Statistically parameterizing and evaluating a positive degree-day model to estimate surface melt in Antarctica from 1979 to 2022

Yaowen Zheng[1], Nicholas R. Golledge[1], Alexandra Gossart[1], Ghislain Picard[2], and Marion Leduc-Leballeur[3]

[1]Antarctic Research Centre, Victoria University of Wellington, Wellington, New Zealand
[2]Univ. Grenoble Alpes, CNRS, Institut des Géosciences de l'Environnement (IGE), UMR 5001, Grenoble, France
[3]Institute of Applied Physics "Nello Carrara", National Research Council, 50019 Sesto Fiorentino, Italy

Correspondence: Yaowen Zheng (yaowen.zheng@vuw.ac.nz)

Abstract.

Surface melt is one of the primary drivers of ice shelf collapse in Antarctica. Surface melting is expected to increase in the future as the global climate continues to warm, because there is a statistically significant positive relationship between air temperature and melt. Enhanced surface melt will impact the mass balance of the Antarctic Ice Sheet (AIS) and, through dynamic feedbacks, induce changes in global mean sea level (GMSL). However, current understanding of surface melt in Antarctica remains limited in past, present and future contexts. Here, we construct a novel cell-level positive degree-day (PDD) model, force it only with 2-m air temperature reanalysis data, and parameterize it spatially by minimizing the error with respect to satellite estimates and SEB model outputs on each computing cell over the period 1979 to 2022. We evaluate the PDD model by performing a goodness-of-fit test and cross-validation. We assess the accuracy of our parameterization method, based on the performance of the PDD model when considering all computing cells as a whole, independently of to the time window chosen for parameterization. We conduct sensitivity experiments by adding $\pm10\%$ to the training data (satellite estimates and SEB model outputs) used for PDD parameterization. We find that the PDD estimates change analogously to the variations in the training data with steady statistically significant correlations, suggesting the applicability of the PDD model to warmer and colder climate scenarios. Within the limitations discussed, we suggest that an appropriately parameterized PDD model can be a valuable tool for exploring Antarctic surface melt beyond the satellite era.

1 Introduction

Surface melting is common and well-studied over the Greenland Ice Sheet (GrIS) (e.g. Mernild et al., 2011; Colosio et al., 2021; Sellevold and Vizcaino, 2021), and is known to play an important role in the net mass balance of the ice sheet and changes in global mean sea level (GMSL), both now and in the past (e.g. Ryan et al., 2019). It is likely to become even more important in the future. Antarctica is currently much colder than Greenland. Antarctic ice shelves show no statistically significant trend for the annual melt days (Johnson et al., 2022) and also no significant increase in melt amount in East Antarctica in the past 40 years (Stokes et al., 2022). However, climate projections have suggested that surface melt will increase in the current century (e.g. Trusel et al., 2015; Kittel et al., 2021; Stokes et al., 2022) – both in terms of area and volume of melting (Trusel et al., 2015; Lee et al., 2017). Studies have suggested that Antarctic surface melt can impact ice sheet mass balance through surface thinning and runoff, surface meltwater draining to the bed, and increasing ice shelf vulnerability (Bell et al., 2018; Stokes et al., 2022). However, these are currently less understood over Antarctica than Greenland, either in the past or at present. This is concerning as surface melting will likely become an increasingly important player to Antarctic environment through this century and the next.

Although the warming taking place over the Antarctic Peninsula has not been consistent over the past two decades (Turner et al., 2016), the global mean surface temperature is predicted to increase (Meinshausen et al., 2011). Moreover, the positive feedback of albedo, in which the absorption of shortwave radiation increases when snow melts to water, amplifies this melting (Lenaerts et al., 2017). However, recent studies have found large inter-annual variability of surface melt in Antarctica with no statistically significant trend (Kuipers Munneke et al., 2012; Johnson et al., 2022). Projecting Antarctic surface melt is therefore still a challenge, partly because of uncertainties introduced by clouds (Kittel et al., 2022), atmospheric rivers (e.g. Clem et al., 2022), or other localized climate phenomena.

Continental-scale spaceborne observations of surface melt are limited to the satellite era (1979–present), meaning that current estimates of Antarctic surface melt are typically derived from surface energy balance (SEB) or positive degree-day (PDD) models. SEB models require diverse and detailed input data that are not always available and require considerable computational resources. The PDD model, by comparison, has fewer input and computational requirements and is therefor suited for exploring surface melt scenarios in the past and future. PDD models calculate surface melt based on the temperature-melt relationship (Hock, 2005). A typical PDD model has two parameters: (1) the threshold temperature (T_0), which controls the decision of melt or no-melt, and (2) the degree-day factor (DDF), which controls the amount of melt.

Although PDD models are empirical, they are often sufficient for estimating melt on a catchment scale (Hock, 2003, 2005) because of their two physical bases: (a) the majority of the heat required for snow and ice melt is primarily a function of near-surface air temperature, and (b) the near-surface air temperature is correlated with the longwave atmospheric radiation, shortwave radiation and sensible heat fluxes (Ohmura, 2001). Wake and Marshall (2015) suggest that Antarctic surface melt can be estimated solely from monthly temperature.

However, as the DDF is related to all terms of the surface energy balance (SEB) (Hock, 2005), a robust PDD model needs to incorporate DDFs that vary spatially and temporally (e.g. Hock, 2003, 2005; van den Broeke et al., 2010), not simply a uniform value that covers a wide region. This is because of the variability of energy partitioning, which is affected by the different climate, seasons and surfaces (Hock, 2003). Topographic influences, such as the gradient of elevation which affects albedo and direct input solar radiation (Hock, 2003), are generally strongest in mountainous terrain, together with seasonal variations in radiation, and can introduce spatial and temporal variabilities of DDF, respectively (Hock, 2005). Spatial and temporal parameterisation of DDF (model calibration), as well as model verification, therefore need to be considered.

Although PDD schemes have been used in many Antarctic numerical ice sheet models (e.g. Winkelmann et al., 2011; Larour et al., 2012) as empirical approximations to compute the ice ablation for the computation of surface mass balance, and in several studies for exploring surface melt in Antarctica, particularly in the Antarctic Peninsula (e.g. Golledge et al., 2010; Barrand et al., 2013; Costi et al., 2018), the spatial variability of PDD parameters are rarely considered. Moreover, compared to PDD model approaches developed(e.g. Reeh, 1991; Braithwaite, 1995) and improved (Fausto et al., 2011; Jowett et al., 2015; Wilton et al., 2017) for Greenland over many decades, such assessments for the PDD approach for the Antarctic domain are limited and a spatially parameterized Antarctic PDD model has not yet been achieved.

In this study, we focus on constructing a computationally efficient cell-level (spatially variable) PDD model to estimate surface melt in Antarctica through the past four decades, by statistically optimizing the parameters of the PDD model individually in each computing cell. We use the European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ECMWF ERA5) (Hersbach et al., 2018a, b) 2-m air temperature as input and compare the simulated presence of melt to satellite estimates of melt days from three satellite products and the Regional Atmospheric Climate Model version 2.3p2 (RACMO2.3p2) surface melt amount simulations. We then examine the distributions of melt days and melt amount from PDD experiments that use varying model parameters against satellite-based and RACMO2.3p2 estimations. Following this, we perform a 3-fold cross validation, together with sensitivity experiments, to evaluate our parameterization method and the PDD model.

2 Data

2.1 Reanalysis data

Table 1. Table of data that we use in this study.

Data type	Time period	Spatial resolution	Temporal resolution	Reference
ERA5 reanalysis data[a]	1979–2021	0.25°× 0.25 ° lon/lat	Hourly	Hersbach et al. (2018b)
Zwally Antarctic drainage basin	–	1000 m	–	Zwally et al. (2012)
Satellite SMMR and SSM/I[b]	1979–2021	25×25 km^2	Daily	Picard and Fily (2006)
Satellite AMSR-E[c]	2002–2011	12.5×12.5 km^2	Daily	Picard et al. (2007)
Satellite AMSR-2[c]	2012–2021	12.5×12.5 km^2	Daily	This study
RACMO2.3p2[d]	1979–2021	27×27 km^2	Monthly	Van Wessem et al. (2018)

[a] The 2-m air temperature data are on single level (Hersbach et al., 2018b). [b] Satellite local acquisition times over Antarctica are around 6 am and 6 pm. [c] Satellite local acquisition times over Antarctica are around 12 am (descending) and 12 pm (ascending). [d] RACMO2.3p2 surface melt simulations.

The dataset we use in this study is the ECMWF ERA5 reanalysis (Hersbach et al., 2018b) (Table 1). It has hourly data for three-dimensional (pressure level) atmospheric fields (Hersbach et al., 2018a) and on a single level for atmosphere and land-surface (Hersbach et al., 2018b). It replaced the previous ECMWF reanalysis product ERA-Interim in 2019 (Hersbach et al., 2020), and has become the new state-of-the-art ECMWF reanalysis product for global and Antarctic weather and climate (Hersbach et al., 2020; Gossart et al., 2019).

The particular ERA5 product we use in this study is the hourly 2-m air temperature data which has been evaluated and used previously for studies in Antarctica (e.g. Gossart et al., 2019; Tetzner et al., 2019; Zhu et al., 2021). Assessments have shown that ERA5 near-surface (or 2-m) air temperature data is a robust tool for exploring Antarctic climate (e.g. Gossart et al., 2019;

Zhu et al., 2021). ERA5 performs better at representing near-surface temperature than its predecessors, the Climate Forecast System Reanalysis (CFSR), and the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) (Gossart et al., 2019). It is continuously being updated and is one of the most state-of-the-art reanalysis datasets available. However, compared to 48 automatic weather station (AWS) observations, it is reported to have a cold bias over the entire continent apart from the winter months (June-July-August) (Zhu et al., 2021). This cold bias is reported at 0.34 $^{\circ}$C annually and at 1.06 $^{\circ}$C during December-January-February (DJF) (Zhu et al., 2021).

2.2 Satellite data

The number of melt days retrieved from the satellite observations is used to parameterize the threshold temperature (T_0) for the PDD model. We use the satellite 42-year daily (once every two days before 1988) Antarctic surface melt dataset produced by Picard and Fily (2006) (Table 1). It contains daily estimates as a binary of melt or no-melt on a 25×25 km southern polar stereographic grid. It is obtained by applying the melt detecting algorithm (Torinesi et al., 2003; Picard and Fily, 2006) on the scanning Multichannel Microwave Radiometer (SMMR) and three Special Sensor Microwave Imager (SSM/I) observed passive-microwave data from the National Snow and Ice Data Center (NSIDC) (Picard and Fily, 2006). SMMR and SSM/I sensors are carried by sun-synchronous orbit satellites observing Earth at least twice per day (Picard and Fily, 2006). For Antarctica, the local acquisition times are around 6 am and 6 pm. The brightness temperature is the daily average of all the passes (those around 6 am and those around 6 pm). There is a reported data gap longer than a month during the period from December 1987 to January 1988 (Torinesi et al., 2003; Johnson et al., 2022), and we find additional missing data during the prolonged summer (from November to March) in 1986/1987 (13 days), 1987/1988 (44 days), 1988/1989 (8 days) and 1991/1992 (9 days), which are significantly longer than the length of the missing data period of the remaining 38 years (zero or one day, Figure A1 in the Appendix A). We therefore omit those periods from our comparison to the satellite estimates.

We also use a more recently developed satellite melt day dataset which uses a similar algorithm as Torinesi et al. (2003); Picard and Fily (2006) on the Advanced Microwave Scanning Radiometer for EOS (AMSR-E) and the Advanced Microwave Scanning Radiometer 2 (AMSR-2) observed passive-microwave data from the Japan Aerospace Exploration Agency (JAXA, Table 1). This dataset is on a 12.5×12.5 km^2 southern polar stereographic grid. It has twice-daily observations over Antarctica covering 2002 to 2011 (AMSR-E) and 2012 to 2021 (AMSR-2, Table 1). These sensors have a local acquisition time over Antarctica of around 12 am (descending) and 12 pm (ascending).

2.3 Regional climate model SEB output

SEB modeling is a physics-based numerical approach used to calculate the surface energy budget in order to estimate how much energy is available for snow/ice melting. A number of studies have used SEB modeling forced by climate model outputs and AWS data to assess surface melting on GrIS and AIS (e.g. Van den Broeke et al., 2011; Zou et al., 2021). To parameterize the DDF for the PDD model, we compare our ERA5 forced numerical experiments to the Antarctic surface melt simulations from the RACMO2.3p2 (Van Wessem et al., 2018). The RACMO2.3p2 simulates Antarctic surface melt by solving the SEB model which is defined as (Van Wessem et al., 2018):

$$Q_M = SW_{\downarrow} + SW_{\uparrow} + LW_{\downarrow} + LW_{\uparrow} + SHF + LHF + G_s \tag{1}$$

where Q_M is the energy available for melting, $SW_{\downarrow}$ and $SW_{\uparrow}$ are the downward and upward shortwave radiative fluxes, $LW_{\downarrow}$ and $LW_{\uparrow}$ are the downward and upward longwave radiative fluxes, SHF and LHF are the sensible and latent turbulent heat fluxes and G_s is the subsurface conductive heat flux (Van Wessem et al., 2018).

The RACMO2.3p2 Antarctic surface melt simulations used here cover the time period from January 1979 to February 2021 with monthly temporal resolution and 27×27 km spatial resolution (Table 1).

2.4 Interpolation and research domain

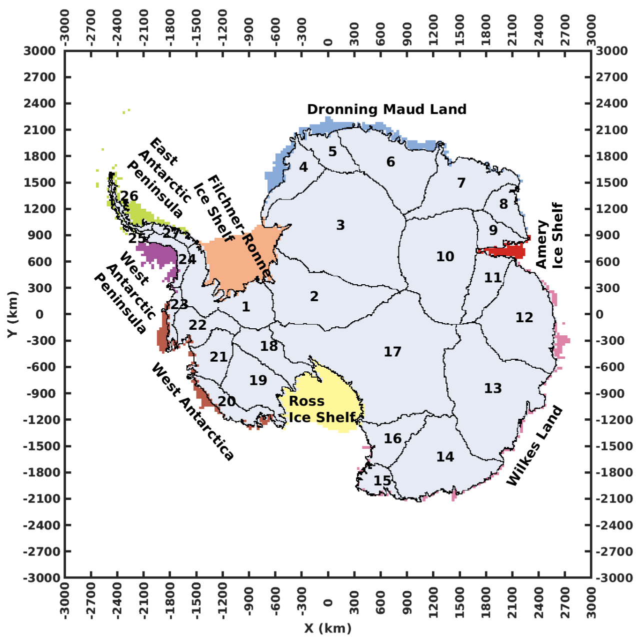

Figure 1. Map of the research domain and 27 Antarctic drainage basins (Zwally et al., 2012) used in this study.

The spatially coarsest dataset used in this study is the ERA5 reanalysis data which is in 0.25° longitude × 0.25° latitude geographic coordinates (Table 1). For consistency with the other data we analyse, we use the southern polar stereographic coordinates instead of the geographic coordinates. We use the Climate Data Operators (CDO) (Schulzweida, 2021) to bilinearly remap ERA5 reanalysis data from longitude-latitude geographic coordinates to NSIDC Sea Ice Polar Stereographic South Projected Coordinate System (NSIDC, 2022) (hereafter "polar stereographic grid"). We use a spatial resolution of 30 km, minimising the number of missing pixels and maximising the resolution. For consistency, we also use CDO to remap all data products used in this study (Table 1) to the same 30×30 km polar stereographic grid. The research domain is shown in Figure 1.

3 Methods

3.1 PDD model

Using an empirical relationship between air temperature and melt, temperature-index models are the most used method for assessing surface melt of ice and snow due to their simplicity as they are only meteorologically forced by the air temperature (Hock, 2005). Not only does the simplicity of the approach enable fast run times and require low computational resources, but the air temperature input data are also much easier to obtain than the full inputs (e.g. radiation fluxes, temperature, wind speed, humidity, ice/ snow density and surface roughness (van den Broeke et al., 2010)) required by the SEB model. If appropriately parameterized, the temperature-index approach offers accurate performance (Ohmura, 2001) and provides a robust surface melt representation.

The PDD model calculates the water equivalent of surface snow melt (M, $\mathrm{mm\ w.e.}$). It integrates the near-surface air temperatures above a predefined threshold, which are multiplied by the empirical DDF ($\mathrm{mm\ w.e.\ ^\circ C^{-1}\ day^{-1}}$) (e.g. Hock, 2005). The adjusted PDD model we use in this study can be written as:

$$\sum_{i=1}^{\mathrm{day}} \mathrm{M} = \frac{1}{24}\mathrm{DDF}\sum_{i=1}^{\mathrm{day}}\sum_{j=1}^{24}\mathrm{T}^\star$$
$$\mathrm{T}^\star = \begin{cases} \mathrm{T}-\mathrm{T}_0 & \text{if } \mathrm{T}-\mathrm{T}_0 > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where T is the hourly temperature and T_0 is the threshold temperature.

3.2 Model parameterisation

3.2.1 Threshold temperature $\mathbf{T_0}$

To parameterize the threshold temperature (T_0) for our PDD model, we firstly focus on the binary melt/no-melt signal. We use the ERA5 2-m air temperature data to force the model and run 151 numerical experiments for T_0 ranging from -10.0 $^\circ$C to +5.0 $^\circ$C with a 0.1 $^\circ$C interval. We define a melt day ($\mathrm{MD}^\star$) as a day in which the daily input of the ERA5 2-m air temperature (T) exceeds the T_0. Note that the T is either the daily mean of 6 am and 6 pm or the daily mean of 12 am and 12 pm depending on the satellite estimates we compare to (detailed in the paragraph below). In each T_0 experiment, we calculate the total number of melt days from 1st April of that year to 31st March of the following year as the "annual number of melt days". The modified

Equation 2 can be written as:

$$\begin{aligned}
&\text{Annual number of melt days} = \sum_{i=t_1}^{t_2} \text{MD}^{\star} \\
&\qquad t_1 = 01 - \text{April} - \text{Year} \\
&\qquad t_2 = 31 - \text{March} - (\text{Year+1}) \\
&\qquad \text{MD}^{\star} = \begin{cases} 1 & \text{if } \text{T} - \text{T}_0 > 0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned} \tag{3}$$

Because the satellite melt day product of SMMR and SSM/I (Table 1) is retrieved from the local acquisition times at around 6am and 6pm, we compute the mean of 6 am and 6 pm ERA5 2-m air temperature data for the input T for the PDD model (Equation 3). For the satellite product from AMSR-E and AMSR-2 (Table 1), we compute the mean of 12am and 12pm ERA5 2-m air temperature data as of their local acquisition times. Next, we calculate the result of Equation 3 for each T_0 experiment.

In order to obtain the optimal T_0, we calculate the root-mean-square error (RMSE) between the time series of the annual number of melt days for the satellite estimates and the model experiments in their overlapped years. As we treat each computing cell individually, all calculations are carried out on each cell independently in each iteration (T_0 experiment). Although these three satellite products have different time periods, we assume their comparability as these satellite products are derived from the same algorithm and threshold (Picard and Fily, 2006). Therefore, we calculate the mean of RMSE between three satellite estimates for each cell. Finally, we define the optimal T_0 of each computing cell where the T_0 experiment has the minimal RMSE. If there are multi T_0 experiments that have same minimal RMSE for their computing cell, we calculate the mean of those T_0 as the optimal T_0 (this only happened on the cells that have very low melt days).

3.2.2 Degree Day Factor DDF

The DDF is a scaling number that controls the amount of melt. It is a lumped parameter that relates to all terms of the SEB (Hock, 2005; Ismail et al., 2023) and is suggested not to be considered as a constant number in PDD models (Ismail et al., 2023). To parameterize the DDF for our PDD model, we substitute the optimal T_0 found in Section 3.2.1 into the Equation 2, and run a series of numerical experiments forced by the hourly ERA5 2-m air temperature data: we firstly set the DDF to 1 $\mathrm{mm\ w.e.\ {}^{\circ}C^{-1}\ day^{-1}}$ then we iterate 291 times with 0.1 $\mathrm{mm\ w.e.\ {}^{\circ}C^{-1}\ day^{-1}}$ increments.

In order to address the optimal DDF, we repeat the calculations for the RMSE between the annual melt amount calculated in each DDF experiment and the melt amount from RACMO2.3p2 simulations for each computing cell. Similarly, we define the optimal DDF where the experiment has the minimal RMSE for each computing cell. If there are multiple DDF experiments that have same minimal RMSE for their computing cell, we calculate the mean of those DDF as the optimal DDF (this only happened on the cells that have very low melt amount).

3.3 Model evaluation

3.3.1 Goodness-of-fit testing

The two-sample Kolmogorov–Smirnov test (hereafter two-sample KS test) has been used in testing for significant difference between two non-Gaussian climatic distributions when parametric tests are inappropriate (e.g. Deo et al., 2009; Zheng et al., 2021). It has also been used as an alternative way to test the dissimilarity of climatic data as a validation of tests on statistical parameters such as the mean (Zheng et al., 2021). The two-sample KS test non-parametrically tests the distributional dissimilarity between two samples by quantifying the distance between two sample-derived empirical distribution functions (Lanzante, 2021). The null hypothesis is that the two samples are from the same continuous distribution. The test result returns a logical index that either accepts or rejects the null hypothesis at the 5% significance level ($p < 0.05$).

Limited by the duration of satellite era and reanalysis data, the time series of annual data for each computing cell is no larger than 45 years with non-normality. To test the goodness-of-fit of the parameterized PDD model, we therefore perform the two-sample KS tests between the time series of annual number of melt days/ melt amount from the satellite estimates/ RACMO2.3p2 and from the parameterized PDD model outputs. We define a 'same distribution cell' as a cell with no statistically significant evidence from the two-sample KS test for the rejection of the null hypothesis (that the two samples are from the same continuous distribution).

3.3.2 K-fold cross-validation

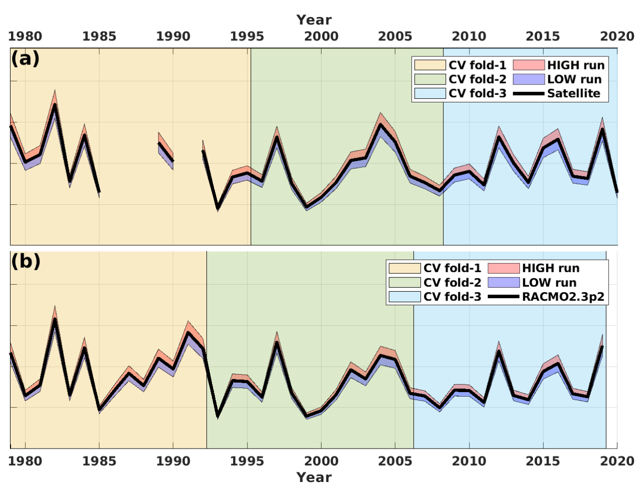

Figure 2. Schematic overview of the time periods for each CV folders and the HIGH, LOW sensitivity experiments.

The cross-validation technique has been developed since the 20th century (Stone, 1974) and has became a standard technique in the field of climate and weather predictions (e.g. Mason, 2008; Maraun and Widmann, 2018). It is especially suitable for statistical models that are calibrated and evaluated on the same data (Maraun and Widmann, 2018).

We consider the spatial variability of PDD parameters by parameterizating the model in each computing cell for the whole time period. However, this does not allow us to explore the variability of the PDD parameters in a temporal sense, as Ismail et al. (2023) suggest that the temporal variability of DDF should also be considered. Due to the short period of the satellite-era and the scarcity of in situ Antarctic surface melt data (Gossart et al., 2019), our PDD model is parameterized and evaluated using the same dataset covering the past four decades.

To therefore assess the temporal dependency of the PDD parameters, we perform an adjusted 3-fold cross-validation (hereafter 3-fold CV). The satellite melt occurrence estimates used in this study cover 38 years (four years have been omitted). Therefore, we sequentially divide the satellite estimates into two 13-year folds and a 12-year fold (Figure 2a). Note that in Section 3.2.1 we calculate the RMSE between the PDD and three satellite estimates on their overlapping period, respectively, and calculate the mean of those three RMSE. However, the second fold has actually only 7 years of overlap between the satellite SMMR and SSM/I, and satellite AMSR-E. Here, we firstly calculate the mean of satellite estimates between their overlapping periods prior to the 3-fold CV and then, we perform the 3-fold CV. The 3-fold CV has three members. the first membercontains the first and second fold used to parameterize the PDD model, and the third foldis used to test the model. In Member 2, we take the first and third fold to parameterize the PDD model and test the model on the second fold. In Member 3, we take the second and third fold to parameterize the PDD model and test the model on the first fold. Similarly, we repeat the calculations for the RACMO2.3p2 surface melt amount but the folds are divided into two 14-year folds and a 13-year fold (Figure 2b).

3.3.3 Sensitivity experiments

Although RACMO2.3p2 is suggested to be one of the best models on reconstructing Antarctic climate, a cold bias of -0.51 K for the near-surface temperatures is also reported (Mottram et al., 2021). However, it is unclear how much this cold bias influences the output of RACMO2.3p2 snowmelt simulations, at least on the spatial scale. Satellite estimates are more direct products for Antarctic surface melt. However, biases in satellite products are likely due to frequent equipment replacements, i.e., 4 times in the period 1979–2005 (Picard and Fily, 2006; Picard et al., 2007).

To explore the sensitivity of PDD parameters and model outputs to biases in both the satellite and RACMO2.3p2 products, we perform two sensitivity experiments. In the first sensitivity experiment, we explore the response of T_0 and the PDD melt-day (and CMS) outputs to perturbations in satellite estimates. We increase/decrease (HIGH/LOW run) satellite CMS estimates by 10% (Figure 2a) for each grid-cell then repeat the T_0 parameterization as described in Section 3.2.1, respectively. In the second sensitivity experiment, we explore the sensitivity of the DDF and the PDD melt amount outputs to perturbations in RACMO2.3p2 melt estimates. We increase/decrease (HIGH/LOW run) the RACMO2.3p2 melt estimates by 10% (Figure 2b) for each grid-cell then repeat the DDF parameterization as described in Section 3.2.2, respectively. Note that in the context of the sensitivity experiments, our optimal parameterization of T_0 and DDF in Section 3.2.1 and Section 3.2.2 constitutes our CONTROL run.

In addition, these sensitivity experiments enable us to explore potential applications of our PDD model to predict Antarctic surface melt in the future. Although our PDD parameters remain stable for the contemporary climate, it is uncertain how they could change in a warmer climate. Exploring the variations in PDD parameters by performing the above sensitivity experiments provides some insights on the model ability to simulate melt under future warming scenarios.

4 Results and discussion

4.1 Optimal PDD parameters

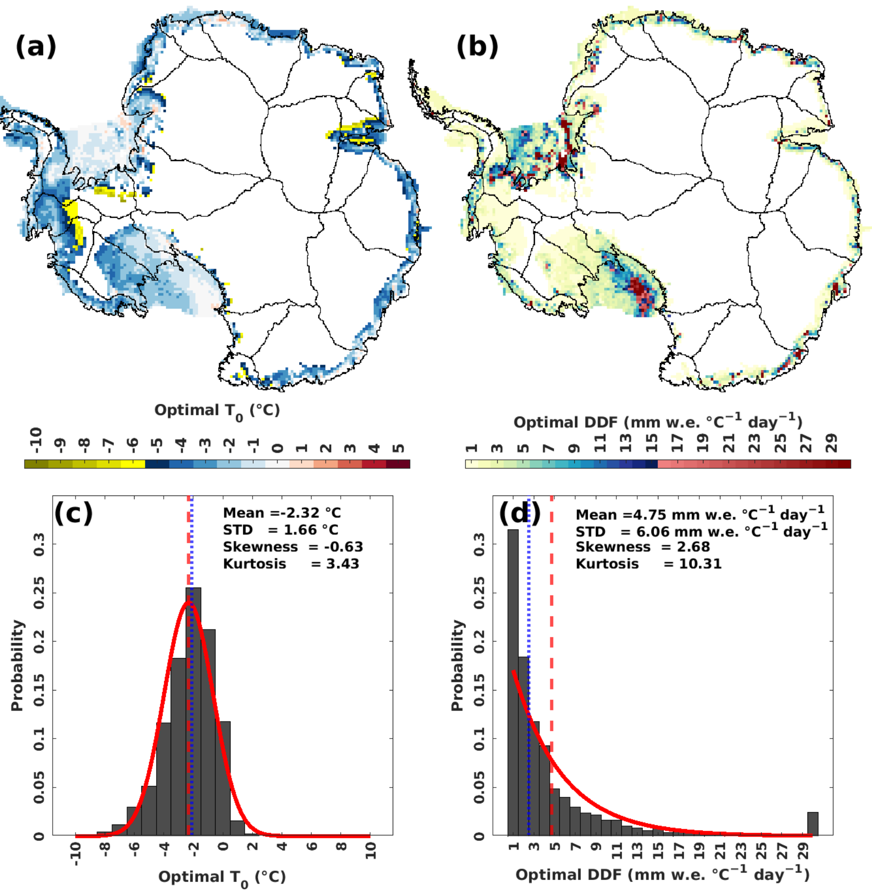

Figure 3. (a) Spatial map of the optimal T_0 (°C) of each computing cell. (b) Spatial map for the optimal DDF ($\mathrm{mm\ w.e.\ {}^{\circ}C^{-1}\ day^{-1}}$) for each computing cell. (c) Probability histogram of the optimal T_0 (°C). Red curve is the fitted normal distribution. Red dashed vertical line is the mean of T_0 for all computing cells. Blue dotted line is the median of T_0 for all computing cells. (d) Probability histogram for the optimal DDF ($\mathrm{mm\ w.e.\ {}^{\circ}C^{-1}\ day^{-1}}$). Red curve is the fitted exponential distribution. Red dashed vertical line is the mean of DDF for all computing cells. Blue dotted line is the median of DDF for all computing cells.

Figure 3a shows the spatial map of the optimal T_0s selected by the minimal RMSE from 151 T_0 experiments on each computing cell (there are 4515 computing cells in total). The optimal T_0 for almost all computing cells are negative. The mean of all optimal T_0 is -2.32 °C. That the dominant number of cells show a negative sign indicates that using T_0 = 0 °C as a melt threshold may significantly underestimate melt events, a finding consistent with other work (Jakobs et al., 2020).

Figure 3c summarizes the statistics of T_0s. The skewness of T_0s is -0.63 indicating a slight left asymmetry of the probability distribution of T_0s. The kurtosis is slightly larger than 3 which is the kurtosis of a normal distribution. We fit a normal distribution with the same mean and standard deviation (STD) (red curve in Figure 3c). That the probability distribution of T_0s is close to the normal distribution is not surprising, given the large sample size of the T_0s (4515 computing cells). There is a small number of cells distributed below -5.5 °C with less than 5% probability (Figure 3c). We highlight these lower-end tail cells with a yellow color in the Figure 3a. These cells are mainly distributed in two areas. One is the interior boundary of the satellite observational area (Figure A2 in the Appendix A) over the drainage basins (e.g. Basin 1, 9, 21 and 22), which is not surprising as the optimal T_0s there may not be significant, given the non-statistically significant ($p \geq 0.05$) temperature-melt correlation over those cells (Figure B1 in the Appendix B). The other area is the central Amery Ice Shelf (Figure 3a).

Figure 3b shows the spatial map of the optimal DDFs identified by the minimal RMSE from 291 DDF experiments on each computing cell. We see a large number of DDFs with relatively low magnitude (from 1 to 4.5 $\mathrm{mm\ w.e.\ °C^{-1}\ day^{-1}}$, colored light yellow), distributed over ice shelves other than the Ross Ice Shelf and Filchner-Ronne Ice Shelf (Figure 3b). We highlight DDFs larger than 15.5 $\mathrm{mm\ w.e.\ °C^{-1}\ day^{-1}}$ in red in Figure 3b. Although the magnitude of the DDF over the cells located in the west Ross Ice Shelf and south-east Filchner-Ronne Ice Shelf may exceed the upper boundary (30 $\mathrm{mm\ w.e.\ °C^{-1}\ day^{-1}}$) of our DDF experiments that we heuristically defined in Section 3.2.2, we do not expand the upper boundary of the DDF or perform more DDF experiments. This is because, (1) the temperature-melt correlations over those cells are not statistically significant ($p \geq 0.05$, Figure B1), therefore the PDD model which is based on the temperature-melt relationship for those cells may not be significant; (2) the total number of those cells is less than 5% of the total number of the computing cells (Figure 3d); (3) surface melting in those cells is negligible under present-day conditions, and even remains negligible in RCP8.5 2100 future projection (Trusel et al., 2015); (4) these parameters are empirically defined by minimizing the RMSE between PDD experiments and satellite estimates/ RACMO2.3p2 simulations, which means the optimal parameters are likely less robust over cells where melt is rare.

Figure 3d summarizes the statistics of DDFs. The probability distribution of the DDFs is asymmetrical and strongly left skewed (Figure 3d). We see that nearly 50% of the DDFs are in the range 1 to 2.5 $\mathrm{mm\ w.e.\ °C^{-1}\ day^{-1}}$. That the majority of the DDFs are low may be associated with the negative T_0s defined in the T_0 experiments. This is because, (1) the parametrization of the T_0 and DDF is sequential. The optimal T_0s are substituted into the Equation 2 (Section 3.2.2) as a predefined variable for the DDF experiments, which means our decision on the optimal T_0 will influence the decision making for the optimal DDF; (2) a low negative optimal T_0 may cause more degrees above the T_0 leading to a low optimal DDF that works in conjunction with the sum of the degrees above a vey low T_0.

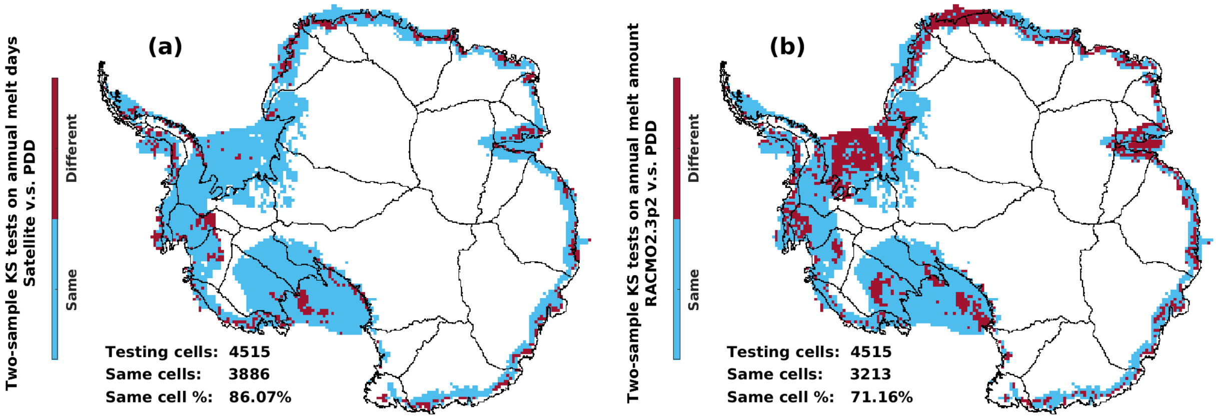

Figure 4. Spatial maps for the two-sample KS test results. The two-sample KS tests are performed individually for each of the 4515 computing cells. The test result "Same" means the tested cell is a same distribution cell where there is no statistically significant evidence for the rejection of the null hypothesis that the testing two samples are from the same continuous distribution (Section 3.3.1). Otherwise, the cell is a different distribution cell ("Different"). (a) is the two-sample KS test results for testing the annual number of melt days between the satellite estimates and the PDD model outputs. (b) is the two-sample KS test results for testing the annual melt amount between the RACMO2.3p2 simulations and the PDD model outputs.

4.2 Model evaluation

4.2.1 Goodness-of-fit

We evaluate the parameterized PDD model outputs (melt day and melt amount) for each computing cell by testing the statistical significance of the similarity between the satellite estimates or RACMO2.3p2 simulations and the PDD model-derived empirical distribution functions. Figure 4 shows the two-sample KS test results for each computing cell. Overall, the parameterized PDD model shows good agreement with the satellite estimates and RACMO2.3p2 simulations both in estimating the annual total of melt days and melt amount, indicated by 86.07% and 71.16% same melt day and amount distribution cells, respectively (Figure 4). Our parameterized PDD model is particularly well-suited for estimating surface melt over the ice shelves in the Antarctic Peninsula, while cells located in other ice shelves, such as the Filchner-Ronne Ice Shelf, ice shelves in Dronning Maud Land, Amery Ice Shelf and Ross Ice Shelf, are either in a good agreement on estimating the surface melt days or amount (Figure 4). That the PDD model performs well in the Antarctic Peninsula is exciting, given the fact that the Antarctic Peninsula is the region of Antarctica experiencing most intense surface melting both at the present (Trusel et al., 2013; Johnson et al., 2022) and in projections of the current century (Trusel et al., 2015).

Next, we evaluate the parameterized PDD model outputs for the whole of Antarctica. Firstly, we evaluate the parameterized optimal T_0 and its related PDD outputs on the surface melt day. To do this, we calculate the cumulative melting surface (CMS) (day km^2) for satellite estimates and PDD outputs, respectively. The CMS which is also known as a melt index (e.g. Trusel et al., 2012), is calculated by multiplying the cell area (km^2) by the total annual melt days (day) in that same cell (Trusel

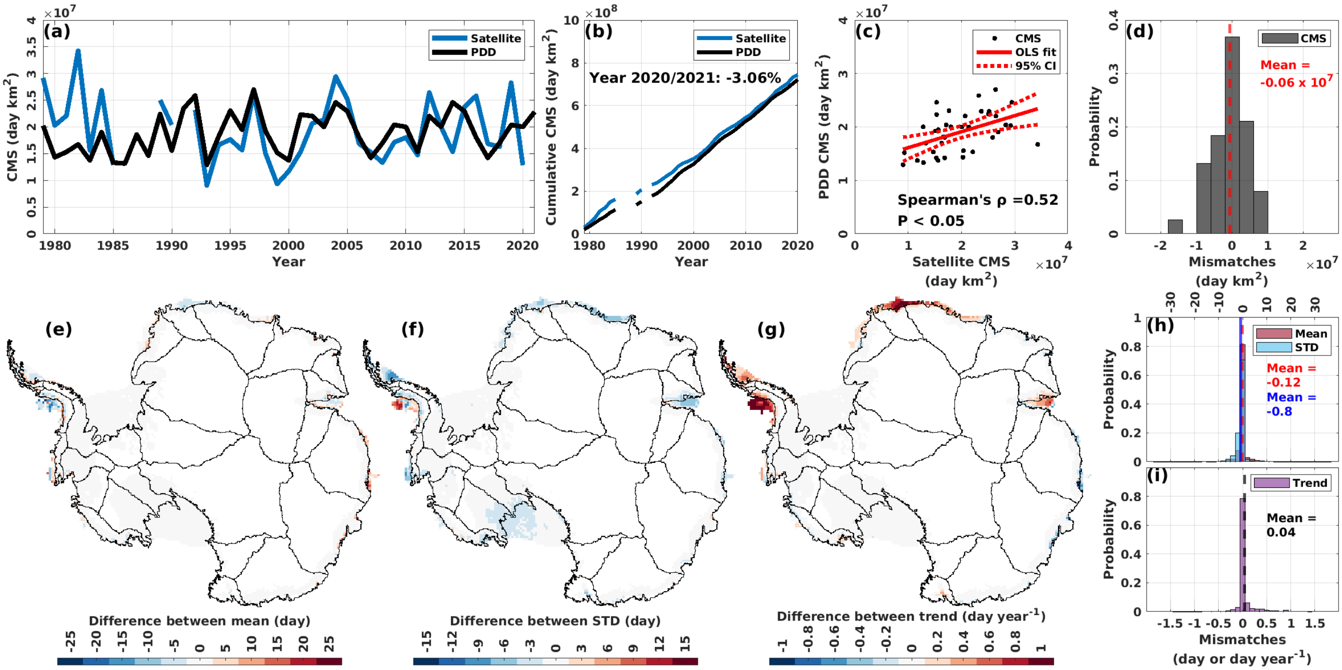

Figure 5. (a) time series for the cumulative melting surface (CMS) (day km^2) for satellite estimates during the period from 1979/1980 to 2020/2021 (with 1986/1987 to 1988/1989 and 1991/1992 omitted), and for PDD outputs during the period from 1979/1980 to 2021/2022. (b) cumulative CMS for satellite estimates and PDD outputs from 1979/1980 to 2020/2021 (with 1986/1987 to 1988/1989 and 1991/1992 omitted). (c) scatter plot and ordinary least squares (OLS) fit between satellite CMS and PDD CMS. (d) probability histogram for the mismatches between the PDD CMS and satellite CMS. Red dashed vertical line indicates the mean of all mismatches. (e) to (g) spatial maps for the differences between mean, standard deviation (STD) and trend of PDD outputs and satellite estimates on the annual melt days (day). Mean, STD and trend for the PDD outputs and satellite estimates are calculated over the period from 1979/1980 to 2020/2021 (with 1986/1987 to 1988/1989 and 1991/1992 omitted), respectively. (h) and (i) probability histograms for the mismatches between the PDD outputs and satellite estimates on mean, STD and trend (histograms for (e) to (g)). Red dashed vertical line indicates the mean of all mismatches between means. Blue vertical line indicates the mean of all mismatches between STDs. Black dashed vertical line indicates the mean of all mismatches between trends. Note that for all panels the satellite estimates from 2002/2003 to 2010/2011 are the average of SMMR and SSM/I, and AMSR-E. The satellite estimates from 2012/2013 to 2020/2021 are the average of SMMR and SSM/I, and AMSR-2.

et al., 2012). We see in Figure 5a that two CMS time series are in a generally good agreement on both the amplitude and the temporal variability, apart from a small number of years including a period from 1979/1980 to 1982/1983, the year 2014/2015, the year 2016/2017 and the year 2019/2020. Although there is a PDD underestimation for the first decade (1980 to 1990), the cumulative CMS of PDD at the end of the 38-year period is in a good agreement with the cumulative CMS of satellite estimates (-3.06% PDD cumulative CMS underestimation compared to the satellite cumulative CMS, Figure 5b). The positive correlation between the satellite CMS and the PDD CMS is statistically significant (Spearman's $\rho = 0.52$, $p < 0.05$, Figure 5c).

The probability histogram for mismatches between the PDD and satellite CMS also indicates a good agreement between the PDD and satellite CMS (Figure 5d). The mismatches are distributed symmetrically to the mean which is approximated to zero.

Globally, we see the the PDD model has the ability to capture the main spatial patterns of surface melt days when compared to the satellite estimates for a majority of the computing cells (Figure 5e, f and g). The computing cells that have relatively large disagreement between the mean annual melt days of PDD outputs and of satellite estimates are mainly located over the ice shelves in the Antarctic Peninsula ($\sim$ -2.5 to -22.5 days), over the Abbot Ice Shelf ($\sim$ -5.5 to -12.5 days over the marine edge and $\sim$ +2.5 to +7.5 days over the interior) and over the Shackleton Ice Shelf ($\sim$ +7.5 to +12.5 days). However, these cells with relatively large disagreement in mean only amount to around 5% of the total computing cells (Figure 5h), and overall for all computing cells, the mean of mismatches in means between the PDD and satellite annual melt days is approximately zero (-0.12 days, Figure 5h). That the PDD model captures the main spatial patterns of melt is not surprising, given the statistically significant positive correlation between surface melt and 2-m air temperature in most of the Antarctic ice shelf and coastal cells used in the calculations (Figure B1).

The computing cells that have relatively large disagreement on STD are mainly located over the Wilkins Ice Shelf ($\sim$ +4.5 to +13.5 days) and over the south of Larsen C Ice Shelf ($\sim$ -7.5 to -10.5 days). Similar to the cells that have relatively large disagreement in their means, these cells amount to only a negligible proportion (less than 5%) of the total number of the computing cells. However, there are around 20% of the computing cells that have -1 to -3 days of STD mismatches (Figure 5h), spatially distributed widely over the eastern Ross Ice Shelf, West Antarctic drainage basins 18 and 19, the Abbot Ice Shelf, ice shelves in Dronning Maud Land, and the Amery Ice Shelf. The mismatches in trend are not symmetrical about zero, both shown by the dominant area of red color (all ice shelves in the Antarctic Peninsula, almost all ice shelves in Dronning Maud Land and nearly the whole Amery Ice Shelf) to blue (some computing cells over the Wilkes Land) in Figure 5g and a slightly right-skewed probability histogram of trend mismatches with a positive mean (+0.04 day year^{-1}, Figure 5i).

Secondly, we evaluate the parameterized optimal DDF and its related PDD outputs on the surface melt amount. Similar to the negative mismatches between PDD and satellite estimates on the CMS for the period from 1979/1980 to 1982/1983 (Figure 5a), negative mismatches of PDD against the RACMO2.3p2 are also present on the annual melt amount for 1982/1983 (Figure 6a). The abnormally extensive melt in 1982/1983 has been reported by previous studies (Zwally and Fiegles, 1994; Liu et al., 2006; Johnson et al., 2022). It is suggested to be driven by the SAM, because of an inverse relationship between the number of melt days in Dronning Maud Land and the southward migration of the southern Westerly Winds (Johnson et al., 2022). The disagreement of the PDD model for this extensive melt event is most likely explained by the absence of any substantial temperature anomaly in the input ERA5 2-m temperature, because of the temperature-dependency of the PDD model (Equation 2) and the temperature-melt relationship (Figure B1). It could also partly be explained by the fact that the PDD parameters were defined based on fitting multi-decadal timeseries between PDD experiments and satellite/ RACMO2.3p2 (Section 3.2.1 and 3.2.2), meaning that some inter/inner- annual signals may not be fully captured.

Apart from the 1982/1983 event, other negative mismatches from PDD are also evident in the period from 1991/1992 to 1992/1993 (Figure 6a). However, we cannot compare this PDD melt amount mismatch period to the PDD CMS mismatch as the year 1991/1992 is omitted for all the analysis related to the satellite estimates due to the missing satellite data. Notwithstanding,

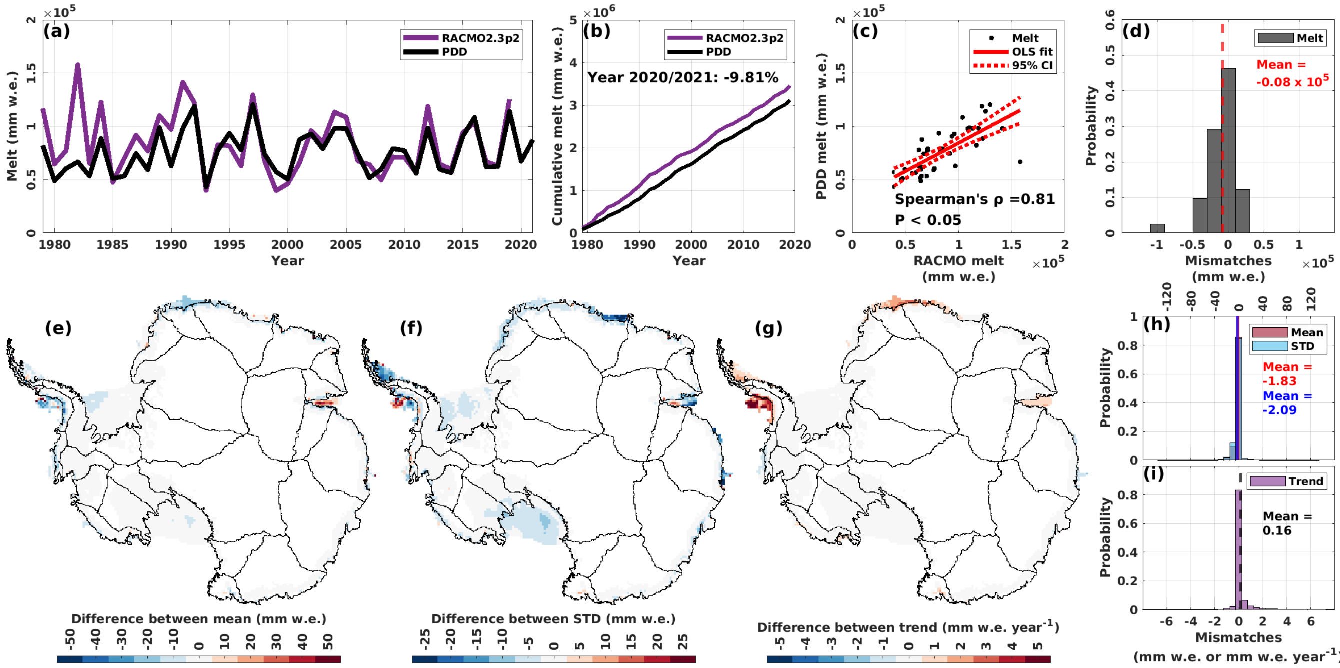

Figure 6. (a) time series for the annual melt amount (mm w.e.) for RACMO2.3p2 simulations during the period from 1979/1980 to 2019/2020, and for PDD outputs during the period from 1979/1980 to 2021/2022. (b) cumulative annual melt amount for RACMO2.3p2 simulations and PDD outputs from 1979/1980 to 2019/2020. (c) scatter plot and ordinary least squares (OLS) fit between satellite annual melt amount and PDD annual melt amount. (d) probability histogram for the mismatches between the PDD annual melt amount and satellite annual melt amount. Red dashed vertical line indicates the mean of all mismatches. (e) to (g) spatial maps for the differences between mean, standard deviation (STD) and trend of PDD outputs and RACMO2.3p2 simulations. Mean, STD and trend for the PDD outputs and RACMO2.3p2 simulations are calculated over the period from 1979/1980 to 2019/2020. (h) and (i) probability histograms for the mismatches between the PDD outputs and RACMO2.3p2 simulations on mean, STD and trend (histograms for (e) to (g)). Red dashed vertical line indicates the mean of all mismatches between means. Blue vertical line indicates the mean of all mismatches between STDs. Black dashed vertical line indicates the mean of all mismatches between trends.

excluding these periods, we see the time series of annual melt amount of the PDD outputs and RACMO2.3p2 simulations are generally in good agreement, especially after 1992/1993 when the two curves start overlapping (Figure 6a) whilst the PDD-satellite CMSs show some disagreement (e.g. 1995/1996, 1999/2000, 2014/2015, 2016/2017 and 2019/2020, Figure 5a). That the PDD is in a good agreement with RACMO2.3p2 on the annual melt amount is also evident by the statistically significant strong positive correlation (Spearman's ρ = 0.81, p < 0.05, Figure 6c). However, the probability histogram of PDD melt mismatches is slightly left-skewed with a negative mean (-0.08×10^5 mm w.e., Figure 6d) and the PDD model underestimates around 9.81 % for the 41-year integrated annual melt amount compared to the RACMO2.3p2 (Figure 6b). Nevertheless, this underestimation on the 41-year integrated annual melt amount is not evolving through the past four decades, as we see in

Figure 6b: the two curves differ in the first decade (i.e. the gap between the two curves is increasing from $\sim$ 1980 to $\sim$ 1990) and becomes parallel for the following three decades.

Figure 6e, f and g show the spatial maps for the difference between the mean, STD and trend of the PDD annual melt amount and RACMO2.3p2 mean annual melt amount for the period from 1979/1980 to 2019/2020. As shown in Figure 6e, f and g, the differences over most of the computing cells are equal to or close to zero, which is similar to the spatial difference maps between the PDD outputs and satellite estimates in Figure 5e, f and g. This indicates that the PDD model has the ability to capture the main spatial patterns of both the surface melt days and amount, when compared to the satellite estimates and RACMO2.3p2 simulations, for the majority of the computing cells. There are less than 5% computing cells with mismatches in the mean of lower than -15 mm w.e. or larger than +15 mm w.e. (Figure 6h). These cells are spatially distributed over the western Antarctic Peninsula, ice shelves in Dronning Maud Land, and the Amery Ice Shelf. For the disagreement on the STD, around 10% of the computing cells mismatch -5 to -15 mm w.e. (Figure 6h). The computing cells that have relatively large disagreement on STD are spatially distributed over the Antarctic Peninsula, ice shelves in eastern Dronning Maud Land, the Amery Ice Shelf and ice shelves in western Wilkes Land (Figure 6f). The mismatch in trends between the PDD and RACMO2.3p2 annual melt amount is similar to the mismatch in trends between the PDD and satellite annual melt days, as they both have the same positive mismatch spatial patterns (Antarctic Peninsula, Dronning Maud Land and Amery Ice Shelf, Figure 5g and Figure 6g) and similar right-skewed probability histograms with positive means (Figure 5i and Figure 6i). This could be explained by other players driving surface melting, such as the Southern Annular Mode (SAM) (Torinesi et al., 2003; Tedesco and Monaghan, 2009; Johnson et al., 2022) which explains $\sim$ 11%–36% of the melt day variability (Johnson et al., 2022). However, these mismatches in trends do not necessarily require that we reject the PDD model, as the trend presented by the PDD model is a reflection of the trend of the input temperature (Figure C1 in the Appendix C), because of the linear relationship between air temperature and surface melt (Figure B1). The disagreement in trends, therefore, is actually between the satellite/RACMO2.3p2 and ERA5 2-m temperature, rather than between the satellite/RACMO2.3p2 and the PDD model itself.

4.2.2 Temporal dependency of the PDD parameters

To evaluate our PDD model in a temporal sense, we perform 3-fold CV for T_0 and DDF (as described in Section 3.3.2), respectively. Table 2 lists the periods for the training folds and testing folds for each T_0 and DDF member. The training fold is used to parameterize the PDD model parameters. For example, in T_0 Member 2, we use the satellite estimates over the periods 1979/1980–1995/1996 (1986/1987–1988/1989 and 1991/1992 are omitted) and 2009/2010–2020/2021 to run 151 T_0 experiments (similar to the Section 3.2.1, but using different time period of satellite estimates) to parameterize the optimal T_0 for Member 2 (see also Figure 2). The testing fold is used to evaluate the PDD model parameterized only on the training fold. For example, in DDF Member 3, the Member 3 DDF is parameterized by the training fold which is over the period from 1993/1994 to 2019/2020 (see also Figure 2). Once the Member 3 DDF is parameterized, we run the PDD model with the Member 3 DDF for the whole 41-year time period. Then we extract the PDD model (the Member 3 DDF PDD model) outputs

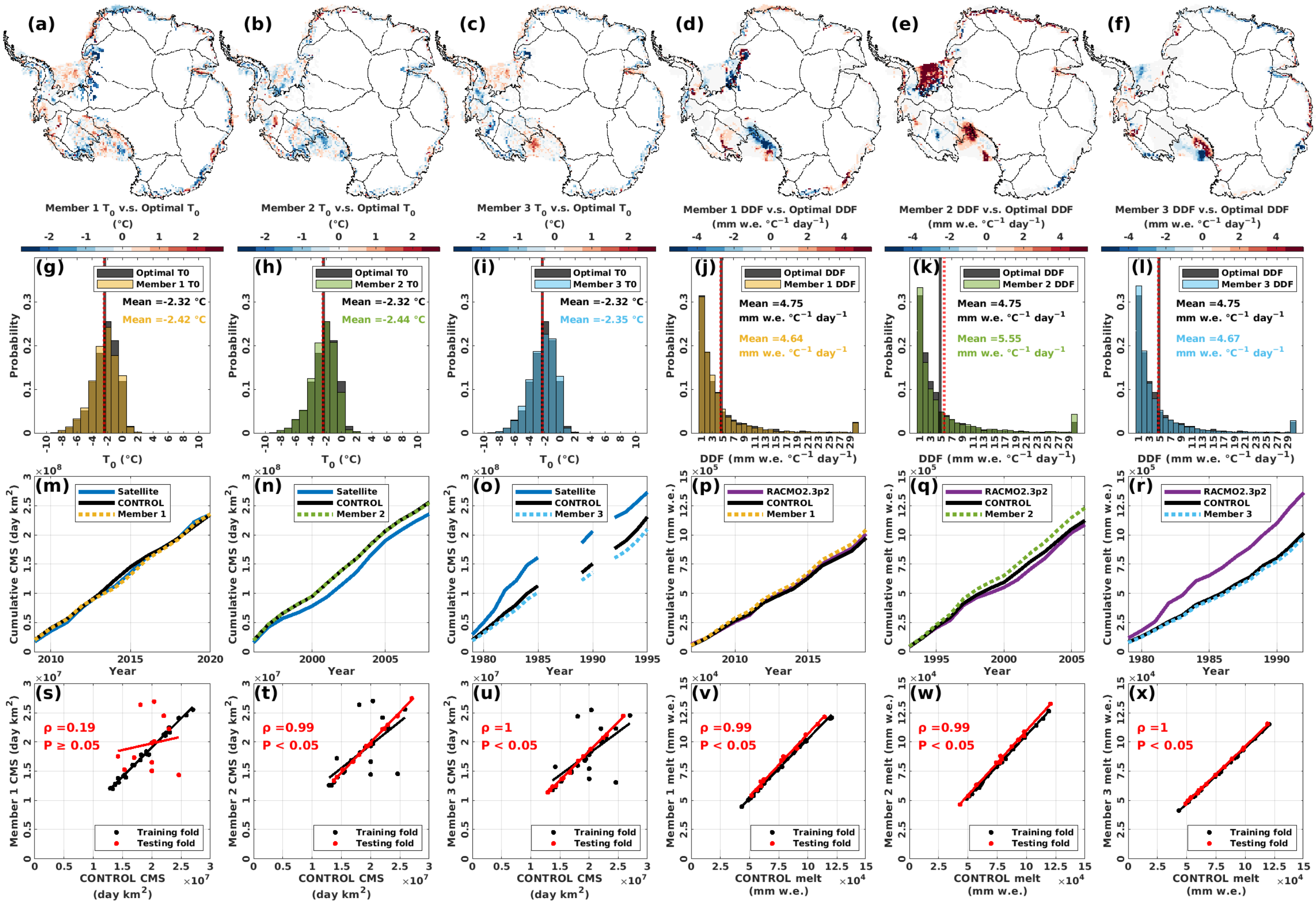

Figure 7. (a) to (f) spatial maps for the differences between the T_0/ DDF parameterized in each member of the T_0/ DDF 3-fold CV and the optimal T_0/ DDF, respectively. (g) to (l) probability histograms for the T_0/ DDF of each T_0/ DDF 3-fold CV and the optimal T_0/ DDF, respectively. Black vertical lines indicate the mean of optimal T_0s/ DDFs. Red dotted vertical lines indicate the mean of T_0/ DDF for each member, respectively. (m) to (r) cumulative CMS/ annual melt amount for satellite estimates/ RACMO2.3p2 simulations, CONTROL (which is the PDD model run with optimal T_0 and DDF) and each member for the period of the testing-fold, respectively. We calculate the difference of cumulative CMS/ annual melt amount between each member and the CONTROL, at the end of the testing fold, respectively. (s) to (x) scatter plots for the CMS/ annual melt amount of each 3-fold CV member against the CONTROL, respectively. The Spearman's ρ and its statistical significance for the testing fold between each member and the CONTROL are calculated, respectively.

for the testing fold period (1979/1980–1992/1993) from the whole 41-year model outputs, for testing (evaluating) the DDF Member 3.

Figure 7 shows the results of the 3-fold CV on T_0 and DDF. We see in Figure 7a to f that there are changes on the value of the T_0 and DDF for a dominant number of the computing cells, depending on the time window (i.e. the training fold) we choose to parameterize the PDD model. Especially for the DDF members, we see conspicuous changes in the values of the DDFs in

Table 2. Periods of the training and testing folds for the T_0 and DDF 3-fold cross-validation, respectively.

Member	Training fold	Testing fold
T_0 CONTROL	1979/1980–2020/2021[a]	–
T_0 Member 1	1979/1980–2008/2009[a]	2009/2010–2020/2021
T_0 Member 2	1979/1980–1995/1996[a] and 2009/2010–2020/2021	1996/1997–2008/2009
T_0 Member 3	1996/1997–2020/2021	1979/1980–1995/1996[a]
DDF CONTROL	1979/1980–2019/2020	–
DDF Member 1	1979/1980–2006/2007	2007/2008–2019/2020
DDF Member 2	1979/1980–1992/1993 and 2007/2008–2019/2020	1993/1994–2006/2007
DDF Member 3	1993/1994–2019/2020	1979/1980–1992/1993

[a] periods from 1986/1987 to 1988/1989 and 1991/1992 are omitted.

the computing cells over the western and southern Ross Ice Shelf, the Filchner-Ronne Ice Shelf and coastal basins 2 and 3 (Figure 7d, e and f), which indicates that a large temporal variability of PDD parameters may exist. However, this indication that a large temporal variability of PDD parameters exists may not be reliable for the western and southern Ross Ice Shelf and coastal basin 2, given that there is no statistically significant evidence for the temperature-melt relationship (Figure B1).

Although we see the parameter changes associated with the time windows for the dominant number of the computing cells, these changes reduce when we look at the whole population of the parameters in each member (Figure 7g to l). It is evident that the probability histogram of the optimal parameters and the probability histogram of each member's parameters are closely comparable, with negligible differences between means (excluding the DDF Member 2 where the differences between means is relatively larger: +0.8 $\mathrm{mm\ w.e.\ ^{\circ}C^{-1}\ day^{-1}}$, Figure 7k).

Next, we evaluate each member's parameters on the testing fold. Firstly, we calculate the cumulative CMS/ annual melt amount for the time windows of the testing folds from the PDD models that are parameterized by the training folds, for each T_0 and DDF members respectively. Overall, the curves of each member are comparable and overlapping with the CONTROL (Figure 7m to r), indicating the temporal consistency of our PDD model, and that the ability of our PDD model in estimating the Antarctic-wide surface melt in terms of the melt occurrence (CMS) and the melt totals (amount) is independent of the time windows chosen for the parameterization. Although the parameters in each computing cells vary through the parameterization time window, the overall performance of the PDD model for all the computing cells as a whole is generally consistent.

Secondly, we calculate the Spearman's ρ and its statistical significance for the testing fold between each member and the CONTROL (Figure 7s to x). Apart from the T_0 Member 1, we see each member's PDD estimates are statistically significantly, strongly ($\rho \geq$ 0.99, p $\leq$ 0.05) correlated with the CONTROL PDD estimates (Figure 7t to x). However, this is not surprising, given the comparable probability distributions of parameters and the indistinguishable cumulative curves between each member's PDD and the CONTROL PDD (Figure 7g to r). Although the T_0 Member 1 PDD estimates and PDD CONTROL estimates are strongly correlated to the training fold (black dots in Figure 7s), which is not surprising as the T_0 Member 1 PDD is parameterized by those PDD CONTROL estimates, the T_0 Member 1 PDD estimates and PDD CONTROL estimates are not statistically significantly correlated ($\rho = 0.19$, $p \geq 0.05$) to the testing fold (red dots, Figure 7s).

To further explore this disagreement in the testing fold, we plot the time series of CMS for satellite estimates, CONTROL estimates and T_0 Member 1 estimates in Figure D1, in the Appendix D. We find that the T_0 Member 1 estimates in the testing fold are likely not unrealistic values. Instead, they are in a good agreement with the satellite estimates over the testing-fold period, as the time series of satellite CMS and Member 1 CMS almost overlap. Therefore the disagreement between the T_0 Member 1 estimates and the CONTROL estimates over the testing-fold period might be the disagreement between the satellite estimates and CONTROL estimates, as the time series of satellite CMS and Member 1 CMS almost overlap. Although the abilities of Member 1 T_0 and optimal T_0 in capturing the cumulative satellite estimates are robust and indistinguishable (Figure 7m), the agreement between the time series of Member 1 T_0 and satellite CMS may suggest that the T_0 parameterized by the Member 1 training fold (which is the period from 1979/1980 to 2008/2009 with 1986/1987–1988/1989 and 1991/1992 omitted) are more robust in capturing the interannual variability of the satellite estimates (for the period from 2009/2010 to 2020/2021) than the optimal T_0 that parameterized by the full 38-year period. However, the data sample that used to parameterize the Member 1 T_0 is only 2/3 the full data length which parameterized the optimal T_0, giving us less confidence on the reliability of the Member 1 T_0s for the full 38-year period.

4.2.3 Sensitivity experiments and implementation to the future predictions

Figure 8 shows the result from our sensitivity experiments. We see changes in the PDD parameters associated with the increase (HIGH run, +10% magnitude of the satellite / RACMO2.3p2 data) and decrease (LOW run, -10% magnitude of the satellite / RACMO2.3p2 data) on the satellite estimates and RACMO2.3p2 simulations (Figure 8a to d). That the T_0 decreases/ increases with the increase/ decrease of the satellite estimates is expected, because a decrease of the threshold temperature is expected to allow more temperatures above the threshold to produce more melt days, and vice versa. The increase/ decrease of the RACMO2.3p2 simulations leads to an increase/ decrease on the DDFs, which is also expected because the T_0 is predefined for the DDF parameterization, thus the sum of the degrees above the T_0 becomes an invariant. Therefore, as a scaling number, the DDF is expected to increase to amplify the sum of the degrees above the T_0 to match the increase of the RACMO2.3p2 melt amount simulations, and vice versa.

Figure 8e shows that the PDD model is less sensitive than the satellite estimates on the low melt scenario, where the PDD estimates only decrease 9.78% for the integrated 38-year CMS when the satellite estimates decrease 10%. Although the PDD model is more sensitive than the satellite estimates on the high melt scenario, where we see that PDD increases 10.84% on the 38-year integrated CMS with the 10% increase of the satellite estimates, this increase in PDD estimates is linear with respect to the increase in satellite estimates, and is of the same proportion (Figure 8e). For the sensitivity experiments on the DDF, we see that the PDD model is less sensitive than the RACMO2.3p2 in both the HIGH and LOW melt scenarios. Taken together, the sensitivity of the PDD model is linear (the correlations do not change much across different sensitivity experiments, Figure 8f and h) and with the same order of magnitude to both the satellite estimates and RACMO2.3p2 simulations, suggesting that the PDD is also applicable to future climate change scenarios where surface melting is predicted to increase (Trusel et al., 2015).

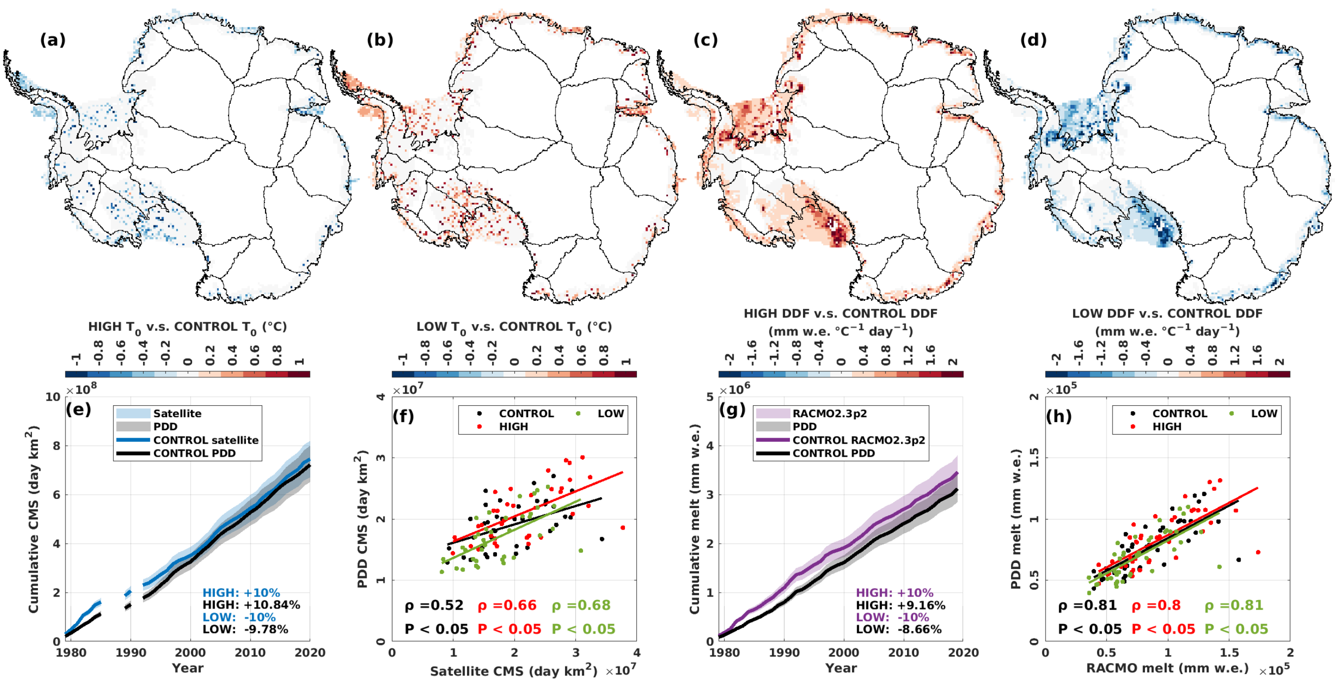

Figure 8. (a) and (b) spatial maps for the difference between the T_0 parameterized in the HIGH/ LOW experiment and the CONTROL (optimal) T_0. (c) and (d) spatial maps for the difference between the DDF parameterized in the HIGH/ LOW experiment and the CONTROL (optimal) DDF. (e) and (g) cumulative CMS/ annual melt amount for the satellite estimates/ RACMO2.3p2 simulations and PDD outputs. Note that the period for (e) is from 1979/1980 to 2020/2021 (with 1986/1987 to 1988/1989 and 1991/1992 omitted). The period for (g) is from 1979/1980 to 2019/2020. The upper and lower boundaries of the semi-transparent shaded areas indicates the HIGH/ LOW satellite estimates and the HIGH/ LOW PDD outputs. The percentage difference annotated in the left-bottom corner is calculated between the HIGH/ LOW and the CONTROL for each variable (by "variable", we mean satellite melt occurrence data/ PDD melt occurrence and amount data/ RACMO2.3p2 melt amount data), respectively. (f) and (h) scatter plots and the Spearman's ρ (with its statistical significance) for PDD outputs and satellite/ RACMO2.3p2, from each sensitivity experiment (HIGH, LOW and CONTROL).

Overall, the PDD model is less sensitive than the satellite estimates and RACMO2.3p2 simulations, which indicates that our PDD model can reduce the bias that the satellite and RACMO2.3p2 have on the melt products, even though their biases are unclear (Picard et al., 2007; Mottram et al., 2021).

4.3 Limitations of the PDD model

The PDD model has the notable advantage of high computational efficiency due to its one-dimensional nature and being solely forced by 2-m air temperature. However, in reality the 2-m air temperature is not the sole driver of Antarctic surface melting (Figure B1). A primary limitation of the PDD model is systematically introduced by the temperature-dependency, making it difficult to accurately estimate surface melt strengthened/ weakened or triggered by other components of the surface energy budget that may accompany katabatic winds (Lenaerts et al., 2017) and climatic phenomena such as the SAM (e.g. Tedesco and Monaghan, 2009; Johnson et al., 2022), El Niño Southern Oscillation (Tedesco and Monaghan, 2009; Scott et al., 2019), föhn winds (e.g. Turton et al., 2020), atmospheric rivers (Wille et al., 2019), sea ice concentrations (Scott et al., 2019), or proximity to dark surfaces such as bare rock (Kingslake et al., 2017). Although we combine observations and model simulations to robustly establish our PDD parameterization and consider the spatial variability of model parameters, the PDD model cannot fully replicate a few of the extensive melt events captured by satellites and RACMO2.3p2 (Figure 5a and Figure 6a).

Besides, the model simply multiplies a scaling number (DDF) by the summation of temperature above a certain threshold (T_0). It lacks the ability to simulate or account for other physical mechanisms such as the meltwater ponding, percolation through the snowpack, refreezing, and so on. As the model is parameterized and calibrated by satellite- and SEB-derived estimates, it is also limited by the various assumptions and shortcomings inherent in those methods. Although we perform a number of cross-validation and sensitivity experiments, due to the scarcity of surface melt data from in situ measurements (Gossart et al., 2019), our PDD output has yet to be confirmed by other datasets.

5 Conclusions

We have constructed a PDD model based on the temperature-melt relationship (e.g. Hock, 2005; Trusel et al., 2015), and used it to estimate surface melt in Antarctica through the past four decades. We parameterized the PDD model by running numerical experiments on each individual computing cell to iterate over various combinations of the threshold temperature and the DDF (Section 3.2). We individually selected an optimal parameter combination by locating the minimal RMSE between the PDD and satellite estimates, and SEB simulations, for each computing cell. We independently performed two-sample KS tests on each computing cell in order to assess the goodness-of-fit for the parameterized PDD model. We also temporally and spatially compared the PDD estimations, satellite estimates and RACMO2.3p2 simulations to evaluate the parameterized PDD model. We found that the PDD model has the ability to capture the main spatial and temporal features for a majority of cells in Antarctica under a range of melt regimes (Section 4.2.1).

As the parameters were parameterized spatially, the PDD is overall in a good agreement with the spatial patterns shown by the satellite and RACMO2.3p2 data, with the exception of an underestimation in the ice shelves of the western Antarctic Peninsula and an overestimation of melt days on Shackleton Ice Shelf and of melt amount on Amery Ice Shelf. The most inadequate estimation was in 1982/1983, during which we found large PDD underestimation on both the melt days and amount. We suggest this underestimation corresponds to SAM-influenced climatic conditions, and that the PDD lacks the ability to accurately capture melt if it arises from effects such as föhn winds or atmospheric rivers that are not reflected in the input ERA5 2-m temperature fields used to force the calculations (e.g. Turton et al., 2020; Wille et al., 2019).

These limitations aside, we found overall high fidelity of PDD model, suggested by the 3-fold cross-validation. Although the PDD parameters vary on the cell-level through the different time window chosen for parameterization, the probability distribution for all computing cells changes negligibly and the overall performance of the PDD model when considering all computing cells is consistent. From the sensitivity experiments, we found the changes of the PDD estimates are comparable to the changes in training data (satellite and RACMO2.3p2 data). The correlations between the PDD estimates and training data exhibit stability regardless of the changes in the training data.

The PDD model can not only relatively accurately estimate surface melt in Antarctica compared with the satellite estimates and more sophisticated SEB model, but it is also highly computationally efficient. These advantages may allow us to use the PDD model to explore Antarctic surface melt in a longer-term context into the future and over periods of the geological past when neither satellite observations nor SEB components are available. This efficiency also allows our model to be employed at a far higher spatial resolution than regional climate models. However, due to the systematical limitations of the PDD model and the scarcity of Antarctic surface melt data available (Gossart et al., 2019), more work is needed, such as model evaluation by independent melt data and discussions of approximations to the physical processes (e.g. refreezing) taking place after surface melting. Nevertheless, PDD models have been used in many numerical ice sheet models for the empirical approximation of surface mass balance computations, due to their unique advantages in terms of their simple temperature-dependency and computational efficiency. We propose that our spatially-parameterized implementation extends the utility of the PDD approach and, when parameterized appropriately, can provide a valuable tool for exploring surface melt in Antarctica in the past, present and future.

Data availability. The ERA5 reanalysis data are available from https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5 (last access: 02 August 2022). The Zwally Antarctic drainage basin (Zwally et al., 2012) data are available from http://imbie.org/imbie-3/drainage-basins/. The satellite SMMR and SSM/I, AMSR-E and AMSR-2 products are available from https://doi.org/10.18709/perscido.2022.09.ds376 (Picard, 2022). The RACMO2.3p2 data are available from https://doi.org/10.5194/tc-12-1479-2018 (Van Wessem et al., 2018). The annually PDD model data (this study) is available in this study. Higher temporal resolution (monthly, daily and hourly) PDD model data (this study) is available by contacting yaowen.zheng@vuw.ac.nz.

Appendix A: Satellite data

The number of melt days and the area of surface melt can be detected using the microwave brightness temperature data since 1979 (e.g. Torinesi et al., 2003; Picard and Fily, 2006). The theoretical basis of this approach is that changes between dry and wet snow can be distinguished by the upwelling microwave brightness temperature change (Chang and Gloersen, 1975). When dry snow is melting, the meltwater at the surface significantly changes the dielectric properties of the surface by increasing absorption and increasing microwave emission (Chang and Gloersen, 1975; Zwally and Fiegles, 1994). By applying an empirical threshold with an appropriate surface melt detecting algorithm (Torinesi et al., 2003), the number of melt days and the spatial extent of surface melt can be detected (e.g. Torinesi et al., 2003; Picard and Fily, 2006). This satellite observational approach has been developed and used for Antarctic surface melt investigations (e.g. Picard and Fily, 2006; Johnson et al., 2022), showing it as a valuable and powerful tool that can be used to study and understand the surface melt frequency in Antarctica on both continental and regional scales (Johnson et al., 2022). However, this approach does not allow melt volume to be retrieved.

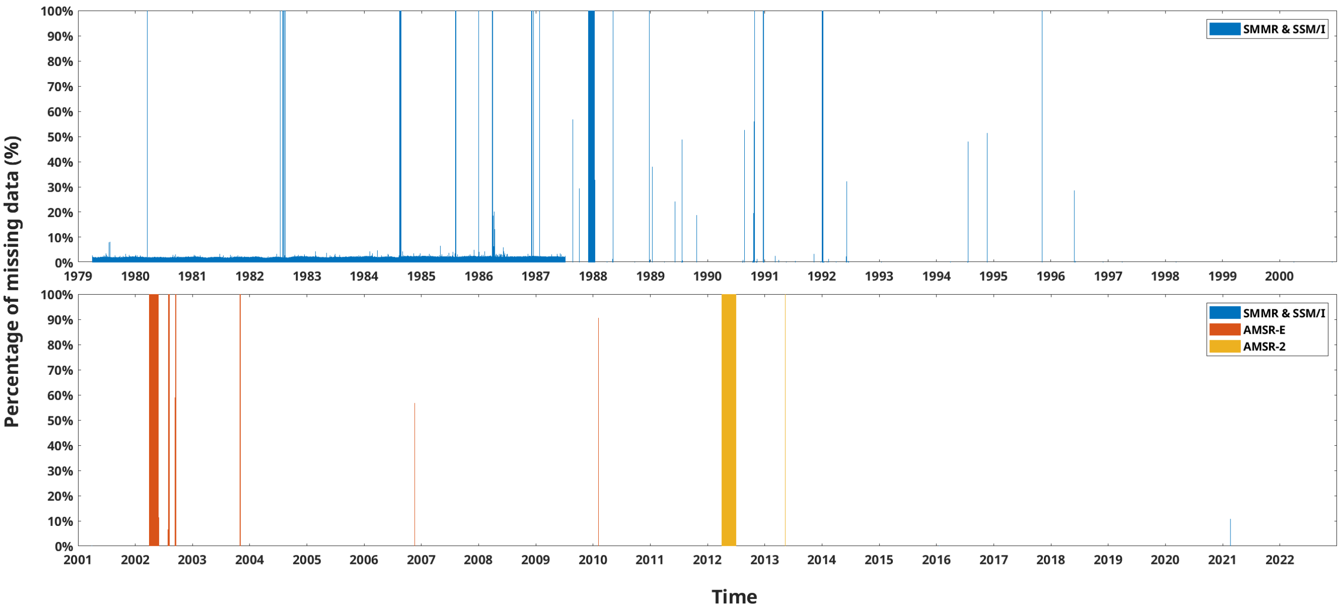

Figure A1. Daily percentage of missing data for satellite estimates. Satellite SMMR and SSM/I covers the period from 1979-04-01 to 2021-03-31. Satellite AMSR-E covers the period from 2002-04-01 to 2011-03-31. Satellite AMSR-2 covers the period from 2012-04-01 to 2021-12-31.

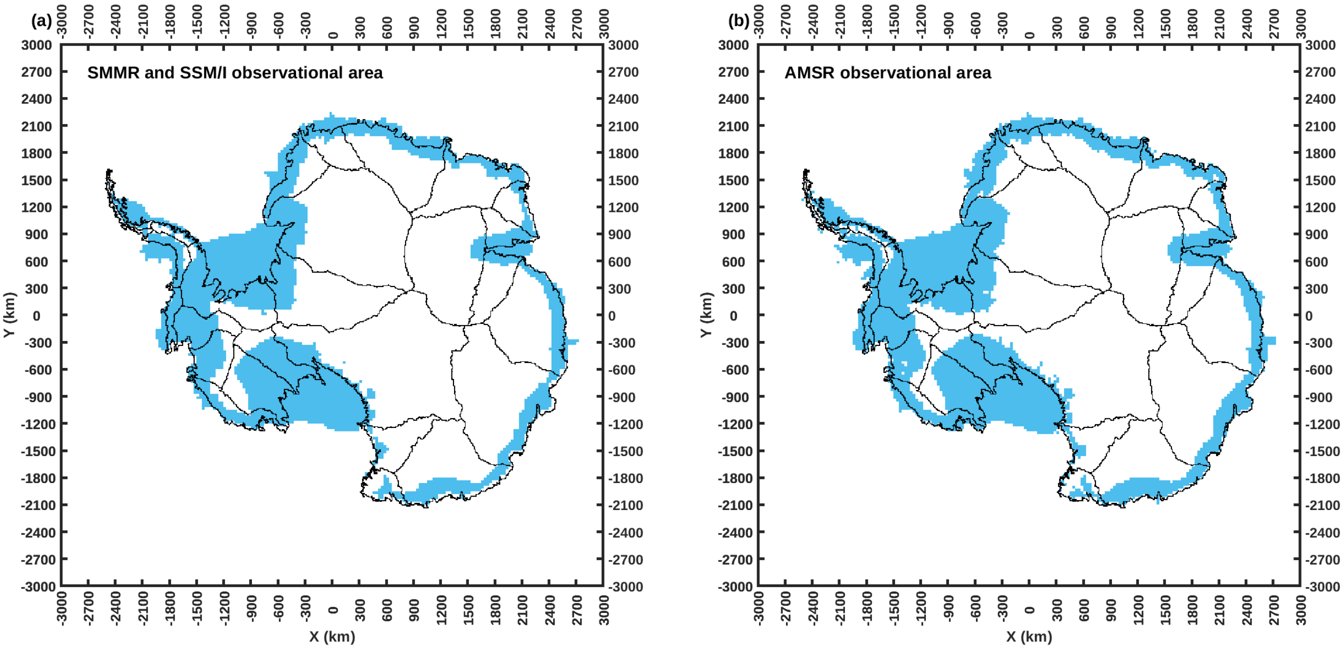

Figure A2. (a) mask of the satellite SMMR and SSM/I observational area. (b) mask of the satellite AMSR (AMSR-E and AMSR-2) observational area. Both masks are bilinearly remapped to the 30×30 km^2 polar stereographic grid.

Appendix B: Temperature-melt relationship

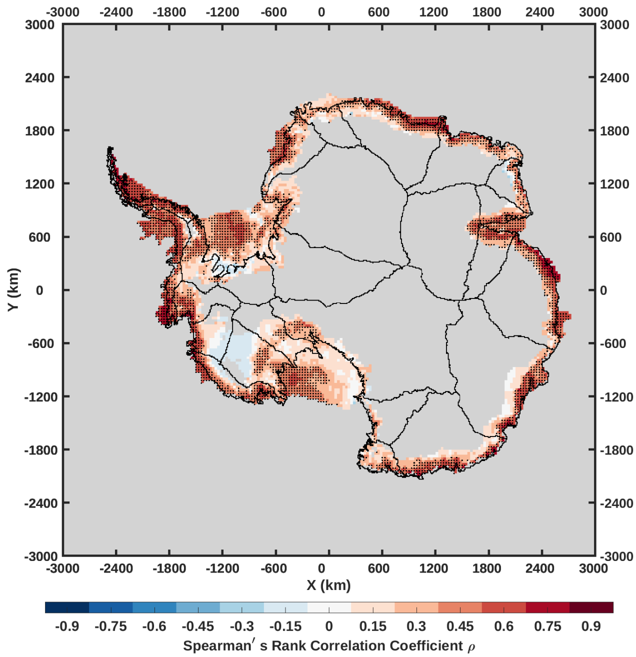

Figure B1. Correlation map between the mean DJF ERA5 2-m air temperature and the RACMO2.3p2 annual surface melt amount for the period from 1979/1980 to 2019/2020. It is calculated by the Spearman's rank correlation coefficient on each cell. Black dots mark the cells where the correlations are statistically significant ($p < 0.05$). Grey cells are either outside our research area (as shown in Figure 1) or have not melted ever during the period.

The positive relationship between 2-m air temperature and surface melt on Antarctic ice shelves (Trusel et al., 2015) allows us to use temperature to empirically estimate Antarctic surface melt via the PDD model. To assess this positive relationship, we calculate the Spearman's rank correlation between the mean summer (DJF) ERA5 2-m air temperature and the RACMO2.3p2 annual surface melt amount for the period from 1979/1980 to 2019/2020. Figure 3 indicates that most of the cells in Antarctic ice shelves and drainage basin coastal zones, apart from the Ross Ice Shelf or nearby basins (17, 18 and 19), have statistically significant ($p < 0.05$) positive correlations. Although the interior basins 19, 20 and 21 show negative correlations without statistical significance ($p \geq 0.05$), the annual melt there is negligible compared to the ice shelves and coastal areas. Overall, the correlation map shows a result consistent with Trusel et al. (2015): Antarctic ice-shelf near-surface temperature and surface melt are positively correlated, which allows us to empirically construct a temperature-index model to explore surface melt in Antarctica and especially Antarctic ice shelves.

Appendix C: ERA5 DJF 2-m temperature trend

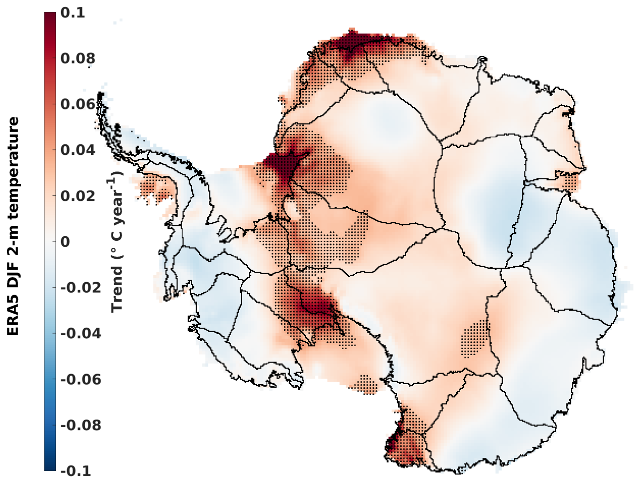

Figure C1. Trend of the mean DJF ERA5 2-m temperature on each computing cell during the period 1979/1980–2019/2020. Black dots mark the trends that are statistically significant ($p < 0.05$).

Appendix D: 3-fold CV T_0 Member 1

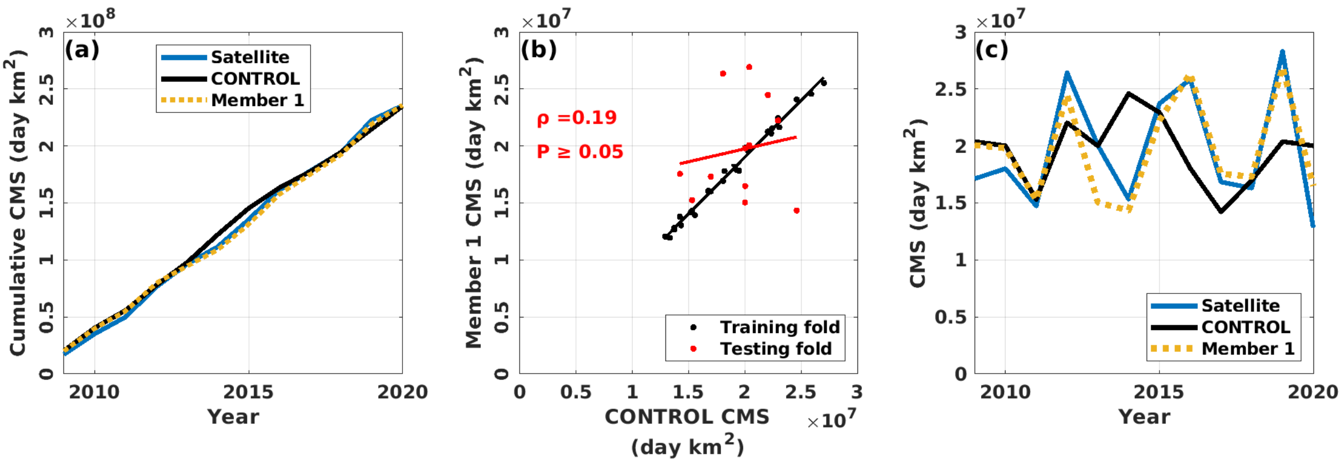

Figure D1. (a) and (b) are same as the Figure 7(m) and (s). (c) time series of the CMS for satellite estimates, CONTROL and Member 1 during the testing fold period.

Author contributions. YZ, NRG and AG conceived the study. YZ performed the analysis and prepared the original draft of the paper. GP and MLL provided satellite products. All authors contributed to writing the paper.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. YZ and NRG are supported by the Royal Society of New Zealand, award RDF-VUW1501. NRG and AG are supported by Ministry for Business Innovation and Employment, Grant/Award Number ANTA1801 ("Antarctic Science Platform"). NRG acknowledges support from Ministry for Business Innovation and Employment, Grant/Award Number RTUV1705 ("NZSeaRise").

References

Barrand, N. E., Vaughan, D. G., Steiner, N., Tedesco, M., Kuipers Munneke, P., Van Den Broeke, M. R., and Hosking, J. S.: Trends in Antarctic Peninsula surface melting conditions from observations and regional climate modeling, Journal of Geophysical Research: Earth Surface, 118, 315–330, 2013.

Bell, R. E., Banwell, A. F., Trusel, L. D., and Kingslake, J.: Antarctic surface hydrology and impacts on ice-sheet mass balance, Nature Climate Change, 8, 1044–1052, 2018.

Braithwaite, R. J.: Positive degree-day factors for ablation on the Greenland ice sheet studied by energy-balance modelling, Journal of Glaciology, 41, 153–160, 1995.

Chang, T. and Gloersen, P.: Microwave emission from dry and wet snow, in: Operational Applications of Satellite Snowcover Observations: The Proceedings of a Workshop Held August 18-20, 1975 at the Waystation, South Lake Tahoe, California, edited by Rango, A., Aeronautics, U. S. N., Administration, S., and University of Nevada, R., NASA SP, Scientific and Technical Information Office, National Aeronautics and Space Administration, https://books.google.co.nz/books?id=jEsCAAAAIAAJ, 1975.

Clem, K. R., Bozkurt, D., Kennett, D., King, J. C., and Turner, J.: Central tropical Pacific convection drives extreme high temperatures and surface melt on the Larsen C Ice Shelf, Antarctic Peninsula, Nature Communications, 13, 1–13, 2022.

Colosio, P., Tedesco, M., Ranzi, R., and Fettweis, X.: Surface melting over the Greenland ice sheet derived from enhanced resolution passive microwave brightness temperatures (1979–2019), The Cryosphere, 15, 2623–2646, 2021.

Costi, J., Arigony-Neto, J., Braun, M., Mavlyudov, B., Barrand, N. E., Da Silva, A. B., Marques, W. C., and Simoes, J. C.: Estimating surface melt and runoff on the Antarctic Peninsula using ERA-Interim reanalysis data, Antarctic Science, 30, 379–393, 2018.

Deo, R. C., Syktus, J., McAlpine, C., Lawrence, P., McGowan, H., and Phinn, S. R.: Impact of historical land cover change on daily indices of climate extremes including droughts in eastern Australia, Geophysical Research Letters, 36, 2009.

Fausto, R. S., Ahlstrøm, A. P., Van As, D., and Steffen, K.: Present-day temperature standard deviation parameterization for Greenland, Journal of Glaciology, 57, 1181–1183, 2011.

Golledge, N. R., Everest, J. D., Bradwell, T., and Johnson, J. S.: Lichenometry on adelaide island, antarctic peninsula: size-frequency studies, growth rates and snowpatches, Geografiska Annaler: Series A, Physical Geography, 92, 111–124, 2010.

Gossart, A., Helsen, S., Lenaerts, J., Broucke, S. V., Van Lipzig, N., and Souverijns, N.: An evaluation of surface climatology in state-of-the-art reanalyses over the Antarctic Ice Sheet, Journal of Climate, 32, 6899–6915, 2019.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., et al.: ERA5 hourly data on pressure levels from 1979 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 2018a.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., et al.: ERA5 hourly data on single levels from 1979 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 10, 2018b.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, 2020.

Hock, R.: Temperature index melt modelling in mountain areas, Journal of hydrology, 282, 104–115, 2003.

Hock, R.: Glacier melt: a review of processes and their modelling, Progress in physical geography, 29, 362–391, 2005.

Ismail, M. F., Bogacki, W., Disse, M., Schäfer, M., and Kirschbauer, L.: Estimating degree-day factors of snow based on energy flux components, The Cryosphere, 17, 211–231, 2023.

Jakobs, C. L., Reijmer, C. H., Smeets, C. P., Trusel, L. D., Van De Berg, W. J., Van Den Broeke, M. R., and Van Wessem, J. M.: A benchmark dataset of in situ Antarctic surface melt rates and energy balance, Journal of Glaciology, 66, 291–302, 2020.

Johnson, A., Hock, R., and Fahnestock, M.: Spatial variability and regional trends of Antarctic ice shelf surface melt duration over 1979–2020 derived from passive microwave data, Journal of Glaciology, 68, 533–546, 2022.

Jowett, A., Hanna, E., Ng, F., Huybrechts, P., Janssens, I., et al.: A new spatially and temporally variable sigma parameter in degree-day melt modelling of the Greenland ice sheet 1870–2013, The Cryosphere Discussions, 9, 5327–5371, 2015.

Kingslake, J., Ely, J. C., Das, I., and Bell, R. E.: Widespread movement of meltwater onto and across Antarctic ice shelves, Nature, 544, 349–352, 2017.

Kittel, C., Amory, C., Agosta, C., Jourdain, N. C., Hofer, S., Delhasse, A., Doutreloup, S., Huot, P.-V., Lang, C., Fichefet, T., et al.: Diverging future surface mass balance between the Antarctic ice shelves and grounded ice sheet, The Cryosphere, 15, 1215–1236, 2021.

Kittel, C., Amory, C., Hofer, S., Agosta, C., Jourdain, N. C., Gilbert, E., Le Toumelin, L., Vignon, É., Gallée, H., and Fettweis, X.: Clouds drive differences in future surface melt over the Antarctic ice shelves, The Cryosphere, 16, 2655–2669, 2022.

Kuipers Munneke, P., Picard, G., Van Den Broeke, M., Lenaerts, J., and Van Meijgaard, E.: Insignificant change in Antarctic snowmelt volume since 1979, Geophysical Research Letters, 39, 2012.

Lanzante, J. R.: Testing for differences between two distributions in the presence of serial correlation using the Kolmogorov–Smirnov and Kuiper's tests, International Journal of Climatology, 41, 6314–6323, 2021.

Larour, E., Seroussi, H., Morlighem, M., and Rignot, E.: Continental scale, high order, high spatial resolution, ice sheet modeling using the Ice Sheet System Model (ISSM), Journal of Geophysical Research: Earth Surface, 117, 2012.

Lee, J. R., Raymond, B., Bracegirdle, T. J., Chadès, I., Fuller, R. A., Shaw, J. D., and Terauds, A.: Climate change drives expansion of Antarctic ice-free habitat, Nature, 547, 49–54, 2017.

Lenaerts, J., Lhermitte, S., Drews, R., Ligtenberg, S., Berger, S., Helm, V., Smeets, C., Van Den Broeke, M., Van De Berg, W. J., Van Meijgaard, E., et al.: Meltwater produced by wind–albedo interaction stored in an East Antarctic ice shelf, Nature climate change, 7, 58–62, 2017.

Liu, H., Wang, L., and Jezek, K. C.: Spatiotemporal variations of snowmelt in Antarctica derived from satellite scanning multichannel microwave radiometer and Special Sensor Microwave Imager data (1978–2004), Journal of Geophysical Research: Earth Surface, 111, 2006.

Maraun, D. and Widmann, M.: Cross-validation of bias-corrected climate simulations is misleading, Hydrology and Earth System Sciences, 22, 4867–4873, 2018.

Mason, S. J.: Understanding forecast verification statistics, Meteorological Applications: A journal of forecasting, practical applications, training techniques and modelling, 15, 31–40, 2008.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C., Riahi, K., et al.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic change, 109, 213–241, 2011.

Mernild, S. H., Mote, T. L., and Liston, G. E.: Greenland ice sheet surface melt extent and trends: 1960–2010, Journal of Glaciology, 57, 621–628, 2011.

Mottram, R., Hansen, N., Kittel, C., van Wessem, J. M., Agosta, C., Amory, C., Boberg, F., van de Berg, W. J., Fettweis, X., Gossart, A., et al.: What is the surface mass balance of Antarctica? An intercomparison of regional climate model estimates, The Cryosphere, 15, 3751–3784, 2021.

NSIDC: National Snow and Ice Data Center Polar Stereographic Grid Definitions, https://nsidc.org/data/polar-stereo/ps_grids.html, 2022.

Ohmura, A.: Physical basis for the temperature-based melt-index method, Journal of applied Meteorology, 40, 753–761, 2001.

Picard, G.: Snow status (wet/dry) in Antarctica from SMMR, SSM/I, AMSR-E and AMSR2 passive microwave radiometers, Dataset, available online at https://doi.org/10.18709/perscido.2022.09.ds376, 2022.

Picard, G. and Fily, M.: Surface melting observations in Antarctica by microwave radiometers: Correcting 26-year time series from changes in acquisition hours, Remote sensing of environment, 104, 325–336, 2006.

Picard, G., Fily, M., and Gallée, H.: Surface melting derived from microwave radiometers: a climatic indicator in Antarctica, Annals of Glaciology, 46, 29–34, 2007.

Reeh, N.: Parameterization of melt rate and surface temperature in the Greenland ice sheet, Polarforschung, 59, 113–128, 1991.

Ryan, J., Smith, L., Van As, D., Cooley, S., Cooper, M., Pitcher, L., and Hubbard, A.: Greenland Ice Sheet surface melt amplified by snowline migration and bare ice exposure, Science Advances, 5, eaav3738, 2019.

Schulzweida, U.: CDO User Guide, https://doi.org/10.5281/zenodo.5614769, 2021.

Scott, R. C., Nicolas, J. P., Bromwich, D. H., Norris, J. R., and Lubin, D.: Meteorological drivers and large-scale climate forcing of West Antarctic surface melt, Journal of Climate, 32, 665–684, 2019.

Sellevold, R. and Vizcaino, M.: First application of artificial neural networks to estimate 21st century Greenland ice sheet surface melt, Geophysical Research Letters, 48, e2021GL092 449, 2021.

Stokes, C. R., Abram, N. J., Bentley, M. J., Edwards, T. L., England, M. H., Foppert, A., Jamieson, S. S., Jones, R. S., King, M. A., Lenaerts, J. T., et al.: Response of the East Antarctic Ice Sheet to past and future climate change, Nature, 608, 275–286, 2022.

Stone, M.: Cross-validation and multinomial prediction, Biometrika, 61, 509–515, 1974.

Tedesco, M. and Monaghan, A. J.: An updated Antarctic melt record through 2009 and its linkages to high-latitude and tropical climate variability, Geophysical Research Letters, 36, 2009.

Tetzner, D., Thomas, E., and Allen, C.: A validation of ERA5 reanalysis data in the Southern Antarctic Peninsula—Ellsworth land region, and its implications for ice core studies, Geosciences, 9, 289, 2019.

Torinesi, O., Fily, M., and Genthon, C.: Variability and trends of the summer melt period of Antarctic ice margins since 1980 from microwave sensors, Journal of Climate, 16, 1047–1060, 2003.

Trusel, L., Frey, K. E., and Das, S. B.: Antarctic surface melting dynamics: Enhanced perspectives from radar scatterometer data, Journal of Geophysical Research: Earth Surface, 117, 2012.

Trusel, L. D., Frey, K. E., Das, S. B., Munneke, P. K., and Van Den Broeke, M. R.: Satellite-based estimates of Antarctic surface meltwater fluxes, Geophysical Research Letters, 40, 6148–6153, 2013.

Trusel, L. D., Frey, K. E., Das, S. B., Karnauskas, K. B., Munneke, P. K., Van Meijgaard, E., and Van Den Broeke, M. R.: Divergent trajectories of Antarctic surface melt under two twenty-first-century climate scenarios, Nature Geoscience, 8, 927–932, 2015.

Turner, J., Lu, H., White, I., King, J. C., Phillips, T., Hosking, J. S., Bracegirdle, T. J., Marshall, G. J., Mulvaney, R., and Deb, P.: Absence of 21st century warming on Antarctic Peninsula consistent with natural variability, Nature, 535, 411–415, 2016.

Turton, J. V., Kirchgaessner, A., Ross, A. N., King, J. C., and Kuipers Munneke, P.: The influence of föhn winds on annual and seasonal surface melt on the Larsen C Ice Shelf, Antarctica, The Cryosphere, 14, 4165–4180, 2020.

van den Broeke, M., Bus, C., Ettema, J., and Smeets, P.: Temperature thresholds for degree-day modelling of Greenland ice sheet melt rates, Geophysical Research Letters, 37, 2010.

Van den Broeke, M., Smeets, C., and Van de Wal, R.: The seasonal cycle and interannual variability of surface energy balance and melt in the ablation zone of the west Greenland ice sheet, The Cryosphere, 5, 377–390, 2011.

Van Wessem, J. M., Van De Berg, W. J., Noël, B. P., Van Meijgaard, E., Amory, C., Birnbaum, G., Jakobs, C. L., Krüger, K., Lenaerts, J., Lhermitte, S., et al.: Modelling the climate and surface mass balance of polar ice sheets using RACMO2–Part 2: Antarctica (1979–2016), The Cryosphere, 12, 1479–1498, 2018.

Wake, L. and Marshall, S.: Assessment of current methods of positive degree-day calculation using in situ observations from glaciated regions, Journal of Glaciology, 61, 329–344, 2015.

Wille, J. D., Favier, V., Dufour, A., Gorodetskaya, I. V., Turner, J., Agosta, C., and Codron, F.: West Antarctic surface melt triggered by atmospheric rivers, Nature Geoscience, 12, 911–916, 2019.

Wilton, D. J., Jowett, A., Hanna, E., Bigg, G. R., Van Den Broeke, M. R., Fettweis, X., and Huybrechts, P.: High resolution (1 km) positive degree-day modelling of Greenland ice sheet surface mass balance, 1870–2012 using reanalysis data, Journal of Glaciology, 63, 176–193, 2017.

Winkelmann, R., Martin, M. A., Haseloff, M., Albrecht, T., Bueler, E., Khroulev, C., and Levermann, A.: The Potsdam Parallel Ice Sheet Model (PISM-PIK) Part 1: Model description, The Cryosphere, 5, 715–726, https://doi.org/10.5194/tc-5-715-2011, 2011.

Zheng, Y., Jong, L. M., Phipps, S. J., Roberts, J. L., Moy, A. D., Curran, M. A., and van Ommen, T. D.: Extending and understanding the South West Western Australian rainfall record using a snowfall reconstruction from Law Dome, East Antarctica, Climate of the Past, 17, 1973–1987, 2021.

Zhu, J., Xie, A., Qin, X., Wang, Y., Xu, B., and Wang, Y.: An assessment of ERA5 reanalysis for antarctic near-surface air temperature, Atmosphere, 12, 217, 2021.

Zou, X., Bromwich, D. H., Montenegro, A., Wang, S.-H., and Bai, L.: Major surface melting over the Ross Ice Shelf part II: surface energy balance, Quarterly Journal of the Royal Meteorological Society, 2021.

Zwally, H. J. and Fiegles, S.: Extent and duration of Antarctic surface melting, Journal of Glaciology, 40, 463–475, 1994.

Zwally, H. J., Giovinetto, M. B., Beckley, M. A., and Saba, J. L.: Antarctic and Greenland drainage systems, GSFC cryospheric sciences laboratory, 2012.