FirnLearn: A Neural-Network-based approach to Firn Density Modeling in Antarctica

Ayobami Ogunmolasuyi[1], Colin R Meyer[1], Ian McDowell[2], Megan Thomson-Munson[3], Ian Baker[1]

[1] *Thayer School of Engineering, Dartmouth College, Hanover, NH 03755*

[2] *Graduate Program of Hydrologic Sciences, University of Nevada, Reno, NV 89557*

[3] *University of Colorado, Boulder, CO 80309*

Correspondence: Ayobami Ogunmolasuyi <ayobami.o.ogunmolasuyi.th@dartmouth.edu>

ABSTRACT. Understanding firn densification is essential for interpreting ice core records, predicting ice sheet mass balance, elevation changes, and future sea-level rise. Current models of firn densification on the Antarctic Ice Sheet (AIS), such as the Herron and Langway (1980) model are either simple semi-empirical models that rely on sparse climatic data and surface density observations or complex physics-based models that rely on poorly understood physics. In this work, we introduce a deep learning technique to study firn densification on the AIS. Our model, FirnLearn, evaluated on 225 cores, shows an average root mean square error (RMSE) of 31 kg m^{-3} and explained variance of 91%. We use the model to generate surface density and the depths to the 550 kg m^{-3} and 830 kg m^{-3} density horizons across the AIS to assess spatial variability. Comparisons with the Herron and Langway (1980) model at ten locations with different climate conditions demonstrate that FirnLearn more accurately predicts density profiles in the second stage of densification and complete density profiles without direct surface density observations. This work establishes deep learning as a promising tool for understanding firn processes and advancing towards a universally applicable firn model.

INTRODUCTION

As snow falls on the surface of the Antarctic Ice Sheet (AIS), it compacts into glacial ice, transitioning through an intermediate stage called firn. Firn has a density that ranges between that of settled snow (300 kg m^{-3}) and glacial ice (917 kg m^{-3}) depending on the network of interconnected pores which exchange air with the atmosphere (Buizer, 2013; Van den Broeke, 2008). Firn densification into glacial ice is influenced by several factors, such as, overburden stress, temperature, grain size, wind, impurity concentration, and water content (Arthern and others, 2010; Hörhold and others, 2012; Kingslake and others, 2022; Baker and Ogunmolasuyi, 2024). Understanding firn densification is important as it affects several processes in ice sheets. Firstly, given that densification changes in response to climatic factors, it causes uncertainty in ice-sheet elevation and mass balance estimates (Helsen and others, 2008; Smith and others, 2020). Secondly, densification results in the closure of the interconnected network of pores, which when closed off, traps gases in the ice. The age difference between the trapped gases and the ice is important for interpreting ice core records (Alley, 2000; Cuffey and Paterson, 2010). Lastly, the pore space within firn columns can serve as storage for meltwater from the warming climate, hence, breaking the link between surface melt, runoff, and sea-level rise (Harper and others, 2012; Forster and others, 2013; Meyer and Hewitt, 2017). Consequently, a comprehensive understanding of firn processes is crucial for accurately predicting ice sheet responses to climate change The-Firn-Symposium-Team (2024).

Firn densification is controlled by microstructural evolution (Anderson and Benson, 1963; Arnaud and others, 2000). It occurs in three stages, each characterized by distinct mechanisms. Initially, grain boundary sliding, vapor transport, and surface diffusion dominate until reaching a density of 550 kg m^{-3} (Anderson and Benson, 1963; Alley, 1987; Gow, 1969; Maeno and Ebinuma, 1983). In the second stage, pore space reduction limits vapor diffusion, giving way for sintering processes and recrystallization until a density of 830 kg m^{-3} is attained (Gow, 1969; Maeno and Ebinuma, 1983). The depth at 830 kg m^{-3} is typically denoted the pore close-off depth. Here, air becomes trapped in bubbles, slowing the densification process as the bubbles are compressed and eventually diffuse into the surrounding ice (Salamatin and others, 1997). Finally, bubble shrinkage and compression become dominant until the density of ice (917 kg m^{-3}) is reached (Bader, 1965). Several studies have been aimed at shedding more light on the microstructural processes in firn (Maeno and Ebinuma, 1983; Freitag and others, 2004; Kipfstuhl and others, 2009; Lomonaco and others, 2011; Burr and others, 2018; Li and Baker, 2021; Ogunmolasuyi and others, 2023). However, a comprehensive understanding of large-scale implications of firn densification requires an integration between the underlying microphysics and modeling. To this end, over four decades of effort has been undertaken to develop firn densification models (Herron and Langway, 1980; Alley, 1987; Barnola and others, 1991; Arnaud and others, 2000; Kaspers and others, 2004; Ligtenberg and others, 2011; Stevens and others, 2020; Meyer and others, 2020; Stevens and others, 2023). These models are either empirical (Herron and Langway, 1980; Barnola and others, 1991; Li and Zwally, 2011) or microphysics-based (Alley, 1987; Arnaud and others, 2000; Morris and Wingham, 2014).

However, knowledge gaps owing to an incomplete understanding of the underlying physics of firn densification still limit the accuracy of microphysics models. Hence, most firn densification models are empirical, predicting density based only on accumulation rate and temperature. These variables are usually obtained from ice core data such as (Buizert and others, 2012), regional climate models such as the Regional Atmospheric Climate (RACMO) (Noël and others, 2018) or long-term weather station data such as the Greenland climate network (GCN) (Steffen and Box, 2001). The models are then used to fit depth-density profiles derived from firn cores, with an assumption known as Sorge's law which states that the accumulation rate, surface density, and the firn column are invariant in time (Bader, 1954) . While these models have served the glaciology community reasonably well, their predictive accuracy can be limited, particularly under conditions that differe significantly differ from those used during their callibration (Lundin and others, 2017; Verjans and others, 2020).

In this study, we explore a novel approach to firn densification modeling based on a statistical analysis of known depth-density profiles as an attempt to improve the firn density estimates of empirical models. In recent years, the utility and significance of machine learning methods have grown. In particular, the ever-growing volume of data combined with hardware and optimization algorithms that allow complex systems to be fitted effortlessly has resulted in advances across various scientific fields, including earth sciences (Camps-Valls and others, 2020; Reichstein and others, 2019), among several other applications. While machine learning techniques, particularly artificial neural networks (ANNs) have seen increasing application in glaciology, including for simulating glacier length (Steiner and others, 2005; Nussbaumer and others, 2012), and modeling glacier flow, evolution and mass balance (Bolibar and others, 2020; Brinkerhoff and others, 2021), less attention has been paid to its implementation in firn densification modeling. Only a few machine learning models have been applied to firn processes. Rizzoli and others (2017) applied clustering techniques to characterize snow facies while Dell and others (2022) used a combination of clustering and classification techniques to identify slush and melt-pond water and Dunmire and others (2021) employed a convolutional neural network to detect buried lakes across the Greenland Ice Sheet. Notably, the only studies that have applied machine learning methods to modeling firn density was done by Li and others (2023), who trained a random forest on radiometer and scatterometer data to derive spatial and temporal variations in Antarctic firn density, and Dunmire and others (2024) who used a random forest to predict ice-shelf effective firn air content.

Here, we present a new steady state densification model: FirnLearn, which takes a deep learning approach to firn densification modeling. FirnLearn simulates the density profile over depth using a deep ANN, fed by density observations from the Surface Mass Balance and Snow on Sea Ice Working Group (SUMup) dataset (Montgomery and others, 2018; Vandecrux and others, 2024), and accumulation rate and temperature data from RACMO (Van Wessem and others, 2014; Noël and others, 2018).

In the next section, we present an overview of the data, brief descriptions of the ANN architecture as well as the evaluation techniques used in this study. In section 3, we present applications of FirnLearn to predicting surface density, depths at 550 kg m^{-3} and 830 kg m^{-3} density horizons, as well as firn air content (FAC). Here, we also discuss the performance of FirnLearn in comparison to the depth-density model of Herron and Langway (1980). FirnLearn maintains a high accuracy and it is robust to outliers, diverse climatic conditions, as well as surface density data. It can also predict surface density, a feature that other models do not share. FirnLearn can be a valuable tool to scientists for better constraining the physics governing firn densification. By examining its predictions, researchers can uncover and analyze complex relationships within firn density data.

METHODS

DATA

The dataset used in this study is based on field observations extracted from the December 2024 release of the SUMup dataset Montgomery and others (2018); Vandecrux and others (2024) and model outputs from RACMO (Van Wessem and others, 2014; Noël and others, 2018) (see figure 1). We combined firn density observations from 2689 locations across the AIS from SUMup with accumulation rate and temperature outputs from RACMO2.3. Given that FirnLearn is a steady-state model, it relies on time-averaged accumulation rate and temperature data, hence we extracted the 1979-2016 average accumulation rate and temperature values from RACMO.

Description	No. of cores
depth < 10 m	2476
depth > 100 m	39
density >= 550 kg m^{-3}	349
density >= 830 kg m^{-3}	71

Table 1. Summary of core depth and density characteristics in SUMup

Density and depth

The snow/firn density subdataset was extracted from SUMup. It contains over 2 million unique measurements of density at different depths across both the Antarctic and Greenland Ice Sheets. These density measurements were obtained using density cutters of different sizes used in snow pits, gravitational methods on ice core sections, neutron-density methods in boreholes, X-ray microfocus computer tomography on snow samples, gamma-ray attenuation in boreholes, pycnometers on snow samples, optical televiewer (OPTV) borehole lagging, and density and conductivity permittivity (DECOMP) (Montgomery and others, 2018). The number of cores at important depths and densities are detailed in table 1

Climate Variables

Accumulation rate: The accumulation rate dataset was obtained from the output of the RACMO2.3 model, containing total precipitation (snowfall and rainfall), runoff, melt, refreezing, and retention. For our accumulation rate input, we average annual surface mass balance (SMB) outputs from RACMO2.3 for 1979-2016. SMB values were converted to meters of water equivalent per year (m w.e. yr^{-1}). For the purpose of this study we assume zero ablation in Antarctica and use SMB as the accumulation rate.

Temperature: For our surface temperature input, we average annual surface temperature outputs from RACMO2.3 for 1979-2016.

Both variables have an initial dimension of (240,262), which we subsequently reshaped into a vector with dimensions (62880,1). The RACMO datasets were combined with the SUMup datasets to create a combined dataset of Antarctic observations containing density, depth, accumulation rate, and temperature.

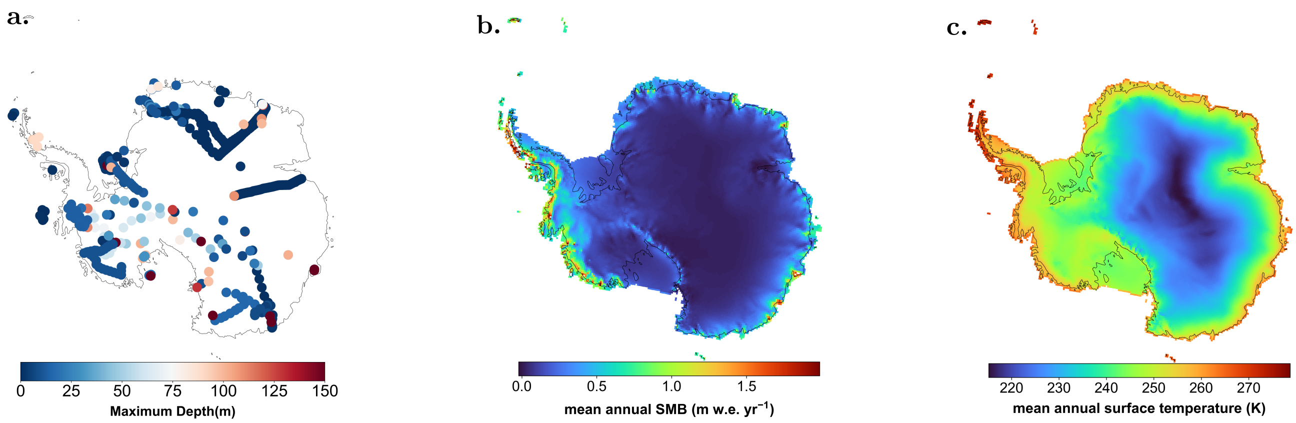

Fig. 1. (a) Locations of the 2689 cores used for density predictions extracted from SUMup colored with their maximum depths(b)surface mass balance and (c) surface temperature from RACMO2.3

FirnLearn model development

In this section, we describe the procedures for preprocessing the input data, and building, training, validating and testing the machine learning models. In the supplement of this paper, we describe other models employed in predicting density profiles, as well as their relative performance.

Neural network architecture

Artificial Neural Networks (ANNs) are nonlinear statistical models that recognize relationships and patterns between the input and output variables of structured data (figure 2) in a manner that models the biological neurons of the human brain(Hatie and others, 2009; O'Shea and Nash, 2015). The structure of an ANN consists of (1) an architecture of node layers containing the input layer that receives the data, the output layer that produces an estimate of the dependent variable, and hidden layers that take in and sum the weighted inputs and produce an output for other hidden layers or the output layer, (2) an optimization algorithm that determines and updates the weights of the connections between the neurons O'Shea and Nash (2015), and (3) an activation function that determines the output of each neuron.

The goal of the training process is to continuously update the weights in every iteration to minimize a loss function, which in most cases, as in our case, is the mean squared error. This cost function is expressed as

$$J(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(\rho_{\text{NN}}(\mathbf{x}_i;\theta) - \rho_{\text{true}}(\mathbf{x}_i)\right)^2, \hat{\theta} = \arg\min_{\theta} J(\theta) \tag{1}$$

where $\hat{\theta}$ represents the optimal parameters of the neural network, θ denotes the sets of parameters (weights) of the neural network that the optimization algorithm is adjusting to minimize the cost function, $J(\theta)$ is the cost function, N is the total number of data points in the dataset. x_i is the input vector for the i-th data point, where i is the index of the current data point that runs from 1 to N, and x is contains the features accumulation rate, temperature and depth. Latitude and longitude are not included as they are already captured in the accumulation rate and temperature. To test this, we included them in a version of our model and found a marginal improvement that was less general. Moreover, when the results from this model were geographically plotted, there were boundary artifacts which may point to scaling issues that cause the NN to overfit to coordinate-specific patterns.

The variables that determine the structure and performance of a model are called hyperparameters and they include the number of neurons per layer, number of layers, activation function, optimizer, learning rate, batch size and number of epochs.

FirnLearn, shown in Figure 2 is a seven-layered ANN coded in python using using tensorflow keras and sci-kit learn. The hyperparameters used to construct FirnLearn were tuned using cross validation to find the best performing combination of hyperparameters, i.e., the hyperparameter combination with the lowest root mean squared error (RMSE - see Evaluation section below). It consists of 1 input layer with 3 neurons corresponding to the number of selected features, accumulation rate, temperature, and depth; 5 hidden layers with 100, 50, 20, 20, 10 neurons, respectively; and 1 output layer corresponding to density. Leaky ReLU was chosen as the activation function for the hidden layers. ReLU, short for Rectified Linear Unit, is a piecewise function that outputs the input value if it is greater than 0. It is given by:

$$f(x) = \begin{cases} \alpha x, & \text{if } x \leqslant 0 \\ x, & \text{if } x > 0 \end{cases} \tag{2}$$

Where α is a small constant (e.g. 0.01) ensuring the gradient is non-zero for negative values.

For the output layer, the sigmoid function was chosen as the activation function. It is represented as

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{3}$$

We used the Adam optimizer technique (Kingma and Ba, 2017) to optimize the weights for gradient descent. We also tuned the learning rate, which determines how much the weights are changed in each iteration. The best performing learning rate was 0.0001 among a starting range of 0.01,0.001 and 0.0001.

FirnLearn predicts firn density (ρ) using the function $\rho = f(A, T, z)$ where A represents the accumulation

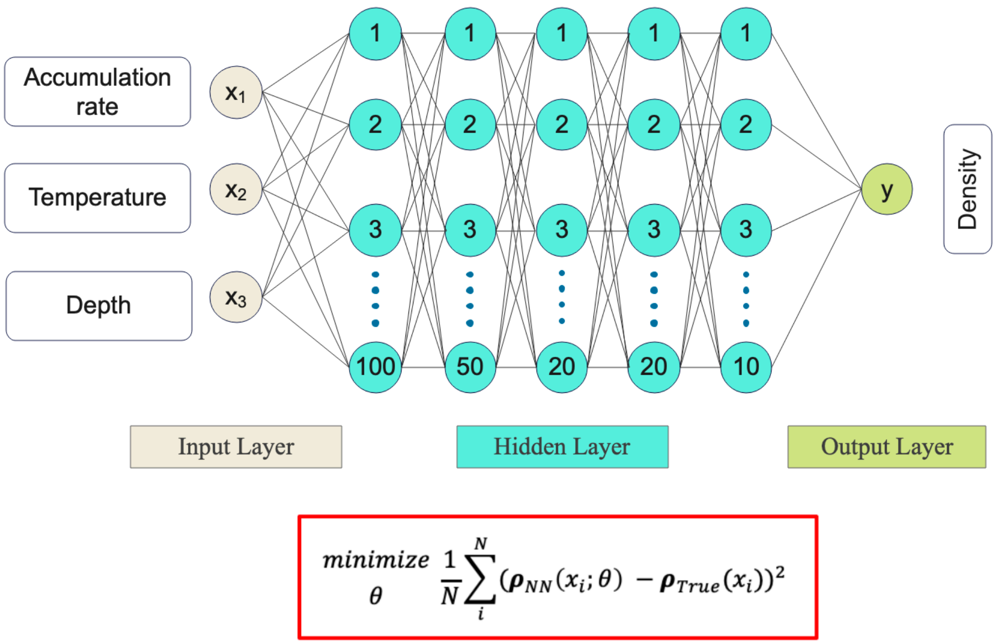

Fig. 2. FirnLearn's Artificial Neural Network Architecture. FirnLearn's goal is to minimize thef cost function (in red box). Here, ρ_{NN} is the density predicted by the neural network, x_i is each individual feature, θ is the weight of the neural network, and ρ_{True} is the observed density value.

rate, T represents temperature, and z represents depth. For surface density predictions (z = 0), the model simplifies to $\rho = f(A, T, 0)$. These inputs capture the primary drivers of firn densification processes, aligning with the benchmark Herron and Langway (1980) model

Training and testing

After the dataset was extracted, we split it into training, testing, and validation sets. The training set is used to train the model, allowing it to learn patterns and relationships in the data. The validation set is used during training to tune model hyperparameters and prevent overfitting by providing an independent evaluation of the models performance. Finally, the testing set is used after training to assess the model's performance on unseen data and evaluate its generalization capabilities. Before training commenced, sklearn's *StandardScaler* was used to normalize the input variables and *MinMaxScaler* was used to normalize the output variable from 0 to 1. We conducted a 72-8-20 split on the cores for the training-testing-validation sets. The validation cores (536 cores) were selected at random, while the testing cores (225 cores) were selected to ensure 9 of the cores were at least 50 m deep, and 1 was below 50 m in order to provide visual representation of FirnLearn's performance with depth, as shown in figure 3. We selected these sites to be representative of the full spread of regions in Antarctica, selecting one site from the Antarctic Peninsula, East Antarctica, West Antarctica, the South Pole and near the Ross Sea. This also let us test a range of surface density values from around 320 kg m^{-3} to greater than 550 kg m^{-3}. Our tests were conducted on one from the Larsen C Ice Shelf (66.58 °S,63.21 °W), Marie Byrd Land (78.12°S, 120°W), Ellsworth land (78.1 °S, 95.65°W) Taylor Dome (77.88°S, 158.46°E), near Vostok Station (82.08°S, 101.97°E), Dumbont D'Urville Station (66.66°S, 140°E), two cores in the Queen Maud Land region [(73.1°S, 39.8°E); 75°S, 0°], and two cores from the South Pole [(88.51°S, 178.53°E), (90°S,0)]. Across all 225 test cores, the RMSE was 30 kg m^{-3} and the explained variance was 98%.

Evaluation

We evaluate our model's performance using several metrics. We use the coefficient of determination r^2 to quantify how well the model predicts the dependent variable (density). It is given by

$$r^2 = 1 - \frac{\sum_{i=1}^{N}(\rho_i - \hat{\rho}_i)^2}{\sum_{i=1}^{N}(\rho_i - \bar{\rho})^2}. \tag{4}$$

where ρ_i are the actual density value at each data point, $\hat{\rho}_i$ is the predicted density value for each data point, and $\bar{\rho}$ is the mean of the actual density values.

The root mean squared error (RMSE) is an average measure of the difference between the observed density and the predicted density, given by

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left(\rho_i - \hat{\rho}_i\right)^2}{N}}, \tag{5}$$

where N is the number of model-observation pairs, ρ_i is the true density value, $\hat{\rho}_i$ is the predicted density value, and $\bar{\rho}_i$ is the mean of the observed density values. We evaluate the RMSE for an independent test set with a split discussed in the 'Training and testing' section. To estimate the difference between modeled and observed surface density and FAC, we use the relative bias metric. The relative bias is given as

$$\text{relative bias} = \frac{\hat{\rho}_i - \rho_i}{\rho_i} \times 100\%. \tag{6}$$

A positive relative bias indicates an overestimation by the model, while negative bias indicates an underestimation by the model.

Herron and Langway, 1980

Herron and Langway (1980), denoted HL in this study, is a benchmark empirical firn densification model, upon which many contemporary models are built due to its foundational assumptions. The assumptions made in HL are: (1) the densification rate is a function of the porosity, and (2) the densification rate has an Arrhenius dependence on the temperature. These assumptions are combined to form the equation

$$\frac{d\rho}{dt} = C(\rho_{ice} - \rho), \tag{7}$$

where ρ_{ice} is the density of ice (917 kg m^{-3}), ρ is the density at a given depth, and C is a site-specific constant. It is worth noting that while $\frac{d\rho}{dt}$ may seem time dependent, the time-dependent change in density is captured in the depth variability of density.

$$C = k \exp\left(-\frac{Q}{RT}\right) A^{a}, \tag{8}$$

where k in equation 8 is a temperature-dependent Arrhenius-type rate constant, a is a constant dependent on the densification mechanism, Q is the Arrhenius activation energy (kJ mol^{-1}), R is the gas constant (8.314 kJ mol^{-1} K^{-1}), T is the mean annual temperature at the site (K), and A is the accumulation rate in waters-equivalent. The site-specific constant, C, for $\rho \leqslant 550$ kg m^{-3}, is given by

$$C = 11 \exp\left(-\frac{10.16}{RT} A\right), \tag{9}$$

and for $\rho > 550$ kg m^{-3}, we have

$$C = 575 \exp\left(-\frac{21.4}{RT} A^{0.5}\right). \tag{10}$$

HL requires a surface density boundary condition. In order to obtain predictions for depth-density profiles, we used both surface density values from observations, as well as surface density predictions from Ligtenberg and others (2011). However, for predicting depths at 550 kg m^{-3} and 830 kg m^{-3}, we used surface density predictions from FirnLearn, which allowed for a direct comparison.

RESULTS AND DISCUSSION

Depth-density profiles

We use FirnLearn and HL to simulate firn profiles at 10 test sites and compare the results in figure 3, tables 2, 3, and 4. This allows us to visually evaluate the difference in performance between HL and FirnLearn. For HL, we evaluated the HL function using two different surface density values, and the two curves are named HL80-observation and HL80-Ligtenberg. For the HL80-observation curves, we use the surface density value directly from the observations, while for the HL80-Ligtenberg curves, we set the surface density values to surface density predictions from Lightenberg and others (2011). The FirnLearn curves, depicted as black lines, are generated by applying function evaluations of the FirnLearn model, taking in specified accumulation rates, temperatures, and depths, i.e. $\rho = f(A, T, z)$. All three models do a good job of predicting the observations. FirnLearn outperforms both HL80 models across the full depth range, in locations c,d,e, and j (figure 3 and table 2). HL80-observation outperforms FirnLearn and HL80-Ligtenberg at locations a,b, f, and g, while FirnLearn performs comparably to HL80-observation at locations h and i.

Of the three model predictions plotted for each core, although FirnLearn perfoms well, HL80-observation leads in the first stage of densification due to accurate surface density inputs in cores b, c, d, f, and g (table 3), with FirnLearn having a comparable performance at these locations. FirnLearn outperforms in the second stage across most locations (table 4 c, d, e, and j) and is comparable or better than HL80-observation in overall performance, while HL80-Ligtenberg generally underperforms consistently due to a greater mismatch in surface density predictions.

A more pronounced discrepancy in performance is evident in figure 3a for the Larsen C ice shelf. This location was included to represent conditions on an ice shelf with meltwater conditions, as well as conditions that may become more frequent with climate change. Here, owing to the accurate surface density input, HL80-observation predicts the density trend with greater accuracy than HL80-Ligtenberg and FirnLearn. On the other hand, at location g, the Dumbont D'Urville Station, FirnLearn underperforms in the third stage of densification (>830 kg m^{-3}). This underperformance is attributed to the limited number of observations under conditions similar to those at this site in Antarctica. As discussed in the introduction and evidenced in the SUMup density dataset, surface density measurements have only been collected for a small percentage of the AIS. Figure 3 shows that without accurate surface density observations, FirnLearn copy 2.jpg copy 2.bb

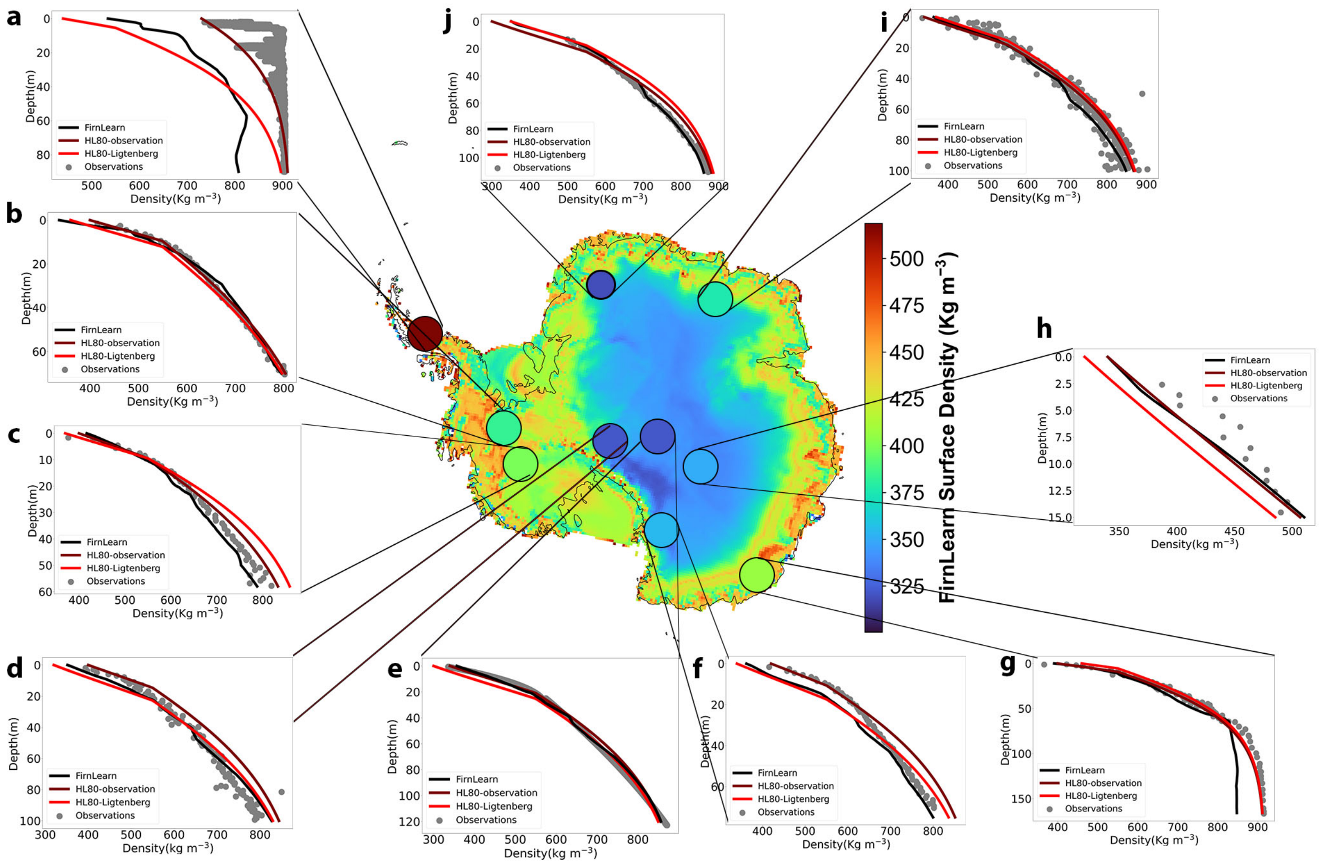

Fig. 3. Depth-density profiles at the 6 test sites. Shown corresponding to each site are the observed density profile in grey, the FirnLearn modeled density in red, and the HL modeled density in black for (a) a location on the Larsen C Ice Shelf, (b) location on the Marie Byrd Land, (c) location near the South Pole,(d) the South Pole, (e) the Taylor dome, and (f) a location near Vostok station

is a better density prediction model than Herron and Langway (1980).

It is worth highlighting that FirnLearn offers the added advantage of providing density information at specific depths for a given site without requiring surface density or density data from previous depths. This characteristic further enhances the speed and utility of FirnLearn in densification research. Additionally, it could be useful in ice core drilling operations for optimized site selection and resource allocation.

Surface Density

We predict surface density across the AIS by putting accumulation rate and temperature from RACMO2.3 (Noël and others, 2018) at $z = 0$ into the trained and validated FirnLearn model. These predictions are based on the equation $\rho = f(A, T, 0)$, where the function f is FirnLearn, A represents the accumulation

Profiles	Core No.	FirnLearn (kg m^{-3})	HL80-Observation (kg m^{-3})	HL80-Ligtenberg (kg m^{-3})
a	659	155.1	59	195.1
b	24	20.4	7.9	20.3
c	25	20.7	21.3	41.6
d	2	25.5	45.1	35.6
e	1937	10.3	14.8	23.6
f	26	56.1	32.2	41.6
g	4435	35.5	15.3	22.2
h	10	22.7	22.6	41.5
i	2556	30.5	30.6	28
j	740	7.9	22.4	33.1

Table 2. RMSE values for the density profile predictions in Fig. 3

Profiles	Core No.	FirnLearn (kg m^{-3})	HL80-Observation (kg m^{-3})	HL80-Ligtenberg (kg m^{-3})
a	659	–	–	–
b	24	87.6	37	117.8
c	25	25.1	18.1	15.9
d	2	33.8	26.4	57.8
e	1937	16.7	17.4	59.1
f	26	66.4	62	73.5
g	4435	21.9	20.1	55.6
h	10	22.7	22.6	41.5
i	2556	43.6	54	37.4
j	740	–	–	–

Table 3. RMSE values for the **first stage of densification** ($<=$ 550 kg m^{-3}) density profile predictions in Fig. 3

Profiles	Core No.	FirnLearn (kg m^{-3})	HL80-Observation (kg m^{-3})	HL80-Ligtenberg (kg m^{-3})
a	659	–	–	–
b	24	20.3	7.6	12.7
c	25	19.7	21.9	44.8
d	2	23	48.5	27.7
e	1937	9	14.3	11.2
f	26	54.4	24.9	34.5
g	4435	36.5	14.2	14.9
h	10	–	–	–
i	2556	27	22.2	25.5
j	740	7.9	19.9	34.1

Table 4. RMSE values for the **second stage of densification** ($>$ 550 kg m^{-3}) density profile predictions in Fig. 3

rate, and T represents the temperature. Across Antarctica, the surface density exhibits a notable spatial variation (Figure 4a). In the interior of East Antarctic, we observe relatively lower values in the range 320–380 kg m^{-3}, reflecting the region's colder surface temperatures. In contrast, we find higher surface density values, exceeding 450 kg m^{-3},along the coastal areas and on ice shelves. We attribute these higher densities to the higher temperatures, higher accumulation rates, and the higher wind speeds prevalent in these regions (McDowell and others, 2020). For the majority of the sites, the relative bias is within $\pm$25%, with only one site having a relative bias above 100% (Figure 4b). For this site in the Southeastern Antarctica, FirnLearn overpredicts the surface density by 174%.

Semi-empirical models require a prescribed surface density boundary condition, making these surface density predictions a key output of FirnLearn. The importance of the surface density boundary condition was underscored by Thompson-Munson and others (2023). They employed two models, the physics-based SNOWPACK Bartelt and Lehning (2002) with a surface density that varies based on atmospheric conditions, and the Community Firn Model configured with a semi- empirical densification equation (CFM-GSFC; Stevens and others (2020)) run with a constant surface density of 350 kg m^{-3}. Their analysis of firn properties across the GrIS revealed that SNOWPACK simulated more variability between firn layers compared to CFM-GSFC, although some of this variability could be associated with SNOWPACK's microstructure dependence. Importantly, our predictions are similar to prior research findings (Kaspers and surface density.jpg surface density.bb

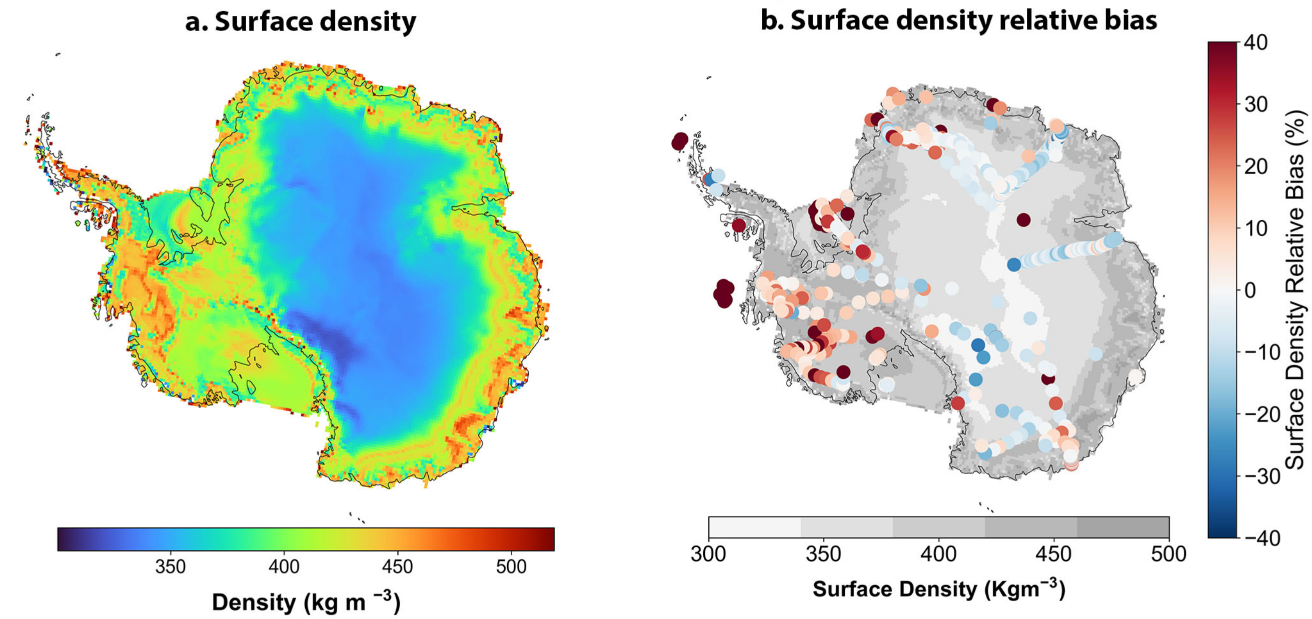

Fig. 4. (a) The FirnLearn predicted surface density field for Antarctica and (b) Relative bias between the predicted surface density and the observed surface density

others, 2004; van den Broeke, 2008; Ligtenberg and others, 2011) that have employed parameterizations based on combinations of surface temperature, accumulation rate, wind speed to derive surface density predictions.

Depths at 550 kg m^{-3} and 830 kg m^{-3}

Figure 5a depicts the depth at 550 kg m^{-3} with a depth range from 0-30 m, with higher values (18-30 m) concentrated in East Antarctica and lower values (0-15 m) prevalent in West Antarctica and along the coast. This spatial distribution is similar to the pattern of surface density, reflecting an expected relationship between the two variables.Figure 5c shows the depth at 830 kg m^{-3} which ranges from 20 - 122 m, with higher values (80 - 122 m) predominantly found in East Antarctica. The spatial distribution of z_{830} is different than z_{550} in that for z_{830}, there are higher values in regions of West Antarctica, the Antarctic Peninsula, and certain coastal areas. This is primarily attributed to the higher accumulation rates, which result in the rapid burial of fresh snow. Consequently, densification rates reduce with depth. These trends align with the trends observed in earlier models such as van den Broeke (2008) and Lightenberg and others (2011). In the vicinity of the major ice shelves, such as the Ross, Filchner-Ronne, Larsen and Amery Ice Shelves , z_{550} ranges from 5 to 13 m in FirnLearn, a close range to both van den Broeke (2008) and

Lightenberg and others (2011), while z_{830} ranges from 60 to 90 m in FirnLearn, 50 to 70m in van den Broeke (2008) and Lightenberg and others (2011). The discrepancy in z_{830} values may stem from limited data availability at these depths. Figure 5b compares the observed and modelled z_{550} for 220 locations with densities beyond 550 kg m^{-3}, while 5d compares the observed and modelled z_{830} for 88 locations with densities beyond 830 kg m^{-3}. For z_{550}, there is a clustering of points along the line of perfect agreement, particularly in the mid-range observed depth values (5-15 m), indicating strong predictive accuracy in this range. FirnLearn exhibits slightly tighter clustering compared to HL80, as reflected in its lower RMSE value (4.1 m vs 4.3 m for HL80). However, at depths below 5 m, there are more points lying above the upper confidence interval for both FirnLearn and HL, indicating that both models typically overestimate the depth values along the coast where the conditions are warmer and wetter, and accumulation rates are higher. This is possibly due to an underestimation of surface density, causing the models to densify slower than in observations. At greater depths, FirnLearn performs better than HL with more HL values lying below the line of perfect agreement. For z_{830}, there is a similar trend with more points above the upper confidence interval for lower to mid-range observed depth values (0-60 m), and a cluster around the line of perfect agreement for the remaining points. FirnLearn typically overpredicts z_{830} as compared to HL, which as mentioned earlier may be a result of the sparsity of data at deeper depths.

Firn Air Content

We explore predictions of FAC, the amount of air-filled pore space within the firn layer, using FirnLearn and HL, and compared them to the FAC from observations. FAC is an important parameter as it is essential for deriving the mass balance estimates of an ice sheet (Helsen and others, 2008; Horlings and others, 2020) improves our understanding and estimates of gas exchange dynamics and climate records (Trudinger and others, 2002; Buizert and others, 2013). It is also essential for estimating meltwater content of ice sheets, hence ice sheet contribution to sea level rise (Medley and others, 2022). To facilitate comparison between FirnLearn and HL, we employed the surface density predictions generated by FirnLearn as the surface density conditions for HL. The FAC is calculated by integrating porosity over the depth of the firn column and is represented as:

$$\mathrm{FAC} = \int_{z_l}^{z_u} \frac{\rho_{ice} - \rho(z)}{\rho_{ice}} dz \tag{11}$$

where ρ_{ice} is the density of ice (917 kg m^{-3}) and $\rho(z)$ is the firn density at a given depth, and the depth interval is set by an upper bound depth z_u and a lower bound depth $z_l = 0$, representing the surface.

and 830 depth.jpg and 830 depth.bb

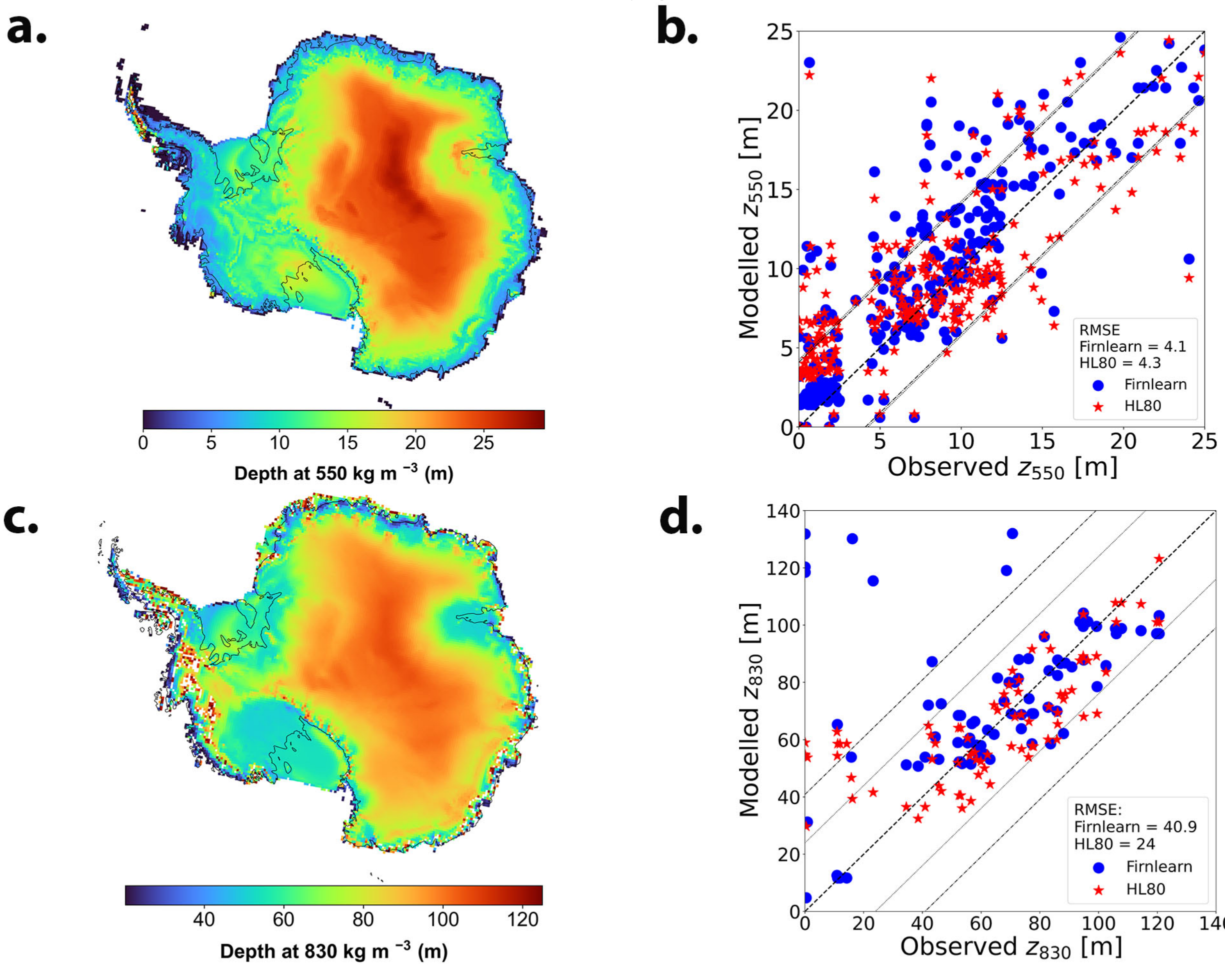

Fig. 5. (a) The predicted depth at 550 kg m^{-3} in meters (b) Comparison of modelled to observed depth at 550 kg m^{-3} (c) The predicted depth at 830 kg m^{-3} in meters (d) Comparison of modelled to observed depth at 830 kg m^{-3}. Here the FirnLearn computed surface density is used for the Herron and Langway (1980) model.

As depicted in figure 6a and table 1, the majority of density cores used in this study are shallow (below 10 m), indicating correspondingly low observed FAC values in figure 6b. Hence, for a direct comparison between the modeled (FirnLearn and HL), and observed FAC, we evaluated the FAC of each core up to its respective maximum depth from SUMup. This results in the difference in FirnLearn's and HL's evaluation of FAC being a reflection of their accuracy in predicting the densities in the first stage of densification. Very little difference is visually observed between the observed FAC and FirnLearn's and HL's predicted FAC (Figures 6 b, c and d). Figures 6 e and f depict the relative bias between FirnLearn's FAC and the observed FAC, and between HL's FAC and the observed FAC respectively. Given that we used FirnLearn's surface density as the boundary condition for HL's FAC calculations, FirnLearn's FAC bias values are similar to HL's FAC bias values, with a bulk relative bias of -5.5% for HL and -5.7% for FirnLearn. The root mean squared error however, shows a better performance in FirnLearn than in HL. In west Antarctica where we have deeper cores, more variation is observed. Specifically, FirnLearn slightly overestimates FAC in these deeper cores, while HL slightly underestimates FAC. To obtain a broader representation of the full firn column, we calculated FAC over a wider accumulation rate (0 to 3 m.w.e.yr $^{-1}$) and temperature (215 to 273 K [-58 to 0°C]) parameter space from the surface to 100 m depth. It is worth noting that some of these conditions are extreme in Antarctica. The heat maps shown in figure 7 depict the firn air content from FirnLearn, Herron and Langway (1980), and the difference between the two. As shown in these figures, FirnLearn and HL produce similar FAC patterns, with FAC being highest at low temperatures, and lowest at low accumulation rates and high temperatures. Relating this to ice sheets, FAC is predicted by both models to be approximately 30 - 50 m on the interior, where accumulation rates and temperatures are low (<0.5 m.w.e yr^{-1} and <225 K [-48°C] respectively), and in coastal regions where accumulation rates could be as high as 2 m.w.e yr^{-1}, and temperatures could be higher than 270 K [-23°C]. In West Antarctica, with accumulation rates between 0.5 and 2 m.w.e yr^{-1}, and temperatures higher than 250K [-23°C, FAC is predicted by FirnLearn and HL to be greater than 40 m.

Figure 7c shows that on average, FirnLearn's FAC is in agreement with HL's FAC, with minor disagreements where FirnLearn underpredicts when compared to HL. The regions with the highest positive differences (FirnLearn $\gg$ Herron and Langway (1980)) are at lower accumulation rates, as indicated by the red hues. Conversely, the regions with the highest negative differences (FirnLearn $\ll$ Herron and Langway (1980)) are at mid to higher accumulation rates, as indicated by the blue hues, a region which coincides with the parameter space of the training data. It is worth noting that conditions where accumulation rates

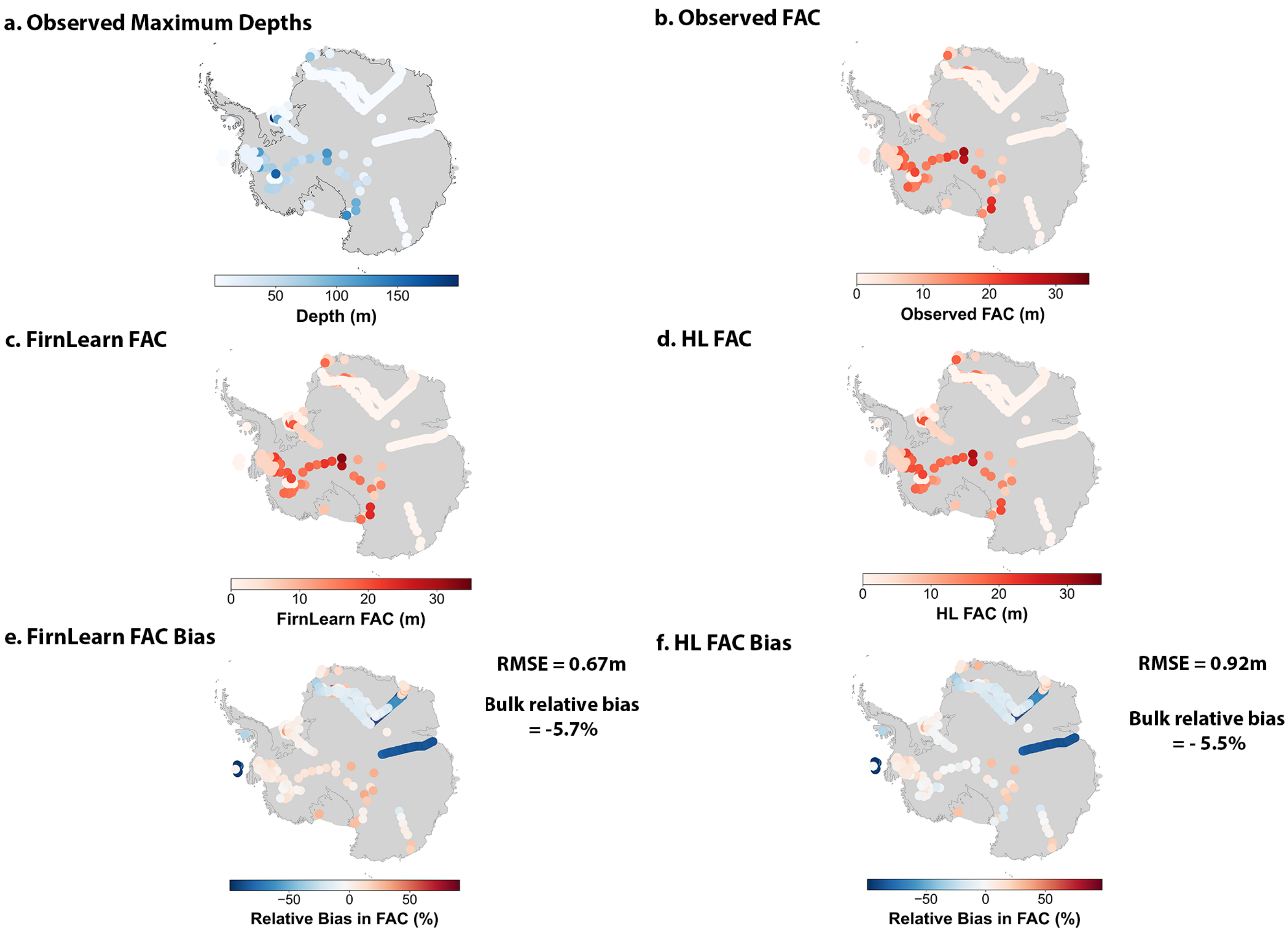

Fig. 6. Firn air content across Antarctica, comparing models to observations and assessing bias: (a) Spatial distribution of 2689 SUMup cores, with shading denoting core depth, (b) Observed FAC from calculated from the densities of the SUMup cores(c) FAC in m, calculated with FirnLearn (d) FAC in m, calculated with Herron and Langway (1980) (e) Relative bias between the FAC calculated with FirnLearn and the observed FAC and (d) Relative bias between the FAC calculated using Herron and Langway (1980) and the observed FAC.

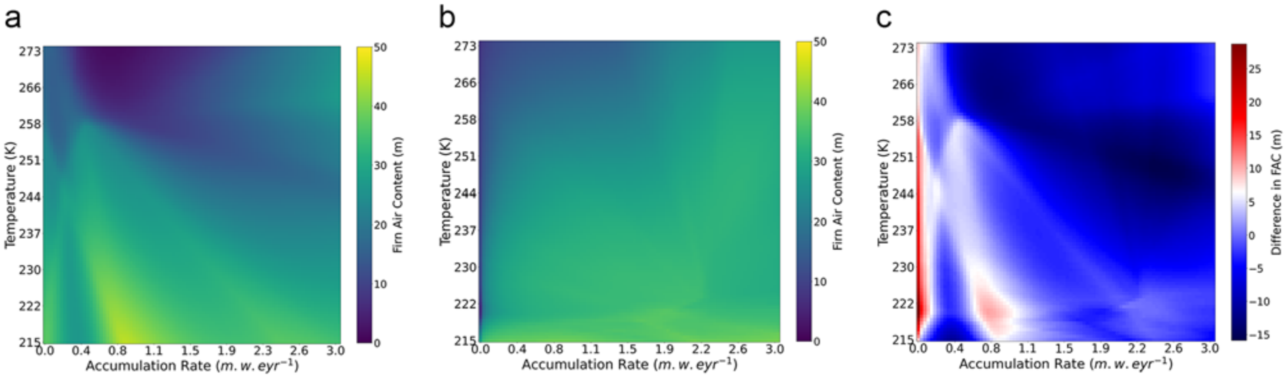

Fig. 7. (a) FAC in m, calculated with FirnLearn, (b) firn air content in meters, calculated with Herron and Langway (1980) (c) Difference in FAC in m, between the FAC calculated using FirnLearn and the FAC calculated using Herron and Langway (1980). The difference is presented as FirnLearn minus HL80.

are very low ($<$ 1 m.w.e. yr $^{-1}$) and temperatures are very high ($>$ 260K [-13°C]) or where accumulation rates are very high ($>$ 1 m.w.e. yr $^{-1}$) and temperatures are very low ($<$ 230K [-43°C]) rarely exist in Antarctica, at least not within its current climate regime. Figure 7 is shown in order to understand FAC values within a wider accumulation rate and temperature parameter space.

LIMITATIONS TO FIRNLEARN

Despite its promising performance, FirnLearn has limitations due to data quality and quantity. As shown in figure 1, the spatial distribution of density observations is notably limited, particularly in East Antarctica. Additionally, as shown in figure 6a, the majority of density observations in the dataset are concentrated at shallow depths. Consequently, the discrepancies between FirnLearn's density predictions and observations increase as depth increases, as evident by the higher RMSE in the predictions of depth at 830 kg m^{-3} compared to the predictions of depth at 550 kg m^{-3} (Fig. 5). Also, as seen in tables 2,3, and 4, FirnLearn still underperforms HL80 in certain conditions. For these conditions, for instance, regions with sparse-depth observations (>830 kg m^{-3}), using a weighted combination of HL80 and FirnLearn predictions could provide better results. This approach would leverage the strengths of both methods, compensating for FirnLearn's inaccuracies at greater depths with HL80's depth performance, while benefiting from FirnLearn's surface density predictions and better performance in stage 2 (550 <= ρ <830 kg m^{-3}). With the growth of the SUMup dataset as more firn data is collected, FirnLearn's density predictions are bound to improve, as it is trained on new data.

FirnLearn is also limited to a steady-state assumption, therefore unable to predict temporal firn density evolution. Density observations from SUMup are collected over several years, and at different periods of the year, leading to knowledge gaps regarding seasonal variability in firn properties. Without explicit time-dependent inputs, FirnLearn struggles to generalize to evolving firn conditions. However, with additional training on newer datasets that include temporal markers, its ability to capture temporal firn density evolution is expected to improve.

It is also worth noting that while FirnLearn's surface density predictions are in line with surface density predictions from (Kaspers and others, 2004; van den Broeke, 2008; Lightenberg and others, 2011), its predictions may still be less reliable in parts of East Antarctica due to the limited number of data points in that region.

Another challenge lies in the lack of interpretability of deep learning models like FirnLearn. These models are effectively 'black boxes', such that it is difficult to understand the underlying processes governing model predictions. However, given the black-box nature, ANNs serve as effective tools in contexts where predictive accuracy outweighs model interpretability, which is likely the case for depth-density profiles in Antarctica at this time. The improved accuracy offered by ANNs holds the potential to produce improved parameters for understanding firn densification physics.

CONCLUSIONS

In this study, we introduced FirnLearn, a new steady-state densification model for the Antarctic firn layer based on deep learning of data from observations and output from the regional atmospheric climate model. Comparison with observations highlights FirnLearn's improved predictability of firn density at intermediate depths, where discrepancies between observations and traditional models, such as Herron and Langway (1980) are more pronounced. In addition, we have used FirnLearn to accurately derive surface density, depth at 550 kg m^{-3} and 830 kg m^{-3} (pore close-off), and firn air content across Antarctica. This study demonstrates the potential of deep learning techniques in improving Antarctic firn density estimates, and strengthens the promising foundation for the development of a generally applicable firn model. In the future, we plan to expand this model by applying it to the Greenland Ice Sheet and coupling it to physics to develop a Physics Informed Neural Network (PINN) which can be applied to both dry and wet firn densification.

DATA AVAILABILITY

FirnLearn's code is available at `https://github.com/ayobamiogunmolasuyi/FirnLearn`. The repository contains all the scripts used to train the models and produce the plots and results. The SUMup dataset is available at `https://github.com/MeganTM/SUMMEDup` while the racmo dataset is here `https://doi.pangaea.de/10.1594/PANGAEA.896940`

SUPPLEMENTAL MATERIAL

The supplement to this article is attached.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation (AO, GRFP-2021295396; CRM, 2024132; IB, 1851094), the Army Research Office (CRM, 78811EG), the National Aeronautics and Space Administration (CRM, EPSCoR-80NSSC21M0329). We thank Brice Noel for his assistance in accessing the RACMO data.

REFERENCES

Alley R (1987) Firn densification by grain-boundary sliding: A first model. *Le Journal De Physique Colloques*, **48**(C1) (doi: 10.1051/jphyscol:1987135)

Alley RB (2000) *The two-mile time machine: Ice cores, abrupt climate change, and our future*. Princeton University Press

Anderson D and Benson C (1963) The densification and diagenesis of snow. *MIT Press*

Arnaud L, Barnola JM and Duval P (2000) Physical modeling of the densification of snow/firn and ice in the upper part of polar ice sheets, in: Physics of ice core records. *Hokkaido University Press*

Arthern RJ, Vaughan DG, Rankin AM, Mulvaney R and Thomas ER (2010) In situ measurements of antarctic snow compaction compared with predictions of models. *Journal of Geophysical Research: Earth Surface*, **115**(F3) (doi: 10.1029/2009jf001306)

Bader H (1954) Sorges law of densification of snow on high polar glaciers. *Journal of Glaciology*, **2**

Bader H (1965) Theory of densification of dry, bubbly glacier ice. *US Cold Regions Research and Engineering Laboratory. Research Report 141*

Baker I and Ogunmolasuyi AO (2024) Mechanical properties of freshwater ice. *The Journal of Physical Chemistry C*, **128**(51), 2160921626 (doi: 10.1021/acs.jpcc.4c05171)

Barnola J, Pimienta P, Raynaud D and YS K (1991) Co2-climate relationship as deduced from the vostok ice core: A re-examination based on new measurements and on a re-evaluation of the air dating. *Tellus B*, **43**(2) (doi: 10.1034/j.1600-3900889.1991.t01-1-00002.x)

Bartelt P and Lehning M (2002) A physical snowpack model for the swiss avalanche warning. *Cold Regions Science and Technology*, **35**(3), 123145 (doi: 10.1016/s0165-232x(02)00074-5)

Bolibar J, Rabatel A, Gouttevin I, Galiez C, Condom T and Sauquet E (2020) Deep learning applied to glacier evolution modelling. *The Cryosphere*, **14**, 565–584

Brinkerhoff D, Aschwanden A and Fahnestock M (2021) Constraining subglacial processes from surface velocity observations using surrogate-based bayesian inference. *Journal of Glaciology*, **67**(263), 385403 (doi: 10.1017/jog.2020.112)

Buizer C (2013) Ice core methods | studies of firn air. *Encyclopedia of Quaternary Science*, 361–372

Buizert C, Martinerie P, Petrenko VV, Severinghaus JP, C M Trudinger EW, Rosen JL, Orsi AJ, Rubino M, Etheridge DM, Steele LP, Hogan C, Laube JC, Sturges WT, Levchenko VA, Smith AM, Levin I, Conway TJ, Dlugokencky EJ, Lang PM, Kawamura K, Jenk TM, White JWC, Sowers T, Schwander J and Blunier T (2012) Gas transport in firn: multiple-tracer characterisation and model intercomparison for neem, northern greenland. *Atmospheric Chemistry and Physics*

Buizert C, Sowers T and Blunier T (2013) Assessment of diffusive isotopic fractionation in polar firn, and application to ice core trace gas records. *Earth and Planetary Science Letters*, **361**, 110119 (doi: 10.1016/j.epsl.2012.11.039)

Burr A, Ballot C, Lhuissier P, Martinerie P, Martin CL and Philip A (2018) Pore morphology of polar firn around closure revealed by x-ray tomography. *The Cryosphere*, **12**(7), 24812500 (doi: 10.5194/tc-12-2481-2018)

Camps-Valls G, Reichstein M, Zhu X and Tuia D (2020) Advancing deep learning for earth sciences: From hybrid modeling to interpretability. *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium* (doi: 10.1109/igarss39084.2020.9323558)

Cuffey KM and Paterson WSB (2010) *The Physics of Glaciers, 4th edition*

Dell RL, Banwell AF, Willis IC, Arnold NS, Halberstadt AR, Chudley TR and Pritchard HD (2022) Supervised classification of slush and ponded water on antarctic ice shelves using landsat 8 imagery corrigendum. *Journal of Glaciology*, **68**(268), 415416 (doi: 10.1017/jog.2022.15)

Dunmire D, Banwell AF, Wever N, Lenaerts JT and Datta RT (2021) Contrasting regional variability of buried meltwater extent over 2 years across the greenland ice sheet. *The Cryosphere*, **15**(6), 29833005 (doi: 10.5194/tc-15-2983-2021)

Dunmire D, Wever N, Banwell A and Lanearts J (2024) Antarctic-wide ice-shelf firn emulation reveals robust future firn air depletion signal for the antarctic peninsula. *Nature Communications Earth & Environment* (doi: 10.1038/s43247-024-01255-4)

Forster RR, Box JE, van den Broeke MR, Miège C, Burgess EW, van Angelen JH, Lenaerts JT, Koenig LS, Paden J, Lewis C and et al (2013) Extensive liquid meltwater storage in firn within the greenland ice sheet. *Nature Geoscience*, **7**(2), 9598 (doi: 10.1038/ngeo2043)

Freitag J, Wilhelms F and Kipfstuhl S (2004) Microstructure-dependent densification of polar firn derived from x-ray microtomography. *Journal of Glaciology*, **50**(169), 243250 (doi: 10.3189/172756504781830123)

Gow AJ (1969) On the rates of growth of grains and crystals in south polar firn. *Journal of Glaciology*, **8**(53), 241252 (doi: 10.1017/s0022143000031233)

Harper J, Humphrey N, Pfeffer WT, Brown J and Fettweis X (2012) Greenland ice-sheet contribution to sea-level rise buffered by meltwater storage in firn. *Nature*, **491**(7423), 240243 (doi: 10.1038/nature11566)

Hatie T, R T and J F (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer

Helsen MM, van den Brooke MR, van de Wal RS, van de Berg WJ and van Meijgaard E (2008) Elevation changes in antarctica mainly determined by accumulation variability. *Science*, **320**

Herron MM and Langway CC (1980) Firn densification: An empirical model. *Journal of Glaciology*, **25**(93), 373385 (doi: 10.3189/S0022143000015239)

Horlings AN, Christianson K, Holschuh N, Stevens CM and Waddington ED (2020) Effect of horizontal divergence on estimates of firn-air content. *Journal of Glaciology*, **67**(262), 287296 (doi: 10.1017/jog.2020.105)

Hörhold M, Laepple J, Freitag J, Bigler M, Fischer H and Kipfstuhl S (2012) On the impact of impurities on the densification of polar firn. *Earth and Planetary Science Letters*, **325-326**, 93–99 (doi: https://doi.org/10.1016/j.epsl.2011.12.022)

Kaspers KA, van de Wal RSW, van den Broeke MR, Schwander J, van Lipzig NPM and Brenninkmeijer CAM (2004) Model calculations of the age of firn air across the antarctic continent. *Atmospheric Chemistry and Physics*, **5** (doi: 10.5194/acp-4-1365-2004,2004)

Kingma DP and Ba J (2017) Adam: A method for stochastic optimization

Kingslake J, Skarbek R, Case E and McCarthy C (2022) Grain-size evolution controls the accumulation dependence of modelled firn thickness. *The Cryosphere*, **16**, 34133430 (doi: https://doi.org/10.5194/tc-16-3413-2022)

Kipfstuhl S, Faria SH, Azuma N, Freitag J, Hamann I, Kaufmann P, Miller H, Weiler K and Wilhelms F (2009) Evidence of dynamic recrystallization in polar firn. *Journal of Geophysical Research: Solid Earth*, **114**(B5) (doi: 10.1029/2008jb005583)

Li J and Zwally H (2011) Modeling of firn compaction for estimating ice-sheet mass change from observed ice-sheet elevation change. *Annals of Glaciology*, **52**(59) (doi: 10.3189/172756411799096321)

Li W, Veldhuijsen SB and Lhermitte S (2023) Machine learning of antarctic firn density by combining radiometer and scatterometer remote sensing data. *EGUsphere[preprint]* (doi: 10.5194/egusphere-2023-1556)

Li Y and Baker I (2021) Observations of the creep of polar firn. *Journal of Glaciology*, **68**(268), 269287 (doi: 10.1017/jog.2021.91)

Lightenberg S, Helsen M and den Broeke MV (2011) An improved semi-empirical model for the densification of antarctic firn. *The Cryosphere*, **5**(4)

Ligtenberg SR, Helsen MM and van den Broeke MR (2011) An improved semi-empirical model for the densification of antarctic firn. *The Cryosphere*, **5**(4), 809819 (doi: 10.5194/tc-5-809-2011)

Lomonaco R, Albert M and Baker I (2011) Microstructural evolution of fine-grained layers through the firn column at summit, greenland. *Journal of Glaciology*, **57**(204), 755762 (doi: 10.3189/002214311797409730)

Lundin JM, STEVENS CM, ARTHERN R, BUIZERT C, ORSI A, LIGTENBERG SR, SIMONSEN SB, CUMMINGS E, ESSERY R, LEAHY W and et al (2017) Firn model intercomparison experiment (firnmice). *Journal of Glaciology*, **63**(239), 401422 (doi: 10.1017/jog.2016.114)

Maeno N and Ebinuma T (1983) Pressure sintering of ice and its implication to the densification of snow at polar glaciers and ice sheets. *The Journal of Physical Chemistry*, **87**(21), 41034110 (doi: 10.1021/j100244a023)

McDowell IE, Albert MR, Lieblappen SA and Keegan KM (2020) Local weather conditions create structural differences between shallow firn columns at summit, greenland and wais divide, antarctica. *Atmosphere*, **11**(12), 1370 (doi: 10.3390/atmos11121370)

Medley B, Neumann T, Zwally H, Smith B and Stevens C (2022) Simulations of firn processes over the greenland and antarctic ice sheets: 19802021. *The Cryosphere*, **16**(10)

Meyer CR and Hewitt IJ (2017) A continuum model for meltwater flow through compacting snow. *Cryosphere*, **11**(6), 2799–2813 (doi: 10.5194/tc-11-2799-2017)

Meyer CR, Keegan KM, Baker I and Hawley RL (2020) A model for french-press experiments of dry snow compaction. *The Cryosphere*, **14**(5), 14491458 (doi: 10.5194/tc-14-1449-2020)

Montgomery L, Koenig L, Alexander P and et al (2018) The sumup dataset: Compiled measurements of surface mass balance components over ice sheets and sea ice with analysis over greenland. *Earth Syst. Sci. Data*, **10**(4), 1959–1985

Morris E and Wingham DJ (2014) Densificatio of polar snow: Measurements, modeling, and implications for altimetry. *Jounal of Geophysical Research-Earth Surface*, **119** (doi: 10.1002/2013JF002898)

Noël B, van de Berg WJ, van Wessem JM, van Meijgaard E, van As D, Lenaerts JTM, Lhermitte S, Munneke PK, Smeets CJPP, van Ulft LH, van de Wal RSW and den Broeke MRV (2018) Modelling the climate and surface mass balance of polar ice sheets using racmo2 part 1: Greenland (19582016). *The Cryosphere*

Nussbaumer S, Steiner D and Zumbühl H (2012) Réseau neuronal et fluctuations des glaciers dans les alpes occidentales

Ogunmolasuyi A, Murdza A and Baker I (2023) The onset of recrystallization in polar firn. *Geophysical Research Letters*, **50**(23) (doi: 10.1029/2023gl103435)

O'Shea K and Nash R (2015) An introduction to convolutional neural networks

Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N and Prabhat (2019) Deep learning and process understanding for data-driven earth system science. *Nature*, **566**(7743), 195204 (doi: 10.1038/s41586-019-0912-1)

Rizzoli P, Martone M, Rott H and Moreira A (2017) Characterization of snow facies on the greenland ice sheet observed by tandem-x interferometric sar data. *Remote Sensing*, **9**(4), 315 (doi: 10.3390/rs9040315)

Salamatin AN, Lipenkov VY and Duval P (1997) Bubbly-ice densification in ice sheets: I. theory. *Journal of Glaciology*, **43**(145), 387396 (doi: 10.3189/S0022143000034961)

Smith B, Fricker HA, Gardner AS, Medley B, Nilsson J, Paolo F, Holschuh N, Adusumilli S, Brunt K, Csatho B, Harbeck K, Markus T, Neumann T, Siegfried MR and Zwally H (2020) Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes. *Science*, **368** (doi: 10.5194/gmd-13-4355-2020)

Steffen K and Box J (2001) Surface climatology of the greenland ice sheet: Greenland climate network. *Journal of Geophysical Research - Atmospheres*, **106**

Steiner D, Walter A and Zumbühl H (2005) The application of a non-linear back-propagation neural network to study the mass balance of grosse aletschgletscher, switzerland. *Journal of Glaciology*, **51**(173), 313323 (doi: 10.3189/172756505781829421)

Stevens CM, Verjans V, Lundin JM, Kahle EC, Horlings AN, Horlings BI and Waddington ED (2020) The community firn model (cfm) v1.0. *Geoscientific Model Development*, **13**(9), 43554377 (doi: 10.5194/gmd-13-4355-2020)

Stevens CM, Lilien DA, Conway H, Fudge TJ, Koutnik MR and Waddington ED (2023) A new model of dry firn-densification constrained by continuous strain measurements near south pole. *Journal of Glaciology*, 115 (doi: 10.1017/jog.2023.87)

The-Firn-Symposium-Team (2024) Firn on ice sheets. *Nature Reviews Earth & Environment* (doi: 10.1038/s43017-023-00507-9)

Thompson-Munson M, Wever N, Stevens CM, Lenaerts JT and Medley B (2023) An evaluation of a physics-based firn model and a semi-empirical firn model across the greenland ice sheet (19802020). *The Cryosphere*, **17**(5), 21852209 (doi: 10.5194/tc-17-2185-2023)

Trudinger CM, Etheridge DM, Rayner PJ, Enting IG, Sturrock GA and Langenfelds RL (2002) Reconstructing atmospheric histories from measurements of air composition in firn. *Journal of Geophysical Research: Atmospheres*, **107**(D24) (doi: 10.1029/2002jd002545)

Van den Broeke M (2008) Depth and density of the antarctic firn layer. *J. Neurosci.*, **40**(2), 432–438

van den Broeke M (2008) Depth and density of the antarctic firn layer. *Arctic, Antarctic, and Alpine Research*, **40** (doi: 10.1657/1523-0430(07-021))

Van Wessem J, C R, Morlighem M, Mouginot J, Rignot E, Medley B, Joughin I, Wouters B, Depoorter MA, Bamber JL, Lenaerts JTM, Van De Berg WJ, Van Den Broeke MR and & Van Meijgaard E (2014) Improved representation of east antarctic surface mass balance in a regional atmospheric climate model. *Journal of Glaciology* (doi: 10.3189/2014jog14j051)

Vandecrux B, Amory C, Ahlstrøm AP, Akers PD, Albert M, Alley RB, Alves de Castro M, Arnaud L, Baker I, Bales R, Benson C, Box JE, Brucker L, Buizert C, Chandler D, Charalampidis C, Cherblanc C, Clerx N, Colgan W, Covi F, Dattler M, Denis G, Derksen C, Dibb JE, Ding M, Dixon D, Eisen O, Fahrner D, Fausto R, Favier V, Fernandoy F, Freitag J, Gerland S, Harper J, Hawley RL, Heuer J, Hock R, Hou S, How P, Humphrey N, Hubbard B, Iizuka Y, Isaksson E, Kameda T, Karlsson NB, Kawakami K, Kjær HA, Kreutz K, Kuipers Munneke P, Lazzara M, Lemeur E, Lenaerts JTM, Lewis G, Ley Lindau FG, Lindsey-Clark J, MacFerrin M, Machguth H, Magand O, Mankoff KD, Marquetto L, Martinerie P, McConnell JR, Medley B, Miège C, Miles KE, Miller O, Miller H, Montgomery L, Morris E, Mosley-Thompson E, Mulvaney R, Niwano M, Oerter H, Osterberg E, Otosaka I, Picard G, Polashenski C, Reijmer C, Rennermalm A, Rutishauser A, Scanlan K, Simoes JC, Simonsen SB, Smeets PCJP, Smith A, Solgaard A, Spencer M, Steen-Larsen HC, Stevens CM, Sugiyama S, Svensson J, Tedesco M, Thomas E, Thompson-Munson M, Tsutaki S, van As D, Van den Broeke MR, van Tiggelen M, Wang Y, Wilhelms F, Winstrup

M, Xiao J and Xiao C (2024) The sumup collaborative database: Surface mass balance, subsurface temperature and density measurements from the greenland and antarctic ice sheets (2024 release). *Arctic Data Center* (doi: 1https://doi.org/0.18739/A2M61BR5M)

Verjans V, Leeson AA, Nemeth C, Stevens CM, Kuipers Munneke P, Noël B and van Wessem JM (2020) Bayesian calibration of firn densification models. *The Cryosphere*, **14**(9), 30173032 (doi: 10.5194/tc-14-3017-2020)