https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Evaluating Arctic sea ice models using simulated microwave brightness temperatures

André Emil Toft Jensen[1], Rasmus Tage Tonboe[2], Hoyeon Shi[1], Suman Singha[1], Mads Hvid Ribergaard[1], Imke Sievers[1], Marcus Huntemann[3], Melody Sandells[4], and Vishnu Nandan[5,6]

[1]Danish Meteorological Institute, Copenhagen, Denmark
[2]National Space Institute, Technical University of Denmark, Lyngby, Denmark
[3]Institute of Environmental Physics, University of Bremen, Bremen, Germany
[4]Geography and Environmental Sciences, Northumbria University, Newcastle, UK
[5]Department of Physics and Technology, UiT The Arctic University of Norway, Tromso, Norway
[6]Department of Geography, University of Calgary, Calgary, Canada

Correspondence: André Toft Jensen (andtoft@space.dtu.dk)

Abstract.

Independently evaluating large-scale sea ice models is challenging because direct observations are sparse and many satellite products are assimilated into the models being evaluated. Moreover, commonly used sea ice concentration products provide limited information on interior pack-ice properties such as ice thickness. To address these issues, we present a framework for evaluating sea ice models directly in satellite microwave radiometer brightness temperature (TB) space. The framework uses an observation operator to simulate all AMSR2 channels except 7.3 GHz. The observation operator couples emission and radiative-transfer models of the ocean, sea ice, snow, and atmosphere for non-melting Arctic conditions. To reduce simulation uncertainty, the snow-scattering parametrisation and multi-year ice scattering properties were constrained using field and airborne observations. We introduce a multi-channel evaluation metric and demonstrate its sensitivity to errors in sea ice concentration, ice-type fractions, thickness, snow depth, and surface temperature. Application to two versions of a pan-Arctic sea ice model showed that improvements and degradations in modelled sea ice thickness are captured by the metric. The metric was significantly correlated with improvements in first-year ice fraction ($R^2 = 0.31$) and ice thickness ($R^2 = 0.14$), but not with snow depth or surface temperature. These results demonstrate that TB-space evaluation provides a complementary and largely independent approach to sea ice model assessment across both the marginal ice zone and interior pack ice.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

1 Introduction

Large-scale sea ice models are crucial for understanding and forecasting sea ice dynamics, simulating climate change, and supporting navigation in polar seas. However, evaluating these models is complicated by the lack of independent validation data. Examples of in situ validation data used for sea ice models include sea ice thickness derived from sonars on submarines or moorings (e.g., Shy et al. 2000; Allard et al. 2018) and sea ice thickness and snow depth derived using mass balance buoys (e.g., Fritzner et al. 2019; Zhang et al. 2023). Airborne observations of sea ice thickness and snow depth have also been used to evaluate sea ice models (e.g., Zhang et al. 2023; Allard et al. 2018). The main limitation of in situ and airborne data is the sparse spatial and temporal coverage. In contrast, many satellites provide excellent coverage, which is advantageous for large-scale modelling. Commonly used satellite-derived variables include sea ice concentration (SIC) or sea ice extent, retrieved from microwave radiometers (e.g., Lavergne et al. 2019), as well as sea ice thickness, retrieved from radar altimeters or low-frequency microwave radiometers (e.g., Ricker et al. 2017). SIC and sea ice extent are widely used for the evaluation of sea ice models; however, they mainly constrain conditions in the marginal ice zone, where variability is highest. More generally, satellite-derived sea ice parameters are often assimilated into the models and therefore cannot be used as independent validation. Instead, models have been assessed by comparing simulations with and without assimilation, or by the Integrated Ice Edge Error metric derived from satellite-observed SIC (e.g., Sievers et al. 2023; Ponsoni et al. 2023; Williams et al. 2026; Smith et al. 2016).

To complement existing evaluation frameworks, we present a methodology in which model output is evaluated in TB space rather than geophysical space. This approach enables evaluation using largely independent observations with extensive coverage, capturing surface conditions both in the marginal ice zone and within the interior ice pack. The method involves simulating TBs from sea ice model output and comparing them with observations from spaceborne radiometers. It is well-suited for microwave radiometers with surface-sensitive channels, such as AMSR2, which is used in this study. Similar satellite radiometers have been operating since the early 1970s and will continue operation with missions such as AMSR3, MWI, and CIMR. Such radiometers provide daily coverage of the polar regions and observe the full sea ice domain. TBs observed by AMSR2 depend on a range of sea ice and snow variables, including SIC, ice and snow density, temperature, brine content, and microstructure (Rückert et al., 2023; Tonboe et al., 2022). Consequently, evaluation in TB space provides a comprehensive assessment of the sea ice and ocean surface state. A comparable approach for climate model validation using observations at 6.9 GHz was presented in Burgard et al. (2020). Our study extends this concept to sea ice model evaluation while exploiting the multi-channel capabilities of the radiometers to maximise the information extracted from the observations. We further propose a simple metric for agreement between simulated and observed TBs, intended to complement conventional per-variable scores.

The challenge in this methodology is that it relies on our ability to accurately calculate or *simulate* the satellite observed TBs from a given geophysical state, including the ocean surface, sea ice, snow, and atmosphere. Physical models of radiative transfer in sea ice and snow exist (Wiesmann and Mätzler, 1999; Tonboe, 2006; Picard et al., 2018), and when coupled with an ocean emissivity and atmospheric radiative transfer model, they form a physics-based *observation operator*, i.e. a function mapping the geophysical state to observation space. However, it is a major challenge that such simulations depend on a detailed

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

description of the sea ice and snow state, since all the necessary parameters are rarely measured in the field or represented in large-scale sea ice models. Furthermore, there is significant uncertainty regarding the scattering processes in snow emission models (Lee and Sohn, 2015). Despite these challenges, satellite observations have been simulated with good accuracy using simplified representations of the geophysical state. In these studies, the state of the atmosphere and ocean is described by standard NWP products. Bulk sea ice parameters such as SIC may come from a sea ice model, and then a detailed layered description of sea ice and snow is achieved by thermodynamic modelling (Tonboe, 2010; Kang et al., 2023) or by introducing an interpolator function before the observation operator (Rückert et al., 2023; Burgard et al., 2020).

Here, we present a framework for evaluating sea ice models in TB space using a physics-based observation operator developed for Arctic sea ice under non-melting spring conditions. To improve the accuracy of the simulated TBs, the observation operator is tuned using snow pit measurements from Cambridge Bay, Nunavut, Canada, and airborne observations collected across the Arctic Ocean during Operation IceBridge. We further introduce a simple multi-channel evaluation metric and investigate its sensitivity to errors in SIC, first-year and multi-year ice fractions, sea ice thickness, snow depth, and surface temperature using controlled experiments. The framework is demonstrated by comparing two versions of a pan-Arctic sea ice model in TB space. By collocating model output and satellite observations with Operation IceBridge observations, we find that improvements in the sea ice model in terms of ice type and thickness are reflected in the proposed TB-space evaluation metric.

2 Data

2.1 AKROSS 2022 Cambridge Bay

To develop the observation operator, we used snow pit measurements collected on first-year ice in the Dease Strait near Cambridge Bay, Canada, in April 2022 during the AKROSS 2022 Cambridge Bay campaign (Meloche et al., 2024). Detailed snow pit profiles were measured at four locations between 20 and 24 April. Temperature, density, and salinity were measured at a vertical resolution of 3 cm directly in the field. The snow microstructure was analysed using X-ray tomography with a μCT scanner. From the tomography images, exponential correlation lengths in the x-, y-, and z-directions were derived at around 1 cm resolution. Under the snow pits, the sea ice thickness was measured, as well as the ice surface temperature, salinity, and in two cases, the salinity of ice plugs 3 cm and 5 cm below the surface, respectively. The AKROSS 2022 Cambridge Bay data used in this study are summarised in Table 1.

2.2 NASA Operation IceBridge

To develop the observation operator and assess the evaluation method, we used airborne measurements of snow depth, sea ice freeboard and thickness, and surface temperature from March to April in 2013, 2014, and 2015, originally sampled during the Operation IceBridge campaign (Kurtz et al., 2013; Studinger, 2020). Snow depth was measured with a snow radar and the total (snow plus ice) freeboard with a laser and optical camera. The ice thickness was then derived by assuming hydrostatic balance and using climatological densities of ice, snow and water. The surface temperature was observed using a Heitronics

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Table 1 Summary of snow pit measurements from the AKROSS 2022 Cambridge Bay campaign used in this study.

	Snow	**Ice below the snow**
Quantity	**Measured at:**	**Measured at:**
Temperature (°C)	3 cm resolution	Surface (+ ice plug in two pits)
Salinity (ppt)	3 cm resolution	Surface (+ ice plug in two pits)
Density ($\mathrm{kg\,m^{-3}}$)	3 cm resolution	No measurement
Exponential correlation length (mm)	Around 1 cm resolution in the x-, y-, and z-directions	No measurement
Thickness (cm)	3 cm layers	Total sea ice thickness

KT-19.85 Series II infrared radiation pyrometer (KT19). Flights were distributed widely over the Central Arctic Ocean, the Lincoln Sea, and the Beaufort Sea, including both first-year and multi-year ice types. We used post-processed snow and ice thickness data from the Round Robin Data Package (RRDP) (Pedersen et al. 2021), where the measurements were averaged over 50 km sections to match the footprint scale of satellite radiometers. We averaged the freeboard and surface temperature data in 50 km segments to match the snow and ice thickness data. We did not filter the KT19 data for clouds or fog.

2.3 DMI-HYCOM-CICE

To demonstrate the proposed evaluation methodology, we used a sea ice model from the Danish Meteorological Institute (DMI), here named DMI-HYCOM-CICE (Ponsoni et al., 2023). DMI-HYCOM-CICE is a coupling of the Hybrid Coordinate Ocean Model (HYCOM) (Chassignet et al., 2007) and the Community Ice CodE model (CICE) (Elizabeth Hunke et al., 2025). The model simulates sea ice with hourly output on a 5 km grid covering most of the Arctic sea ice domain, except the Pacific south of the Bering Strait. Within each grid cell, the diversity of the sea ice is described by six ice-thickness categories. Each ice category is then described by a range of parameters, such as cell area fraction and snow and ice volume. Vertically, each category is resolved by one snow layer and seven ice layers. The sea ice model outputs many parameters, some of which are irrelevant to the emission modelling, e.g., the stress tensor, and some are only relevant under melting conditions, e.g., melt pond fraction. Other variables may influence emissions but are ignored in this study, such as the area fraction of level and ridged ice. The variables used in this study are shown in Table 2.

We compared two versions of DMI-HYCOM-CICE. The first is the current operational configuration (internal version "v7"; hereafter the *baseline model*), while the second is an updated configuration (internal version "Smedv2"; hereafter the *updated model*). A major change in the update was the removal of the previous scaling factor of 0.9 applied to incoming surface shortwave radiation. The scaling was implemented to offset the well-known warm bias in the surface temperatures of ERA5 (Wang et al., 2019; Tian et al., 2024). This led to higher melting rates during summer in the new setup and is expected to strongly affect sea ice properties. Furthermore, while the old setup only uses ERA5 as atmospheric forcing, the updated model uses the CARRA1 west reanalysis (Schyberg et al., 2020), which covers the seas surrounding Greenland. Finally, the CICE code was updated from v6.3.0 to v6.5.0, including modifications to GitHub repos up to 1 March 2024, and similar for HYCOM v2.3.01.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Table 2 DMI-HYCOM-CICE data used in this study. x and y refer to the horizontal dimensions, and z to the vertical. "Per category" means per sea ice thickness category as described in the Section 2.3. "Per grid-cell" just means one number per grid-cell.

Quantity	Dimensions	Per category / Per grid-cell
Sea ice concentration (%)	x, y	Per category
First-year ice fraction (%)	x, y	Per grid-cell
Sea ice thickness (m)	x, y	Per category
Snow depth (m)	x, y	Per category
Snow/ice surface temperature (°C)	x, y	Per grid-cell
Snow temperature (°C)	x, y	Per category
Sea ice temperature (°C)	x, y, z	Per category
Sea surface temperature (°C)	x, y	Per grid-cell
Sea ice salinity (ppt)	x, y, z	Per category
Sea surface salinity (ppt)	x, y	Per grid-cell

Normally, DMI-HYCOM-CICE is assimilated with remote sensing products of SIC and sea surface temperatures, but to achieve a cleaner comparison, we used model runs without data assimilation. The models were initialised on 1 January 2013 and run until and including April 2013. Here, we used data from March and April to compare with the OIB observations. The applied data are made available (see the *Data availability* section).

2.4 AMSR2

In this study, we simulated AMSR2 (Advanced Microwave Scanning Radiometer 2) observations of TBs. AMSR2 observes at a constant inclination angle of 55° in vertical and horizontal polarisation at frequencies 6.9, 7.3, 10.7, 18.7, 23.8, 36.5, and 89.0 GHz. The 7.3 GHz channels are not part of the applied ocean emissivity and atmospheric radiative transfer model (Section 3.1.4), and therefore they were excluded in this work. The instrument is flown onboard the polar-orbiting satellite GCOM-W, and due to its swath width of 1450 km, it provides daily coverage of the Arctic. We used the level 1R product (JAXA, 2012) where all channels have been resampled and centred to match the 6.9 GHz footprint, which has a size of 35 km × 62 km. We corrected the TB observations following JAXA (2015), which calibrates the instrument to its predecessor, AMSR-E.

2.5 ERA5

To specify the atmospheric state for the observation operator, we used reanalysis data from the ERA5 data set (ECMWF Reanalysis v5) (Hersbach et al., 2020) obtained on a horizontal resolution grid of $0.25° \times 0.25°$. We used the following variables: Total column water vapour, total column cloud liquid water, skin temperature, and 10-meter wind speeds in u and v components.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

2.6 Snow depth, ice thickness, and interface temperatures from SIMBA buoys

The observation operator setup includes interpolation of snow and ice temperature profiles (see Section 3.2.2). To interpolate realistic profiles, we fitted an empirical parameter, the *apparent snow thermal conductivity*, to snow depth, ice thickness, snow-ice interface temperature, and air-snow interface temperature derived from SIMBA buoys between 2012 and 2023 (Preußer et al., 2025). We selected Arctic buoys and restricted the analysis to data from March and April.

3 Methods

In this section, we first describe the observation operator setup. Then we present the details of how the observation operator was built and tuned to observations. Finally, we introduce our chosen evaluation metric.

3.1 The observation operator setup

3.1.1 Overview

The observation operator consists of the following main processes (see the flowchart, Figure 1):

1. The input is sea ice and snow data from the model or observations, which are passed through a sea ice and snow interpolator function to generate the required input for the emission model. A tailored interpolator function is necessary for each kind of input data, which in this study is (1) the snow pit data from Cambridge Bay, (2) the OIB data, or (3) the DMI-HYCOM-CICE data.

2. The complete sea ice and snow profiles are input to the sea ice and snow emission model MEMLS (Microwave Emission Model of Layered Snowpacks) extended to sea ice (Section 3.1.2). MEMLS outputs the upwelling TB from the snow/ice surface, which we use to compute the snow/ice surface emissivity ϵ and effective emitting temperature T_{eff} following Tonboe (2023):

$$\epsilon = 1 - \frac{T_b(T_{sky} = 100\,\mathrm{K}) - T_b(T_{sky} = 0\,\mathrm{K})}{100\,\mathrm{K}}, \qquad T_{eff} = \frac{T_b(T_{sky} = 0\,\mathrm{K})}{\epsilon},$$

   where T_{sky} is the atmosphere downwelling TB. The ϵ and T_{eff} are computed for each radiometer channel and each surface type within the grid cell. More on this in the following sections. The calculation of ϵ and T_{eff} allows us to couple the surface emission model to the atmosphere radiative transfer model (RTM).

3. ϵ and T_{eff} of each ice surface type are re-gridded, together with the ERA5 data, to the satellite swath.

4. The W&M atmosphere RTM (Section 3.1.4) computes the TOA TB, i.e., the AMSR2 observation, given the surface ϵ and T_{eff} and the atmospheric state given by ERA5. In this step, we also calculate the ocean surface ϵ and T_{eff} using the W&M ocean surface emission model (Section 3.1.4). The RTM is run per surface type. The types of surfaces are: ocean, first-year/multi-year ice and the different ice thickness categories in DMI-HYCOM-CICE.

5. Finally, the total TB is computed as the area-weighted average.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

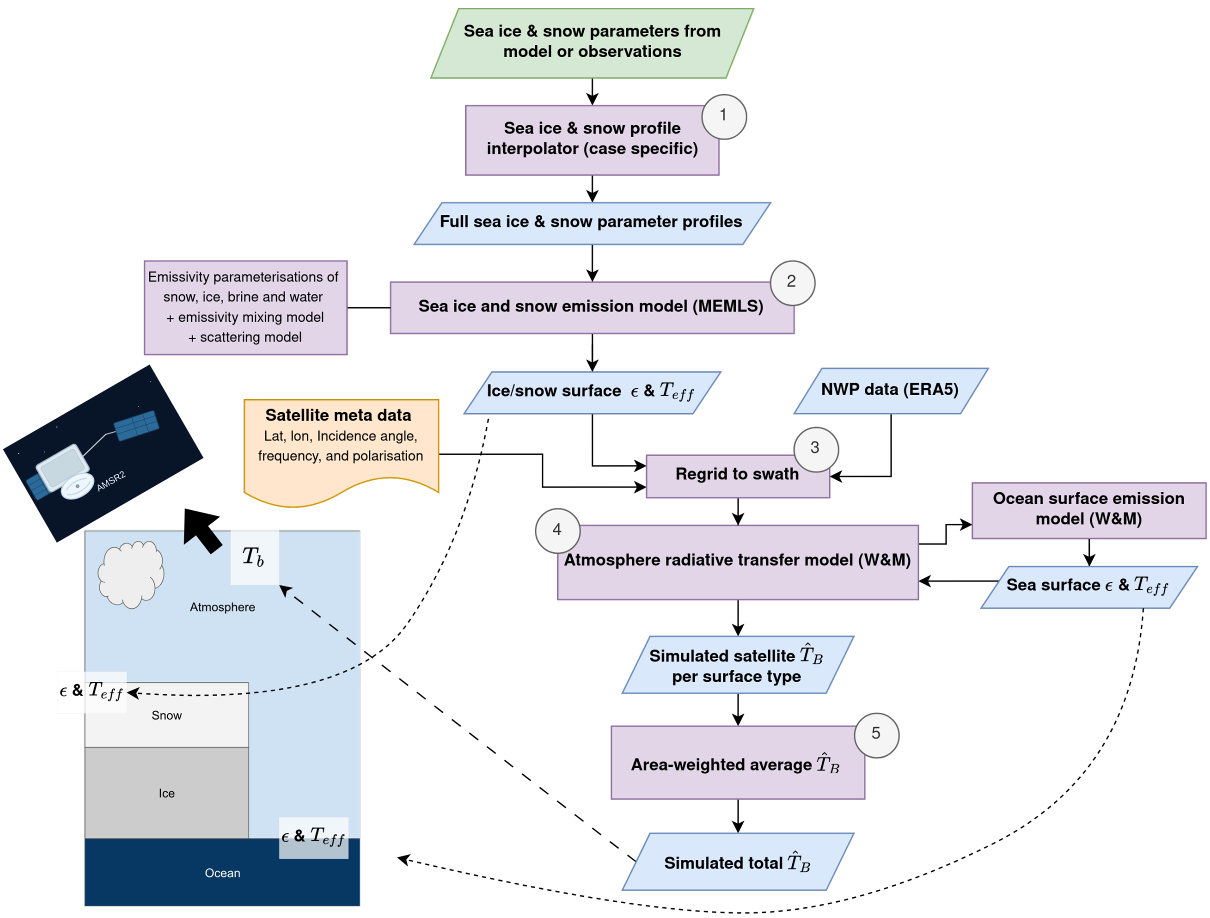

Figure 1. Observation operator flow chart. The numbering matches the description in Section 3.1.1.

3.1.2 Sea ice and snow emission model, MEMLS

MEMLS is a physical microwave emission model based on six-flux theory, originally developed for terrestrial snowpacks (Wiesmann and Mätzler, 1999), and later extended to include sea ice (Tonboe, 2006). MEMLS considers a 1-D profile of layered sea ice and snow on top of a thick ocean surface. For each layer, the user specifies the input variables: layer thickness, temperature, salinity in ppt or brine volume fraction, density, wetness, exponential correlation length l_{ex} (a statistical measure of the distance between scatterers) and the layer type (snow type, or first-year ice or multi-year ice). From this, MEMLS computes the permittivity, absorption coefficient and scattering coefficient of each layer and the refraction and reflection between layers using physical (or optional empirical) models.

For modelling the scattering coefficient of snow, MEMLS has several options, and the choice of scattering model is important (Kang et al., 2023). Here, we used the physical model of volume scattering, the Improved Born Approximation, derived for spheres and thin oblate spheroids with isotropic orientation (Mätzler, 1998; Mätzler and Wiesmann, 1999). We determine the

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

final scattering coefficient as a weighted average of the six-flux scattering coefficient of both shapes, where the weight w_γ is tuned to observations from the Cambridge Bay dataset (see Section 3.2.1) and assumed constant across all simulations.

Modelling the permittivity of sea ice and snow is not straightforward, as these are mixed anisotropic media of ice, air, and brine. To calculate the effective permittivity of sea ice, we used the *Polder and van Santen* two-phase mixing model assuming a host material with ellipsoidal inclusions (Jones and Friedman, 2000). The result depends on the shape and orientation of the ellipsoids, but in MEMLS the media is assumed to be isotropic, and therefore we can only model randomly isotropically oriented inclusions. We took ice as the host material. First, brine inclusions are embedded in the ice, assuming an aspect ratio of 10^3, corresponding to long needle-like shapes. The resulting ice-brine mixture is then treated as the host medium, into which air inclusions are embedded with an assumed aspect ratio of 1, corresponding to spherical shapes. The permittivity of dry snow is calculated as described in Wiesmann and Mätzler (1999). The bottom layers of snow on first-year ice contain brine. To mix brine and dry snow, we used an empirical mixing model described in Geldsetzer et al. (2009). The empirical parameters are based on measurements of permittivity at 50 MHz, and then the model is theoretically extrapolated to AMSR2 frequencies. Calculating the permittivity of brine-wetted snow using the theoretical approach that we applied for sea ice resulted in significantly different values for the complex permittivity. We encourage more research to constrain the estimates at these frequencies.

We adopted the MEMLS code from Rückert et al. (2023). This implementation is written in Julia (Bezanson et al., 2017), making it a fast model, which is necessary to deal with the large grid of the sea ice model. We only changed the formulation of the snow scattering coefficient as mentioned, implemented the empirical permittivity mixing model for brine-wetted snow as mentioned, and changed the model to take brine volume fraction as input rather than salinity, as described below.

3.1.3 Conversion between sea ice and snow salinity and brine volume fraction

The sea ice and snow permittivity parametrisations depend on brine volume fraction (v_b), but in the field, brine content is measured in terms of salinity (S). Therefore, *per layer* conversion formula between S and v_b are needed. The brine volume fraction of sea ice ($v_{b,i}$) is calculated as a function of sea ice temperature, porosity, and salinity following Cox and Weeks (1983). The brine volume fraction of snow ($v_{b,s}$) is calculated as a function of snow density (ρ_s), temperature, and salinity (Drinkwater and Crocker, 1988):

$$v_{b,s} = \frac{v_{b,i}\rho_s}{(1 - v_{b,i})\rho_i + v_{b,i}\rho_b}. \tag{1}$$

In this case, we calculate $v_{b,i}$ assuming zero porosity. The pure ice density ρ_i and brine density ρ_b are calculated following Pounder (1965) and Cox and Weeks (1975), respectively:

$$\rho_i = 917 - 0.1403T, \qquad \rho_b = 1000.0 + 0.8S_b, \tag{2}$$

where T is the temperature in °C and S_b is the brine salinity in ppt. In thermodynamic equilibrium, the brine salinity increases when the temperature of an ice-brine system drops (see e.g. the phase diagram in Assur 1960). We calculated S_b following

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Vant et al. (1978) (we allow the temperature to go to the freezing point whereas the authors put 2 °C as the limit):

$$S_b(T) = \begin{cases} 1.725 - 18.756T - 0.3964T^2, & -8.2\,^\circ\mathrm{C} \leq T, \\ 57.041 - 9.929T - 0.16204T^2 - 0.002396T^3, & -22.9 \leq T < -8.2\,^\circ\mathrm{C}, \\ 242.94 + 1.5299T + 0.0429T^2, & -36.8 \leq T < -22.9\,^\circ\mathrm{C}, \\ 508.18 + 14.535T + 0.2018T^2, & -43.2 \leq T < -36.8\,^\circ\mathrm{C}. \end{cases} \tag{3}$$

3.1.4 Ocean emissivity and atmosphere radiative transfer model, W&M

The applied ocean emissivity and atmosphere RTM are modules of the ocean surface observation operator by Wentz and Meissner (2000). We name them the W&M ocean emissivity model and the W&M RTM. Both are semi-empirical models: The ocean emissivity model calculates the ocean emissivity from wind speed, surface temperature, and surface salinity (we ignored the wind direction component, as it did not improve the simulations). The atmosphere RTM computes AMSR-E/AMSR2 observations given surface parameters ϵ and T_{eff} and vertically integrated atmospheric variables from ERA5: total column water vapour and total column cloud liquid water. In addition, the original model parametrised the effect of ozone by sea surface temperature. We replaced the sea surface temperature with the area-weighted average of the sea surface temperature and the sea ice/snow surface temperature. We used a Python implementation by the authors of Andersen et al. (2006) and edited the RTM to take user-provided sea ice ϵ and T_{eff} to run it coupled with MEMLS as described in Section 3.1.1.

3.2 Building the observation operator

The observation operator is built in three steps:

1. We simulate AMSR2 observations using the snow pit measurements from Cambridge Bay and tune the weight w_γ (associated with snow scattering) against observations. Because the ice in Cambridge Bay is exclusively first-year ice, the results are independent of the parameter learned in step 2.

2. We simulate AMSR2 observations using the OIB data and then tune the assumed value of the exponential correlation length (l_{ex}) in the freeboard section of multi-year ice.

3. Finally, we set up the observation operator to work on output from the pan-Arctic sea ice model. This involved the parameters learned in steps 1 and 2.

All three steps involve inferring snow and ice parameters where necessary, i.e., building a profile interpolator as introduced in Section 3.1.1. To infer parameters, we use thermodynamic arguments or climatologies/observations as described in the following sections.

3.2.1 Step 1: Using snow pit measurements from Cambridge Bay to tune the scattering formulation

Step 1 involved tuning the snow scattering parametrisation; specifically, the coefficient w_γ. We used the snow pit measurements from Cambridge Bay (see Section 2.1) to simulate AMSR2 TBs and tune to observations. To do so, three issues were addressed:

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

i) the snow l_{ex} was measured in the x, y and z-directions, but MEMLS models scattering assuming isotropic distributions of scatterers, ii) sea ice profiles were not measured, and iii) the snow pits were located in a fjord more narrow than the satellite footprint.

First, we estimated the snow l_{ex} as interpreted by MEMLS by projecting the anisotropic structure onto the line of sight given by the incidence angle θ:

$$l_{ex} = \sin^2(\theta) \cdot l_{ex}^{horizontal} + \cos^2(\theta) \cdot l_{ex}^{vertical} \tag{4}$$

where $l_{ex}^{horizontal}$ is the correlation length measured in the horizontal direction and $l_{ex}^{vertical}$ in the vertical. Here, $l_{ex}^{horizontal}$ is taken as the average l_{ex} measured in the x and y-directions. To get l_{ex} at the same resolution as the rest of the data, we interpolate and extrapolate the data linearly and then take the average across each layer.

Secondly, we inferred the sea ice parameter profiles. In each snow pit, they measured the thickness of the sea ice, the surface temperature, and the surface salinity. We interpolated the temperature of the sea ice linearly to the bottom, where it was assumed that the ice is $-1.8\,^{\circ}\mathrm{C}$ (the freezing temperature of sea water). To interpolate the salinity, we fitted a function describing winter sea ice salinity to simulation results from Vancoppenolle et al. (2006) (details in Appendix A). The porosity of the sea ice was assumed to be 0.5 % to match typical first-year ice densities of $920\,\mathrm{kg\,m^{-3}}$ (Timco and Weeks, 2010). The density of sea ice was then computed as a function of temperature, salinity, and porosity following Cox and Weeks (1983), Equation 15. The sea ice l_{ex} was calculated using Equation 4 with $l_{ex}^{horizontal} = 0.16\,\mathrm{mm}$ and $l_{ex}^{vertical} = 0.56\,\mathrm{mm}$ as measured in Tonboe et al. (2026). The measured snow and interpolated sea ice parameters are plotted in Figure 2.

Thirdly, we addressed the issue that the Dease Strait is narrow compared to the satellite footprint. Satellite measurements that cover the measurement sites are disturbed by land surfaces. Therefore, instead of comparing with satellite observations that directly cover the snow pits, we found suitable observations in the wider Queen Maude Gulf (Figure 3). The sea ice in this region is fairly laterally homogeneous, and therefore, we assume that the surface emission does not vary significantly between these locations. This assumption is supported by the fact that AMSR2 TBs from the same day varied relatively little across the fjord. This is shown for the 36 GHz horizontally polarised channel in Figure 3. For each day of in situ data, we selected four AMSR2 observations.

The observation operator is then used to simulate the AMSR2 TBs corresponding to each snow pit with atmospheric profiles as if they were located in the satellite footprint centres. Hence, each snow pit is simulated with four different atmospheric profiles. ERA5 profiles are selected by the nearest neighbour method. The observation operator is run with varying values of w_γ, and we find the best fit by minimising a χ^2-like quantity:

$$\chi^2 = \sum_{i,c} \frac{(y_{i,c} - \hat{y}_{i,c})^2}{{\sigma_c}^2} \tag{5}$$

where i is the sample and c the channel. σ_c is the uncertainty of a given channel, which is estimated as the standard deviation of the observations shown in Figure 3; see the values in Table 3. The 89 GHz channels were not used for fitting as we achieved relatively poor results for these channels (Figure 6). The best fit gave $\hat{w}_\gamma = 0.646$, i.e., the snow scatterers are modelled as 64.6 % spheres and 35.4 % thin oblate spheroids.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

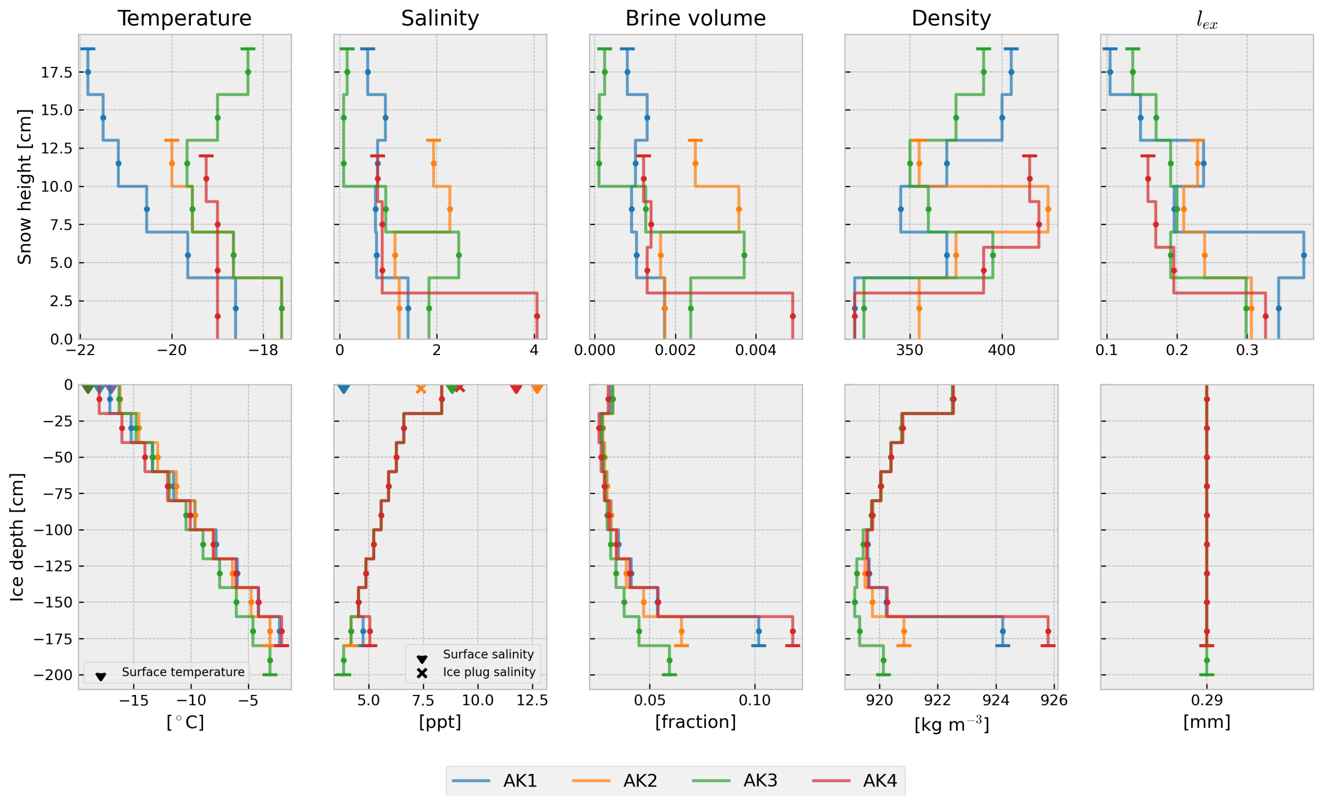

Figure 2. AKROSS 2022 Cambridge Bay snow pit and sea ice parameter profiles. Top: Snow pit profiles measured at the four sites. Bottom: Sea ice surface and plug measurements and interpolated profiles at the four sites.

Table 3 Estimated simulation uncertainty per channel (σ_c) when simulated using the Cambridge Bay snow pit data.

Channel	V06	H06	V10	H10	V18	H18	V23	H23	V36	H36
σ_c [K]	0.55	1.37	0.54	1.62	0.71	1.96	0.76	2.14	0.98	2.35

3.2.2 Step 2: Using airborne measurements from the OIB campaign to tune l_{ex} in the freeboard of multi-year ice

In this step, we constrained the l_{ex} in the freeboard of multi-year sea ice. This parameter is required in the interpolation function for the sea ice model and is important for the microwave scattering. We opted to tune it to the AMSR2 observations, since there is little observational evidence of it.

The OIB data introduced in Section 2.2, contains surface temperature (T_{sfc}), sea ice thickness (h_i), total freeboard (F_t), and snow depth (h_s) collocated with AMSR2 TBs. Additionally, we estimated first-year ice and multi-year fractions using AMSR2 observations and the NASA Team algorithm (Cavalieri et al., 1984) with tie-points derived for AMSR-E in Ivanova et al. (2015). We assumed 100 % SIC at all points. Figure 4 shows the locations of collocated OIB samples coloured by first-year ice

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

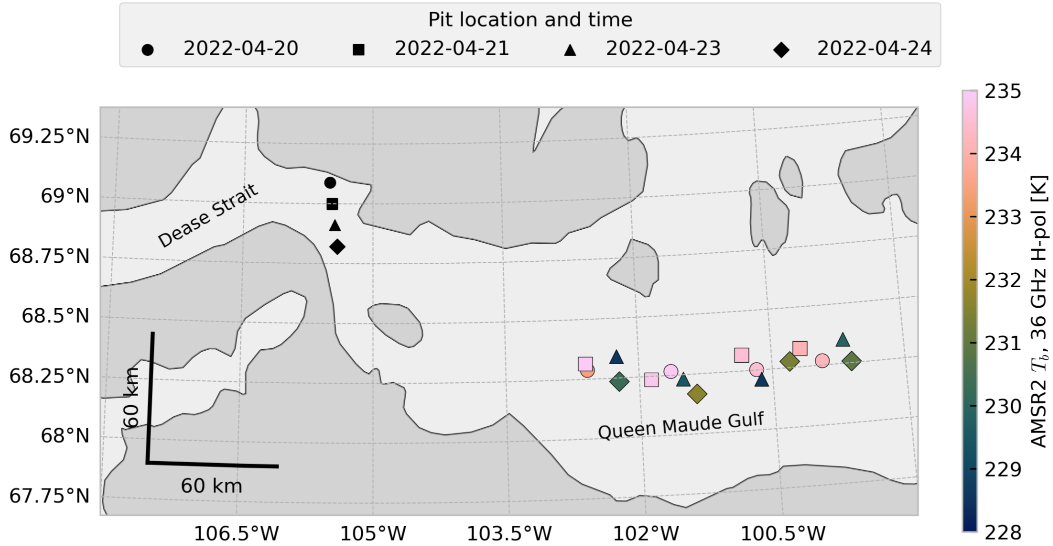

Figure 3. AKROSS 2022 Cambridge Bay pit locations (black markers) and the footprint centres of selected AMSR2 observations. AMSR2 observations are colour-coded by the TB of the 36 GHz horizontally polarised channel. The marker types indicate date.

260 fraction. The ERA5 profiles were selected by the nearest neighbour method. To describe the surface temperature, we used the ERA5 skin temperature field since it better matches the spatial scale of the AMSR2 footprint than the KT19 data. However, the ERA5 skin temperature is corrected based on the KT19 data (details in Appendix B).

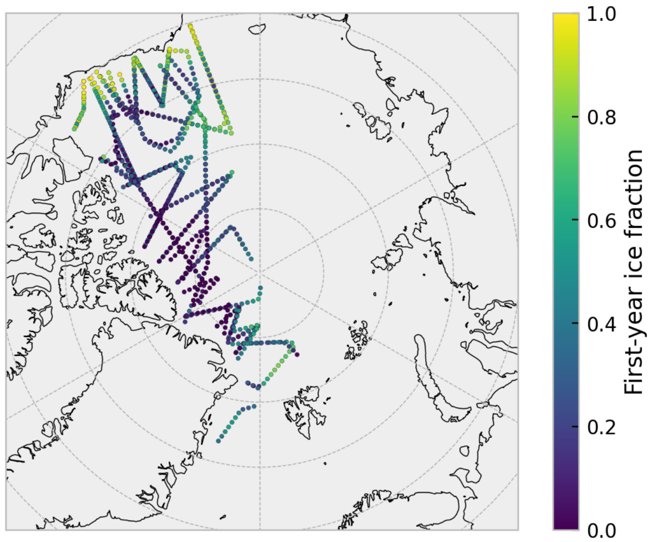

Figure 4. The locations of the resampled OIB data coloured by first-year ice fraction derived from AMSR2 observations.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

The interpolator aims to reconstruct physically consistent snow and ice profiles from the limited set of bulk variables (ice type, snow depth, total freeboard, ice thickness, and surface temperature), using a combination of thermodynamic constraints and empirically derived parametrisations. The interpolator parameters are shown in Table 4, and example profiles of a first- and multi-year ice column are shown in Figure 5. We interpolate to 10 ice and 4 snow layers. Further increasing the vertical resolution of the ice and snow packs did not significantly affect the simulated TBs. Snow layers are classified as either wind slab or depth hoar, and sea ice layers as freeboard or draft. The wind slab layers are placed over the depth hoar layers, and the relative fraction of the total thickness of the wind slab and the depth hoar is fixed. The ice freeboard (F_i) is the total freeboard minus the snow depth. The porosity and correlation length profiles are set with prescribed values for each type of layer.

The salinity of first-year ice is again approximated to the modelled profiles in Vancoppenolle et al. (2006) (see appendix A). The multi-year ice salinity profile is calculated as a linearly increasing function from zero at the top of the ice draft to match the bulk salinity (ppt) observed in Kovacs (1996):

$$S_{bulk}^{MYI} = 1.85 + \frac{8.02179}{\max(h_i, 200)^2}, \tag{6}$$

where h_i is in cm (the "max" function was introduced to ensure realistic values). Snow on multi-year ice is assumed to be fresh (without salt), while the brine volume fraction in first-year ice snow is determined as an exponential function decreasing with height above the ice surface z:

$$v_{b,s}(z) = v_0 e^{-z/\tau_{v_b}}. \tag{7}$$

v_0 and τ_{v_b} are approximated to the climatological profile of brine volume fraction shown in Nandan et al. (2017), giving $v_0 = 3.37\,\%$ and $\tau_{v_b} = 5.4\,\mathrm{cm}$.

The temperature is determined by linear interpolation between the snow surface temperature T_s, the temperature of the ice-snow interface T_{si} and the temperature of the ice-ocean interface $T_w = -1.8\,^{\circ}\mathrm{C}$. T_{si} is computed from a steady-state equation assuming constant temperature gradients in snow and ice (Tonboe, 2023):

$$T_{si} = \frac{k_i h_s T_w - k_s h_i T_s}{k_i h_s + k_s h_i}, \tag{8}$$

where k_i and k_s are thermal conductivities. Approximate values for k_i and k_s are $2\,\mathrm{W\,K^{-1}\,m^{-1}}$ (Pringle et al., 2006) and $0.3\,\mathrm{W\,K^{-1}\,m^{-1}}$ (Sturm et al., 1997), respectively. Here, we assumed $k_i = 2.1\,\mathrm{W\,K^{-1}\,m^{-1}}$, and then fitted k_s to buoy observations from Preußer et al. (2025) (see Appendix C). Hence, k_i and k_s should here be interpreted as "apparent" and not actual thermal conductivities. Finally, the density of each layer is computed following Cox and Weeks (1983).

The AMSR2 TBs were simulated for a first-year ice and a multi-year ice surface, and the total was calculated as the area-weighted average. Again, we tuned by minimising the χ^2-like quantity (Equation 5). Now, the channel uncertainties (σ_c) were estimated as the standard deviation of the errors of a first-guess simulation run on the OIB data. We obtained $l_{ex} = 0.61\,\mathrm{mm}$ in the freeboard of multi-year ice. Simulation results are seen in Section 4.2.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Table 4 Parameters used for sea ice and snow profile interpolation as described in Section 3.2.2. (*) The value of l_{ex} in multi-year ice freeboard was determined by fitting to observed TBs as described in the text.

Quantity	Value		Source / justification
	First-year ice	**Multi-year ice**	
Depth hoar fraction of total snow pack	0.2	0.5	King et al. (2020)
Snow wind slab porosity	0.64	0.67	To match density measurements in King et al. (2020)
Snow depth hoar porosity	0.73		To match density measurements in King et al. (2020)
Ice freeboard porosity	0.5 %	10 %	To match typical densities (Timco and Weeks, 2010)
Ice draft porosity	0.5 %		To match typical densities (Timco and Weeks, 2010)
l_{ex}, wind slab snow	0.16 mm		The AKROSS Cambridge Bay 2022 data
l_{ex}, depth hoar snow	0.32 mm		The AKROSS Cambridge Bay 2022 data
l_{ex}, ice freeboard	0.29 mm	0.61 mm*	Tonboe et al. (2026) and Equation 4
l_{ex}, ice draft	0.29 mm	0.29 mm	Tonboe et al. (2026) and Equation 4
Apparent ice thermal conductivity, k_i	$2.1\,\mathrm{W\,K^{-1}\,m^{-1}}$		Pringle et al. (2006) (appendix C)
Apparent snow conductivity, k_s	$0.388\,\mathrm{W\,K^{-1}\,m^{-1}}$		Preußer et al. (2025) (appendix C)
Snow bottom v_b	3.37 %	0	Nandan et al. (2017)
Salinity exponential vertical decay scale, τ_S	5.4 cm		Nandan et al. (2017)
w_γ	0.646		$\hat{w}_\gamma$ described in Section 3.2.1

3.2.3 Step 3: Setting up the observation operator for the pan-Arctic sea ice model, DMI-HYCOM-CICE

The final step was to design the profile interpolator for the output of DMI-HYCOM-CICE (the sea ice model introduced in Section 2.3). Note that grid cells with temperatures above the freezing point are excluded as the observation operator is not valid under melting conditions.

The model provides a relatively detailed description of the sea ice state, but still, a profile interpolator is required to infer parameters that were not modelled. We chose to stick with the native model vertical resolution of 7 ice layers, though it is lower than what we found to converge in the OIB study. The height of the ice freeboard is approximated following Alexandrov et al. (2010):

$$F_i = \frac{\rho_w - \rho_i}{\rho_w} h_i - \frac{\rho_s}{\rho_w} h_s, \tag{9}$$

where ρ_s is a climatological bulk snow density of $324\,\mathrm{kg\,m^{-3}}$ on first-year ice and $320\,\mathrm{kg\,m^{-3}}$ on multi-year ice, and ρ_i is a climatological bulk ice density of $916.7\,\mathrm{kg\,m^{-3}}$ for first-year ice and $882\,\mathrm{kg\,m^{-3}}$ for multi-year ice. The density of sea water (ρ_w) is $1025\,\mathrm{kg\,m^{-3}}$. As mentioned, first-year ice fraction was provided *per grid cell* in the model, but we want the fraction of ice type *per category*. Thus, we assume the thin ice categories to be first-year and the thick ice to be multi-year ice; see Appendix D for the exact equation. When dividing ice categories this way, we sometimes obtained unrealistically high salinity in the upper layers of multi-year ice. Therefore, we overwrite the modelled multi-year ice salinity in the same way

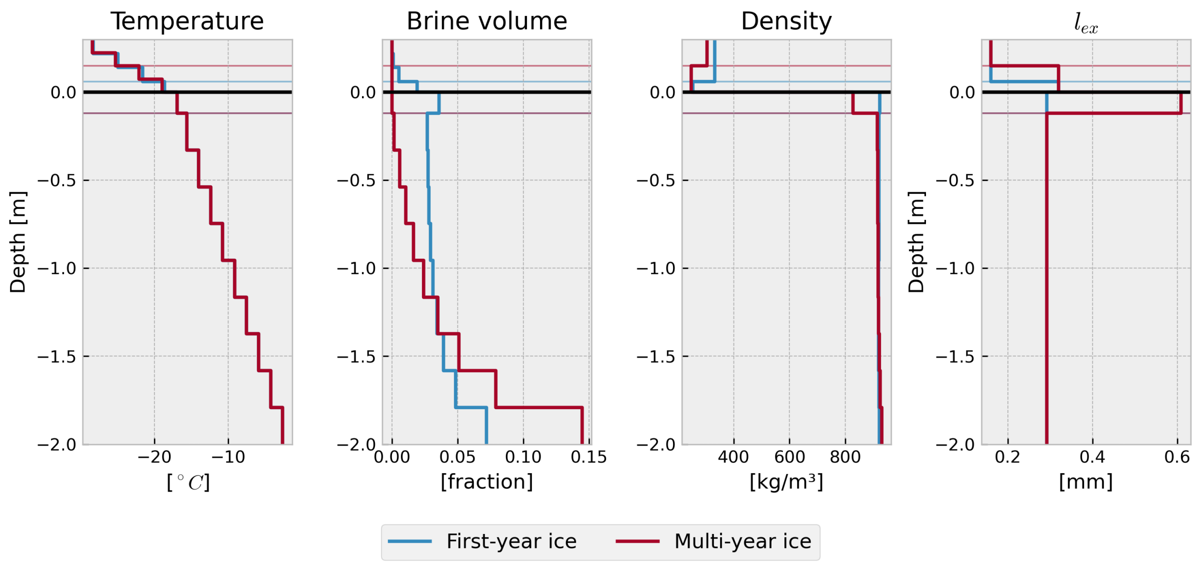

Figure 5. Examples of parameter profiles interpolated by the OIB study interpolator. The plots show a first-year ice (blue) and multi-year ice (red) column with input parameters snow depth, ice thickness, ice freeboard, and surface temperature: 30 cm, 200 cm, 12 cm, and -30 °C. The profiles are located with the ice-snow interface at depth zero, and the freeboard and "depth hoar, wind slab" interface are shown with vague horizontal lines.

as described in Step 2 (Equation 6). This ensures physically realistic input to the emission model, but implies that this part of the thermodynamic state is prescribed rather than evaluated. Sea ice porosity and correlation length are not predicted by DMI-HYCOM-CICE, and therefore they were inferred using the values in Table 4.

DMI-HYCOM-CICE only has a single snow layer, which we split into four layers for a more detailed description. First, the ice-snow interface temperature is computed using Equation 8, but with h_s, h_i, T_s, and T_w being replaced by the half snow pack depth, half top ice layer thickness, and the temperature of the snow pack and top ice layer. Values of k_s and k_i are the same as in Step 2. Then, snow temperature, salinity, porosity, and correlation length are inferred as described in Step 2. The densities of both ice and snow are computed following Cox and Weeks (1983).

After interpolating the snow and ice profile, we run MEMLS to compute ϵ and T_{eff} for each ice thickness category and for each ice type (first-year / multi-year). These values, along with the ERA5 data, are re-gridded onto a matching AMSR2 swath using Gaussian averaging of the 8 nearest neighbours with $\sigma = 14$ km chosen to approximately match the AMSR2 footprint size. Then, the TBs are simulated for each surface type and the total TBs are found by taking the area-weighted average.

3.3 Assumptions and limitations of the observation operator

The proposed observation operator relies on several simplifying assumptions. First, all simulations assume horizontally homogeneous conditions within each satellite footprint, such that sub-footprint variability in sea ice, snow, and atmospheric

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

properties is not explicitly resolved. This assumption is somewhat alleviated by the fact that sub-footprint variability is partly captured through the ice thickness distribution in DMI-HYCOM-CICE. Second, MEMLS is based on a one-dimensional, plane-parallel representation of the snow–ice column, which neglects lateral heterogeneity and three-dimensional scattering effects. Third, ice and snow layers are assumed to be smooth, which may not be a good assumption in regions dominated by deformed or ridged ice.

In addition, key components of the emission modelling depend on parametrisations that introduce uncertainty. The scattering formulation is tuned using a limited number of in situ observations (Section 3.2.1), which may affect the generality of the results across regions and conditions. Also, little validation of the permittivity mixing models for sea ice and snow at these frequencies exists, and the model assumptions may be crude. The empirical mixing model for brine-wetted snow was only tested at 50 MHz, and so the theoretical extrapolation to high frequencies may be imprecise. The Polder and van Santen mixing model used for sea ice was developed for two-phase systems, but sea ice is a three-phase system. Furthermore, brine inclusions in sea ice and snow are anisotropic, which is not taken into account.

Additional uncertainties may arise from the atmospheric state, which is taken from NWP reanalysis data, as well as from limitations in the atmospheric radiative transfer model.

Finally, several geophysical properties required by the observation operator are not provided by the sea ice model and must therefore be inferred through interpolation or empirical parametrisations. These introduce additional uncertainties that cannot be clearly attributed to errors in the sea ice model itself.

3.4 Application of the observation operator in model evaluation

Here, we describe how we used the observation operator to do model evaluation. The operator simulates TBs, which should be compared to observations to result in a *divergence* or *error*:

$$e_{T_b,c} = \hat{T}_{b,c} - T_{b,c}. \tag{10}$$

where c is the radiometer channel index. e_{T_b} will be mainly a sum of errors from the observation operator and from the sea ice model. Errors in any channel tells something different about the sea ice model performance, but it is hard to interpret all channels at the same time. One may define any kind of clever metric based on these errors, but we suggest a simple one: the "weighted TB RMSE", defined as:

$$\mathrm{wRMSE}_{T_b} = \sqrt{\sum_c \frac{w_c (e_{T_b,c})^2}{\sum_c w_c}} \tag{11}$$

where the weights $w_c = \frac{1}{\sigma_c^2}$ are the inverse of the channel uncertainty squared. σ_c was estimated by the RMSE of the simulations using OIB observations (Table 5). Lower values of wRMSE$_{T_b}$ indicate a better agreement between the simulated and observed brightness temperatures. This metric convolves the information from all channels into one number per sample. This convolution results in a loss of information, but makes the results more robust and easy-to-interpret.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

4 Results and Discussion

4.1 Simulations using snow pit measurements from Cambridge Bay

Figure 6 shows the simulation results using the AKROSS 2022 Cambridge Bay data as part of Step 1 of building the observation operator (Section 3.2.1). These are the results after tuning $\hat{w}_\gamma$ to 0.646. Simulations of vertically polarised (V-pol) channels show reasonable accuracy, though TBs at 6 GHz, 36 GHz, and 89 GHz are overestimated. Simulations in horizontally polarised (H-pol) channels are - outside the 89 GHz channel - within the range of the observations. Simulations of the 89 GHz channels show more variability due to differences in both atmospheric (given by the width of the coloured lines) and surface conditions (given by the difference between the coloured lines). Generally, there is a large spread in TB between snow pits and relatively limited variation due to atmospheric effects. These results suggest that the observation operator captures the emission physics well, despite the assumptions listed in Section 3.3 and the spatial mismatch between the snow pits and the satellite footprints. The variation in TBs between snow pits indicates a high sensitivity to local snow and ice conditions, suggesting a complex coupling between satellite observation and the distribution of sea ice and snow properties within the satellite footprint. This coupling highlights the need to better account for sub-footprint variability in future studies.

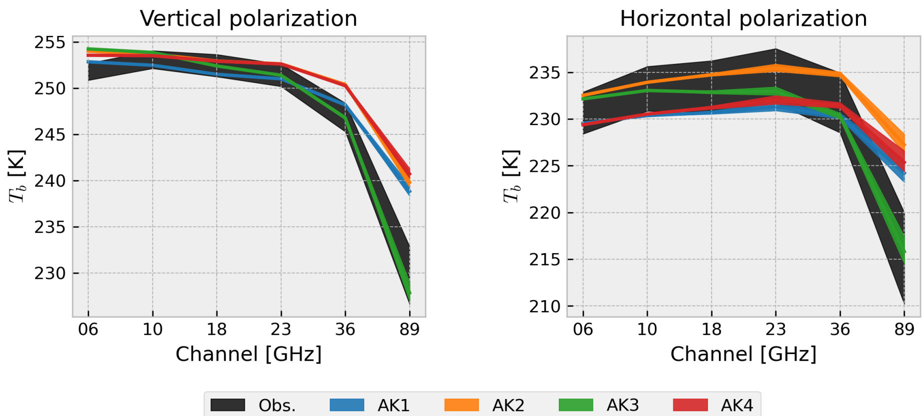

Figure 6. Simulations using the AKROSS 2022 Cambridge Bay snow pit data and corresponding observations. The black shadings show the min-max range of all observations in the gulf (see locations in Figure 3), and the coloured shadings show the min-max range simulated from each snow pit given different atmospheric profiles matching the satellite footprints.

4.2 Simulations using airborne measurements from Operation IceBridge

Here we present the TB simulations using airborne observations from OIB as input to the observation operator (Section 3.2.2). The performance statistics are shown in Table 5. The biases are small relative to the RMSEs. R^2 values increase with frequency, except for the 89 GHz channels, for which the R^2 is slightly lower than the 23 GHz and 36 GHz channels. RMSEs also increase with frequency, which is explained by the larger variation in the higher frequency channels. V-pol channels are simulated better

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

than H-pol channels at low frequencies, but worse at high frequencies. Of these simulations, the average wRMSE$_{T_b}$ was 4.1 K with a standard deviation of 1.6 K.

Table 5 Performance metrics of simulations using the OIB data as input to the observation operator.

Stat	V06	H06	V10	H10	V18	H18	V23	H23	V36	H36	V89	H89
Bias [K]	0.26	0.76	2.31	1.56	0.97	-0.76	0.25	-0.00	-4.99	-3.12	-2.67	-0.42
RMSE [K]	2.62	3.95	3.96	3.94	5.13	4.51	5.36	4.88	7.71	5.70	6.69	5.28
R^2	0.36	0.30	0.37	0.54	0.73	0.81	0.83	0.86	0.89	0.91	0.82	0.81

Selected channels of observed versus simulated TBs are plotted as scatter plots in Figure 7 coloured by first-year ice fraction. These plots highlight clear differences between channels: TB variability increases with frequency, and the contrast between first-year and multi-year ice becomes more pronounced. The increasing separation between first-year and multi-year ice enhances the variance in TB, which contributes to higher R^2 values at higher frequencies seen in Table 5.

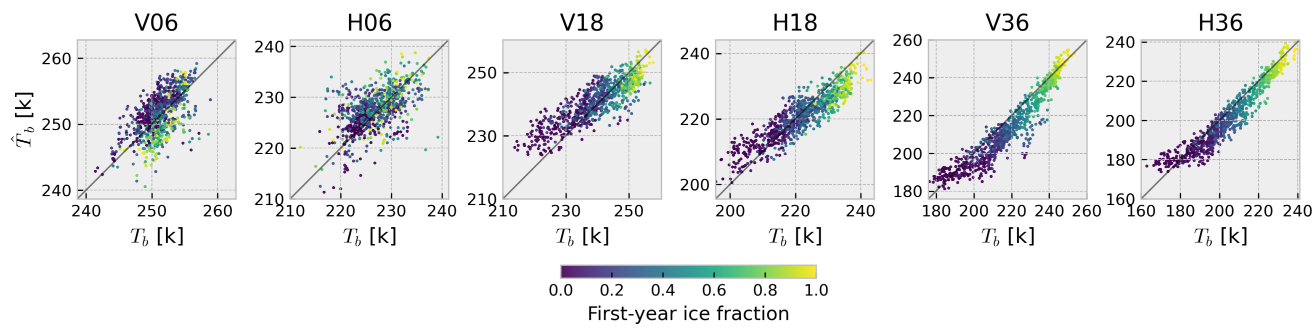

Figure 7. Simulated versus observed AMSR2 TBs using OIB data as input to the observation operator for selected channels. Colours indicate the first-year ice fraction as determined by the NASA Team algorithm.

These results demonstrate that AMSR2 TBs can be simulated with reasonable accuracy despite the simplified representation of the ice and snow landscape, spatial mismatch between the AMSR2 footprint and OIB observations, and the limited set of input parameters. The simulation errors are ideally independent of the absolute size of geophysical parameters. Figures with TB simulation errors plotted against snow depth and ice thickness are found in the Appendix E; most channels show little dependence of the simulation error on snow depth and ice thickness. However, some non-linear dependencies remain, particularly in the 36 and 89 GHz channels.

4.3 wRMSE$_{T_b}$, sensitivity tests

We tested the sensitivity of wRMSE$_{T_b}$ - our proposed evaluation metric - to errors in bulk sea ice parameters: SIC, first-year ice fraction, sea ice thickness, snow depth, and surface temperature. To do so, we used the interpolator function designed for the OIB data (Section 3.2.2) as well as the Equation 9 to estimate the ice freeboard height from snow depth and ice thickness. We

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

define perturbations relative to a reference state with 100 % SIC, where the ice is described by one of two ice-snow profiles: A first-year ice column and a multi-year ice column with ice thicknesses of 1.5 meters and 3.0 meters, snow depths of 10 cm and 25 cm, and surface temperatures of $-15\,^{\circ}\mathrm{C}$ and $-25\,^{\circ}\mathrm{C}$ respectively. Both profiles are combined with the same atmospheric profile. The sensitivities are tested by perturbing one parameter at a time and computing the wRMSE$_{T_b}$ with the unperturbed state as the reference. The magnitude of the perturbations is chosen to be representative of typical sea ice model errors shown later. The results are seen in Figure 8.

The wRMSE$_{T_b}$ responds almost linearly to all five types of perturbation, but depends on the reference ice columns. The non-linearity in response to snow depth appears when the snow depth approaches zero. On the scale of errors selected here, errors in SIC result in the largest wRMSE$_{T_b}$ by far, explained by the large contrast in emissivity between ice and ocean surfaces. Errors in other parameters produce more moderate increases in wRMSE$_{T_b}$. These results show how our error metric ideally responds to different kinds of sea ice model errors, but note that in sea ice models, sea ice parameters are often correlated and so are sea ice model errors; some correlations amplify and others dampen the wRMSE$_{T_b}$ signal.

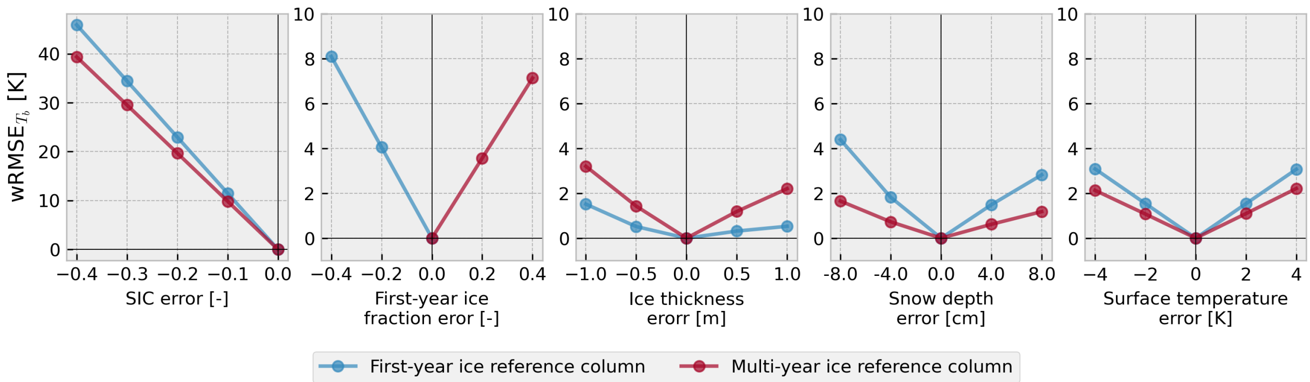

Figure 8. The sensitivity of the wRMSE$_{T_b}$ to errors in bulk sea ice parameters. The x-axes show the sizes of the errors, and the y-axes show the wRMSE$_{T_b}$ when compared to the reference state. The reference state was a 100 % ice-covered surface, where the ice was modelled as either 100 % first-year (blue) or multi-year (red) sea ice column covered with snow. See the main text for details.

4.4 Comparison of the two sea ice model versions in TB space

In this Section we present a demonstration of sea ice model evaluation in TB space utilising the observation operator. The demonstration is a comparison of two versions of the sea ice model named DMI-HYCOM-CICE (introduced in Section 2.3).

From the model output, we simulated AMSR2 TBs as described in Section 3.2.3 and compared these to the observations. Lower TB errors (e_{T_b}) indicate better agreement between simulated radiances and observations, and thus improved model performance. To further diagnose the results, we run the model matching in time and space with the OIB observations. In this way, we check if the simulated TB errors correspond to errors in the geo-physical parameters predicted by the sea ice models. To match the sea ice model, OIB, and satellite data, we took sea ice model predictions for 12 UTC and chose the swath between

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

11 and 13 UTC that best covered the OIB flight track. We first zoom in on the results for one day and then consider all data matches for 2013.

4.4.1 Results for 21 March 2013

Figure 9 shows TB simulation errors for two AMSR2 channels along the swath matching the first OIB flight track (21 March 2013), together with the resulting wRMSE$_{T_b}$ and SIC errors. SIC errors are computed as model minus observed, where the observed sea ice concentration was derived with the simple implementation of the NASA Team algorithm (see Section 3.2.2). The maps also show the locations of the OIB observations and the flight direction (black dots and arrow). Simulation errors vary between channels as they contain different information, but a common pattern is observed: warm biases north of Greenland transitioning to cooler biases in the Beaufort Sea. wRMSE$_{T_b}$ aggregates errors across all AMSR2 channels (12 in this case) into one number, making it easier to analyse. The wRMSE$_{T_b}$ is large (>16 K) in both models at the edge of the sea ice east of Greenland coinciding with the derived SIC errors. In the interior of the sea ice area, the wRMSE$_{T_b}$ varies up to around 16 K. In the baseline model, we find a region in the Beaufort Sea at the edge of the swath with notably high wRMSE$_{T_b}$ values which is not seen in the updated model. We see that the wRMSE$_{T_b}$ metric informs about the model performance at the ice edge *and* interior; however, SIC errors are mainly found at the sea ice edge.

Continuing along the flight track in Figure 9, we plot the wRMSE$_{T_b}$ and sea ice model errors of first-year ice fraction, ice thickness, snow depth, and surface temperature in Figure 10. In particular, the updated model underestimated the ice thickness in the first part of the track, while the baseline model overestimated the ice thickness in the latter part. This is reflected in the wRMSE$_{T_b}$ statistic, which tends to be higher for the updated model in the first part and lower in the last compared to the baseline model. The magnitude of the first-year ice fraction error is also reflected in the wRMSE$_{T_b}$ signal while the snow height error seems to have less of an effect. The surface temperatures in the two models are very similar, since the same atmospheric forcing is applied over most of the domain. We generalise these results in the following section.

4.4.2 Results for all of 2013

Now we present results from all nine flight track match-ups in 2013 to show the correspondence between wRMSE$_{T_b}$ and the model errors of geophysical variables. A time series of this is found in Appendix F, Figure F1. Here, we show that wRMSE$_{T_b}$ is related to the accuracy of the sea ice model.

Figure 11 compiles the match-ups into scatter plots of ΔwRMSE$_{T_b}$, i.e., difference of wRMSE$_{T_b}$ between the two model versions, and $\Delta|e_x|$, the absolute prediction error difference, where x is some geophysical variable:

$$\Delta\mathrm{wRMSE}_{T_b} = \mathrm{wRMSE}_{T_b}(\text{updated model}) - \mathrm{wRMSE}_{T_b}(\text{baseline model}), \quad (12)$$

$$\Delta|e_x| = |e_x(\text{updated model})| - |e_x(\text{baseline model})| \quad (13)$$

That is, if the updated model is better than the baseline model, then $\Delta|e_x| < 0$ and vice versa. We expect that this is reflected in ΔwRMSE$_{T_b}$ and, therefore, that there is a positive correlation between these two quantities. Figure 11 shows scatter plots of ΔwRMSE$_{T_b}$ versus $\Delta|e_x|$, where x is either first-year ice fraction, sea ice thickness, snow depth, or surface temperature.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

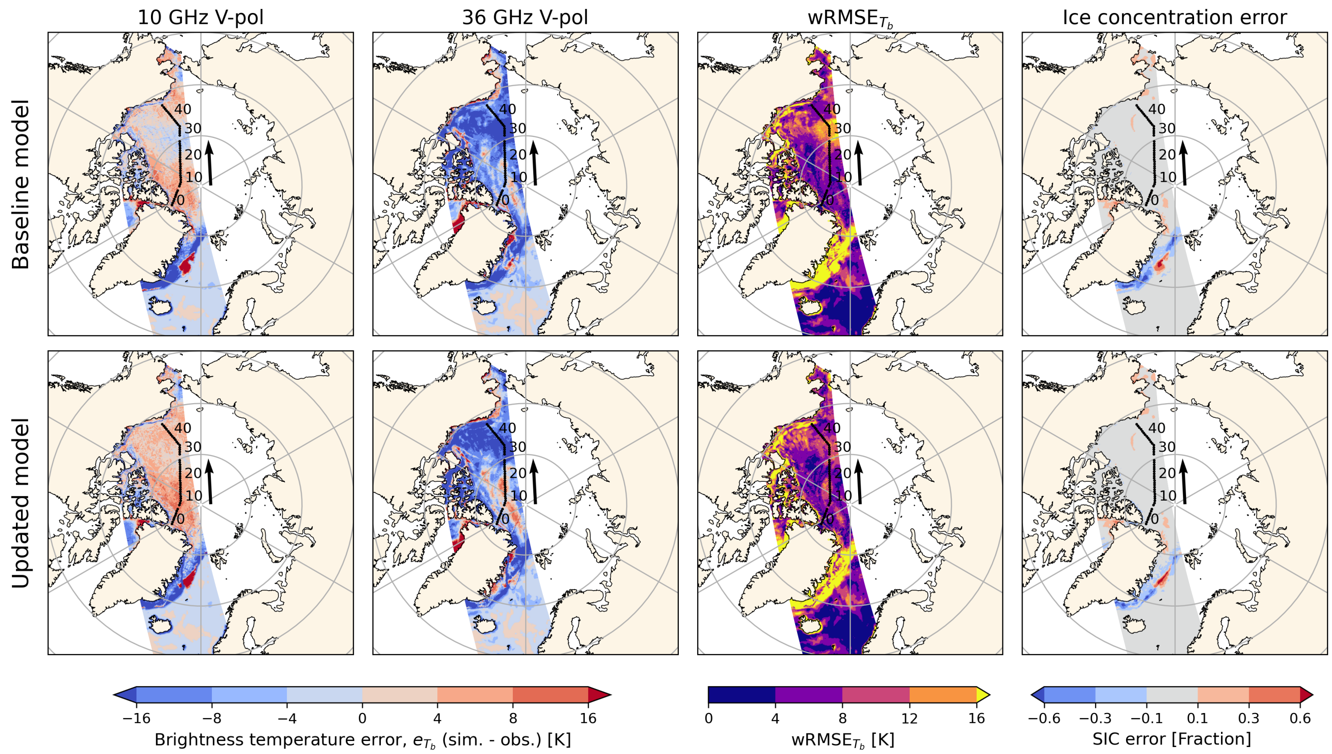

Figure 9. Example swath from 21 March 2013 showing TB errors simulated using DMI-HYCOM-CICE output as input to the observation operator. We show simulation errors for some two selected channels and the resulting wRMSE$_{T_b}$. Also, the SIC error is shown for comparison. The resampled OIB data locations are plotted with black dots and indices. These indices link to the x-axis in Figure 10. The arrows show the flight direction.

Correlation coefficients and their associated p-values are calculated using the *pearsonr* function from SciPy (Virtanen et al., 2020).

We find that ΔwRMSE$_{T_b}$ is significantly correlated with the improvements/degradations in first-year ice fraction with $R^2 = 0.31$ and also with the improvements/degradations ice thickness with $R^2 = 0.14$. In contrast, the difference in snow depth errors and surface temperature errors between the two versions of the sea ice model were too small to significantly impact the ΔwRMSE$_{T_b}$ signal.

The correlations found in Figure 11 bring evidence to our suggested methodology, by showing that the divergence of simulated and observed TBs is related to the sea ice model errors. In the synthetic experiment (Section 4.3), we found that SIC had the largest impact on our evaluation metric, but since the OIB data only cover high SIC regions, we could not confirm this relation in our real experiment here. Both the synthetic experiment and this comparison of two sea ice model versions show that the first-year ice fraction has a strong impact on the TB divergence. Accordingly, inaccurate modelling of this parameter

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

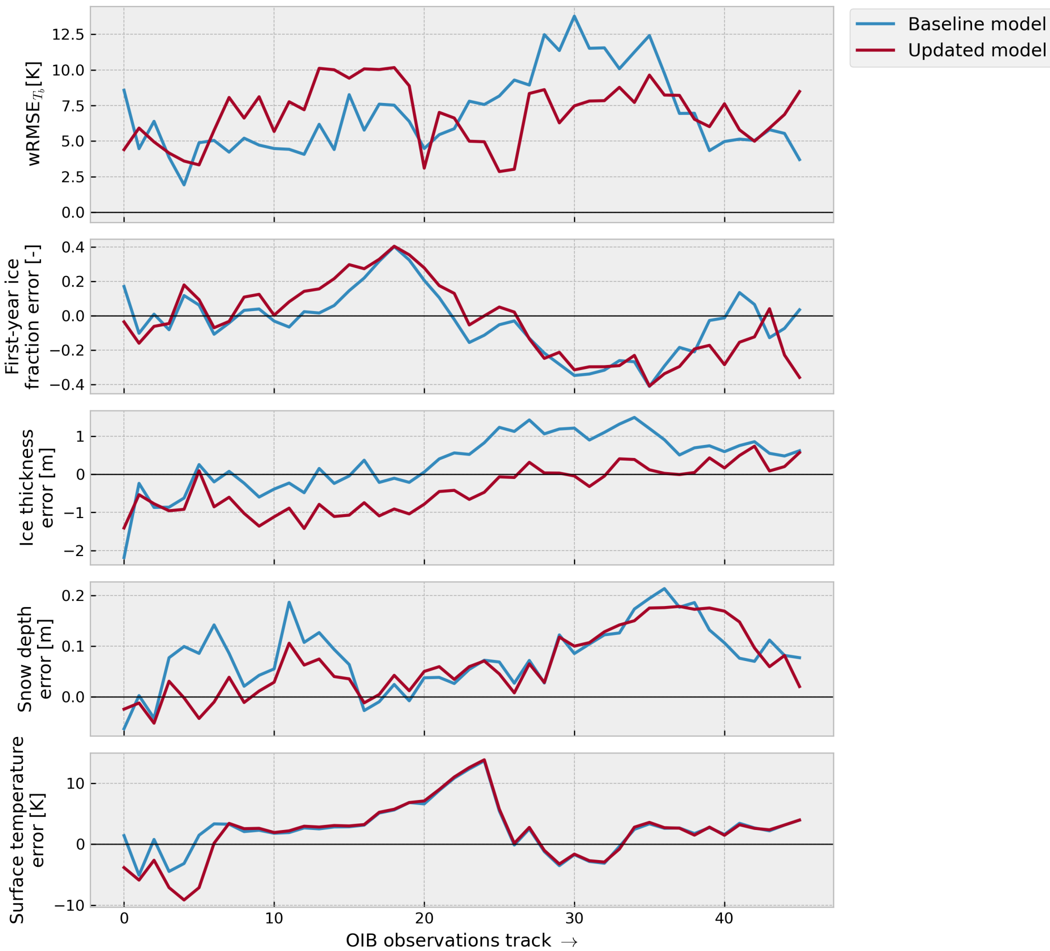

Figure 10. wRMSE$_{T_b}$ and sea ice model errors along the OIB flight track from 21 March 2013. The errors are defined as model minus observation. The flight track is drawn in Figure 9.

may hinder a reliable evaluation of other parameters. In this case, DMI-HYCOM-CICE models the sea ice type accurately enough, so that there is a clear relation between the wRMSE$_{T_b}$ and sea ice thickness.

4.5 Final remarks

We emphasise that this framework is not designed to retrieve direct information about individual sea ice parameters. Instead, the TB errors and the wRMSE$_{T_b}$ provide a measure of the overall accuracy of the sea ice model, as they depend on multiple model parameters simultaneously. Consequently, the framework is better suited for assessing the overall performance of a sea ice model rather than its ability to represent specific variables such as sea ice thickness. A key strength of the method is its

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

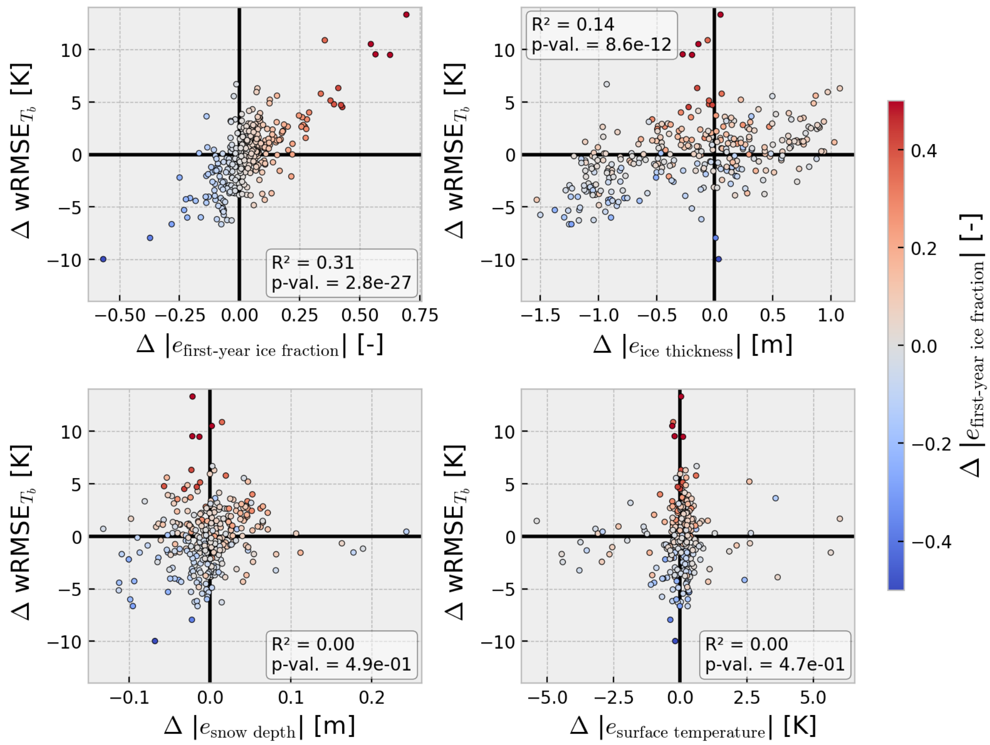

Figure 11. Scatter plot of ΔwRMSE$_{T_b}$ and $\Delta|e_x|$ for all OIB-CICE match-ups in 2013. Dots are coloured by $\Delta|e_{\text{first-year ice fraction}}|$ to show how this quantity dominates the ΔwRMSE$_{T_b}$ signal here.

extensive spatial coverage, as well as its provision of a largely independent metric for model evaluation in both the marginal ice zone and the ice interior.

It should be noted that the observation operator must be adapted to the specific sea ice model used. If the model does not provide the parameters required to simulate the target TBs, these must be inferred from climatology or additional model simulations. This is challenging and may introduce additional uncertainty. Therefore, sea ice models may benefit from explicitly representing the parameters required by the observation operator. For AMSR2 channels, these parameters include, in particular, the salinity, temperature, microstructure and density of both sea ice and snow. Including these parameters should also improve the thermodynamic representation of sea ice in simulations (Bitz and Lipscomb, 1999).

5 Conclusions

In this study, we presented a framework for evaluating sea ice models in TB space using a physics-based observation operator. By simulating satellite-observed TBs from model output and comparing them directly with observations, the approach complements traditional evaluation methods that rely on geophysical variables such as sea ice concentration or thickness. The

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

methodology enables model assessment using largely independent observations with total polar coverage, thereby addressing limitations related to sparse in situ data and the non-independence of assimilated satellite products.

The observation operator was constructed for Arctic sea ice under freezing conditions and simulated all AMSR2 channels except the 7.3 GHz channels. It used MEMLS coupled to a RTM for the atmosphere and ocean. The observation operator was tuned using detailed snow pit measurements made on first-year ice and airborne observations of sea ice and snow parameters from across the Arctic Ocean. Despite the necessary simplifications in the representation of snow and ice properties, the operator was able to reproduce AMSR2 TBs with reasonable accuracy across multiple frequencies and polarisations (Figure 6 and Table 5).

We introduced a simple multi-channel evaluation metric, the weighted TB RMSE (wRMSE$_{T_b}$), which aggregates information across frequencies into a single, interpretable quantity. Synthetic sensitivity experiments showed how the metric responds to errors in SIC, first-year ice fraction, sea ice thickness, snow depth, and surface temperature (Figure 8). We found that the metric is very sensitive to SIC errors and moderately sensitive to the other error types. Application to two versions of a pan-Arctic sea ice model demonstrated that model improvements are reflected in reduced TB errors and thus the evaluation metric, wRMSE$_{T_b}$. In particular, we found significant correlations between the wRMSE$_{T_b}$ and model improvements in ice type and ice thickness with $R^2 = 0.31$ and $R^2 = 0.14$ respectively. The change in snow depth and surface temperature between the model versions were not large enough to be correlated to the wRMSE$_{T_b}$.

The results highlight that TB-space evaluation provides a comprehensive measure of model performance, reflecting the combined influence of multiple physical properties of the snow–ice system. As such, the approach is not intended to diagnose individual parameters in isolation, but rather to assess the overall realism of the simulated surface state. The method is particularly valuable in the interior of the ice pack, where conventional metrics provide limited constraints.

Several limitations remain. The observation operator relies on simplified representations of snow and ice structure, as well as parametrisations or assumptions that introduce uncertainty. Furthermore, some required geophysical properties must be inferred rather than directly obtained from the sea ice model, complicating the attribution of errors. Especially, the microstructure of air-inclusions in the freeboard of multi-year ice is uncertain but has a large impact on the surface emission. Future work should aim to improve the representation of sub-footprint variability, refine parametrisations of sea ice and snow permittivity and scattering, and extend the framework to melting conditions.

Overall, the proposed TB-space evaluation framework represents a promising and scalable tool for the assessment of sea ice models. With the continued availability of radiometers from current and upcoming missions, it offers a pathway toward more comprehensive and independent evaluation of sea ice models across spatial and temporal scales.

Data availability. The AKROSS 2022 Cambridge Bay data is available at https://doi.org/10.5281/zenodo.11205157. The RRDP data is available at https://doi.org/10.6084/m9.figshare.6626549.v7. The OIB KT19 data is available at https://nsidc.org/data/iakst1b/versions/2. The DMI-HYCOM-CICE data is available at https://doi.org/10.11583/DTU.32685171. The AMSR2 L1R data is available at https://gportal.jaxa.jp. The ERA5 data can be downloaded at https://cds.climate.copernicus.eu/datasets/reanalysis-era5-single-levels. The data derived from SIMBA buoys is available at https://doi.pangaea.de/10.1594/PANGAEA.973193.

Author contributions. AJ wrote the paper and did all the data analysis. AJ, RT, HS, and SS developed the idea for the paper. MHR and IS contributed expertise on DMI-HYCOM-CICE and provided perspectives on sea ice model evaluation. MHR ran the sea ice model experiments. MH helped with using and understanding MEMLS. MS was part of the AKROSS team. RT, HS, SS, MHR, IS, MH, MS, and VN gave feedback on the writing.

Competing interests. Some authors are members of the editorial board of The Cryosphere.

Disclaimer. (publisher's text)

Acknowledgements. We would like to thank the Japan Aerospace Exploration Agency for the AMSR2 brightness temperature observations, the European Centre for Medium-Range Weather Forecasts for the NWP analysis (ERA5), NASA's National Snow and Ice Data Center for the Operation IceBridge observations of ice and snow, the Arctic PASSION partners for the buoy data, and the AKROSS 2022 Cambridge Bay team for the snow pit and sea ice measurements. We would also like to thank Andreas Weismann and Christian Mätzler for the development of MEMLS, Thomas Meissner and Frank J. Wentz for the development of the ocean and atmosphere RTM, and the consortium behind the CICE model development.

Financial support. AJ, HS, SS, MR, and IS were founded by the Danish State through the National Centre for Climate Research.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

References

Alexandrov, V., Sandven, S., Wahlin, J., and Johannessen, O. M.: The relation between sea ice thickness and freeboard in the Arctic, The Cryosphere, 4, 373–380, https://doi.org/10.5194/tc-4-373-2010, 2010.

Allard, R. A., Farrell, S. L., Hebert, D. A., Johnston, W. F., Li, L., Kurtz, N. T., Phelps, M. W., Posey, P. G., Tilling, R., Ridout, A., and Wallcraft, A. J.: Utilizing CryoSat-2 sea ice thickness to initialize a coupled ice-ocean modeling system, Advances in Space Research, 62, 1265–1280, https://doi.org/10.1016/j.asr.2017.12.030, 2018.

Andersen, S., Tonboe, R., Kern, S., and Schyberg, H.: Improved retrieval of sea ice total concentration from spaceborne passive microwave observations using numerical weather prediction model fields: An intercomparison of nine algorithms, Remote Sensing of Environment, 104, 374–392, https://doi.org/10.1016/j.rse.2006.05.013, 2006.

Assur, A.: Composition of Sea Ice and its Tensile Strength, Research Report 44, U. S. Army Cold Regions Research and Engineering Laboratory, Wilmette, Illinois, 1960.

Bezanson, J., Edelman, A., Karpinski, S., and Shah, V. B.: Julia: A Fresh Approach to Numerical Computing, SIAM Review, 59, 65–98, https://doi.org/10.1137/141000671, 2017.

Bitz, C. M. and Lipscomb, W. H.: An energy-conserving thermodynamic model of sea ice, Journal of Geophysical Research: Oceans, 104, 15 669–15 677, https://doi.org/10.1029/1999JC900100, 1999.

Burgard, C., Notz, D., Pedersen, L. T., and Tonboe, R. T.: The Arctic Ocean Observation Operator for 6.9 GHz (ARC3O) – Part 1: How to obtain sea ice brightness temperatures at 6.9 GHz from climate model output, The Cryosphere, 14, 2369–2386, https://doi.org/10.5194/tc-14-2369-2020, 2020.

Cavalieri, D. J., Gloersen, P., and Campbell, W. J.: Determination of sea ice parameters with the NIMBUS 7 SMMR, Journal of Geophysical Research: Atmospheres, 89, 5355–5369, https://doi.org/10.1029/JD089iD04p05355, 1984.

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., Baraille, R., and Bleck, R.: The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system, Journal of Marine Systems, 65, 60–83, https://doi.org/10.1016/j.jmarsys.2005.09.016, 2007.

Cox, G. F. N. and Weeks, W. F.: Brine Drainage and Initial Salt Entrapment in Sodium Chloride Ice., Research Report 345, COLD REGIONS RESEARCH AND ENGINEERING LABORATORY (CRREL), Hanover, New Hampshire, 1975.

Cox, G. F. N. and Weeks, W. F.: Equations for Determining the Gas and Brine Volumes in Sea-Ice Samples, Journal of Glaciology, 29, 306–316, https://doi.org/10.3189/S0022143000008364, 1983.

Drinkwater, M. R. and Crocker, G. B.: Modelling Changes in Scattering Properties of the Dielectric and Young Snow-Covered Sea Ice at GHz Requencies, Journal of Glaciology, 34, 274–282, https://doi.org/10.3189/S0022143000007012, 1988.

Elizabeth Hunke, Richard Allard, David A. Bailey, Philippe Blain, David Clemens-Sewall, Anthony Craig, Frederic Dupont, Alice DuVivier, Jim Edwards, Robert Grumbine, David Hebert, Marika Holland, Nicole Jeffery, Jean-Francois Lemieux, Robert Osinski, Jacob Poulsen, Till Rasmussen, Mads Ribergaard, Lettie Roach, Andrew Roberts, Anton Steketee, Matthew Turner, Michael Winton, Denise Worthen, and Bin Zhao: CICE-Consortium/CICE: CICE Version 6.6.2, https://doi.org/10.5281/ZENODO.1205674, 2025.

Fritzner, S., Graversen, R., Christensen, K. H., Rostosky, P., and Wang, K.: Impact of assimilating sea ice concentration, sea ice thickness and snow depth in a coupled ocean–sea ice modelling system, The Cryosphere, 13, 491–509, https://doi.org/10.5194/tc-13-491-2019, 2019.

Geldsetzer, T., Langlois, A., and Yackel, J.: Dielectric properties of brine-wetted snow on first-year sea ice, Cold Regions Science and Technology, 58, 47–56, https://doi.org/10.1016/j.coldregions.2009.03.009, 2009.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Herrmannsdörfer, L., Müller, M., Shupe, M. D., and Rostosky, P.: Surface temperature comparison of the Arctic winter MOSAiC observations, ERA5 reanalysis, and MODIS satellite retrieval, Elementa: Science of the Anthropocene, 11, 00 085, https://doi.org/10.1525/elementa.2022.00085, 2023.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., De Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., De Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Ivanova, N., Pedersen, L. T., Tonboe, R. T., Kern, S., Heygster, G., Lavergne, T., Sørensen, A., Saldo, R., Dybkjær, G., Brucker, L., and Shokr, M.: Inter-comparison and evaluation of sea ice algorithms: towards further identification of challenges and optimal approach using passive microwave observations, The Cryosphere, 9, 1797–1817, https://doi.org/10.5194/tc-9-1797-2015, 2015.

JAXA: GCOM-W/AMSR2 L1R Brightness Temperature, https://doi.org/10.57746/EO.01GS73ANT4D0Y20R133TZ5BZ7P, 2012.

JAXA: Intercomparison results between AMSR2 and TMI/AMSR-E/GMI, https://www.eorc.jaxa.jp/AMSR/datacatalog/tb/pdf/150326_AMSR2_XcalResults.pdf, 2015.

Jones, S. B. and Friedman, S. P.: Particle shape effects on the effective permittivity of anisotropic or isotropic media consisting of aligned or randomly oriented ellipsoidal particles, Water Resources Research, 36, 2821–2833, https://doi.org/10.1029/2000WR900198, 2000.

Kang, E., Sohn, B., Tonboe, R. T., Noh, Y., Kwon, I., Kim, S., Maturilli, M., Kim, H., and Liu, C.: Explicitly determined sea ice emissivity and emission temperature over the Arctic for surface-sensitive microwave channels, Quarterly Journal of the Royal Meteorological Society, 149, 2011–2030, https://doi.org/10.1002/qj.4492, 2023.

King, J., Howell, S., Brady, M., Toose, P., Derksen, C., Haas, C., and Beckers, J.: Local-scale variability of snow density on Arctic sea ice, The Cryosphere, 14, 4323–4339, https://doi.org/10.5194/tc-14-4323-2020, 2020.

Kovacs, A.: Sea Ice Part 1. Bulk salinity versus ice floe thickness, CRREL Report 96-7, 1996.

Kurtz, N. T., Farrell, S. L., Studinger, M., Galin, N., Harbeck, J. P., Lindsay, R., Onana, V. D., Panzer, B., and Sonntag, J. G.: Sea ice thickness, freeboard, and snow depth products from Operation IceBridge airborne data, The Cryosphere, 7, 1035–1056, https://doi.org/10.5194/tc-7-1035-2013, 2013.

Lavergne, T., Sørensen, A. M., Kern, S., Tonboe, R., Notz, D., Aaboe, S., Bell, L., Dybkjær, G., Eastwood, S., Gabarro, C., Heygster, G., Killie, M. A., Brandt Kreiner, M., Lavelle, J., Saldo, R., Sandven, S., and Pedersen, L. T.: Version 2 of the EUMETSAT OSI SAF and ESA CCI sea-ice concentration climate data records, The Cryosphere, 13, 49–78, https://doi.org/10.5194/tc-13-49-2019, 2019.

Lee, S. and Sohn, B.: Retrieving the refractive index, emissivity, and surface temperature of polar sea ice from 6.9 GHz microwave measurements: A theoretical development, Journal of Geophysical Research: Atmospheres, 120, 2293–2305, https://doi.org/10.1002/2014JD022481, 2015.

Meloche, J., Sandells, M., Rutter, N., Löwe, H., Scharien, R. K., Essery, R., and Jaggi, M.: AKROSS 2022 Cambridge Bay Dataset, https://doi.org/10.5281/ZENODO.11205157, 2024.

Mätzler, C.: Improved Born approximation for scattering of radiation in a granular medium, Journal of Applied Physics, 83, 6111–6117, https://doi.org/10.1063/1.367496, 1998.

Mätzler, C. and Wiesmann, A.: Extension of the Microwave Emission Model of Layered Snowpacks to Coarse-Grained Snow, Remote Sensing of Environment, 70, 317–325, https://doi.org/10.1016/S0034-4257(99)00047-4, 1999.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Nandan, V., Geldsetzer, T., Yackel, J., Mahmud, M., Scharien, R., Howell, S., King, J., Ricker, R., and Else, B.: Effect of Snow Salinity on CryoSat-2 Arctic First-Year Sea Ice Freeboard Measurements, Geophysical Research Letters, 44, https://doi.org/10.1002/2017GL074506, 2017.

Pedersen, L. T., Saldo, R., Ivanova, N., Kern, S., Heygster, G., Tonboe, R., Huntemann, M., Ozsoy, B., Ardhuin, F., and Kaleschke, L.: Reference dataset for sea ice concentration, https://doi.org/10.6084/M9.FIGSHARE.6626549.V7, 2021.

Picard, G., Sandells, M., and Löwe, H.: SMRT: an active–passive microwave radiative transfer model for snow with multiple microstructure and scattering formulations (v1.0), Geoscientific Model Development, 11, 2763–2788, https://doi.org/10.5194/gmd-11-2763-2018, 2018.

Ponsoni, L., Ribergaard, M. H., Nielsen-Englyst, P., Wulf, T., Buus-Hinkler, J., Kreiner, M. B., and Rasmussen, T. A. S.: Greenlandic sea ice products with a focus on an updated operational forecast system, Frontiers in Marine Science, 10, https://doi.org/10.3389/fmars.2023.979782, 2023.

Pounder, E. R.: The physics of ice, Pergamon Press, 1965.

Preußer, A., Nicolaus, M., and Hoppmann, M.: Snow depth, sea ice thickness and interface temperatures derived from measurements of SIMBA buoys deployed in the Arctic Ocean and Southern Ocean between 2012 and 2023, https://doi.org/10.1594/PANGAEA.973193, 2025.

Pringle, D. J., Trodahl, H. J., and Haskell, T. G.: Direct measurement of sea ice thermal conductivity: No surface reduction, Journal of Geophysical Research: Oceans, 111, 2005JC002 990, https://doi.org/10.1029/2005JC002990, 2006.

Ricker, R., Hendricks, S., Kaleschke, L., Tian-Kunze, X., King, J., and Haas, C.: A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data, The Cryosphere, 11, 1607–1623, https://doi.org/10.5194/tc-11-1607-2017, 2017.

Rückert, J. E., Huntemann, M., Tonboe, R. T., and Spreen, G.: Modeling Snow and Ice Microwave Emissions in the Arctic for a Multi-Parameter Retrieval of Surface and Atmospheric Variables From Microwave Radiometer Satellite Data, Earth and Space Science, 10, e2023EA003 177, https://doi.org/10.1029/2023EA003177, 2023.

Schyberg, H., Yang, X., Køltzow, M., and Amstrup, B.: Arctic regional reanalysis on single levels from 1991 to present, https://doi.org/10.24381/CDS.713858F6, 2020.

Shy, T. L., Walsh, J. E., Chapman, W. L., Lynch, A. H., and Bailey, D. A.: Sea-ice model validation using submarine measurements of ice draft, Annals of Glaciology, 31, 307–312, https://doi.org/10.3189/172756400781820110, 2000.

Sievers, I., Rasmussen, T. A. S., and Stenseng, L.: Assimilating CryoSat-2 freeboard to improve Arctic sea ice thickness estimates, The Cryosphere, 17, 3721–3738, https://doi.org/10.5194/tc-17-3721-2023, 2023.

Smith, G. C., Roy, F., Reszka, M., Surcel Colan, D., He, Z., Deacu, D., Belanger, J., Skachko, S., Liu, Y., Dupont, F., Lemieux, J., Beaudoin, C., Tranchant, B., Drévillon, M., Garric, G., Testut, C., Lellouche, J., Pellerin, P., Ritchie, H., Lu, Y., Davidson, F., Buehner, M., Caya, A., and Lajoie, M.: Sea ice forecast verification in the Canadian Global Ice Ocean Prediction System, Quarterly Journal of the Royal Meteorological Society, 142, 659–671, https://doi.org/10.1002/qj.2555, 2016.

Studinger, M.: IceBridge KT19 IR Surface Temperature, Version 2, https://doi.org/10.5067/UHE07J35I3NB, 2020.

Sturm, M., Holmgren, J., König, M., and Morris, K.: The thermal conductivity of seasonal snow, Journal of Glaciology, 43, 26–41, https://doi.org/10.3189/S0022143000002781, 1997.

Tian, T., Yang, S., Høyer, J. L., Nielsen-Englyst, P., and Singha, S.: Cooler Arctic surface temperatures simulated by climate models are closer to satellite-based data than the ERA5 reanalysis, Communications Earth & Environment, 5, 111, https://doi.org/10.1038/s43247-024-01276-z, 2024.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Timco, G. and Weeks, W.: A review of the engineering properties of sea ice, Cold Regions Science and Technology, 60, 107–129, https://doi.org/10.1016/j.coldregions.2009.10.003, 2010.

Tonboe, R. T.: Sea ice emission modelling, in: Thermal microwave radiation - Applications for remote sensing, Institution of Engineering and Technology, 2006.

Tonboe, R. T.: The simulated sea ice thermal microwave emission at window and sounding frequencies, Tellus A, 62, 333–344, https://doi.org/10.1111/j.1600-0870.2010.00434.x, 2010.

Tonboe, R. T.: Modeling Microwave Emission and Scattering from Snow-Covered Sea Ice, in: Sea ice: Physics and Remote Sensing, Wiley, 2. edn., https://doi.org/10.1002/9781119828242.ch12, 2023.

Tonboe, R. T., Nandan, V., Mäkynen, M., Pedersen, L. T., Kern, S., Lavergne, T., Øelund, J., Dybkjær, G., Saldo, R., and Huntemann, M.: Simulated Geophysical Noise in Sea Ice Concentration Estimates of Open Water and Snow-Covered Sea Ice, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 15, 1309–1326, https://doi.org/10.1109/JSTARS.2021.3134021, 2022.

Tonboe, R. T., Nandan, V., Huntemann, M., Stroeve, J., Scharien, R., Yackel, J., Kaleschke, L., Shi, H., and Casal, T.: Estimation of sea ice air-bubble and brine pocket distribution for scattering and emission model parametrization, https://doi.org/10.5194/egusphere-2026-1440, 2026.

Vancoppenolle, M., Fichefet, T., and Bitz, C. M.: Modeling the salinity profile of undeformed Arctic sea ice, Geophysical Research Letters, 33, 2006GL028 342, https://doi.org/10.1029/2006GL028342, 2006.

Vant, M. R., Ramseier, R. O., and Makios, V.: The complex-dielectric constant of sea ice at frequencies in the range 0.1–40 GHz, Journal of Applied Physics, 49, 1264–1280, https://doi.org/10.1063/1.325018, 1978.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., Van Der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, , Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., Van Mulbregt, P., SciPy 1.0 Contributors, Vijaykumar, A., Bardelli, A. P., Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C. N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D. A., Hagen, D. R., Pasechnik, D. V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G. A., Ingold, G.-L., Allen, G. E., Lee, G. R., Audren, H., Probst, I., Dietrich, J. P., Silterra, J., Webber, J. T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J. L., De Miranda Cardoso, J. V., Reimer, J., Harrington, J., Rodríguez, J. L. C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N. J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P. A., Lee, P., McGibbon, R. T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T. J., Robitaille, T. P., Spura, T., Jones, T. R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y. O., and Vázquez-Baeza, Y.: SciPy 1.0: fundamental algorithms for scientific computing in Python, Nature Methods, 17, 261–272, https://doi.org/10.1038/s41592-019-0686-2, 2020.

Wang, C., Graham, R. M., Wang, K., Gerland, S., and Granskog, M. A.: Comparison of ERA5 and ERA-Interim near-surface air temperature, snowfall and precipitation over Arctic sea ice: effects on sea ice thermodynamics and evolution, The Cryosphere, 13, 1661–1679, https://doi.org/10.5194/tc-13-1661-2019, 2019.

Wentz, F. J. and Meissner, T.: AMSR Ocean Algorithm, Algorithm Theoretical Basis Document, 2000.

Wiesmann, A. and Mätzler, C.: Microwave Emission Model of Layered Snowpacks, Remote Sensing of Environment, 70, 307–316, https://doi.org/10.1016/S0034-4257(99)00046-2, 1999.

Williams, N., Wang, Y., and Counillon, F.: Enhancing sea ice knowledge through assimilation of sea ice thickness from ENVISAT and CS2SMOS, The Cryosphere, 20, 853–873, https://doi.org/10.5194/tc-20-853-2026, 2026.

Zhang, Q., Luo, H., Min, C., Xiu, Y., Shi, Q., and Yang, Q.: Evaluation of Arctic Sea Ice Thickness from a Parameter-Optimized Arctic Sea Ice–Ocean Model, Remote Sensing, 15, 2537, https://doi.org/10.3390/rs15102537, 2023.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Appendix A: Parametrisation of the salinity profile of first-year sea ice

Here we describe how the salinity profile of first-year ice was inferred when simulating TBs from the AKROSS 2022 Cambridge Bay data and the OIB data. We assumed the profile follows this parametrisation:

Two depth points are defined:

$$p_1 = c_{p_1}, \qquad p_2 = h_i - c_{p_2}, \tag{A1}$$

where h_i is the total ice thickness and c_{p_1} and c_{p_2} are model parameters to be fitted. Then the salinity at depth z is calculated:

$$S(z) = \begin{cases} \left[f_0(z/s) - f_0(0)\right] s + f_0(0), & \text{if } p_1 \geq p_2 \\ a_0 + a_1 z + \mathbf{1}_{z \leq p_1} a_2 (z - p_1)^2 + \mathbf{1}_{z > p_2} a_2 (z - p_2)^2, & \text{otherwise} \end{cases} \tag{A2}$$

where a_0, a_1 and a_2 are fitting coefficients, $s = \frac{h_i}{c_{p_1} + c_{p_2}}$, and

$$f_0(z) = a_0 + a_1 z + a_2 (z - p_1)^2, \tag{A3}$$

The first branch in Equation A2 is used when the ice is sufficiently thin that the quadratic shoulder regions overlap ($p_1 \geq p_2$). Finally, non-negativity is enforced: $S(z) = \max(S(z), 0)$.

We digitized plots of modelled salinity profiles in Vancoppenolle et al. (2006) and fitted $S(z)$ to the January, February and March profiles to get:

$$a_0 = 7.13\,\mathrm{PSU}, \quad a_1 = -1.73\,\mathrm{PSU\,m^{-1}}, \quad a_2 = 91.85\,\mathrm{PSU\,m^{-2}}, \quad c_{p_1} = 0.22\,\mathrm{m}, \quad c_{p_2} = 0.13\,\mathrm{m} \tag{A4}$$

Profiles of different thickness are plotted along with the digitized Vancoppenolle et al. (2006) profiles for January, February, March and April in Figure A1. The parametrisation also reproduces the April profile well, despite the April profile not being included in the fitting procedure.

Appendix B: Comparison of ERA5 and OIB KT19 surface temperatures

When simulating TBs using OIB data (Section 3.2.2), we used ERA5 skin temperature data instead of OIB KT19 measurements, since ERA5 data better match the spatial scale of AMSR2 footprints. However, ERA5 skin/surface temperatures are known to be warm-biased (Wang et al., 2019), which we also found when comparing KT19 and ERA5 surface temperatures. The warm bias is suspected to be mainly explained by snow depth (Herrmannsdörfer et al., 2023), and therefore we made a linear correction from snow depth by comparing ERA5 and KT19 surface temperatures. A linear regression of the ERA5–KT19 temperature difference against snow depth is shown in Figure B1. The correction using snow depth (h_s) in meters is:

$$T_{sfc} = T_{sfc,\mathrm{ERA5}} - (25.1 \cdot h_s - 0.99) \tag{B1}$$

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

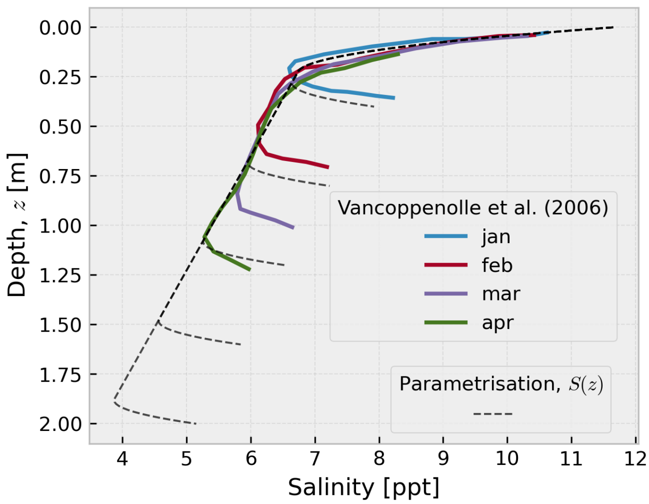

Figure A1. First-year ice salinity computed using the parametrisation along with modelled profiles from Vancoppenolle et al. (2006).

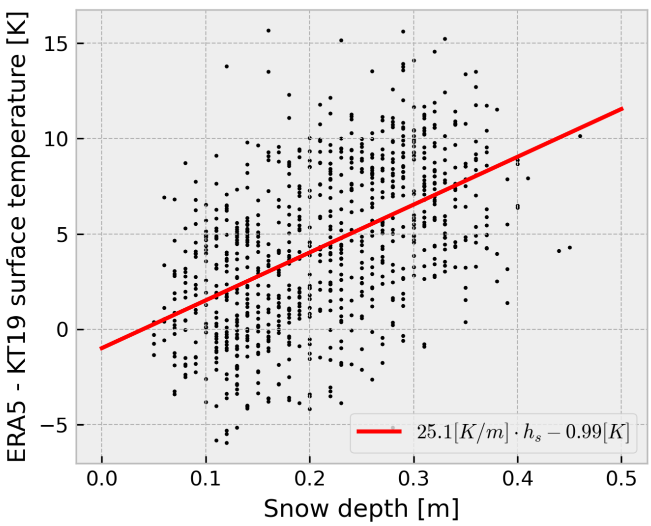

Figure B1. Difference between ERA5 and KT19 surface temperatures as a function of snow depth. The solid line shows the fitted linear regression.

Appendix C: Ice-snow interface temperature fit to SIMBA observations

In Section 3.2.2, we presented a method (Equation 8) for estimating the ice-snow interface temperature (T_{si}), which involved 695 *apparent* snow and ice thermal conductivities (k_s and k_i). To obtain an empirical estimate of k_s, we fitted Equation 8 to estimates of snow depth, ice thickness, ice-snow interface temperature and snow-air interface temperature derived from SIMBA buoys (Preußer et al., 2025). We assumed $k_i = 2.1\,\mathrm{W\,K^{-1}\,m^{-1}}$ and fitted k_s using observations from April and March by minimising the mean absolute error between modelled and observed ice–snow interface temperatures. Only Arctic buoys were considered. The resulting fit giving $k_s = 0.388\,\mathrm{W\,K^{-1}\,m^{-1}}$ is seen in Figure C1.

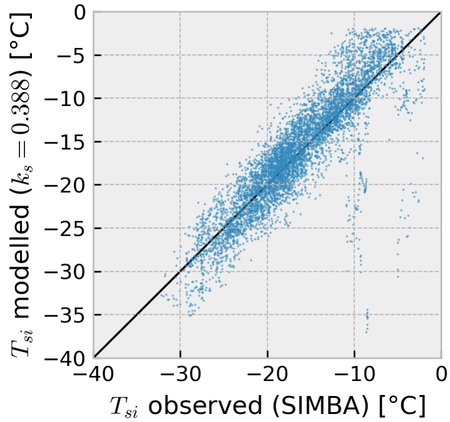

Figure C1. Modelled versus observed ice-snow interface temperature (T_{si}) after fitting the apparent snow thermal conductivity (k_s).

Appendix D: Calculation of first-year and multi-year ice area fraction from CICE output

As explained in Section 2.3, DMI-HYCOM-CICE provides first-year ice fraction *per grid cell* (F_{FY}) and not *per sea ice thickness category* ($f_{\mathrm{FY},k}$), which we want. Ice type fraction is here the fraction of the total ice area - *not* the grid cell area. We infer $f_{FY,k}$ with the following method, which labels the thinnest ice categories in a cell as first-year and the thickest as multi-year ice while preserving F_{FY}. First, the multi-year ice fraction per category ($f_{MY,k}$) is computed:

$$f_{\mathrm{MY},k} = \begin{cases} 0 & x_k \leq 0 \\ 1 & x_k \geq 1 \\ x_k & \text{otherwise} \end{cases}, \qquad \text{where } x_k = \left(\sum_{i=1}^{k} f_i - F_{\mathrm{FY}} \sum_{i=1}^{6} f_i\right) f_i^{-1}. \tag{D1}$$

Here, f_i is the sea ice area fraction of a specific ice category, where $i = 1$ is the thinnest ice and $i = 6$ is the thickest. $\sum_{i=1}^{6} f_i$ gives the total sea ice concentration. The first-year ice fraction per category is then simply $f_{\mathrm{FY},k} = 1 - f_{\mathrm{MY},k}$.

Appendix E: State dependence of observation operator errors: Snow depth and sea ice thickness

Here, we show the dependence of the observation-operator on snow depth and sea-ice thickness. These are supplementary figures to the results in Section 4.2. Ideally, the TB simulation errors (e_{T_b}) are independent of the absolute value of the snow

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

depth and ice thickness. Figures E1 and E2 show that this is approximately true for most channels, but the 36 GHz and 89 GHz channels in particular do show some non-linear dependencies.

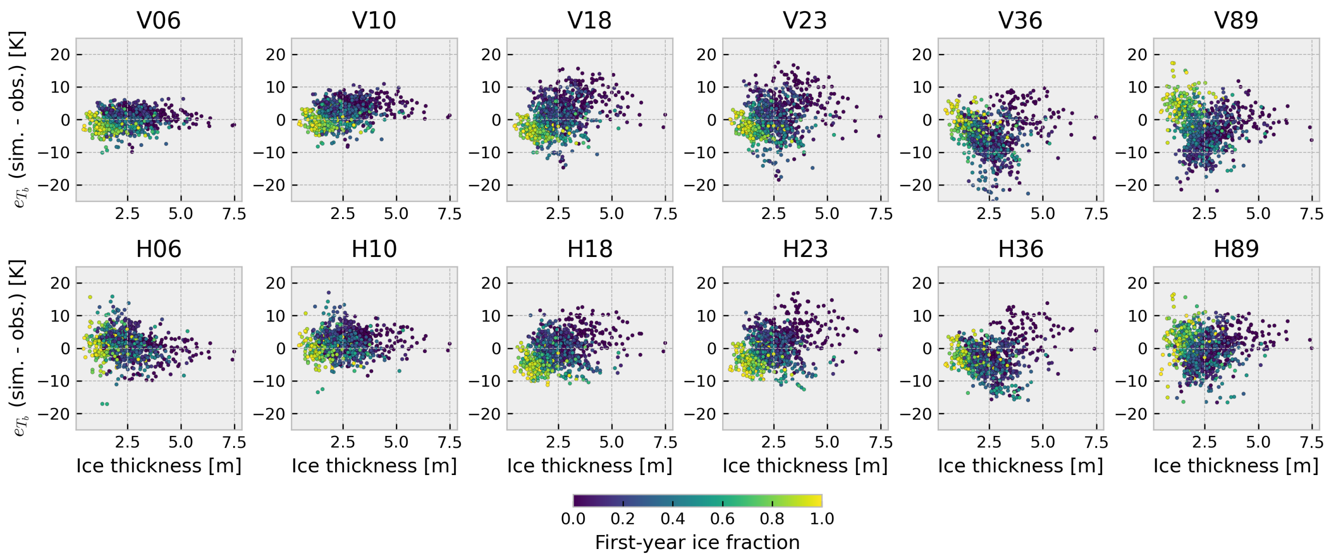

Figure E1. TB simulation error e_{T_b} versus sea ice thickness - complementary figures to the results in Section 4.2.

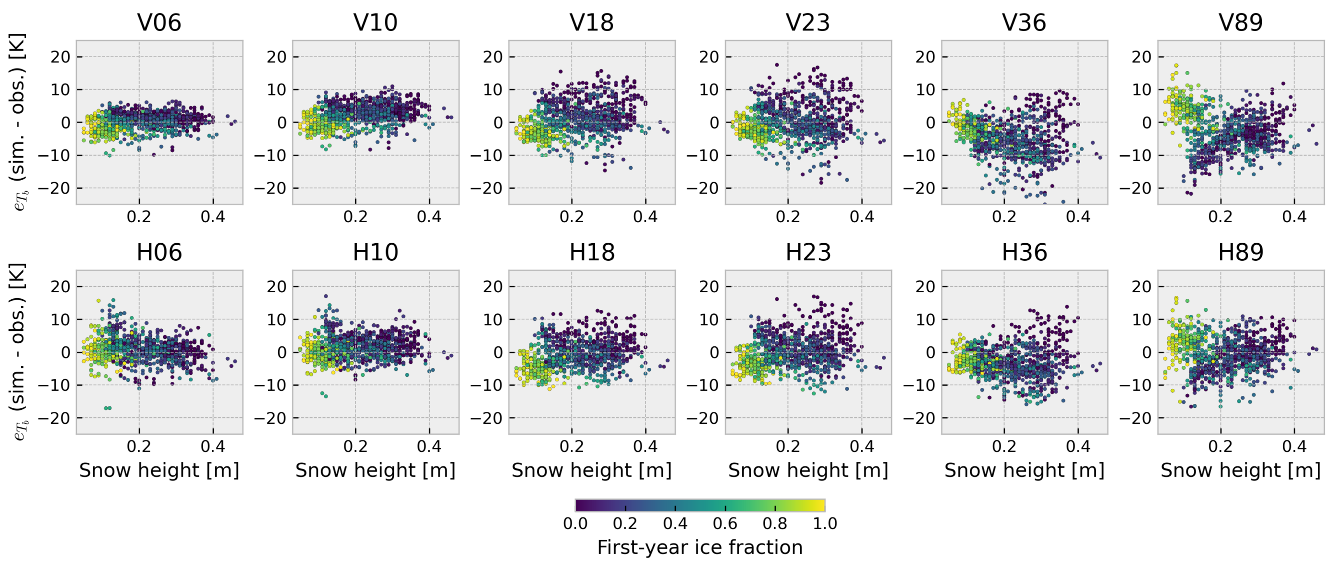

Figure E2. TB simulation error e_{T_b} versus snow depth - complementary figures to the results in Section 4.2.

https://doi.org/10.5194/egusphere-2026-3613
Preprint. Discussion started: 11 August 2026

Appendix F: Sea ice model errors along all OIB flight tracks

Figure F1 is similar to Figure 10, but includes data from all the relevant OIB flights in 2013 shown on a timeline. Over-/under-
715 estimations of first-year ice fractions are clearly reflected in the wRMSE$_{T_b}$ metric, e.g., the spikes in wRMSE$_{T_b}$ on the 23 and 24 March 2013 coincide with poor estimates of first-year ice fraction. There is also some correspondence between prediction errors of ice thickness and the evaluation metric, e.g., the pattern of ice thickness errors during 2013-03-22 is clearly reflected in the wRMSE$_{T_b}$.

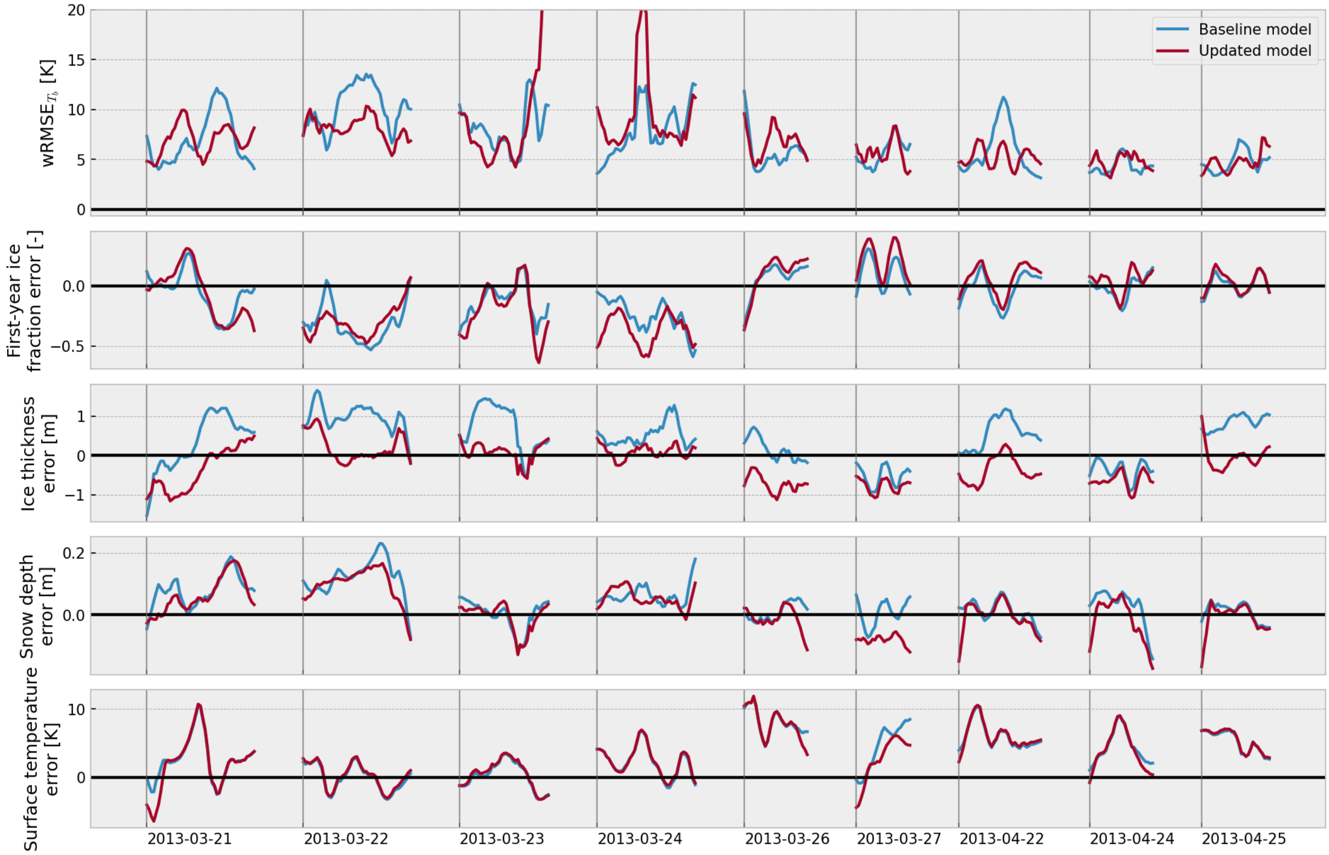

Figure F1. Concatenated OIB track results from all of 2013. Similar to Figure 10 but including all OIB tracks from 2013. The tracks are concatenated with some spacing along the x-axis, with dates shown. The data was smoothed along the track with a box-car window of size 5.