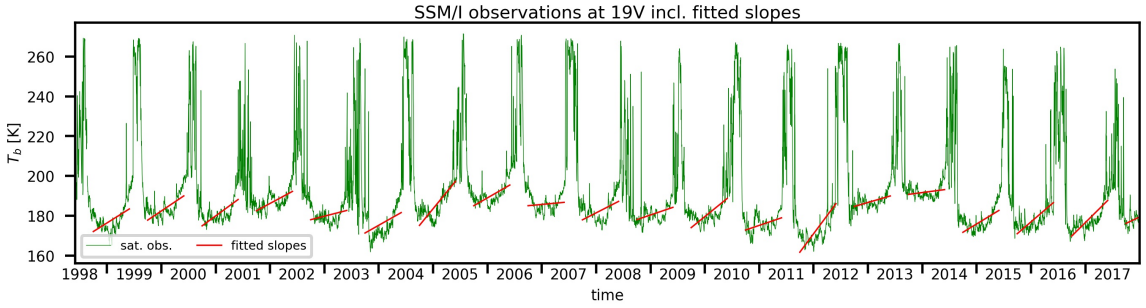


Figure C.2: *Linear segments (red) fitted to the observed T_b signal at 19V, per accumulation season (assumed to run from October up to and including May). Slopes are upward without exception.*

D

Meteo Exeriment Ensemble Spread

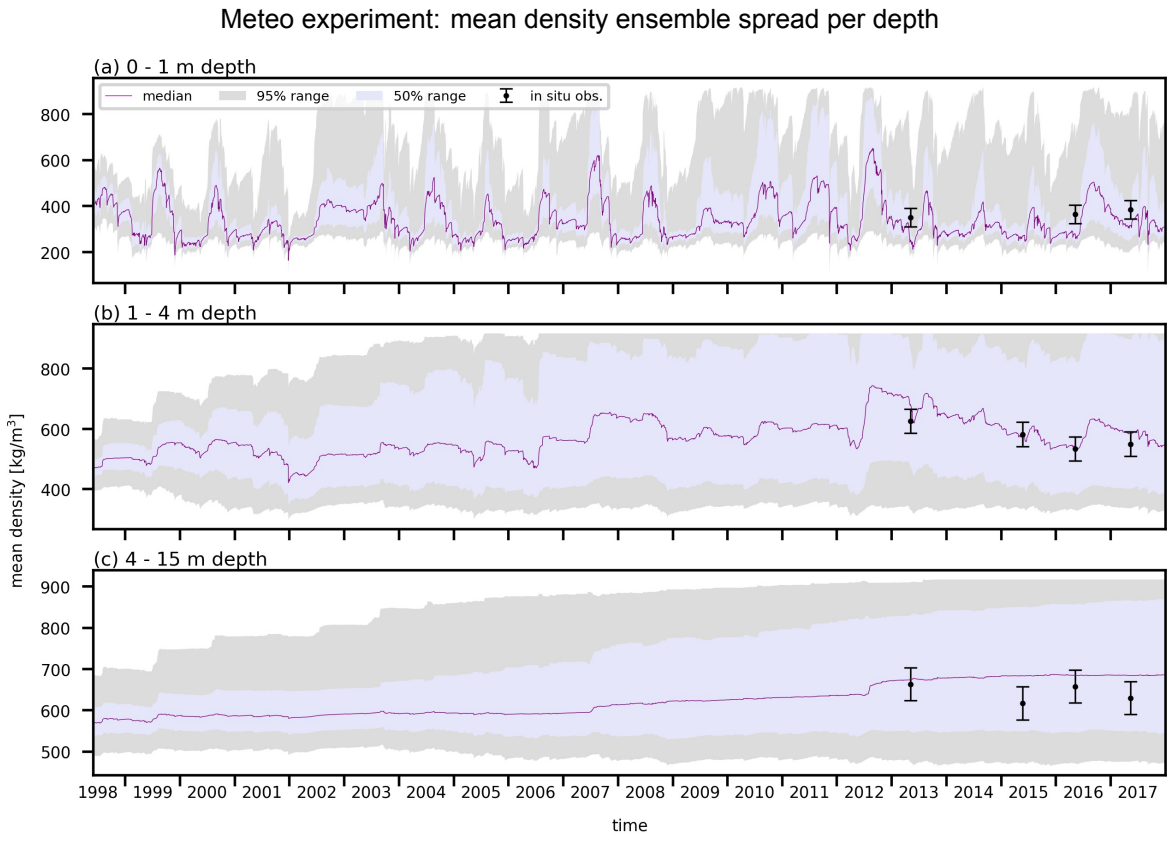


Figure D.1: *Ensemble spread of mean modelled densities for 0 - 1 m (**a**), 1 - 4 m (**b**) and 4 - 15 m (**c**) depth of meteo experiment SNOWPACK output. In-situ observations come from firn cores and include 40 kg/m^3 error bars. The median modelled time series were constructed by taking the ensemble's median value at every data point.*

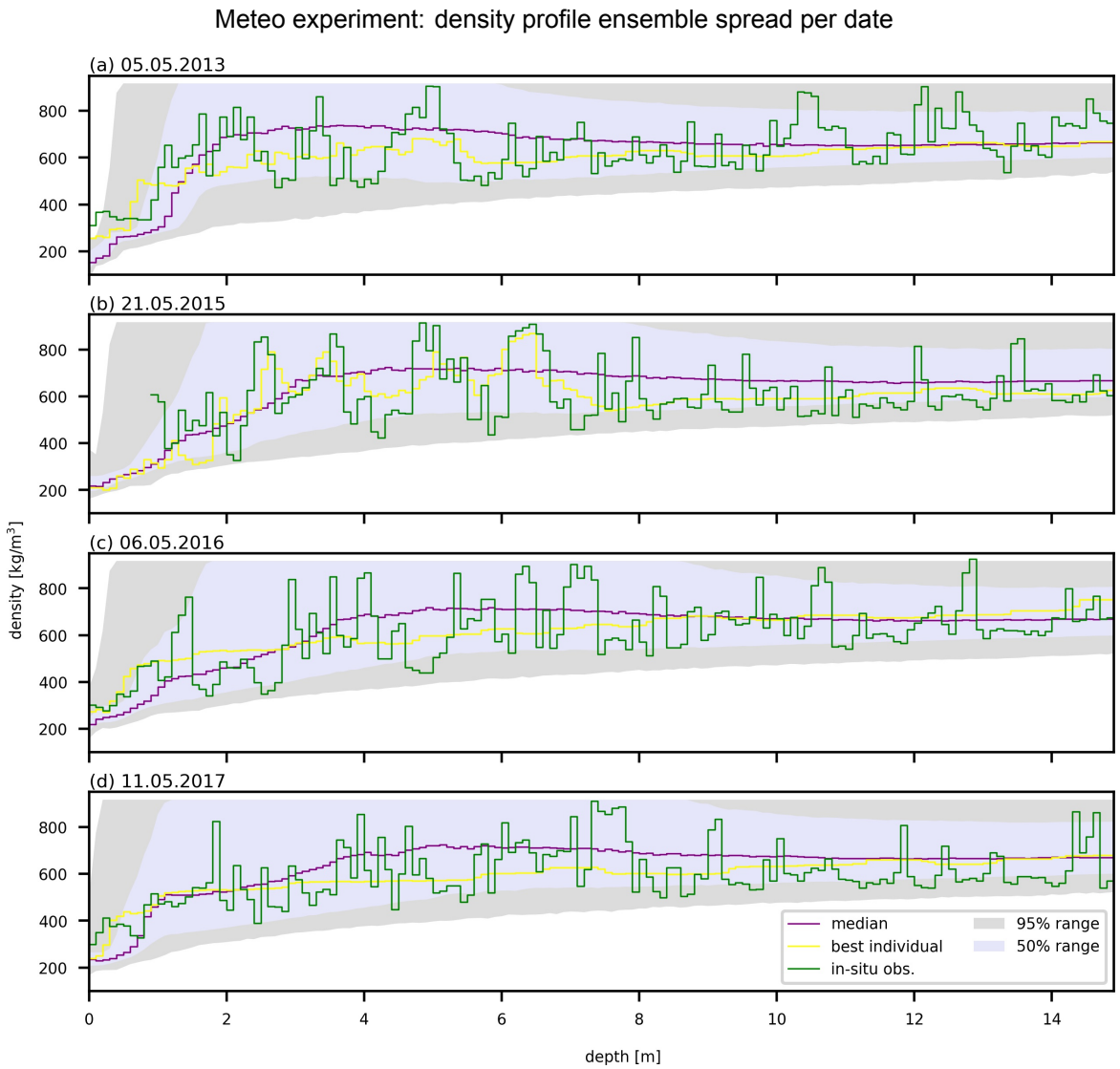


Figure D.2: *Comparison of modelled density profile ensemble spread with an actual firn core observations in May 2013 (**a**), 2015 (**b**), 2016 (**c**) and 2017 (**d**). The median modelled profiles were constructed by taking the ensemble's median density value at every depth. Best ensemble members in terms of min. RMSE plotted in yellow. All modelled and observed profiles were converted to have a depth resolution of 10 cm.*

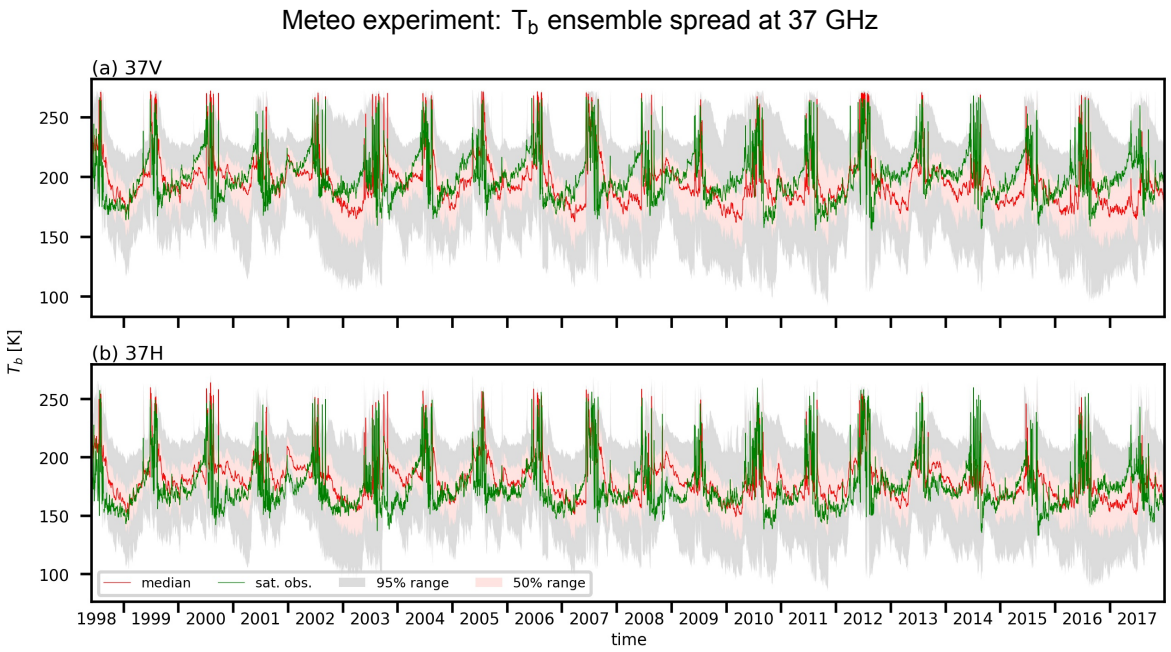


Figure D.3: *Comparison of 37 GHz modelled T_b ensemble spread with independent satellite observations at V (**a**) and H (**b**) polarisations. The median modelled series were constructed by taking the ensemble's median T_b value at every data point (data points are daily). Only 37 GHz T_b observations were modelled for the meteo experiment (no 19 GHz).*

E

Undisturbed Exeriment at H Pol

Complementary to the 37V and 19V ensemble spread of fig. 3.11. At 37H the mean 95% spread is equal to 71.5 K, RMSE of the median modelled series vs. observations 32.2 K and on an individual basis ensemble member's RMSE range from 20.3 to 53.4 K. At 19H the mean 95% spread is equal to 69.8 K, RMSE of the median modelled series vs. observations 36.1 K and on an individual basis ensemble member's RMSE range from 24.0 to 77.22 K.

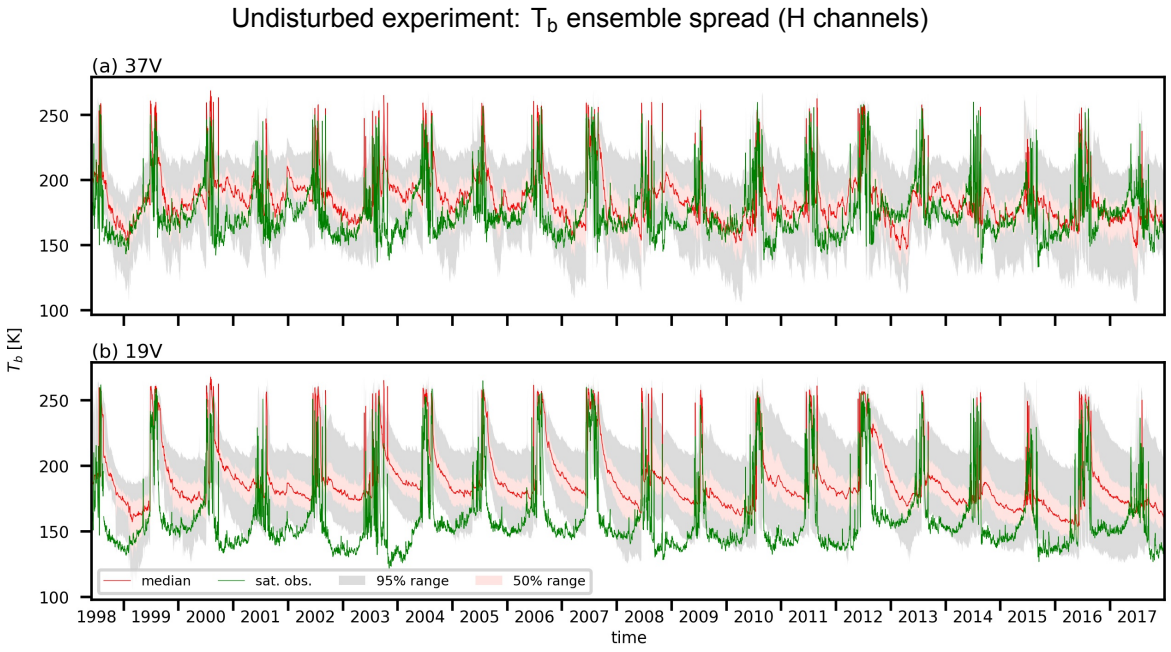


Figure E.1: *Comparison of modelled T_b ensemble spread with independent satellite observations at 37H (**a**) and 19H (**b**) for the undisturbed experiment. The median modelled series were constructed by taking the ensemble's median T_b value at every data point (data points are daily).*

Bibliography

Baroni, G. and Tarantola, S. (2012). A general probabilistic framework for uncertainty and sensitivity analysis of deterministic models. *iEMSs 2012 International Congres*.

Bartelt, P. and Lehning, M. (2002). A physical SNOWPACK model for the Swiss avalanche warning Part I: numerical model. *Cold Regions Science and Technology*, page 23.

Baxter, R. J. (1968). Percus–yevick equation for hard spheres with surface adhesion. *The Journal of Chemical Physics*, 49(6):2770–2774.

Bell, C., Mair, D., Burgess, D., Sharp, M., Demuth, M., Cawkwell, F., Bingham, R., and Wadham, J. (2008). Spatial and temporal variability in the snowpack of a High Arctic ice cap: implications for mass-change measurements. *Annals of Glaciology*, 48:159–170.

Bell, R. E., Chu, W., Kingslake, J., Das, I., Tedesco, M., Tinto, K. J., Zappa, C. J., Frezzotti, M., Boghosian, A., and Lee, W. S. (2017). Antarctic ice shelf potentially stabilized by export of meltwater in surface river. *Nature*, 544(7650):344–348.

Box, J. E., Van As, D., and Steffen, K. (2017). Greenland, Canadian and Icelandic land-ice albedo grids (2000–2016). *Geological Survey of Denmark and Greenland Bulletin*, pages 53–56.

Brucker, L., Picard, G., Arnaud, L., Barnola, J.-M., Schneebeli, M., Brunjail, H., Lefebvre, E., and Fily, M. (2011). Modeling time series of microwave brightness temperature at Dome C, Antarctica, using vertically resolved snow temperature and microstructure measurements. *Journal of Glaciology*, 57(201):171–182.

Brucker, L., Picard, G., and Fily, M. (2010). Snow grain-size profiles deduced from microwave snow emissivities in Antarctica. *Journal of Glaciology*, 56(197):514–526.

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 Smeets, C. J. P. P. (2015). Changing surface–atmosphere energy exchange and refreezing capacity of the lower accumulation area, West Greenland. *The Cryosphere*, 9(6):2163–2181.

Colbeck, S. (1997). A review of sintering in seasonal snow. *CRREL Report*, 97(10).

Colgan, W., Sommers, A., Rajaram, H., Abdalati, W., and Frahm, J. (2015). Considering thermal□viscous collapse of the Greenland ice sheet. *Earth's Future*, 3(7):252–267.

DeConto, R. M. and Pollard, D. (2016). Contribution of Antarctica to past and future sea-level rise. *Nature*, 531(7596):591–597.

Dupont, F., Picard, G., Royer, A., Fily, M., Roy, A., Langlois, A., and Champollion, N. (2014). Modeling the microwave emission of bubbly ice: Applications to blue ice and superimposed ice in the antarctic and arctic. *IEEE Transactions on Geoscience and Remote Sensing*, 52(10):6639–6651.

Fausto, R. S., Box, J. E., Vandecrux, B., van As, D., Steffen, K., MacFerrin, M. J., Machguth, H., Colgan, W., Koenig, L. S., McGrath, D., Charalampidis, C., and Braithwaite, R. J. (2018). A Snow Density Dataset for Improving Surface Boundary Conditions in Greenland Ice Sheet Firn Modeling. *Frontiers in Earth Science*, 6:51.

Fettweis, X., Box, J. E., Agosta, C., Amory, C., Kittel, C., Lang, C., van As, D., Machguth, H., and Gallée, H. (2017). Reconstructions of the 1900–2015 Greenland ice sheet surface mass balance using the regional climate MAR model. *The Cryosphere*, 11(2):1015–1033.

Fettweis, X., Tedesco, M., van den Broeke, M., and Ettema, J. (2011). Melting trends over the Greenland ice sheet (1958–2009) from spaceborne microwave data and regional climate models. *The Cryosphere*, 5(2):359–375.

Forster, R. R., Box, J. E., van den Broeke, M. R., Miège, C., Burgess, E. W., van Angelen, J. H., Lenaerts, J. T. M., Koenig, L. S., Paden, J., Lewis, C., Gogineni, S. P., Leuschen, C., and McConnell, J. R. (2014). Extensive liquid meltwater storage in firn within the Greenland ice sheet. *Nature Geoscience*, 7(2):95–98.

Gardner, A. S. and Sharp, M. J. (2010). A review of snow and ice albedo and the development of a new physically based broadband albedo parameterization. *Journal of Geophysical Research*, 115(F1):F01009.

Günther, D., Marke, T., Essery, R., and Strasser, U. (2019). Uncertainties in Snowpack Simulations—Assessing the Impact of Model Structure, Parameter Choice, and Forcing Data Error on Point□Scale Energy Balance Snow Model Performance. *Water Resources Research*, 55(4):2779–2800.

Hallikainen, M. T., Ulaby, F. T., and Van Deventer, T. E. (1987). Extinction behavior of dry snow in the 18-to 90-ghz range. *IEEE Transactions on Geoscience and Remote Sensing*, GE-25(6):737–745.

IPCC (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. *IPCC*.

IPCC (2019). IPCC Special Report on the Ocean and Cryosphere in a Changing Climate [H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]. *IPCC*.

Jansen, M. J. (1999). Analysis of variance designs for model output. *Computer Physics Communications*, 117(1-2):35–43.

Kontu, A., Lemmetyinen, J., Vehviläinen, J., Leppänen, L., and Pulliainen, J. (2017). Coupling snowpack-modeled grain size parameters with the hut snow emission model. *Remote Sensing of Environment*, 194:33–47.

Kuipers Munneke, P., Ligtenberg, S. R., Van Den Broeke, M. R., and Vaughan, D. G. (2014). Firn air depletion as a precursor of Antarctic ice-shelf collapse. *Journal of Glaciology*, 60(220):205–214.

Kuipers Munneke, P., van den Broeke, M. R., Lenaerts, J. T. M., Flanner, M. G., Gardner, A. S., and van de Berg, W. J. (2011). A new albedo parameterization for use in climate models over the Antarctic ice sheet. *Journal of Geophysical Research*, 116(D5):D05114.

Lazzaro, A., Wismer, A., Schneebeli, M., Erny, I., and Zeyer, J. (2015). Microbial abundance and community structure in a melting alpine snowpack. *Extremophiles*, 19(3):631–642.

Lehning, M., Bartelt, P., Brown, B., and Fierz, C. (2002a). A physical SNOWPACK model for the Swiss avalanche warning Part III: meteorological forcing, thin layer formation and evaluation. *Cold Regions Science and Technology*, page 16.

Lehning, M., Bartelt, P., Brown, B., Fierz, C., and Satyawali, P. (2002b). A physical SNOWPACK model for the Swiss avalanche warning Part II. Snow microstructure. *Cold Regions Science and Technology*, page 21.

Lenaerts, J., Lhermitte, S., Drews, R., Ligtenberg, S., Berger, S., Helm, V., Smeets, C., Broeke, M., van de Berg, W., van Meijgaard, E., Eijkelboom, M., Eisen, O., and Pattyn, F. (2017). Meltwater produced by wind–albedo interaction stored in an East Antarctic ice shelf. *Nature Climate Change*, 7(1):58–62.

Liang, D., Xu, X., Tsang, L., Andreadis, K. M., and Josberger, E. G. (2008). The Effects of Layers in Dry Snow on Its Passive Microwave Emissions Using Dense Media Radiative Transfer Theory Based on the Quasicrystalline Approximation (QCA/DMRT). *IEEE Transactions on Geoscience and Remote Sensing*, 46(11):3663–3671.

Linow, S., Hörhold, M. W., and Freitag, J. (2012). Grain-size evolution of polar firn: a new empirical grain growth parameterization based on X-ray microcomputer tomography measurements. *Journal of Glaciology*, 58(212):1245–1252.

Löwe, H. and Picard, G. (2015). Microwave scattering coefficient of snow in MEMLS and DMRT-ML revisited: the relevance of sticky hard spheres and tomography-based estimates of stickiness. *The Cryosphere*, 9(6):2101–2117.

Lüthi, M. P., Ryser, C., Andrews, L. C., Catania, G. A., Funk, M., Hawley, R. L., Hoffman, M. J., and Neumann, T. A. (2015). Heat sources within the Greenland Ice Sheet: dissipation, temperate paleo-firn and cryo-hydrologic warming. *The Cryosphere*, 9(1):245–253.

MacFerrin, M., Machguth, H., As, D. v., Charalampidis, C., Stevens, C. M., Heilig, A., Vandecrux, B., Langen, P. L., Mottram, R., Fettweis, X., Broeke, M. R. v. d., Pfeffer, W. T., Moussavi, M. S., and Abdalati, W. (2019). Rapid expansion of Greenland's low-permeability ice slabs. *Nature*, 573(7774):403–407.

MacFerrin, M., Stevens, C., Abdalati, W., and Waddington, E. (In Preparation). The Firn Compaction Verification and Reconnaissance (FirnCover) dataset. *In Preparation*.

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. (2016). Greenland meltwater storage in firn limited by near-surface ice formation. *Nature Climate Change*, 6(4):390–393.

Magnusson, J., Wever, N., Essery, R., Helbig, N., Winstral, A., and Jonas, T. (2015). Evaluating snow models with varying process representations for hydrological applications. *Water Resources Research*, 51(4):2707–2723.

Maslanik, J. and Stroeve, J. (2018). DMSP SSM/I-SSMIS Daily Polar Gridded Brightness Temperatures, Version 3. *NASA National Snow and Ice Data Center Distributed Active Archive Center*.

Mätzler, C. (1998). Improved born approximation for scattering of radiation in a granular medium. *Journal of Applied Physics*, 83:6111–6117.

Montgomery, L., Koenig, L., and Alexander, P. (2018). The SUMup dataset: compiled measurements of surface mass balance components over ice sheets and sea ice with analysis over Greenland. *Earth System Science Data*, 10(4):1959–1985.

Montpetit, B., Royer, A., Roy, A., Langlois, A., and Derksen, C. (2013). Snow Microwave Emission Modeling of Ice Lenses Within a Snowpack Using the Microwave Emission Model for Layered Snowpacks. *IEEE Transactions on Geoscience and Remote Sensing*, 51(9):4705–4717.

Mosley-Thompson, E., McConnell, J. R., Bales, R. C., Li, Z., Lin, P.-N., Steffen, K., Thompson, L. G., Edwards, R., and Bathke, D. (2001). Local to regional-scale variability of annual net accumulation on the Greenland ice sheet from PARCA cores. *Journal of Geophysical Research: Atmospheres*, 106(D24):33839–33851.

Mottram, R., B. Simonsen, S., Høyer Svendsen, S., Barletta, V. R., Sandberg Sørensen, L., Nagler, T., Wuite, J., Groh, A., Horwath, M., Rosier, J., Solgaard, A., Hvidberg, C. S., and Forsberg, R. (2019). An Integrated View of Greenland Ice Sheet Mass Changes Based on Models and Satellite Observations. *Remote Sensing*, 11(12):1407.

Mätzler, C. (2002). Relation between grain-size and correlation length of snow. *Journal of Glaciology*, 48(162):461–466.

Noël, B., van de Berg, W. J., van Wessem, J. M., van Meijgaard, E., van As, D., Lenaerts, J. T. M., Lhermitte, S., Kuipers Munneke, P., Smeets, C. J. P. P., van Ulft, L. H., van de Wal, R. S. W., and van den Broeke, M. R. (2018). Modelling the climate and surface mass balance of polar ice sheets using RACMO2 – Part 1: Greenland (1958–2016). *The Cryosphere*, 12(3):811–831.

Obleitner, F. and Lehning, M. (2004). Measurement and simulation of snow and superimposed ice at the Kongsvegen glacier, Svalbard (Spitzbergen). *Journal of Geophysical Research: Atmospheres*, 109(D4).

Pfeffer, W. T., Meier, M. F., and Illangasekare, T. H. (1991). Retention of Greenland runoff by refreezing: Implications for projected future sea level change. *Journal of Geophysical Research*, 96(C12):22117.

Phillips, T., Rajaram, H., and Steffen, K. (2010). Cryo-hydrologic warming: A potential mechanism for rapid thermal response of ice sheets: RAPID THERMAL RESPONSE OF ICE SHEETS. *Geophysical Research Letters*, 37(20):n/a–n/a.

Picard, G., Brucker, L., Fily, M., Gallée, H., and Krinner, G. (2009). Modeling time series of microwave brightness temperature in Antarctica. *Journal of Glaciology*, 55(191):537–551.

Picard, G., Brucker, L., Roy, A., Dupont, F., Fily, M., Royer, A., and Harlow, C. (2013). Simulation of the microwave emission of multi-layered snowpacks using the Dense Media Radiative transfer theory: the DMRT-ML model. *Geoscientific Model Development*, 6(4):1061–1078.

Picard, G., Royer, A., Arnaud, L., and Fily, M. (2014a). Influence of meter-scale wind-formed features on the variability of the microwave brightness temperature around dome c in antarctica. *The Cryosphere*, 8(3):1105–1119.

Picard, G., Royer, A., Arnaud, L., and Fily, M. (2014b). Influence of meter-scale wind-formed features on the variability of the microwave brightness temperature around Dome C in Antarctica. *The Cryosphere*, 8(3):1105–1119.

Picard, G., Sandells, M., and Löwe, H. (2018). SMRT: an active–passive microwave radiative transfer model for snow with multiple microstructure and scattering formulations (v1.0). *Geoscientific Model Development*, 11(7):2763–2788.

Pollard, D., DeConto, R. M., and Alley, R. B. (2015). Potential Antarctic Ice Sheet retreat driven by hydrofracturing and ice cliff failure. *Earth and Planetary Science Letters*, 412:112–121.

Raleigh, M. S., Lundquist, J. D., and Clark, M. P. (2015). Exploring the impact of forcing error characteristics on physically based snow simulations within a global sensitivity analysis framework. *Hydrology and Earth System Sciences*, 19(7):3153–3179.

Reijmer, C. H., van den Broeke, M. R., Fettweis, X., Ettema, J., and Stap, L. B. (2012). Refreezing on the Greenland ice sheet: a comparison of parameterizations. *The Cryosphere*, 6(4):743–762.

Riihelä, A., King, M. D., and Anttila, K. (2019). The surface albedo of the Greenland Ice Sheet between 1982 and 2015 from the CLARA-A2 dataset and its relationship to the ice sheet's surface mass balance. *The Cryosphere*, 13(10):2597–2614.

Robinson, A., Calov, R., and Ganopolski, A. (2012). Multistability and critical thresholds of the Greenland ice sheet. *Nature Climate Change*, 2(6):429–432.

Roy, A., Picard, G., Royer, A., Montpetit, B., Dupont, F., Langlois, A., Derksen, C., and Champollion, N. (2013). Brightness temperature simulations of the canadian seasonal snowpack driven by measurements of the snow specific surface area. *IEEE Transactions on Geoscience and Remote Sensing*, 51(9):4692–4704.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010). Variance based sensitivity analysis of model output - design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270.

Sobol, I. M. (1993). Sensitivity analysis for nonlinear mathematical models. *Mathematical Models and Computer Exp*, 1(4):407–414.

Steffen, K., Box, J., and Abdalati, W. (1996). Greenland Climate Network: GC-Net. *CRREL Special Report on Glaciers, Ice Sheets and Volcanoes*, 96(27):98–103.

Steger, C. R., Reijmer, C. H., van den Broeke, M. R., Wever, N., Forster, R. R., Koenig, L. S., Kuipers Munneke, P., Lehning, M., Lhermitte, S., Ligtenberg, S. R. M., Miège, C., and Noël, B. P. Y. (2017). Firn Meltwater Retention on the Greenland Ice Sheet: A Model Comparison. *Frontiers in Earth Science*, 5.

Sturm, M. and Benson, C. (2004). Scales of spatial heterogeneity for perennial and seasonal snow layers. *Annals of Glaciology*, 38:253–260.

Tsang, L., Pan, J., Liang, D., Li, Z., Cline, D. W., and Tan, Y. (2007). Modeling active microwave remote sensing of snow using dense media radiative transfer (dmrt) theory with multiple-scattering effects. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4):990–1004.

van As, D., Box, J. E., and Fausto, R. S. (2016). Challenges of Quantifying Meltwater Retention in Snow and Firn: An Expert Elicitation. *Frontiers in Earth Science*, 4.

van den Broeke, M. R., Enderlin, E. M., Howat, I. M., Kuipers Munneke, P., Noël, B. P. Y., van de Berg, W. J., van Meijgaard, E., and Wouters, B. (2016). On the recent contribution of the Greenland ice sheet to sea level change. *The Cryosphere*, 10(5):1933–1946.

van Pelt, W. J. J., Pohjola, V. A., and Reijmer, C. H. (2016). The Changing Impact of Snow Conditions and Refreezing on the Mass Balance of an Idealized Svalbard Glacier. *Frontiers in Earth Science*, 4.

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. (2018). Drivers of Firn Density on the Greenland Ice Sheet Revealed by Weather Station Observations and Modeling. *Journal of Geophysical Research: Earth Surface*, 123(10):2563–2576.

Vandecrux, B., Mottram, R., Langen, P. L., Fausto, R. S., Olesen, M., Stevens, C. M., Verjans, V., Leeson, A., Ligtenberg, S., Kuipers Munneke, P., Marchenko, S., van Pelt, W., Meyer, C. R., Simonsen, S. B., Heilig, A., Samimi, S., Marshall, S., Machguth, H., MacFerrin, M., Niwano, M., Miller, O., Voss, C. I., and Box, J. E. (2020). The firn meltwater Retention Model Intercomparison Project (RetMIP): evaluation of nine firn models at four weather station sites on the Greenland ice sheet. *The Cryosphere*, 14(11):3785–3810.

Wang, W., Zender, C. S., van As, D., Smeets, P. C. J. P., and van den Broeke, M. R. (2016). A Retrospective, Iterative, Geometry-Based (RIGB) tilt-correction method for radiation observed by automatic weather stations on snow-covered surfaces: application to Greenland. *The Cryosphere*, 10(2):727–741.

Wever, N., Fierz, C., Mitterer, C., Hirashima, H., and Lehning, M. (2014). Solving Richards Equation for snow improves snowpack meltwater runoff estimations in detailed multi-layer snowpack model. *The Cryosphere*, 8(1):257–274.

Wiesmann, A., Fierz, C., and Mätzler, C. (2000). Simulation of microwave emission from physically modeled snowpacks. *Annals of Glaciology*, 31:397–405.

Wiesmann, A. and Mätzler, C. (1999). Microwave Emission Model of Layered Snowpacks. *Remote Sensing of Environment*, 70(3):307–316.