# Scoring Antarctic surface mass balance in climate models to refine future projections

Tessa Gorte[1], Jan T. M. Lenaerts[1], and Brooke Medley[2]

[1]Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder

[2]Cryospheric Sciences Laboratory, National Aeronautics and Space Administration's Goddard Space Flight Center

**Abstract.** An increase of Antarctic Ice Sheet (AIS) surface mass balance (SMB) has the potential to mitigate future sea level rise that is driven by enhanced solid ice discharge from the ice sheet. For climate models, AIS SMB provides a difficult challenge, as it is highly susceptible to spatial, seasonal and interannual variability.

Here we use a reconstructed data set of AIS snow accumulation as "true" observational data, to evaluate the ability of the CMIP5 and CMIP6 suites of models in capturing the mean, trends, temporal variability and spatial variability in SMB over the historical period (1850-2000). This gives insight into which models are most reliable for predicting SMB into the future. We found that the best scoring models included the National Aeronautics and Space Administration's GISS models and the Max Planck Institute fr Meteorologie's MPI models for CMIP5 and one of the National Center for Atmospheric Research's CESM2 models and one MPI model for CMIP6.

Using a scoring system based on SMB mean value, trend, and temporal variability across the AIS, as well as spatial SMB variability, we selected a subset of the top 10th percentile of models to refine 21st century (2000-2100) AIS-integrated SMB projections to $2372 \pm 282$ Gt yr$^{-1}$, $2452 \pm 286$ Gt yr$^{-1}$, and $2588 \pm 291$ Gt yr$^{-1}$ for Representative Concentration Pathways (RCPs) 2.6, 4.5, and 8.5, respectively. We also reduced the spread in AIS-integrated mean SMB by 79%, 79%, and 74% in RCPs 2.6, 4.5, and 8.5, respectively.

Notably, we find that there is no improvement from CMIP5 to CMIP6 in overall score. In fact, CMIP6 performed slightly worse on average compared to CMIP5 at capturing the aforementioned SMB criteria. Our results also indicate that model performance scoring is affected by internal variability, which is illustrated by the fact that the range in overall score between ensemble members within the CESM1 Large Ensemble is comparable to the range in overall score between CESM1 model simulations within the CMIP5 model suite. However, we also find that a higher horizontal resolution does not yield to a conclusive improvement in score.

## 1 Introduction

Surface mass balance (SMB) is the rate of accumulation of mass on the surface of the ice sheet and is characterized predominantly by precipitation and sublimation, and also includes runoff and blowing snow terms (Lenaerts et al., 2019). We neglect blowing snow and runoff and estimate SMB as precipitation minus sublimation (Lenaerts et al., 2012). Ignoring these terms, AIS SMB can be estimated as SMB = precipitation - sublimation. As SMB variability is dominated by that of AIS precipitation, which is subject to high spatial and temporal variability (Bromwich et al., 2011), SMB is also highly variable from year to year (Monaghan and Bromwich, 2008).

Over longer ($\sim$100-1000 year) time scales, AIS SMB was assumed – until recently – to be relatively constant. Frezzotti et al. (2013) found that current SMB values are not anomalously high compared to the past  1000 years. Monaghan et al. (2006) found no discernible trend in AIS snowfall in the period 1957-2003. More recent studies, adding more annually-resolved SMB records covering the period 1800 to present and improving the spatial extrapolation, contested those earlier findings (Thomas et al. (2017); Medley and Thomas (2019)). These studies found that, integrated over the AIS, SMB has been increasing at a rate of $0.4 \pm 0.1$ Gt yr$^{-2}$ over the last 200 years, although the trends show substantial regional variability. Several studies have provided additional evidence of regional variations in SMB trends, with strong SMB increase in some areas (Philippe et al. (2016); Thomas et al. (2015); Thomas et al. (2017)), and no SMB increase, or even SMB decrease, in other areas (Burgener et al., 2013). Synoptic-scale variability induces a strong regional variability of the SMB (Fyke et al. (2017); Marshall et al. (2017)). Additionally, as the atmosphere is projected to warm both globally and especially in the polar regions, the atmosphere is expected to be able to hold more moisture per the Clausius-Clapeyron relation. As such, SMB is expected to show an overall increase. In recent decades, this forced SMB response is undetectable due to the significant natural SMB variability (Previdi and Polvani, 2016). Teasing apart the forced response from natural SMB variability requires longer SMB time series – on the order of centuries. In 2017, Thomas et al. found no significant SMB trend over the last 1000 years. In 2019, however, Medley & Thomas found that, over the past 200 years, there is a statistically significant SMB increase that can be derived from ice core measurements.

Despite its importance for AIS MB and GMSL, there are only few robust observations of SMB across the continent. A lack of regular spatial and temporal distribution of observations has led to many efforts to model SMB using both regional and global climate models (RCMs and GCMs, respectively). Because the AIS is so large, predicting SMB out onto timescales from decades to centuries requires the use of GCMs (Gallée et al., 2013). Some GCMs have been shown to capture positive precipitation and SMB trends (Palerme et al. (2014); Lenaerts et al. (2016)), but many of those models tend to overestimate annual precipitation values likely due to poor representation of coastal topography as previous studies have shown this to be a significant factor in how precipitation is represented of the AIS (Genthon et al., 2009). This allows the atmospheric moisture to penetrate too far inland and leads to excessive precipitation on much of the grounded AIS, while underestimating precipitation nearby the coasts (Palerme et al. (2017)). This inability to reproduce modern observations brings into question the models' ability to accurately project future changes.

While past research by Palerme et al. (2014) compared model output to observations using CloudSat and ERA-Interim, their observational data sets only spanned a short period (2006-2011). The limited climatology of AIS precipitation combined with its highly temporally variable nature means that large limitations exist to enable a comparison. Barthel et al. (2019) investigated the Ice Sheet Model Intercomparison Project for CMIP6 to determine a recommendation of which models to use for ice sheet model forcings based on best captured current Antarctic climate relative to observations and their ability to project certain metrics into the future. The object of this paper is similar in that Barthel et al. (2019) use scoring criteria to refine model selection specifically for ice sheet model forcing. Their work differs in that their criteria look more at the large-scale circulation patterns around ice sheets and the data set to which they compare models consists of large-scale fields reanalysis fields. Additionally, they don't then use this subselection of models to constrain future projections. In this work, we use a data set that specifically accounts for AIS SMB using recent advancements in synthesizing ice cores and reanalysis products. These reconstructed data sets now allow for a new avenue to investigate the ability of GCMs to capture SMB into the more distant past (Medley and Thomas, 2019) – an avenue that we leverage for climate model evaluation of AIS SMB to compare the suite of CMIP5 and CMIP6 climate models to this new SMB reconstruction.

In this work, we leverage the availability of that new avenue for climate model evaluation of AIS SMB, and compare the suite of CMIP5 and CMIP6 climate models to that new SMB reconstruction.

## 2 Data

### 2.1 SMB Reconstructions

To improve upon model estimates, several groups have combined ice core data with models to create spatio-temporally robust SMB data sets (Monaghan et al. (2006), Thomas et al. (2017), Medley and Thomas (2019)). In this paper, we use the AIS SMB reconstruction generated by Medley and Thomas (2019). The authors synthesize SMB time series from an extensive ice-core database with reanalysis-derived spatial coherence patterns to generate a continent-wide AIS SMB data set. While Medley and Thomas (2019) compared three reanalysis products, they also show that MERRA-2 performed better than the other two reconstructed products in matching observations. As such, we will use the MERRA-2 based data set as a proxy for all three reconstructions and refer to it as "reconstruction."

For this work, we investigate AIS SMB in GCMs. GCMs have, compared to RCMs, relatively low horizontal resolution, which makes it difficult for them to reproduce the detailed AIS SMB. RCMs have been shown to be more accurate in capturing AIS SMB (Agosta et al., 2019); however, due to their high resolution, RCMs are also relatively computationally expensive to run for long periods ($\sim$100s of years). Because one of the goals of this paper is to investigate the future of SMB over Antarctica, we analyze GCMs for their ability to simulate these long-term climate effects. As RCMs are by definition regional, they need boundary forcings, which adds an additional layer of complexity and a source of uncertainty to running RCMs into the long-term future. An additional reason we choose to analyze GCMs is simply to figure out which GCMs perform best at capturing these SMB phenomena. There has been extensive work investigating SMB in RCMs (e.g., Agosta et al. (2019); van Wessem et al. (2017); Lenaerts et al. (2012)), but comparably little looking at GCMs. To investigate the global coupled response to future SMB changes, one needs GCMs. As such, this work is aimed to inform the modeling community who is interested in global ramifications of changing AIS mass balance, and the ice sheet modeling community who needs AIS SMB input for running dynamical ice sheet models (Seroussi et al., 2019 in TC). Several recent studies, such as Barthel et al. (2019), Krinner et al. (2014), and Beaumet et al. (2019) have investigated the impacts of thermodynamical phenomena such as sea level pressure, zonal wind speed, and near-surface temperatures as well as phenomena like sea ice extent on AIS SMB, but have not scored climate models on their performance on SMB specifically. Here, we develop scoring criteria that assess AIS

SMB exclusively, and focus less on the mechanisms behind SMB variability and change. To get a comprehensive look at how well global climate models capture SMB, we compared the suites of CMIP5 and CMIP6 models to the reconstruction.

## 2.2 Climate Models

We used all applicable CMIP5 and CMIP6 model outputs, of which there were 81 models and 42 independent models (i.e. different model physics and/or resolutions) respectively, for the historical simulations (1850-2005). As for the future simulations (2006-2100), we focused on CMIP5 only, since there are few CMIP6 models available as of yet, and CMIP5 and CMIP6 scenarios are similar. We only had available output for 30 CMIP5 models, 19 of which are independent, for the future simulations. See Tables 1-3 in Supplementary Material for a list of models and their resolutions. The future simulations include three different forcing scenarios: Representative Concentration Pathway (RCP) 2.6, RCP4.5, and RCP8.5. RCP2.6 represents a low emission scenario, RCP4.5 a mid-range emission scenario, and RCP8.5 a high emission scenario through the 21st century (van Vuuren et al., 2011).

We downloaded CMIP5 and CMIP6 precipitation and evaporation/sublimation output at monthly time resolution and, after calculating SMB as precipitation - evaporation/sublimation, converted an annual time scale and integrated across the grounded AIS using the Ice Sheet Mass Balance Inter-comparison Exercise Team's (IMBIE Team) ice sheet mask.

# 3 Methods

We formulated five criteria on which to score the historical runs of the models. Three of the criteria are based on the AIS-integrated SMB: mean, trends, variability – and two are based on AIS SMB spatial patterns: modes of SMB variability, and variance explained by these modes. As the models' abilities to capture SMB are presented in the format of a "score card," judging the models against each criterion will be hereinafter referred to as "scoring". These criteria were determined having in mind the following questions: (1) do the models adequately simulate several SMB observed characteristics in the recent past, and (2) are the models that perform well adequately simulating SMB for the right reasons? All five criteria are weighted equally in the final scoring.

## 3.1 AIS-integrated SMB criteria

To score the models based on AIS-integrated SMB, we took the mean SMB across the AIS for every year that the reconstruction overlapped the models (1850-2000) to generate a single 151-year, AIS-integrated time series. We then split the time series into three aspects: the mean value of the SMB time series values (mean value referring to the value obtained by integrating SMB over the entire AIS), the time series linear trend, and the time series interannual variability.

To score the time series mean value, we assigned a score, x, for how many x-times the reconstruction uncertainty was required for the entire time series to be within the reconstruction uncertainty. The minimum possible score, then, is one, for a model that represents SMB within $1\times$ the reconstruction uncertainty. Fig. 1 illustrates that a model that fits entirely within $1\times$ the reconstruction uncertainty (dark purple) – MPI ESM LR – would receive a score of 1. A model that fits within $2\times$ the

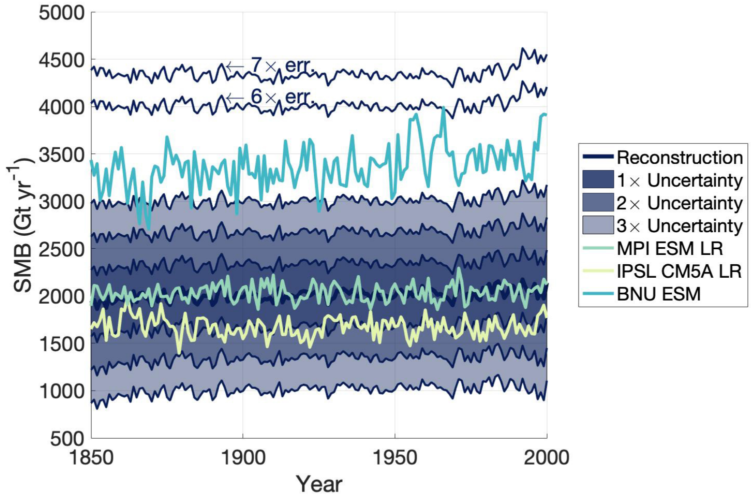


**Figure 1.** Time series of the reconstructed AIS-integrated SMB time series (purple) with 1×, 2×, and 3× the uncertainty in dark purple, medium purple, and light purple, respectively. Three model AIS-integrated SMB time series, MPI ESM LR (green), IPSL CM5A LR (yellow), and BNU ESM (cyan) have been plotted as well to demonstrate different model scoring. MPI ESM LR is entirely captured within 1× the reconstruction uncertainty and, thus, receives a score of 1. IPSL CM5A LR is entire captured within 2× the uncertainty so its score for this criterion is 2. BNU ESM is fully captured within 7× the uncertainty.

reconstruction uncertainty (medium purple) – IPSL CM5A LR – would receive a score of 2. A poorer scoring model, BNU ESM, would receive a score of 6.

Similarly, for the time series trend, we assigned a score of x based on how many x-times the reconstructed trend uncertainty was required to capture the model trend. We looked at multiple time "slices" to investigate how well the models performed at capturing century-scale (100+ year) versus multi-decadal (50 year) SMB trends. To achieve this goal, we analyzed trends from 1850-2000, 1900-2000, and 1950-2000. The first two of these three time slices confirm the robustness of the trends with longer periods for trend analysis. The last time slice, 1950-2000, allows us to view SMB in the context of significant anthropogenic warming. However, the large interannual variability overwhelms the signal at shorter period lengths, which results in large uncertainty bounds. By looking at several time slices, we ensure consistency between the model and reconstruction over different intervals. It is equally important to confirm that pre-1950, the trends are relatively small. We performed a Monte Carlo simulation wherein we assumed a normal distribution where the standard deviation of the distribution is equal to the reconstruction uncertainty of possible SMB values for each year. We then created 10,000 potential SMB time series by choosing SMB values based on that normal distribution for each year and recalculated the trend for each of these time series. Our uncertainty, then, was the standard deviation of this range of trends, similar to Medley and Thomas (2019).

For temporal variability, if a model should greatly underestimate the mean value, for example, the variability about that mean value will also likely be underestimated. To ensure that we are not double-counting the impact of SMB mean value (because this is already covered by the first scoring criterium), we calculate the variability about the normalized time series. To detrend and normalize each time series, then, to separate the SMB variability from its mean value, we performed the following analysis:

$$\text{normalized SMB} = \frac{\text{SMB} - \text{mean SMB}}{\text{mean SMB}}. \tag{1}$$

We then calculated the standard deviation of each time series and assigned a score, x, based on how many x-times the reanalysis standard deviation were required to capture the model standard deviation. For this criterion, we used the original MERRA-2 reanalysis precipitation minus evaporation data (1980-2019). Likely due to sampling only 53 ice core sites, the reconstruction produced a relatively low variability record. The reconstructed variability at any location can only be as large as the maximum variability in the ice cores. Thus, undersampling regions of stronger interannual variability will dampen the variability signal in the reconstruction. Analyses of the AIS-integrated SMB mean value and trend show that the reconstruction is generally in line with the literature (Medley and Thomas, 2019).

### 3.2 Spatial SMB criteria

To ensure model performance was not solely based on AIS-integrated SMB values, we also analyzed the spatial SMB variability. To do so, we performed an empirical orthogonal function (EOF) analysis on annual data from 1850-2005. EOF analysis maps the spatial pattern of a variable associated with the highest temporal variance of another variable.Here we apply EOF analysis to the spatial pattern of sea level pressure associated to the highest variability in annual SMB integrated over the AIS for the period 1850-2000. By breaking this criterion down into two main factors, (1) spatial variability and (2) variance explained, both of which are considered as separate scoring criteria, we aim to determine the models' abilities to accurately capture the modes of variability as well as how much variance each EOF mode explained.

In the reconstruction, the top three modes of variability collectively explain roughly 76% of the total variance explained. The fourth mode explains only about 6% of the total variance and all other modes explain <5% of the total variance. As such, we only include the top three modes in our analysis. To avoid manually sorting the top three modes of variability for all 53 models, we generated difference maps between each of the top three reconstructed modes and each of the top three modes for each model: 9 difference maps for each model. For each grid point, we took the absolute value of the difference between the model and the reconstruction. We then summed those differences to generate a single number ("difference number") that represented the difference between the model and the reconstruction in terms of spatial variability. Mathematically, this looks like:

$$\text{difference number} = \sum_{lat}\sum_{lon}\left|\text{reconstruction}_{lat,lon} - \text{model}_{lat,lon}\right| \tag{2}$$

We did this for all nine combinations of model and reconstruction maps for the top three modes variability ($\text{model}_1$:$\text{reconstruction}_1$, $\text{model}_1$:$\text{reconstruction}_2$, $\text{model}_1$:$\text{reconstruction}_3$, $\text{model}_2$:$\text{reconstruction}_1$, $\text{model}_2$:$\text{reconstruction}_2$, etc.). For reconstruction mode 1 ($\text{reconstruction}_1$), then, we matched which model mode best represented this spatial variability by sorting the model modes based on the smallest difference number. We did this for each reconstruction mode (excluding previously matched model modes) to sort the modes based on the smallest difference. Summing the absolute value of these differences yielded a single number that explained how different a given model was from the reconstruction for each mode of variability. The score, then, for the variability of SMB is the total difference of all the top 3 modes.

Because the variance explained is also important for gauging how well models are performing at recreating the observed spatial patterns, we also summed the difference in variance explained for the top three sorted modes of variability for each model. Because the modes were sorted based on difference for the maps, each mode kept its variance explained to preserve the accuracy of the models regarding the dominance of each spatial pattern.

### 3.3 Final Scoring

After compiling scores for all five of the aforementioned scoring criteria, we removed any outliers by calculating the 1.5 quartile range of the data and neglecting models that fell outside of that range. We then normalized each set of scores to be on a scale from one to ten to ensure that each criterion was equally weighted. After this normalization, the outliers for any given criterion were retroactively assigned a score of ten for that criterion. The total score, then, is the average of all five sets of normalized scores. Because the scores are based on the difference between the reconstruction and the models, higher scores indicate poorer model performance.

### 3.4 Future Projections

To refine the scope of what we predict for AIS SMB in the future, we created a subset of models that had a final score in the top 10th percentile (90th percentile and above) of CMIP5 and CMIP6. For our future projections, we investigated the impacts of SMB under three different forcing scenarios: RCPs 2.6, 4.5, and 8.5. Because CMIP6 uses a different future forcing scenario mechanism (Shared Socioeconomic Pathways), CMIP5 and CMIP6 future projections are not directly comparable. As such, we focused on the CMIP5 suite of models and their future projections. To that end, we compared the top scoring CMIP5 models that could be projected out under the three RCP forcings (of which there are four) to the entire scope of CMIP5. We ran a Monte Carlo simulation in which four random CMIP5 models were selected 100,000 times. Those 100,000 sets of four random scores were compared to the four best scoring model scores using a two-sided t-test. From this, we found that, to a 95% confidence level, we can reject the null hypothesis that the four best scoring models are not statistically significantly different from any random four CMIP5 or CMIP6 models.

Using this subset of best scoring models, we calculated the projected AIS-integrated mean value and trend in three different warming scenarios, RCPs 2.6, 4.5, and 8.5, out to 2100. To see if and how the models respond differently to different warming scenarios, we also calculated the AIS-integrated SMB sensitivity as

$$\text{Sensitivity} = \frac{\Delta\text{SMB}}{\Delta\text{T}}. \tag{3}$$

## 4 Results

The final overall scores are an average of all the scores from all five criteria. After performing the analysis outlined in the Methods section the top 90th percentile overall scoring models were determined to be GISS E2 H CC, GISS E2 R CC, GISS

E2 R, MPI ESM LR, MPI ESM MR, and MPI ESM P from CMIP5 and CESM FV2 and MPI ESM2 LR from CMIP6. These eight models have been added in retroactively to figures 2-3 for comparison of their performance in each scoring criterion relative to the rest of the CMIP model suites.

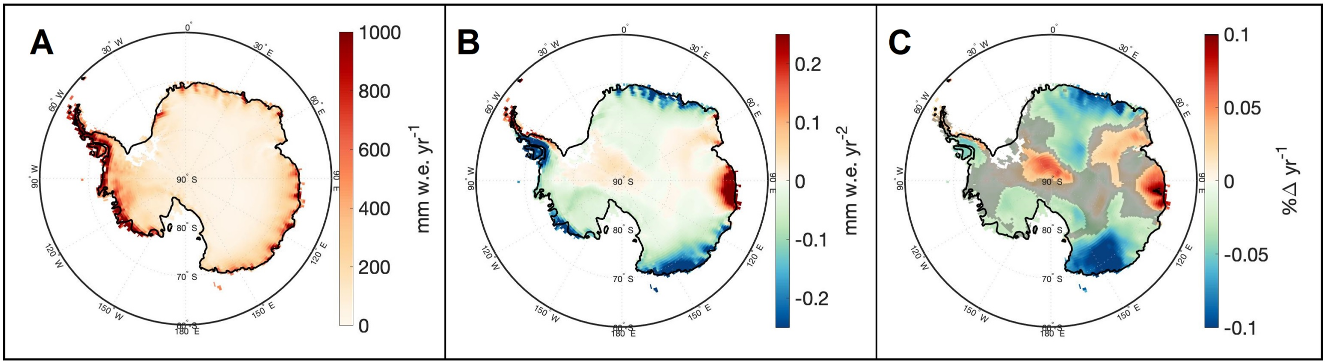


**Figure 2.** A spatial map of **A** the temporal average from 1801-2000 of the reconstructed AIS SMB, **B** the linear trend from 1801-2000 of the reconstructed AIS SMB, and **C** the relative SMB trend in percent SMB change per year. Non-shaded regions in panel **C** denote areas that are statistically significant.

Along with higher SMB values, the coastal regions of East Antarctica and the Antarctic Peninsula also show the highest absolute SMB trends (Fig. 2**B**). This reconstruction also highlights large portions of East Antarctica as well as the Antarctic Peninsula as the regions with the most significant SMB trends from 1801-2000 (Fig. 2**C**). Taking the spatial average but keeping the temporal information yields the AIS-integrated, reconstructed SMB time series shown in Fig. 3**C** (black).

Panel (A) in Fig. 3 shows an example box plot for a suite of models in yellow and the reconstructed observations in black and grey. Panel (B) in Fig. 3 shows a box plot of the temporal average of the spatially integrated AIS SMB for CMIP5 and CMIP6. The interquartile range of AIS-integrated SMB in the CMIP5 models is between 1727 and 2282 Gt yr$^{-1}$ compared to the CMIP6 models whose interquartile range is between 1728 and 2196 Gt yr$^{-1}$. The best eight models range from 1909 to 2461 Gt yr$^{-1}$ for the temporal average AIS-integrated SMB mean value.

The reconstructed AIS SMB ranges from 1800 $\pm$ 338 Gt yr$^{-1}$ from 1850-1900 to 2039 $\pm$ 333 Gt yr$^{-1}$ from 1950-2000. All but one of the eight of best scoring models are fully captured within the reconstructed uncertainty for the entire 150 year time series. The reconstruction and best scoring models all show generally increasing SMB from 1850-2000, albeit with large interannual variability. Both the trend and variability are analyzed in follow-up evaluations and scoring.

While the reconstructed SMB time series and eight best scoring models show a generally increasing trend, the same is not true for all CMIP5 or CMIP6 models (Fig. 4). Looking at multiple time "slices" allows us to investigate if models capture the reconstructed SMB trends for the whole time series compared to more recent decades. Here, we looked at three time slices: the entire overlapping time series from 1850-2000, the last century from 1900-2000, and the last 50 years from 1950-2000. The reconstructed linear SMB trends for the three time slices are 0.52 $\pm$ 0.27 Gt yr$^{-2}$ (1850-2000), 0.56 $\pm$ 0.38 Gt yr$^{-2}$

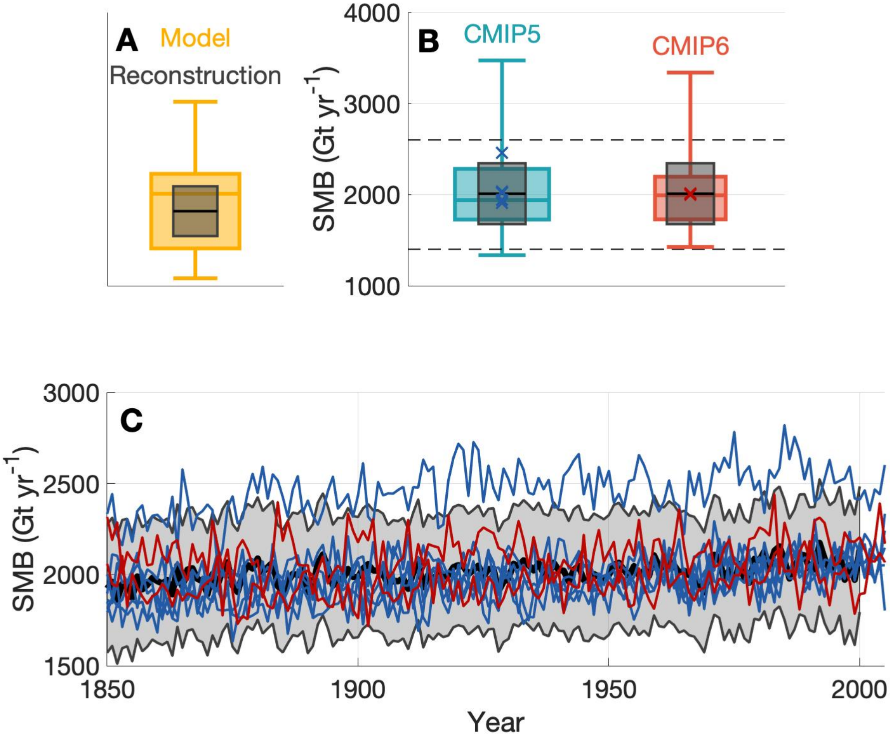


**Figure 3. A** An example of a box plot for model data (yellow) and reconstructed data (black and grey). The yellow shaded box shows the models' interquartile range while the whiskers extend to capture the entire distribution of modeled data. The line going through the box plot shows the median model value. The grey shaded box shows the reconstructed uncertainty around the reconstructed value shown as a black line. **B** A box plot of spatially integrated, temporally averaged (1850-2000) AIS SMB for CMIP5 (aqua) and CMIP6 (red). The dark blue x's associated with the CMIP5 box and the red x's associated with the CMIP6 box represent the eight best scoring models: GISS E2 H CC, GISS E2 R CC, GISS E2 R, MPI ESM LR, MPI ESM MR, and MPI ESM P from CMIP5 and CESM FV2 and MPI ESM2 LR from CMIP6. The black dashed lines indicate the lower and upper bounds of the time series plot in the bottom of Figure 3. **C** A time series of spatially integrated SMB for the reconstruction (black) and its uncertainty (shaded grey) with the best eight scoring models: GISS E2 H CC, GISS E2 R CC, GISS E2 R, MPI ESM LR, MPI ESM MR, and MPI ESM P from CMIP5 (dark blue) and CESM FV2 and MPI ESM2 LR from CMIP6 (red)

(1900-2000), and $1.0 \pm 1.3$ Gt yr$^{-2}$ (1950-2000). That implies that for all but the last time slice, 1950-2000, the reconstruction uncertainty trends are exclusively positive.

Looking at all of the CMIP5 and CMIP6 models, the median linear trend is positive for all three time slices and the trend interquartile ranges are from -0.8 to +1.8 Gt yr$^{-2}$ for 1850-2000, -0.6 to +1.7 Gt yr$^{-2}$ for 1900-2000, and 0.8 to +2.7 Gt yr$^{-2}$ for 1950-2000. For CMIP5, median trends for these time slices are 0.88 Gt yr$^{-2}$, 0.66 Gt yr$^{-2}$, and 1.8 Gt yr$^{-2}$ for 1850-2000, 1900-2000, and 1950-2000 respectively. For CMIP6, median trends for these time slices are 0.05 Gt yr$^{-2}$, 0.46 Gt yr$^{-2}$, and

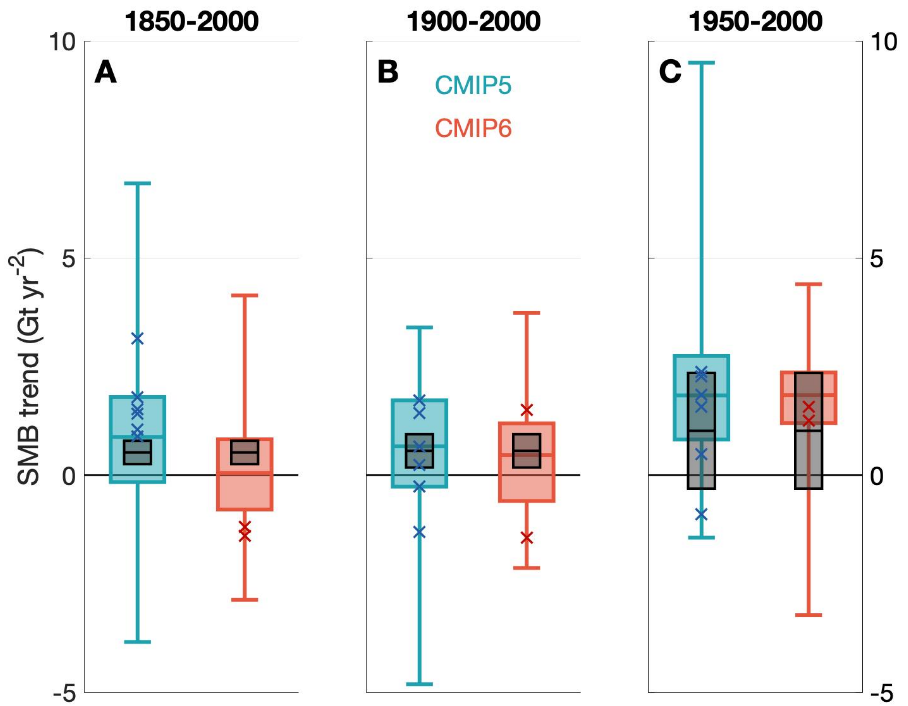


**Figure 4.** Box plots of the linear trends in spatially integrated AIS SMB in CMIP5 (blue) and CMIP6 (red) for the periods **A** from 1850 to 2000; **B** from 1900 to 2000 ; and **C** from 1950 to 2000. In all three panels, the grey boxes denote the reconstructed uncertainty around the reconstructed trend (black line). The eight best scoring models are represented by dark blue x's if they are among the CMIP5 suite of models or red x's if they are among the CMIP6 suite.

1.8 Gt yr$^{-2}$ for 1850-2000, 1900-2000, and 1950-2000 respectively. The eight best scoring models range from -1.4 to +3.1 Gt yr$^{-2}$, -1.4 to +1.7 Gt yr$^{-2}$, and -0.9 to +2.4 Gt yr$^{-2}$ for the same respective time spans. The spread in the eight best scoring models reduces the total spread by 57%, 62%, and 70%, respectively. For the first two time slices, the reconstructed trend and uncertainty are captured within the interquartile range for all CMIP5 models. For 1950-2000, the models tend to overestimate the reconstructed trend.

Apart from its trend magnitude and sign, SMB variability is also important for accurately representing SMB, and can be indicative of the relevant SMB driving mechanisms. Figure 5**A**-**B** shows the average detrended and normalized variability for CMIP5 and CMIP6 models as well as the reconstruction plotted as a normal distribution. The detrended and normalized interannual variability in SMB in the reconstruction ranges between ∼-20% to 20%, while SMB in all the models varies between ∼-15 to 15%. Figure 5**C** shows a box plot the standard deviations of the normalized and detrended time series. The normalization process made it such that the standard deviations are calculated in % of variability about the mean value of the

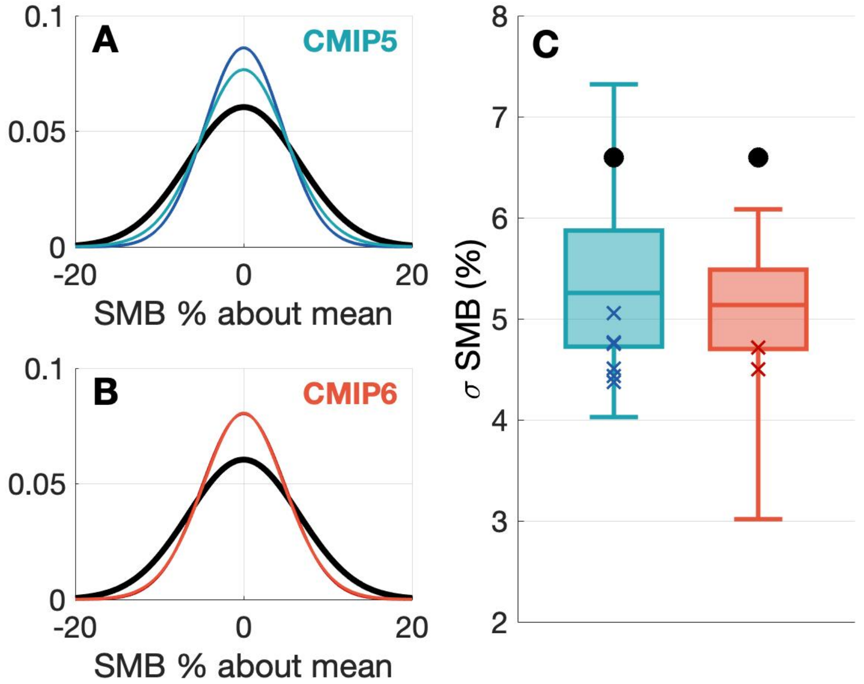


**Figure 5.** Gaussian distributions of SMB where the standard deviation is that of the SMB time series for the reconstruction (black) and **A** all CMIP5 models in light blue and the best scoring CMIP5 models in dark blue and **B** all CMIP6 models in light red and all CMIP6 models in red (the two Gaussians, here, are largely indistinguishable by eye as they overlap almost entirely). **C** Box plots of the CMIP5 (blue) and CMIP6 (red) SMB time series standard deviations. The black dots show the standard deviation of the reconstruction.

time series. The standard deviation for the normalized and detrended SMB in the reanalysis is about 6.6% compared to the best eight models which range between 4.4% to 5.1%. (For comparison, the reconstructed normalized and detrended SMB standard deviation is about 2.9%.) Most CMIP5 and CMIP6 models underestimate SMB variability. The CMIP5 and CMIP6 models' standard deviations range from 4.0% to 7.3% and from 3.0% to 6.1%, respectively (Fig. 5**C**).

Just as temporal SMB variability is important for accurately capturing AIS SMB, spatial variations in SMB are also important in AIS SMB representation in models as precipitation is not distributed uniformly. To look at the spatial variability in SMB, we performed EOF analysis and plotted looked at the top three modes of variability which collectively account for 76.3% of the total spatial variability.

Separated out, the top three modes of variability in the reconstruction from EOF analysis explain 39%, 26%, and 12% of the total variability, respectively (Fig. 6). High values on the EOF map indicate regions that explain large amounts of the variability in AIS SMB. The top mode of variability in the reconstruction shows a dipole pattern from the Antarctic Peninsula to the Ross Sea region. Mode 2 of the reconstruction EOF shows a strong signal over the entire Antarctic Peninsula and toward the Ross

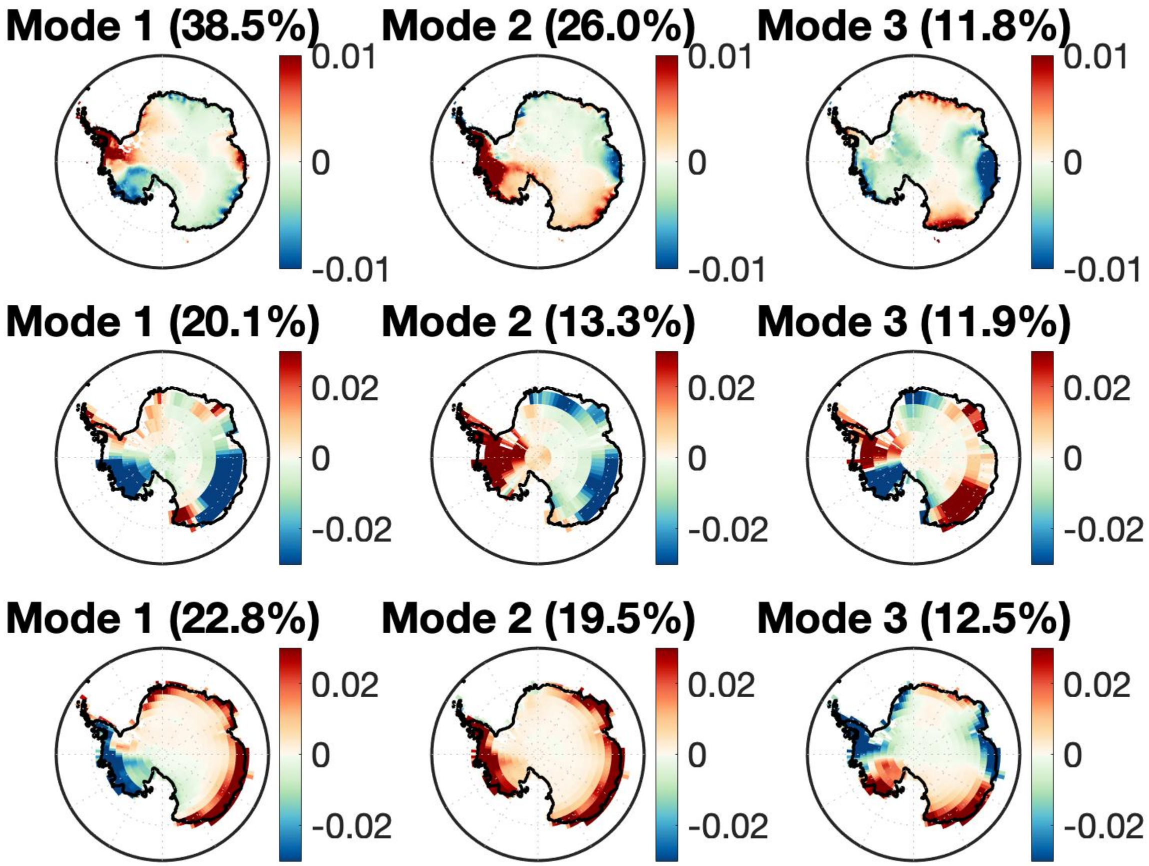


**Figure 6.** EOF analysis plots of the top 3 modes of variability for **A** the reconstruction, **B** a relatively high scoring model (CMCC CM), and **C** a low scoring model (CESM1 WACCM). Note that the scale for the model EOFs is 3× that of the reconstructed EOF.

Ice Shelf region of West Antarctica. The third mode of variability shows a strong signal in Wilkes Land (East Antarctic region), near the Davis Sea, and two opposite, weaker signals in Dronning Maud Land (Atlantic sector) and Adélie land (Pacific sector). This signal is reflective of the linear trend in SMB as seen in Fig. 2**B**.

As a example of the comparison, one of the better scoring models for the EOF map criterion, CMCC CM, also shows a dipole between the Antarctic Peninsula and the Ross Sea region for the top mode as well as strong variance signal around the Antarctic Peninsula for mode 2 and a quadrupolar pattern for mode 3. However, even the better scoring models tend to overestimate the magnitude of the variance particularly around the coast even when they capture the general spatial patterns. CESM1 WACCM, one of the poorer performing models with regard to this metric, generally overestimates the variance everywhere in all three of the top modes. The top mode for this model reflects an East/West Antarctic SMB dipole and mode 2 shows a strong, unidirectional signal across the entire AIS, though mode 3 seems to reflect the same quadrupolar pattern as seen in the reconstruction, albeit with a much higher magnitude.

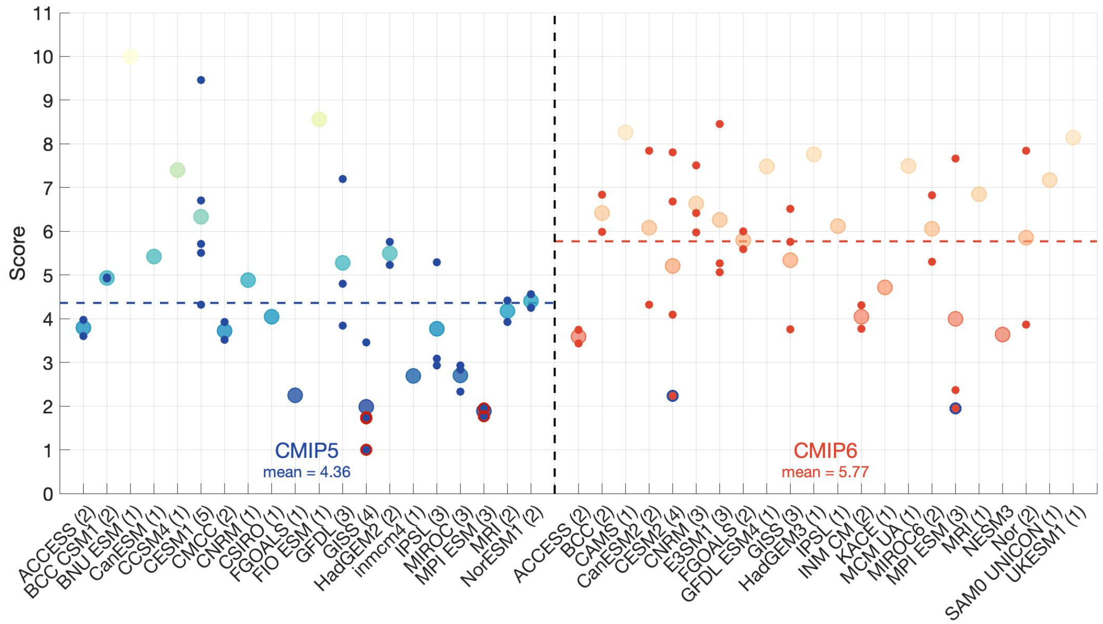


**Figure 7.** The scores for all CMIP5 and CMIP6 models. The large dots show the average score for all model groupings. Models are grouped by similar model physics and have in parenthesis the number of models in the grouping after the name. Each model grouping has all model scores plotted as small blue/red dots for CMIP5/6 with the model average plotted in the larger dots. Models that have no like models are followed by a one in parenthesis and only have a larger dot. The eight best scoring models (above the 90th percentile) are denoted with red outlines if they are among the CMIP5 suite of models – GISS E2 H CC, GISS E2 R CC, GISS E2 R, MPI ESM LR, MPI ESM MR, and MPI ESM P – or with blue outlines if they are among the CMIP6 suite of models – CESM FV2 and MPI ESM2 LR. Note that the overall scores for two of the GISS models and three of the MPI models in CMIP5 are almost exactly equal so outlines overlap almost completely.

Models that score above the 90th percentile make up the subset of best scoring models. Eight models – GISS E2 H CC, GISS E2 RCC, GISS E2 R, MPI ESM LR, MPI ESM MR, and MPI ESM P from CMIP5 and CESM FV2 and MPI ESM LR from CMIP6 – comprise this top 90th percentile. The two CMIP6 models as well as MPI ESM P GISS E2 R from CMIP5 do not appear in the future projections analysis as CMIP6 does not follow the same RCP structure as CMIP5 and the MPI ESM P model does not contain the necessary information to perform the analysis. The poorest performing models include BNU ESM, CESM FASTCHEM, and FIO ESM. The mean model score is 4.36 for CMIP5 and 5.77 for CMIP6. CMIP5 and CMIP6 scores were normalized together such that all scores are on the same scale and are directly comparable. With that, there is not much change from CMIP5 to CMIP6.

With this subset of the eight best performing models, we then refined future projections of AIS SMB in terms of mean value, trend, and variability. Because there are currently an insufficient number of future model runs available for CMIP6, our projection efforts were solely based on CMIP5. Comparing the difference in SMB projections between RCPs allows us a look into the different potential sea level changes caused by different amounts of warming. In CMIP5, there are 25 model outputs for RCP2.6 and 32 model outputs for RCPs 4.5 and 8.5.

As stated earlier, both mean value and trend of AIS SMB have significant implications for future projections of sea level change. The spatially integrated AIS SMB (i.e. SMB mean value) has been increasing from 1850-2000 (Fig. 3) and is projected to continue to increase for the following hundred years to 2100 in all three warming scenarios (Fig. 8).

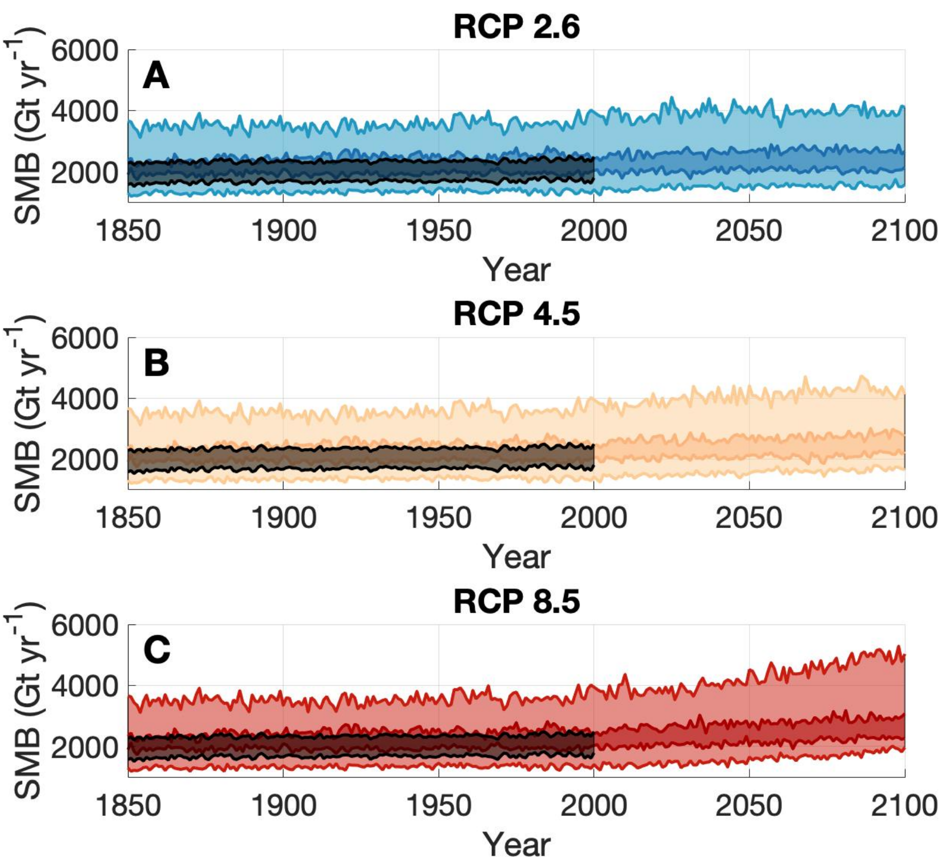


**Figure 8.** Time series for the reconstruction with uncertainty bounds (grey), all CMIP5 models (light) and best scoring CMIP5 models (dark) for **A** RCP2.6 (blue), **B** RCP4.5 (yellow), and **C** RCP8.5 (red).

From 2070-2100, spatially integrated AIS SMB is projected to be 2751 $\pm$ 570 Gt yr$^{-1}$ for RCP2.6, 2948 $\pm$ 581 Gt yr$^{-1}$ for RCP4.5, and 3307 $\pm$ 663 Gt yr$^{-1}$ for RCP8.5 for all CMIP5 models where the associated uncertainties are 1-$\sigma$ of all models between 2070-2100 (for a list of projected SMB and related variable values for all models and the best scoring models across the RCPs, see supplementary). The subset of eight best scoring models have lower projections and smaller spread at 2372 ± 282 Gt $yr^{-1}$ for RCP2.6, 2452 ± 286 Gt $yr^{-1}$ for RCP4.5, and 2588 ± 291 Gt $yr^{-1}$ for RCP8.5 on average between 2070-2100. The ranges of the best eight scoring models reduced the spread by 79%, 79%, and 74% for RCPs 2.6, 4.5, and 8.5, respectively. The mean value of modeled SMB increases with increasing warming scenarios for all CMIP5 models and the subset of the eight best scoring models. Similarly to the mean value increasing with increasing warming, the projected SMB trend also increases with increased warming (Fig. 9). As such, the stronger the emission scenario, the larger the projected response in AIS SMB with regard to both mean value and trend.

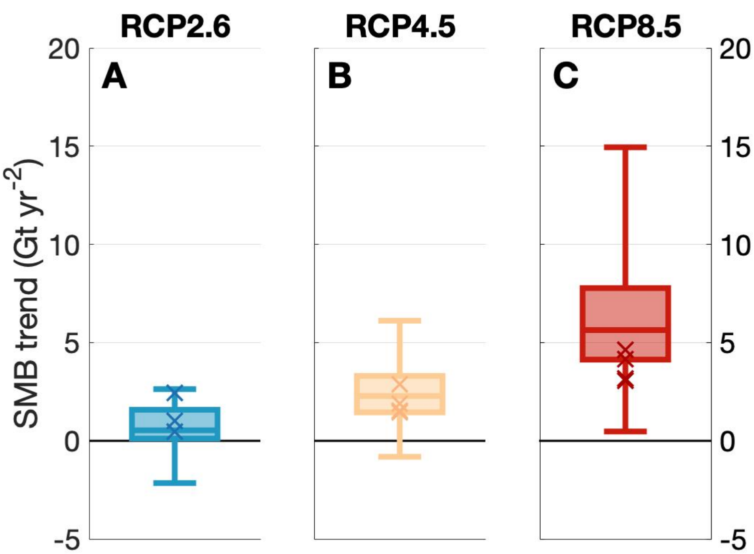


**Figure 9.** Box plots of the linear trend in spatially integrated AIS SMB from 2050-2100 for **A** RCP2.6 (blue), **B** RCP4.5 (yellow), and **C** RCP8.5 (red). The four darker x's denote the four models – GISS E2 H CC, GISS E2 R CC, MPI ESM LR, and MPI ESM MR – among the eight best scoring models with the appropriate and necessary information for direct comparison of future projections.

For the entirety of the 21st century, 2000-2100, most CMIP5 climate models project positive SMB trends in all forcing scenarios (Fig. 9). For RCP2.6, all CMIP5 models project a median trend of 0.53 Gt $yr^{-2}$ and a range of -2.15 to +2.63 Gt $yr^{-2}$. For RCPs 4.5 and 8.5, the median trends are 2.28 Gt $yr^{-2}$ and 5.64 Gt $yr^{-2}$ with ranges of -0.81 to +6.11 Gt $yr^{-2}$ and 0.47 to 14.9 Gt $yr^{-2}$, respectively.

The best scoring models range from 0.47 to 2.45 Gt $yr^{-2}$, 1.44 to 2.88 Gt $yr^{-2}$, and 3.06 to 4.63 Gt $yr^{-2}$ for RCPs 2.6, 4.5, and 8.5, respectively. For RCPs 2.6 and 4.5, the best scoring model trend projections lie close to or within the interquartile range for all CMIP5 models. As the warming scenarios strengthen, the four of the eight best scoring models projected into the future move closer to the lower end of the overall CMIP5 interquartile range in trend. Some of the differences in these concentration pathways can be described by the modeled SMB sensitivity to different atmospheric $CO_2$ emission scenarios.

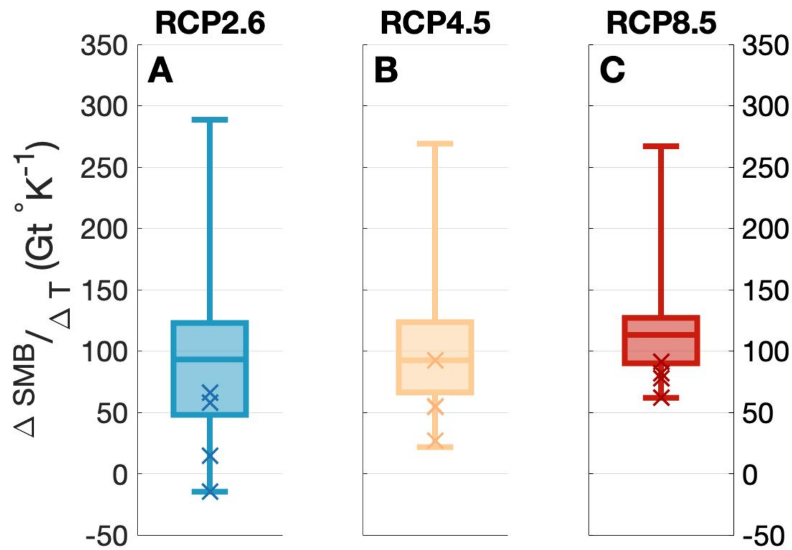


**Figure 10.** Box plots of all CMIP5 models' projected SMB sensitivity to temperature changes ($\Delta$SMB/$\Delta$T) for **A** RCP2.6, **B** RCP4.5, and **C** RCP8.5. The five darker x's denote the four models – GISS E2 H CC, GISS E2 R CC, MPI ESM LR, and MPI ESM MR – among the eight best scoring models with the appropriate and necessary information for direct comparison of future projections.

Box plots of modeled SMB sensitivity to changes in temperature (i.e. how much SMB will change per degree warming) are shown in Fig. 10. The projected sensitivity medians for RCPs 2.6, 4.5, and 8.5 are 101.7 Gt $^{\circ}$K$^{-1}$, 111.2 Gt $^{\circ}$K$^{-1}$, and 128.2 Gt $^{\circ}$K$^{-1}$, respectively. These results are not statistically significantly different from one another, indicating no significant more-than-linear SMB increase in enhanced warming scenarios.

## 5 Discussion

### 5.1 EOF Analysis

Mode 1 of the reconstruction EOF shows a dipolar pattern across the Antarctic Peninsula and Ross Ice Shelf region of West Antarctica. This dipole corresponds to variability in precipitation generated by variations in the track and strength of the Amundsen Sea Low. The Amundsen Sea Low, a dominant synoptic phenomenon that drives a significant amount of the circulation variability in West Antarctica and on the Antarctic Peninsula (Turner et al., 2013), is marked by high precipitation around the coast of the Antarctic Peninsula (Grieger et al., 2016). Changes in the Amundsen Sea Low synoptic pattern, then, represent the dominant cause of variability in the reconstruction SMB. The depth of the ASL is strongly influenced by the phase of the Southern annular mode (SAM) with positive (negative) mean sea level pressure anomalies when the SAM is negative (positive) (Turner et al., 2013).

Looking at mode 2, previous work by Hosking et al. (2013) and Turner et al. (2013) (among others) have shown that variability in the Amundsen Sea Low is responsible for high precipitation variability in West Antarctica and on the Antarctic Peninsula. Because this region dominates the overall AIS precipitation signal (as East Antarctica sees little snowfall by comparison), a variable Amundsen Sea Low signal, here, would explain the EOF pattern reflected in mode 2 of the reconstruction. Additional work highlighted in the supplementary material indicates that variability in sea level pressure in the Amundsen Sea region may be playing a large role in the AIS SMB spatial variability patterns.

### 5.2 Impact of Internal Variability in Model Scoring: CESM Large Ensemble

The CESM Large Ensemble (CESM-LENS) is an experiment wherein the Community Earth System Model Version 1 (CESM) is run 40 times with random temperature perturbations at the level of round-off error applied in 1920 (Kay et al., 2015). Because of its large number of ensemble members, the CESM-LENS experiment is useful for quantifying the role of internal variability. Only 35 of the original 40 ensemble members contain the necessary information for assessing AIS SMB. Figure 4 in Supplementary shows the final scores of the five CESM simulations that are included in the CMIP5 suite of models as well as the final scores of the CESM-LENS experiment. The final scores for the CESM-LENS model runs are calculated the same way for all model criteria except for AIS-integrated trend. Because these runs only differ after 1920, we only use the third time slice (1950-2000) to assess the quality of trend reproduction.

The final scores of the five CMIP5 CESM model runs range from 3.99 to 9.74 while the final scores of the 35 CESM-LENS runs range from 1.32 to 5.96. Given that the scores range by 5.74 and 4.65 for the CMIP5 CESM runs and the CESM-LENS runs, respectively, it is reasonable to conclude that internal variability plays as significant a role in determining final score as do model parameterizations.

A major caveat of this finding, though, is that the CESM-LENS runs and the reconstruction only overlap from 1920-2000. This will likely most significantly impact the assessment of the trend and EOF analyses.

With that, internal variability plays a significant role in our AIS SMB assessment. Some models within the CMIP5 and CMIP6 frameworks, such as CESM1-CAM5, have many ensemble members. However, not all models – and even not all model versions – have multiple ensemble members. As such, performing a direct comparison of the models using the ensemble mean would not necessarily yield an accurate result as models with more ensemble members would have their final score shifted significantly while the same is not true for models with a single ensemble member. For considering using GCMs for AIS SMB analysis, then, we strongly suggest taking into account the fact that internal variability could be playing a strong role in some models final score, and that the number of ensemble members available should be considered along with the final score.

### 5.3 Impact of Model Resolution in Model Scoring

The CMIP5 and CMIP6 models vary in resolution from about $0.75°\times0.75°$ to $3°\times3°$ (Tables 1-3 in Supplementary). Figure 5 in Supplementary shows a scatter plot of resolution versus total score. Resolution, here, is the latitudinal resolution multiplied by the longitudinal resolution such that a model with latitude/longitude resolutions 0.9375°/1.25° would have a resolution of

1.1719°. A linear regression yields a correlation of R = -0.40 with 95% confidence intervals of -0.62 and -0.17. From this, there is a statistically significant negative correlation between resolution and total model score, signaling that, perhaps contrary to intuition, lower-resolution models score equally well, if not better, than higher resolution models. This result might be skewed by the fact that lower-resolution models include better physics to represent AIS SMB than higher-resolution models. However, when comparing total scores from the same model run at different resolutions, we find a consistent result: the relative high-resolution CESM CAM5, IPSL CM5A MR, MPI ESM MR, CESM2, CESM2 WACCM, and MPI ESM2 HR all perform worse than their coarser resolution counterparts – CESM CAM5 FV2, IPSL CM5A LR, MPI ESM LR, CESM2 FV2, CESM2 WACCM FV2, and MPI ESM2 LR. Because so many models close to 1°/1° resolution and there is large spread in these models' final scores, we also divided the models into two groups, finer and coarser than 1.25°/1.25°, and performed the same regression analysis. Figure 6 in Supplementary shows the coarser resolution models have a correlation of R = -0.14 with 95% confidence intervals of -0.51 and 0.24 while finer resolution models have a correlation of R = -0.06 with 95% confidence intervals of -0.38 and 0.26. From this, we conclude that there is no significant correlation between model resolution and total score.

## 6 Conclusions

In this paper, we tested the ability of the suite of models in CMIP5 to capture SMB reconstructed from ice cores and reanalysis products by scoring them using a series of criteria: AIS-integrated mean value, trend, and variability, as well as the spatial variability patterns. This scoring system is designed as a guide for choosing what GCMs to focus on studying for future SMB projections. Using this scoring system, we found that the top 90th percentile models were GISS E2 H CC, GISS E2 R CC, GISS E2 R, MPI ESM LR, MPI ESM MR, and MPI ESM P of CMIP5 and CESM FV2 and MPI ESM2 LR of CMIP6. A similar study in Agosta et al. (2015) found ACCESS1-3, ACCESS1-0, CESM BGC, CESM CAM5, NorESM1-M, and EC-Earth to most accurately capture AIS sea level pressure, 850 hPa air temperature, precipitable water, and ocean conditions – all of which impact AIS SMB to varying degrees. They focused their investigation into more atmospheric and oceanic dynamics (sea ice extent, sea surface temperature, sea surface pressure, precipitable water, 850 hPa temperature) and were comparing models directly to a reanalysis product. Barthel et al. (2019), another study with a similar goal of analyzing SMB performance among GCMs selected CCSM4, MIROC ESM CHEM, and NorESM1-M as their top three performing models for Antarctica. They ruled out both the GISS and MPI modeling groups due to their initial selection criteria and were also looking more at the impacts thermodynamical processes on SMB.

Our SMB mean value estimates are comparable to Agosta et al. (2019), who found a mean SMB value of roughly 2100 $\pm$ 100 Gt yr$^{-1}$ for the grounded AIS using ERA-Interim products. The SMB trends are also in line with Medley and Thomas (2019) over the 20th century. Unlike previous studies, we use a reconstructed data set based on ice core reanalysis, not RCMs. Also of note is the fact that this data set and the GCMs we use for comparison allow us to investigate much longer time periods (150 years), enhancing the robustness of long-term AIS SMB trends. Using this reconstruction, we are able to refine estimates of SMB mean value and SMB trend by the end of the 21st century using CMIP5 by assigning scores to the models and creating a subset of the most accurate models historically. Also unlike previous studies, we analyze both CMIP5 and the early models of CMIP6 together allowing for direct comparison between the two suites of models. The scores for all CMIP5 models are, on average, better than the average score of the currently released CMIP6 models.

All scores are equally weighted to avoid issues with coincidental good or bad performance. Having a spread of criteria against which we score the models limits the possibility that models are recreating one aspect well for the wrong reasons. This scoring method does well in determining simple and consistent criteria to score the accuracy of modeled SMB. In contrast, it struggles to recognize any difference in the importance of individual criteria as they are all weighted equally and also only reflects a few, simple scoring metrics. The criteria were chosen such that they all carry equal weight which we justify by arguing that not meeting any one of the criteria to within a reasonable degree would significantly impact future SMB estimates.

Using the top eight best scoring models, four of which we were able to project out to 2100 under three different RCPs, we refined future SMB predictions to $2372 \pm 282$ Gt yr$^{-1}$ for RCP2.6, $2452 \pm 286$ Gt yr$^{-1}$ for RCP4.5, and $2588 \pm 291$ Gt yr$^{-1}$ for RCP8.5. Over the 21st century this translates to 8.6 cm, 9.6 cm, and 11 cm of GMSL rise buffering in RCPs 2.6, 4.5, and 8.5, respectively, for all of CMIP5. Our result of these best scoring models projecting AIS SMB at the lower end of the overall CMIP5 interquartile range in trend is in contrast to Palerme et al. (2017) who found that, especially considering RCPs 2.6 and 4.5, the CMIP5 models that best captured snowfall change rates tended to predict higher snowfall rates into the 21st century. Additionally, model trends were refined to 0.47 to 2.45 Gt yr$^{-2}$ for RCP2.6, 1.44 to 2.88 Gt yr$^{-2}$ for RCP4.5, and 3.06 to 4.63 Gt yr$^{-2}$ for RCP8.5. Comparing the projected change in SMB per degree warming between the emission scenarios gives median sensitivities of $64 \pm 80$ Gt °K$^{-1}$, $57 \pm 33$ Gt °K$^{-1}$, and $78 \pm 15$ Gt °K$^{-1}$ for RCPs 2.6, 4.5, and 8.5, respectively, for the best scoring models. However, these results are not statistically significantly different from one another across forcing scenarios and indicate that there is no difference in the sensitivity response to changes in temperature between the three forcing scenarios. Given that the best performing models show lower AIS-integrated SMB values and trends compared to the entire CMIP5 spread indicates less sea level rise mitigation from increasing SMB than is implied by looking at all CMIP5 models.

The major limitations of this work stem from the subjective selection of scoring criteria. While each model is scored based on the same criteria, each criterion is chosen specifically to gauge model performance for capturing AIS SMB. As such, these criteria may be ill suited for looking at other variables and, thus, other metrics could yield very different results. Another caveat of this work is that we are only capable of analyzing the CMIP6 models that have been released. As this analysis and the release of CMIP6 are concurrent, this limits the number of models we can reasonably analyze due to time constraints. Additional CMIP6 models may have different results and may skew the comparison between CMIP5 and CMIP6 significantly. Similarly, due to the small number of CMIP6 models released at this point, using statistical analyses becomes moot as the top 90% of models constitutes the single, best scoring model. One final major caveat with this work is the relatively narrow scope of just looking at AIS SMB. Because we refined our criteria at the outset of our experiment to solely reflect model performance with regard to capturing SMB and didn't include outside factors like synoptic weather patterns, sea ice or sea surface conditions (Krinner et al. (2014); Kittel et al. (2018)), there are potentially some wider model biases that we are missing that could affect SMB projections. In our analysis, we make the significant assumption that the past ability to capture SMB correlates to higher skill in projecting AIS SMB into the future. However, model biases in some of the larger physical drivers – and how those biases change into the future – will significantly impact future AIS SMB trajectory.

## Appendix A

### A1

*Author contributions.* T. G. and J. T. M. L. conceptualized and initiated this work. T. G. performed the analysis, discussed the results with J. T. M. L., and wrote the paper. B. M. provided the reconstructions and guidance on using and interpreting them. All authors reviewed the paper before submission.

*Competing interests.* The authors declare no competing interests.

*Acknowledgements.* T. G. and J. T. M. L. acknowledge support from the National Aeronatics and Space Administration (NASA), Grant 80NSSC17K0565 (NASA Sea Level Team 2017–2020).

## References


Agosta, C., Fettweis, X., and Datta, R.: Evaluation of the CMIP5 models in the aim of regional modelling of the Antarctic surface mass balance, The Cryosphere, 9, 2311–2321, https://doi.org/10.5194/tc-9-2311-2015, www.the-cryosphere.net/9/2311/2015/, 2015.

Agosta, C., Amory, C., Kittel, C., Orsi, A., Favier, V., Gallée, H., van den Broeke, M. R., Lenaerts, J. T. M., van Wessem, J. M., and Fettweis, X.: Estimation of the Antarctic surface mass balance using MAR (1979–2015) and identification of dominant processes, The Cryosphere Discussions, pp. 1–22, https://doi.org/10.5194/tc-2018-76, 2019.

Barthel, A., Agosta, C., Little, C. M., Hatterman, T., Jourdain, N. C., Goelzer, H., Nowicki, S., Seroussi, H., Straneo, F., and Bracegirdle, T. J.: CMIP5 model selection for ISMIP6 ice sheet model forcing: Greenland and Antarctica, The Cryosphere Discussions, pp. 1–34, https://doi.org/10.5194/tc-2019-191, 2019.

Beaumet, J., Déqué, M., Krinner, G., Agosta, C., and Alias, A.: Effect of prescribed sea surface conditions on the modern and future Antarctic surface climate simulated by the ARPEGE atmosphere general circulation model, The Cryosphere, 13, 3023–3043, https://doi.org/10.5194/tc-13-3023-2019, https://www.the-cryosphere.net/13/3023/2019/, 2019.

Bromwich, D. H., Nicolas, J. P., and Monaghan, A. J.: An Assessment of precipitation changes over antarctica and the southern ocean since 1989 in contemporary global reanalyses, Journal of Climate, 24, 4189–4209, https://doi.org/10.1175/2011JCLI4074.1, 2011.

Burgener, L., Rupper, S., Koenig, L., Forster, R., Christensen, W. F., Williams, J., Koutnik, M., Miège, C., Steig, E. J., Tingey, D., Keeler, D., and Riley, L.: An observed negative trend in West Antarctic accumulation rates from 1975 to 2010: Evidence from new observed and simulated records, Journal of Geophysical Research Atmospheres, 118, 4205–4216, https://doi.org/10.1002/jgrd.50362, 2013.

Frezzotti, M., Scarchilli, C., Becagli, S., Proposito, M., and Urbini, S.: A synthesis of the Antarctic surface mass balance during the last 800 yr, Cryosphere, 7, 303–319, https://doi.org/10.5194/tc-7-303-2013, 2013.

Fyke, J., Lenaerts, J. T., and Wang, H.: Basin-scale heterogeneity in Antarctic precipitation and its impact on surface mass variability, National Snow and Ice Data Center, 558, 2595–2609, https://doi.org/10.1175/2011JCLI4074.1, http://nsidc.org/cryosphere/quickfacts/icesheets.html, 2017.

Gallée, H., Trouvilliez, A., Agosta, C., Genthon, C., Favier, V., and Naaim-Bouvet, F.: Transport of Snow by the Wind: A Comparison Between Observations in Adélie Land, Antarctica, and Simulations Made with the Regional Climate Model MAR, Boundary-Layer Meteorology, 146, 133–147, https://doi.org/10.1007/s10546-012-9764-z, 2013.

Genthon, C., Krinner, C., and Castebrunet, H.: Antarctic precipitation and climate-change predictions: Horizontal resolution and margin vs plateau issues, Annals of Glaciology, 50, 55–60, https://doi.org/10.3189/172756409787769681, 2009.

Grieger, J., Leckebusch, G. C., and Ulbrich, U.: Net precipitation of Antarctica: Thermodynamical and dynamical parts of the climate change signal, Journal of Climate, 29, 907–924, https://doi.org/10.1175/JCLI-D-14-00787.1, 2016.

Hosking, J. S., Orr, A., Marshall, G. J., Turner, J., and Phillips, T.: The Influence of the Amundsen–Bellingshausen Seas Low on the Climate of West Antarctica and Its Representation in Coupled Climate Model Simulations, Journal of Climate, 26, 6633–6648, https://doi.org/10.1175/JCLI-D-12-00813.1, http://journals.ametsoc.org/doi/10.1175/JCLI-D-12-00813.1, 2013.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J.-F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L., and Vertenstein, M.: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability, Bulletin of the American Meteorological Society, 96, 1333–1349, https://doi.org/10.1175/BAMS-D-13-00255.1, http://journals.ametsoc.org/doi/10.1175/BAMS-D-13-00255.1, 2015.

Kittel, C., Amory, C., Agosta, C., Delhasse, A., Doutreloup, S., Huot, P.-V., Wyard, C., Fichefet, T., and Fettweis, X.: Sensitivity of the current Antarctic surface mass balance to sea surface conditions using MAR, The Cryosphere, 12, 3827–3839, https://doi.org/10.5194/tc-12-3827-2018, https://doi.org/10.5194/tc-12-3827-2018, 2018.

Krinner, G., Largeron, C., Ménégoz, M., Agosta, C., and Brutel-Vuilmet, C.: Oceanic Forcing of Antarctic Climate Change: A Study Using a Stretched-Grid Atmospheric General Circulation Model, Journal of Climate, 27, 5786–5800, https://doi.org/10.1175/JCLI-D-13-00367.1, http://journals.ametsoc.org/doi/10.1175/JCLI-D-13-00367.1, 2014.

Lenaerts, J. T., Van Den Broeke, M. R., Van De Berg, W. J., Van Meijgaard, E., and Kuipers Munneke, P.: A new, high-resolution surface mass balance map of Antarctica (1979-2010) based on regional atmospheric climate modeling, Geophysical Research Letters, 39, 1–5, https://doi.org/10.1029/2011GL050713, 2012.

Lenaerts, J. T., Vizcaino, M., Fyke, J., van Kampenhout, L., and van den Broeke, M. R.: Present-day and future Antarctic ice sheet climate and surface mass balance in the Community Earth System Model, Climate Dynamics, 47, 1367–1381, https://doi.org/10.1007/s00382-015-2907-4, 2016.

Lenaerts, J. T., Medley, B., van den Broeke, M. R., and Wouters, B.: Observing and Modeling Ice Sheet Surface Mass Balance, Reviews of Geophysics, https://doi.org/10.1029/2018RG000622, 2019.

Marshall, G. J., Thompson, D. W., and van den Broeke, M. R.: The Signature of Southern Hemisphere Atmospheric Circulation Patterns in Antarctic Precipitation, Geophysical Research Letters, 44, 580–11, https://doi.org/10.1002/2017GL075998, 2017.

Medley, B. and Thomas, E. R.: Increased snowfall over the Antarctic Ice Sheet mitigated twentieth-century sea-level rise, Nature Climate Change, 9, 34–39, https://doi.org/10.1038/s41558-018-0356-x, http://dx.doi.org/10.1038/s41558-018-0356-x, 2019.

Monaghan, A. and Bromwich, D.: Global warming at the poles, Nature Geoscience, 1, 728, https://doi.org/10.1038/ngeo346http://10.0.4.14/ngeo346, 2008.

Monaghan, A. J., Bromwich, D., Fogt, R., Wang, S.-H., Mayewski, P., Dixon, D., Ekaykin, A., Frezzotti, M., Goodwin, I., Isaksson, E., Kaspari, S., Morgan, V., Oerter, H., Van Ommen, T., Van der Veen, C., and Wen, J.: Insignificant Change in Antarctic Snowfall Since the International Geophysical Year Andrew, Science, 313, 827–831, https://doi.org/10.5061/dryad.5t110.Supplementary, 2006.

Palerme, C., Kay, J. E., Genthon, C., L'Ecuyer, T., Wood, N. B., and Claud, C.: How much snow falls on the Antarctic ice sheet?, Cryosphere, 8, 1577–1587, https://doi.org/10.5194/tc-8-1577-2014, 2014.

Palerme, C., Genthon, C., Claud, C., Kay, J. E., Wood, N. B., and L'Ecuyer, T.: Evaluation of current and projected Antarctic precipitation in CMIP5 models, Climate Dynamics, 48, 225–239, https://doi.org/10.1007/s00382-016-3071-1, 2017.

Philippe, M., Tison, J. L., Fjøsne, K., Hubbard, B., Kjær, H. A., Lenaerts, J. T., Drews, R., Sheldon, S. G., De Bondt, K., Claeys, P., and Pattyn, F.: Ice core evidence for a 20th century increase in surface mass balance in coastal Dronning Maud Land, East Antarctica, Cryosphere, 10, 2501–2516, https://doi.org/10.5194/tc-10-2501-2016, 2016.

Previdi, M. and Polvani, L. M.: Anthropogenic impact on Antarctic surface mass balance, currently masked by natural variability, to emerge by mid-century, Environ. Res. Lett, 11, 94 001, https://doi.org/10.1088/1748-9326/11/9/094001, 2016.

Thomas, E. R., Hosking, J. S., Tuckwell, R. R., Warren, R. A., and Ludlow, E. C.: Twentieth century increase in snowfall in coastal West Antarctica, Geophysical Research Letters, 42, 9387–9393, https://doi.org/10.1002/2015GL065750, 2015.

Thomas, E. R., Melchior Van Wessem, J., Roberts, J., Isaksson, E., Schlosser, E., Fudge, T. J., Vallelonga, P., Medley, B., Lenaerts, J., Bertler, N., Van Den Broeke, M. R., Dixon, D. A., Frezzotti, M., Stenni, B., Curran, M., and Ekaykin, A. A.: Regional Antarctic snow accumulation over the past 1000 years, Climate of the Past, 13, 1491–1513, https://doi.org/10.5194/cp-13-1491-2017, 2017.

Turner, J., Phillips, T., Hosking, J. S., Marshall, G. J., and Orr, A.: The amundsen sea low, International Journal of Climatology, 33, 1818–1829, https://doi.org/10.1002/joc.3558, 2013.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J. F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J., and Rose, S. K.: The representative concentration pathways: An overview, Climatic Change, 109, 5–31, https://doi.org/10.1007/s10584-011-0148-z, 2011.

van Wessem, J. M., van de Berg, W. J., Noël, B. P. Y., van Meijgaard, E., Birnbaum, G., Jakobs, C. L., Krüger, K., Lenaerts, J. T. M., Lhermitte, S., Ligtenberg, S. R. M., Medley, B., Reijmer, C. H., van Tricht, K., Trusel, L. D., van Ulft, L. H., Wouters, B., Wuite, J., and van den Broeke, M. R.: Modelling the climate and surface mass balance of polar ice sheets using RACMO2, part 2: Antarctica (1979–2016), The Cryosphere Discussions, pp. 1–35, https://doi.org/10.5194/tc-2017-202, 2017.