https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

Emulating high-resolution distributions of snow depth and density on Arctic sea ice

Ayush Prasad[1], Ioanna Merkouriadi[1], Aleksi Nummelin[2], Jack Landy[3], Arttu Jutila[1], and Glen E. Liston[4]

[1]Earth Observation Research, Finnish Meteorological Institute, Helsinki, Finland
[2]Marine Research, Finnish Meteorological Institute, Helsinki, Finland
[3]Earth Observation Group, Department of Physics and Technology, UiT The Arctic University of Norway, Tromsø, Norway
[4]Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins, CO, USA

Correspondence: Ayush Prasad (ayush.prasad@fmi.fi)

Abstract. Snow on Arctic sea ice significantly influences regional and global climate processes, yet accurately modelling its spatial and temporal variability remains challenging. We introduce a physics-guided Long Short-Term Memory (LSTM) emulator developed to efficiently simulate pan-Arctic snow density and depth distributions. The emulator is trained on high-resolution SnowModel-3D simulations that resolve wind-driven snow redistribution over sea-ice surface topography, and is applied over the Lagrangian parcel framework of SnowModel-LG for pan-Arctic inference. The topographic input required for inference comes from CryoSat-2 surface roughness estimates at a pan-Arctic scale. We assessed the emulator's predictive accuracy across multiple years (2014–2021) and various seasonal conditions. Compared with airborne snow observation campaigns, the emulated snow depth has an RMSE of 0.08–0.10 m. By representing the influence of sea-ice topography on snow redistribution, which is not accounted for in SnowModel-LG, the emulator adds a process that brings the simulated snow depth closer to these airborne observations. While seasonal and interannual trends in snow density and depth were effectively captured, emulator accuracy varied during complex melt periods. Our emulator provides an efficient tool for enhancing pan-Arctic snow modelling capabilities in climate studies.

1 Introduction

Snow on sea ice plays a crucial role in the climate system due to its unique thermal and optical properties (Sturm et al., 2002; Perovich et al., 2015) and understanding the spatial distribution and temporal evolution of snow on sea ice is essential for climate modelling and remote sensing applications. For example, accurate retrieval of sea-ice thickness from satellite altimetry requires knowledge of snow depth and density. Uncertainties in snow cover properties can lead to sea ice thickness errors of up to 50 % (Carret et al., 2025; Landy et al., 2020). Historically, however, data on Arctic snow cover have been sparse and large-scale climate models and observations have struggled to capture its variability.

Snow on sea ice is difficult to represent in models due to several challenges. First, snow accumulation and redistribution processes create highly heterogeneous snow distributions at small scales (Sturm et al., 2002; Liston et al., 2018). Snow depth can vary by tens of centimetres within just a few metres because of wind-drift, ridging of sea ice and other factors. Such fine-scale variability is well beyond the resolution of typical climate models or satellite sensors, which operate at tens to hundreds

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

of kilometres, leading to a loss of detail. Second, the remote Arctic environment limits in situ observations, making it hard to constrain or validate models. Field measurements of snow on sea ice are infrequent and localised to specific parts of the Arctic, so model forcing and evaluation often rely on imperfect data. Third, the complexity of processes of snow on sea ice is a challenge. An accurate model must simulate snowfall accumulation, wind packing and blowing-snow redistribution, snow density evolution, melt–freeze cycles and formation of snow ice. Capturing all these processes with high fidelity requires detailed physics and high temporal resolution. In addition, numerical models are often forced with atmospheric reanalysis data, in which winter precipitation remains a known challenge, potentially introducing substantial uncertainties into the simulations. Finally, there is a significant computational constraint. High-resolution, process-rich models are computationally expensive to run over large areas or long time periods. The Arctic sea-ice zone spans millions of square kilometres and resolving snow processes at even 1 km scale would be prohibitive in a long-term simulation. As a result, current large-scale models use simplifications that sacrifice physical detail. Many climate and operational sea-ice models treat the snow cover with minimal layers or use climatological assumptions (e.g. a fixed snow density) rather than dynamic simulation. One widely used snow climatology, for example, is the Warren et al. (1999) dataset, based on outdated drifting station data. For decades this was the primary pan-Arctic snow depth reference. While simple to implement in, e.g., sea-ice thickness retrievals from satellite altimetry, such approaches cannot capture year-to-year or regional variations and may be increasingly inaccurate under changing sea-ice conditions.

To improve trustworthiness, specialised physics-based snow models have been developed for Arctic conditions. SnowModel (Liston and Elder, 2006) is one such state-of-the-art process model that simulates snow evolution and distribution in a multilayer framework. In its full 3D configuration (hereafter SnowModel-3D), it can simulate snow redistribution by wind over detailed sea-ice surface topography at metre-scale horizontal resolution (Liston et al., 2018). However, a major drawback is the intense computational demand, which makes running SnowModel-3D at high resolution over the entire Arctic impractical. To address this, a Lagrangian variant, SnowModel-LG, was developed specifically for pan-Arctic simulations (Liston et al., 2020). SnowModel-LG uses a Lagrangian approach in which individual 1D SnowModel instances are used to simulate snow density and depth on sea-ice floes along their satellite-derived tracks. However, compared to SnowModel-3D, it operates at coarser effective resolution and it lacks the lateral processes, such as snow redistribution, and does not account for sea-ice surface topography. SnowModel-LG predicts more textured patterns of snow depth variability across the Arctic than other model-based or remotely sensed snow depth products (Zhou et al., 2021), but it is not clear whether these are more or less realistic.

Besides SnowModel, other modelling efforts exist with varying complexity. For example, the NASA Eulerian Snow on Sea Ice Model (NESOSIM) (Petty et al., 2018) provides daily estimates of snow depth and density across the Arctic Ocean by modelling snow accumulation and redistribution in a simplified framework. NESOSIM uses two vertical layers and parameterised processes (accumulation, wind packing, blowing snow losses, etc.), forced by reanalysis snowfall and winds and satellite-derived ice drift, to simulate the winter snow budget on sea ice. This model runs efficiently on a coarse grid ($\sim$100 km) but, by design, omits finer-scale processes and does not explicitly simulate the spring/summer melt season.

In this paper, we address the gap between detailed but expensive SnowModel-3D simulations and efficient but simplified pan-Arctic models by extending our previous work (Prasad et al., 2024). In Prasad et al. (2024), we developed and evaluated a physics-guided Long Short-Term Memory (LSTM) network emulator trained on SnowModel-3D outputs for five regional

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

Arctic domains. Here, we extend that emulator to the full pan-Arctic scale by adopting the Lagrangian parcel framework of SnowModel-LG and incorporating satellite-derived sea-ice surface roughness from CryoSat-2 as a topographic input.

The emulator predicts the snow state for each parcel through time using an autoregressive, parcel-based inference logic. At each time step, the LSTM uses its own prior prediction as input for the next step, preserving temporal dependencies (e.g. the accumulation from one day affecting the next) in a manner analogous to a physical model. The parcel-based design handles spatial heterogeneity naturally: each parcel experiences its own meteorological forcing, drift trajectory and topographic conditions, so the emulator adapts to regional differences. By training the emulator on SnowModel-3D simulations from five diverse Arctic regions and applying it over the SnowModel-LG parcel trajectories, we enable pan-Arctic snow estimation that incorporates the subgrid-scale snow redistribution process, and the impact of sea-ice surface topography on it, which are both absent in SnowModel-LG itself.

2 Methodology

In this study, we develop a forecasting framework that integrates high-resolution snow simulations with a moving parcel approach and a physics-guided Long Short-Term Memory (LSTM) network to predict the evolution of snow density and depth on sea ice over extended time periods. Our approach builds on the methodology established in our previous work (Prasad et al., 2024), where we developed and evaluated a physics-guided LSTM emulator for SnowModel at regional scales. In this paper, we extend that framework to the full pan-Arctic domain. Figure 1 summarises the overall pipeline. To avoid confusion, we distinguish between two variants of SnowModel used in this work:

- **SnowModel-3D** refers to the original, high-resolution version of SnowModel (Liston and Elder, 2006), which simulates snow evolution and redistribution in a multilayer, spatially distributed framework including surface topography. We use SnowModel-3D at 1-m horizontal resolution over individual sea-ice floes to generate the training data for the emulator.

- **SnowModel-LG** refers to the Lagrangian variant (Liston et al., 2020), in which individual 1D SnowModel instances are used to simulate snow density and depth on sea ice floes along their satellite-derived tracks. SnowModel-LG simulates snow density and depth along these parcel trajectories.

The emulator is trained on SnowModel-3D outputs and applied over the pan-Arctic domain using the parcel framework similar to SnowModel-LG. The following sections describe the methodology in detail.

2.1 SnowModel-3D: Training Data Generation

To generate training data for the LSTM emulator, we ran SnowModel-3D at high spatial resolution over individual sea-ice floes (Fig. 1a), following the approach described in Prasad et al. (2024). SnowModel-3D simulates the full three-dimensional evolution of the snowpack, including accumulation, wind-driven redistribution, sublimation, densification and melt–freeze processes within a multilayer framework (Liston and Elder, 2006).

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

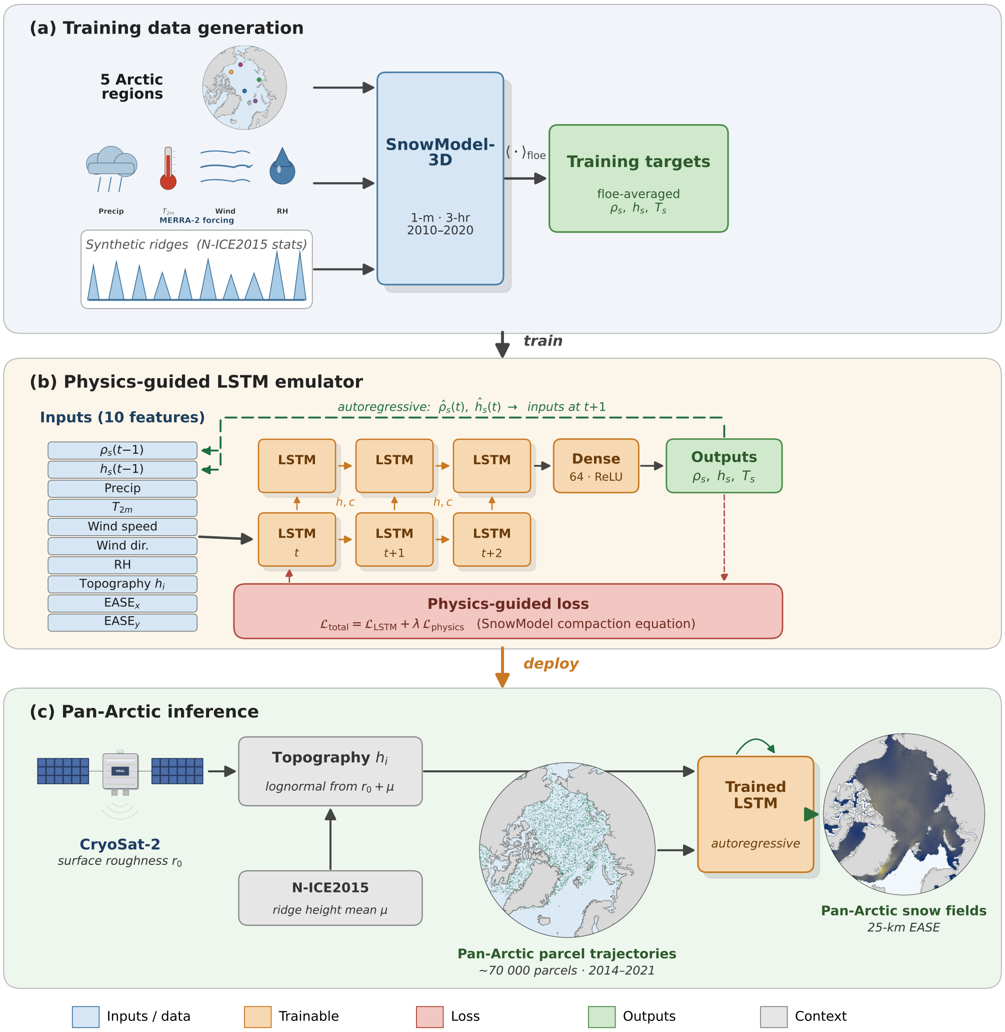


Figure 1. Overview of the physics-guided LSTM emulator framework. (a) SnowModel-3D is run at 1-m resolution over individual sea-ice floes in five Arctic regions, driven by MERRA-2 forcing and synthetic ridge topography generated from N-ICE2015 statistics, the floe-averaged outputs serve as training targets. (b) The physics-guided LSTM emulator takes ten parcel-level inputs (previous snow density and depth, precipitation, 2-m temperature, 10-m wind speed and direction, 2-m relative humidity, topography and the two EASE-grid coordinates) and predicts snow density, depth and temperature at each time step. The emulator is used autoregressively during inference, with the predicted snow density and depth at time t fed back as inputs at time $t+1$. A physics-guided loss augments the standard mean squared error (MSE) term with a compaction-equation residual from SnowModel. (c) For pan-Arctic inference, CryoSat-2 surface roughness combined with the mean log surface height from N-ICE2015 produces parcel-level topography. The trained emulator is then applied autoregressively along $\sim$70 000 annual SnowModel-LG parcel trajectories over 2014–2021 to produce pan-Arctic snow fields on the 25-km EASE grid.

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

To produce the training data for the emulator, SnowModel-3D was configured at 1-m horizontal resolution over individual floe domains of 1500×1500 m size, with a temporal resolution of 3 hours to capture diurnal variations. We used meteorological forcing from five climatologically distinct Arctic regions: the Central Arctic, Beaufort Sea, Chukchi Sea, Laptev Sea and Barents Sea. The simulation period was from August 2010 to August 2020, and daily outputs were saved for emulator training. The meteorological forcing was derived from NASA's Modern Era Retrospective Analysis for Research and Application Version 2 (MERRA-2; Gelaro et al., 2017), which provides 10-m wind speed and direction, 2-m air temperature, 2-m relative humidity and total water-equivalent precipitation.

A key feature of SnowModel-3D is its ability to simulate wind-driven snow redistribution over sea-ice surface topography. However, high-resolution topographic data for sea-ice surfaces are not widely available. In Prasad et al. (2024), we addressed this by synthetically generating sea-ice ridge topography based on observed ridge statistics. Specifically, we generated synthetic ridge fields by randomly placing ridges across each floe domain, drawing ridge heights, widths and spacings from the statistical distributions (minimum, maximum and standard deviation) reported in Sudom et al. (2011). Each ridge is represented as a triangular cross-section with a given sail height and base width, oriented at random angles across the domain. By varying the ridge density (number of ridges per unit area), we produce topographic fields needed to drive wind redistribution in SnowModel-3D.

Our aim is to capture the influence of hundred-metre-scale sea-ice topography on snow properties while creating an emulator designed for applications well above this local scale of the ice topography (>1 km). Therefore, we average the SnowModel-3D snow properties across each ice floe domain and train the emulator on these spatially averaged outputs, allowing it to learn the influence of metre-scale topography on snow properties as a subgrid-scale process.

2.2 Pan-Arctic Extension: Topography from CryoSat-2

For the pan-Arctic application, the emulator requires a topographic input for each Lagrangian parcel at each time step. Since observed high-resolution topography is not available at the pan-Arctic scale, we derive surface roughness estimates from CryoSat-2 satellite altimetry and convert these into topographic fields on the same physical scale as the ridge topography used to generate the training data (Fig. 1c). The CryoSat-2 roughness estimates are obtained from the Lognormal Altimeter Retracker Model radar waveform retracking algorithm applied to Level 1b observations from the ESA Baseline E ICE processor, as described in Landy et al. (2020) and Landy et al. (2026). Briefly, each detected CryoSat-2 waveform is fit with a physical model for the expected radar echo return from a rough sea-ice surface. The model assumes that the sea-ice surface dominates the reflected echo and that the surface can be closely represented with a lognormal height distribution (Landy et al., 2020). One of the parameters obtained from the radar model fit is the "macro-scale surface topography", r_0, which represents the standard deviation of the sea-ice surface over approximately the hundred-metre length scale of the sensor footprint. All valid sea-ice topography observations are sampled to a 25-km EASE-Grid (EPSG:3408) on a monthly timescale, and then used for LSTM model predictions. Typically there are several hundred CryoSat-2 observations per 25-km grid cell. Sea-ice thickness from the same CryoSat-2 retrieval, on the same grid and monthly timescale, is used as an upper bound on the generated surface heights as described below.

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

To generate sea-ice surface topography (h_i) for each parcel, we use three fields: the CryoSat-2 roughness r_0 and sea-ice thickness described above, and the N-ICE2015 field campaign dataset. Within the footprint, the surface height h is taken to follow a lognormal distribution, so that $\ln(h)$ is normally distributed:

$$\ln(h) \sim \mathcal{N}(\mu, \sigma^2). \tag{1}$$

Here μ is the mean of $\ln(h)$, obtained from the N-ICE2015 dataset, and σ is the spread of $\ln(h)$. Because σ is a property of the logarithm it is dimensionless, whereas r_0 is the standard deviation of the height distribution itself and carries units of metres, with values between 0 and 1 m (Landy et al., 2026). The two are therefore related through the lognormal moments rather than being equal, and we obtain σ from r_0 and the mean height $\bar{h} = \exp(\mu + \sigma^2/2)$ by inverting

$$r_0 = \bar{h}\sqrt{\exp(\sigma^2) - 1}. \tag{2}$$

Each 25-km grid cell is then assigned the mean of the fitted height distribution,

$$h_i = \exp\left(\mu + \frac{\sigma^2}{2}\right), \tag{3}$$

so that two cells with the same CryoSat-2 roughness receive the same topography. The heights are in metres throughout and are not rescaled, so they remain on the same physical scale as the ridge topography used to generate the SnowModel-3D training data. Heights are constrained not to exceed the corresponding CryoSat-2 sea-ice thickness, since a surface feature cannot be taller than the ice it sits on.

This procedure generates a monthly topographic field on the EASE-Grid (EPSG:3408) at 25-km resolution across the pan-Arctic domain. The field is kept at monthly resolution and is not interpolated in time. Each Lagrangian parcel is assigned the value for the current month at its grid location, so within a month the topography a parcel experiences changes only through its own drift, and the field itself updates step-wise from one month to the next.

2.3 SnowModel-LG: Pan-Arctic Parcel Framework

For pan-Arctic inference, we adopt the Lagrangian parcel framework of SnowModel-LG (Liston et al., 2020). SnowModel-LG tracks approximately 70 000 individual snow parcels that drift with sea ice across the Arctic. Each parcel carries its own snow state (density, depth).

From the SnowModel-LG parcel trajectories, we extract a moving parcel dataset spanning 2014 to 2021. Sea-ice parcels are identified by applying a threshold of 15 % on sea-ice concentration. For each active parcel, key variables such as snow density ($sden$), spatial coordinates ($EASE\_x$, $EASE\_y$) and the CryoSat-2-derived topographic index are extracted (Table 1). The moving parcel framework preserves the spatial heterogeneity of the snowpack and provides the sequential data required for LSTM-based forecasting. Figure 2a illustrates a sample of parcel drift trajectories over a single simulation year, showing the Lagrangian motion of ice parcels across the Arctic and the evolution of emulated snow depth along each track. Figure 2b shows the emulated snow depth field for February 2018 together with the monthly-mean sea-ice drift vectors, illustrating how the parcels move across the Arctic while carrying their snow state.

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

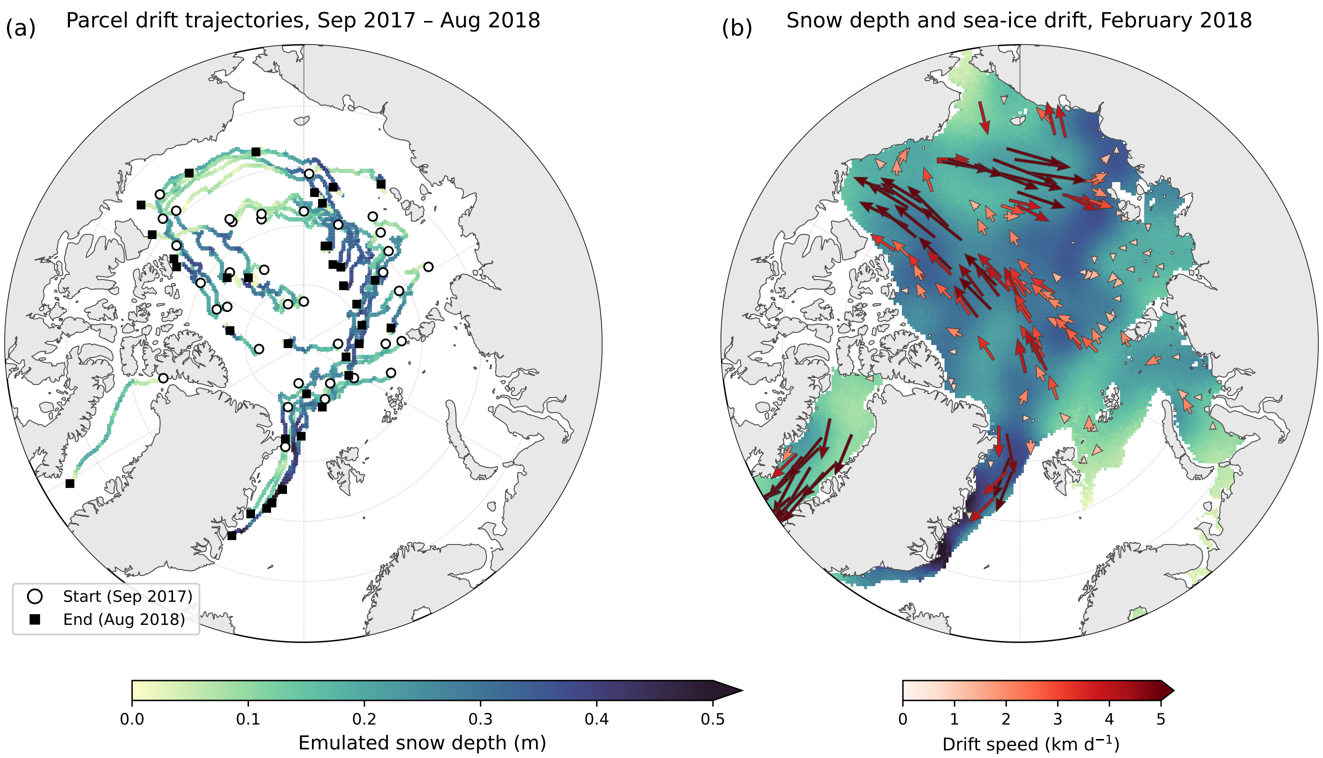


Figure 2. The Lagrangian parcel framework. (a) Sample parcel drift trajectories from September 2017 to August 2018. Each coloured line shows the path of a single Lagrangian parcel, with the colour indicating the emulated snow depth along the track. White circles mark starting positions (September 2017) and black squares mark end positions (August 2018). (b) Emulated snow depth for February 2018 overlaid with monthly-mean sea-ice drift vectors. Arrow colour indicates drift speed and arrow length is scaled 3× for visibility. Both panels use the same snow depth scale.

2.4 LSTM Architecture

The core of our forecasting framework is a Long Short-Term Memory (LSTM) network, which is designed for modelling long-term dependencies in sequential data. Traditional recurrent neural networks (RNNs) often suffer from vanishing or exploding gradients, which limits their ability to capture and remember long-term trends. LSTMs address this issue by incorporating memory cells and gating mechanisms (input, forget and output gates) (Hochreiter and Schmidhuber, 1997).

The LSTM architecture (Fig. 1b) follows the design in Prasad et al. (2024), with modifications to adapt it for the pan-Arctic application. The input layer receives ten features for each moving parcel: the snow density and snow depth from the previous time step, precipitation, 2-m air temperature, 10-m wind speed and direction, 2-m relative humidity, the CryoSat-2-derived topographic index and two spatial coordinates ($EASE\_x$, $EASE\_y$). Two LSTM layers, each containing 128 hidden units, capture the temporal dynamics of the parcel-wise time series. Dropout layers (Srivastava et al., 2014) with a rate of 0.25 are

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

Table 1. Input variables driving SnowModel-3D simulations and inputs for the LSTM emulator.

Source	Variable
MERRA-2	Precipitation 2-m Air Temperature 10-m Wind Speed 10-m Wind Direction 2-m Relative Humidity
CryoSat-2 / N-ICE2015	Surface Topography
Parcel State	Snow Density, Snow Depth (previous time step) Spatial Coordinates (EASE_x, EASE_y)

applied after each LSTM layer to prevent overfitting. The output from the LSTM is passed through a dense layer with 64 units and a Rectified Linear Unit (ReLU) activation (Nair and Hinton, 2010), followed by a final output layer that predicts snow density, snow depth and snow temperature. Snow density and snow depth are the predicted snow properties evaluated in this study and are both carried forward autoregressively, while snow temperature is an auxiliary output used only in constructing the physics-guided loss function.

2.5 Physics-Guided Loss Function

We developed a loss function to incorporate the physics of snowpack formation in the LSTM model. As snow accumulates, the lower layers of the snowpack are subjected to increasing pressure due to the weight of the snow above. This pressure causes the snow grains to compact. This process of density increase due to compaction is included in SnowModel's SnowPack component, which provides key equations explaining the snow compaction dynamics (Liston and Elder, 2006). The equation for density increase due to compaction is given by

$$\frac{\partial \rho_s}{\partial t} = C_\rho \, A_1 \, h_w^* \, \rho_s \exp\big(-B(T_f - T_s)\big) \exp\big(-A_2 \rho_s\big), \tag{4}$$

where ρ_s is the snow density (kg m^{-3}), t represents time (s), $h_w^* = \frac{1}{2} h_w$ is the weight of snow defined as half of the snow water equivalent h_w, $T_f = 273.15$ K is the freezing temperature of water, and T_s is the average snow temperature (K). C_ρ is the snow density rate adjustment factor applied in SnowModel, with $C_\rho = 5.0$. Constants A_1, A_2 and B are derived from empirical studies: $A_1 = 0.0013$ m^{-1} s^{-1}, $A_2 = 0.021$ m^3 kg^{-1} and $B = 0.08$ K^{-1}. Both exponential terms decay, since the compactive viscosity of snow increases with both density and decreasing temperature, so denser and colder snow compacts more slowly.

The snow water equivalent, h_w (m), is defined as

$$h_w = \frac{\rho_{s,\mathrm{pred}}}{\rho_w} \cdot \zeta_{s,\mathrm{pred}}, \tag{5}$$

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

where ρ_w is the density of water ($\mathrm{kg\,m^{-3}}$) and ζ_s is the snow depth (m). In discrete form, the physically expected density at the next time step is

$$\rho_{s,\text{phys}}(t+\Delta t) = \rho_{s,\text{phys}}(t) + \tfrac{1}{2}\Delta t\, C_\rho\, A_1\, \frac{\rho_{s,\text{pred}}(t)}{\rho_w}\, \zeta_{s,\text{pred}}(t)\, \rho_{s,\text{pred}}(t) \exp\!\big(-B(T_f - T_{s,\text{pred}}(t))\big) \exp\!\big(-A_2 \rho_{s,\text{pred}}(t)\big). \tag{6}$$

To align the LSTM model's predictions with these physical principles, we constructed a physics-informed loss function, $\mathcal{L}_{\text{physics}}$. This loss function is the mean squared error (MSE) between the predicted snow density and the density expected from the SnowPack compaction equation. The predicted snow depth and snow temperature enter through that expected density, since $\rho_{s,\text{phys}}$ in Eq. (6) is evaluated using $\zeta_{s,\text{pred}}$ and $T_{s,\text{pred}}$, rather than being constrained by separate error terms. The combined loss function is therefore expressed as

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{LSTM}} + \lambda\, \mathcal{L}_{\text{physics}}, \tag{7}$$

where $\mathcal{L}_{\text{LSTM}}$ is the mean squared error between the emulator's predicted snow density, depth and temperature and the corresponding floe-averaged SnowModel-3D training targets, and λ is a regularisation parameter that balances the contribution of the physics-informed loss in the overall objective function. Specifically,

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{LSTM}} + \lambda\big[\text{MSE}(\rho_{s,\text{pred}}, \rho_{s,\text{phys}})\big]. \tag{8}$$

The two terms are evaluated in different spaces, so their normalisation has to be stated for λ to be interpretable. $\mathcal{L}_{\text{LSTM}}$ is computed on standardised targets, so that snow density, depth and temperature contribute comparably despite differing by orders of magnitude in physical units. The compaction residual is evaluated in physical units, since A_1, A_2 and B are defined against $\mathrm{kg\,m^{-3}}$, m and K, and is then divided by the standard deviation of the density target before entering Eq. (8). Both terms are therefore dimensionless and λ expresses the physics contribution as a fraction of the data term.

2.6 Inference and Autoregressive Forecasting

Once trained on SnowModel-3D outputs, the LSTM model is applied for inference over the pan-Arctic domain using the SnowModel-LG parcel framework (Fig. 1c). The model is applied in an autoregressive manner to generate multi-year forecasts. The inference begins by loading the parcel dataset that contains sea-ice concentration, meteorological variables, CryoSat-2-derived topographic data and spatial coordinates for each parcel. An initial state file provides the starting snow density and depth values for each parcel at the beginning of the forecast period.

The trained model is loaded and set to evaluation mode. Hidden states and cell states for the LSTM are initialised to zero. At each time step, the model first determines which parcels are active based on a sea-ice concentration threshold of 15 %. For active parcels, input features are constructed by stacking current meteorological conditions, topographic information and the snow density and depth from the previous time step. The LSTM processes these inputs and produces a forecast for the current time step while updating its internal states.

A special procedure is followed for parcels that become active after a period of inactivity. In such cases, a nearest-neighbour search (using a KDTree (Bentley, 1975)) is performed on the spatial coordinates from the previous time step to identify the

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

most relevant state among the nearby parcels that were active. The corresponding hidden state, cell state, snow density and snow depth are then assigned to the reactivated parcel to ensure continuity in the forecast.

As the forecast progresses, the predicted snow density and depth from the previous time step are fed back into the model as inputs for the next time step, thereby extending the prediction horizon. This autoregressive forecasting method inherently leads to the accumulation of errors over time. In our framework, this error accumulation is a natural outcome of using previous predictions as inputs and is directly reflected in the model outputs.

2.7 Model Training

The network weights are learned by minimising the loss function of Eq. (8) with the Adam optimiser (Kingma and Ba, 2015) at a learning rate of 0.001, using a batch size of 32. Following Prasad et al. (2024), the data are split by region rather than by time, using leave-one-region-out cross-validation: four of the five Arctic regions are used for training and the fifth is held out, rotated so that each region serves as the held-out set in turn. A region-based split was chosen over a time-based one because the distinct climatic and geographical characteristics of each region influence performance more strongly than temporal patterns, so holding out a region tests generalisation to unseen conditions rather than to unseen dates. We train for up to 150 epochs with early stopping after 10 epochs without improvement on the held-out region to prevent overfitting (Goodfellow et al., 2016). The five regions provide the five folds, and the same split is used for hyperparameter selection: the number of LSTM layers, the number of hidden units, the dropout rate and the physics-loss weight λ were chosen by random search (Bergstra and Bengio, 2012) over 50 combinations, each scored by the mean error across the five held-out regions (Goodfellow et al., 2016). This selected the architecture reported in Sect. 2.4, with $\lambda = 0.1$. The selected configuration was then trained three times with different random seeds, and all results reported in Sect. 3.1 onwards are the mean across those three runs.

The SnowModel-3D training simulations span August 2010 to August 2020, while the pan-Arctic inference covers 2014 to 2021, so the meteorological forcing overlaps for 2014–2020. The training targets are floe-scale simulations over 1500×1500 m domains in five regions and are therefore distinct from the pan-Arctic parcel fields on which the emulator is applied, but only 2021 lies entirely outside the period spanned by the training simulations.

Model performance is evaluated by comparing the LSTM forecasts against the SnowModel-LG outputs over the full forecast period. Standard metrics, including the Root Mean Squared Error (RMSE), are computed to quantify the accuracy of the predictions. In addition, spatial and temporal analyses are conducted to assess how well the model captures the evolution of snow density across the pan-Arctic domain. The evaluation also includes an assessment of the cumulative error in the autoregressive forecasts, which arises because each prediction is fed back as input to the next time step.

2.8 Evaluation Datasets

In addition to the intercomparison with SnowModel-LG, we evaluated the emulator against three independent observational datasets. These are airborne snow radar measurements from Operation IceBridge (OIB) and from a set of Alfred Wegener Institute campaigns, and in situ snow density and snow depth observations collected during the Multidisciplinary drifting

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

Observatory for the Study of Arctic Climate (MOSAiC) expedition. This subsection describes each dataset and the processing applied prior to comparison. The corresponding evaluation results are presented in Sect. 3.3 and Sect. 3.4.

2.8.1 Airborne snow radar campaigns

The first evaluation dataset is NASA's Operation IceBridge (MacGregor et al., 2021), which conducted extensive airborne surveys of Arctic sea ice using snow radar instruments. We use 42 OIB survey flights spanning March to May of 2015–2019 (Kurtz et al., 2016). The second set comprises three Alfred Wegener Institute airborne campaigns carrying a similar snow radar instrument (Jutila et al., 2022) between 2017 and 2019. These are the PAMARCMIP2017 campaign (Jutila et al., 2021c), one flight in the RESURV79 campaign in 2018 (Jutila et al., 2021a) and the IceBird Winter 2019 campaign (Jutila et al., 2021b). These datasets provide radar-derived snow depth estimates on sea ice from airborne platforms. Because the airborne observations are much finer than the model grid, with the OIB products averaged to approximately 40 m along-track and the Alfred Wegener Institute observations having a footprint of approximately 2 m and an along-track spacing of 5–6 m, a single 25 km model grid cell may contain hundreds to thousands of individual measurements spanning a wide range of snow depths due to sub-grid-scale variability from ridges, leads and bare ice patches. A direct comparison between individual observations and the corresponding 25 km model value is therefore not meaningful. Following Merkouriadi et al. (2025), we averaged the airborne snow depth measurements onto the same 25 km EASE-Grid (EPSG:3408) used by SnowModel-LG, retaining only grid cells containing more than 50 observations. For SnowModel-LG, which is already provided on this grid, the model value was looked up directly at the corresponding grid cell and date. For the emulator, whose output is on Lagrangian parcels, we first gridded the parcel predictions onto the EASE grid using concentration-weighted averaging, then performed the same grid-cell lookup. This weighting follows the same principle as the parcel-to-grid conversion applied within SnowModel-LG (Sect. 2.2.9 of Liston et al., 2020), where each parcel contributes to a grid cell in proportion to its ice concentration. The two procedures are not identical, since SnowModel-LG also weights each parcel by its fractional area of overlap with the grid cell, whereas we assign each parcel to the grid cell containing it.

2.8.2 MOSAiC in situ snow density

The MOSAiC expedition (Nicolaus et al., 2022) deployed the research vessel Polarstern on a year-long drift through the central Arctic from October 2019 to September 2020, providing an unprecedented dataset of snow properties along the drift trajectory. We use only the snow density measured with density cutters at multiple depth layers within the snowpack (Macfarlane et al., 2023). In total, 958 individual layer measurements were obtained across 209 unique snow profiles, spanning latitudes 79–89° N. For comparison with the models, which output bulk snow density rather than depth-resolved profiles, we computed the depth-weighted average density for each profile, $\bar{\rho} = \sum_i \rho_i \Delta z_i / \sum_i \Delta z_i$, where ρ_i and Δz_i are the density and thickness of each measured layer, respectively. The resulting depth-averaged densities ranged from 102 to 564 $\mathrm{kg\,m^{-3}}$ with a mean of 328 $\mathrm{kg\,m^{-3}}$. Because MOSAiC measurements were collected along a single ship drift rather than as a spatially extensive survey, grid-cell averaging is not appropriate here and each in situ profile was instead matched to the nearest SnowModel-LG grid cell and the nearest emulator parcel in space and time.

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

2.8.3 MOSAiC Magnaprobe snow depth

Snow depth was measured along MOSAiC transects throughout the drift period with a Magnaprobe automated snow depth probe (Sturm and Holmgren, 2018), and we use the published transect dataset (Itkin et al., 2021). To match the spatial support of the 25-km gridded model output, individual probe points were averaged within each EASE grid cell and day, yielding 102 matched cell-day means spanning October 2019 to September 2020. As with the snow density comparison, SnowModel-LG values were looked up at the corresponding grid cell and date, while the emulator parcel output was first gridded onto the EASE grid using concentration-weighted averaging before the cell-day lookup.

3 Results

We present a comprehensive evaluation of our LSTM-based snow emulator across the pan-Arctic domain for multiple years. We first examine the emulator's pan-Arctic output fields (Sect. 3.1), followed by quantitative comparisons against SnowModel-LG simulations (Sect. 3.2). We then evaluate the emulator against independent airborne campaign observations (Sect. 3.3) and MOSAiC in situ measurements (Sect. 3.4).

3.1 Emulator Output

To illustrate the spatial patterns produced by the emulator, we present pan-Arctic maps of snow density and snow depth averaged over the period 2014–2021. Figure 3a shows the mean emulated snow density across the Arctic, revealing the expected gradient from lower densities in the central Arctic to higher densities near coastal and marginal ice zones where wind-driven compaction is more pronounced.

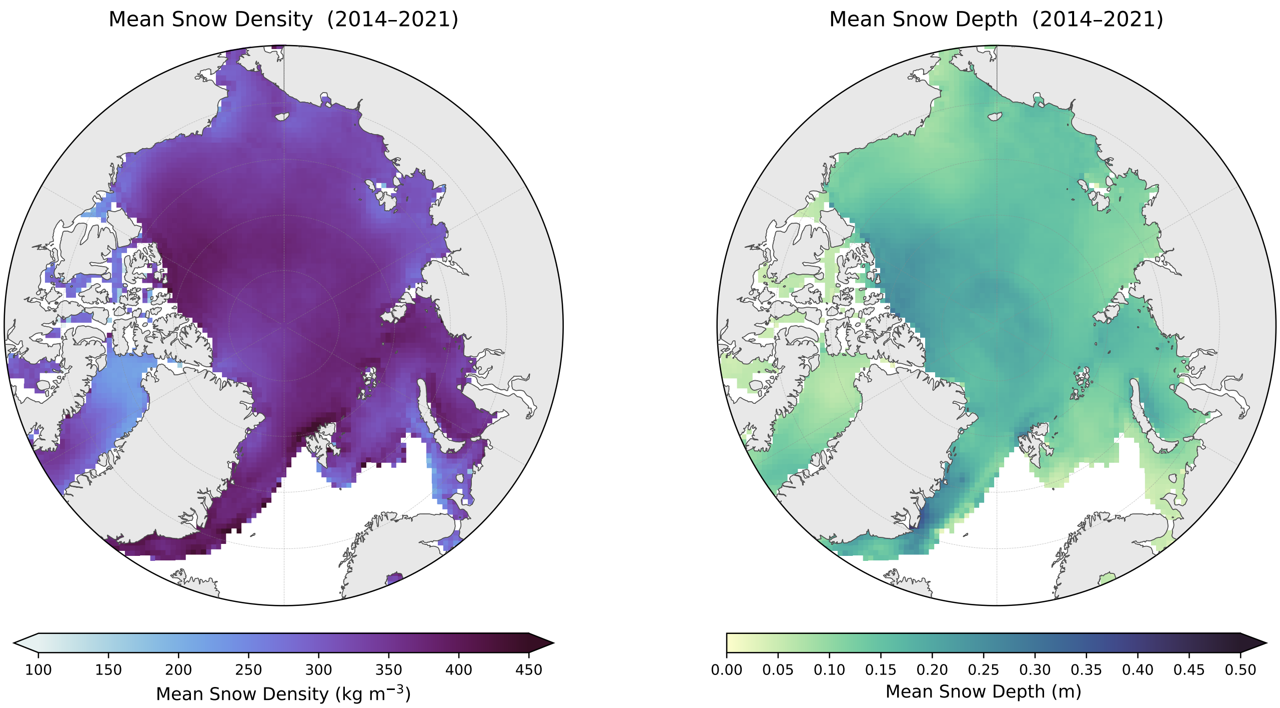
Similarly, Fig. 3b shows the mean emulated snow depth, which exhibits higher values in the central Arctic and north of the Canadian Archipelago where snow typically accumulates over a longer duration each season on multi-year ice. High snow depth values are also observed between Greenland and Svalbard and are attributed to the increased precipitation in that region due to the high frequency of storms (Graham et al., 2019).

Figures 4 and 5 present seasonal composites of snow density and snow depth, respectively, averaged over all years. The four seasons (DJF, MAM, JJA, SON) trace the seasonal evolution of the snowpack: winter (DJF) shows widespread moderate densities, spring (MAM) exhibits continued densification, summer (JJA) is characterised by reduced coverage and higher densities in remaining snow, and autumn (SON) marks the onset of new snow accumulation with lower densities. Complete month-by-month maps for each year are provided in the Supplement, showing the full spatiotemporal evolution of both snow density (Fig. S1) and snow depth (Fig. S2).

3.2 Comparison with SnowModel-LG

We compared the predicted snow densities and depths with those obtained from SnowModel-LG simulations over the entire pan-Arctic domain for multiple years (2014–2021). Although this comparison does not serve as evaluation, since the emulator

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026



(a) (b)

Figure 3. Mean emulated (a) snow density and (b) snow depth across the pan-Arctic domain, averaged over 2014–2021. The polar stereographic projection shows density gradients from the central Arctic toward peripheral seas, and higher snow depths in the central Arctic and Canadian Archipelago where snow accumulates over multi-year ice.

is trained with SnowModel-3D, it is still valuable because SnowModel-3D and SnowModel-LG apply the same physics in the evolution of snow properties regardless of the topography driven effects (only included in SnowModel-3D) that should be responsible for some of the differences. Table 2 summarises the RMSE of the emulator predictions for each year, where each year is aggregated by calendar year (January to December). The emulator has an overall RMSE of 76.76 kg m^{-3}, demonstrating consistent performance across different years with modest interannual variability. The year 2015 shows the highest RMSE (96.69 kg m^{-3}), while 2021 has the lowest (66.23 kg m^{-3}). This variability reflects natural fluctuations in meteorological conditions and snow evolution dynamics captured effectively by our emulator.

Additionally, Table 2 lists the RMSE by month, showing seasonal variations in prediction accuracy. Prediction differences are highest during summer months (July and August), coinciding with melting events and increased snow density variability. Conversely, June exhibits the lowest RMSE (41.31 kg m^{-3}), reflecting stable early summer snow conditions in better agreement with SnowModel-LG. Table 3 presents the corresponding RMSE values for snow depth predictions, with an overall RMSE of

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

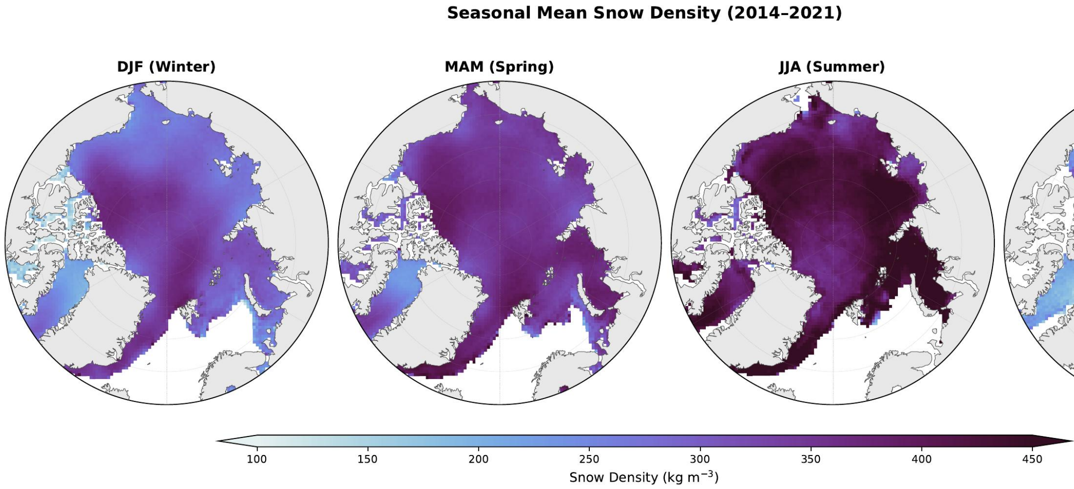


Figure 4. Seasonal mean snow density (kg m^{-3}) from the emulator, averaged over 2014–2021. Panels show DJF (winter), MAM (spring), JJA (summer) and SON (autumn).

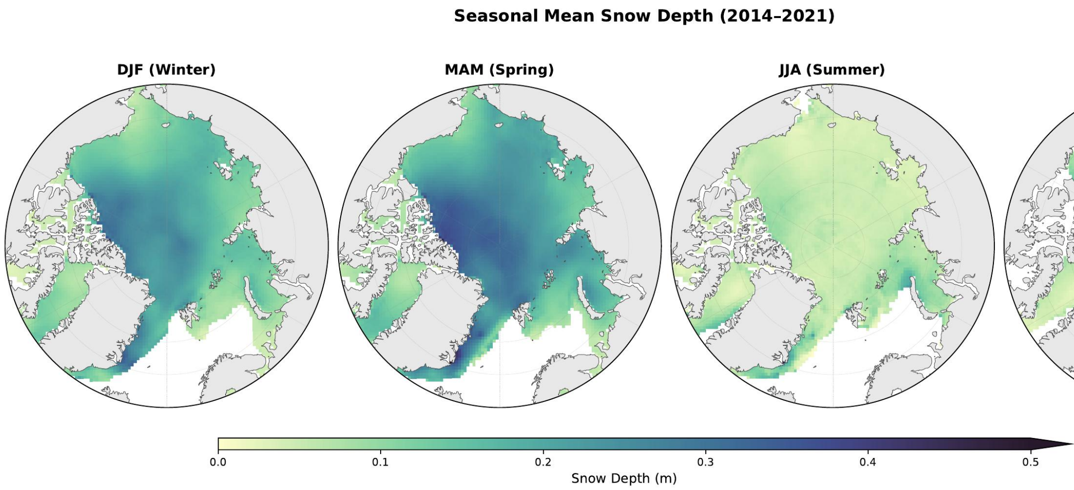


Figure 5. Seasonal mean snow depth (m) from the emulator, averaged over 2014–2021. Panels show DJF (winter), MAM (spring), JJA (summer) and SON (autumn).

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

Table 2. RMSE of snow density predictions (kg m^{-3}) between the emulator and SnowModel-LG. Yearly values are aggregated by calendar year (January to December), and the monthly values are climatologies averaged over 2014–2021.

Year	2014	2015	2016	2017	2018	2019	2020	2021	**Overall**			
RMSE	72.06	96.69	77.65	71.24	74.05	73.59	74.66	66.23	**76.76**			
Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
RMSE	73.90	73.58	75.56	74.46	69.52	41.31	104.99	88.69	80.33	78.57	72.66	72.45

Table 3. RMSE of snow depth predictions (m) between the emulator and SnowModel-LG. Yearly values are aggregated by calendar year (January to December), and the monthly values are climatologies averaged over 2014–2021.

Year	2014	2015	2016	2017	2018	2019	2020	2021	**Overall**			
RMSE	0.07	0.07	0.07	0.07	0.09	0.06	0.04	0.04	**0.07**			
Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
RMSE	0.07	0.08	0.09	0.10	0.10	0.05	0.00	0.01	0.03	0.05	0.05	0.06

0.07 m. Unlike snow density, snow depth differences are smallest in July–August because most snow has melted and depths are near zero, and are largest in late winter and spring (March–May) when snow depths are deepest.

Figure 6 provides a visual comparison of mean snow depth and snow density fields from the emulator and SnowModel-LG for 2018, showing close spatial agreement. Multi-year comparisons of mean snow density and snow depth distributions are provided in the Supplement (Figs. S3 and S4), with the difference maps showing that deviations between the emulator and SnowModel-LG remain small across years.

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

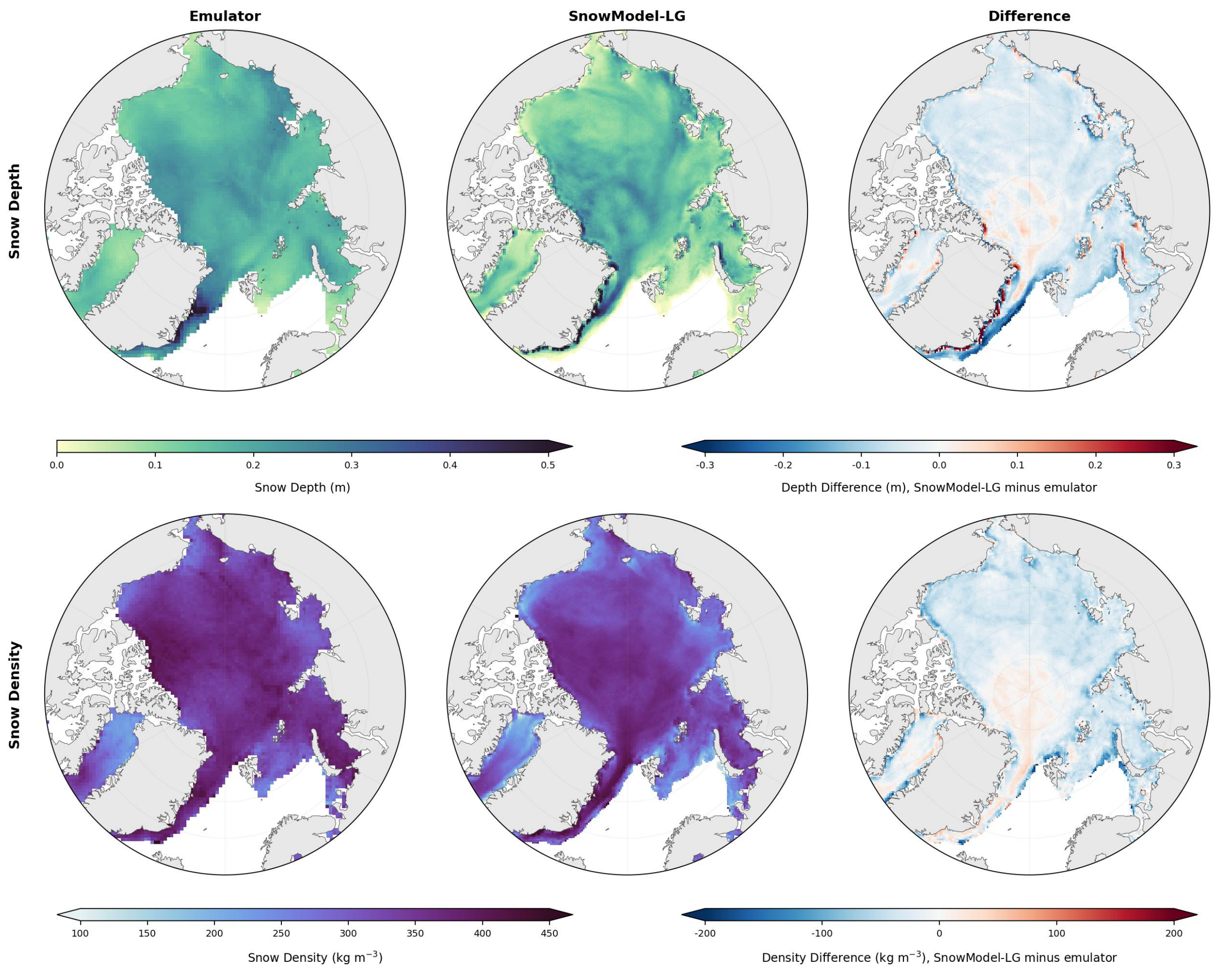


Figure 6. Visual comparison of annual mean snow depth (top row) and snow density (bottom row) from the emulator (left), SnowModel-LG (centre), and their difference, computed as SnowModel-LG minus emulator (right), for 2018.

3.3 Comparison with airborne campaigns

Table 4 summarises the comparison against the airborne campaigns described in Sect. 2.8. Against OIB data, the emulator achieves an RMSE of 0.100 m compared to 0.139 m for SnowModel-LG. Against the combined Jutila et al. IceBird campaigns (2017–2019), the emulator achieves an RMSE of 0.080 m compared to 0.134 m for SnowModel-LG. The emulator shows closer agreement with the airborne snow depths than SnowModel-LG in both comparisons (Fig. 7). We do not interpret this as the emulator being a more accurate version of SnowModel-LG, since the two simulate different things. Rather, the emulator

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

Table 4. RMSE of snow depth (m) for SnowModel-LG and the emulator evaluated against airborne campaign observations. Observations are averaged onto 25 km EASE grid cells following Merkouriadi et al. (2025).

Model	Operation IceBridge	Airborne campaigns (2017–2019)
SnowModel-LG	0.139	0.134
Emulator	0.100	0.080

additionally represents the small-scale snow redistribution over sea-ice topography that SnowModel-LG does not account for, learned from the high-resolution SnowModel-3D outputs together with the auxiliary CryoSat-2 surface topography estimates, and this added process brings the simulated snow depth closer to the airborne observations. Our SnowModel-LG RMSE values (0.13–0.14 m) are in a similar range to those reported by Merkouriadi et al. (2025) for SnowModel-LG evaluated against the same airborne campaigns (RMSE of 0.11–0.14 m for IceBird, their Figure 3).

3.4 Comparison with MOSAiC in situ observations

Figure 8 shows the spatial distribution of the MOSAiC observations described in Sect. 2.8, overlaid on the SnowModel-LG mean density field for the same period. Table 5 summarises the evaluation statistics. The sample sizes differ slightly between the two models because a few profiles fell outside the active SnowModel-LG grid cells and could not be matched. SnowModel-LG achieves an RMSE of 159.6 $\mathrm{kg\,m^{-3}}$, while the emulator achieves an RMSE of 139.2 $\mathrm{kg\,m^{-3}}$. The elevated RMSE values for both models reflect the inherent challenge of comparing point-scale in situ measurements with 25-km gridded or parcel-based model output. Local-scale snow variability driven by ice topography and wind redistribution is not resolved at the model scale.

Figure 9a presents the monthly mean snow density from the in situ observations alongside the matched SnowModel-LG and emulator values. Both models capture the broad seasonal trend of increasing density from autumn through spring, although they underestimate the observed variability during winter months. The emulator tracks the seasonal evolution with comparable fidelity to SnowModel-LG, confirming that it provides a physically reasonable representation of snow density even against independent field measurements.

We additionally evaluated snow depth against the MOSAiC Magnaprobe observations described in Sect. 2.8. SnowModel-LG and the emulator produce nearly identical bulk performance (Fig. 9b), with RMSE values of 0.106 m and 0.105 m respectively, and a comparable negative bias of -0.07 m relative to the observed grid-cell mean of 0.17 m (Table 6). The emulator shows a higher correlation with the observations ($R = 0.74$ versus $R = 0.70$ for SnowModel-LG), indicating that it better tracks the spatial variability of snow depth along the drift even though its mean depths are slightly lower. Both models therefore underestimate snow depth on level central-Arctic ice during MOSAiC, a known limitation of MERRA-2-forced reconstructions over thick-snow regions (Merkouriadi et al., 2017, 2020), but the emulator preserves the spatial variability of the snowpack as well as, or slightly better than, SnowModel-LG.

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

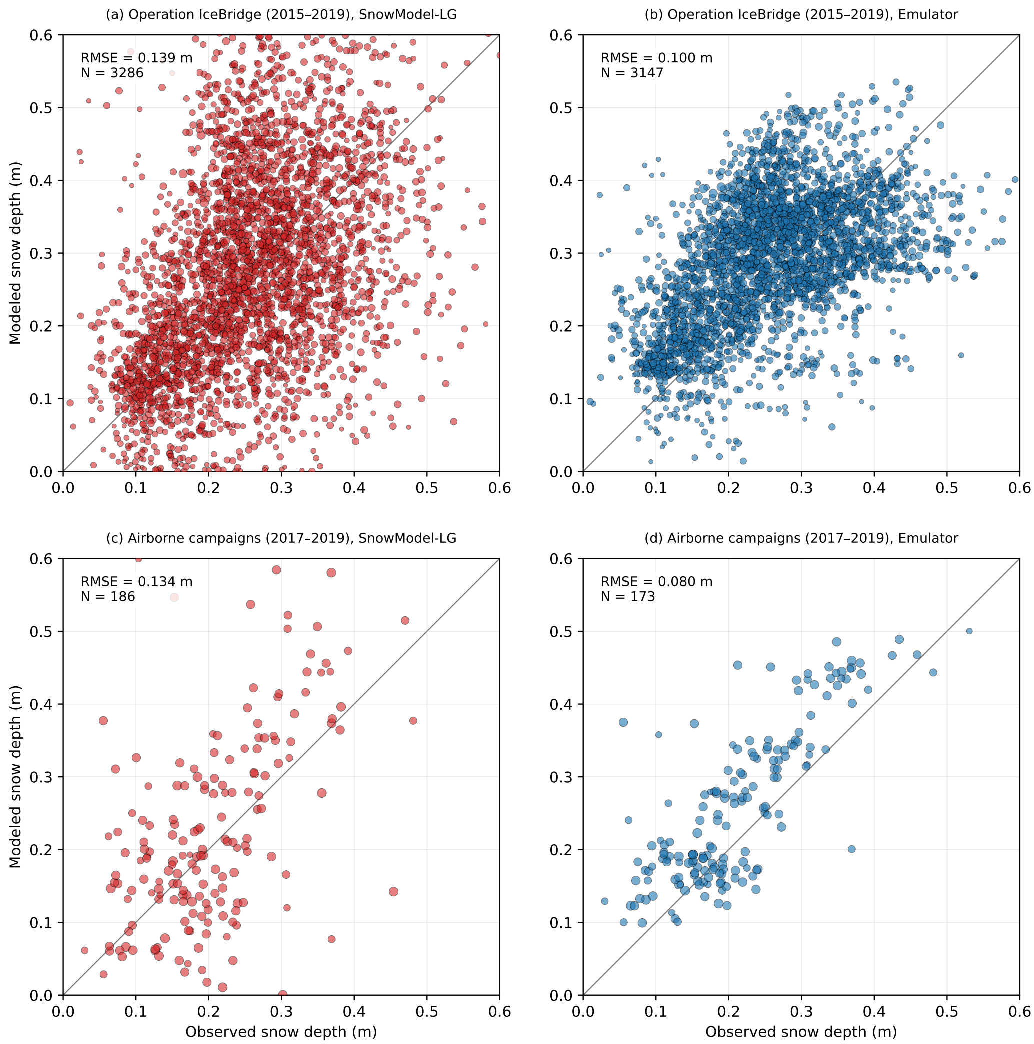


Figure 7. Scatter plots of grid-averaged observed snow depth versus modelled snow depth for (a, b) Operation IceBridge (2015–2019) and (c, d) the combined IceBird/PAMARCMIP/RESURV79 campaigns (2017–2019). Left column shows SnowModel-LG, right column shows the emulator. Each point represents one 25 km EASE grid cell. Point size is proportional to the number of airborne observations within the cell.

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

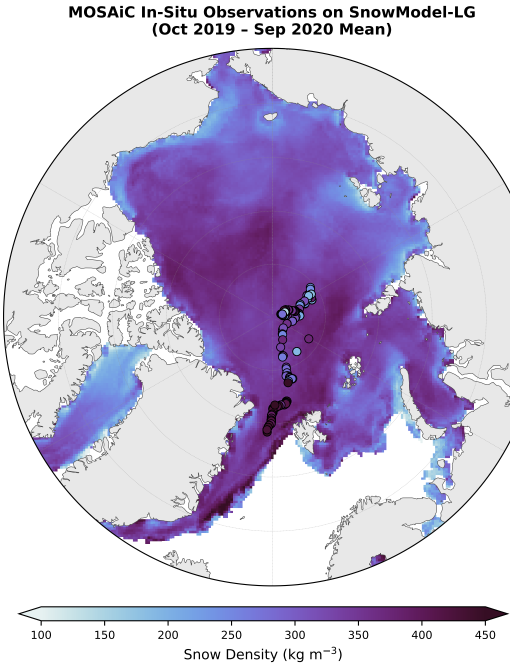


Figure 8. Spatial distribution of MOSAiC in situ snow density observations (coloured circles) overlaid on the SnowModel-LG mean snow density field for October 2019 – September 2020. Both the background field and observation markers use the same colour scale.

Table 5. Validation of snow density (kg m^{-3}) against MOSAiC in situ density cutter observations (October 2019 – September 2020).

Model	N	RMSE
SnowModel-LG	204	159.6
Emulator	209	139.2

Table 6. Validation of snow depth (m) against MOSAiC observations aggregated to 25-km EASE grid-cell daily means (October 2019 – September 2020).

Model	N	RMSE	MAE	Bias	R
SnowModel-LG	102	0.106	0.079	−0.066	0.70
Emulator	102	0.105	0.082	−0.074	0.74

3.5 Computational cost

One of the motivations for developing the emulator is to make the topographic effect on snow, which only SnowModel-3D resolves, available at the pan-Arctic scale, where SnowModel-3D itself cannot be run directly. Table 7 summarises the cost

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

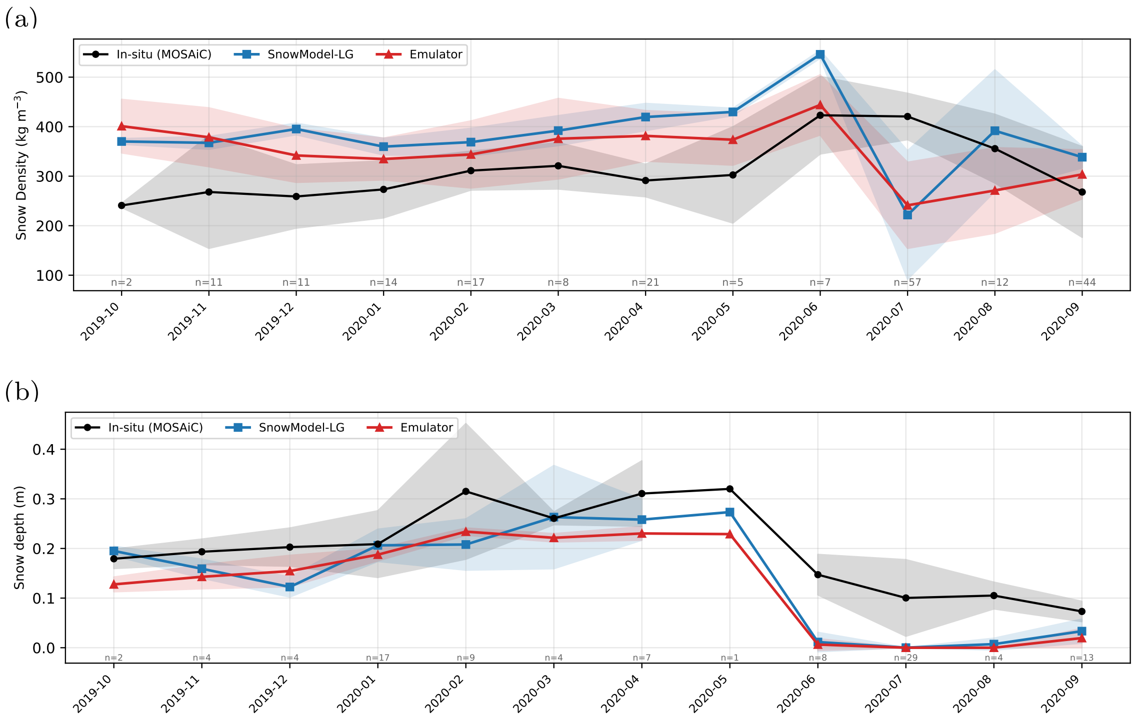


Figure 9. Monthly mean (a) snow density and (b) snow depth from MOSAiC in situ observations (black), SnowModel-LG (blue) and the emulator (red) over the MOSAiC drift period (October 2019 – September 2020). Density is from density-cutter profiles and depth is from Magnaprobe transects aggregated to 25-km EASE-Grid cell daily means. Shaded regions indicate ± 1 standard deviation across all matched profiles within each month. In (b), numbers below each tick indicate the number of matched days, and in May 2020 the MOSAiC expedition stayed within one EASE-Grid cell, yielding zero standard deviation.

of one simulated pan-Arctic year for the three approaches. SnowModel-LG simulates one year over its ~70 000 Lagrangian parcels in about 8.3 min and the emulator produces the same pan-Arctic year in about 45 min on a single NVIDIA A100 GPU. The emulator is therefore of the same order of magnitude as SnowModel-LG, while additionally carrying the sea-ice topographic effect that SnowModel-LG omits entirely.

The relevant comparison is with SnowModel-3D, which is the only model among the three that has physical processes resolving wind-driven snow redistribution over sea-ice topography. In Prasad et al. (2024), SnowModel-3D required about 8 CPU-hours to simulate a single 1500×1500 m floe domain at 1 m resolution over a 10-year period, that is roughly 0.8 CPU-hours per floe-year for 2.25×10^6 grid cells. The pan-Arctic sea-ice zone covers of the order of 10^7 km^2, which is about 4×10^6 such floe domains. Because SnowModel is a serial, per-grid-cell process model whose cost scales linearly with the number of grid cells and time steps (Liston et al., 2018; Prasad et al., 2024), tiling the basin at 1 m would require of the order of 3×10^6 CPU-hours, or several hundred CPU-years, for each simulated year. A direct pan-Arctic SnowModel-3D simulation at

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

Table 7. Approximate computational cost of one simulated pan-Arctic year. The SnowModel-3D figure is an order-of-magnitude estimate derived by linear scaling from the per-floe runtime reported in Prasad et al. (2024) (about 0.8 CPU-hours per 1500×1500 m floe-year at 1 m resolution) to the $\sim 10^7$ km^2 pan-Arctic sea-ice zone.

Model	**Runtime (one year)**	**Resolves ice topography**
SnowModel-LG	$\sim$8.3 min	No
Emulator	$\sim$45 min (1 A100 GPU)	Yes (learned from SnowModel-3D)
SnowModel-3D	$\sim 3 \times 10^6$ CPU-hours (est.)	Yes

this resolution is thus impractical with current process-based models, whereas the emulator reproduces its topographic effect in tens of minutes.

4 Discussion

The results of our emulator demonstrate consistency with SnowModel-LG simulations across multiple years and seasonal variations. The general patterns of snow depth and density were similar between the two: elevated north of the Canadian Arctic and in the Fram Strait, reduced in the peripheral seas where sea ice is younger. The seasonality of differences varied due to the changing complexities of snow dynamics between freezing and melting periods.

The most obvious methodological difference between the emulator and SnowModel-LG is the inclusion of sea-ice-topography-driven depth and density evolution in the former but not the latter. In simple terms, the emulator and SnowModel-LG receive the same snowfall amount, as prescribed by the meteorological forcing. The key difference is how this snow is distributed after deposition. SnowModel-LG assumes that snowfall is spread uniformly over the sea ice surface, whereas the emulator accounts for the redistribution of snow across sea-ice topography, as learned from SnowModel-3D. This redistribution affects both averages of snow depth and density within an ice floe. Consequently, average snow depth and density predicted by the emulator are expected to differ from SnowModel-LG, i.e., redistributed snow can become locally deeper and denser, and because snow depth is a prognostic variable, denser snow with the same snow water equivalent will correspond to a shallower snowpack. Therefore, several of the differences may be attributed to this: patterns of snow are generally smoother, with less texture, in the emulator, and there are elevated differences in the RMSE values for density during summer months. These seasonal differences emphasise the influence of non-linear thermodynamic and radiative feedbacks that drive snow densification during the melt season.

The emulator's overall performance remains robust across years (2014–2021), and by additionally representing the topographic effect on snow redistribution it shows closer agreement with the airborne observations than SnowModel-LG. The autoregressive inference framework enables the emulator to maintain temporal continuity, effectively capturing the persistence of snow states across multi-year timescales. However, a gradual accumulation of forecast error remains inherent to this approach, particularly in prolonged free-run experiments, a common limitation of data-driven emulators.

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

The computational cost of the emulator (Sect. 3.5) is comparable to that of SnowModel-LG and slightly higher, at about 45 min against 8.3 min for one pan-Arctic year. The value of the emulator is therefore not that it is faster than SnowModel-LG, but that it delivers a fundamentally different product. SnowModel-LG cannot represent snow redistribution over sea-ice topography at all, whereas the emulator reproduces this SnowModel-3D effect at a cost that remains in the tens of minutes. The alternative of obtaining the same topographic effect directly from SnowModel-3D at the pan-Arctic scale is impractical by several orders of magnitude. The emulator should thus be understood as making a previously unattainable simulation tractable, rather than as an acceleration of SnowModel-LG.

Moreover, the emulator's architecture has flexibility for further development. Incorporating additional physics-based constraints, such as energy balance terms or radiative transfer corrections, could further improve melt-season predictive skill. Similarly, coupling the emulator with satellite data streams from additional satellites could provide additional modalities to enable near-real-time monitoring of snow parameters.

5 Conclusions

We presented an LSTM-based emulator for pan-Arctic snow on sea ice, trained on high-resolution SnowModel-3D simulations and applied over the Lagrangian parcel trajectories of SnowModel-LG. By combining an autoregressive, parcel-based inference framework with a physics-guided loss function, the emulator reproduces snow density and depth distributions across the Arctic at a cost of tens of minutes per pan-Arctic year, far below that of a direct SnowModel-3D simulation at the same scale. The use of CryoSat-2-derived surface roughness as a topographic input enables the emulator to account for the influence of sea-ice topography on snow redistribution at the pan-Arctic scale.

The main outcome of this study is that the influence of sea-ice surface topography on snow redistribution, which has until now only been resolved at metre scales by SnowModel-3D, can be carried up to the pan-Arctic scale through a learned emulator and a satellite-derived roughness proxy. The comparison against airborne snow radar from Operation IceBridge and IceBird campaigns shows that including this topographic information brings the emulated snow depth to within 0.08–0.10 m RMSE of the airborne observations, while the comparison against MOSAiC, where the snow is deeper and the ice surface is comparatively level, shows that the emulator and SnowModel-LG behave similarly and share a bias that is consistent with known limitations of the underlying atmospheric reanalysis.

Resolving sea-ice topography at the pan-Arctic scale would normally require running SnowModel-3D over the entire basin, which is not practical with current process-based models, and the emulator makes this computationally tractable. This makes it feasible to incorporate machine learning-emulated snow physics into large-ensemble climate experiments, sea-ice data assimilation systems and satellite altimetry-based thickness retrievals, which have until now had to rely either on a fixed snow climatology or on pan-Arctic snow products that do not include the ice surface beneath the snow.

Code and data availability. The observational and model datasets used in this study are available as follows.

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

The MERRA-2 atmospheric reanalysis used to force SnowModel-3D and to drive the emulator (Gelaro et al., 2017), from which we use 10-m wind speed and direction, 2-m air temperature, 2-m relative humidity and total water-equivalent precipitation, is distributed by the NASA Global Modeling and Assimilation Office through the Goddard Earth Sciences Data and Information Services Center (https://disc.gsfc.nasa.gov/datasets?project=MERRA-2).

The SnowModel-LG parcel trajectories and gridded snow density and depth fields (Liston et al., 2020) are available from the NASA National Snow and Ice Data Center as *Lagrangian Snow Distributions for Sea-Ice Applications, Version 1* (https://doi.org/10.5067/27A0P5M6LZBI).

The Operation IceBridge snow depth retrievals are available as *IceBridge Sea Ice Freeboard, Snow Depth, and Thickness Quick Look, Version 1* (Kurtz et al., 2016), https://doi.org/10.5067/GRIXZ91DE0L9. The Alfred Wegener Institute airborne snow depth data are archived at PANGAEA for the PAMARCMIP2017 campaign (Jutila et al., 2021c), https://doi.org/10.1594/PANGAEA.932668, the RESURV79 campaign (Jutila et al., 2021a), https://doi.org/10.1594/PANGAEA.932702, and the IceBird Winter 2019 campaign (Jutila et al., 2021b), https://doi.org/10.1594/PANGAEA.932790.

The MOSAiC Magnaprobe snow depth transects (Itkin et al., 2021) are available at PANGAEA, https://doi.org/10.1594/PANGAEA.937781. The MOSAiC snow density measurements are part of the snow database described by Macfarlane et al. (2023), https://doi.org/10.1038/s41597-023-02273-1.

Author contributions. AP, IM and AN designed the study. AP developed the emulator. JL provided the CryoSat-2 surface-roughness data. AJ provided the airborne snow-radar data. GEL developed SnowModel and contributed the SnowModel-3D and SnowModel-LG framework. IM and AN supervised the work. All authors discussed the results and reviewed the manuscript.

Competing interests. The authors declare that they have no competing interests.

Acknowledgements. The authors acknowledge CSC – IT Center for Science, Finland, for the computational resources used in this work. Generative AI tools (OpenAI Codex) was used to assist with code development and with editing of the manuscript. All AI-assisted output was reviewed by the authors.

Financial support. AP and IM were supported by the Research Council of Finland (grant no. 341550). JCL acknowledge support from the air–snow–ice–ocean INTERActions transforming Atlantic Arctic Climate (INTERAAC) project under the Research Council of Norway, RCN (grant no. 328957) and the Detecting episodes of Arctic sea ice Mass Imbalance (DynAMIC) project under the Research Council of Norway, RCN (grant no. 343069). JCL further acknowledges funding from the European Research Council SI/3D grant (#101077496) which supported the preparation of CryoSat-2 sea ice roughness data. AJ was supported by the Research Council of Finland project *SNOWBALL* (369647).

https://doi.org/10.5194/egusphere-2026-4727
Preprint. Discussion started: 1 September 2026

References

Bentley, J. L.: Multidimensional binary search trees used for associative searching, Communications of the ACM, 18, 509–517, https://doi.org/10.1145/361002.361007, 1975.

Bergstra, J. and Bengio, Y.: Random Search for Hyper-Parameter Optimization, Journal of Machine Learning Research, 13, 281–305, https://jmlr.org/papers/v13/bergstra12a.html, 2012.

Carret, A., Fleury, S., Di Bella, A., Landy, J., Lawrence, I., Kurtz, N., Laforge, A., Bouffard, J., and Parrinello, T.: A multi-frequency altimetry snow depth product over Arctic sea ice, Scientific Data, 12, 104, https://doi.org/10.1038/s41597-024-04343-4, 2025.

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., Randles, C. A., Darmenov, A., Bosilovich, M. G., Reichle, R., Wargan, K., Coy, L., Cullather, R., Draper, C., Akella, S., Buchard, V., Conaty, A., da Silva, A. M., Gu, W., Kim, G.-K., Koster, R., Lucchesi, R., Merkova, D., Nielsen, J. E., Partyka, G., Pawson, S., Putman, W., Rienecker, M., Schubert, S. D., Sienkiewicz, M., and Zhao, B.: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Journal of Climate, 30, 5419–5454, https://doi.org/10.1175/JCLI-D-16-0758.1, 2017.

Goodfellow, I., Bengio, Y., and Courville, A.: Deep Learning, MIT Press, https://www.deeplearningbook.org, 2016.

Graham, R. M., Itkin, P., Meyer, A., Sundfjord, A., Spreen, G., Smedsrud, L. H., Liston, G. E., Cheng, B., Cohen, L., Divine, D., Fer, I., Fransson, A., Gerland, S., Haapala, J., Hudson, S. R., Johansson, A. M., King, J., Merkouriadi, I., Peterson, A. K., Provost, C., Randelhoff, A., Rinke, A., Rösel, A., Sennéchael, N., Walden, V. P., Duarte, P., Assmy, P., Steen, H., and Granskog, M. A.: Winter storms accelerate the demise of sea ice in the Atlantic sector of the Arctic Ocean, Scientific Reports, 9, https://doi.org/10.1038/s41598-019-45574-5, 2019.

Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Computation, 9, 1735–1780, https://doi.org/10.1162/neco.1997.9.8.1735, 1997.

Itkin, P., Webster, M., Hendricks, S., Oggier, M., Jaggi, M., Ricker, R., Arndt, S., Divine, D. V., von Albedyll, L., Raphael, I., Rohde, J., and Liston, G. E.: Magnaprobe snow and melt pond depth measurements from the 2019–2020 MOSAiC expedition, https://doi.org/10.1594/PANGAEA.937781, 2021.

Jutila, A., King, J., Ricker, R., Hendricks, S., Helm, V., and Binder, T.: Airborne high-altitude snow depth on sea ice during aircraft flight P6_211_RESURV79_2018_1804100301, Version 1, https://doi.org/10.1594/PANGAEA.932702, 2021a.

Jutila, A., King, J., Ricker, R., Hendricks, S., Helm, V., Binder, T., and Haas, C.: Airborne snow depth on sea ice during the IceBird Winter 2019 campaign in the Arctic Ocean, Version 1, https://doi.org/10.1594/PANGAEA.932790, 2021b.

Jutila, A., King, J., Ricker, R., Hendricks, S., Helm, V., Binder, T., and Herber, A.: Airborne snow depth on sea ice during the PAMAR-CMIP2017 campaign in the Arctic Ocean, Version 1, https://doi.org/10.1594/PANGAEA.932668, 2021c.

Jutila, A., King, J., Paden, J., Ricker, R., Hendricks, S., Polashenski, C., Helm, V., Binder, T., and Haas, C.: High-Resolution Snow Depth on Arctic Sea Ice From Low-Altitude Airborne Microwave Radar Data, IEEE Transactions on Geoscience and Remote Sensing, 60, 4300716, https://doi.org/10.1109/TGRS.2021.3063756, 2022.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, in: 3rd International Conference on Learning Representations (ICLR), https://doi.org/10.48550/arXiv.1412.6980, 2015.

Kurtz, N., Studinger, M., Harbeck, J., Onana, V., and Yi, D.: IceBridge Sea Ice Freeboard, Snow Depth, and Thickness Quick Look, Version 1, https://doi.org/10.5067/GRIXZ91DE0L9, 2016.

Landy, J. C., Petty, A. A., Tsamados, M., and Stroeve, J. C.: Sea ice roughness overlooked as a key source of uncertainty in CryoSat-2 ice freeboard retrievals, Journal of Geophysical Research: Oceans, 125, e2019JC015820, https://doi.org/10.1029/2019JC015820, 2020.

Landy, J. C., de Rijke-Thomas, C., Nab, C., Lawrence, I., Glissenaar, I. A., Mallett, R. D. C., Fredensborg Hansen, R. M., Petty, A., Tsamados, M., Macfarlane, A. R., and Braakmann-Folgmann, A.: Anticipating CRISTAL: an exploration of multi-frequency satellite altimeter snow depth estimates over Arctic sea ice, 2018–2023, The Cryosphere, 20, 183–208, https://doi.org/10.5194/tc-20-183-2026, 2026.

Liston, G. E. and Elder, K.: A Distributed Snow-Evolution Modeling System (SnowModel), Journal of Hydrometeorology, 7, 1259–1276, https://doi.org/10.1175/JHM548.1, 2006.

Liston, G. E., Polashenski, C., Rösel, A., Itkin, P., King, J., Merkouriadi, I., and Haapala, J.: A Distributed Snow-Evolution Model for Sea-Ice Applications (SnowModel), Journal of Geophysical Research: Oceans, 123, 3786–3810, https://doi.org/10.1002/2017JC013706, 2018.

Liston, G. E., Itkin, P., Stroeve, J., Tschudi, M., Stewart, J. S., Pedersen, S. H., Reinking, A. K., and Elder, K.: A Lagrangian Snow-Evolution System for Sea-Ice Applications (SnowModel-LG): Part I—Model Description, Journal of Geophysical Research: Oceans, 125, e2019JC015 913, https://doi.org/10.1029/2019JC015913, 2020.

Macfarlane, A. R., Schneebeli, M., Dadic, R., Tavri, A., Immerz, A., Polashenski, C., Krampe, D., Clemens-Sewall, D., Wagner, D. N., Perovich, D. K., Hannula, H.-R., Raphael, I., Matero, I., Regnery, J., Smith, M. M., Nicolaus, M., Jaggi, M., Oggier, M., Webster, M. A., Lehning, M., Kolabutin, N., Itkin, P., Naderpour, R., Pirazzini, R., Hämmerle, S. H., Arndt, S., and Fons, S.: A Database of Snow on Sea Ice in the Central Arctic Collected during the MOSAiC Expedition, Scientific Data, 10, 398, https://doi.org/10.1038/s41597-023-02273-1, 2023.

MacGregor, J. A., Boisvert, L. N., Medley, B., Petty, A. A., Harbeck, J. P., Bell, R. E., Blair, J. B., Blanchard-Wrigglesworth, E., Buckley, E. M., Christoffersen, M. S., Cochran, J. R., Csathó, B. M., De Marco, E. L., Dominguez, R. T., Fahnestock, M. A., Farrell, S. L., Gogineni, S. P., Greenbaum, J. S., Hansen, C. M., Hofton, M. A., Holt, J. W., Jezek, K. C., Koenig, L. S., Kurtz, N. T., Kwok, R., Larsen, C. F., Leuschen, C. J., Locke, C. D., Manizade, S. S., Martin, S., Neumann, T. A., Nowicki, S. M. J., Paden, J. D., Richter-Menge, J. A., Rignot, E. J., Rodríguez-Morales, F., Siegfried, M. R., Smith, B. E., Sonntag, J. G., Studinger, M., Tinto, K. J., Truffer, M., Wagner, T. P., Woods, J. E., Young, D. A., and Yungel, J. K.: The Scientific Legacy of NASA's Operation IceBridge, Reviews of Geophysics, 59, e2020RG000 712, https://doi.org/10.1029/2020RG000712, 2021.

Merkouriadi, I., Cheng, B., Graham, R. M., Rösel, A., and Granskog, M. A.: Critical role of snow on sea ice growth in the Atlantic sector of the Arctic Ocean, Geophysical Research Letters, 44, 10 479–10 485, https://doi.org/10.1002/2017GL075494, 2017.

Merkouriadi, I., Cheng, B., Hudson, S. R., and Granskog, M. A.: Effect of frequent winter warming events (storms) and snow on sea-ice growth–a case from the Atlantic sector of the Arctic Ocean during the N-ICE2015 campaign, Annals of Glaciology, 61, 164–170, https://doi.org/10.1017/aog.2020.25, 2020.

Merkouriadi, I., Jutila, A., Liston, G. E., Preußer, A., and Webster, M. A.: Investigating snow sinks on level sea ice: A case study in the western Arctic, Journal of Glaciology, 71, e66, https://doi.org/10.1017/jog.2025.34, 2025.

Nair, V. and Hinton, G. E.: Rectified Linear Units Improve Restricted Boltzmann Machines, in: Proceedings of the 27th International Conference on Machine Learning (ICML), pp. 807–814, https://www.cs.toronto.edu/~fritz/absps/reluICML.pdf, 2010.

Nicolaus, M., Perovich, D. K., Spreen, G., Granskog, M. A., von Albedyll, L., Angelopoulos, M., Anhaus, P., Arndt, S., Belter, H. J., Bessonov, V., Birnbaum, G., Brauchle, J., Calmer, R., Cardellach, E., Cheng, B., Clemens-Sewall, D., Dadic, R., Damm, E., de Boer, G., Demir, O., Dethloff, K., Divine, D. V., Fong, A. A., Fons, S., Frey, M. M., Fuchs, N., Gabarró, C., Gerland, S., Goessling, H. F., Gradinger, R., Haapala, J., Haas, C., Hamilton, J., Hannula, H.-R., Hendricks, S., Herber, A., Heuzé, C., Hoppmann, M., Høyland, K. V., Huntemann, M., Hutchings, J. K., Hwang, B., Itkin, P., Jacobi, H.-W., Jaggi, M., Jutila, A., Kaleschke, L., Katlein, C., Kolabutin, N., Krampe, D., Kristensen, S. S., Krumpen, T., Kurtz, N., Lampert, A., Lange, B. A., Lei, R., Light, B., Linhardt, F., Liston, G. E., Loose, B., Macfarlane, A. R., Mahmud, M., Matero, I. O., Maus, S., Morgenstern, A., Naderpour, R., Nandan, V., Niubom, A., Oggier, M., Oppelt, N., Pätzold, F., Perron, C., Petrovsky, T., Pirazzini, R., Polashenski, C., Rabe, B., Raphael, I. A., Regnery, J., Rex, M., Ricker, R., Riemann-Campe, K., Rinke, A., Rohde, J., Salganik, E., Scharien, R. K., Schiller, M., Schneebeli, M., Semmling, M., Shimanchuk, E., Shupe, M. D., Smith, M. M., Smolyanitsky, V., Sokolov, V., Stanton, T., Stroeve, J., Thielke, L., Timofeeva, A., Tonboe, R. T., Tavri, A., Tsamados, M., Wagner, D. N., Watkins, D., Webster, M., and Wendisch, M.: Overview of the MOSAiC expedition: Snow and sea ice, Elementa: Science of the Anthropocene, 10, 000 046, https://doi.org/10.1525/elementa.2021.000046, 2022.

Perovich, D., Holland, M., and Hunke, E.: Improving Representation of Snow on Sea Ice in Climate Models, Eos, 96, https://doi.org/10.1029/2015EO039419, 2015.

Petty, A. A., Webster, M., Boisvert, L., and Markus, T.: The NASA Eulerian Snow on Sea Ice Model (NESOSIM) v1.0: initial model development and analysis, Geoscientific Model Development, 11, 4577–4602, https://doi.org/10.5194/gmd-11-4577-2018, 2018.

Prasad, A., Merkouriadi, I., and Nummelin, A.: Modeling snow on sea ice using physics-guided machine learning, Environmental Data Science, 3, e28, https://doi.org/10.1017/eds.2024.40, 2024.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R.: Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research, 15, 1929–1958, https://jmlr.org/papers/v15/srivastava14a.html, 2014.

Sturm, M. and Holmgren, J.: An Automatic Snow Depth Probe for Field Validation Campaigns, Water Resources Research, 54, 9695–9701, https://doi.org/10.1029/2018WR023559, 2018.

Sturm, M., Perovich, D. K., and Holmgren, J.: Thermal conductivity and heat transfer through the snow on the ice of the Beaufort Sea, Journal of Geophysical Research: Oceans, 107, 8043, https://doi.org/10.1029/2000JC000409, 2002.

Sudom, D., Timco, G., Sand, B., and Fransson, L.: Analysis of first-year and old ice ridge characteristics, in: Proceedings of the 21st International Conference on Port and Ocean Engineering under Arctic Conditions (POAC 2011), pp. 732–742, Montreal, QC, Canada, https://doi.org/10.4224/23004497, 2011.

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I., and Colony, R.: Snow Depth on Arctic Sea Ice, Journal of Climate, 12, 1814–1829, https://doi.org/10.1175/1520-0442(1999)012<1814:SDOASI>2.0.CO;2, 1999.

Zhou, L., Stroeve, J., Xu, S., Petty, A., Tilling, R., Winstrup, M., Rostosky, P., Lawrence, I. R., Liston, G. E., Ridout, A., Tsamados, M., and Nandan, V.: Inter-comparison of snow depth over Arctic sea ice from reanalysis reconstructions and satellite retrieval, The Cryosphere, 15, 345–367, https://doi.org/10.5194/tc-15-345-2021, 2021.