https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

The Cryosphere Discussions
Open Access
EGU

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.

Using a scoring system based on SMB magnitude, 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 2295 $\pm$ 1222 Gt yr^{-1}, 2382 $\pm$ 1316 Gt yr^{-1}, and 2648 $\pm$ 1530 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 78%, 75%, and 78% in RCPs 2.6, 4.5, and 8.5, respectively.

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). Integrated over the grounded Antarctic ice sheet (AIS), the blowing snow and runoff terms are negligibly small (Lenaerts et al., 2012a). 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 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

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

and Thomas (2019)). These studies found that, integrated over the AIS, SMB has been increasing at a rate of 0.4 ± 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). The strong regional variability suggests an important impact of variations in synoptic-scale patterns around the AIS (Fyke et al. (2017); Marshall et al. (2017)). Additionally, as the atmosphere has been warming over large parts of the AIS and can hold more moisture per the Clausius-Clapeyron relation, SMB is expected to show an overall increase. Driven by the same mechanism, models indicate that AIS SMB will increase even further over the next century and beyond (Palerme et al., 2014).

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 due to poor representation of coastal topography. 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 (Lenaerts et al. (2012b)). 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 version 6 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 concept of this paper is very similar, but we use a different observational data set for comparison as well as different scoring criteria. 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). 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 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.

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

2 Data

2.1 SMB Reconstructions

In this paper, we use the AIS SMB reconstruction generated by Medley & Thomas (2019). In their study of AIS SMB, they synthesized ice core records using three different atmospheric reanalysis products: the Climate Forecast System Reanalysis (CFSR), the European Centre for Medium-Range Weather Forecasts 'Interim' (ERA-Interim), and the Modern-Era Retrospective Analysis for Research Applications Version 2 (MERRA-2). To generate the reconstructions, Medley & Thomas, 2019 used 53 ice core records that spanned the entire 19th and 20th centuries. While more ice core records are available across the AIS, they stipulated that the records be annually resolved and must cover the years 1980-1988 to provide sufficient overlap with the reanalysis products that cover 1979/80-2016. To integrate the reanalyses with the ice core records, they created a field of shared variance using coefficient of determination, r^2, for the AIS. Using this spatial field, they weighted each ice core spatially to generate the 200-year data set. They performed bias correction to the overall SMB magnitude of each of the three reanalyses that form the basis of the reconstructions by using observations within the reanalysis time frame and calculating:

$$\text{bias correction} = \frac{\text{model-observations}}{\text{model}} \tag{1}$$

for each grid cell. The reconstruction uncertainty accounted for both measurement error and uncertainty in spatial sampling. The measurement uncertainty is the root mean square error (RMSE) between the ice core records and the reanalyses time series at the grid cell of the ice core. Similarly, they calculated spatial sampling uncertainty is based on the RMSE between the reanalyses and an internal reconstruction that uses the reanalysis time series rather than the ice core records. The total uncertainty, then, is the square root sum of squares of the two sources of uncertainty and varies in both time and space.

While the original three reanalysis products differ substantially in their representations of SMB both spatially and temporally, the three ice core forced reconstructions show very good spatial agreement (Medley and Thomas, 2019). Because of their agreement, we can use any of the three reconstructions interchangeably. As MERRA-2 performed better than the other two reconstructed products in matching observations (Medley & Thomas 2019), we will use it as a proxy for all three reconstructions and refer to it as "reconstruction."

Global climate models tend to show higher skill at representing interannual variability compared to regional climate models (Medley and Thomas, 2019). As such, we can make the most direct comparisons to the reconstruction with global climate models. To get a comprehensive look at how well global climate models capture SMB, we compared the suite of CMIP5 models to the reconstruction.

2.2 Climate Models

We used all applicable CMIP5 and CMIP6 model outputs, of which there were 53 models and 28 independent models (i.e. different model physics and resolutions) respectively, for the historical simulations (1850-2005). We only had available output for 30 CMIP5 models, 19 of which are independent, for the future simulations (2006-2100). See Tables 1-3 in Supplementary

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019


The Cryosphere Discussions

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 data with monthly resolution in units of kg m^{-2} s^{-1}. After calculating SMB as precipitation - evaporation/sublimation, we converted these to annual time scales and integrated them across the AIS using the Ice Sheet Mass Balance Inter-comparison Exercise Team's (IMBIE Team) AIS grounded ice sheet masks and units of Gt yr^{-1} by multiplying each grid cell by its area, converting s^{-1} to yr^{-1}, and converting kg to Gt (1 Gt = 10^{12} kg)Shepherd et al. (2012). We interpolated the IMBIE Team's AIS mask using the nearest sample grid point and applied it to all data sets.

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 magnitude of the SMB time series values, the time series trend, and the time series variability.

To score the time series magnitude, 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. For example, if a model time series was fully captured within 2× the reconstruction uncertainty, the model would receive a score of 2.

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 1850-2000. The reconstructed trend uncertainties were calculated by performing a Monte Carlo simulation assuming a normal distribution of SMB values centered around the reconstructed SMB value with a standard deviation of the reconstruction uncertainty for each year. From those distributions, we generated 10,000 simulated SMB time series based on the reconstruction and calculated the trends for each. The standard deviation in trend, then, is the reconstructed trend uncertainty.

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

To score the time series variability, we detrended and normalized each time series to separate the SMB trend from its absolute magnitude using:

$$\text{normalized SMB} = \frac{\text{SMB} - \text{mean SMB}}{\text{mean SMB}}. \quad (2)$$

We then calculated the standard deviation of each time series and assigned a score, x, based on how many x-times the reconstruction standard deviation were required to capture the model standard deviation.

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. EOF analysis, as applied to these annual data, involves finding what spatial SMB patterns explain the highest variance in the AIS-integrated SMB time. By breaking this criterion down into two main factors, we were able to determine the models' abilities to accurately capture the modes of variability as well as how much variance each mode explained.

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. We then sorted the top modes of variability for each model based on smallest difference thus giving the models the "benefit of the doubt." 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 normalized each set of scores to be on a scale from one to ten to ensure that each criterion was equally weighted. 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

We weighted all scores from the five scoring criteria equally on a scale from 1 to 10 with lower scores indicating better performance. The final score, then, is the sum of all the individual scores, which is renormalized on a scale of 1 to 10 with lower scores still indicating better performance. To refine the scope of what we predict for SMB in the future, we used a subset

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

of models that had a final score in the top 10th percentile of CMIP5 and compared them to the entire scope of CMIP5. There are currently insufficient CMIP6 models to create a similar subset and future projection analysis so all future analysis is restricted to CMIP5.

Using this subset of best scoring models, we calculated the projected AIS-integrated magnitude 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

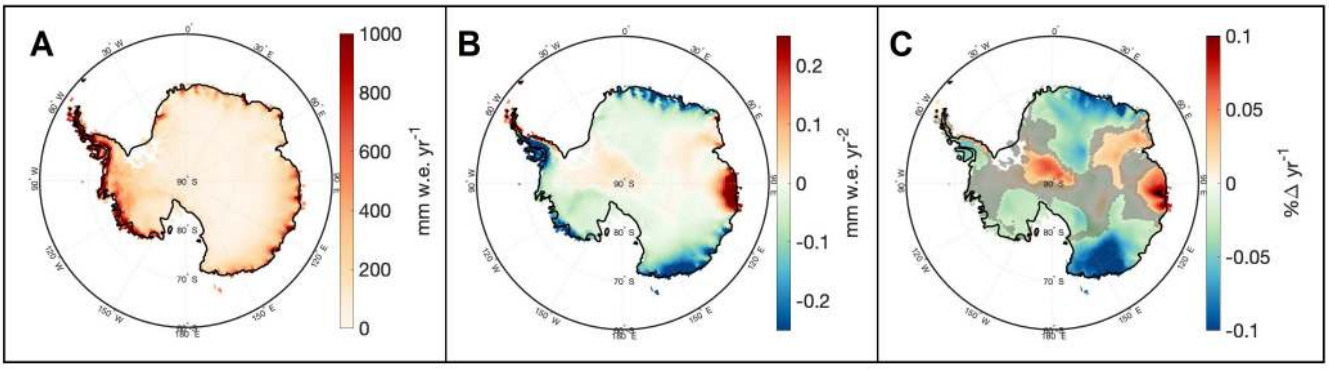

Figure 1. A spatial map of **A** the temporal average from 1801-2000 of the reconstructed AIS SMB, **B** the linear trend from 1801-2000 of the reconstructed AIS SMB, and **C** the relative SMB trend in percent SMB change per year. Non-shaded regions in panel **C** denote areas that are statistically significant.

The reconstructed AIS SMB averaged from 1801-2000 shows higher SMB values around the coastal areas, particularly in the Antarctic Peninsula and West Antarctic regions (Fig. 1**A**). The highest absolute SMB trends are around the coastal regions of East Antarctica and the Antarctic Peninsula (Fig. 1**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. 1**C**). Taking the spatial average but keeping the temporal information yields the AIS integrated, reconstructed SMB time series shown in Fig. 2**C** (black).

Panel (A) in Fig. 2 shows an example box plot for a suite of models in yellow and the reconstructed observations in black and grey. Panel (B) in Fig. 2 shows a box plot of the temporal average of the spatially integrated AIS SMB for CMIP5 and CMIP6. The average AIS-integrated SMB in the CMIP5 models range between 1335 and 3472 Gt yr^{-1} compared to the CMIP6 models which range between 1471 and 3339 Gt yr^{-1}. The interquartile ranges for CMIP5 and CMIP6 are 1727 to 2282 Gt yr^{-1} and 1728 to 2229 Gt yr^{-1}, respectively, with means of 1940 Gt yr^{-1} and 2115 Gt yr^{-1}, respectively.

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

The Cryosphere
Discussions
Open Access
EGU

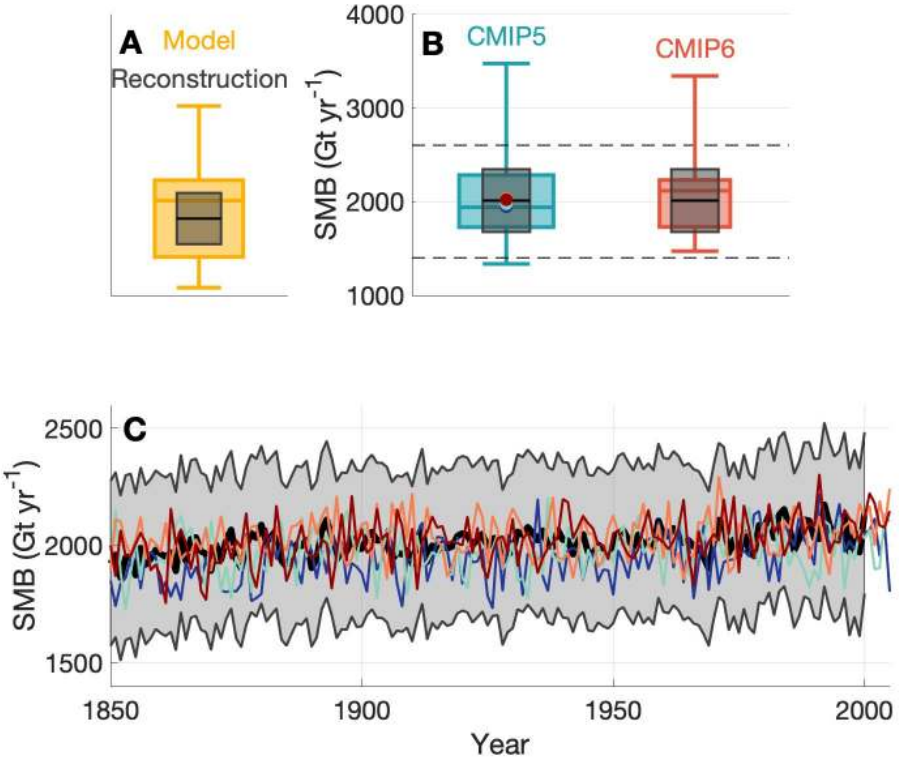

Figure 2. A An example of a box plot for model data (yellow) and reconstructed data (black and grey). The yellow shaded box shows the models' interquartile range while the whiskers extend to capture the entire distribution of modeled data. The line going through the box plot shows the median model value. 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, green, coral, and dark red dots represent the four best scoring models: GISS E2 H, GISS E2 R, MPI ESM LR, and MPI ESM MR, respectively. The black dashed lines indicate the lower and upper bounds of the time series plot in the bottom of Figure 2. **C** A time series of spatially integrated SMB for the reconstruction (black) and its uncertainty (shaded grey) with the best four scoring models: GISS E2 H (dark blue), GISS E2 R (green), MPI ESM LR (coral), and MPI ESM MR (dark red).

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 four of best scoring models are captured within the reconstructed uncertainty for the entire 150 year time series. The
170 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 four best scoring models show a generally increasing trend, the same is not true for all CMIP5 models (Fig. 3). 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 over-

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

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

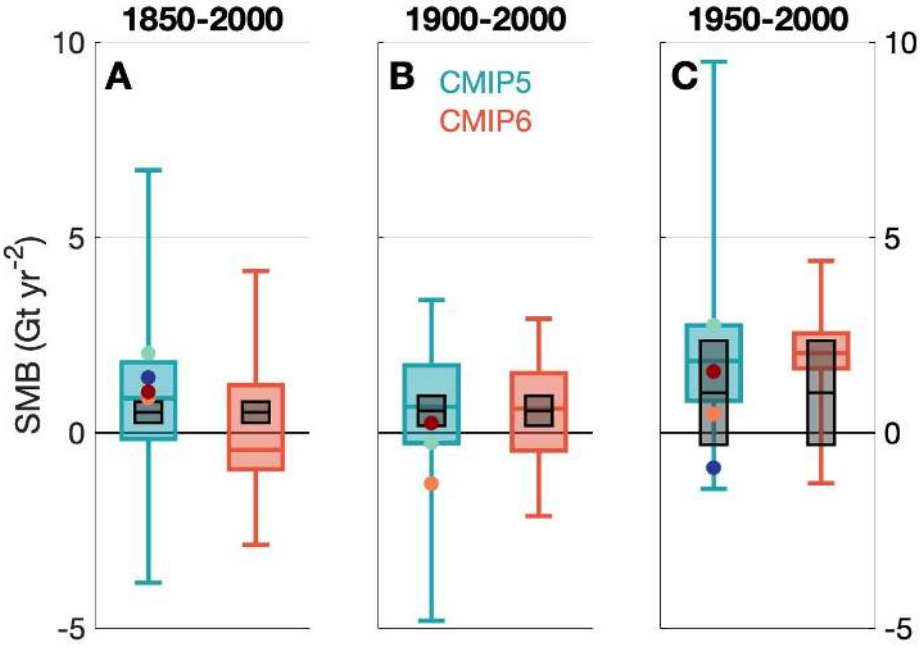

Figure 3. Box plots of the linear trends in spatially integrated AIS SMB in CMIP5 (blue) and CMIP6 (red) for the periods **A** from 1850 to 2000; **B** from 1900 to 2000 ; and **C** from 1950 to 2000. In all three panels, the grey boxes denote the reconstructed uncertainty around the reconstructed trend (black line). The four best scoring models are shows in dark blue, green, coral, and red, with the colors corresponding to the same models as in Figure 2.

Looking at all of the CMIP5 models, the median linear trend is positive and trends range in absolute minimum to absolute maximum from -3.8 to +6.7 Gt yr^{-2} for 1850-2000, -4.8 to +3.4 Gt yr^{-2} for 1900-2000, and -1.4 to +9.5 Gt yr^{-2} for 1950-2000 with median trends of 0.88 Gt yr^{-2}, 0.66 Gt yr^{-2}, and 1.8 Gt yr^{-2}, respectively. The four best scoring models range from -2.5 to +0.81 Gt yr^{-2}, -0.92 to +3.4 Gt yr^{-2}, and -0.33 to +4.4 Gt yr^{-2} for the same respective time spans. The spread in the four best scoring models reduces the total spread by 31%, 52%, and 43%, 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. The four best scoring models are at the lower end of the model estimates and the two MPI ESM models are captured within the reconstructed uncertainty.

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

Similar to CMIP5, the median linear trend is positive for the latter two time slices for CMIP6. The median linear trend in the first time slice, however, is negative in CMIP6, implying that more than half the CMIP6 models produce a negative SMB trend over the 151-year time series. Also similar to CMIP5, the first two time slices also capture the reconstructed trend and uncertainty in the interquartile range while the models tend to overestimate the trend for 1950-2000. The spread in trend in the CMIP6 models is significantly lower than for CMIP5 models which follows as there are fewer models. The trends for CMIP6 range from -2.9 to +4.1 Gt yr^{-2} with a median trend of -0.44 Gt yr^{-2} for 1850-2000, -1.8 to +2.9 Gt yr^{-2} with a median trend of 0.68 Gt yr^{-2} for 1900-2000, and 1.2 to 4.4 Gt yr^{-2} with a median trend of 1.9 Gt yr^{-2} for 1950-2000.

Apart from its trend magnitude and sign, SMB variability is also important for accurately representing SMB and determining the impact on sea level it may have in any given year. SMB variability also is indicative of the relevant drivers behind SMB.

The detrended and normalized interannual variability in SMB in the reconstruction ranges between -6.4-8.0%, while SMB in the best four models varies between ∼-10-10% (Fig. 4**A**-**D**). All CMIP5 and CMIP6 models overestimate SMB variability. The CMIP5 and CMIP6 models range from overestimates of 144% to 261% and 151% to 217% of the reconstruction standard deviation, respectively (Fig. 4**E**).

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 melt and discharge are not distributed equally. To look at the spatial variability in SMB, we performed EOF analysis and plotted looked at the top three modes of variability which 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. 5). The top mode of variability in the reconstruction shows a dipole pattern from the Antarctic Peninsula to the Ross Sea region. This dipole corresponds to variability in precipitation generated by variations in the track and strength of the Amundsen Sea Low. The Amundsen Sea Low, which represents the pole of circulation variability in Antarctica (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). The second mode of variability represents high variability in West Antarctica and the Antarctic Peninsula. This could be caused by the topography in these regions which can induce large amounts of snowfall. 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. 1**B**. See supplemental for further EOF analysis of sea level pressure variability.

By 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 tended 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

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

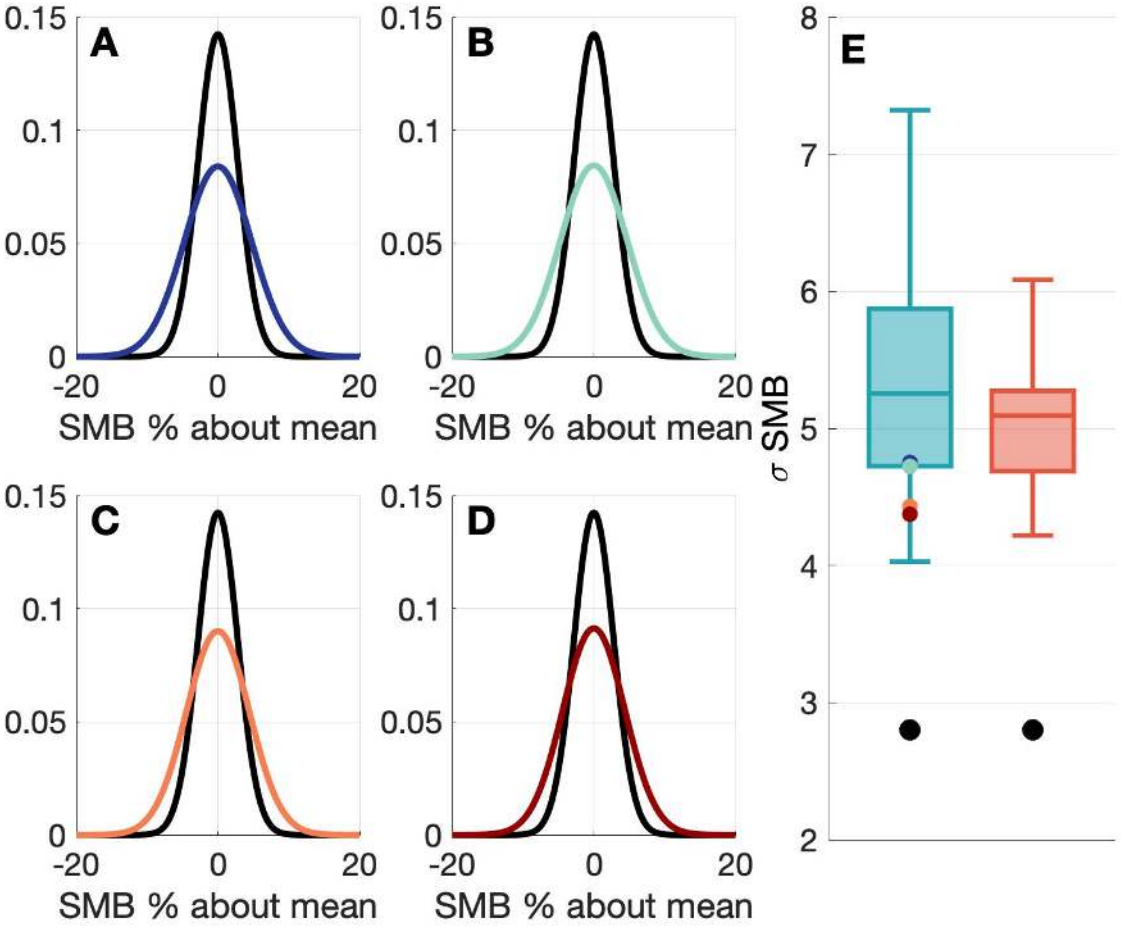

Figure 4. Gaussian distributions of SMB where the standard deviation is that of the SMB time series for the reconstruction (black) and **A** GISS E2 H (dark blue) **B** GISS E2 R (green) **C** MPI ESM LR (coral) and **D** MPI ESM MR (dark red). **E** Box plots of the CMIP5 (blue) and CMIP6 (red) SMB time series standard deviations. The black dots show the standard deviation of the reconstruction.

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 to a much higher magnitude.

225 Models that score above the 90th percentile make up the subset of best scoring models. Five models – GISS E2 H, GISS E2 R, MPI ESM LR, MPI ESM MR, and MPI ESM P – qualify for this status as there is a three-way tie for third, but as MPI ESM P does not have the necessary information for future projections, it is neglected. Similar to CMIP5, the GISS models in CMIP6 are also among the best performing model in the small sample size. The poorest performing models include CESM

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

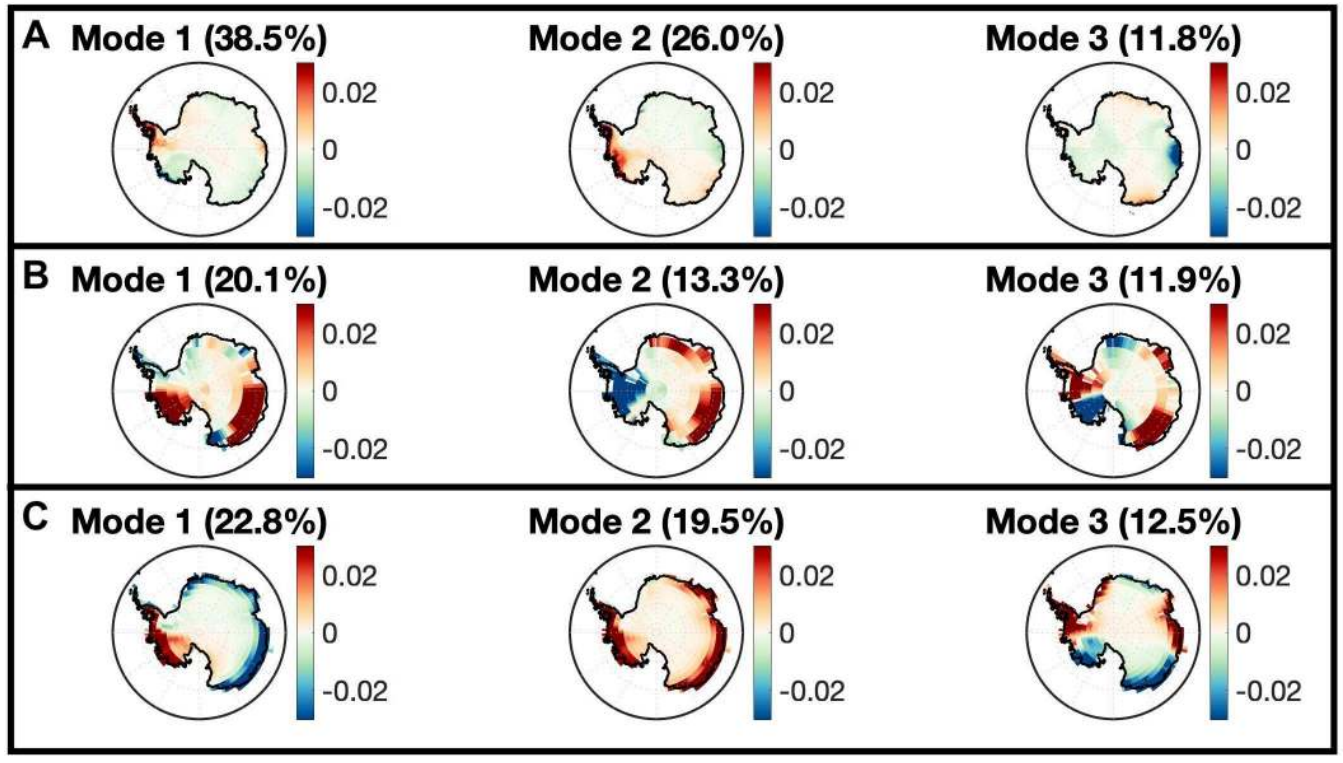

Figure 5. EOF analysis plots of the top 3 modes of variability for **A** the reconstruction, **B** a relatively high scoring model (CMCC CM), and **C** a low scoring model (CESM1 WACCM).

FASTCHEM, BNU ESM, and FIO ESM in CMIP5, and CanESM2 in CMIP6. The mean model score is 3.7 for CMIP5 and 4.5 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. In fact, the scores increase from CMIP5 to CMIP6 albeit with a small sample size of models for CMIP6.

With this subset of the four best performing models, we then refined future projections of AIS SMB in terms of magnitude, 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. Future CMIP5 projections are created in the context of warming scenarios called Representative Concentration Pathways (RCPs). The RCPs we used to investigate SMB projections are RCP2.6, RCP4.5, and RCP8.5 which have progressively higher CO_2 concentration projections and, thus, higher projected global warming. Comparing the difference in SMB projections between these 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.

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

The Cryosphere
Discussions
Open Access
EGU

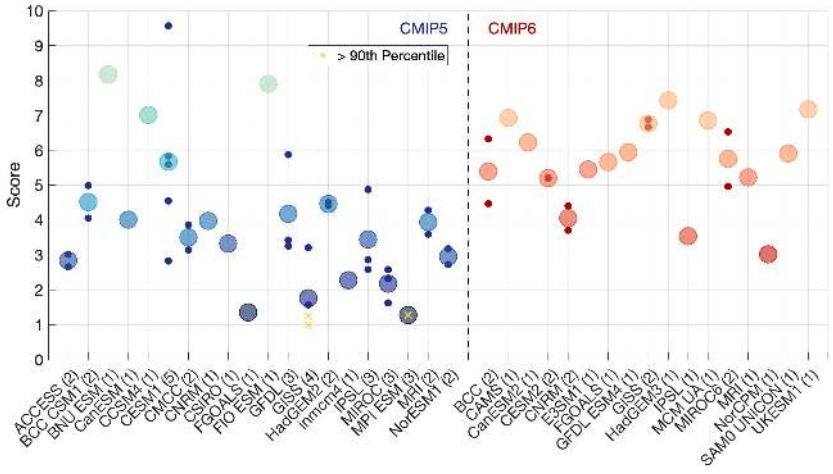

Figure 6. The scores for all CMIP5 and CMIP6 models. The large dots show the average score for all model groupings. Models are grouped by similar model physics and have in parenthesis the number of models in the grouping after the name. Each model grouping has all model scores plotted as small blue/red 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 four best scoring models (above the 90th percentile) are denoted with yellow x's instead of blue dots.

As stated earlier, both magnitude and trend of AIS SMB have significant implications for future projections of sea level change. The spatially integrated AIS SMB (i.e. SMB magnitude) has been increasing from 1850-2000 (Fig. 2) and is projected to continue to increase for the following hundred years to 2100 in all three warming scenarios (Fig. 7).

From 2070-2100, spatially integrated AIS SMB is projected to be 2295 $\pm$ 1222 Gt yr^{-1} for RCP2.6, 2382 $\pm$ 1316 Gt yr^{-1} for RCP4.5, and 2648 $\pm$ 1530 Gt yr^{-1} for RCP8.5 for all CMIP5 models (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 four best scoring models have lower projections and smaller spread at 2246 $\pm$ 268 Gt yr^{-1} for RCP2.6, 2358 $\pm$ 331 Gt yr^{-1} for RCP4.5, and 2495 $\pm$ 335 Gt yr^{-1} for RCP8.5 on average between 2070-2100. The magnitude of modeled SMB increases with increasing warming scenarios for all CMIP5 models and the subset of the four best scoring models. Similarly to the magnitude increasing with increasing warming, the projected SMB trend also increases with increased warming (Fig. 8). As such, the stronger the emission scenario, the larger the projected response in AIS SMB with regard to both magnitude and trend.

For the entirety of the 21st century, 2000-2100, most CMIP5 climate models project positive SMB trends in all forcing scenarios (Fig. 8). 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 0.34 to 2.09 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

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

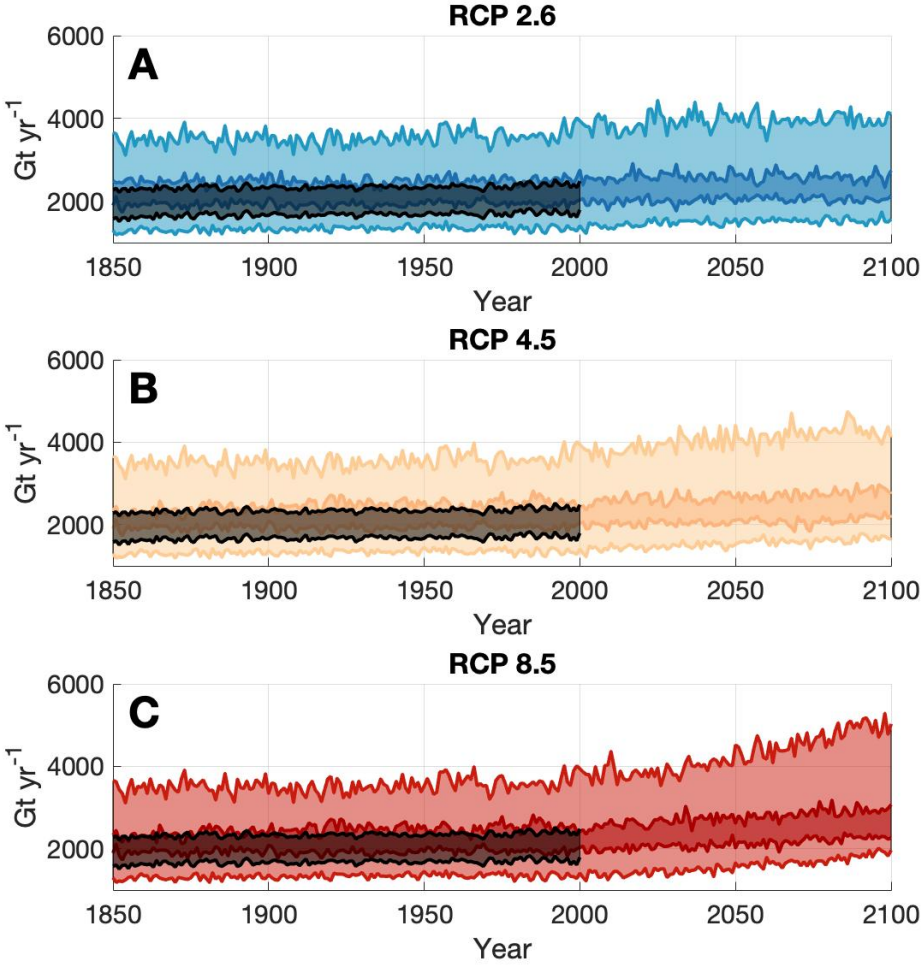

Figure 7. Time series for the reconstruction with uncertainty bounds (grey), all CMIP5 models (light) and best scoring CMIP5 models (dark) for **A** RCP2.6 (blue), **B** RCP4.5 (yellow), and **C** RCP8.5 (red).

range for all CMIP5 models. The best four model projections are near or below the lower bound of the interquartile range for RCP8.5. Some of the differences in these concentration pathways can be described by the modeled SMB sensitivity to different atmospheric CO_2 emission scenarios.

Box plots of modeled SMB sensitivity to changes in temperature (i.e. how much SMB will change per degree warming) show that SMB responds differently in different warming scenarios (Fig. 9). The CMIP5 models project that each warming scenario with higher CO_2 concentrations will see greater SMB sensitivity to increases in temperature than those with lower

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

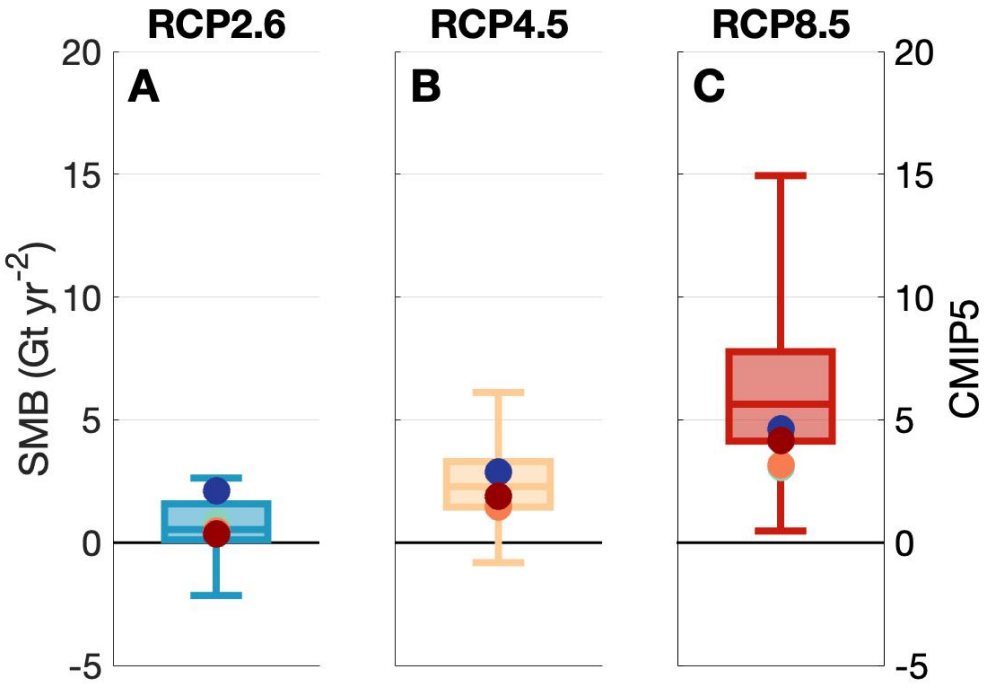

Figure 8. Box plots of the linear trend in spatially integrated AIS SMB from 2050-2100 for **A** RCP2.6 (blue), **B** RCP4.5 (yellow), and **C** RCP8.5 (red). The four larger, colored dots represent the four best scoring models: GISS E2 H (dark blue), GISS E2 R (green), MPI ESM LR (coral), and MPI ESM MR (dark red).

CO_2 concentrations. While the ranges differ from scenario to scenario, the projected sensitivity medians for RCPs 2.6, 4.5, and 8.5 are 101.7 Gt $°C^{-1}$, 111.2 Gt $°C^{-1}$, and 128.2 Gt $°C^{-1}$, respectively.

The different responses to the warming scenarios indicates that the concentration of carbon dioxide in the atmosphere has a coupled role in AIS SMB. As seen in figures 7 and 8, the greater the CO_2 concentration, the larger the AIS SMB response. From Fig. 9, we can also say that the greater the CO_2 concentration, the more sensitive to warming AIS SMB is meaning that continued warming past the end of the 21st century will have increasingly greater effects on AIS SMB.

5 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 SMB prediction.

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

The Cryosphere Discussions
Open Access
EGU

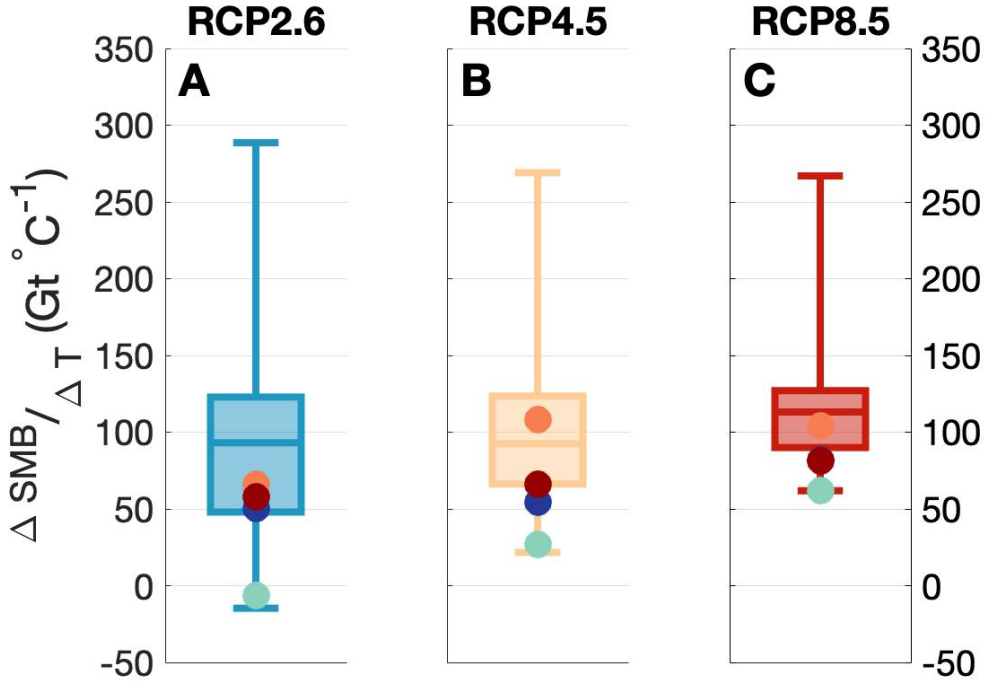

Figure 9. Box plots of all CMIP5 models' projected SMB sensitivity to temperature changes (ΔSMB/ΔT) for **A** RCP2.6, **B** RCP4.5, and **C** RCP8.5. The four larger, colored dots represent the four best scoring models: GISS E2 H (dark blue), GISS E2 R (green), MPI ESM LR (coral), and MPI ESM MR (dark red).

Our SMB mean value estimates are comparable to Agosta et al. (2019) who found a mean SMB value of roughly 2100 ± 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. 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, lower 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

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

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 four best scoring models, we refined future SMB predictions to 2246 $\pm$ 268 Gt yr^{-1} for RCP2.6, 2358 $\pm$ 331 Gt yr^{-1} for RCP4.5, and 2495 $\pm$ 335 Gt yr^{-1} for RCP8.5. Over the 21st century this translates to 1.82 cm, 3.49 cm, and 6.57 cm of GMSL rise buffering in RCPs 2.6, 4.5, and 8.5, respectively, for all of CMIP5. The best models, which show lower AIS-integrated SMB over the 21st century compared to all of CMIP5, project 1.30 cm, 3.22 cm, and 5.05 cm of GMSL rise buffering in RCPs 2.6, 4.5, and 8.5, respectively. Additionally, model trends were refined to 0.9 $\pm$ 0.1 Gt yr^{-2} for RCP2.6, 1.9 $\pm$ 1.0 Gt yr^{-2} for RCP4.5, and 3.8 $\pm$ 1.8 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 °C^{-1}, 57 $\pm$ 33 Gt °C^{-1}, and 78 $\pm$ 15 Gt °C^{-1} for RCPs 2.6, 4.5, and 8.5, respectively, for the best scoring models. Combined, these data tell us that for stronger emission scenarios, the AIS SMB response will be stronger in both magnitude and trend.

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.

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.

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

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).

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 2019

References

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.

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.

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.

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, 2012a.

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.

Lenaerts, J. T. M., van den Broeke, M. R., Déry, S. J., van Meijgaard, E., van de Berg, W. J., Palm, S. P., and Sanz Rodrigo, J.: Modeling drifting snow in Antarctica with a regional climate model: 1. Methods and model evaluation, Journal of Geophysical Research: Atmospheres, 117, n/a–n/a, https://doi.org/10.1029/2011jd016145, 2012b.

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.

https://doi.org/10.5194/tc-2019-240
Preprint. Discussion started: 17 December 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.

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.

Shepherd, A., Ivins, E. R., A, G., Barletta, V. R., Bentley, M. J., Bettadpur, S., Briggs, K. H., Bromwich, D. H., Forsberg, R., Galin, N., Horwath, M., Jacobs, S., Joughin, I., King, M. A., Lenaerts, J. T. M., Li, J., Ligtenberg, S. R. M., Luckman, A., Luthcke, S. B., McMillan, M., Meister, R., Milne, G., Mouginot, J., Muir, A., Nicolas, J. P., Paden, J., Payne, A. J., Pritchard, H., Rignot, E., Rott, H., Sørensen, L. S., Scambos, T. A., Scheuchl, B., Schrama, E. J. O., Smith, B., Sundal, A. V., van Angelen, J. H., van de Berg, W. J., van den Broeke, M. R., Vaughan, D. G., Velicogna, I., Wahr, J., Whitehouse, P. L., Wingham, D. J., Yi, D., Young, D., and Zwally, H. J.: A Reconciled Estimate of Ice-Sheet Mass Balance, Science, 338, 1183 LP – 1189, https://doi.org/10.1126/science.1228102, http://science.sciencemag.org/content/338/6111/1183.abstract, 2012.

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.