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 melting is one of the primary drivers of ice shelf collapse in Antarctica and 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 melting. 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, the current understanding of surface melt in Antarctica remains limited in terms of the uncertainties of quantifying surface melt and understanding the driving processes of surface melt in past, present, and future contexts. Here, we construct a novel grid cell-level spatially-distributed positive degree-day (PDD) model, forced with 2-m air temperature reanalysis data, and spatially parameterized 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 the time window chosen for parameterization. We conduct a sensitivity experiment by adding ±10% to the training data (satellite estimates and SEB model outputs) used for PDD parameterization, and a sensitivity experiment by adding constant temperature perturbations (+1 °C, +2 °C, +3 °C, +4 °C, and +5 °C) to the 2-m air temperature field to force the PDD model. We find that the PDD melt extent and amounts change analogously to the variations in the training data with steady statistically significant correlations, and the PDD melt amounts increase nonlinearly with the temperature perturbations, demonstrating the consistency of our parameterization and the applicability of the PDD model to warmer 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 ice sheet net mass balance 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 statistically significant negative trend for the annual melt days (Banwell et al., 2023) and 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 that can increase ice shelf vulnerability, as meltwater can pond, drain and further contribute to the structural weakness of ice shelves (Glasser and Scambos, 2008; Bell et al., 2018; Stokes et al., 2022). However, the roles of surface meltwater production in relation to ice shelf hydrofracture, surface rivers acting as buffers and ice shelf surface hydrology, are currently less understood over Antarctica than Greenland (Bell et al., 2018). This is concerning as surface melting will likely become an increasingly important player in the Antarctic environment through this century and the next. Surface melting will not only impact the dynamics of the ice shelves and ice sheet through meltwater production (e.g. Bell et al., 2018), but will also impact the habitat of the Antarctic biodiversity (Lee et al., 2017).

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 are employed in Regional Climate Models such as the Regional Atmospheric Climate MOdel (RACMO) (Van Wessem et al., 2018) and Modèle Atmosphérique Régional (MAR) (Agosta et al., 2019). PDD models are employed in ice sheet models such as the SImulation COde for POLythermal Ice Sheets (SICOPOLIS) (Nowicki et al., 2013), Ice Sheet System Model (ISSM) (Larour et al., 2012), and Parallel Ice Sheet Model (PISM) (Winkelmann et al., 2011). 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 therefore better 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 meltwater production.

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 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). Spatial and temporal variability in DDF can result from topographic variation, such as the gradient of elevation which affects albedo and direct input solar radiation (Hock, 2003), and seasonal variations in radiation. 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 is 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) (Van Wessem et al., 2018) surface melt amount simulations. We also use the same data and method to parameterize a spatially uniform PDD model. We then examine the distributions of melt days and melt amount from PDD outputs against satellite melt day estimates and RACMO2.3p2 melt amount simulations, respectively. 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 $^{\circ}\mathrm{C}$ annually and at 1.06 $^{\circ}\mathrm{C}$ 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). The dataset contains daily estimates as a binary of melt or no-melt on a $25\times25\ \mathrm{km}^2$ southern polar stereographic grid. The dataset is obtained by applying the melt detecting algorithm (Torinesi et al., 2003; Picard and Fily, 2006) to detect the presence of surface liquid water 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 $12.5\times12.5\ \mathrm{km}^2$ 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

To parameterize the DDF for the PDD model, we compare our ERA5 forced numerical experiments to the Antarctic surface melt simulations from RACMO2.3p2 (Van Wessem et al., 2018). 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).

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

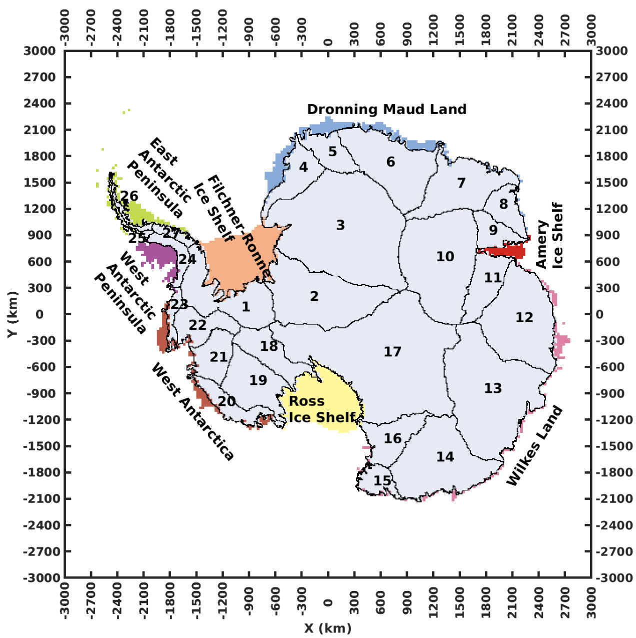

Figure 1. The research domain and 27 Antarctic drainage basins (Zwally et al., 2012) used in this study.

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 commonly 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. However, because of the temperature dependency, the robustness of the temperature-index approach is therefore attributed to the temperature-melt correlation.

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:

$$\begin{aligned} &\sum_{i=1}^{\mathrm{day}} \mathrm{M} = \frac{1}{24}\mathrm{DDF}\sum_{i=1}^{\mathrm{day}}\sum_{j=1}^{24}\mathrm{T}^{\star} \\ &\mathrm{T}^{\star} = \begin{cases} \mathrm{T}-\mathrm{T_0} & \text{if } \mathrm{T}-\mathrm{T_0} > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned} \tag{2}$$

where T is the hourly temperature and $\mathrm{T_0}$ is the threshold temperature.

3.2 Model parameterisation

3.2.1 Threshold temperature $\mathbf{T_0}$

To parameterize the threshold temperature ($\mathrm{T_0}$) 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 $\mathrm{T_0}$ ranging from -10.0 $^{\circ}\mathrm{C}$ to +5.0 $^{\circ}\mathrm{C}$ with a 0.1 $^{\circ}\mathrm{C}$ 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 $\mathrm{T_0}$. 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 $\mathrm{T_0}$ 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:

$$
\begin{aligned}
&\text{Annual number of melt days} = \sum_{i=t_1}^{t_2} \mathrm{MD}^\star \\
&\qquad\qquad t_1 = 01 - \text{April} - \text{Year} \\
&\qquad\qquad t_2 = 31 - \text{March} - (\text{Year+1}) \\
&\qquad\quad \mathrm{MD}^\star = \begin{cases} 1 & \text{if } \mathrm{T} - \mathrm{T_0} > 0 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
\tag{3}
$$

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 $\mathrm{T_0}$ experiment.

In order to obtain the optimal $\mathrm{T_0}$, 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 the overlapping years. As we treat each computing cell individually, all calculations are carried out on each cell independently in each iteration ($\mathrm{T_0}$ experiment). Although these three satellite products have different time periods (Table 1), 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 $\mathrm{T_0}$ of each computing cell where the $\mathrm{T_0}$ experiment has the minimal RMSE. If there are multiple $\mathrm{T_0}$ experiments that have same minimal RMSE for their computing cell, we calculate the mean of those $\mathrm{T_0}$ as the optimal $\mathrm{T_0}$ (this only happens on the cells that have very low melt days).

3.2.2 Degree Day Factor DDF

The DDF is a scaling parameter that controls the meltwater production and is related to all terms of the SEB (Hock, 2005). To parameterize the DDF for our PDD model, we substitute the optimal $\mathrm{T_0}$ 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 determine 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

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. We use two-sample Kolmogorov–Smirnov test (hereafter two-sample KS test) to evaluate the dissimilarity between the PDD results and RACMO2.3p2 melt volume outputs at a confidence level of 5%. 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

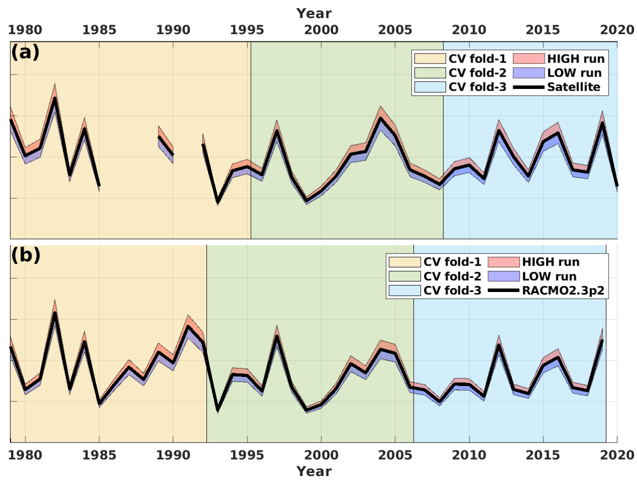

Figure 2. Schematic overview of the time periods for each CV folders and the HIGH, LOW sensitivity experiments. (a) is for satellite estimates and PDD melt day calculations. (b) is for RACMO2.3p2 simulations and PDD melt amount calculations.

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 and Table 2). 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 be-

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.

tween 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 independent members. In Member 1, we take the first and second fold to parameterize the PDD model and test the model on the third fold. 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 RACMO2.3p2 surface melt amount but the folds are divided into two 14-year folds and a 13-year fold (Figure 2b and Table 2).

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 the inconsistency in the characteristics of satellite sensors caused by 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 cumulative melting surface (CMS) outputs to perturbations in satellite estimates. 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 et al., 2012). 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) 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.

To assess the applicability of our PDD model in simulating melt under warmer climate scenarios, we conduct temperature-melt sensitivity experiments. To do this, we add constant temperature perturbations of +1 °C, +2 °C, +3 °C, +4 °C, and +5 °C to the whole 43-year (1979/1980 to 2021/2022) ERA5 2-m air temperature field to force our PDD model.

4 Results and discussion

4.1 Optimal PDD parameters

Figure 3a shows the spatial distribution of the optimal T_0 values selected through 151 T_0 experiments conducted on each computing cell, based on the minimal RMSE criterion. The mean of all optimal T_0 is -2.32 °C. The majority of cells have a negative T_0, indicating that using T_0 = 0 °C as a melt threshold may substantially underestimate melt events, a finding consistent with other work (Jakobs et al., 2020).

The probability distribution of T_0 across all grid cells is approximately normal (Figure 3c). There is a small number of cells distributed below -5.5 °C which is around 1.96 standard deviations lower than the mean (-5.57 °C, 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). We speculate that this feature may be related to the presence of local rocks (e.g., Fricker et al., 2021; Spergel et al., 2021), or it could be a result of frequent surface melt events over the central Amery Ice Shelf (as suggested by the low T_0 value), which are likely to have a low intensity (as indicated by the low DDF value).

Figure 3b shows the spatial map of the optimal DDFs identified for each computing cell. We show that 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

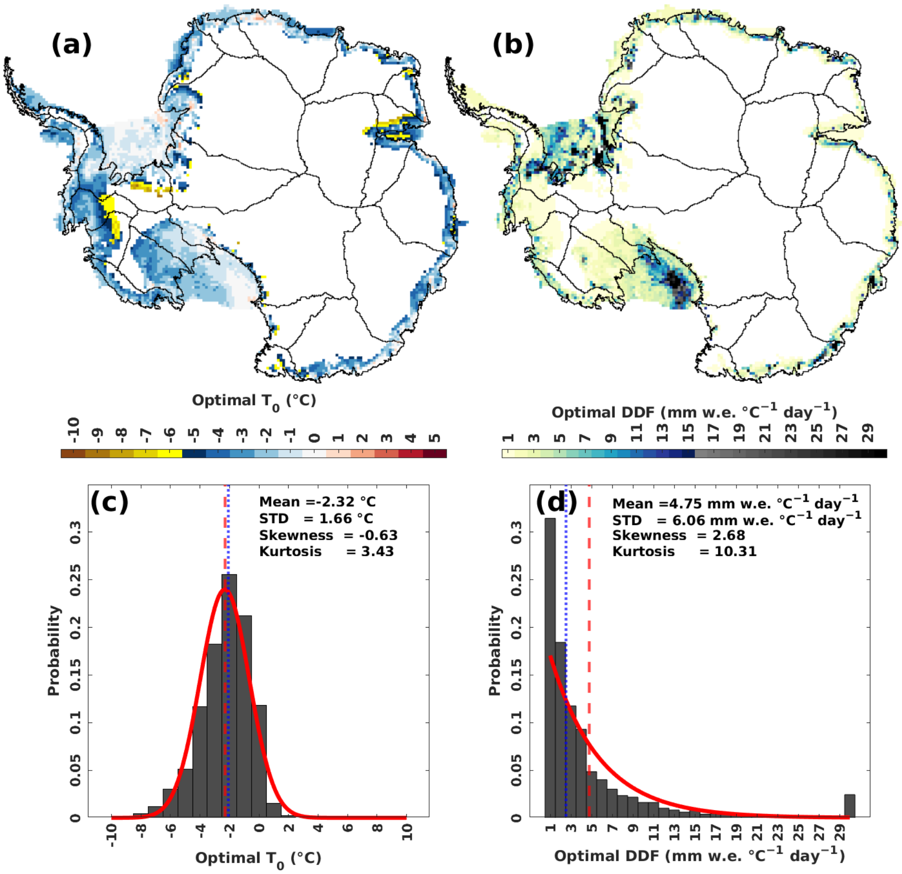

Figure 3. (a) The optimal T_0 ($^\circ$C) of each computing cell. (b) The optimal DDF ($\mathrm{mm\ w.e.\ ^\circ C^{-1}\ day^{-1}}$) for each computing cell. (c) Probability histogram of the optimal T_0 ($^\circ$C). Red curve is the fitted normal distribution. Red dashed vertical line is the mean of 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.

is rare. Figure 3d summarizes the statistics of DDFs. The probability distribution of the DDFs is asymmetrical and strongly right-skewed (Figure 3d).

We also use the same method and data to parameterize a spatially uniform PDD (hereafter, "uni-PDD") model (one T_0 and DDF for all computing cells, Appendix C). For convenience, we name the grid cell-level spatially-distributed PDD “dist-PDD”. The optimal T_0 for uni-PDD is -2.6 $^\circ$C and the optimal DDF is 1.9 $\mathrm{mm\ w.e.\ ^\circ C^{-1}\ day^{-1}}$ (Figure C1 in the Appendix C).

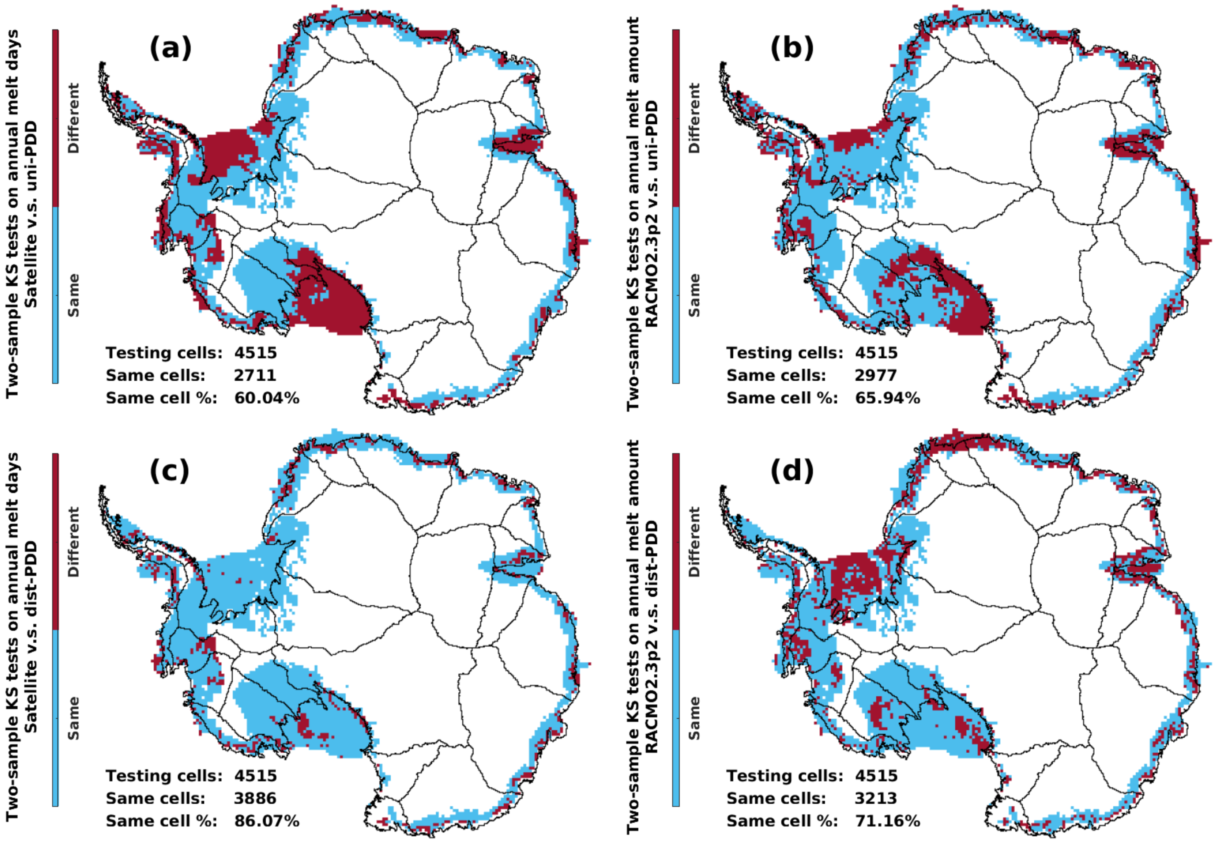

Figure 4. The two-sample KS test results. The two-sample KS tests are performed individually for each of the 4515 computing cells. The test result "Same" means the tested cell is a same distribution cell where there is no statistically significant evidence for the rejection of the null hypothesis that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

4.2 Model evaluation

4.2.1 Goodness-of-fit

We evaluate the parameterized dist-PDD and uni-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 dist-PDD/ uni-PDD model-derived empirical distribution functions. Figure 4 shows the two-sample KS test results for each computing cell. The dist-PDD model improves the proportion of cells with the same distribution for melt days/ amount from 60.04%/ 65.94% to 86.07%/ 71.16%, respectively, compared to the uni-PDD model. Overall, the dist-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 (Figure 4c and d). Our dist-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, do not perform as well for both the surface melt days and amount (Figure 4c and d). It is especially encouraging that the PDD model performs well in the Antarctic Peninsula, given the fact that it is the region of Antarctica experiencing most intense surface melting both at the present (Trusel et al., 2013; Johnson et al., 2022) and in future projections (Trusel et al., 2015).

Table 3. Summary of the statistics for Figure 5c. The Spearman's ρ and P-value for dist-PDD/ uni-PDD CMS with the satellite CMS. Slope, R^2, RMSE and P-value for the Ordinary Least Squares (OLS) fit between dist-PDD/ uni-PDD CMS and satellite CMS. Note that 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. All the statistics are calculated over the period from 1979/1980 to 2020/2021 (with 1986/1987 to 1988/1989 and 1991/1992 omitted).

Member	Spearman's ρ	P-value	OLS slope	R^2	RMSE (day km^2)	P-value
uni-PDD v.s. satellite	0.4881	$P < 0.05$	0.3421	0.208	4.09×10^6	$P < 0.05$
dist-PDD v.s. satellite	0.5203	$P < 0.01$	0.3004	0.229	3.38×10^6	$P < 0.05$

Next, we evaluate the parameterized dist-dist-PDD/ uni-PDD model outputs for the whole of Antarctica. Firstly, we evaluate the parameterized optimal T_0 and its related dist-PDD/ uni-PDD outputs on the surface melt day. To do this, we calculate the CMS (day km^2) for satellite estimates and dist-PDD/ uni-PDD outputs, respectively. We show that in Figure 5a that the dist-PDD and satellite CMS time series are generally in good agreement on both the amplitude and the temporal variability, apart from a small number of years including from 1979/1980 to 1982/1983, the year 2014/2015, the year 2016/2017 and the year 2019/2020. Although there is a dist-PDD underestimation of cumulative CMS for the first decade (1980 to 1990), the cumulative CMS of dist-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 dist-PDD CMS is strongly statistically significant (Spearman's ρ = 0.5203, $p < 0.01$, Table 3). The probability histogram for biases between the dist-PDD and satellite CMS also indicates a good agreement between the dist-PDD and satellite CMS (Figure D1 in the Appendix D). The biases are distributed symmetrically around the mean which is approximated to zero (Figure D1).

Globally, we show that the accuracy of the PDD models on estimating the surface melt days has improved from the uni-PDD model to the dist-PDD model (Table 3 and Figure 5), and the dist-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 5). The computing cells that have relatively large disagreement between the mean annual melt days of dist-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

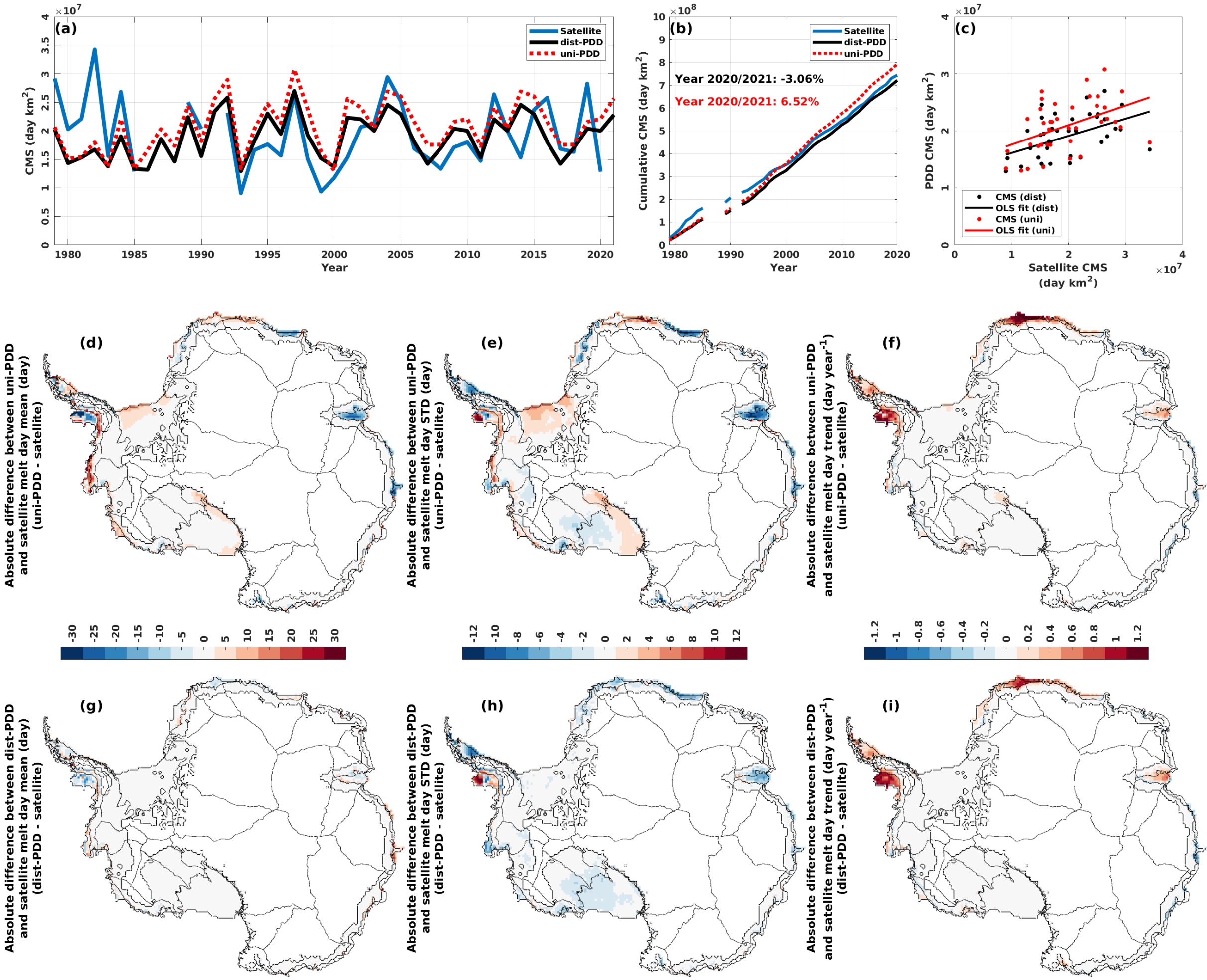

Figure 5. (a) time series for the cumulative melting surface (CMS) (day km^2) for satellite estimates during the period from 1979/1980 to 2020/2021 (with 1986/1987 to 1988/1989 and 1991/1992 omitted), and for dist-PDD/ uni-PDD outputs during the period from 1979/1980 to 2021/2022. (b) cumulative CMS for satellite estimates and dist-PDD/ uni-PDD outputs from 1979/1980 to 2020/2021 (with 1986/1987 to 1988/1989 and 1991/1992 omitted). (c) scatter plot and Ordinary Least Square (OLS) fit between satellite CMS and dist-PDD/ uni-PDD CMS. (d) to (i) absolute differences between mean, standard deviation (STD) and trend of dist-PDD/ uni-PDD outputs and satellite estimates on the annual melt days. Mean, STD and trend for the dist-PDD/ uni-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. 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.

+12.5 days). However, these cells with large absolute differences experience frequent surface melt (Figure D2a and d in the Appendix D), meaning that the relative differences in melt are low (Figure D2g). In addition, these cells only amount to around 5% of the total computing cells (Figure D1b), and overall for all computing cells, the mean of average differences between the dist-PDD and satellite annual melt days is approximately zero (-0.12 days, Figure D1b). It is not surprising that the dist-PDD model captures the main spatial patterns of melt, 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 absolute differences 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 absolute differences in their means, the relative differences are low (Figure D2h) and these cells amount to only a negligible proportion (less than 5%) of the total number of the computing cells (Figure D1b). However, there are around 20% of the computing cells that have -1 to -3 days of STD biases (Figure D1b), 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 (Figure 5h). The biases 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 5i and a slightly right-skewed probability histogram of trend biases with a positive mean (+0.04 day year^{-1}, Figure D1c).

Table 4. Summary of the statistics for Figure 6c. The Spearman's ρ and P-value for dist-PDD/ uni-PDD melt amount with the RACMO2.3p2 melt amount. Slope, R^2, RMSE and P-value for the Ordinary Least Squares (OLS) fit between dist-PDD/ uni-PDD melt amount and RACMO2.3p2 melt amount. All the statistics are calculated over the period from 1979/1980 to 2019/2020.

Member	Spearman's ρ	P-value	OLS slope	R^2	RMSE (mm w.e.)	P-value
uni-PDD v.s. RACMO2.3p2	0.7052	$P < 0.01$	0.9416	0.091	2.16×10^4	$P < 0.01$
dist-PDD v.s. RACMO2.3p2	0.8052	$P < 0.01$	0.5307	0.55	1.42×10^4	$P < 0.01$

Secondly, we evaluate the parameterized optimal DDF and the simulated surface melt amount. Similar to the negative biases between the dist-PDD and the satellite estimates for the CMS for the period from 1979/1980 to 1982/1983 (Figure 5a), the negative biases of dist-PDD against RACMO2.3p2 are also present when compared to 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 Southern Annular Mode (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 dist-PDD model for this extensive melt event is most likely explained by the absence of any substantial temperature anomaly in the ERA5 2-m temperature input (Figure E1 in the Appendix E), 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 dist-PDD parameters were defined based on fitting multi-decadal timeseries between dist-PDD experiments and satellite/ RACMO2.3p2 (Section 3.2.1 and 3.2.2), meaning that some inter/intra- annual signals may not be fully captured.

Apart from the 1982/1983 event, other negative biases from dist-PDD are also evident in the period from 1991/1992 to 1992/1993 (Figure 6a). However, we cannot compare this dist-PDD melt amount bias period to the dist-PDD CMS bias as

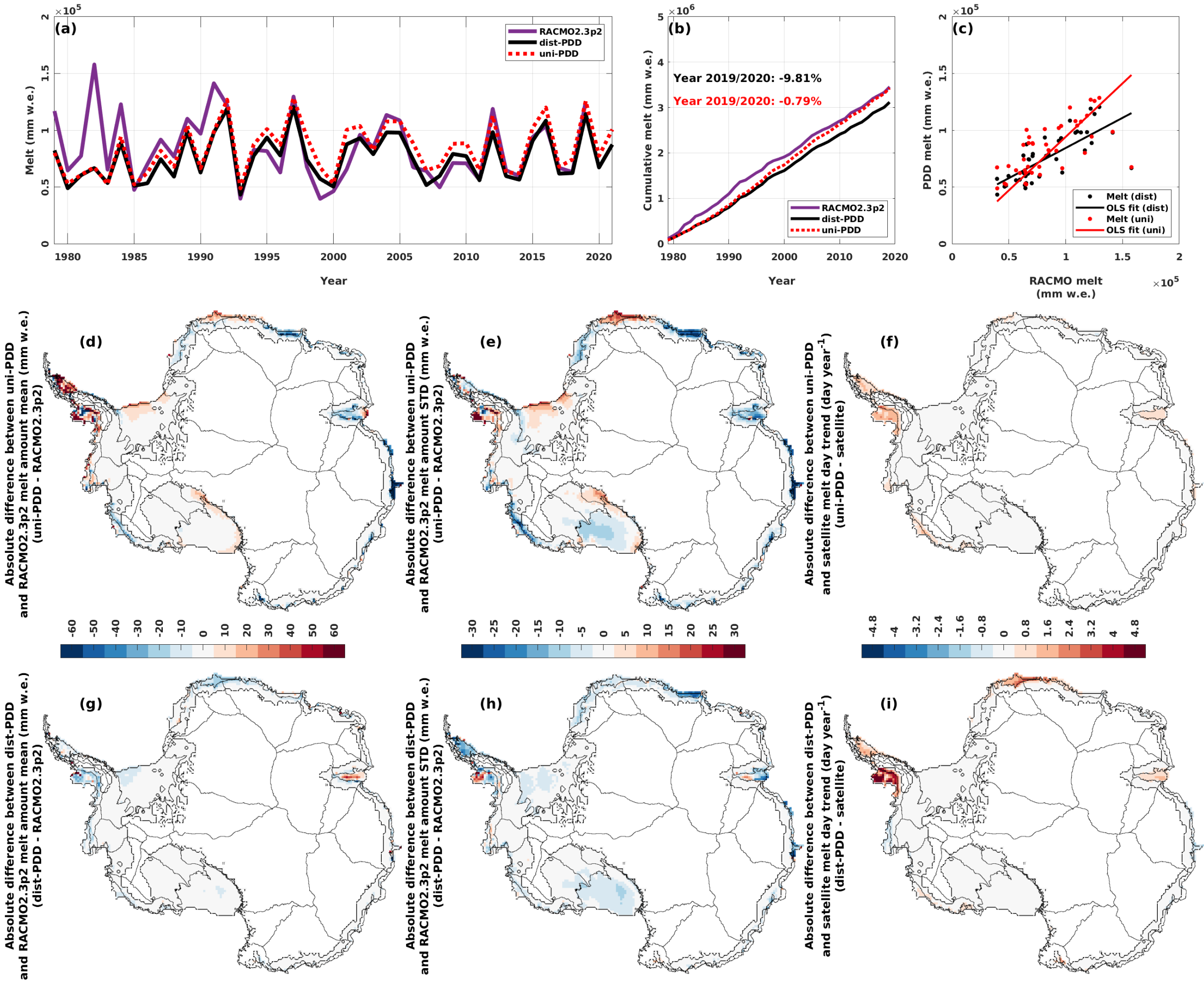

Figure 6. (a) time series for the annual melt amount (mm w.e.) for RACMO2.3p2 simulations during the period from 1979/1980 to 2019/2020, and for dist-PDD/ uni-PDD outputs during the period from 1979/1980 to 2021/2022. (b) cumulative annual melt amount for RACMO2.3p2 simulations and dist-PDD/ uni-PDD outputs from 1979/1980 to 2019/2020. (c) scatter plot and Ordinary Least Squares (OLS) fit between satellite annual melt amount and dist-PDD/ uni-PDD annual melt amount. (d) to (i) absolute differences between mean, standard deviation (STD) and trend of dist-PDD/ uni-PDD outputs and RACMO2.3p2 simulations on the annual melt amount. Mean, STD and trend for the dist-PDD/ uni-PDD outputs and satellite estimates are calculated over the period from 1979/1980 to 2019/2020, respectively.

the year 1991/1992 is omitted for all the analysis related to the satellite estimates due to the missing satellite data. Excluding these periods, the time series of annual melt amount of the dist-PDD outputs and RACMO2.3p2 simulations are generally in good agreement, especially after 1992/1993 when the two curves start to overlap (Figure 6a) whilst the dist-PDD-satellite CMSs show some disagreement (e.g. 1995/1996, 1999/2000, 2014/2015, 2016/2017 and 2019/2020, Figure 5a). It is also evident by the statistically significant strong positive correlation (Spearman's ρ = 0.8052, $p < 0.01$, Table 4) that the dist-

PDD is in a good agreement with RACMO2.3p2 annual melt amount. However, the probability histogram of dist-PDD melt biases is slightly left-skewed with a negative mean (-0.08 × 10^5 mm w.e., Figure D3 in the Appendix D) and the dist-PDD model underestimates around 9.81% for the 41-year integrated annual melt amount compared to RACMO2.3p2 (Figure 6b). Nevertheless, this underestimation on the 41-year integrated annual melt amount does not change through the past four decades, as we show that 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. Although the 41-year integrated annual melt amounts for 2019/2020 between uni-PDD and RACMO2.3p2 show very good agreement (-0.79%, as shown in Figure 6b), the two cumulative curves are not parallel. The uni-PDD curve diverges from the RACMO2.3p2 curve for around 15 years and then converges to RACMO2.3p2 for the rest of the time period (as shown in Figure 6b). This indicates that the uni-PDD model is not sufficiently flexible to accurately estimate surface melt amount.

Figure 6d to i show the spatial maps for the difference between the mean, STD and trend of the dist-PDD/ uni-PDD annual melt amount and RACMO2.3p2 mean annual melt amount for the period from 1979/1980 to 2019/2020. The spatial maps for the mean, STD and trend of the dist-PDD/ uni-PDD annual melt amount and RACMO2.3p2 mean annual melt amount for the same period are shown in Figure D4 in the Appendix D. Consistent with the PDD melt day estimates, using the dist-PDD model improves the accuracy of estimating surface melt amount compared to using spatially uniform PDD parameters. As shown in Figure 6g, h and i, 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 dist-PDD outputs and satellite estimates in Figure 5g, h and i. This indicates that the dist-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. Less than 5% of the total number of all computing cells are 15 mm w.e. below or above the bias on mean (Figure 6g). These cells are 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 total number of the computing cells bias -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 6h). The bias in trends between the dist-PDD and RACMO2.3p2 annual melt amount is similar to the bias in trends between the dist-PDD and satellite annual melt days, as they both have the same positive spatial bias patterns (Antarctic Peninsula, Dronning Maud Land and Amery Ice Shelf, Figure 5i and Figure 6i) and similar right-skewed probability histograms with positive means (Figure D1c and Figure D3c). This could be explained by other players driving surface melting, such as the 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 biases in trends are a reflection of the trend of the input temperature (Figure D5 in the Appendix D), because of the correlation 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 dist-PDD model itself.

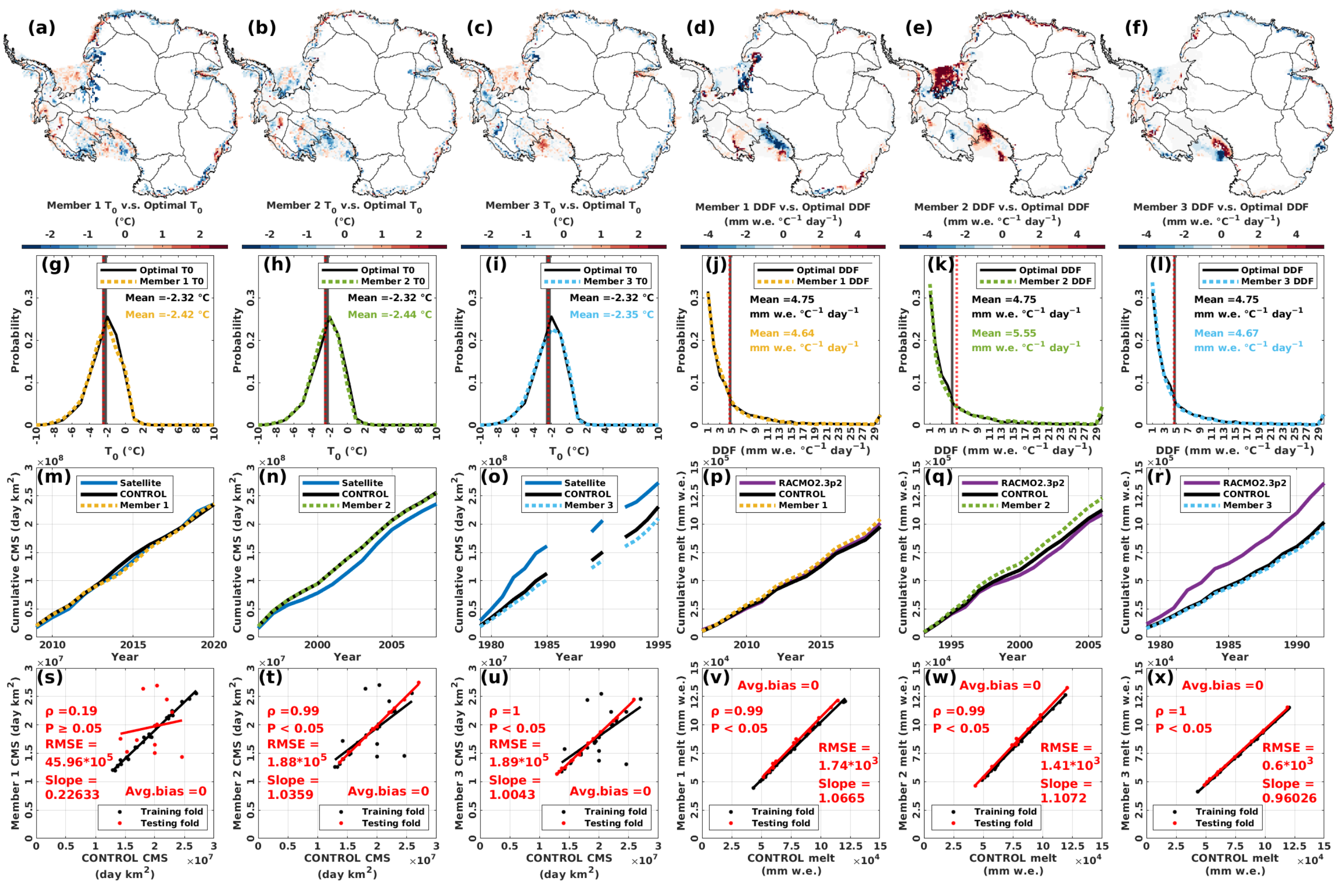

Figure 7. (a) to (f) differences between the T_0/ DDF parameterized in each member of the T_0/ DDF 3-fold CV and the optimal T_0/ DDF, respectively. (g) to (l) probability distributions for the T_0/ DDF of each T_0/ DDF 3-fold CV and the optimal T_0/ DDF, respectively. Black vertical lines indicate the 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, and the slope, RMSE and average bias for the OLS fit, for the testing fold between each member and the CONTROL are calculated, respectively. This analysis is based on dist-PDD.

4.2.2 Temporal dependency of the dist-PDD parameters

To evaluate our dist-PDD model in a temporal sense, we perform 3-fold CV for T_0 and DDF (as described in Section 3.3.2), respectively.

Figure 7 shows the results of the 3-fold CV on T_0 and DDF. We show that 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 dist-PDD model. Especially for the DDF members, we show that conspicuous changes in the values of the DDFs in 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 dist-PDD parameters may exist. However, this indication 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 show that 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 dist-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 dist-PDD model, and that the ability of our dist-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 dist-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 show that each member's dist-PDD estimates are significantly ($\rho \geq 0.99$, $p \leq 0.05$) correlated with the CONTROL dist-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 dist-PDD and the CONTROL dist-PDD (Figure 7g to r). The T_0 Member 1 dist-PDD estimates and dist-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 dist-PDD is parameterized by those dist-PDD CONTROL estimates. The T_0 Member 1 dist-PDD estimates and dist-PDD CONTROL estimates are not 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 F1, in the Appendix F. 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 explain 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 is parameterized by the full 38-year period. However, the data sample used to parameterize the Member 1 T_0 is only 2/3 the full data length used to estimate 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

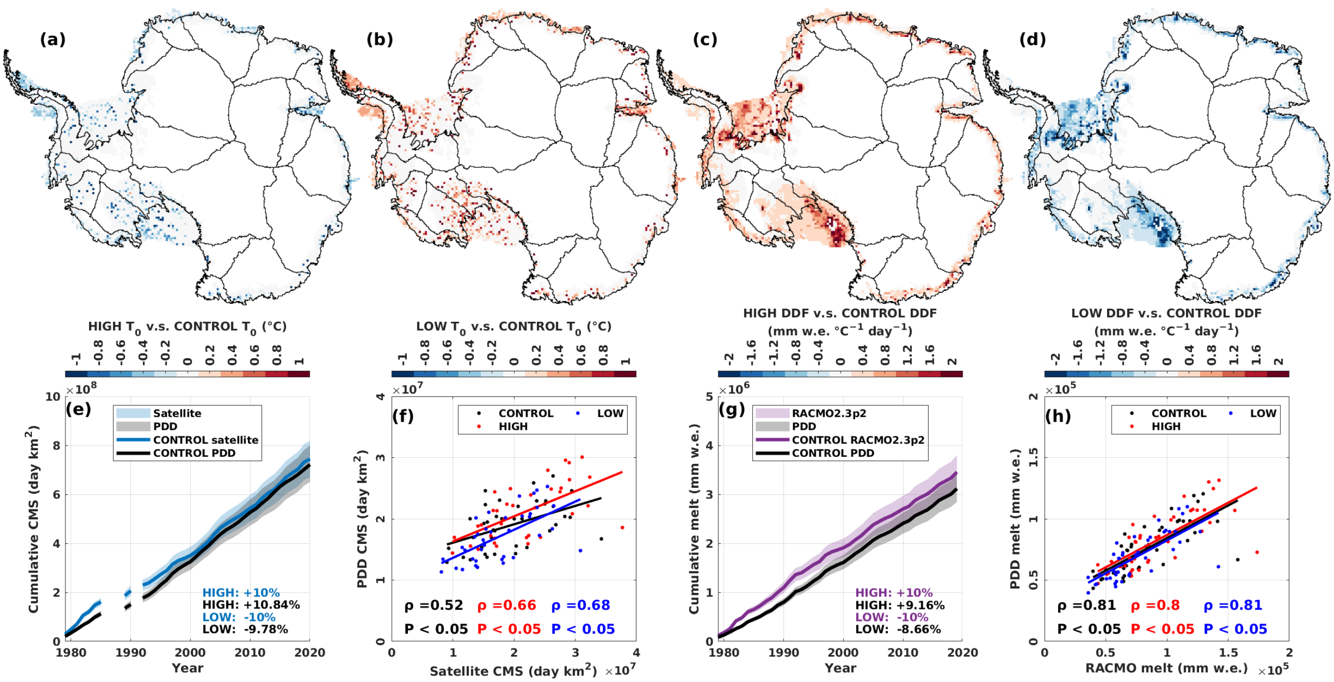

Figure 8. (a) and (b) difference between the T_0 parameterized in the HIGH/ LOW experiment and the CONTROL (optimal) T_0. (c) and (d) spatial maps for the difference between the DDF parameterized in the HIGH/ LOW experiment and the CONTROL (optimal) DDF. (e) and (g) cumulative CMS/ annual melt amount for the satellite estimates/ RACMO2.3p2 simulations and dist-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 indicate the HIGH/ LOW satellite estimates and the HIGH/ LOW dist-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/ dist-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 dist-PDD outputs and satellite/ RACMO2.3p2, from each sensitivity experiment (HIGH, LOW and CONTROL). This analysis is based on dist-PDD.

Figure 8 shows the result from our sensitivity experiments. We show that changes in the dist-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). It is expected that the T_0 decreases/ increases with the increase/ decrease of the satellite estimates, because a decrease of the threshold temperature is expected to increase the occurrence of temperatures above the threshold to produce more melt days, and vice versa. The increase/ decrease of 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 RACMO2.3p2 melt amount simulations, and vice versa.

Figure 8e shows that the dist-PDD model is less sensitive to the low melt scenario than the satellite estimates, as the dist-PDD estimates only decrease by 9.78% for the integrated 38-year CMS while the satellite estimates decrease by 10%. Although the dist-PDD model is more sensitive to the high melt scenario than the satellite estimates, where we show that dist-PDD increases by 10.84% on the 38-year integrated CMS with the 10% increase of the satellite estimates, this increase in dist-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 show that the dist-PDD model is less sensitive than RACMO2.3p2 in both the HIGH and LOW melt scenarios. Taken together, the sensitivity of the dist-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 our parameterization method is consistent to both the high and low melt scenarios.

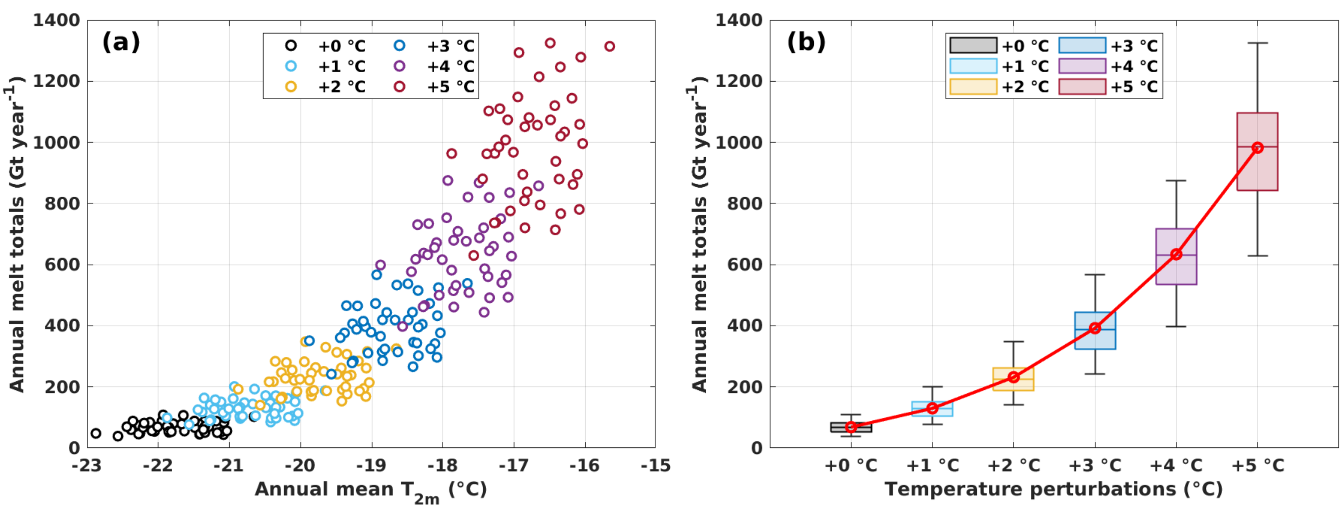

Figure 9. (a) scatter plot between annual mean 2-m air temperature (T_{2m}) and Antarctic annual melt totals for each temperature-melt sensitivity experiment for the period from 1979/1980 to 2021/2022. (b) boxplot of Antarctic annual melt totals for each temperature-melt sensitivity experiment for the period from 1979/1980 to 2021/2022.

Figure 9 shows the results from our temperature-melt sensitivity experiments. We show that a nonlinear increase in our dist-PDD estimates of Antarctic surface melt totals as the temperature perturbation gradually rises from +0 °C to +5 °C. It is not surprising that both the mean and standard deviation increase, given the anticipated nonlinear growth in melt volume resulting from the expansion of both the melt area and amount. The nonlinearity of temperature-melt sensitivity of our dist-PDD model is consistent with the nonlinearity temperature-melt relationship that is reported by other studies (Trusel et al., 2015; Bell et al., 2018; Banwell et al., 2023), further implying the applicability of our dist-PDD model to warmer climate scenarios.

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 dist-PDD parameterization and consider the spatial variability of model parameters, the dist-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 dist-PDD output has yet to be confirmed by other datasets.

5 Conclusions

We have constructed a PDD model with spatially varying parameters (dist-PDD) and with spatially uniform parameters (uni-PDD) based on the temperature-melt relationship (e.g. Hock, 2005; Trusel et al., 2015), and used them to estimate surface melt in Antarctica through the past four decades. We parameterized the dist-PDD and uni-PDD models 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 dist-PDD/ uni-PDD and satellite estimates, and SEB simulations, for each/ all computing cell(s). We independently performed two-sample KS tests on each computing cell in order to assess the goodness-of-fit for the parameterized dist-PDD and uni-PDD models. We also temporally and spatially compared the dist-PDD/ uni-PDD estimations, satellite estimates and RACMO2.3p2 simulations to evaluate the parameterized dist-PDD/ uni-PDD model. We found that our dist-PDD model improves the accuracy of Antarctic surface melt estimates compared to the uni-PDD setting, and 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 dist-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 of melt days and amounts 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 dist-PDD underestimation on both the melt days and amount. We suggest this underestimation corresponds to SAM-influenced climatic conditions, and that the dist-PDD lacks the ability to accurately capture melt if it arises from effects such as föhn winds that are not reflected in the input ERA5 2-m air temperature fields used to force the calculations (e.g. Turton et al., 2020).

These limitations aside, we found overall high fidelity of dist-PDD model, suggested by the 3-fold cross-validation. Although the dist-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 dist-PDD model when considering all computing cells is consistent. From the sensitivity experiments, we found the changes of the dist-PDD estimates are comparable to the changes in training data (satellite and RACMO2.3p2 data). The correlations between the dist-PDD estimates and training data exhibit stability regardless of the changes in the training data.

The dist-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 dist-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/ and https://earth.gsfc.nasa.gov/cryo/data/polar-altimetry/antarctic-and-greenland-drainage-systems (last access: 18 July 2023). 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). RACMO2.3p2 data are available from https://doi.org/10.5194/tc-12-1479-2018 (Van Wessem et al., 2018). The annual dist-PDD and uni-PDD models data from this study are available at https://doi.org/10.5281/zenodo.7131459. Data with higher temporal resolution (monthly, daily, and hourly) for dist-PDD and uni-PDD models from this study can be obtained 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.

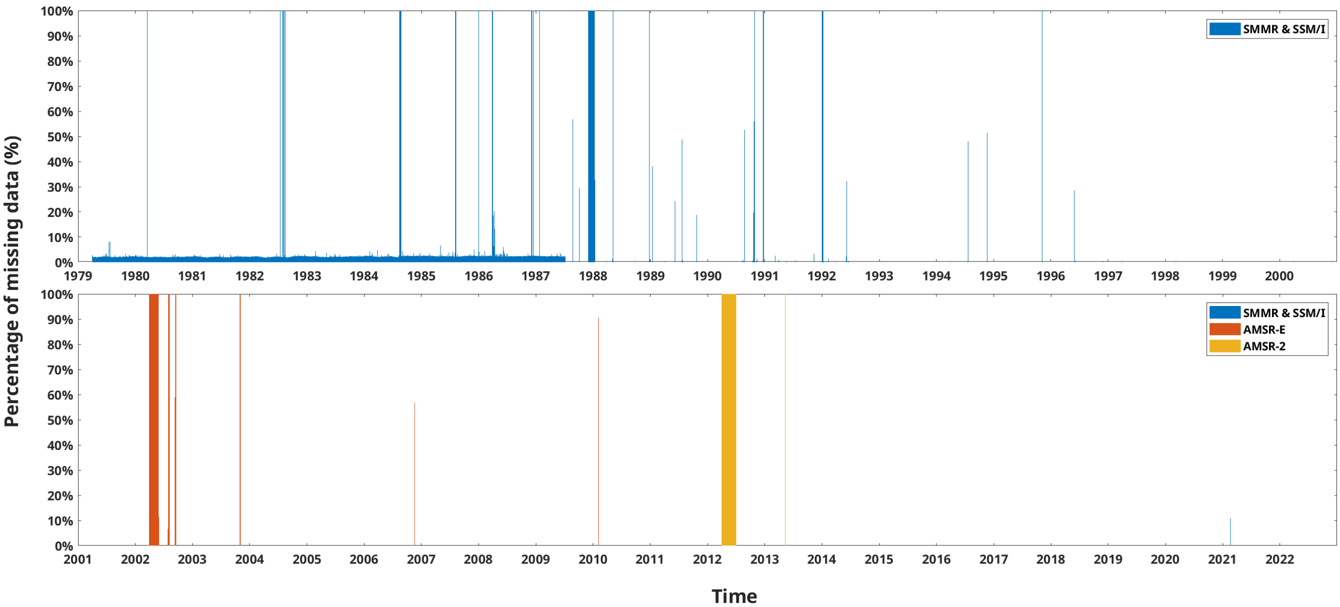

Figure A1. Daily percentage of missing data for satellite estimates. Satellite SMMR and SSM/I covers the period from 1979-04-01 to 2021-03-31. Satellite AMSR-E covers the period from 2002-04-01 to 2011-03-31. Satellite AMSR-2 covers the period from 2012-04-01 to 2021-12-31.

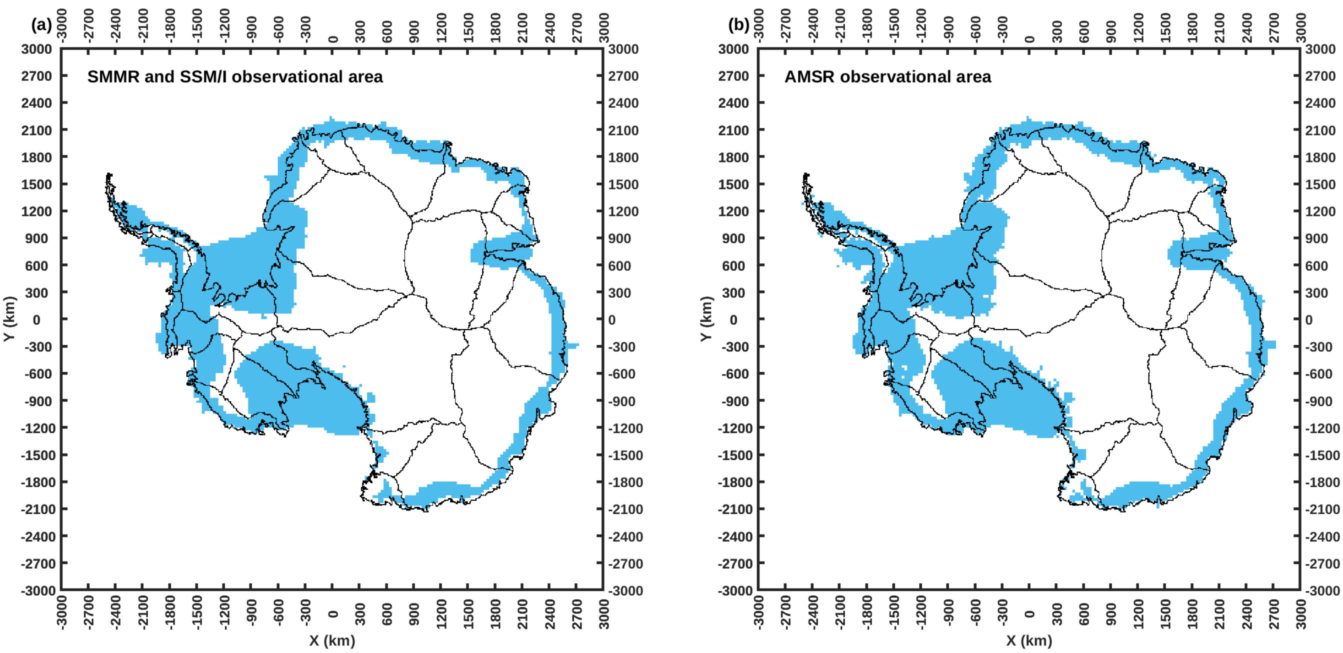

Figure A2. (a) mask of the satellite SMMR and SSM/I observational area. (b) mask of the satellite AMSR (AMSR-E and AMSR-2) observational area. Both masks are bilinearly remapped to the 30×30 km^2 polar stereographic grid.

Appendix B: Temperature-melt relationship

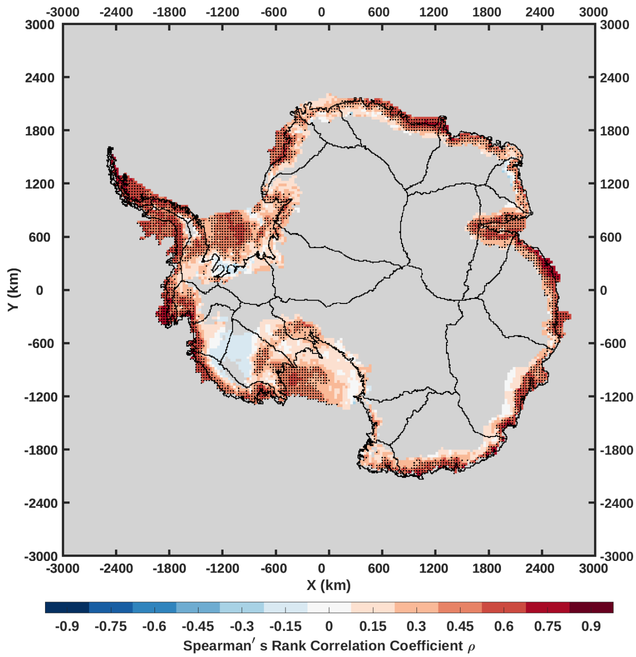

Figure B1. Correlation map between the mean DJF ERA5 2-m air temperature and RACMO2.3p2 annual surface melt amount for the period from 1979/1980 to 2019/2020. It is calculated by the Spearman's rank correlation coefficient on each cell. Black dots mark the cells where the correlations are statistically 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 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.

505 ## Appendix C: Spatially uniform PDD model

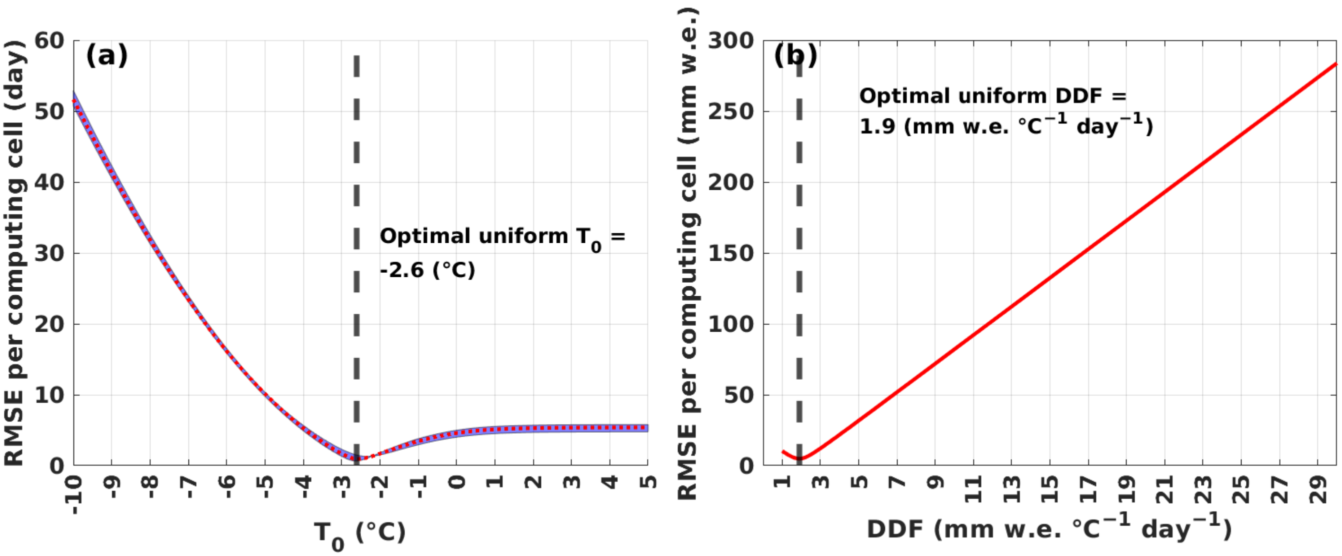

Figure C1. (a) red dotted curve is the average of the RMSE across all satellites along each uni-PDD T_0 experiment. In each uni-PDD T_0 experiment, we calculate the RMSE between the time series of annual sum of melt days over all computing cells between uni-PDD model and each satellite estimate. Blue envelope covers the span of the three individual satellite results. Black vertical dash line marks the optimal uni-PDD T_0 suggested by the minimal RMSE. (b) red curve is the RMSE along each uni-PDD DDF experiment. In each uni-PDD DDF experiment, we calculate the RMSE between the time series of annual sum of melt amount over all computing cells between uni-PDD model and RACMO2.3p2. Black vertical dash line marks the optimal uni-PDD DDF suggested by the minimal RMSE.

Appendix D: PDD model evaluation

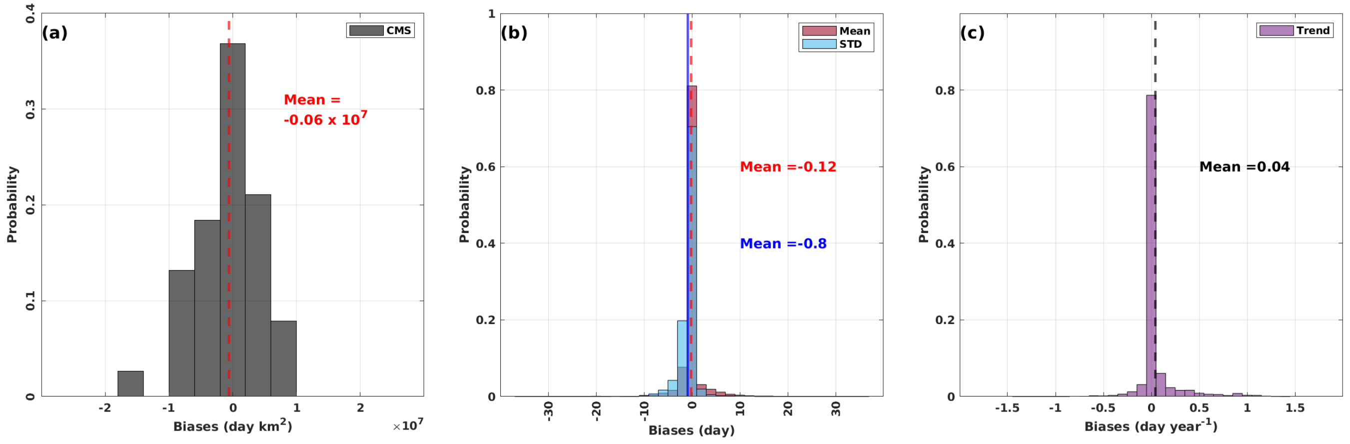

Figure D1. (a) probability histogram for the biases between the dist-PDD and satellite CMS. Red dashed vertical line indicates the mean of all biases. (b) and (c) probability histograms for the biases between the dist-PDD outputs and satellite estimates on mean, STD and trend. Red dashed vertical line indicates the mean of all biases between means. Blue vertical line indicates the mean of all biases between STDs. Black dashed vertical line indicates the mean of all biases 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.

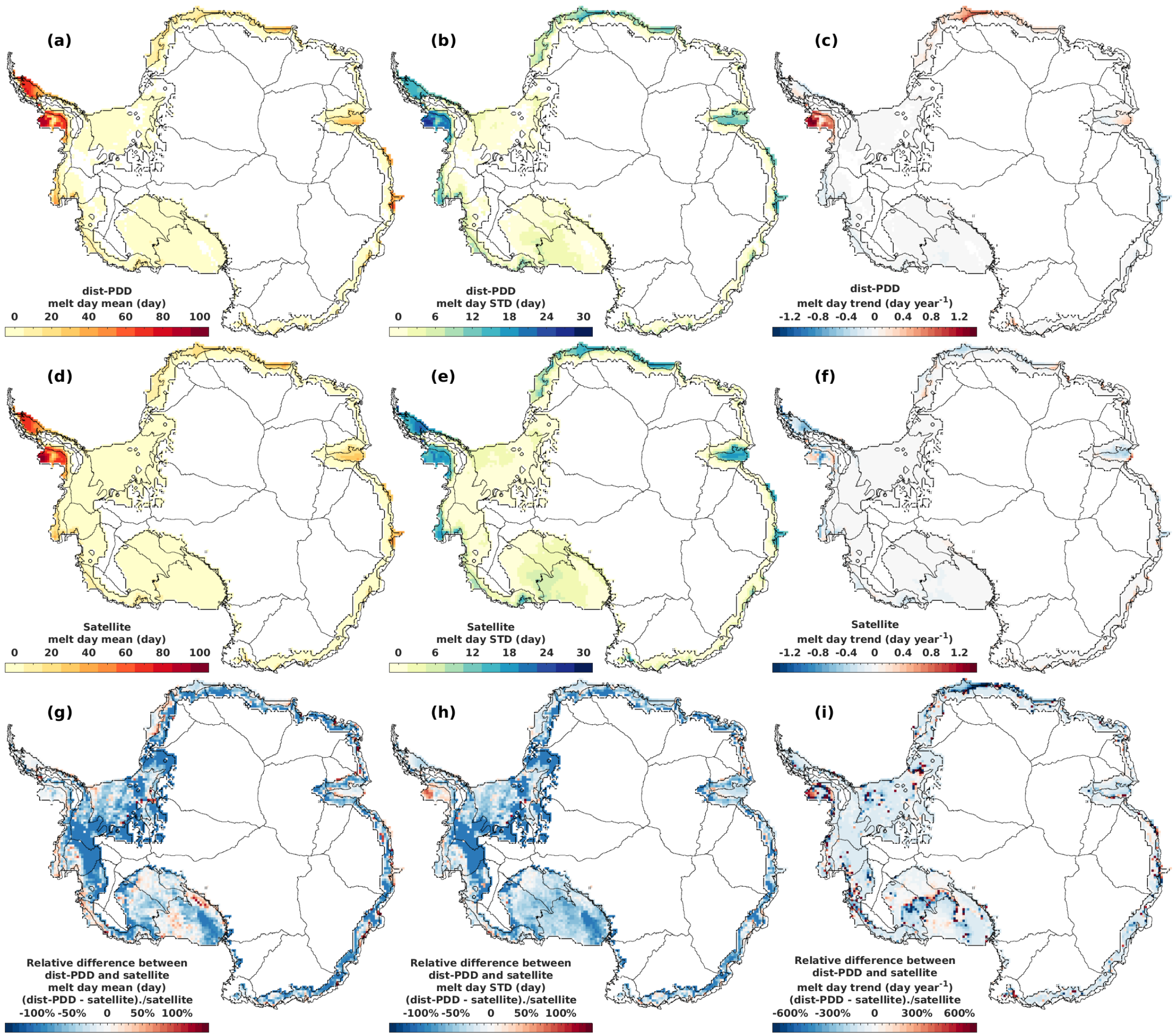

Figure D2. (a) to (f) mean, STD and trend of dist-PDD/ satellite melt days for the period 1979/1980 to 2020/2021, respectively. (g) to (i) relative difference between dist-PDD and satellite melt day mean, STD and trend for the period 1979/1980 to 2020/2021, respectively. Note that for all panels the 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. For all panels the period 1986/1987, 1987/1988, 1988/1989 and 1991/1992 are omitted.

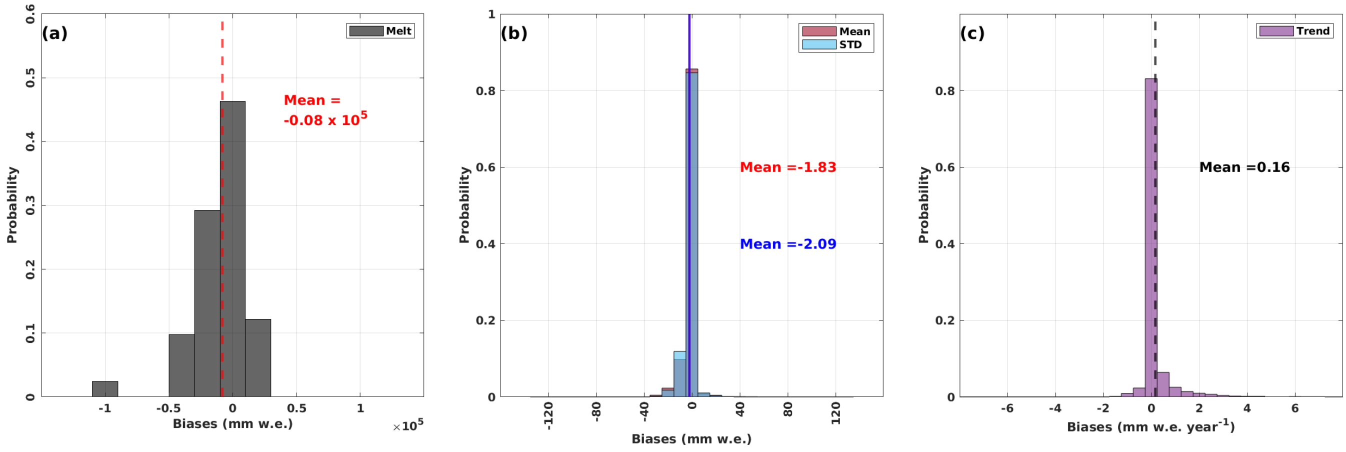

Figure D3. (a) probability histogram for the biases between the dist-PDD and RACMO2.3p2 melt amounts. Red dashed vertical line indicates the mean of all biases. (b) and (c) probability histograms for the biases between the dist-PDD outputs and RACMO2.3p2 simulations on mean, STD and trend. Red dashed vertical line indicates the mean of all biases between means. Blue vertical line indicates the mean of all biases between STDs. Black dashed vertical line indicates the mean of all biases between trends.

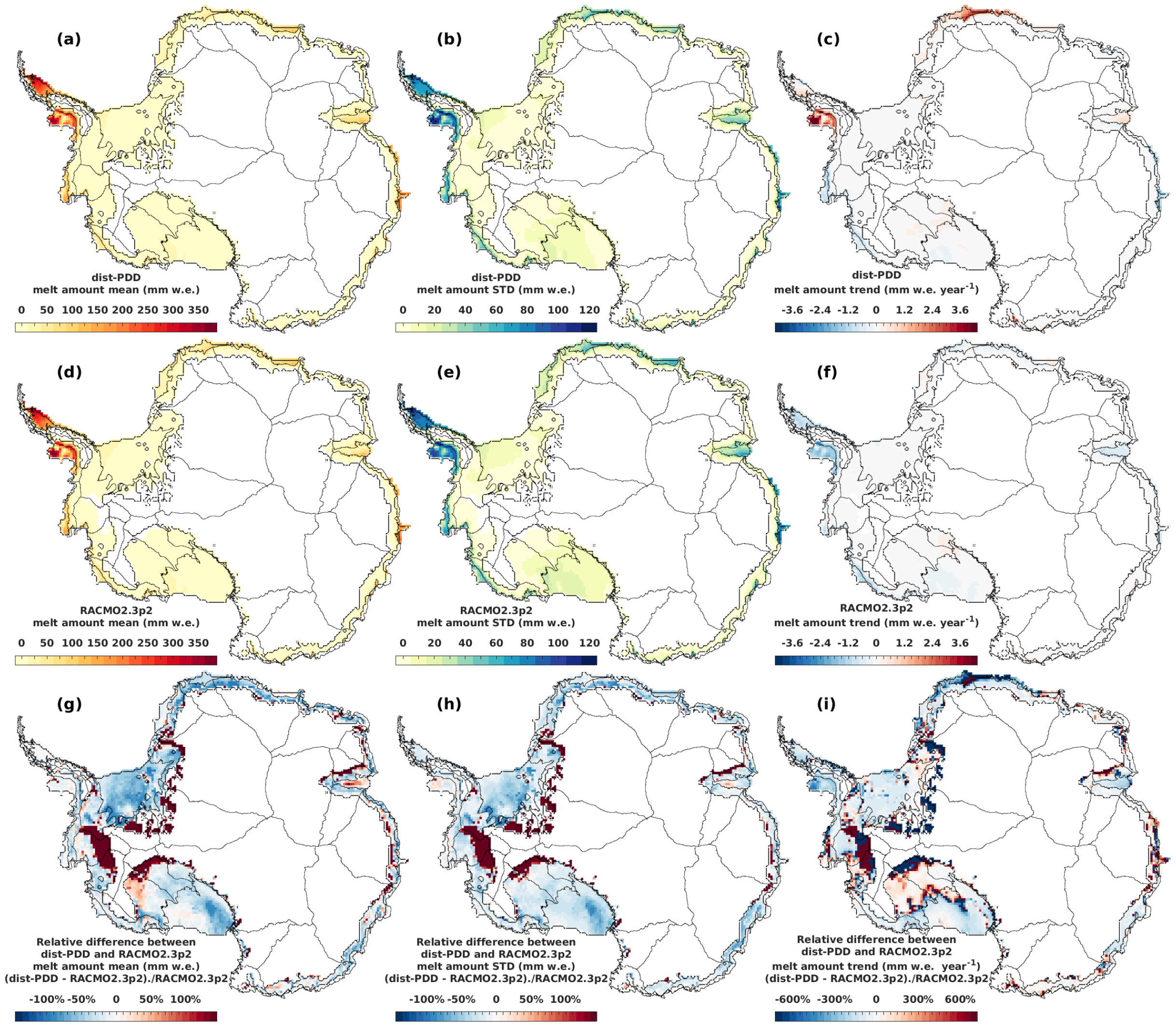

Figure D4. (a) to (f) mean, STD and trend of dist-PDD/ RACMO2.3p2 melt amounts for the period 1979/1980 to 2019/2020, respectively. (g) to (i) relative difference between dist-PDD and RACMO2.3p2 melt amount mean, STD and trend for the period 1979/1980 to 2019/2020, respectively.

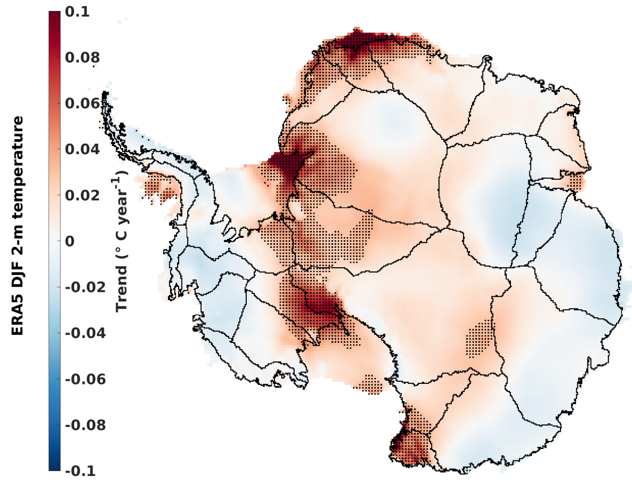

Figure D5. Trend of the mean DJF ERA5 2-m air temperature on each computing cell during the period 1979/1980–2019/2020. Black dots mark the trends that are statistically significant ($p < 0.05$).

Appendix E: 1982/1983 event

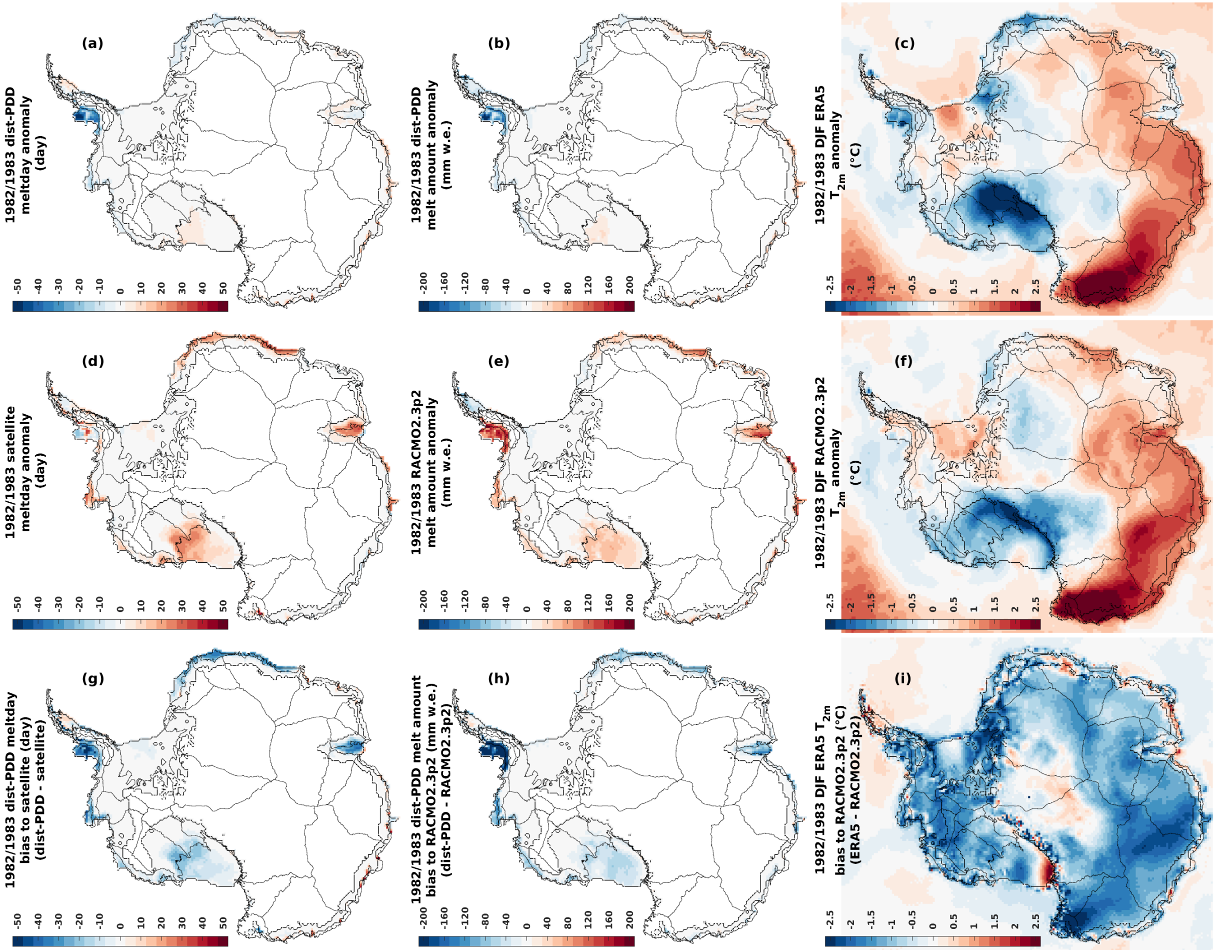

Figure E1. (a) and (d) 1982/1983 dist-PDD/ satellite meltday anomaly to the dist-PDD/ satellite mean meltday over the period 1979/1980–2020/2021 (with 1982/1983, 1986/1987, 1987/1988, 1988/1989 and 1991/1992 omitted). (g) absolute differences between 1982/1983 dist-PDD and satellite meltday. (b) and (e) 1982/1983 dist-PDD/ RACMO2.3p2 melt amount anomaly to the dist-PDD/ RACMO2.3p2 mean melt amount over the period 1979/1980–2019/2020 (with 1982/1983 omitted). (h) absolute differences between 1982/1983 dist-PDD and RACMO2.3p2 melt amount. (c) and (f) 1982/1983 DJF ERA5/ RACMO2.3p2 2-m air temperature anomaly to the DJF ERA5/ RACMO2.3p2 mean 2-m air temperature over the period 1979/1980–2019/2020 (with 1982/1983 omitted). (i) absolute differences between 1982/1983 DJF EAR5 and RACMO2.3p2 2-m air temperature. 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.

Figure E1d and e suggest that there is a positive surface melt anomaly in the ice shelves around Amundsen Sea, Ross Ice Shelf, Amery Ice Shelf, and ice shelves in Dronning Maud Land during the period 1982/1983. However, our dist-PDD model does not capture this event (Figure E1a and b). Our dist-PDD model shows significant negative bias in both surface melt days and amounts compared to satellite estimates and RACMO2.3p2 simulations for this 1982/1983 event (Figure E1g and h).

Both ERA5 and RACMO2.3p2 exhibit similar spatial patterns for the 1982/1983 DJF 2-m air temperature anomaly (Figure E1c and f). Although RACMO2.3p2 is forced by ERA5 2-m air temperature, its 2-m air temperature is consistently warmer than that of ERA5 during the 1982/1983 DJF period. This is particularly noticeable in the computing cells over the ice shelves around the Amundsen Sea, Ross Ice Shelf, Amery Ice Shelf, and Dronning Maud Land, where we show that significant negative biases for dist-PDD surface melt days and amounts compared to satellite and RACMO2.3p2. These cells also align with the cells where negative ERA5 2-m air temperature biases towards RACMO2.3p2 are found.

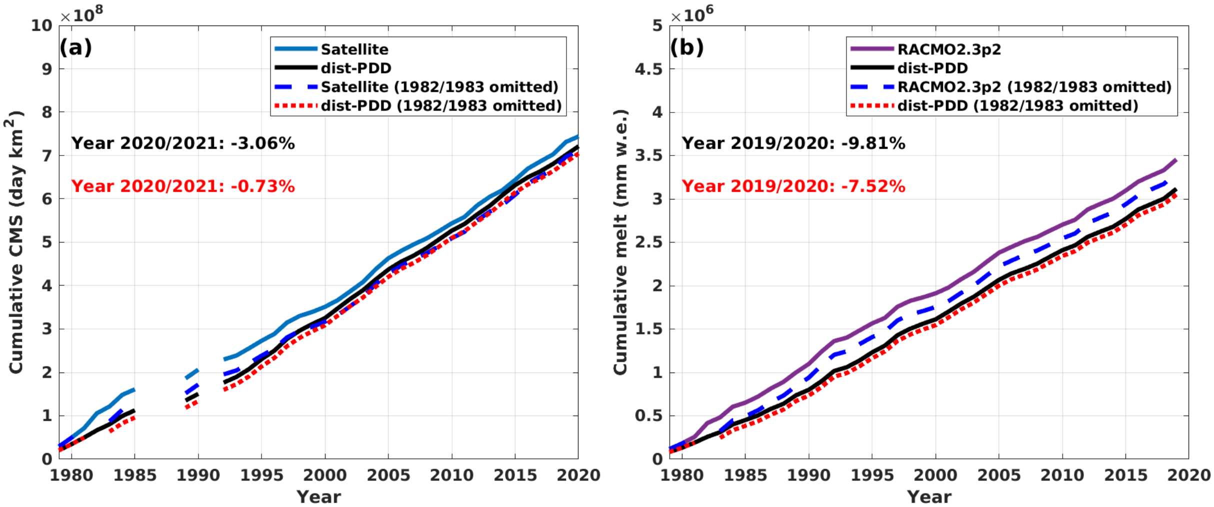

Figure E2. (a) cumulative CMS for satellite estimates and dist-PDD/ dist-PDD (1982/1983 omitted) outputs from 1979/1980 to 2020/2021 (with 1986/1987 to 1988/1989 and 1991/1992 omitted. (b) cumulative annual melt amount for RACMO2.3p2 simulations and dist-PDD/ dist-PDD (1982/1983 omitted) outputs from 1979/1980 to 2019/2020.

We then assess the goodness-of-fit of the dist-PDD model after removing the 1982/1983 period for dist-PDD, satellite, and RACMO2.3p2. The exclusion of the 1982/1983 period significantly improves the accuracy of the dist-PDD model in comparison to satellite and RACMO2.3p2 (Figure E2). Although there is a slight negative bias of dist-PDD (excluding 1982/1983) cumulative CMS compared to satellite data (excluding 1982/1983) in the first decade, the two cumulative CMS curves converge after approximately the first decade and almost overlap for the rest of the time period (Figure E2a). Similarly, the cumulative melt curves for dist-PDD (excluding 1982/1983) and RACMO2.3p2 (excluding 1982/1983) show a slight divergence in the first decade but remain parallel for the rest of the time period (Figure E2b). By the end of the integration period, the relative difference between dist-PDD and satellite CMS decreased from -3.06% to -0.73% (Figure E2a), while the relative difference between dist-PDD and RACMO2.3p2 melt amounts decreased from -9.81% to -7.52% (Figure E2b). These improvements are consistent across correlations and OLS linear regression analyses, as shown in Table E1, indicating the enhanced perfor-

Table E1. The Spearman's ρ and P-value for dist-PDD/ dist-PDD (1982/1983 omitted) CMS/ melt amounts with the satellite CMS/ RACMO2.3p2 melt amounts. Slope, R^2, RMSE and P-value for the Ordinary Least Squares (OLS) fit between dist-PDD/ dist-PDD (1982/1983 omitted) CMS/ melt amounts and satellite CMS/ RACMO2.3p2 melt amounts. Note that 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. All the dist-PDD with satellite statistics are calculated over the period from 1979/1980 to 2020/2021 (with 1986/1987 to 1988/1989 and 1991/1992 omitted). All the dist-PDD with RACMO2.3p2 statistics are calculated over the period from 1979/1980 to 2019/2020.

Member	Spearman's ρ	P-value	OLS slope	R^2	RMSE (day km^2/ mm w.e.)	P-value
dist-PDD v.s. satellite	0.5203	$P < 0.01$	0.3004	0.229	3.38×10^6	$P < 0.01$
dist-PDD[a] v.s. satellite[a]	0.5778	$P < 0.01$	0.3894	0.325	3.19×10^6	$P < 0.01$
dist-PDD v.s. RACMO2.3p2	0.8052	$P < 0.01$	0.5307	0.55	1.42×10^4	$P < 0.01$
dist-PDD[a] v.s. RACMO2.3p2[a]	0.8486	$P < 0.01$	0.6582	0.712	1.15×10^4	$P < 0.01$

[a] 1982/1983 is omitted.

mance of the dist-PDD model in estimating both surface melt days and amounts compared to satellite and RACMO2.3p2 after excluding the 1982/1983 period.

On the basis of this additional experimentation we are able to confidently conclude that our model is accurate for the vast majority of the time series, and that any previously apparent bias was almost entirely due to the anomalous conditions of a single year.

Appendix F: 3-fold CV T_0 Member 1

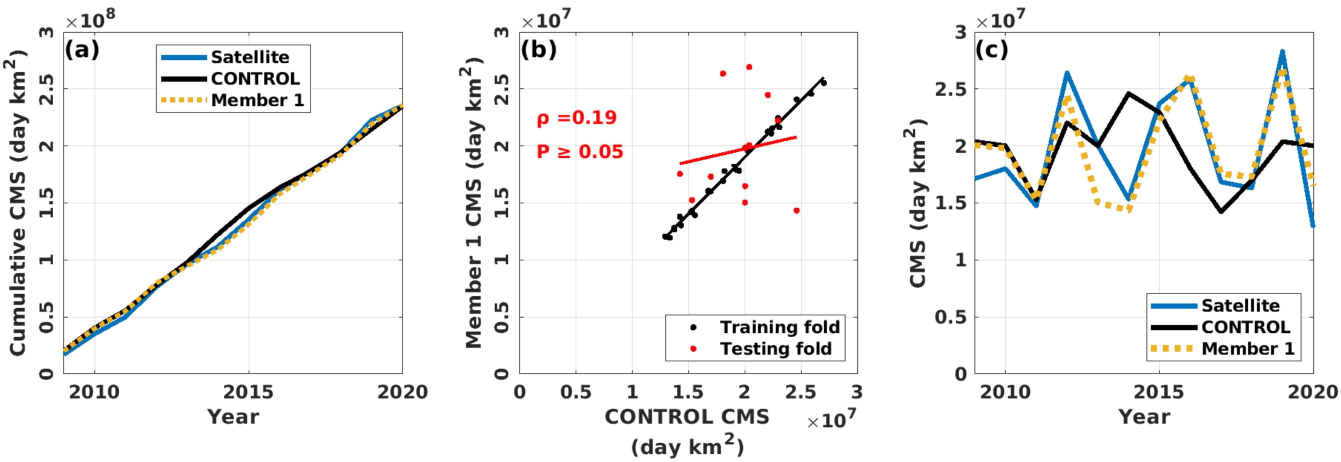

Figure F1. (a) and (b) are same as the Figure 7(m) and (s). (c) time series of the CMS for satellite estimates, CONTROL and Member 1 during the testing fold period.

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

Agosta, C., Amory, C., Kittel, C., Orsi, A., Favier, V., Gallée, H., van den Broeke, M. R., Lenaerts, J., van Wessem, J. M., van de Berg, W. J., et al.: Estimation of the Antarctic surface mass balance using the regional climate model MAR (1979–2015) and identification of dominant processes, The Cryosphere, 13, 281–296, 2019.

Banwell, A. F., Wever, N., Dunmire, D., and Picard, G.: Quantifying Antarctic-Wide Ice-Shelf Surface Melt Volume Using Microwave and Firn Model Data: 1980 to 2021, Geophysical Research Letters, 50, e2023GL102 744, 2023.

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.

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.

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.

Fricker, H. A., Arndt, P., Brunt, K. M., Datta, R. T., Fair, Z., Jasinski, M. F., Kingslake, J., Magruder, L. A., Moussavi, M., Pope, A., et al.: ICESat-2 meltwater depth estimates: application to surface melt on Amery Ice Shelf, East Antarctica, Geophysical Research Letters, 48, e2020GL090 550, 2021.

Glasser, N. and Scambos, T. A.: A structural glaciological analysis of the 2002 Larsen B ice-shelf collapse, Journal of Glaciology, 54, 3–16, 2008.

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.

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.

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.

Nowicki, S., Bindschadler, R. A., Abe-Ouchi, A., Aschwanden, A., Bueler, E., Choi, H., Fastook, J., Granzow, G., Greve, R., Gutowski, G., et al.: Insights into spatial sensitivities of ice mass response to environmental change from the SeaRISE ice sheet modeling project I: Antarctica, Journal of Geophysical Research: Earth Surface, 118, 1002–1024, 2013.

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.

Spergel, J. J., Kingslake, J., Creyts, T., van Wessem, M., and Fricker, H. A.: Surface meltwater drainage and ponding on Amery Ice Shelf, East Antarctica, 1973–2019, Journal of Glaciology, 67, 985–998, 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.

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.

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 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.

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.

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.