

Physically-Informed Super-Resolution Downscaling of Antarctic Surface Melt

de Roda Husman, Sophie; Hu, Zhongyang; van Tiggelen, Maurice; Dell, Rebecca; Bolibar, Jordi; Lhermitte, Stef; Wouters, Bert; Munneke, Peter Kuipers

DOI

[10.1029/2023MS004212](https://doi.org/10.1029/2023MS004212)

Publication date

2024

Document Version

Final published version

Published in

Journal of Advances in Modeling Earth Systems

Citation (APA)

de Roda Husman, S., Hu, Z., van Tiggelen, M., Dell, R., Bolibar, J., Lhermitte, S., Wouters, B., & Munneke, P. K. (2024). Physically-Informed Super-Resolution Downscaling of Antarctic Surface Melt. *Journal of Advances in Modeling Earth Systems*, 16(7), Article e2023MS004212. <https://doi.org/10.1029/2023MS004212>

Important note

To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.



RESEARCH ARTICLE

10.1029/2023MS004212

Physically-Informed Super-Resolution Downscaling of Antarctic Surface Melt

Special Collection:

Machine learning application to Earth system modeling

These authors Sophie de Roda Husman and Zhongyang Hu contributed equally to this work.

Sophie de Roda Husman¹ , Zhongyang Hu² , Maurice van Tiggelen² , Rebecca Dell³ , Jordi Bolibar^{1,4} , Stef Lhermitte^{1,5} , Bert Wouters¹, and Peter Kuipers Munneke² 

¹Department of Geoscience & Remote Sensing, Delft University of Technology, Delft, The Netherlands, ²Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, The Netherlands, ³Scott Polar Research Institute (SPRI), University of Cambridge, Cambridge, UK, ⁴Institut des Géosciences de l'Environnement, Université Grenoble Alpes, CNRS, IRD, G-INP, Grenoble, France, ⁵Department of Earth & Environmental Sciences, Catholic University of Leuven, Leuven, Belgium

Key Points:

- Our method downscales Antarctic surface melt from a regional climate model, employing a physically-informed super-resolution architecture
- The super-resolution model relies on physical information derived from remote sensing data, specifically surface albedo and elevation
- Incorporating physical information boosts model generalization, enabling accurate high-resolution surface melt beyond the training region

Correspondence to:

S. de Roda Husman,
S.deRodaHusman@tudelft.nl

Citation:

de Roda Husman, S., Hu, Z., van Tiggelen, M., Dell, R., Bolibar, J., Lhermitte, S., et al. (2024). Physically-informed super-resolution downscaling of Antarctic surface melt. *Journal of Advances in Modeling Earth Systems*, 16, e2023MS004212. <https://doi.org/10.1029/2023MS004212>

Received 10 JAN 2024
Accepted 25 JUN 2024

Author Contributions:

Conceptualization: Zhongyang Hu, Peter Kuipers Munneke
Methodology: Sophie de Roda Husman, Zhongyang Hu, Jordi Bolibar, Stef Lhermitte, Bert Wouters, Peter Kuipers Munneke
Resources: Maurice van Tiggelen, Rebecca Dell

Abstract Because Antarctic surface melt is mostly driven by local processes, its simulation necessitates high-resolution regional climate models (RCMs). However, the current horizontal resolution of RCMs (≈ 25 –30 km) is inadequate for capturing small-scale melt processes. To address this limitation, we present SUPREME (SUPer-RESolution-based Melt Estimation over Antarctica), a deep learning method to downscale surface melt to 5.5 km resolution using a physically-informed super-resolution model. The physical information integrated into the model originates from observations tied to surface melt, specifically remote sensing-derived albedo and elevation. These remote sensing data, in addition to a Regional Atmospheric Climate Model (RACMO) run at 27 km resolution, account for the diverse drivers of surface melt across Antarctica, facilitating effective generalization beyond the training region of the Antarctic Peninsula. A comparison of SUPREME with a dynamically downscaled RACMO run at 5.5 km over the Antarctic Peninsula shows high accuracy, with average yearly RMSE and bias of 5.5 mm w.e. yr^{-1} and 4.5 mm w.e. yr^{-1} , respectively. Validation at five automatic weather stations reveals SUPREME's marked improvement with substantially lower average RMSE (81 mm w.e.) compared to RACMO 27 km (129 mm w.e.). Beyond the training region, SUPREME aligns more closely with remote sensing products associated with surface melt than super-resolution models lacking physical constraints. While further validation of SUPREME is needed, our study highlights the potential of super-resolution techniques with physical constraints for high-resolution surface melt monitoring in Antarctica, providing insights into the impacts of localized melting on processes affecting ice shelf integrity such as hydrofracturing.

Plain Language Summary To improve surface melt monitoring in Antarctica, high-resolution climate models are essential. Existing models, like the Regional Atmospheric Climate Model (RACMO), do not have a fine enough spatial resolution to capture small-scale melt processes. To overcome this, we introduce SUPREME (SUPer-RESolution-based Melt Estimation over Antarctica), a method that refines surface melt data to a higher resolution of 5.5 km using an advanced super-resolution model. We enhance this model with physical information derived from observations directly related to surface melt, specifically using albedo and elevation data from remote sensing. By incorporating these observations, along with RACMO data at 27 km resolution, we account for the various triggers of surface melt across Antarctica. SUPREME accurately predicts high-resolution surface melt beyond the training region of the Antarctic Peninsula, showing promising results compared to existing melt observations. Further validation is needed, but this approach, combining super-resolution techniques and remote sensing data, holds potential for accurate surface melt monitoring in Antarctica. This may advance our understanding of the impacts of localized features on processes affecting ice shelf integrity such as meltwater-induced hydrofracturing.

© 2024 The Author(s). Journal of Advances in Modeling Earth Systems published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Reductions in both thickness (Gudmundsson et al., 2019) and extent (Greene et al., 2022) of floating ice shelves have contributed to an increasing mass loss of the Antarctic Ice Sheet (IMBIE, 2018; Shepherd et al., 2012). A recent study by Davison et al. (2023) revealed that over 40% of Antarctic ice shelves have lost mass between 1997 and 2021. While around two-thirds of this volume loss was attributed to basal melting (Davison et al., 2023), surface melt is anticipated to play a more significant role in ice shelf shrinkage and weakening in the coming

Supervision: Stef Lhermitte, Bert Wouters, Peter Kuipers Munneke
Validation: Sophie de Roda Husman, Zhongyang Hu
Visualization: Sophie de Roda Husman
Writing – original draft: Sophie de Roda Husman
Writing – review & editing: Zhongyang Hu, Maurice van Tiggelen, Rebecca Dell, Jordi Bolibar, Stef Lhermitte, Bert Wouters, Peter Kuipers Munneke

decades under continued atmospheric warming in all future climate scenarios (Gilbert & Kittel, 2021; Trusel et al., 2015). The presence of surface meltwater reduces albedo (Lenaerts et al., 2017) and firm air content (Kuipers Munneke et al., 2014), creating positive feedback loops that generate additional surface melt and consequently increase meltwater ponding. This ponding poses a threat to ice shelf stability through surface meltwater-induced flexure and hydrofracturing (Banwell et al., 2019; Banwell & Macayeal, 2015; Scambos et al., 2009).

Nonetheless, accurately quantifying meltwater volumes on the Antarctic Ice Sheet is posing a significant challenge. Direct assessment of melt from ice core stratigraphy (e.g., Abram et al., 2013; Das & Alley, 2008) and in-situ surface energy balance (SEB) observations (e.g., Jakobs et al. 2020; Kuipers Munneke, Van den Broeke et al., 2012) provide robust quantitative melt records, but these records are sparse. Remote sensing observations, while capable of offering measurements continent-wide, usually provide only binary data indicating whether melting has occurred or not, without quantifying the meltwater volume (de Roda Husman et al., 2022, 2024). Trusel et al. (2013) made a notable exception by employing the Quick Scatterometer (QuikSCAT) to measure annual meltwater volumes. Unfortunately, QuikSCAT ceased operation in 2009. Recently, also Banwell et al. (2023) made an effort to derive meltwater volumes from passive and active microwave sensors, but only by applying a physics-based snow model to the binary melt observations derived from remote sensing data, meltwater volumes could be estimated. Hence, besides in-situ observations and a few remote sensing methods, our understanding of past, present, and future surface melt volumes relies heavily on regional climate models (RCMs) such as the Regional Atmospheric Climate Model (RACMO) (van Wessem et al., 2018) and Modèle Atmosphérique Régional (MAR) (Agosta et al., 2019). Nevertheless, with a spatial resolution of typically 25–30 km, RCMs face limitations in capturing the intricate interplay between melting and albedo, and have difficulty representing small-scale, high-melt features like blue ice, rocks, and ponds (Arthur et al., 2022; Kingslake et al., 2017; Lenaerts et al., 2017).

Hence, there is a necessity for downscaling techniques that reproduce the critical details of surface melt that are currently lacking in RCMs. A frequently employed technique is statistical downscaling, a method that enhances the resolution of RCM variables, typically by leveraging their linear correlation with elevation. Recently, statistical downscaling has been employed to generate high-resolution surface mass balance (SMB) variables for Antarctica (Gallée et al., 2011), including snowfall (Ghilain et al., 2022) and surface melt (Noël et al., 2023a). Similarly, over Greenland, SMB components have been subject to statistical downscaling in studies by Hanna et al. (2005), Hanna et al. (2008, 2011), Franco et al. (2012), Noël et al. (2016), and Tedesco et al. (2023). However, in areas where the correlation of an SMB component with elevation is weak, statistical downscaling may not provide additional benefits (van de Berg et al., 2020). Another commonly used method, dynamical downscaling, entails running a model in a nested domain with higher spatial resolution and represents a prevalent design approach for many RCMs (Box & Rinke, 2003; Fettweis et al., 2017; Noël et al., 2016). Dynamical downscaling requires that physical parameterizations are adjusted in a finer grid (Hourdin et al., 2017; Schmidt et al., 2017). It relies on physical formulations rather than empirical relationships and correlations, but it comes with a significant computational cost (Fyke et al., 2018). Hence, both statistical and dynamical downscaling methods have their limitations, prompting us to explore an alternative approach for downscaling surface melt from an RCM: deep learning.

In this study, we introduce SUPREME (SUPER-RESolution-based Melt Estimation over Antarctica), a method to refine RACMO from a 27 km horizontal resolution to a more detailed 5.5 km resolution, employing a physically-informed deep learning super-resolution technique. Super-resolution techniques have been successfully used to refine image details (Dong et al., 2016; Goodfellow et al., 2014; Shi et al., 2016). Their potential for geophysical application has been demonstrated in Antarctic studies, including the downscaling of global climate models (van der Meer et al., 2023), bedrock topography (Leong & Horgan, 2020), and sea ice motion (Petrou et al., 2018). Super-resolution methods distinguish themselves by their adeptness at preserving intricate details without being constrained by potentially limiting linear statistical assumptions or high computational costs for predictions. Despite such advantages, super-resolution models may not generalize well beyond their training region (Jiang et al., 2022). This could be particularly problematic for its application to surface melt, given the notable variations in melt triggers and patterns observed across Antarctica (de Roda Husman et al., 2022; Hu et al., 2022). To enhance spatial transferability, alongside RACMO 27 km, we include observations that are physically associated with surface melt—specifically, albedo and elevation from remote sensing data—into our super-resolution model. Albedo is linked to surface reflectivity, while elevation accounts for altitude-related temperature variations—both critical factors in determining the extent and intensity of surface melt (Giesen & Oerlemans, 2012). The

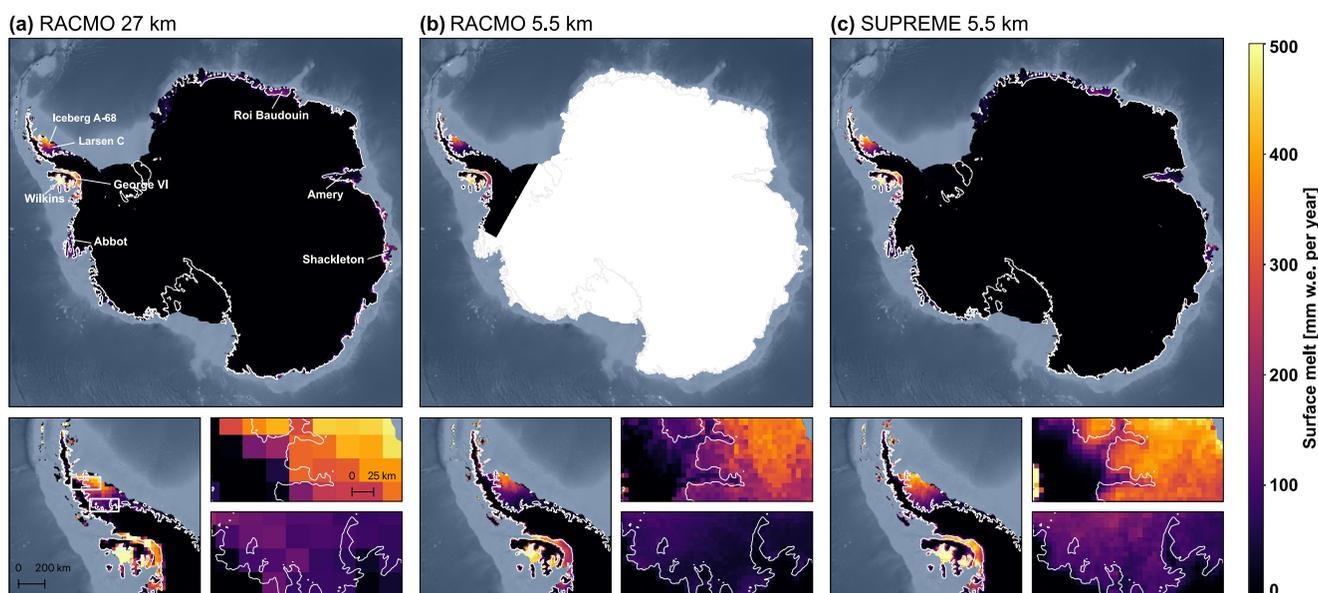


Figure 1. Surface melt volume comparison for the 2017–2018 melt season: (a) RACMO 27 km, (b) RACMO 5.5 km (van Wessem et al., 2016), and (c) SUPREME (this study). Upper panels display Antarctic-wide scale, while lower panels provide close-ups of the Antarctic Peninsula. Since RACMO 5.5 km data are solely available for the Antarctic Peninsula, the white region indicates no data.

incorporation of remote sensing data into the super-resolution model involves transitioning from a single-image super-resolution model, which depends solely on the low-resolution counterpart of the desired high-resolution output (Yang et al., 2014), to a multi-image super-resolution model that integrates multiple input features (Kawulok et al., 2019).

2. Data

2.1. RACMO 27 km and 5.5 km

We employ surface melt volume data from two RACMO data sets, both belonging to version 2.3p2, with horizontal resolutions of approximately 27 km (van Wessem et al., 2018) and 5.5 km (van Wessem et al., 2016). The RACMO 27 km data set serves as one of the input features for the super-resolution model, while the RACMO 5.5 km data set, covering the Antarctic Peninsula exclusively, serves as the reference data (or “ground truth”) for training the super-resolution model.

RACMO 27 km combines the atmospheric dynamics from the High-Resolution Limited Area Model (HIRLAM) (Undén et al., 2002) with the physical parameterizations of the European Center for Medium-Range Weather Forecasts (ECMWF) global model (ECMWF, 2009). It is coupled bidirectionally with a multi-layer snow model that accounts for processes like melting, percolation, and refreezing in the snow (Ettema et al., 2010). Additionally, the model incorporates a snow albedo scheme based on the evolution of snow grain size (Kuipers Munneke et al., 2011) and a drifting snow scheme that simulates the redistribution and sublimation of suspended snow particles (Lenaerts et al., 2012). For more detailed technical information about RACMO 27 km, we direct the reader to van Wessem, Reijmer, Lenaerts, et al. (2014), van Wessem, Reijmer, Morlighem, et al. (2014). In 2018, RACMO 5.5 km was introduced as a designated run at 5.5 km over the Antarctic Peninsula. For this high-resolution run, RACMO was forced by ERA-Interim, using an updated surface topography for the Antarctic Peninsula (van Wessem et al., 2016).

Both RACMO 27 and 5.5 km data are available for the period spanning from 1979 to the present day. In our study, we focus on daily surface melt outputs between 2001 and 2019, measured in millimeters water equivalent (mm w.e.). While RACMO 27 km provides coverage across the entire Antarctic region (see Figure 1a), RACMO 5.5 km is limited to the Antarctic Peninsula (Figure 1b). In addition to differences in spatial coverage and horizontal resolutions, RACMO 5.5 km simulates lower surface melt values compared to RACMO 27 km. For the Antarctic Peninsula over the period from 1979 to 2014, the annual surface melt volume showed a notable

disparity, with RACMO 5.5 km recording 34 Gt yr^{-1} compared to RACMO 27 km's 57 Gt yr^{-1} (Kuipers Munneke, Picard, et al., 2012; van Wessem et al., 2016). Also throughout our study period from 2001 to 2019, RACMO 5.5 km reports melt values that are lower compared to RACMO 27 km (see Figure A1), which is likely attributed to a cloud cover underestimation (van Wessem et al., 2016). As our SUPREME model is calibrated on the RACMO 5.5 km reference data set, this underestimation compared to RACMO 27 km is anticipated to persist. In follow-up studies, there is potential for retraining our developed super-resolution architecture using high-resolution surface melt reference data other than RACMO 5.5 km, such as from upcoming versions of RACMO or other RCMs like MAR.

2.2. Physical Constraints

2.2.1. Albedo

In addition to RACMO 27 km, albedo and elevation are additional variables input into the super-resolution model. We use daily white-sky albedo, obtained from the MCD43A3 product on a spatial resolution of 500 m, based on data from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Schaaf et al., 2002). This MCD43A3 product matches those in previous Antarctic studies (e.g., Lenaerts et al. 2017; Tollenaar et al., 2022). The data set is stored in Google Earth Engine (GEE; Gorelick et al., 2017) and consists of daily observations, which are generated as 16-day running means. To align with the resolution of the RACMO 5.5 km product, we employ bicubic interpolation to reproject the data to the Antarctic Polar Stereographic projection, and average the data to upsample it to a 5.5 km resolution. If albedo values are unavailable due to cloud cover, we substitute the pixel with the monthly median albedo value of that specific pixel.

However, since cloud cover can reach up to 90% on the Antarctic ice shelves (Lachlan-Cope, 2010), the use of the daily albedo input feature is limited. Therefore, the albedo Q5 input feature acts as an additional data source, highlighting the most extreme melt patterns on an annual basis. We compute the 5th percentile (Q5) albedo for each pixel during every melt season, defined as spanning from July 1 to June 30 of the following year.

2.2.2. Elevation

The final input feature for the super-resolution model is a static elevation map sourced from the TanDEM-X PolarDEM. This map provides gap-free elevation data for 2013–2014 over Antarctica at a spatial resolution of 90 m (Wessel et al., 2021). Similar to the MODIS albedo data, we employ bicubic interpolation to reproject the data set onto the Antarctic Polar Stereographic projection, and apply averaging to upsample the data to a resolution of 5.5 km. We neglect elevation changes and assume constant topography since the elevation changes over the studied time period are relatively small and are not expected to significantly impact atmospheric dynamics.

2.3. Validation Data

2.3.1. Automatic Weather Stations

SUPREME is assessed against five automatic weather stations (AWS): AWS14, AWS15, AWS17, AWS18, and Neumayer (refer to Figure A2 for their locations). The five selected AWS have multiyear data records without any gaps and have experienced surface melt during our study period. AWS14 and AWS15 are located at the center of the Larsen C Ice Shelf, AWS17 on the remnant of the Larsen B Ice Shelf, AWS18 at the grounding line of the Larsen C Ice Shelf, and Neumayer at the Ekström Ice Shelf. Meteorological observations from the five AWS are forced into a SEB model, enabling the calculation of surface melt. The SEB model and configuration are the same as used by Jakobs et al. (2020). In addition, the measured longwave radiation is corrected for window heating as described in Smeets et al. (2018), and measured shortwave radiation is corrected for the zero offset as described in Foken (2021).

2.3.2. RACMO 2 km

The SUPREME melt product is compared with another downscaled RACMO product besides RACMO 5.5 km, referred to as RACMO 2 km, as developed by Noël et al. (2023a). RACMO 2 km was obtained through a statistical-downscaling approach applied to RACMO 27 km, resulting in a spatial resolution of 2 km. RACMO

2 km data are available on a daily resolution, but in this study, we only use annual melt values (from July 1 to June 30 for the years spanning from 2001 to 2019).

2.3.3. QuikSCAT

We compare SUPREME's annual melt volumes with QuikSCAT data for eight melt seasons, spanning from 2001–2002 to 2008–2009. QuikSCAT observations for Antarctica are available from 1999 to 2009 at a 4.45 km resolution (Trusel et al., 2013). The product relies on the annual sum of backscatter intensity reduction, calibrated with AWS observations.

2.3.4. Passive Microwave Remote Sensing

We also compare SUPREME to data from PMW sensors, also known as radiometers. PMWs provide binary melt data (i.e., melt or no-melt) that allow for the calculation of cumulative melt days within a melt season. We use cumulative melt days using a data set from Picard and Fily (2006), which contains daily observations that distinguish between melt and no-melt on a 25 by 25 km² spatial resolution. This binary melt product was generated by applying a melt detection algorithm (Picard & Fily, 2006; Torinesi et al., 2003) to data from the scanning Multichannel Microwave Radiometer (SMMR) and three Special Sensor Microwave Imager (SSM/I) instruments (Picard & Fily, 2006).

2.3.5. Optical Remote Sensing

Finally, we compare SUPREME with a melt-slush product derived from optical remote sensing (Dell et al., 2022, 2024a). This product is particularly effective in detecting melt presence in regions with low firn air content, where the meltwater is clearly visible in optical imagery, in contrast to microwave sensors. Dell et al. (2022, 2024a) developed a product for surface melt and slush based on the visible and near-infrared (NIR) bands of Landsat 8 observations covering the Antarctic ice shelves. The slush and melt data set (Dell et al., 2024a) offers monthly data spanning from 2013 to 2021, resulting in six overlapping melt seasons with SUPREME (from 2013–2014 to 2018–2019). To streamline the comparison, we aggregate the monthly data to create a single optical remote sensing product for each melt season, indicating the number of months with either melt or slush per pixel. The melt and slush data set, with a 30 m spatial resolution, is too detailed for a straightforward visual comparison with SUPREME's 5.5 km resolution. To address this, we apply maximum interpolation to downscale the optical remote sensing product to a 5.5 km resolution. This process involves assigning the value representing the highest number of months within each 5.5 km pixel area.

3. Methods

3.1. Preprocessing

The super-resolution model is trained on the Antarctic Peninsula, which is divided into 13 training patches as shown in Supplementary Figure A3. These patches, created for memory efficiency, are sized at 297 km by 297 km, which is the lowest common multiple of 5.5 km (for RACMO 5.5 km, albedo, and elevation) and 27 km (for RACMO 27 km): that is, 5.5 km multiplied by 54, and 27 km by 11 both equal 297 km. We gather input features including daily RACMO 27 km, daily albedo, annual Q5 albedo, and a static elevation map for each of the training patches, along with the daily RACMO 5.5 km reference data. To ensure that edge information receives equal consideration as central data, we apply padding. For RACMO 5.5 km, this entails adding 10 pixels (equivalent to a 55 km border), while for RACMO 27 km, two pixels are added (equivalent to a 54 km border). After the training phase, the patches are clipped to maintain their central dimensions of 297 km by 297 km. After the padding, the preprocessed data are split into three sets: training (2001–2006), validation (2007–2010), and testing (2011–2019). The validation set is used during training to tune hyperparameters, while the testing set is reserved for evaluating the final performance of the trained model on unseen data. The split was primarily driven by the availability of AWS records, predominantly spanning from 2011 onwards. The distribution of the training, validation, and testing data sets are statistically similar, as confirmed by an Anderson-Darling test. This statistical assessment is designed to determine similarity in distribution (Scholz & Stephens, 1987). The Anderson-Darling test revealed that the data sets surpass a p-value of 0.05, indicating consistent distributions among the data sets.

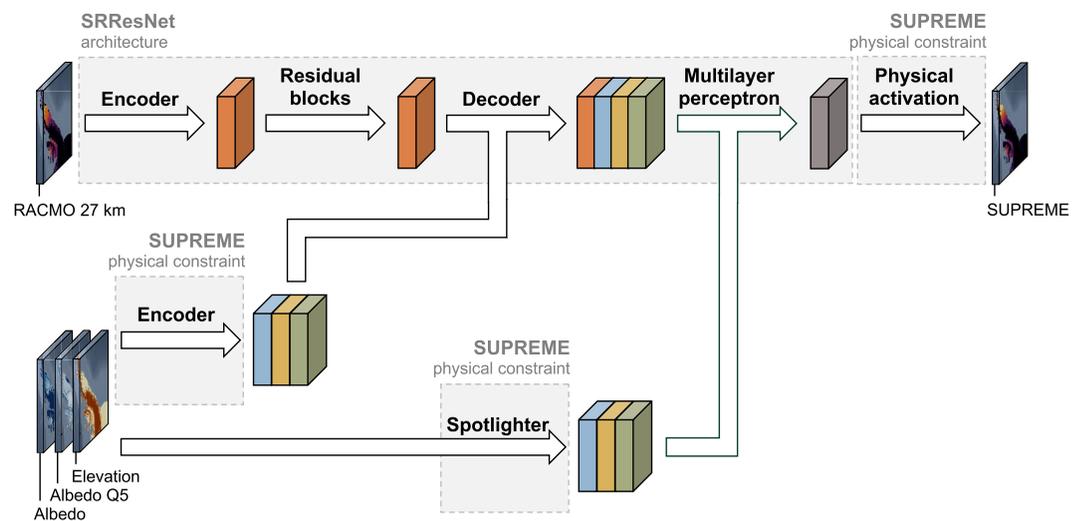


Figure 2. Overview of super-resolution model architecture. The model incorporates the SRResNet architecture along with three additional physical constraints. The processed data were color-coded: orange for daily RACMO 27 km, blue for daily albedo, yellow for annual 5th percentile albedo, and green for elevation.

3.2. SUPREME Model Architecture

3.2.1. SRResNet Architecture

Our customized neural network architecture (Figure 2) builds upon the SRResNet, a neural network architecture specifically designed for image super-resolution tasks (Ledig et al., 2017). In the first step of the SRResNet, the RACMO 27 km data undergoes encoding through a single convolution operation. Specifically, a 3×3 filter is applied to the RACMO 27 km data to compute weighted sums, resulting in dimensionality reduction. The outcome of this encoding step produces feature maps that emphasize the large-scale spatial melt patterns, effectively translating the data from spatial melt values into abstract features.

The next step of the SRResNet involves the implementation of 16 residual blocks applied to the encoded RACMO 27 km (He et al., 2016). These blocks are constructed to allow the combination of the output layer of one layer with that of another layer situated deeper within the block. This design facilitates learning by enabling the model to concentrate on recognizing specific residual details, which pertain to the differences between the current prediction and the desired output (i.e., RACMO 5.5 km). Within the 16 residual blocks, three operations are carried out. First, a convolutional layer with 3×3 kernels and 64 feature maps extracts spatial patterns from the input. Subsequently, batch normalization is applied to stabilize and accelerate training. Finally, an element-wise summation operation combined the output of a layer with the input of a deeper layer through a skip connection, allowing data to bypass multiple layers and address the vanishing gradients problem. This phenomenon occurs when gradients become extremely small during training, hindering the ability of deep neural networks to effectively learn from data (He et al., 2016).

After the encoding steps and residual block processing, we merge the RACMO 27 km data with the encoded albedo and elevation data (Section 3.2.2). Subsequently, the data undergoes decoding, a process that transforms compressed or abstract features back into a more understandable format, effectively upsampling the data. Once again, we follow a strategy similar to SRResNet, employing six deconvolution layers with 3×3 kernels, complemented by Parametric Rectified Linear Unit (PReLU) as the activation function. The decoded data are combined with the spotlighted data from SUPREME (Section 3.2.2). Following this, we implement the final step in the standard SRResNet architecture: the multilayer perceptron, which involves connecting every neuron in one layer to every neuron in the next layer through a single convolution operation with a 1×1 kernel. Although this step forms the final stage in the conventional SRResNet, as well as in many other neural networks, we introduce an additional step involving a physical activation function, as detailed in Section 3.2.2.

3.2.2. Physical Constraints

The SRResNet is enriched using three physical constraints, each performing distinct tasks: the “encoder” adds daily albedo, annual Q5 albedo, and elevation data for feature extraction and dimensionality reduction, the “spotlighter” uses 3D-convolutions to analyze albedo and elevation patterns in latitude, longitude, and time, and the “physical activation” scales melt values based on elevation, leveraging the knowledge that lower elevations are associated with higher melt values (see Figure 2).

For the first physical constraint, the “encoder,” we apply a similar procedure to the encoding of RACMO 27 km data, as outlined in the SRResNet architecture, applying a convolution operation and PReLU activation function to extract meaningful patterns from the albedo and elevation data. Subsequently, we downscale the data to match the spatial resolution of the encoded RACMO 27 km data. This is accomplished through 2×2 average pooling, wherein the input data are divided into 2×2 regions, and the average of the values within each region is calculated. These encoded albedo and elevation data are combined with the encoded RACMO 27 km to undergo a decoding step explained in the SRResNet architecture section.

The “spotlighter” constrains the super-resolution model by the same data used in the “encoder,” which undergoes a 3D convolution for analysis across longitude, latitude, and time. The 3D convolution functions as a specialized filter, emphasizing important aspects of the data in both spatial and temporal domains. The resulting output is subsequently merged with the decoded RACMO 27 km, albedo, and elevation data. This step aids in capturing meaningful spatiotemporal features which are passed to the multilayer perceptron.

Recognizing the negative correlation between elevation and surface melt (Trusel et al., 2012), and the positive feedback between albedo and surface melt (Jakobs et al., 2019), we apply two correction functions that adjust pixels based on elevation and albedo, referred to as “physical activation.” The physical activation function for elevation increases surface melt values for pixels with low elevation, and reduces surface melt values for pixels at higher elevations. In the re-scaling process, each high-resolution surface melt pixel generated as the final output of the SRResNet architecture is multiplied by a custom sigmoid-like function. These sigmoid functions are created by analyzing the albedo Q5 and elevation data for annual surface melt values from RACMO 5.5 km, RACMO 27 km, and QuikSCAT. Based on this comparison, the sigmoid functions are plotted using visual inspection and serve as initial guides (see Supplementary Figure A6). During training, the weights of the model are updated based on learned weighting functions. The custom sigmoid-like functions have the following shape:

$$Out(lon, lat) = In(lon, lat) \times \frac{A}{1 + B \times e^{X(lon, lat) - C}} \quad (1)$$

in which $In(lon, lat)$ is the final output of the SRResNet architecture, and the parameters A is set to 1.36, B is set to 0.005, and C is set to 1.0.

After adjusting the elevation values, we apply the albedo physical activation function, once more using a custom sigmoid-like function (Equation 1). In this context, the parameter $In(lon, lat)$ denotes the output of the elevation physical activation function, while $X(lon, lat)$ represents annual albedo Q5 data, varying both spatially and on a yearly basis. The values assigned to the albedo Q5 physical-activation functions are as follows: A is set to 1.0, B is set to 0.05, and C is set to 37.0.

3.3. Training

During training, the RMSprop optimizer (Reddy et al., 2018) is employed, starting with a learning rate of 10^{-3} . An exponential learning rate decay is implemented to dynamically adjust the learning rate in later training epochs. The models are trained with a batch size of 16 over a span of 30 epochs. Early stopping is implemented to mitigate potential model overfitting, whereby it monitors training mean squared error (MSE) loss. If the MSE loss did not improve by at least the specified minimum threshold of 5×10^{-4} (mm w.e. yr^{-1})² for five consecutive epochs, the training process concludes prematurely. We experimented with alternative hyperparameters, such as learning rate, loss function, and optimizer, using the validation data set, but found that the presented parameters produced the best results.

To evaluate the importance of physical constraints, we train different models: one incorporating all three physical constraints (referred to as SUPREME), another without any physical constraint (referred to as SRResNet), and

additional models with exclusive use of each individual physical constraint—SUPREME_{only encoder}, SUPREME_{only spotlighter}, and SUPREME_{only physical activation}. Additionally, we train two models to evaluate the importance of the remote sensing input features on the downscaled surface melt product, which we denote as SUPREME_{no albedo} and SUPREME_{no elevation}. In SUPREME_{no albedo}, we omit the albedo input features, retaining only RACMO 27 km and elevation as inputs. In SUPREME_{no elevation}, we exclude elevation, keeping only RACMO 27 km and albedo. In SRResNet, we exclude all remote sensing features, using only RACMO 27 km as an input feature.

3.4. Post-Processing

Using the trained models from Section 3.3, we produce the SUPREME results and the results for which one or more input features or physical constraints were excluded. The output from the super-resolution models includes patches with padding, which are subsequently clipped to form patches with dimensions of 54×54 pixels. Despite the applied padding on patch edges, some patches exhibit pronounced transitions, resulting in a “checkerboard” effect. To address this, we reposition all patches 27 km left and right relative to their original positions and recalculate the results. Subsequently, we average the outcomes from the three runs to obtain the final SUPREME results. Finally, we sum all the individual days to create aggregated products per melt season, measured in mm w. e. per year. Due to memory constraints, we evaluate only the melt products per melt season rather than at a higher temporal resolution.

3.5. Evaluation

Evaluating SUPREME poses a considerable challenge. There is no observational data set providing Antarctic-wide surface melt data at such a high spatial resolution for the past two decades. Consequently, we have to rely on various available products to assess the performance of SUPREME. We compare SUPREME with in-situ data (AWS), another downscaled RACMO product (RACMO 2 km), and various remote sensing products (QuikSCAT, PMW remote sensing, and a melt-slush product from optical remote sensing), as was described in Section 2.3.

Moreover, we also conduct a comparative analysis of SUPREME against RACMO 5.5 km, focusing exclusively on the Antarctic Peninsula, where RACMO 5.5 km data are available. We compare SUPREME and RACMO 5.5 km for the testing period from 2010 to 2019 and employ four key metrics: coefficient of determination (R^2), Root Mean Square Error (RMSE), bias, and Pearson Correlation Coefficient (PCC). Firstly, R^2 is a statistical measure indicating the proportion of the variance in the reference data that can be explained by the prediction data in a regression model. Secondly, the RMSE gives us an understanding of the average size of the differences between predicted and observed values, effectively quantifying the overall performance of the model by taking into account both the bias and variance of the errors. A lower RMSE signifies a better fit of the model to the data. Thirdly, bias refers to the systematic deviation of the super-resolution model predictions from the observed values. The bias provides insight into any consistent overestimation or underestimation tendencies present in the model's predictions compared to the reference data. Finally, the PCC evaluates the linear relationship between predicted and observed values, with a scale ranging from -1 to 1 . A PCC of 1 indicates a perfect positive linear relationship, signifying that the model accurately captures the temporal variability of the data. These four metrics collectively offer a comprehensive evaluation of our model's performance. While R^2 measures how well a regression model fits the observed data, RMSE provides the overall accuracy, bias evaluates the systematic errors or deviations in predictions compared to the true value, and PCC assesses the linear relationship.

4. Results

In the results section, we begin by analyzing the performance of SUPREME on the testing data set. At the end of Section 4.1, we demonstrate that the testing metrics closely align with the training metrics, indicating that the model generalizes effectively. Subsequently, in the remaining results (Sections 4.2–4.5), we incorporate the training region and period for further analysis, while consistently indicating instances where data were included in the training of the super-resolution model.

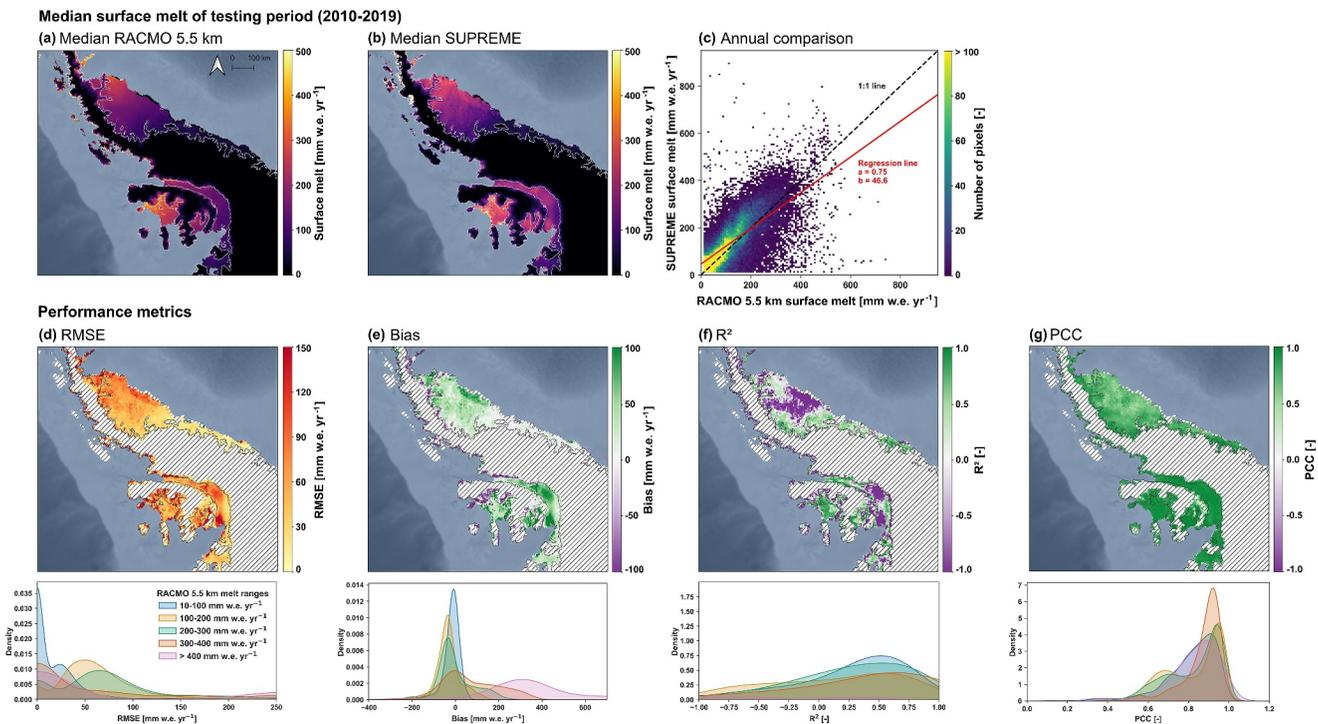


Figure 3. Comparison of RACMO 5.5 km and SUPREME over the Antarctic Peninsula for the testing period (2010–2019), with median melt values for RACMO 5.5 km in panel (a), median melt values for SUPREME in panel (b), and a scatter plot with annual RACMO 5.5 km and SUPREME melt values and regression line in panel (c). The performance metrics with the root mean square error (RMSE), bias, coefficient of determination (R^2), and Pearson correlation coefficient (PCC) are shown in panels (d–g), respectively, where the upper panels show spatial maps and the bottom panels show the distribution of the data in kernel density estimate (KDE) plots, displaying the data per melt regime based on the median RACMO 5.5 km melt values. Pixels with surface melt rates below $10 \text{ mm w.e. yr}^{-1}$ and pixels elevated more than $1,700 \text{ m}$ above sea level are excluded from the analyses, as their inclusion could artificially enhance the performance metrics.

4.1. Performance on Testing Data Set

SUPREME demonstrates strong performance during the testing melt seasons from 2010 to 2019 when compared to the reference data from RACMO 5.5 km, as illustrated by the performance metrics displayed in Figure 3. Notably, SUPREME exhibits slightly higher melt values in specific regions, including the northern part of the Larsen C Ice Shelf, the collapsed A-68 iceberg in 2017 (Braakmann-Folgmann et al., 2022), George VI, and Wilkins ice shelves (see Figures 3a and 3b). For the locations of iceberg A-68, Wilkins, and George VI ice shelves see Figure 1. These areas align with locations exhibiting higher RMSE (Figure 3d) and bias (Figure 3e) values. RMSE and bias values vary noticeably among different melt regimes (i.e., 10–100, 100–200, 200–300, 300–400, and over $400 \text{ mm w.e. yr}^{-1}$). The highest RMSE and bias are found for the high melt regime (exceeding $400 \text{ mm w.e. yr}^{-1}$), with median values of 103 and $325 \text{ mm w.e. yr}^{-1}$, respectively. This indicates that SUPREME's performance could be improved, particularly for high melt rates. The lowest R^2 values (Figure 3f) are observed for both low and high melt regimes, specifically at the center of the Larsen C Ice Shelf, and the center and south of George VI Ice Shelves, with values around -1 , indicating that SUPREME does not adequately capture the variability in surface melt observed in RACMO 5.5 km.

The PCC (Figure 3g) is less dependent on the melt regimes and has a median value of 0.86, indicating a high level of correlation between the annual melt values predicted by RACMO 5.5 km and SUPREME for most of the Antarctic Peninsula. At the grounding zone of the Larsen C Ice Shelf and the George VI Ice Shelf, the PCC surpasses 0.90, underlining a particularly strong correlation in these regions.

To assess the performance of SUPREME and RACMO 5.5 km during the training (2001–2002 to 2006–2007) and non-training (2007–2008 to 2018–2019) periods, we conducted a paired t -test. We compared the yearly mean RMSE values between RACMO 5.5 km and SUPREME for the six melt seasons of the training period against two segments of the non-training period: the first segment (2007–2008 to 2012–2013) and the second segment (2013–

2014 to 2018–2019). The results indicate no significant difference in RMSE between RACMO 5.5 km and SUPREME for both the training and non-training periods. Table A1 presents the yearly average performance metrics of surface melt for the Antarctic Peninsula and Larsen C during the training (2001–2002 to 2006–2007) and non-training (2007–2008 to 2018–2019) periods.

4.2. Comparison of SUPREME and RACMO 27 km

4.2.1. Temporal Analysis

Figure 4 shows consistent year-to-year area-integrated variability among RACMO 27 km and SUPREME (with a PCC >0.95 for all study regions), despite notable quantitative differences in the modeled values. Between 2001–2002 and 2018–2019, the RACMO 27 km model projects an annual surface melt over the entire Antarctic Ice Sheet of 98.5 and a standard deviation of the inter-annual variability of ± 22.0 Gt, whereas the SUPREME model reports lower numbers at 79.5 ± 14.9 Gt yr^{-1} (Figure 4a). For the Antarctic Peninsula (Figure 4b) and Larsen C (Figure 4c), RACMO 27 km yields higher annual melt values (37.3 ± 10.7 Gt and 13.6 ± 4.6 Gt, respectively) than SUPREME (34.6 ± 9.1 Gt and 12.4 ± 4.1 Gt, respectively). When comparing RACMO 27 km and SUPREME in the remainder of the ice sheet, we observe similar annual surface melt values for most ice shelves. However, a notable exception is Shackleton (Figure 4f), where RACMO 27 km reports approximately 2 Gt more surface melt (6.0 ± 2.2 Gt) compared to SUPREME (4.3 ± 1.4 Gt).

4.2.2. Spatial Analysis

To get a better understanding of when and where RACMO 27 km tends to result in higher surface melt values than SUPREME, we can examine the spatial patterns of Larsen C, Roi Baudouin, Amery, Shackleton, and Abbot ice shelves (Figure 5). These five ice shelves are selected for closer examination due to their relatively high surface melt rates.

The lower surface melt values in SUPREME compared to RACMO 27 km presented in Figure 4 are not uniform; rather, surface melt exhibits localized patterns of both lower and higher melt values. Lower surface melt values for SUPREME are observed over the Larsen C Ice Shelf, especially during the melt season of 2011–2012, as well as across all melt seasons over the Shackleton Ice Shelf. One exception is the northern part of the Shackleton Ice Shelf, where SUPREME shows a peak in surface melt with an average annual value of approximately 600 mm w.e., while RACMO 27 km has an average of around 160 mm w.e. annually. These high melt values for SUPREME can likely be attributed to the extremely low albedo values in the vicinity of this area (see Figure A5), connected to a persistent polynya adjacent to the Shackleton Ice Shelf (Nihashi & Ohshima, 2015). With median albedo values of this open water of around 0.1, SUPREME most likely misidentifies this phenomenon as intense surface melt. For specific areas within the grounding zone, SUPREME consistently demonstrates higher melt values than RACMO 27 km. This is especially evident during the high melt season of 2012–2013 for the center part of the grounding zone of Roi Baudouin, the eastern part of Amery, and the northern part of Abbot ice shelves.

4.2.3. Data Spread Analysis

Figure 6 quantifies the surface melt values of RACMO 27 km and SUPREME, distinguishing between values on the floating ice shelf and in the grounding zone (see Figure A4 for precise locations). The surface melt range in RACMO 27 km and SUPREME is assessed using the 5th (Q5) and 95th (Q95) percentiles. Notably, there is a consistently wider range of surface melt in SUPREME compared to RACMO 27 km, observed in both the ice shelf and grounding zone. In contrast, the mean and median values for both ice shelves and grounding zones are generally higher in RACMO 27 km compared to SUPREME (except for the Amery Ice Shelf), reaffirming the lower melt volumes in SUPREME in comparison to RACMO 27 km.

Examining the floating ice shelves, the most substantial discrepancy between RACMO 27 km and SUPREME is observed on the Shackleton Ice Shelf. RACMO 27 km shows a Q5–Q95 range of 73–234 mm w.e. yr^{-1} , while SUPREME exhibits a wider range spanning from 2 to 659 mm w.e. yr^{-1} (Figure 6d). The elevated melt values (Q95) in SUPREME for the Shackleton Ice Shelf are also visually evident in Figure 5, particularly in the northern sector of the shelf. Also, on the grounding zone, SUPREME consistently displays a larger Q5–Q95 range compared to RACMO 27 km. The most pronounced differences are seen in the cases of Amery (Figure 6c) and

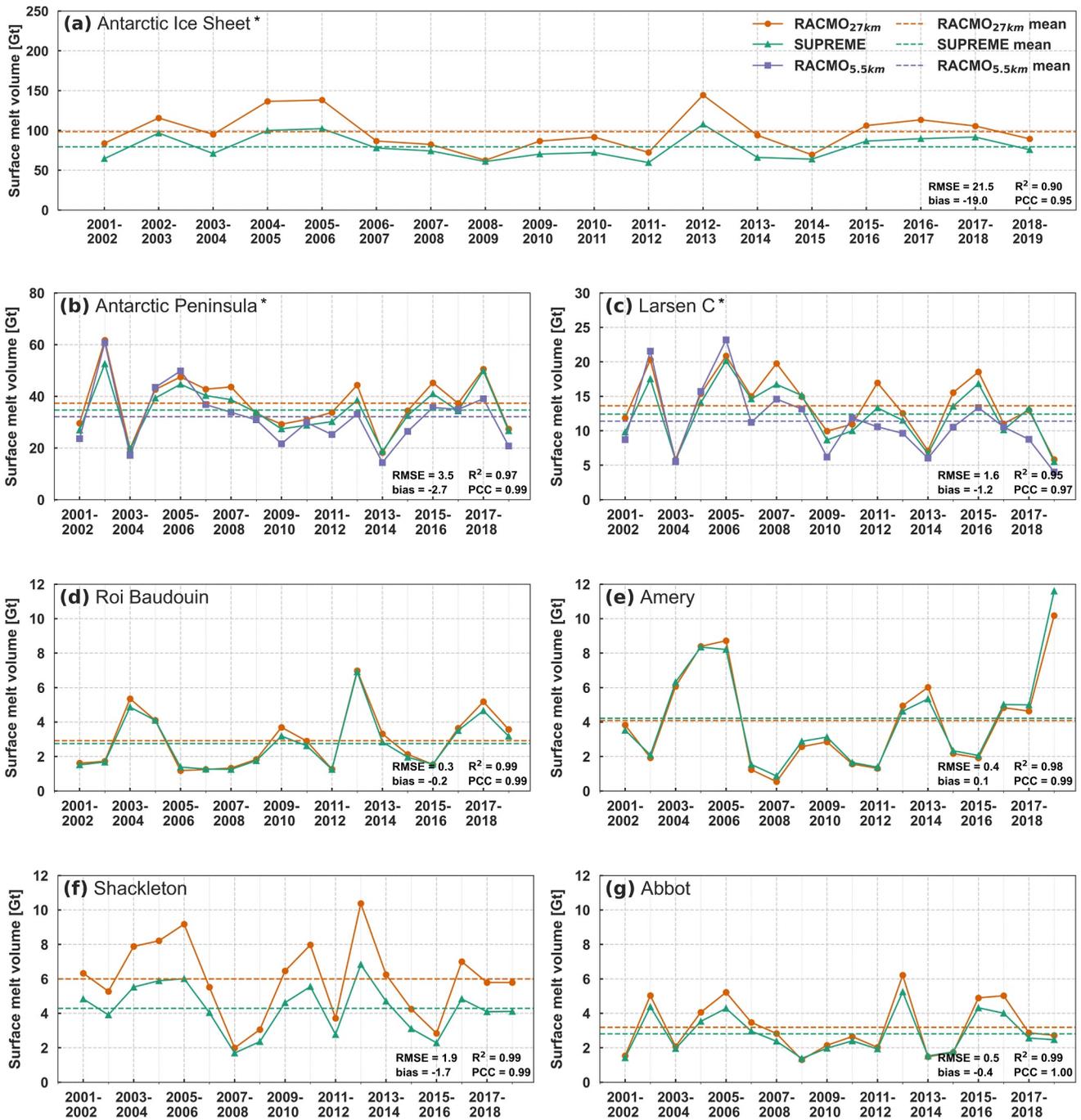


Figure 4. Temporal evolution of surface melt for the melt seasons from 2001 to 2019, represented by RACMO 27 km (in orange), SUPREME (in green), and RACMO 5.5 km (in purple), with dashed lines representing the 2001–2019 mean. RACMO 5.5 km data are exclusively available for the Antarctic Peninsula (b) and Larsen C (c). For the Antarctic-wide time series (a), and specific locations including Roi Baudouin (d), Amery (e), Shackleton (f), and Abbot (g), only RACMO 27 km and SUPREME data are depicted. Caution is advised when interpreting the data sets marked with an asterisk (a–c), as a part of these data were used in training SUPREME. The extent of the selected regions can be found in Figure A4. Relevant statistics used to compare RACMO 27 km and SUPREME, including root mean square error (RMSE), bias, coefficient of determination (R^2), and Pearson correlation coefficient (PCC), are listed.

Abbot (Figure 6e). RACMO 27 km shows a Q95 value of 195 mm w.e. yr^{-1} for Amery, while SUPREME yields a higher value of 385 mm w.e. yr^{-1} . The contrast is even more pronounced for Abbot, with RACMO 27 km at 165 mm w.e. yr^{-1} and SUPREME at 711 mm w.e. yr^{-1} .

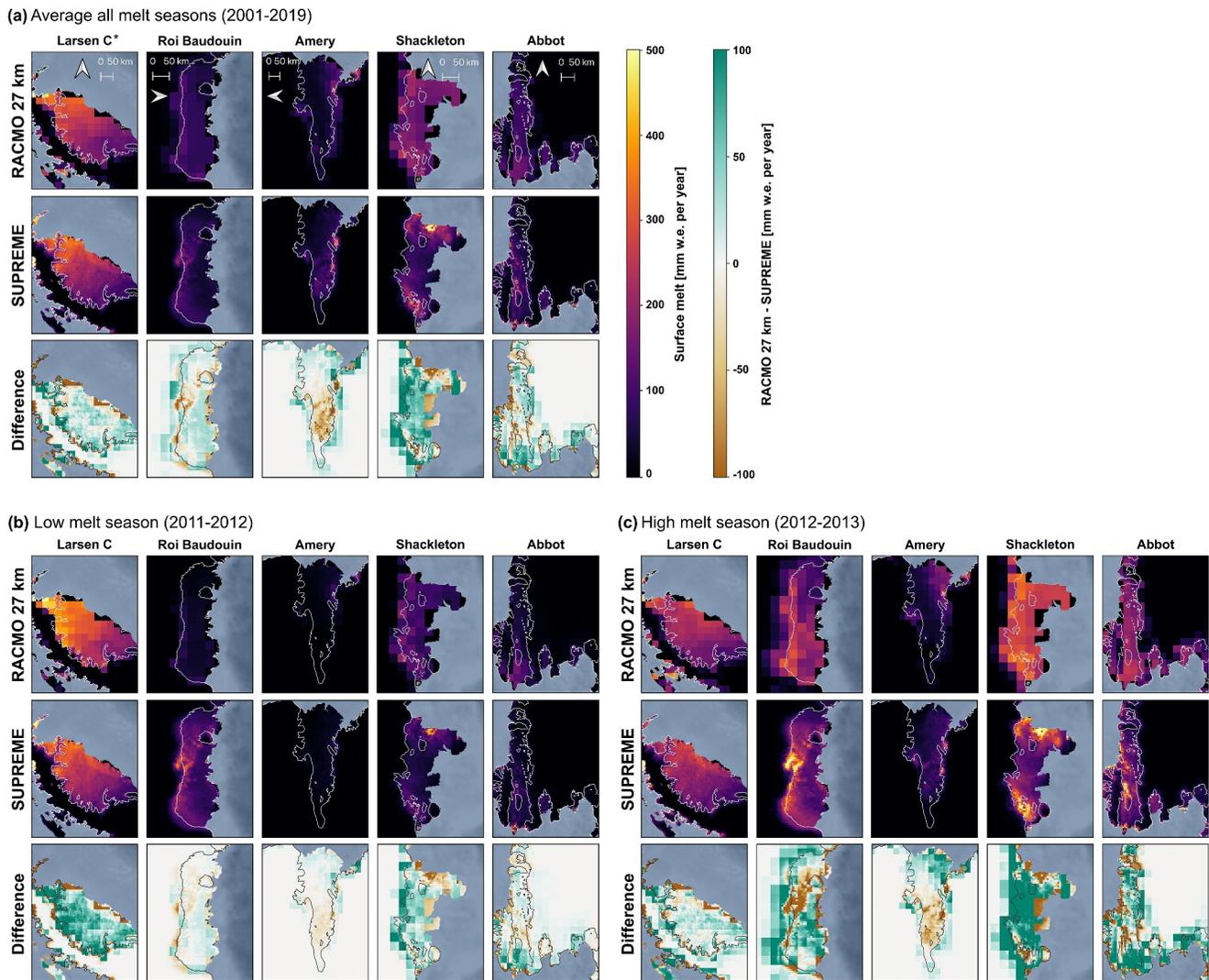


Figure 5. Surface melt for (a) the average of melt seasons of 2001–2019, (b) low melt season 2011–2012, and high melt season 2012–2013 for RACMO 27 km, SUPREME, and the difference between RACMO 27 km and SUPREME. Caution is advised when interpreting the data sets marked with an asterisk (i.e., Larsen C, panel (a), as a part of these data were used in training SUPREME).

4.3. Comparison of SUPREME and AWS

Figure 7 presents a comparison between RACMO 5.5 km, RACMO 27 km, and SUPREME with five AWS. Notably, SUPREME exhibits a lower RMSE than RACMO 27 km for all five AWS, and a lower absolute bias for AWS14, AWS15, and AWS18. The largest difference between SUPREME and RACMO 27 km is observed for AWS18 (Figure 7d). For this AWS, SUPREME yields an RMSE of 87 mm w.e. yr⁻¹ and a bias of 23 mm w.e. yr⁻¹, compared to the 244 mm w.e. yr⁻¹ RMSE and -216 mm w.e. yr⁻¹ bias for RACMO 27 km. Due to its coarse resolution, RACMO 27 km fails to resolve the surface melt in areas with highly variable topography, such as around AWS18. This AWS is situated near the grounding line of the Larsen C ice shelf, east of the Antarctic Peninsula mountain range, and is prone to föhn events that contribute to surface melt (Kuipers Munneke et al., 2018; Wiesenecker et al., 2018).

4.4. Comparison of SUPREME and Other Surface Melt Products

Given that RACMO 5.5 km data are limited to the Antarctic Peninsula, a comprehensive Antarctic-wide assessment between SUPREME and RACMO 5.5 km data set is not feasible. Here, we proceeded to compare SUPREME with other melt products, that is, RACMO 2 km, QuikSCAT, PMW remote sensing, and an optical

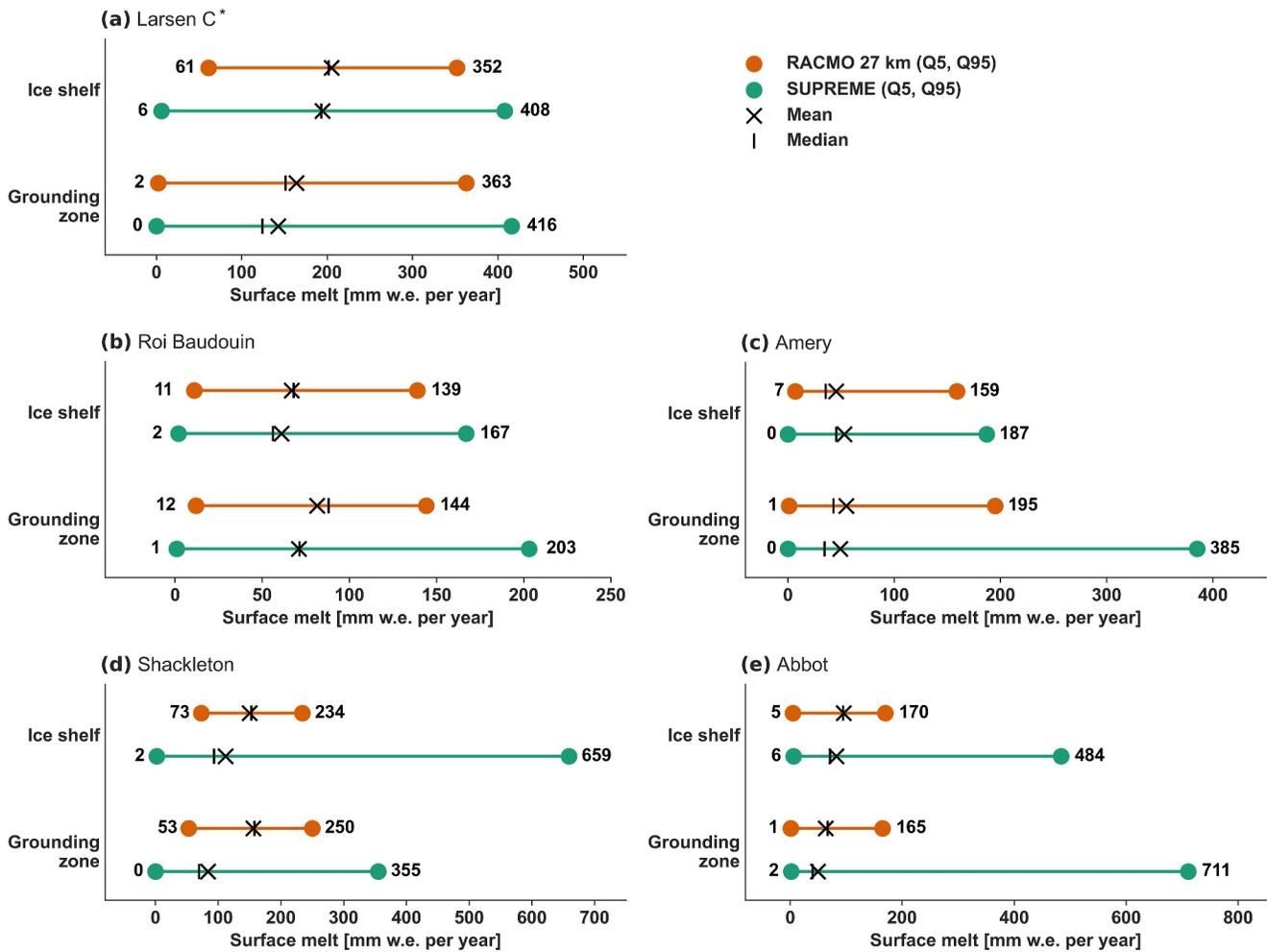


Figure 6. The range of surface melt values for the ice shelves and grounding zones, showing the mean (crosses), median (vertical lines), 5th (Q5), and 95th (Q95) percentile values, for the melt seasons spanning from 2001 to 2019. The data are presented for both RACMO 27 km (in orange) and SUPREME (in green) at the ice shelf and grounding zone of (a) Larsen C, (b) Roi Baudouin, (c) Amery, (d) Shackleton, and (e) Abbot. Caution is advised when interpreting the data sets marked with an asterisk (i.e., Larsen C in panel (a)), as a part of these data were used in training SUPREME.

remote sensing product that provides information on both melt and slush. Both RACMO 2 km and QuikSCAT offer quantitative data on surface melt in mm w.e., whereas PMW and optical remote sensing techniques provide binary outputs without specific quantification of melt volume.

Each of the validation data sets has distinct underlying methodologies (as was explained in Section 2.3), resulting in large disparities in local surface melt patterns, as illustrated in Figure 8. In addition, Figure 9 illustrates the comparison between SUPREME and the validation products using both a scatter plot and several performance metrics. In the period from 2001 to 2009 (Figures 8a and 9a), the Larsen C Ice Shelf should be considered with caution, as this data set was used during training of SUPREME. For the Roi Baudouin Ice Shelf, both QuikSCAT and PMW data reveal higher surface melt levels in the eastern region (150 mm w.e. yr⁻¹ and 35 melt days yr⁻¹, respectively) compared to the western area (75 mm w.e. yr⁻¹ and 20 melt days yr⁻¹, respectively). This gradient, however, is not as clearly discernible in the RACMO 27 km, RACMO 2 km, or SUPREME data sets. Also, the R² values are low when comparing SUPREME to QuikSCAT (−2.18), and slightly higher but still very low for RACMO 27 km (−1.77) and RACMO 2 km (−0.98), as illustrated in Figure 9a. For the Amery Ice Shelf, both SUPREME and RACMO 2 km display some high surface melt values (250 mm w.e. yr⁻¹), particularly on the southern part of the ice shelf. Likewise, QuikSCAT and PMW also show increased melt in the southern region (150 mm w.e. yr⁻¹ and 40 melt days yr⁻¹, respectively), but the precise locations with the highest melt differ. QuikSCAT shows the highest melt values toward the southeast, whereas PMW tends toward the center-south,

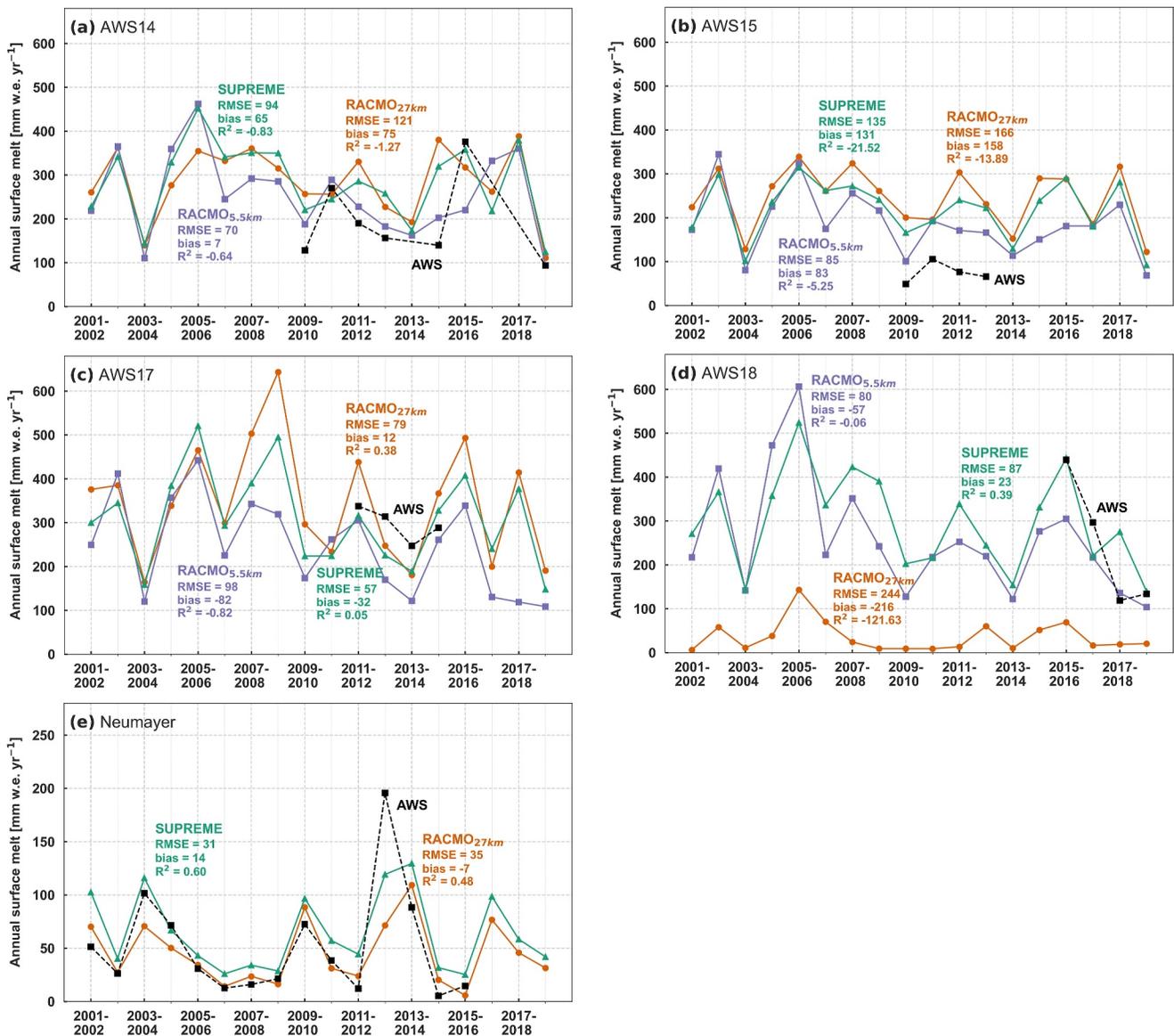


Figure 7. Surface melt per melt season by RACMO 27 km (in orange), SUPREME (in green), RACMO 5.5 km (in purple), and automatic weather stations (AWS) (in black). Data are collected from five automatic weather stations: AWS14 (a), AWS15 (b), AWS17 (c), AWS18 (d), and Neumayer (e). The Root Mean Squared Error (RMSE), bias, and coefficient of determination (R^2) are represented by the color of the respective model in comparison to AWS measurements. Figure A2 displays the locations of the AWS.

gradients that are not present in SUPREME. It is worth noting that RACMO 27 km and QuikSCAT and RACMO 27 km report surface melt below 200 mm w.e. yr⁻¹ for Amery, while SUPREME and RACMO 2 km register values exceeding 500 mm w.e. yr⁻¹. The smoother representation of surface melt by RACMO 27 km and QuikSCAT, attenuating peak values, is likely a result of their coarser spatial resolutions (i.e., 27 km for RACMO 27 km, and 25 km as a native spatial resolution for QuikSCAT). Likewise, the Shackleton Ice Shelf displays elevated melt values in SUPREME and RACMO 2 km, exceeding 400 mm w.e. at the southern grounding line. Moreover, in the northern region, SUPREME shows melting rates exceeding 500 mm w.e. yr⁻¹. The high surface melt values at the southern grounding line of the Shackleton Ice Shelf are also observed in other validation products, exceeding 250 mm w.e. yr⁻¹ in QuikSCAT, and 40 melt days yr⁻¹ in PMW. For the Abbot Ice Shelf, the most substantial melt values are concentrated in the southern region bordering the Amundsen Sea. Here, SUPREME depicts surface melt values exceeding 400 mm w.e. yr⁻¹, whereas RACMO 27 km, RACMO 2 km, and QuikSCAT melt values remain under 200 mm w.e. yr⁻¹.

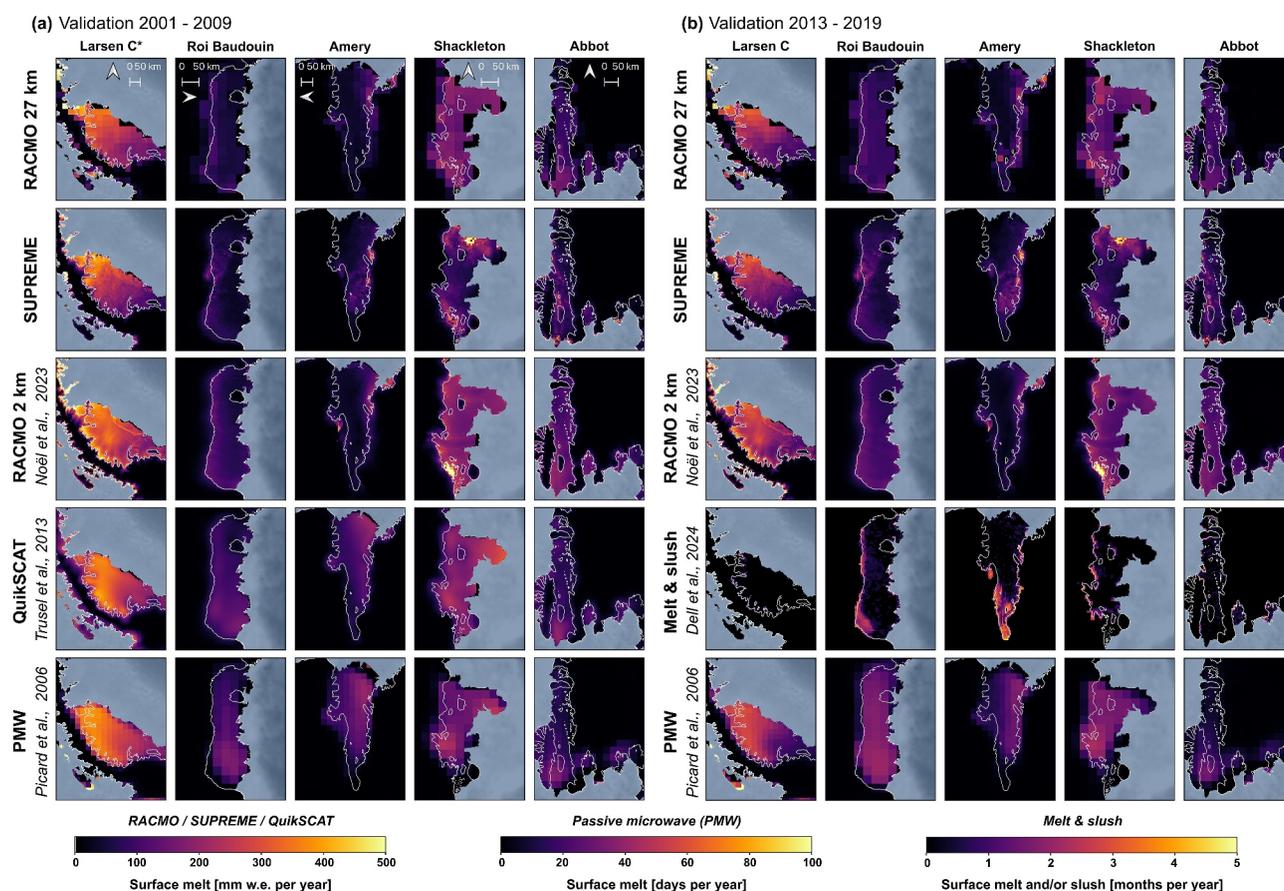
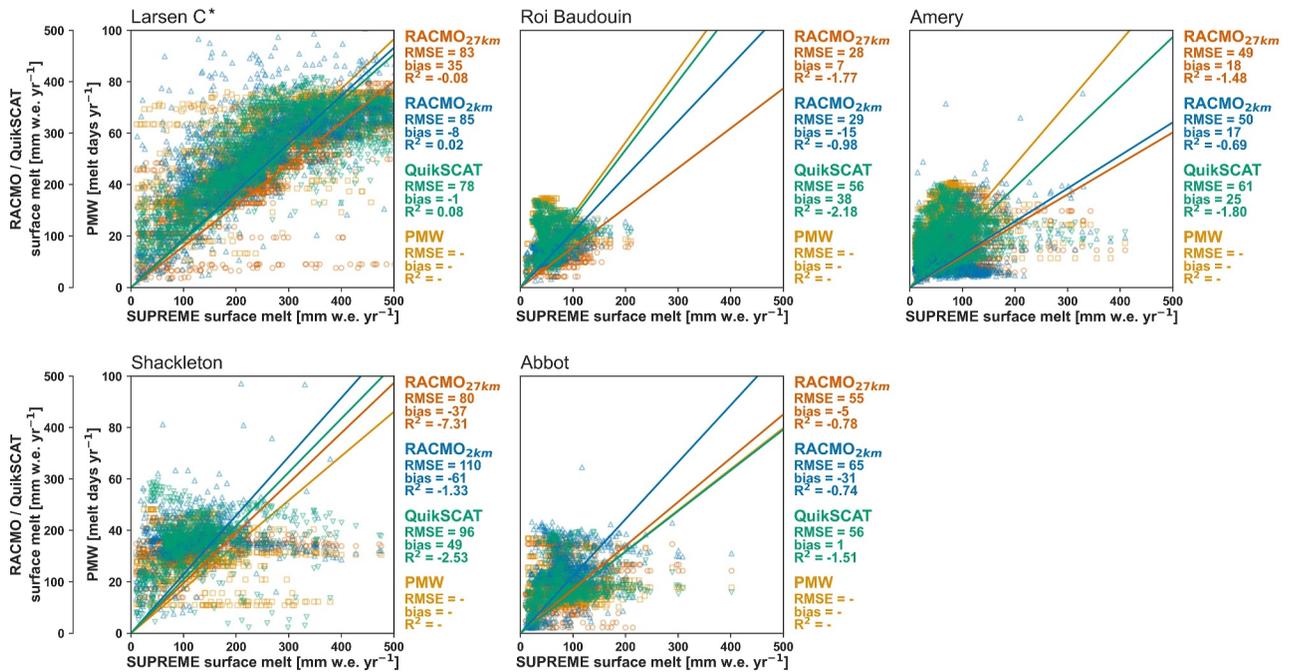


Figure 8. Spatial comparison of SUPREME to validation data sets across two distinct periods: (a) 2001–2009 and (b) 2013–2019, encompassing Larsen C, Roi Baudouin, Shackleton ice shelves, and Mary Byrd Land. For the 2001–2009 period, SUPREME is compared to RACMO 27 km, RACMO 2 km, QuikSCAT, and passive microwave (PMW) remote sensing (note that the latter employs a different unit than the other methods). In the 2013–2019 period, SUPREME is compared to RACMO 27 km, RACMO 2 km, a surface melt and slush product from optical remote sensing, and PMW remote sensing (both optical and PMW measurements employ different units than the other methods). Caution is advised when interpreting the data set marked with an asterisk (i.e., SUPREME, Larsen C, 2001–2009), as a part of these data were used during the training of SUPREME.

Moving to the second comparison period (2013–2019) in Figures 8b and 9b, we compare SUPREME to RACMO 27 km, RACMO 2 km, a melt and slush product derived from optical imagery, and PMW. As noted in Section 2.3, since optical imagery detects meltwater solely when slushy or ponded water is present, they do not represent wet snow that may later refreeze within the firn layer. Therefore, SUPREME is expected to show similar melt patterns to the optical melt and slush product only in locations characterized by a low firn air content, where meltwater accumulates at the surface. Figure 8b shows that the melt and slush product predominantly reveals high values around the grounding line, indeed often corresponding to locations with supraglacial melt lake presence (e.g., Stokes et al., 2019). Some of these locations align with high melt values detected in SUPREME, such as the central part of the grounding line of the Roi Baudouin Ice Shelf and the southern part of the grounding line of the Amery Ice Shelf, where surface melt values range between 250 and 500 mm w.e. yr⁻¹. For the Roi Baudouin Ice Shelf, the RMSE values are relatively low compared to the other ice shelves, standing at 33 mm w.e. yr⁻¹ for RACMO 27 km and 38 mm w.e. yr⁻¹ for RACMO 2 km, as illustrated in Figure 9b. However, there are also instances where there is not a clear overlap between surface melt in the melt and slush product and SUPREME, as observed over the western part of the Amery Ice Shelf. Here, in the optical remote sensing product, melt or slush is detected on average for at least three months per year, while RACMO 27 km, SUPREME, RACMO 2 km, and the PMW product show low melt values around 50–100 mm w.e. yr⁻¹ or 5 melt days yr⁻¹. Similar to the 2001–2009 period, there are also variations in melt patterns between PMW and SUPREME for 2013–2019, with PMW displaying a smoother surface melt pattern with fewer high-resolution details compared to SUPREME. For both

(a) Validation 2001-2009



(b) Validation 2013-2019

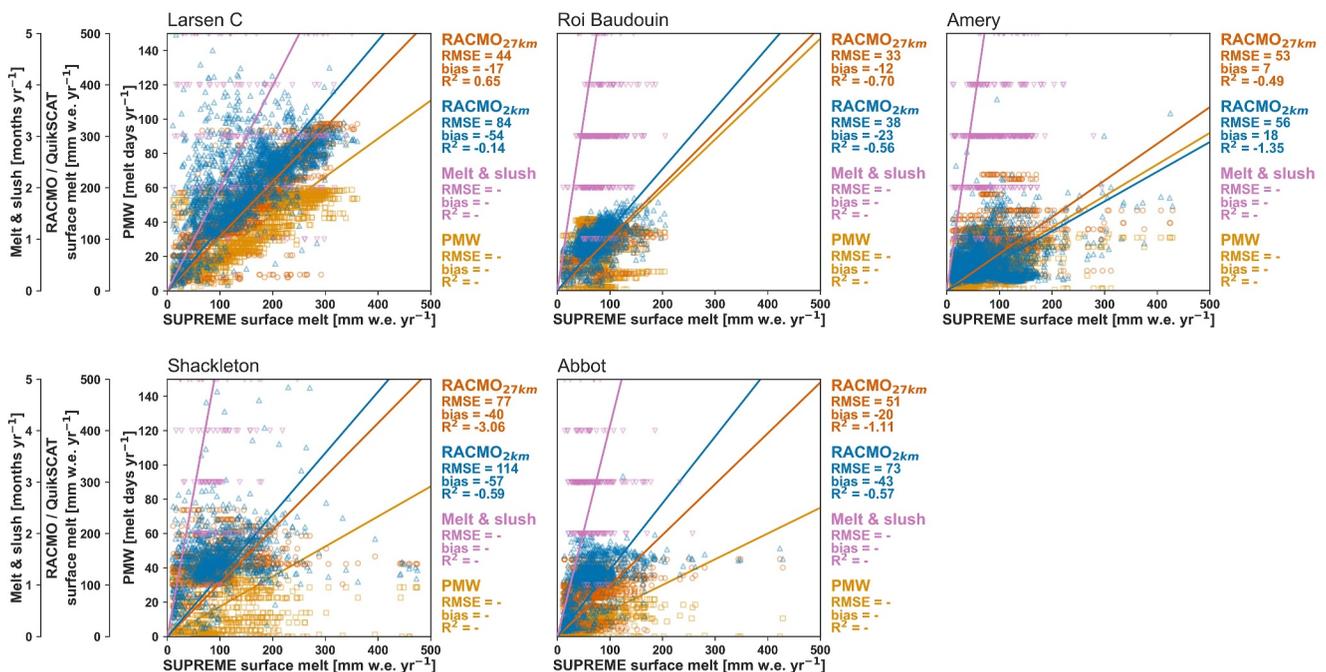


Figure 9. Comparison of SUPREME to validation data sets across two distinct periods: (a) 2001–2009 and (b) 2013–2019, encompassing Larsen C, Roi Baudouin, Shackleton ice shelves, and Mary Byrd Land, complementing Figure 8. The scatter plots compare SUPREME with validation products (i.e., RACMO 27 km, RACMO 2 km, QuikSCAT, passive microwave (PMW) remote sensing, and surface melt and slush product), with each line representing the least-square fit. The Root Mean Squared Error (RMSE), bias, and coefficient of determination (R^2) are represented by the color of the respective validation data in comparison to SUPREME, for PMW and the melt and slush product only the R^2 is included. Caution is advised when interpreting the data sets marked with an asterisk (i.e., Larsen C, panel (a)), as a part of these data were used in training SUPREME. Figure A4 shows the spatial extents analyzed.

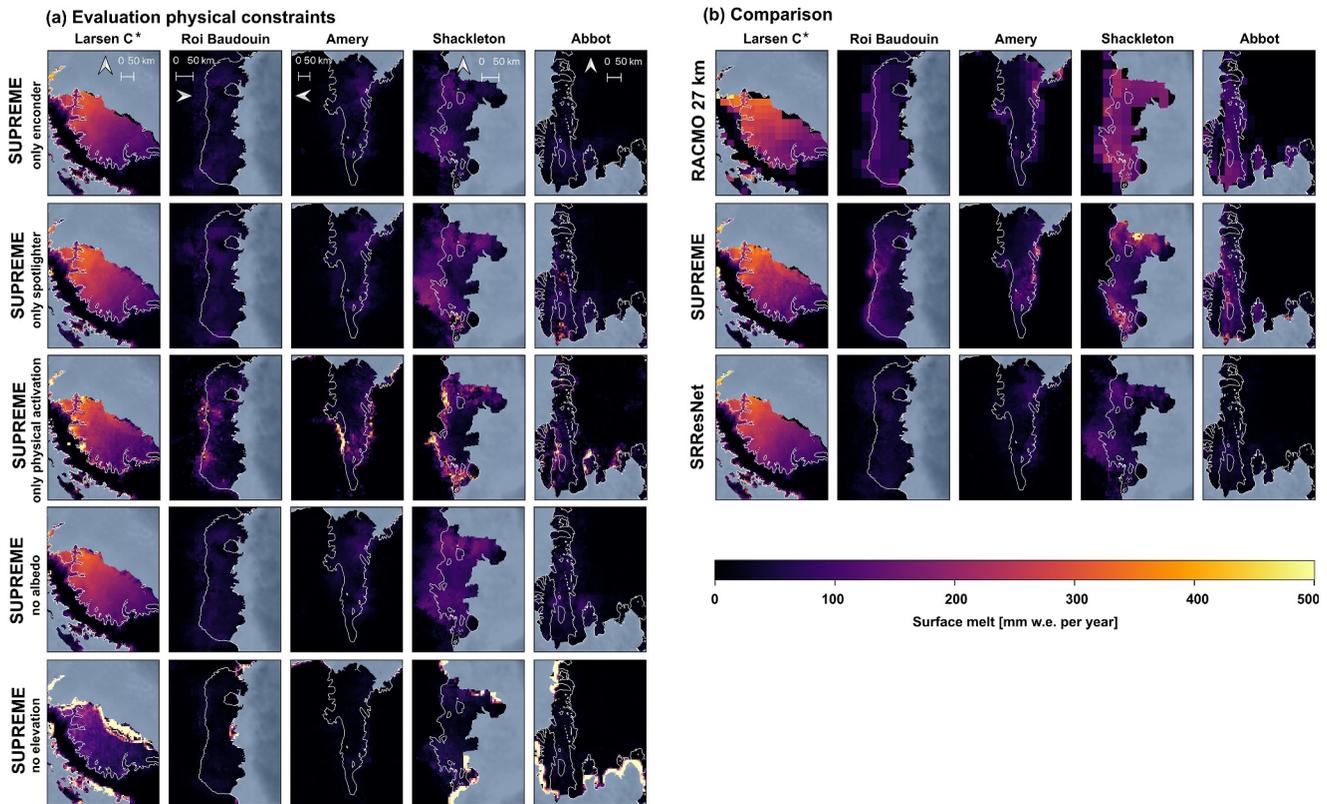


Figure 10. Evaluating the performance of surface melt for the average of melt seasons of 2001–2019 across different model architectures. (a) SUPREME with individual physical constraints (only encoder, only spotlighter, only physical activation), and SUPREME with the exclusion of albedo and elevation input features from the model architectures, respectively. (b) Comparative data featuring RACMO 27 km, SUPREME with all physical constraints and input features, and SRResNet without any physical constraint and with both albedo and elevation excluded from the model architecture. Caution is advised when interpreting the data sets marked with an asterisk (i.e., Larsen C in panels (a) and b)), as a part of these data were used in training SUPREME.

the Shackleton and Abbot ice shelves, we observe relatively large RMSE values, ranging from 51 to 114 mm w.e. yr⁻¹.

4.5. Importance of Physical Constraints

In Sections 4.2, 4.3, and 4.4, we exclusively presented SUPREME, which uses the SRResNet model architecture along with three physical constraints to incorporate remote sensing data. To assess the importance of these individual physical constraints, Figure 10a presents models trained exclusively with each physical constraint (SUPREME_{only encoder}, SUPREME_{only spotlighter}, SUPREME_{only physical activation}). Additionally, models with all physical constraints but only one of the remote sensing input features are depicted (SUPREME_{no albedo}, SUPREME_{no elevation}). Figure 10b showcases SUPREME with all three physical constraints alongside SRResNet without any physical constraints. Additionally, Figure A7 shows the mean, median, and spread of the data sets presented in Figure 10, and Table A2 shows the performance metrics of the different super-resolution models compared to RACMO 27 km.

The physical constraints responsible for integrating encoded albedo and elevation data into the model architecture (SUPREME_{only encoder}), incorporating spotlighted albedo and elevation data (SUPREME_{only spotlighter}), and applying the physical activation function (SUPREME_{only physical activation}) demonstrate a similar melt pattern as SUPREME and RACMO 27 km over the Larsen C Ice Shelf, exhibiting a south-north melt trend ranging from roughly 100–400 mm w.e. yr⁻¹. However, beyond the Antarctic Peninsula, SUPREME_{only encoder}, SUPREME_{only spotlighter}, and SUPREME_{only physical activation} compare less favorably to SUPREME and RACMO 27 km. The median annual melt values over the Roi Baudouin, Amery, Shackleton, and Abbot ice shelves are higher for SUPREME (averaging 61 mm w.e. yr⁻¹) and RACMO 27 km (averaging 88 mm w.e. yr⁻¹), compared to SUPREME_{only encoder} (averaging 37 mm w.e. yr⁻¹), SUPREME_{only spotlighter} (averaging

50 mm w.e. yr⁻¹), and SUPREME_{only physical activation} (averaging 44 mm w.e. yr⁻¹) (see Figure A7). In comparison to RACMO 27 km on an Antarctic-wide scale, SUPREME exhibits an RMSE of 32 mm w.e. yr⁻¹, while SUPREME_{only encoder} shows a slightly higher value at 33 mm w.e. yr⁻¹. However, SUPREME_{only spotlihter} and SUPREME_{only physical activation} exhibit notably higher RMSE values of 55 mm w.e. yr⁻¹ and 53 mm w.e. yr⁻¹ respectively (see Table A2).

Similar observations apply to models that lack either albedo (SUPREME_{no albedo}) or elevation (SUPREME_{no elevation}), as they do not generalize well beyond the Antarctic Peninsula. These models also tend to underestimate surface melt compared to SUPREME and RACMO 27 km, as shown by their median annual melt values over the Roi Baudouin, Amery, Shackleton, and Abbot ice shelves for SUPREME_{no albedo} (averaging 49 mm w.e. yr⁻¹) and SUPREME_{no elevation} (averaging 23 mm w.e. yr⁻¹) (see Figure A7). When compared to RACMO 27 km across Antarctica, SUPREME shows an RMSE of 32 mm w.e. yr⁻¹, akin to SUPREME_{no albedo}. However, SUPREME_{no elevation} displays a notably higher RMSE of 60 mm w.e. yr⁻¹ (see Table A2).

Besides the finding that models lacking physical constraints or input features display less melt on the ice shelves compared to RACMO 27 km and SUPREME, there are also discrepancies in spatial patterns. The SUPREME model, incorporating all input features, shows heightened melt values in the grounding zones, aligning with observations from the validation data sets (Figure 8). However, models without albedo data (i.e., SUPREME_{no albedo} and SRResNET) indicate surface melt extending up to 50 km inland. This pattern persists with SUPREME_{only encoder} and SUPREME_{only spotlihter}. This is most clearly visible in the case of the Shackleton Ice Shelf, where values range between 50 and 150 mm w.e. yr⁻¹, whereas RACMO 27 km, SUPREME, and the validation data sets in Figure 8 demonstrate values close to zero mm w.e. yr⁻¹. Therefore, incorporating both albedo and elevation data is crucial for accurately representing melt patterns outside the training region.

5. Discussion

Our study illustrated that the newly developed super-resolution model, incorporating physical constraints derived from albedo and elevation observations, outperforms models lacking such constraints. Notably, the single-image super-resolution model, SRResNet, which was trained solely on RACMO 27 km, and the models lacking at least one physical constraint, exhibited poor performance outside of the Antarctic Peninsula. The melt patterns observed in these models differed notably from those seen in RACMO 27 km, RACMO 2 km, or QuikSCAT, with most showing less surface melt (Figure A7). On the other hand, in some cases, models lacking albedo (i.e., SUPREME_{no albedo}) or albedo and elevation (i.e., SRResNet) led to increased melting further inland, which was not present in other melt products. This poor performance underscored the need to integrate both albedo and elevation data in an efficient manner into the super-resolution model in order to produce an accurate downscaled surface melt product. The incorporation of additional input features, such as climate variables like wind speed and direction, could potentially enhance the super-resolution product even further.

Integrating remote sensing data through deep learning into the downscaling of RCMs presents considerable promise for multiple research purposes. Our approach could serve as a proof-of-concept and be extended to other studies. Firstly, besides RACMO 27 km, our developed super-resolution algorithm could be extended to other RCMs like MAR (Agosta et al., 2019). Secondly, the potential application of super-resolution techniques extends across diverse data sets, including observations, model simulations, and emulated data (Vandal et al., 2017). Lastly, the super-resolution architecture is not limited to surface melt; it can enhance various other RCM variables. A critical advancement for Antarctica could involve applying such a super-resolution technique not only to surface melt but also to basal melt (e.g., Burgard et al., 2023), bed topography (e.g., Cai et al., 2023), or sea ice motion (e.g., Petrou et al., 2018). In this study, albedo and elevation are identified as relevant remote sensing data for downscaling surface melt, given their correlation with surface melt. However, downscaling other variables from RCMs may require the incorporation of completely different remote sensing data, depending on the physical processes that influence those variables. For example, in the case of downscaling surface temperature, Li et al. (2019) demonstrated the value of including elevation, reflectance from optical remote sensing, vegetation indices, and elevation in the downscaling process.

When assessing the downscaled surface melt product developed with the SUPREME architecture, we observe favorable comparisons against RACMO 5.5 km, AWS, and various remote sensing products associated with surface melt. Over the Antarctic Peninsula, SUPREME performed well, exhibiting a bias of -22 mm w.e. yr⁻¹

for melt pixels when compared to RACMO 5.5 km. Additionally, in comparison to AWS, SUPREME exhibited an average RMSE of 81 mm w.e. yr^{-1} , outperforming RACMO 27 km, which had an average RMSE of 129 mm w.e. yr^{-1} . However, evaluating SUPREME's performance beyond the Antarctic Peninsula presented challenges. We conducted comparisons with RACMO 2 km, QuikSCAT, PMW remote sensing, and an optical remote sensing-derived melt and slush product. The substantial discrepancies between these data sets made direct assessments challenging. Although some melt patterns, such as elevated surface melt along the grounding zones of the studied ice shelves, were consistent with most of the other products, others, like the heightened melt values north of the Shackleton Ice Shelf, were distinctive to SUPREME and appeared to be an artifact due to low albedo values from the nearby ocean. Conducting follow-up studies to further compare SUPREME against validation data, such as with RACMO data downscaled using different techniques (e.g., a more detailed comparison with the statistically-downscaled RACMO product by Noël et al., 2023a), other RCMs like MAR (Agosta et al., 2019), or quantitative meltwater products from PMW (currently available for Greenland from Zheng et al., 2022), may provide additional insights into meltwater volumes on Antarctica. Additionally, while this study focused on annual surface melt volumes due to memory constraints, future research is planned to study the intra-annual surface melt dynamics of SUPREME and compare them across various validation data sets.

Opting for deep learning over traditional statistical or dynamical downscaling methods to enhance the resolution of RCMs, SUPREME uses the capabilities of deep learning to capture complex patterns and non-linear dependencies in the data. While statistical and dynamical downscaling methods have played instrumental roles in climate research (e.g., Ahmed et al., 2013; Noël et al., 2016, 2023a; Themeal et al., 2012), deep learning offers an alternative downscaling approach by incorporating diverse data sets without the need for statistical linear relationships and integrating good generalization capabilities for efficient climate parameter downscaling. However, significant steps still need to be taken, as the inherent “black box” nature of deep learning poses challenges in interpreting deep learning models. Anticipated efforts in enhancing the interpretability of the model are expected to refine its overall effectiveness and integration into climate research methodologies (Guidotti et al., 2018; La Rocca & Perna, 2022; Savage, 2022).

6. Conclusion

The relatively coarse spatial resolution in RCMs is insufficient for accurately capturing small-scale variations in surface melt across the Antarctic Ice Sheet. In response, we adapted the super-resolution architecture of SRResNet with physical constraints to create SUPREME, a physically-constrained super-resolution model that incorporates remote sensing data—specifically albedo and elevation—into the architecture. Using the SUPREME architecture, the surface melt component from RACMO at 27 km was downscaled to 5.5 km resolution. SUPREME achieved excellent out-of-sample performance within the training region, encompassing the Antarctic Peninsula, as validated by comparisons with AWS and RACMO 5.5 km. Outside the training region, SUPREME exhibited year-to-year variations that closely mirrored those of RACMO 27 km for individual ice shelves and on an Antarctic-wide scale, while also providing additional spatial details absent in RACMO 27 km. Moreover, comparing SUPREME with other melt products, like RACMO 2 km, QuikSCAT, PMW remote sensing, and optical remote sensing, showed similar melt patterns. This study underscores the potential of multi-image super-resolution in downscaling climate variables. Beyond the input features and model architecture examined in this research, numerous other combinations deserve exploration to further enhance the resolution of RCMs.

Appendix A: Additional Figures

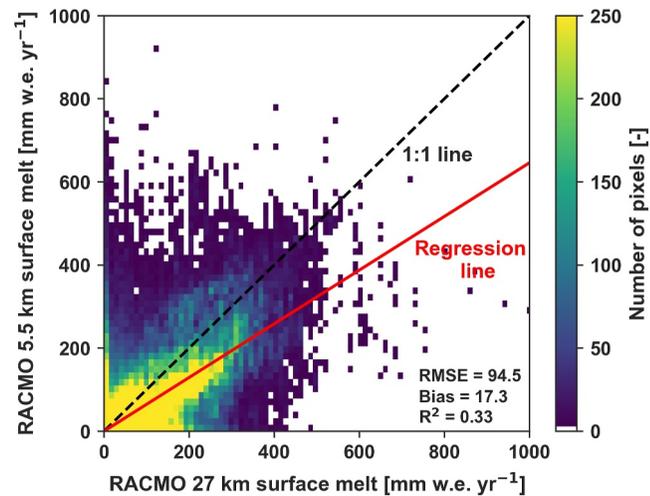


Figure A1. Relation between RACMO 27 km and RACMO 5.5 km surface melt. The red line indicates the regression line, while the black dashed line shows the 1:1 line, highlighting RACMO 5.5 km's tendency to underestimate surface melt compared to RACMO 27 km. RMSE and bias are in mm w.e. yr^{-1} .



Figure A2. Locations of the five AWS used in this study. For additional details on the AWS, see Jakobs et al., 2020. Elevation in the background map is from the Reference Elevation Model of Antarctica (REMA) (Howat et al., 2019).

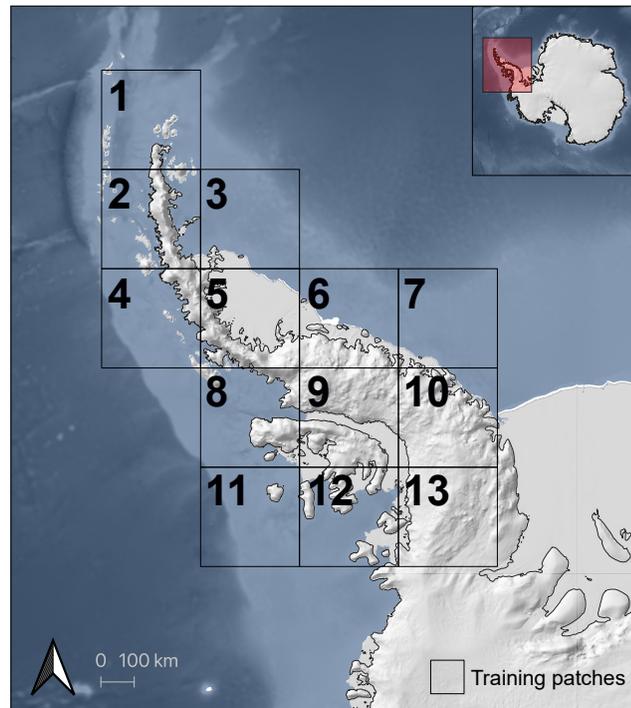


Figure A3. Location of the 13 patches used in training the super-resolution models, all positioned on the Antarctic Peninsula. The red box in the upper right corner highlights the specific location of the Antarctic Peninsula, with a background image of the Cryosat-2 elevation map (Helm et al., 2014).

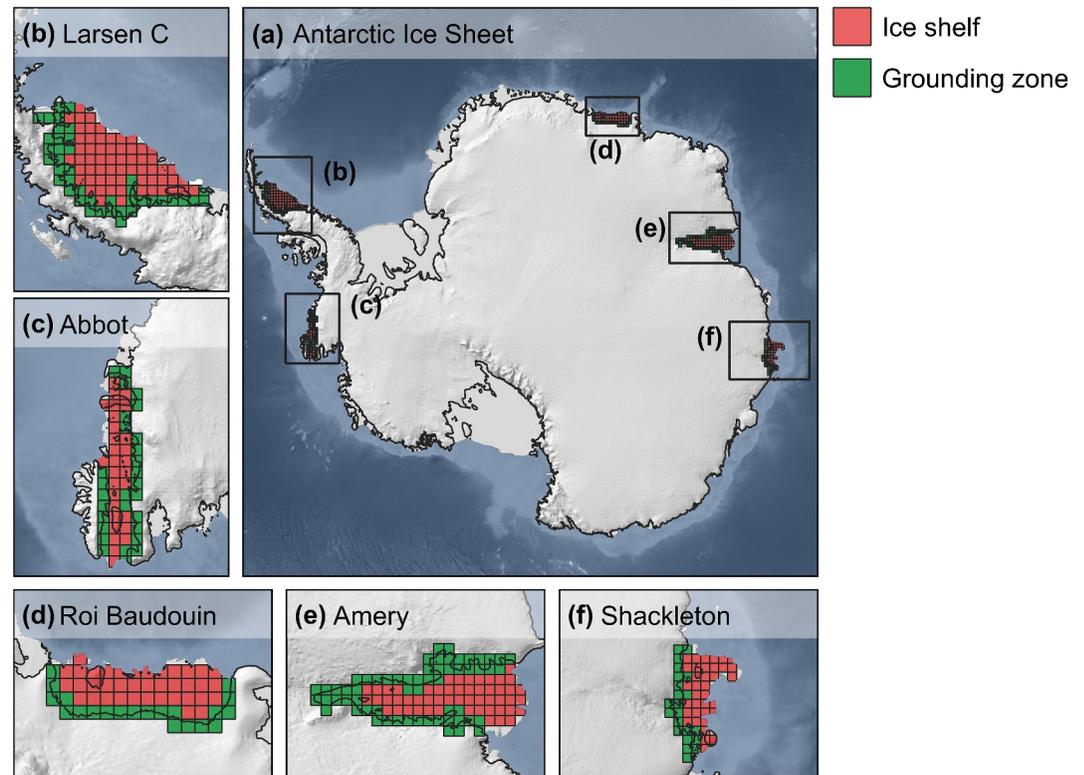


Figure A4. Overview of the five highlighted regions in this study, featuring an Antarctic-wide perspective (a) along with specific views of Larsen C (b), Abbot (c), Roi Baudouin (d), Amery (e), and Shackleton (f) ice shelves. The pixels in this figure have a 27 km resolution matching the RACMO 27 km resolution, and for data with 5.5 km resolution (albedo, elevation, RACMO 5.5 km), all data within these pixels were considered. Region-wide assessments involved selecting all pixels, ice shelf-wide assessments only considered the red pixels, and grounding zone assessments focused solely on the green ones.

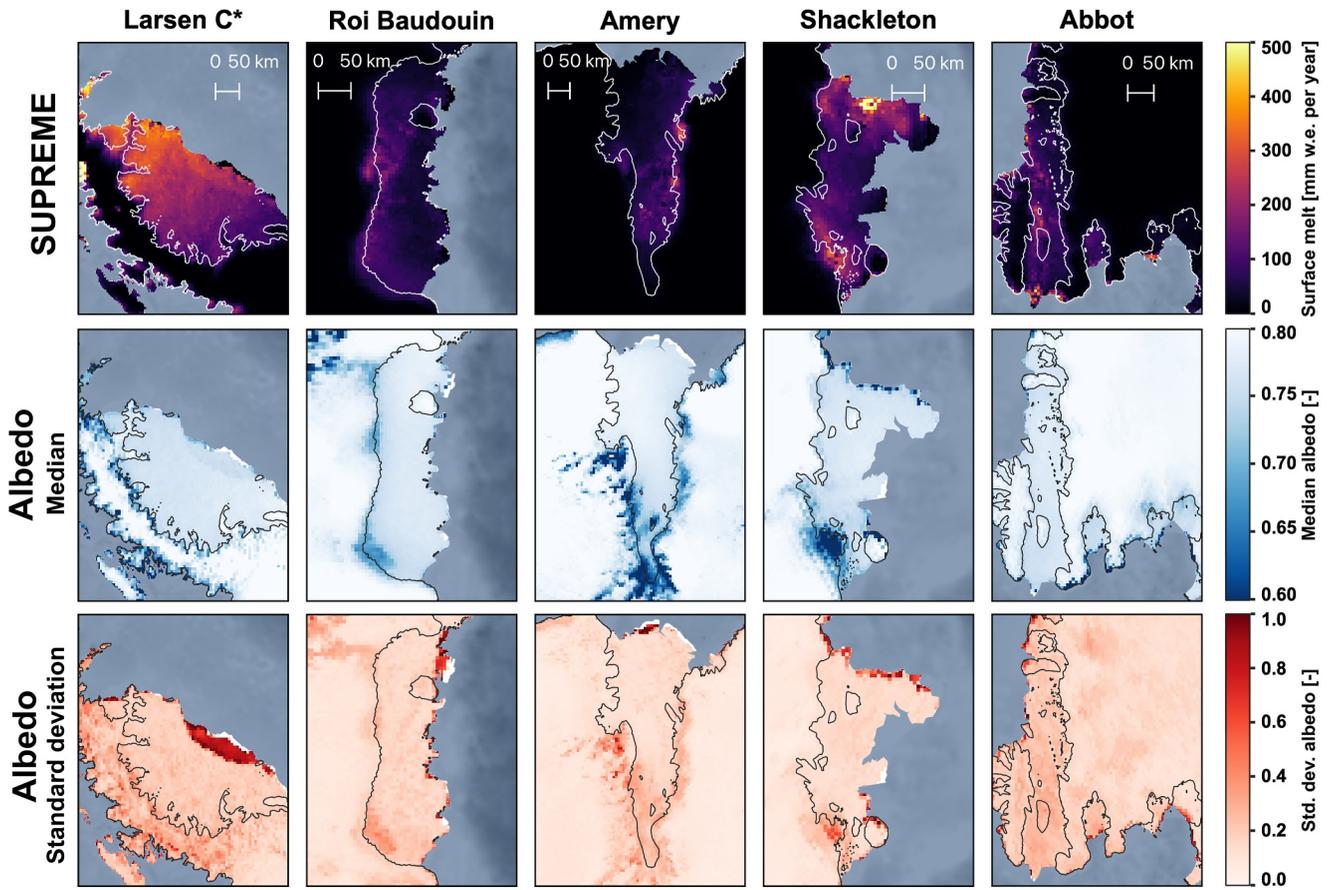


Figure A5. The mean surface melt for SUPREME (a), along with the median (b) and standard deviation (c) of albedo for 2001–2019. Caution is advised when interpreting the data sets marked with an asterisk (i.e., Larsen C), as a part of these data were used in training SUPREME.

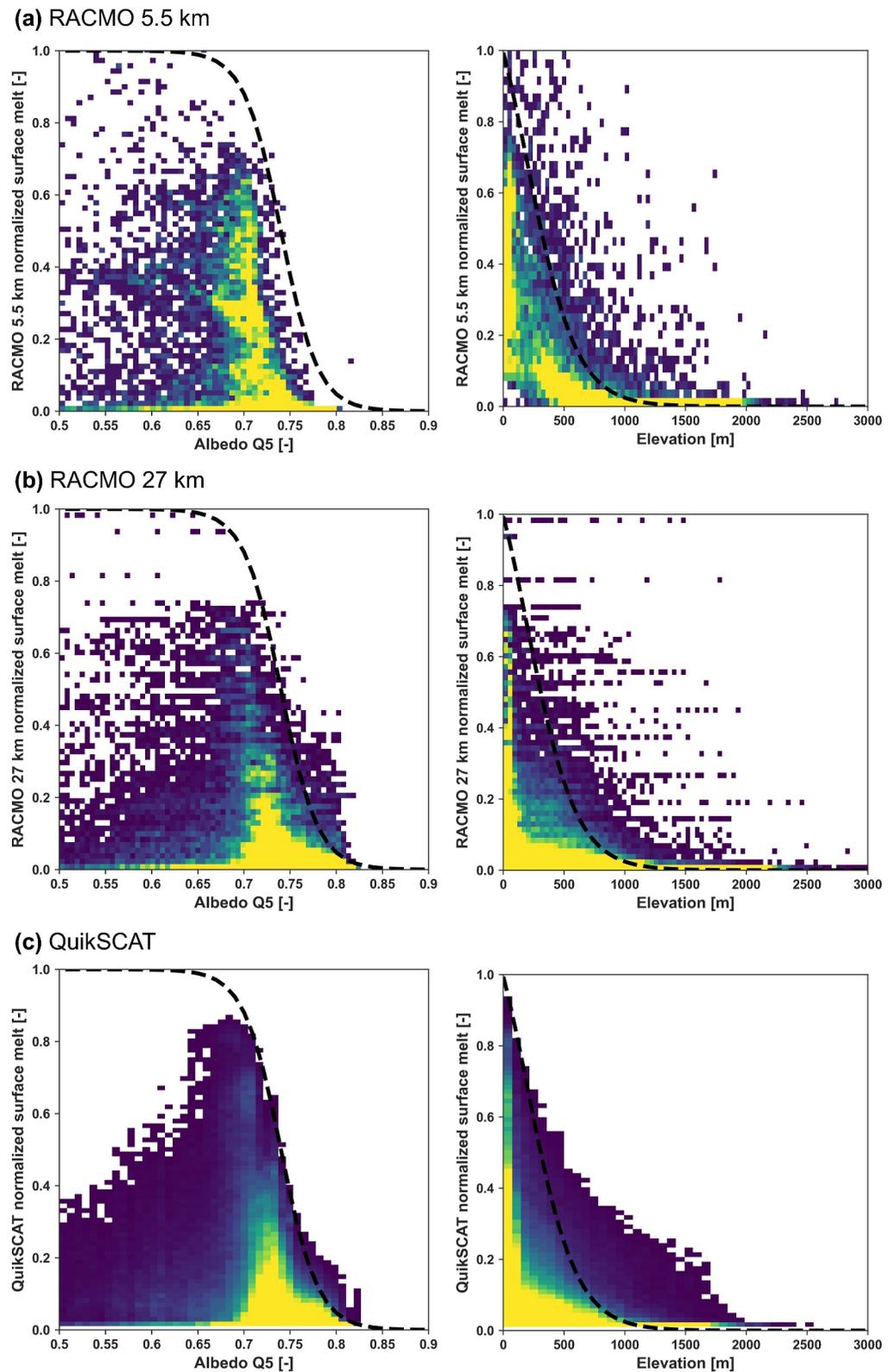


Figure A6. The two physical activation functions (albedo Q5 and elevation) implemented in the SUPREME architecture are depicted by black dashed lines. The weighted scatter plot represents normalized surface melt data for (a) RACMO 5.5 km between 2001 and 2019 on the Antarctic Peninsula, (b) RACMO 27 km between 2001 and 2019 Antarctic-wide, and (c) QuikSCAT between 2001 and 2009 Antarctic-wide.

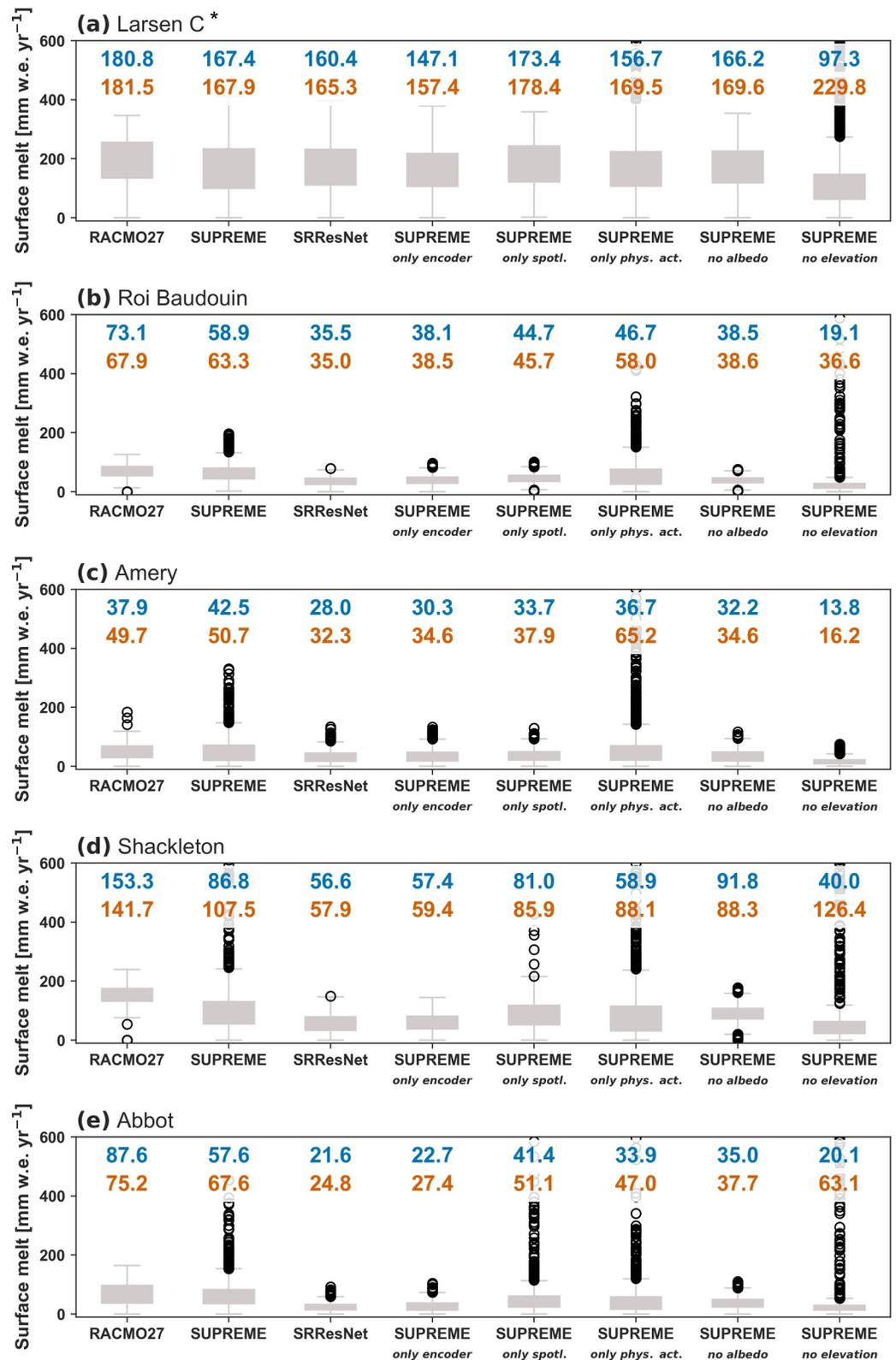


Figure A7. Boxplots of yearly average surface melt for the melt seasons of 2001–2019 at every pixel across different model architectures, with median (in blue) and mean (in orange) values in mm w.e. yr⁻¹. Corresponding spatial maps are provided in Figure 10, and the spatial extents analyzed are shown in Figure A4. Caution is advised when interpreting the data sets marked with an asterisk (i.e., Larsen C in panel (a)), as a part of these data were used in training SUPREME.

Table A1
Analysis of the Yearly Average RMSE and Bias of SUPREME Compared to RACMO 5.5 km and RACMO 27 km Within the Training Period (2001–2006) and the Subsequent Period (2007–2019)

Comparison	RMSE	Bias	R ²
(a) Data used during training (2001–2006)			
SUPREME versus RACMO 5.5 km (Antarctic Peninsula)	4.8	−1.4	0.96
SUPREME versus RACMO 27 km (Antarctic Peninsula)	4.4	−3.4	0.99
SUPREME versus RACMO 5.5 km (Larsen C)	2.6	−0.7	0.89
SUPREME versus RACMO 27 km (Larsen C)	1.6	−1.2	0.97
(b) Data not used during training (2007–2019)			
SUPREME versus RACMO 5.5 km (Antarctic Peninsula)	5.5	4.5	0.84
SUPREME versus RACMO 27 km (Antarctic Peninsula)	2.8	−2.1	0.96
SUPREME versus RACMO 5.5 km (Larsen C)	2.5	1.8	0.73
SUPREME versus RACMO 27 km (Larsen C)	1.5	−1.1	0.93

Note. RMSE and bias are in mm w.e. yr^{−1}. The spatial extents analyzed are shown in Figure A4.

Table A2
Antarctic-Wide Performance Metrics of Deep Learning Model Architectures Compared to RACMO 27 km

Model architecture	RMSE	Bias	R ²
SUPREME	32.1	6.7	0.77
SRResNet	33.0	9.1	0.79
SUPREME _{only encoder}	32.5	8.1	0.80
SUPREME _{only spotlighter}	54.5	−1.6	0.51
SUPREME _{only physical activation}	53.3	2.06	0.59
SUPREME _{no albedo}	32.1	8.4	0.80
SUPREME _{no elevation}	60.4	8.4	0.41

Note. RMSE and bias are in mm w.e. yr^{−1}.

Data Availability Statement

Antarctic-wide RACMO 27 km data at a 3-hourly resolution can be obtained from van Wessem et al. (2022). Additionally, RACMO 5.5 km data, as discussed in van Wessem et al. (2016), are accessible from Wallis et al. (2023). RACMO 2 km data from Noël et al. (2023b) can be acquired from Noël et al. (2023b). Automatic weather station data are accessible from Jakobs et al. (2020). In this research, the automatic weather station data from Jakobs et al. (2020) were updated and provided through personal communication with Maurice van Tiggelen. The passive microwave data, as discussed in Picard and Fily (2006), can be downloaded from the following link: <https://snow.univ-grenoble-alpes.fr/melting/>. The melt and slush data can be obtained from Dell et al. (2024b). All scripts and data used in this study are accessible at <https://github.com/SdeRodaHusman/SUPREME>, and the data are also archived at 4TU.ResearchData (de Roda Husman et al., 2024). For additional details regarding the code and data, please feel free to contact us via email at S.deRodaHusman@tudelft.nl.

References

- Abram, N. J., Mulvaney, R., Wolff, E. W., Triest, J., Kipfstuhl, S., Trusel, L. D., et al. (2013). Acceleration of snow melt in an Antarctic Peninsula ice core during the twentieth century. *Nature Geoscience*, 6(5), 404–411. <https://doi.org/10.1038/ngeo1787>
- Agosta, C., Amory, C., Kittel, C., Orsi, A., Favier, V., Gallée, H., et al. (2019). Estimation of the Antarctic surface mass balance using the regional climate model MAR (1979–2015) and identification of dominant processes. *The Cryosphere*, 13(1), 281–296. <https://doi.org/10.5194/tc-13-281-2019>
- Ahmed, K. F., Wang, G., Silander, J., Wilson, A. M., Allen, J. M., Horton, R., & Anyah, R. (2013). Statistical downscaling and bias correction of climate model outputs for climate change impact assessment in the US northeast. *Global and Planetary Change*, 100, 320–332. <https://doi.org/10.1016/j.gloplacha.2012.11.003>

Acknowledgments

The authors thank editor-in-chief Stephen Griffies, associate editor Nick Golledge, as well as Vincent Verjans and two anonymous reviewers for their constructive comments.

- Arthur, J. F., Stokes, C. R., Jamieson, S. S., Rachel Carr, J., Leeson, A. A., & Verjans, V. (2022). Large interannual variability in supraglacial lakes around East Antarctica. *Nature Communications*, *13*(1), 1–12. <https://doi.org/10.1038/s41467-022-29385-3>
- Banwell, A. F., & Macayeal, D. R. (2015). Ice-shelf fracture due to viscoelastic flexure stress induced by fill/drain cycles of supraglacial lakes. *Antarctic Science*, *27*(6), 587–597. <https://doi.org/10.1017/s0954102015000292>
- Banwell, A. F., Wever, N., Dunmire, D., & Picard, G. (2023). Quantifying Antarctic-wide ice-shelf surface melt volume using microwave and firm model data: 1980 to 2021. *Geophysical Research Letters*, *50*(12), e2023GL102744. <https://doi.org/10.1029/2023gl102744>
- Banwell, A. F., Willis, I. C., Macdonald, G. J., Goodsell, B., & MacAyeal, D. R. (2019). Direct measurements of ice-shelf flexure caused by surface meltwater ponding and drainage. *Nature Communications*, *10*(1), 1–10. <https://doi.org/10.1038/s41467-019-08522-5>
- Box, J. E., & Rinke, A. (2003). Evaluation of Greenland ice sheet surface climate in the HIRHAM regional climate model using automatic weather station data. *Journal of Climate*, *16*(9), 1302–1319. <https://doi.org/10.1175/1520-0442-16.9.1302>
- Braakmann-Folgmann, A., Shepherd, A., Gerrish, L., Izzard, J., & Ridout, A. (2022). Observing the disintegration of the A68A iceberg from space. *Remote Sensing of Environment*, *270*, 112855. <https://doi.org/10.1016/j.rse.2021.112855>
- Burgard, C., Jourdain, N. C., Mathiot, P., Smith, R. S., Schäfer, R., Caillet, J., et al. (2023). Emulating present and future simulations of melt rates at the base of Antarctic ice shelves with neural networks. *Journal of Advances in Modeling Earth Systems*, *15*(12). <https://doi.org/10.1029/2023MS003829>
- Cai, Y., Wan, F., Lang, S., Cui, X., & Yao, Z. (2023). Multi-branch deep neural network for bed topography of Antarctica super-resolution: Reasonable integration of multiple remote sensing data. *Remote Sensing*, *15*(5), 1359. <https://doi.org/10.3390/rs15051359>
- Das, S. B., & Alley, R. B. (2008). Rise in frequency of surface melting at Siple Dome through the Holocene: Evidence for increasing marine influence on the climate of West Antarctica. *Journal of Geophysical Research*, *113*(D2). <https://doi.org/10.1029/2007jd008790>
- Davison, B. J., Hogg, A. E., Gourmelen, N., Jakob, L., Wuite, J., Nagler, T., et al. (2023). Annual mass budget of Antarctic ice shelves from 1997 to 2021. *Science Advances*, *9*(41), eadi0186. <https://doi.org/10.1126/sciadv.adi0186>
- Dell, R., Banwell, A. F., Willis, I. C., Arnold, N. S., Halberstadt, A. R. W., Chudley, T. R., & Pritchard, H. D. (2022). Supervised classification of slush and ponded water on Antarctic ice shelves using Landsat 8 imagery. *Journal of Glaciology*, *68*(268), 401–414. <https://doi.org/10.1017/jog.2021.114>
- Dell, R., Willis, I., Arnold, N., Banwell, A., & de Roda Husman, S. (2024a). Substantial contribution of slush to meltwater area across Antarctic ice shelves. *Nature Geoscience*, *17*(7), 624–630. <https://doi.org/10.1038/s41561-024-01466-6>
- Dell, R., Willis, I., Arnold, N., Banwell, A., & de Roda Husman, S. (2024b). Research data supporting “Substantial contribution of slush to meltwater area across Antarctic ice shelves” [Dataset]. *Apollo - University of Cambridge Repository*. <https://doi.org/10.17863/CAM.108421>
- de Roda Husman, S., Hu, Z., van Tiggelen, M., Dell, R., Bolibar, J., Lhermitte, S., et al. (2024). Physically-informed super-resolution downscaling of Antarctic surface melt [Dataset]. *4TU.ResearchData*. <https://doi.org/10.4121/87607637-09d9-4147-8125-c85ecc901a88>
- de Roda Husman, S., Hu, Z., Wouters, B., Munneke, P. K., Veldhuijzen, S., & Lhermitte, S. (2022). Remote sensing of surface melt on Antarctica: Opportunities and challenges. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *16*, 2462–2480. <https://doi.org/10.1109/jstars.2022.3216953>
- de Roda Husman, S., Lhermitte, S., Bolibar, J., Izeboud, M., Hu, Z., Shukla, S., et al. (2024). A high-resolution record of surface melt on Antarctic ice shelves using multi-source remote sensing data and deep learning. *Remote Sensing of Environment*, *301*, 113950. <https://doi.org/10.1016/j.rse.2023.113950>
- Dong, C., Loy, C. C., & Tang, X. (2016). Accelerating the super-resolution convolutional neural network. In *Computer vision—eccv 2016: 14th european conference, amsterdam, The Netherlands, october 11–14, 2016, proceedings, part ii 14* (pp. 391–407).
- ECMWF, E. (2009). IFS documentation CY33r1 (Tech. Rep.). Part IV, Tech. rep., ECMWF. Retrieved from <http://www.ecmwf.int/research>
- Ettema, J., Van den Broeke, M., Van Meijgaard, E., Van de Berg, W., Box, J., & Steffen, K. (2010). Climate of the Greenland ice sheet using a high-resolution climate model—Part 1: Evaluation. *The Cryosphere*, *4*(4), 511–527. <https://doi.org/10.5194/tc-4-511-2010>
- Fettweis, X., Box, J. E., Agosta, C., Amory, C., Kittel, C., Lang, C., et al. (2017). Reconstructions of the 1900–2015 Greenland ice sheet surface mass balance using the regional climate MAR model. *The Cryosphere*, *11*(2), 1015–1033. <https://doi.org/10.5194/tc-11-1015-2017>
- Foken, T. (2021). *Springer handbook of atmospheric measurements*. Springer Nature.
- Franco, B., Fettweis, X., Lang, C., & Ericum, M. (2012). Impact of spatial resolution on the modelling of the Greenland ice sheet surface mass balance between 1990–2010, using the regional climate model MAR. *The Cryosphere*, *6*(3), 695–711. <https://doi.org/10.5194/tc-6-695-2012>
- Fyke, J., Sergienko, O., Löfverström, M., Price, S., & Lenaerts, J. T. (2018). An overview of interactions and feedbacks between ice sheets and the Earth system. *Reviews of Geophysics*, *56*(2), 361–408. <https://doi.org/10.1029/2018rg000600>
- Gallée, H., Agosta, C., Gential, L., Favier, V., & Krinner, G. (2011). A downscaling approach toward high-resolution surface mass balance over Antarctica. *Surveys in Geophysics*, *32*(4), 507–518. https://doi.org/10.1007/978-94-007-2063-3_13
- Ghilain, N., Vannitsem, S., Dalaiden, Q., Goosse, H., De Cruz, L., & Wei, W. (2022). Large ensemble of downscaled historical daily snowfall from an earth system model to 5.5 km resolution over Dronning Maud Land, Antarctica. *Earth System Science Data*, *14*(4), 1901–1916. <https://doi.org/10.5194/essd-14-1901-2022>
- Giesen, R., & Oerlemans, J. (2012). Calibration of a surface mass balance model for global-scale applications. *The Cryosphere*, *6*(6), 1463–1481. <https://doi.org/10.5194/tc-6-1463-2012>
- Gilbert, E., & Kittel, C. (2021). Surface melt and runoff on Antarctic ice shelves at 1.5 C, 2 C, and 4 C of future warming. *Geophysical Research Letters*, *48*(8), e2020GL091733. <https://doi.org/10.1029/2020gl091733>
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, *27*.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Greene, C. A., Gardner, A. S., Schlegel, N.-J., & Fraser, A. D. (2022). Antarctic calving loss rivals ice-shelf thinning. *Nature*, *609*(7929), 948–953. <https://doi.org/10.1038/s41586-022-05037-w>
- Gudmundsson, G. H., Paolo, F. S., Adusumilli, S., & Fricker, H. A. (2019). Instantaneous Antarctic ice sheet mass loss driven by thinning ice shelves. *Geophysical Research Letters*, *46*(23), 13903–13909. <https://doi.org/10.1029/2019gl085027>
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, *51*(5), 1–42. <https://doi.org/10.1145/3236009>
- Hanna, E., Huybrechts, P., Cappelen, J., Steffen, K., Bales, R. C., Burgess, E., et al. (2011). Greenland Ice Sheet surface mass balance 1870 to 2010 based on Twentieth Century Reanalysis, and links with global climate forcing. *Journal of Geophysical Research*, *116*(D24). <https://doi.org/10.1029/2011jd016387>
- Hanna, E., Huybrechts, P., Janssens, I., Cappelen, J., Steffen, K., & Stephens, A. (2005). Runoff and mass balance of the Greenland ice sheet: 1958–2003. *Journal of Geophysical Research*, *110*(D13). <https://doi.org/10.1029/2004jd005641>

- Hanna, E., Huybrechts, P., Steffen, K., Cappelen, J., Huff, R., Shuman, C., et al. (2008). Increased runoff from melt from the Greenland ice sheet: A response to global warming. *Journal of Climate*, *21*(2), 331–341. <https://doi.org/10.1175/2007jcli1964.1>
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778).
- Helm, V., Humbert, A., & Miller, H. (2014). Elevation and elevation change of Greenland and Antarctica derived from CryoSat-2. *The Cryosphere*, *8*(4), 1539–1559. <https://doi.org/10.5194/tc-8-1539-2014>
- Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., et al. (2017). The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, *98*(3), 589–602. <https://doi.org/10.1175/bams-d-15-00135.1>
- Howat, I. M., Porter, C., Smith, B. E., Noh, M.-J., & Morin, P. (2019). The Reference Elevation Model of Antarctica. *The Cryosphere*, *13*(2), 665–674. <https://doi.org/10.5194/tc-13-665-2019>
- Hu, Z., Sun, Y., Kuipers Munneke, P., Lhermitte, S., & Zhu, X. (2022). Towards a spatially transferable super-resolution model for downscaling Antarctic surface melt. In *NeurIPS 2022 workshop on tackling climate change with machine learning*. Retrieved from <https://www.climatechange.ai/papers/neurips2022/32>
- IMBIE. (2018). Mass balance of the Antarctic ice sheet from 1992 to 2017. *Nature*, *558*(7709), 219–222. <https://doi.org/10.1038/s41586-018-0179-y>
- Jakobs, C. L., Reijmer, C. H., Kuipers Munneke, P., König-Langlo, G., & Van Den Broeke, M. R. (2019). Quantifying the snowmelt–albedo feedback at Neumayer station, east Antarctica. *The Cryosphere*, *13*(5), 1473–1485. <https://doi.org/10.5194/tc-13-1473-2019>
- Jakobs, C. L., Reijmer, C. H., Smeets, C. P., Trusel, L. D., Van De Berg, W. J., Van Den Broeke, M. R., & Van Wessem, J. M. (2020). A benchmark dataset of in situ Antarctic surface melt rates and energy balance. *Journal of Glaciology*, *66*(256), 291–302. <https://doi.org/10.1017/jog.2020.6>
- Jiang, J., Shu, Y., Wang, J., & Long, M. (2022). Transferability in deep learning: A survey. *arXiv preprint arXiv:2201.05867*.
- Kawulok, M., Benecki, P., Piechaczek, S., Hryneczenko, K., Kostrzewa, D., & Nalepa, J. (2019). Deep learning for multiple-image super-resolution. *IEEE Geoscience and Remote Sensing Letters*, *17*(6), 1062–1066. <https://doi.org/10.1109/lgrs.2019.2940483>
- Kingslake, J., Ely, J. C., Das, I., & Bell, R. E. (2017). Widespread movement of meltwater onto and across Antarctic ice shelves. *Nature*, *544*(7650), 349–352. <https://doi.org/10.1038/nature22049>
- Kuipers Munneke, P., Ligtenberg, S. R., Van Den Broeke, M. R., & Vaughan, D. G. (2014). Firm air depletion as a precursor of Antarctic ice-shelf collapse. *Journal of Glaciology*, *60*(220), 205–214. <https://doi.org/10.3189/2014jog13j183>
- Kuipers Munneke, P., Luckman, A., Bevan, S., Smeets, C., Gilbert, E., Van den Broeke, M., et al. (2018). Intense winter surface melt on an Antarctic ice shelf. *Geophysical Research Letters*, *45*(15), 7615–7623. <https://doi.org/10.1029/2018gl077899>
- Kuipers Munneke, P., Picard, G., Van Den Broeke, M., Lenaerts, J., & Van Meijgaard, E. (2012). Insignificant change in Antarctic snowmelt volume since 1979. *Geophysical Research Letters*, *39*(1). <https://doi.org/10.1029/2011gl0150207>
- Kuipers Munneke, P., Van den Broeke, M., King, J., Gray, T., & Reijmer, C. (2012). Near-surface climate and surface energy budget of Larsen C ice shelf, Antarctic Peninsula. *The Cryosphere*, *6*(2), 353–363. <https://doi.org/10.5194/tc-6-353-2012>
- Kuipers Munneke, P., Van den Broeke, M., Lenaerts, J., Flanner, M., Gardner, A., & Van de Berg, W. (2011). A new albedo parameterization for use in climate models over the Antarctic Ice Sheet. *Journal of Geophysical Research*, *116*(D5), D05114. <https://doi.org/10.1029/2010jd015113>
- Lachlan-Cope, T. (2010). Antarctic clouds. *Polar Research*, *29*(2), 150–158. <https://doi.org/10.1111/j.1751-8369.2010.00148.x>
- La Rocca, M., & Perna, C. (2022). Opening the black box: Bootstrapping sensitivity measures in neural networks for interpretable machine learning. *Stats*, *5*(2), 440–457. <https://doi.org/10.3390/stats502026>
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., et al. (2017). Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4681–4690).
- Lenaerts, J., Lhermitte, S., Drews, R., Ligtenberg, S., Berger, S., Helm, V., et al. (2017). Meltwater produced by wind–albedo interaction stored in an East Antarctic ice shelf. *Nature Climate Change*, *7*(1), 58–62. <https://doi.org/10.1038/nclimate3180>
- Lenaerts, J., Van den Broeke, M., Déry, S., Van Meijgaard, E., Van de Berg, W., Palm, S. P., & Sanz Rodrigo, J. (2012). Modeling drifting snow in Antarctica with a regional climate model: 1. Methods and model evaluation. *Journal of Geophysical Research*, *117*(D5). <https://doi.org/10.1029/2011jd016145>
- Leong, W. J., & Horgan, H. J. (2020). DeepBedMap: A deep neural network for resolving the bed topography of Antarctica. *The Cryosphere*, *14*(11), 3687–3705. <https://doi.org/10.5194/tc-14-3687-2020>
- Li, W., Ni, L., Li, Z.-L., Duan, S.-B., & Wu, H. (2019). Evaluation of machine learning algorithms in spatial downscaling of MODIS land surface temperature. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *12*(7), 2299–2307. <https://doi.org/10.1109/jstars.2019.2896923>
- Nihashi, S., & Ohshima, K. I. (2015). Circumpolar mapping of Antarctic coastal polynyas and landfast sea ice: Relationship and variability. *Journal of Climate*, *28*(9), 3650–3670. <https://doi.org/10.1175/jcli-d-14-00369.1>
- Noël, B., van de Berg, W. J., Machguth, H., Lhermitte, S., Howat, I., Fettweis, X., & Van Den Broeke, M. R. (2016). A daily, 1 km resolution data set of downscaled Greenland ice sheet surface mass balance (1958–2015). *The Cryosphere*, *10*(5), 2361–2377. <https://doi.org/10.5194/tc-10-2361-2016>
- Noël, B., van Wessem, J. M., Wouters, B., Trusel, L., Lhermitte, S., & van den Broeke, M. R. (2023a). Higher Antarctic ice sheet accumulation and surface melt rates revealed at 2 km resolution. *Nature Communications*, *14*(1), 7949. <https://doi.org/10.1038/s41467-023-43584-6>
- Noël, B., van Wessem, M., Wouters, B., Trusel, L., Lhermitte, S., & van den Broeke, M. (2023b). Data set: “Higher Antarctic ice sheet accumulation and surface melt rates revealed at 2 km resolution”. *Zenodo*. <https://doi.org/10.5281/zenodo.10007855>
- Petrou, Z. I., Xian, Y., & Tian, Y. (2018). Towards breaking the spatial resolution barriers: An optical flow and super-resolution approach for sea ice motion estimation. *ISPRS Journal of Photogrammetry and Remote Sensing*, *138*, 164–175. <https://doi.org/10.1016/j.isprsjrs.2018.01.020>
- Picard, G., & Fily, M. (2006). Surface melting observations in Antarctica by microwave radiometers: Correcting 26-year time series from changes in acquisition hours. *Remote Sensing of Environment*, *104*(3), 325–336. <https://doi.org/10.1016/j.rse.2006.05.010>
- Reddy, S., Reddy, K. T., & Vallikumari, V. (2018). Optimization of deep learning using various optimizers, loss functions and dropout. *International Journal of Recent Technology and Engineering*, *7*, 448–455.
- Savage, N. (2022). Breaking into the black box of artificial intelligence. *Nature*. <https://doi.org/10.1038/d41586-022-00858-1>
- Scambos, T., Fricker, H. A., Liu, C.-C., Bohlander, J., Fastook, J., Sargent, A., et al. (2009). Ice shelf disintegration by plate bending and hydrofracture: Satellite observations and model results of the 2008 Wilkins ice shelf break-ups. *Earth and Planetary Science Letters*, *280*(1–4), 51–60. <https://doi.org/10.1016/j.epsl.2008.12.027>
- Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., et al. (2002). First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*(1–2), 135–148. [https://doi.org/10.1016/s0034-4257\(02\)00091-3](https://doi.org/10.1016/s0034-4257(02)00091-3)

- Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J.-C., Hannay, C., et al. (2017). Practice and philosophy of climate model tuning across six US modeling centers. *Geoscientific Model Development*, *10*(9), 3207–3223. <https://doi.org/10.5194/gmd-10-3207-2017>
- Scholz, F. W., & Stephens, M. A. (1987). K-Sample Anderson–Darling tests. *Journal of the American Statistical Association*, *82*(399), 918–924. <https://doi.org/10.2307/2288805>
- Shepherd, A., Ivins, E. R., Barletta, V. R., Bentley, M. J., Bettadpur, S., Briggs, K. H., et al. (2012). A reconciled estimate of ice-sheet mass balance. *Science*, *338*(6111), 1183–1189. <https://doi.org/10.1126/science.1228102>
- Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., et al. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1874–1883).
- Smeets, P. C. J. P., Kuipers Munneke, P., van As, D., van den Broeke, M. R., Boot, W., Oerlemans, H., et al. (2018). The K-transect in west Greenland: Automatic weather station data (1993–2016). *Arctic Antarctic and Alpine Research*, *50*(1), S100002. <https://doi.org/10.1080/15230430.2017.1420954>
- Stokes, C. R., Sanderson, J. E., Miles, B. W., Jamieson, S. S., & Leeson, A. A. (2019). Widespread distribution of supraglacial lakes around the margin of the East Antarctic Ice Sheet. *Scientific Reports*, *9*(1), 13823. <https://doi.org/10.1038/s41598-019-50343-5>
- Tedesco, M., Colosio, P., Fettweis, X., & Cervone, G. (2023). A computationally efficient statistically downscaled 100 m resolution Greenland product from the regional climate model MAR. *The Cryosphere Discussions*, *2023*(12), 1–27. <https://doi.org/10.5194/tc-17-5061-2023>
- Thiemeal, M. J., Gobiet, A., & Heinrich, G. (2012). Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change*, *112*(2), 449–468. <https://doi.org/10.1007/s10584-011-0224-4>
- Tollenaar, V., Zekollari, H., Lhermitte, S., Tax, D. M., Debaille, V., Goderis, S., et al. (2022). Unexplored Antarctic meteorite collection sites revealed through machine learning. *Science Advances*, *8*(4), eabj8138. <https://doi.org/10.1126/sciadv.abj8138>
- Torinesi, O., Fily, M., & Genthon, C. (2003). Variability and trends of the summer melt period of Antarctic ice margins since 1980 from microwave sensors. *Journal of Climate*, *16*(7), 1047–1060. [https://doi.org/10.1175/1520-0442\(2003\)016<1047:vatotms>2.0.co;2](https://doi.org/10.1175/1520-0442(2003)016<1047:vatotms>2.0.co;2)
- Trusel, L., Frey, K., Das, S., Munneke, P. K., & Van Den Broeke, M. (2013). Satellite-based estimates of Antarctic surface meltwater fluxes. *Geophysical Research Letters*, *40*(23), 6148–6153. <https://doi.org/10.1002/2013gl058138>
- Trusel, L., Frey, K. E., & Das, S. B. (2012). Antarctic surface melting dynamics: Enhanced perspectives from radar scatterometer data. *Journal of Geophysical Research*, *117*(F2). <https://doi.org/10.1029/2011Jf002126>
- Trusel, L., Frey, K. E., Das, S. B., Karnauskas, K. B., Munneke, P. K., Van Meijgaard, E., & Van Den Broeke, M. R. (2015). Divergent trajectories of Antarctic surface melt under two twenty-first-century climate scenarios. *Nature Geoscience*, *8*(12), 927–932. <https://doi.org/10.1038/ngeo2563>
- Undén, P., Rontu, L., Jarvinen, H., Lynch, P., Calvo Sánchez, F. J., Cats, G., et al. (2002). HIRLAM-5 scientific documentation.
- Vandal, T., Kodra, E., Ganguly, S., Michaelis, A., Nemani, R., & Ganguly, A. R. (2017). DeepSD: Generating high resolution climate change projections through single image super-resolution. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1663–1672).
- van de Berg, W. J., van Meijgaard, E., & van Ulft, L. H. (2020). The added value of high resolution in estimating the surface mass balance in southern Greenland. *The Cryosphere*, *14*(6), 1809–1827. <https://doi.org/10.5194/tc-14-1809-2020>
- van der Meer, M., de Roda Husman, S., & Lhermitte, S. (2023). Deep learning regional climate model emulators: A comparison of two downscaling training frameworks. *Journal of Advances in Modeling Earth Systems*, *15*(6), e2022MS003593. <https://doi.org/10.1029/2022ms003593>
- van Wessem, J., Ligtenberg, S., Reijmer, C., Van De Berg, W., Van Den Broeke, M., Barrand, N., et al. (2016). The modelled surface mass balance of the Antarctic Peninsula at 5.5 km horizontal resolution. *The Cryosphere*, *10*(1), 271–285. <https://doi.org/10.5194/tc-10-271-2016>
- van Wessem, J., Reijmer, C., Lenaerts, J., Van de Berg, W., Van den Broeke, M., & Van Meijgaard, E. (2014). Updated cloud physics in a regional atmospheric climate model improves the modelled surface energy balance of Antarctica. *The Cryosphere*, *8*(1), 125–135. <https://doi.org/10.5194/tc-8-125-2014>
- van Wessem, J., Reijmer, C., Morlighem, M., Mougnot, J., Rignot, E., Medley, B., et al. (2014). Improved representation of East Antarctic surface mass balance in a regional atmospheric climate model. *Journal of Glaciology*, *60*(222), 761–770. <https://doi.org/10.3189/2014jog14j051>
- van Wessem, J. M., Van De Berg, W. J., Noël, B. P., Van Meijgaard, E., Amory, C., Birnbaum, G., et al. (2018). Modelling the climate and surface mass balance of polar ice sheets using RACMO2—Part 2: Antarctica (1979–2016). *The Cryosphere*, *12*(4), 1479–1498. <https://doi.org/10.5194/tc-12-1479-2018>
- van Wessem, J. M., van den Broeke, M. R., & Lhermitte, S. (2022). Data set: Yearly RACMO2.3p2 variables [Dataset]. *Zenodo*. <https://doi.org/10.5281/zenodo.6602723>
- Wallis, B., Hogg, A., van Wessem, J. M., & van den Broeke, M. (2023). Data for widespread seasonal speed-up of west Antarctic Peninsula glaciers from 2014–2021 (version 2) [Dataset]. *Zenodo*. <https://doi.org/10.5281/zenodo.7521416>
- Wessel, B., Huber, M., Wohlfart, C., Bertram, A., Osterkamp, N., Marschalk, U., et al. (2021). TanDEM-X PolarDEM 90 m of Antarctica: Generation and error characterization. *The Cryosphere*, *15*(11), 5241–5260. <https://doi.org/10.5194/tc-15-5241-2021>
- Wiesenkcker, J. M., Kuipers Munneke, P., Van den Broeke, M. R., & Smeets, C. P. (2018). A multidecadal analysis of Föhn winds over Larsen C ice shelf from a combination of observations and modeling. *Atmosphere*, *9*(5), 172. <https://doi.org/10.3390/atmos9050172>
- Yang, C.-Y., Ma, C., & Yang, M.-H. (2014). Single-image super-resolution: A benchmark. In *Computer vision—eccv 2014: 13th european conference, zürich, Switzerland, september 6–12, 2014, proceedings, part iv 13* (pp. 372–386).
- Zheng, L., Cheng, X., Shang, X., Chen, Z., Liang, Q., & Wang, K. (2022). Greenland Ice Sheet daily surface melt flux observed from space. *Geophysical Research Letters*, *49*(6), e2021GL096690. <https://doi.org/10.1029/2021gl096690>