Response to editor and reviewers:

Dear editor and reviewers, we would like to thank you for your comments on our manuscript. To facilitate readability, our responses to reviewers are displayed in blue and modifications in the manuscript are highlighted in red. These suggested changes, together with additional minor corrections, are also displayed in red in the attached revised manuscript.

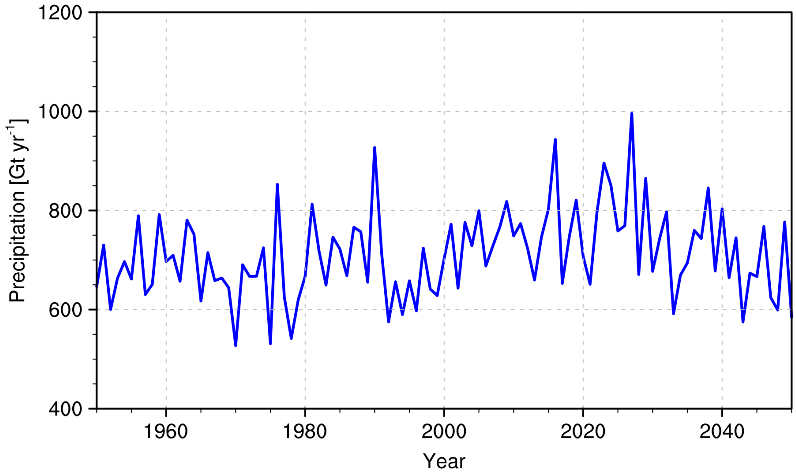
Reviewer #1

In this paper, the authors introduce present-day historical (1950-2014) global model simulation data generated by the Community Earth System Model version 2 (CESM2), which can be utilized to force polar regional climate models (RCMs) like RACMO2 used in this study. If "stand-alone" CESM2 can provide realistic climate forcing data for polar RCMs, it means that such RCMs are allowed to conduct a seamless model calculation from past to the present and future without any bias corrections. This kind of seamless simulation by a polar RCM is a state-of-the-art challenge, so that it can provide more realistic information related to possible future changes in the physical conditions of polar ice sheets as well as terrestrial climate system (e.g., sea level rise). Here, the authors perform dynamical downscaling of the CESM2 data using RACMO2 in the Greenland ice sheet (GrIS) and try to prove the effectiveness of CESM2 through validating GrIS SMB simulated by RACMO2 (equipped with the statistical downscaling postprocessing). This reviewer thinks that this considerable challenge is deserved to be published in the journal The Cryosphere as a brief communication if it is addressed in an appropriate manner. Overall, this paper is well written and structured; however, this reviewer suggests the following points to be considered before the publication. Please note that page and line numbers are denoted by "P" and "L", respectively.

Specific comments (major)

1) P. 1, L. 9-10 (and Sect. 2.1): According to the paper by Van Kampenhout et al. (2019b), which I read before reviewing this manuscript, the CESM2 simulation by Van Kampenhout et al. (2019b) was conducted following the so-called AMIP (Atmospheric Model Intercomparison Project)-run procedure. Did the authors use the same procedure/data as those presented by Van Kampenhout et al. (2019b)? If YES, the authors cannot argue "This means that, for the first time, an Earth System Model (CESM2), without assimilating observations, can be used to reconstruct historical GrIS SMB and the mass loss acceleration that started in the 1990s." in my opinion. It is because Van Kampenhout et al. (2019b) prescribe ocean and sea ice data at monthly intervals in their CESM2 simulation following the AMIP protocol, which is a kind of observation data assimilation (I mean observed ocean physical conditions can drive changes in atmospheric conditions in the model, although the atmosphere-ocean interaction would not be so strong in the model). If NO, there is no doubt that this study is amazing, and I would like to congratulate for the achievement. Anyway, please clarify this point in Sect. 2.1.
No, in this work we do not use an AMIP configuration as in Van Kampenhout et al. (2019). Instead, we enable full atmosphere-ocean coupling in CESM2. This means that sea ice and sea surface temperature evolve freely. Only land ice is kept fixed. This is now clarified in the revised manuscript in P2 L27-28: "Here, we use a full atmosphere-ocean coupling in CESM2, i.e. including sea ice dynamics and sea surface temperature evolution while excluding land ice dynamics (e.g. calving)."

2) P. 6, L. 32; P. 7, L. 2: Why does CESM2-forced RACMO2.3p2 show the significant positive trend of total precipitation since 1990, which is not shown in the ERA-forced run? Please discuss. The apparent 1991-2014 positive trend in precipitation is caused by internal decadal variability in the CESM2 climate forcing. In fact, the trend becomes insignificant in the longer term (see additional figure hereunder). This is now clarified in the revised manuscript in P7 L19-20: "In addition, the positive precipitation trend disappears when extending time series using a CESM2-based SSP8.5 scenario (not shown), demonstrating that the latter trend originates from internal decadal variability."



Specific comments (minor)

P. 2, L. 24: If possible, please indicate/mention the GrIS ice discharge simulated by the ice sheet model CISM2.1 incorporated in CESM2, which might be of interest to readers of The Cryosphere. While CESM2 does include CISM2.1, land ice is held constant in the historical simulation. This is now clarified in P2 L27-28: "Here, we use a full atmosphere-ocean coupling in CESM2, i.e. including sea ice dynamics and sea surface temperature evolution while excluding land ice dynamics (e.g. calving)."

P. 2, L. 25-26: Please indicate data sources for "atmospheric greenhouse gas emissions (CO2 and CH4), aerosol concentrations, and land cover use". Our CESM2 simulation uses time series of atmospheric greenhouse gas and aerosol emissions and land use field following the CMIP6 standards discussed in Eyring et al. (2016). We included the reference in the revised manuscript.

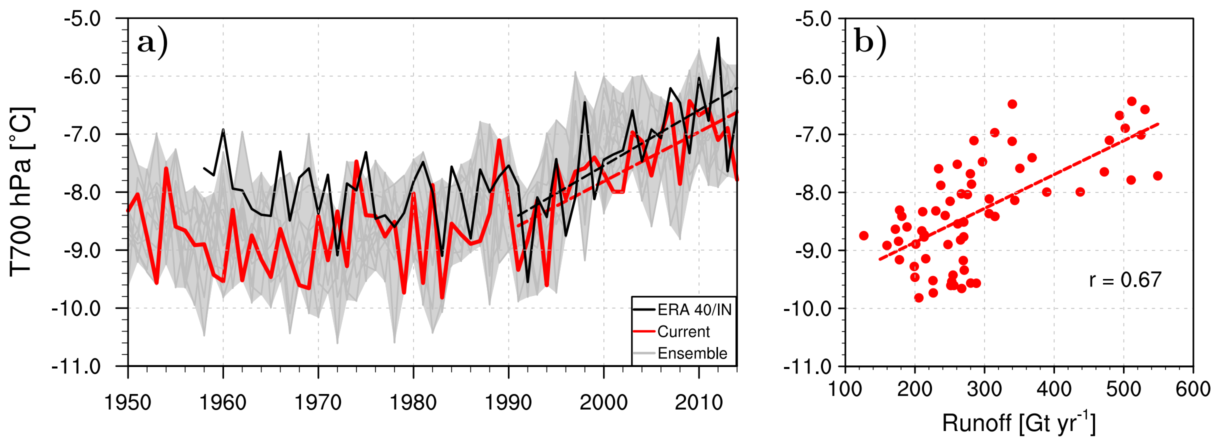
P. 3, L. 4: What kind of impurities do the authors consider in RACMO2 applied in the GrIS? And, how do the authors give concentrations of the impurities in the model? Impurities consist of soot concentration that is set to 0.1 ppmv as a constant in time and space on the RACMO2 grid at 11 km. This is clarified in P3 L8-10: "In line with in situ measurements (Doherty et al., 2010), impurity concentration (soot) in RACMO2 is prescribed as a constant in time and space at 0.1 ppmv (Noël et al., 2018)."

P. 3, L. 15: Please indicate data source of sea surface temperature and sea ice extent used here (maybe from the parent CESM2 simulation results?). Indeed we prescribe the sea ice extent/cover and sea surface temperature from the CESM2 simulation. See also our response to specific comment #1. This is now clarified in P3 L19: "Sea surface temperature and sea ice extent/cover are also prescribed from the CESM2 forcing every 6 hours."

P. 3, L. 19-21: This sentence is a bit difficult to understand to me. Does it mean the 5 % lowest bare ice albedo from MCD43A3 is 0.30? This means that the 5% lowest bare ice albedo in MODIS (i.e. that can locally be lower than 0.30 for bare ice or exceed 0.55 for firn) are clipped between 0.30 and 0.55 in RACMO2. This is now reformulated in P3 L25-26 as: "[…] 5% lowest surface albedo records for the period 2000-2015, clipped between 0.30 for bare ice and 0.55 for bright […]".

P. 5, L. 30; P. 6, L. 1: I think the results from RACMO2.1 forced by HadGEM2 is not necessary in this paper, because it can confuse readers who do not know much about RACMO2. If the authors think this part is really important for this paper, they should at least indicate key differences between RACMO2.1 and RACMO2.3p2 briefly in Sect. 2.2. Also, brief introduction of HadGEM2 would be needed as well. P. 6, L. 15-16: Same as the above comment. We deem that the comparison is valuable and shows how ESMs climate forcing has improved in time. To keep the manuscript concise and because the differences between the various RACMO2 model versions and associated forcing have been previously discussed in Van Angelen et al. (2013a,b) and Noël et al. (2015; 2018), we prefer to directly refer the reader to those publications. We included the following sentence in P6 L11-13: "For additional information about the HadGEM2-forced RACMO2.1 simulation and settings, we refer the reader to Van Angelen et al. (2013a); key differences between RACMO2.1, RACMO2.3p1 and p2 are discussed in Noël et al. (2015; 2018)."

P. 6, L. 24: Please indicate quantitatively how realistic the simulated T700 is.
Good suggestion. We now include T700 time series from ERA-40 (1958-1978) and ERA-Interim (1979-2014) averaged over the region 60-80ºN 20-80ºW (black line in revised Fig. 4a hereunder). Over the period 1958-2014, T700 from "our" CESM2 member (red line) is 0.6ºC colder than the reanalysis. For the ERA-Interim period only (1979-2014), the cold bias drops to 0.4ºC. The recent (1991-2014) warming trend (dashed black line) is well reproduced by the CESM2 forcing (dashed red line). This is now clarified in the revised manuscript in P7 L6-10: "Compared to T_{700} derived from ERA-40 (1958-1978) and ERA-Interim (1979-2014; black line in Fig. 4a), the current CESM2 simulation shows a cold bias of 0.6ºC over 1958-2014. For the ERA-Interim period, the bias decreases to 0.4ºC. All CESM2 members show a similar warming trend after 1991, in line with the reanalysis data (dashed black line), highlighting the ability of CESM2 to represent the recent climate of Greenland. As in Fettweis et al. […]". The figure caption has been modified accordingly.



P. 6, L. 28-29: This reviewer agrees with the authors' point that the attempt mentioned here is very interesting as long as the CESM2 data used in this study is not from the AMIP-type simulation. Please also see my first major comment. See our response to specific comment #1.

Figure 3: Can the authors briefly comment on why CESM2-forced RACMO2.3p2 could not simulate the 2012 extreme melt, which is simulated successfully by the ERA-forced run? I think this point is related to "physical drivers of the warming" (P. 6, L. 28), and any comments/suggestions by the authors will be informative for readers. Note that CESM2 does not assimilate observational data, in contrast to the reanalysis. The only forcing prescribed in CESM2 is greenhouse gas, aerosol emissions and land use cover. As a result, only the climate can be compared (e.g. the recent warming), not the weather (e.g. the 2012 melt event). In other words, CESM2-forced RACMO2 produces the right variability as e.g. expressed by extreme melt years (e.g. 2005 and 2011) that are realistic in magnitude but not necessarily in timing. This is clarified in P6 L32-33 and P7 L1: "It is important to note that, compared to forcing by reanalyses that assimilate observations, the CESM2-forced simulation produces extreme melt years (e.g. 2005 and 2011; Fig. 3b) that are realistic in magnitude but not necessarily in timing (e.g. the observed 2012 melt peak; Fig. 3a)."

Reviewer #2

The authors present the results of one dynamically downscaled Earth System Model (ESM) simulation over the Greenland Ice Sheet (GrIS) and present the resulting historical surface mass balance (SMB) output from their regional climate model RACMO. After dynamical downscaling of the ESM input, the SMB is furthermore statistically downscaled to a nominal horizontal resolution of 1km. In general, the authors are doing a very good job in keeping their sentence and paragraph structure easy to follow and all their figures are well presented. Therefore, the manuscript is good to read.

Scientific assessment Overall, it's hard to make a case for how the study in its present form will benefit the wider cryospheric and climate community. The point of the authors here is to create a scientific foundation for additional papers that they want to write on the future contribution of the GrIS to sea level rise via (surface) mass loss. Overall, 21st century simulations of the GrIS climate and SMB would be very beneficial for the community, however, the presented analysis currently lacks the needed depth to be considered a valuable contribution to the field. Therefore, I would encourage the authors to consider the following points.

1) The authors present only one RCM simulation forced with one GCM/ESM run to create a foundation for a future paper on 21st century GrIS climate projections. However, in its current form, the paper lacks a consideration of the inter-model spread between all of the different GCMs in the CMIP5/6 model domain a consideration of how the authors made their specific selection for the one run they choose out of their CESM2 ensemble. Fettweis et al (2013) for example analyse all the CMIP5 models over the current climate, selectively find the most suitable boundary forcings and create a downscaled RCM ensemble for multiple emission scenarios and models. This point is unfortunately omitted in this study. This study assesses the ability of CESM2 (CMIP6 version) to represent the climate and SMB of the GrIS after applying dynamical (RACMO2) and statistical downscaling. The reason for choosing CESM2 is that our institute is actively involved in the improvement of the model for studies over Greenland and Antarctica, in collaboration with the National Centre for Atmospheric Research (NCAR, Boulder, USA). We have now made this motivation specific in the introduction in P2 L8-10: "The reason for selecting CESM2 as the climate forcing for RACMO2 stems from the active involvement of the Institute for Marine and Atmospheric research Utrecht (IMAU) in the development and improvement of the model for studies over both the Greenland and Antarctic ice sheets."

Of course, running multiple members of the CESM2 historical ensemble would be of added value, but doing so in a transient fashion at this high resolution is computationally prohibitive. We have now made this clear in the text in P3 L26-29: "The current study uses the climate forcing of one out of the twelve members of the CESM2 historical ensemble. Forcing RACMO2 with other CESM2 members would have been ideal, but doing so in a transient fashion and at high spatial and temporal resolution is computationally prohibitive. Instead, we select one member that offers the 6-hourly climate forcing required to drive RACMO2 while being representative of other CESM2 members (see Section 4.3 and Fig. 4a)."

This ensemble member was selected because it had 6-hourly forcing available and is representative of other members; Fig. 4 shows that there is no reason to believe the results would be different if another member had been chosen. Based on these considerations, we judge that our conclusions on the quality of the CESM2 climate forcing are robust.

2) The authors focus their analysis only on the GrIS surface mass balance If this study should become a standalone piece of work without the promised future projections, then the authors should be highly encouraged to consider at least a subset of other parameters to validate their single-simulation analysis to exclude the likelihood of compensating biases leading to a "correct" SMB due to "false" physical reasons - (a) Surface energy budget vs. observations (b) Albedo vs. observations (c) Temperature and/or cloud properties vs. observations. We decided to limit the evaluation to SMB measurements, as the ability of CESM2 to represent key surface processes (including the near surface climate and the surface energy budget, SEB) has been addressed in other recent publications that emerged from the CESM2 development phase, e.g. Van Kampenhout et al. (2019) and Sellevold et al. (2019). In addition, direct comparison to daily in situ measurements (e.g. PROMICE, GC-NET) of (a) SEB components, (b) snow albedo, (c) near-surface temperature and cloud properties is not appropriate since ESMs, as opposed to reanalysis, do not assimilate observations and hence cannot reproduce the actual weather and exact timing of extremes (as in e.g. 2010 and 2012). See also our response to Reviewer #1 on Figure 3. We therefore deem the good agreement with in situ SMB measurements in different regions of the GrIS, characterized by very different climate conditions, to be a solid model evaluation, especially in view of the excellent agreement with temporal mass loss from GRACE.

3) If the reader assesses the novelty based on what the authors highlight "…for the first time an ESM (CESM2) can be used to reconstruct historical SMB…" then the science of the paper would need to be judged either on the claim (a) that is "the first time" or (b) that the "historical SMB" is more accurate than from other model setups. We have chosen option (a), as to the authors' knowledge no ESM-forced RCM simulation has ever accurately simulated the SMB before the 1990s and reproduced the post-1991 mass loss in close agreement with GRACE. We point out that Reviewer #1 agrees with this: *"If NO, there is no doubt that this study is amazing, and I would like to congratulate for the achievement."*

4) However, (a) e.g. Fettweis et al. (2013) as a benchmark already show that GCMs/ESMs can be used to force RCMs over the historical period and roughly get the magnitude of the SMB components right. (b) The most accurate "historical SMB" does not come from this model setup, but rather from regional climate models that downscale observation-based reanalysis data (e.g. RCM with ERA-I or ERA-5). The presented results (Figure 3) unsurprisingly show that CESM2-RACMO does not capture the interannual SMB variability and extremes (e.g. melt in 2012) which is expected with GCM boundary forcings. However, it means that the accuracy of historical SMB representation is also not an advancement of the scientific knowledge. The fact that no additional bias

correction in the forcing field is required to obtain accurate SMB is novel. We also disagree with the statement that "the accuracy of historical SMB representation is also not an advancement of the scientific knowledge". The reduced uncertainty in historical SMB reconstruction from ESM forcing as shown here is the **only way** to assess the reliability of future climate projections.

Recommendations The reviewer would like to encourage the authors to either add significant extra analysis to their current model and study setup to create a solid foundation for their promised future attribution studies, or potentially add the presented analysis to their upcoming future projections altogether. The authors could potentially consider some of the following points/questions when considering the next steps for their analysis post-review.

1) Given the limited amount of future GrIS mass loss studies with RCMs and GCM forcing, the scientific interest of the presented approach lies in the actual future projections, not necessarily on the historical SMB reconstructions due to obvious limitations when using GCM/ESM boundary conditions.
Please see our previous responses to scientific assessment #1, 3 and 4.

2) How representative is this one CESM2 run compared to the spread in CMIP5/6 simulations? Other recent studies have found great uncertainties in future GrIS projections using RCMs to downscale GCMs/ESMs which is/are not really discussed yet in the manuscript. What if the authors would force RACMO with other GCMs? How well does the current setup represent the surface energy budget, temperature, albedo, cloud properties?
Please see our previous responses to scientific assessment #1 and 2. In addition, assessing uncertainties in future projections is beyond the scope of this study that focuses on the ability of the CESM2 climate forcing to represent the present-day SMB of the GrIS.

3) If forcing RACMO with other GCMs is technically not feasible, then one approach would be to force RACMO with additional ensemble members presented in Figure 4. The robustness of the SMB and potential underlying compensating errors can hardly be assessed by only one simulation.
Please see our previous response to scientific assessment #1.

Minor comments
P1.L9: "without assimilating observations" is this correct? The methods of the paper claim that RACMO uses satellite albedo to constrain the surface albedo. Please clarify. Good point, of course we meant that **CESM2** is not constrained by observations. This is now clarified in P1 L9-11 as follows: "This means that, for the first time, climate forcing from an Earth System Model (CESM2), that assimilates no observations, can be used without additional corrections to reconstruct historical GrIS SMB […]".

P3.L19: "bare ice albedo is prescribed from … MODIS.." – please see first minor comment and clarify. P3.L28 Also in the statistical downscaling technique the authors use observed MODIS albedo. Please see the first comment on how this fits with the claim that this study doesn't use assimilated observations.
Please see the answer above in P1 L9.

P3.L32-33: Does it only change the runoff and SMB or also improve the statistical comparison?
Statistical downscaling aims at resolving narrow marginal glaciers, ablation zones, and associated large SMB gradients not resolved by the 11 km grid, as well as correcting for the bare ice albedo bias in RACMO2. As a result, statistical downscaling primarily increases marginal runoff, which improves the SMB agreement with observations. The method is presented in detail in Noël et al. (2016).

P4.L24: "due to the high quality of the CESM2 climate" but also e.g. P1.L5 "good comparison" and P5.L6 "shows excellent agreement" and at other points in the manuscript - these are quite colloquial expressions with little scientific meaning. What does a "high quality" climate in a GCM mean? The manuscript doesn't even currently evaluate the CESM2 climate for example. Good point. In P1 L5 we deem that evaluation statistics should not be listed in the abstract, the "good agreement" and associated statistics are elaborated in more detail in Sections 3 and 4. In P5 L6 we feel that mass loss derived from combined observed ice discharge and modelled SMB of 3,299 Gt yr^{-1} is indeed in "excellent agreement" with GRACE estimates of 3,290 Gt yr^{-1}. Concerning the "high quality" statement, we decided to remove the sentence in P4 L23-25.

P4.L25ff: But what about other parameters such as the surface energy budget, temperature and clouds? How does it compare to recent circulation and cloud anomalies over Greenland which have been shown to be important for future projections? Upper atmospheric temperature (T700) in the CESM2 forcing is now evaluated using ECMWF reanalyses in Fig. 4a. See also our response to scientific assessment #2. Addressing circulation and cloud anomalies is beyond the scope of this study: this work assesses the ability of the CESM2 climate forcing to reconstruct the present-day SMB of the GrIS.

P5.L6-8: The acceleration (i.e. dSMB/dt) is likely not discussed here but rather a "total mass loss".
Thank you for pointing this out, we meant "mass loss" rather than "mass loss acceleration". This is now corrected in P5 L14 as follows: "[…] realistically capture the recent Greenland mass loss (2003-2014) (Bamber et al., 2018)."

P5.L30-32ff: ad HadGEM; "did not accurately reproduce SMB". a) Throughout this study the reader is often left in the dark as to "Why?" certain numbers or results are mentioned, and why certain processes behave the way they do. At the moment, the paper is an ensemble of nice figures and easy-to-follow text, but the study and the reader would highly benefit if the authors would more often dig into the question of "Why?" some processes and numbers are reported here and apparently deemed important for the reader. b) This would also be a good point to address the matter why HadGEM and CESM2 produce such different SMB/ME/RU results (+-50%)? Is it due to differences in the lateral forcings/ the internal RACMO physics/ circulation / cloud physics? Hofer et al. (2019) for example show the large spread in GrIS SMB that can result from different GCM forcing.
To address the latter, we now include this sentence in P6 L26-28: "The reason is that, unlike CESM2 (Van Kampenhout et al., 2019a), the HadGEM2 forcing had a strong, systematic warm bias of ~1ºC (Van Angelen et al., 2013a), resulting in overestimated meltwater runoff and thus underestimated SMB (Fig. 2d)." Regarding the first comment, the topic of this paper is how development of CESM2 in particular (and therewith ESMs in general) has led to much improved representation of (downscaled) GrIS SMB. Back in its time, HadGEM2 was too warm over Greenland and required corrections to obtain an acceptable GrIS SMB. The main message of this paper is that this kind of corrections is now no longer required and even recent (mass loss) trends are captured correctly. We feel that for a short communication, this presents sufficient advance of the state-of-the-art to warrant publication. At the same time, we deem that it is beyond the scope of this paper to analyse problems in a -now obsolete- GCM.

P7.L8-9 "can reliably reproduce … variability of historical SMB" – When looking at Figure 3 the GCM forced SMB reconstruction clearly lacks the ability to reproduce the interannual SMB variability and extremes shown in Figure 3A when RACMO is forced by reanalysis. Just as an example, the extreme melt summer of 2012 accurately captured in Figure 3A is not present in Figure 3B, therefore the reader considers this to be a doubtful assumption. See our response to Reviewer #1 on Figure 3.

P7.L3-4: unclear phrasing "is for 60%" Thank you, this has been reformulated as follows in P7 L20: "[…] Fig. 3c shows that 60% of the recent mass loss acceleration in CESM2-forced RACMO2.3p2 is caused by decreased SMB […]".

P7.L7-10: What are the uncertainties coming from the lack of a multi model forcing (e.g. Fettweis et al. (2013).
See our response to scientific assessment #1.

Figure 1: How does it compare during melt season? How does the SEB compare to the observational networks of PROMICE, DMI and/or GCNET? See our response to scientific assessment #2.

Figure 2: Please clarify the choice of HadGEM and not other GCMs? If it is feasible to force RACMO with other GCMs then please consider analysing the intermodel spread of the GrIS climate when RACMO is forced by other GCMs. See our response to scientific assessment #1.

Brief communication: CESM2 climate forcing (1950-2014) yields realistic Greenland ice sheet surface mass balance

Brice Noël[1], Leonardus van Kampenhout[1], Willem Jan van de Berg[1], Jan T. M. Lenaerts[2], Bert Wouters[1,3], and Michiel R. van den Broeke[1]

[1]Institute for Marine and Atmospheric research (IMAU), Utrecht University, Utrecht, Netherlands
[2]Department of Atmospheric and Oceanic Sciences, University of Colorado Boulder, Boulder CO, USA
[3]Department of Geoscience and Remote Sensing, Delft University of Technology, Delft, Netherlands

Correspondence: Brice Noël (b.p.y.noel@uu.nl)

Abstract.

We present a reconstruction of historical (1950-2014) surface mass balance (SMB) of the Greenland ice sheet (GrIS) using a high-resolution regional climate model (RACMO2; ∼11 km) to dynamically downscale the climate of the Community Earth System Model version 2 (CESM2; ∼111 km). After further statistical downscaling to 1 km spatial resolution, evaluation using in situ SMB measurements and remotely sensed GrIS mass change shows good agreement, including the recently observed acceleration in surface mass loss (2003-2014). Comparison with an ensemble of eight previously conducted RACMO2 simulations forced by climate reanalysis demonstrates that the current product accurately reproduces the long term average and variability of individual SMB components, and captures the recent increase in meltwater runoff that accelerated GrIS mass loss. This means that, for the first time, climate forcing from an Earth System Model (CESM2), that assimilates no observations, can be used without additional corrections to reconstruct historical GrIS SMB and the mass loss acceleration that started in the 1990s. This paves the way for attribution studies of future GrIS mass loss projections and contribution to sea level rise.

1 Introduction

A common approach to project the future surface mass balance (SMB) of the Greenland ice sheet (GrIS) is to force a regional climate model (RCM), typically running at 5 to 10 km horizontal resolution, at the lateral and top boundaries with the output of an Earth System Model (ESM; ∼100 km) (Van Angelen et al., 2013a; Fettweis et al., 2013; Mottram et al., 2017). However, ESMs from the fifth phase of the Climate Model Intercomparison Project (CMIP5) do not accurately represent the contemporary large-scale climate of the Greenland region (Rae et al., 2012; Fettweis et al., 2013). The reason is that ESMs do not assimilate nor prescribe climatic observations as do global climate reanalyses (Uppala et al., 2005; Dee et al., 2011) and RCMs (Fettweis et al., 2017; Mottram et al., 2017; Niwano et al., 2018; Noël et al., 2018). For instance, ESMs fail at capturing the recent summertime Arctic atmospheric circulation change (Hanna et al., 2018), making projections of GrIS mass loss and contribution to sea level rise highly uncertain (Delhasse et al., 2018). Consequently, climate forcing from CMIP5 ESMs still requires dedicated bias correction before being used to force RCMs over the GrIS (Rae et al., 2012; Fettweis et al., 2013; Van Angelen et al., 2013a). An alternative approach is to directly use outputs of ESMs to estimate GrIS SMB; however, most ESMs do not have (sophisticated) snow models that consider meltwater retention in firn, while their coarse spatial resolution does not accurately resolve the large SMB gradients at the GrIS margins (Lenaerts et al., 2019).

Here, we use the historical climate (1950-2014) of a CMIP6 model, the Community Earth System Model version 2.1 (CESM2; $\sim$111 km), to force the lateral and top boundaries of the Regional Atmospheric Climate Model version 2.3p2 (RACMO2; $\sim$11 km). The reason for selecting CESM2 as the climate forcing for RACMO2 stems from the active involvement of the Institute for Marine and Atmospheric research Utrecht (IMAU) in the development and improvement of the model for studies over both the Greenland and Antarctic ice sheets. To obtain a meaningful comparison with in situ observations, the resulting SMB field is then statistically downscaled to 1 km over the GrIS and peripheral glaciers and ice caps (Fig. 1a) (Noël et al., 2018). We show that, without additional corrections, CESM2 climate forcing yields a realistic reconstruction of historical GrIS SMB (1950-2014), including the recent acceleration in mass loss. This is unexpected for an ESM which is exclusively driven by prescribing greenhouse gas (CO_2 and CH_4) and aerosol emissions, and may herald more accurate projections of GrIS contribution to future sea level rise. Section 2 describes CESM2 and RACMO2, including model initialisation, forcing set-up, as well as observational and model data sets used for evaluation. Section 3 evaluates the CESM2-forced RACMO2 product using in situ and remotely sensed measurements. Model comparison to previous RACMO2 simulations is discussed in Section 4, as well as representation of recent trends in SMB components and mass loss. Conclusions are drawn in Section 5.

2 Methods

2.1 Community Earth System Model: CESM2

CESM2.1, hereafter referred to as CESM2, is an ESM that simulates mutual interactions between atmosphere-ocean-land systems on the global scale. The model incorporates the Community Atmosphere Model version 6 (CAM6) (Gettelman et al., 2019), resolving global atmospheric dynamics and physics; the Parallel Ocean Program model version 2.1 (POP2.1) (Smith et al., 2010) and the Los Alamos National Laboratory Sea Ice Model version 5.1 (CICE5.1) (Bailey et al., 2018), modelling global oceanic circulation and sea-ice evolution. These are coupled with the Community Land Model version 5 (CLM5) (Lawrence, 2019) and the Community Ice Sheet Model version 2.1 (CISM2.1) (Lipscomb et al., 2019) simulating land-atmosphere interactions and ice dynamics. Here, we use a full atmosphere-ocean coupling in CESM2, i.e. including sea ice dynamics and sea surface temperature evolution while excluding land ice dynamics (e.g. calving). The model is run at 1° spatial resolution ($\sim$111 km) and only prescribes atmospheric greenhouse gas (CO_2 and CH_4) and aerosol emissions as well as land cover use (Eyring et al., 2016). CESM2 has been extensively tested and adapted to realistically reproduce the contemporary climate and SMB of the GrIS (Van Kampenhout et al., 2017, 2019b), detailed model description and latest updates are provided in Van Kampenhout et al. (2019a).

2.2 Regional Atmospheric Climate Model: RACMO2

RACMO2 is an RCM that is specifically adapted to simulate the climate of polar ice sheets (Noël et al., 2018; Van Wessem et al., 2018). The model incorporates the dynamical core of the High Resolution Limited Area Model (HIRLAM) (Undèn et al., 2002) and the physics package cycle CY33r1 of the European Centre for Medium-range Weather Forecasts Integrated Forecast System (ECMWF-IFS, 2008). It includes a multi-layer snow module that simulates melt, liquid water percolation and retention, refreezing and runoff (Ettema et al., 2010), and accounts for dry snow densification following Ligtenberg et al. (2011). RACMO2 implements an albedo scheme that calculates snow albedo based on prognostic snow grain size, cloud optical thickness, solar zenith angle and impurity concentration in snow (Kuipers Munneke et al., 2011). In line with in situ measurements (Doherty et al., 2010), impurity concentration (soot) in RACMO2 is prescribed as a constant in time and space at 0.1 ppmv (Noël et al., 2018). The model simulates drifting snow erosion and sublimation following Lenaerts et al. (2012). The latest model version RACMO2.3p2 accurately simulates the contemporary climate and SMB of the GrIS when forced by ERA-40 (1958-1978) and ERA-Interim (1979-present) climate reanalyses (Uppala et al., 2005; Dee et al., 2011) and statistically downscaled to 1 km (see Section 2.4). For detailed model description and latest updates, we refer to Noël et al. (2018, 2019).

2.3 Model initialisation and set-up

Here, we conduct a CMIP6-style historical simulation (1950-2014) using RACMO2.3p2 at 11 km horizontal resolution (Noël et al., 2018) to dynamically downscale the output of CESM2 prescribed in a 24 grid-cell wide relaxation zone at the model lateral boundaries. Forcing consists of atmospheric temperature, pressure, specific humidity, wind speed and direction being prescribed on a 6-hourly basis at the 40 model atmospheric levels. Upper atmosphere relaxation is implemented (Van de Berg and Medley, 2016). Sea surface temperature and sea ice extent/cover are also prescribed from the CESM2 forcing every 6 hours. RACMO2.3p2 has typically 40 to 60 active snow layers that are initialised in January 1950 using temperature and density profiles derived from the offline IMAU Firn Densification Model (IMAU-FDM) (Ligtenberg et al., 2018). Glacier outlines and surface topography are prescribed from a down-sampled version of the 90 m Greenland Ice Mapping Project (GIMP) Digital Elevation Model (DEM) (Howat et al., 2014). Bare ice albedo is prescribed from the 500 m MODerate-resolution Imaging Spectroradiometer (MODIS) 16-day Albedo product (MCD43A3), as the 5% lowest surface albedo records for the period 2000-2015, clipped between 0.30 for bare ice and 0.55 for bright ice covered by perennial firn in the accumulation zone. The current study uses the climate forcing of one out of the twelve members of the CESM2 historical ensemble. Forcing RACMO2 with other CESM2 members would have been ideal, but doing so in a transient fashion and at high spatial and temporal resolution is computationally prohibitive. Instead, we select one member that offers the 6-hourly climate forcing required to drive RACMO2 while being representative of other CESM2 members (see Section 4.3 and Fig. 4a).

2.4 Statistical downscaling

Following Noël et al. (2016), the historical simulation at 11 km, hereafter referred to as CESM2-forced RACMO2.3p2, is further statistically downscaled to a 1 km ice mask and topography derived from the 90 m GIMP DEM (Howat et al., 2014).

In brief, the downscaling procedure corrects individual SMB components (except for precipitation), i.e. primarily meltwater runoff, for elevation and ice albedo biases on the relatively coarse model grid at 11 km resolution. These corrections reconstruct individual SMB components on the 1 km GrIS topography using daily-specific gradients estimated at 11 km, and minimise the remaining runoff underestimation using a down-sampled 1 km MODIS 16-day ice albedo product averaged for 2000-2015. Precipitation, including snowfall and rainfall, is bi-linearly interpolated from the 11 km onto the 1 km grid without additional corrections (Noël et al., 2018). Statistical downscaling proves essential to resolve narrow ablation zones, outlet glaciers and ice caps at the GrIS margins that significantly contribute to contemporary mass loss of Greenland land ice (Noël et al., 2017, 2019). For instance, applying statistical downscaling increases GrIS-wide runoff by 55 Gt yr^{-1} (+23%) on average for the period 1950-2014, resulting in a SMB decrease of 56 Gt yr^{-1} (-13%).

2.5 Evaluation data sets

For evaluation, we use a compilation of in situ SMB measurements derived from 182 stakes, snow pits (Bales et al., 2009) and airborne radar campaign (Overly et al., 2016) in the GrIS accumulation area (182 records; white dots in Fig. 1a), and collected at 213 sites (Machguth et al., 2016) in the ablation zone (1073 records; yellow dots in Fig. 1a). In addition, combined modelled SMB and glacial discharge estimates (Mouginot et al., 2019) are compared to mass changes from GRACE over the period 2003-2014 (Wouters et al., 2013). The CESM2-forced RACMO2.3p2 historical simulation is also compared to SMB and individual components from an ensemble of eight previous RACMO2 simulations (Van Angelen et al., 2013a, b; Noël et al., 2015, 2016, 2018, 2019), using different climate forcing (ERA-reanalysis or the ESM HadGEM2) at various spatial resolutions (1, 5.5 and 11 km). These simulations are listed and further compared in Tables A1 and A2 for the overlapping model period 1960-2012.

3 Surface mass balance evaluation and uncertainty

Figure 1a shows annual mean SMB from CESM2-forced RACMO2.3p2, statistically downscaled to 1 km. As is the case with state-of-the-art reanalysis-forced simulations (Mottram et al., 2017; Fettweis et al., 2017; Niwano et al., 2018; Noël et al., 2018, 2019), it accurately captures the extensive inland accumulation area, and narrow ablation zones, outlet glaciers and ice caps fringing the GrIS margins (Fig. 1a). The model shows very good agreement with multi-year averaged SMB observations in the accumulation zone (R^2 = 0.89; Fig. 1b), with a small bias and RMSE of -20.5 mm w.e. and 63.3 mm w.e. Interestingly, these statistics are on par with the most recent RACMO2.3p2 run forced by ERA-reanalysis and statistically downscaled to 1 km (Noël et al., 2018), hereafter referred to as ERA-forced RACMO2.3p2.

In the ablation zone, CESM2-forced RACMO2.3p2 agrees reasonably well with ablation measurements: R^2 = 0.61 vs. 0.72 (Noël et al., 2018) (Fig. 1c). The model shows larger bias and RMSE relative to ERA-forced RACMO2.3p2 (+0.06 m w.e. and +0.18 m w.e.). As CESM2 does not assimilate nor prescribe climatic observations, a larger bias was expected. Good agreement with observations can be partly attributed to dynamical downscaling in RACMO2, that results in realistic SMB gradients if appropriate climate forcing is prescribed (Noël et al., 2018); and to statistical downscaling, as it minimises SMB bias by enhancing runoff in marginal ablation zones (Noël et al., 2016). On the regional scale, CESM2-forced and ERA-forced RACMO2.3p2 simulations show no significant difference in SMB and components for the period 1958-2014 (not shown), i.e. mean difference (CESM2-forced minus ERA-forced) lower than one standard deviation of the 1958-2014 period.

We follow Noël et al. (2017) to estimate the SMB uncertainty (σ). Mean accumulation (20.5 mm w.e.; Fig. 1b) and ablation biases (180.0 mm w.e.; Fig. 1c) are accumulated over the long-term (1958-2014) accumulation and ablation zone of the GrIS, with an area of $\sim$1,521,400 km^2 and $\sim$179,400 km^2 respectively:

$$\sigma = \sqrt{(bias^{\,\text{abl.}} \times area^{\,\text{abl.}})^2 + (bias^{\,\text{acc.}} \times area^{\,\text{acc.}})^2} \tag{1}$$

This yields σ = 48 Gt yr^{-1}. A similar value (43 Gt yr^{-1}) is obtained for the downscaled ERA-forced RACMO2.3p2 product. Integrated over the ice sheet, CESM2-forced and ERA-forced simulations agree very well, with an average cumulative SMB of 365 $\pm$ 48 Gt yr^{-1} and 357 $\pm$ 43 Gt yr^{-1} for the period 1958-2014. Figure 1d compares modelled and remotely sensed (GRACE) cumulative mass change for the period 2003-2014, respectively. Modelled mass change (-3,299 $\pm$ 1,240 Gt; blue box), estimated as cumulative SMB over the GrIS, peripheral ice caps and tundra region (2,970 $\pm$ 1,097 Gt; black box) minus glacial discharge (6,269 $\pm$ 143 Gt; orange box), shows excellent agreement with GRACE (-3,290 $\pm$ 1,434 Gt; red box). This highlights the ability of CESM2-forced RACMO2.3p2 to also realistically capture the recent Greenland mass loss (2003-2014) (Bamber et al., 2018).

4 Surface mass balance variability

Figures 2a and c show annual mean SMB components simulated by CESM2-forced RACMO2.3p2 at 1 km (horizontal bars) for the periods 1960-1990 and 1991-2012. Likewise, Figs. 2b and d show annual mean mass balance (MB; blue), i.e. SMB (black) minus glacial discharge (orange), simulated by CESM2-forced RACMO2.3p2 on the original model grid (11 km) and statistically downscaled to 1 km for the two periods. Ice discharge and associated uncertainties (1972-2014) are from Mouginot et al. (2019). Prior to 1972, ice discharge and uncertainties are assumed constant at the 1972 value. Error bars represent the inter-annual variability in SMB components estimated as one standard deviation around the mean. Boxes show the range of modelled SMB components derived from the ensemble of RACMO2 simulations forced by ERA-reanalysis. Annual mean SMB components and corresponding inter-annual variability for the ensemble RACMO2 simulations are listed in Tables A1 (1960-1990) and A2 (1991-2012).

4.1 Approximate mass balance: 1960-1990

In the period 1960-1990, the mass balance of the GrIS was close to zero (Van den Broeke et al., 2016) or slightly negative (Mouginot et al., 2019). Figures 2a, b and Table A1 show that downscaled CESM2-forced RACMO2.3p2 reproduces, within one standard deviation, SMB and components obtained from eight previous reanalysis-forced RACMO2 simulations at various spatial resolutions. For instance, precipitation (701 $\pm$ 98 Gt yr^{-1}) and runoff (242 $\pm$ 40 Gt yr^{-1}) compare well with ERA-forced RACMO2.3p2 (Noël et al., 2018), i.e. 712 ± 73 Gt yr^{-1} and 257 ± 53 Gt yr^{-1}, resulting in similar SMB of 428 and 423 Gt yr^{-1} (-1%) (Fig. 2a and Table A1). This highlights the ability of the CESM2 forcing to capture realistic Greenland SMB before mass loss started in the 1990s.

Figure 2b shows that SMB on the 11 km grid falls well within ERA-forced simulations at similar resolution (black box). Through statistical downscaling, SMB at 1 km decreases by 13% from 485 to 428 Gt yr^{-1}, in line with other simulations (Fig. 2b and Table A1). Combining average GrIS-integrated SMB with glacial discharge (458 Gt yr^{-1}; 1960-1990), CESM2-forced RACMO2.3p2 results in slightly negative mass balance (-31 Gt yr^{-1}; Table A1). A previous attempt using RACMO2.1 forced by the climate of the HadGEM2 ESM (Van Angelen et al., 2013a) did not accurately represent GrIS-integrated SMB components (dark green dots in Figs. 2a,b). While precipitation was generally well represented (685 ± 82 Gt yr^{-1}), runoff was overestimated by ∼50% compared to ERA-forced RACMO2.3p2 (Table A1). As a result, SMB was underestimated by ∼40%, driving an unrealistic mass loss of 189 Gt yr^{-1} over the period 1960-1990. For additional information about the HadGEM2-forced RACMO2.1 simulation and settings, we refer the reader to Van Angelen et al. (2013a); key differences between RACMO2.1, RACMO2.3p1 and p2 are discussed in Noël et al. (2015, 2018).

4.2 Mass loss: 1991-2012

In the two decades following 1990 (1991-2012), the GrIS experienced accelerated mass loss (Bamber et al., 2018; Shepherd et al., 2019), primarily driven by a decrease in SMB (-138 Gt yr^{-1} with respect to 1960-1990; Fig. 2c and Table A2) combined with an increase in glacial discharge (+26 Gt yr^{-1}). Figure 2c shows that CESM2-forced RACMO2.3p2 similarly reproduces the recent SMB decrease resulting from enhanced surface runoff (+138 Gt yr^{-1} or +57%) compared to the ERA-forced RACMO2.3p2 simulation (+100 Gt yr^{-1} or +38%). This pronounced runoff increase stems from enhanced surface melt (+163 Gt yr^{-1} or +36%) exceeding the increase in meltwater retention and refreezing in the firn (+36 Gt yr^{-1} or +14%). Precipitation does not substantially change after 1991, in line with the ensemble ERA-forced RACMO2 simulations (Table A2). We conclude that CESM2-forced RACMO2.3p2 captures the post-1990 SMB decrease that tipped the GrIS into a state of sustained mass loss (195 Gt yr^{-1} and 170 Gt yr^{-1} for downscaled CESM2-forced and ERA-forced RACMO2.3p2 for 1991-2012; Fig. 2d). In contrast, SMB components in HadGEM2-forced RACMO2.1 remain largely overestimated compared to other simulations (Table A2), particularly runoff and melt (Fig. 2c), resulting in overestimated mass loss for 1991-2012 (240 ± 80 Gt yr^{-1}; Fig. 2d). The reason is that, unlike CESM2 (Van Kampenhout et al., 2019a), the HadGEM2 forcing had a strong, systematic warm bias of ∼1^{o}C (Van Angelen et al., 2013a), resulting in overestimated meltwater runoff and thus underestimated SMB (Fig. 2d).

4.3 Time series and trends

Figures 3a and b show time series of individual GrIS-integrated SMB components for the period 1950-2014 as modelled by the latest, state-of-the-art ERA-forced RACMO2.3p2 run at 5.5 km horizontal resolution (Noël et al., 2019) and the current CESM2-forced RACMO2.3p2 simulation, both statistically downscaled to 1 km. It is important to note that, compared to forcing by reanalyses that assimilate observations, the CESM2-forced simulation produces extreme melt years (e.g. 2005 and

2011; Fig. 3b) that are realistic in magnitude but not necessarily in timing (e.g. the observed 2012 melt peak; Fig. 3a). For 1960-1990, the two products show similar and insignificant trends in total precipitation (2.4 $\pm$ 1.6 Gt yr^{-2}; p-value = 0.14) and runoff (1.1 $\pm$ 1.0 Gt yr^{-2}; p-value = 0.27) (Fig. 3b). After 1991, CESM2-forced RACMO2.3p2 reproduces the significant (p-value = 0.0001) positive runoff trend (10.4 $\pm$ 2.2 Gt yr^{-2}; Fig. 3b) similar to the ERA-forced simulation (8.8 $\pm$ 2.1 Gt yr^{-2}; Fig. 3a). The runoff trend in CESM2-forced RACMO2.3p2 is no coincidence. Figure 4a shows the atmospheric temperature at 700 hPa (T_{700}) from the current CESM2 simulation (red) and from 11 additional ensemble members (grey). Compared to T_{700} derived from ERA-40 (1958-1978) and ERA-Interim (1979-2014; black line in Fig. 4a), the current CESM2 simulation shows a cold bias of 0.6^{o}C over 1958-2014. For the ERA-Interim period, the bias decreases to 0.4^{o}C. All CESM2 members show a similar warming trend after 1991, in line with the reanalysis data (dashed black line), highlighting the ability of CESM2 to represent the recent climate of Greenland. As in Fettweis et al. (2013), we find a clear correlation (r = 0.67; Fig. 4b) between CESM2-forced RACMO2.3p2 runoff at 1 km and T_{700} from the CESM2 simulation (red). This means that the post-1990 runoff increase would have been obtained irrespective of the selected CESM2 member. Physical drivers of the warming trend in the CESM2 forcing are currently being investigated and will be discussed in a forthcoming publication. Compared to the ERA-forced run, the more pronounced runoff trend in CESM2-forced RACMO2.3p2 results from a significant (p-value = 0.016) positive trend in rainfall (1.3 $\pm$ 0.3 Gt yr^{-2} vs. 0.3 $\pm$ 0.2 Gt yr^{-2}) for a similar melt acceleration (11.9 $\pm$ 3.0 Gt yr^{-2} vs. 10.9 $\pm$ 3.0 Gt yr^{-2}). Total precipitation in the CESM2-forced RACMO2.3p2 simulation shows a significant (p-value = 0.002) positive trend (5.8 $\pm$ 1.7 Gt yr^{-2}; Fig. 3b) in contrast to a negative trend in the ERA-forced run (-1.9 $\pm$ 1.8 Gt yr^{-2}; Fig. 3a). However, the latter trend stems from decadal variability as it becomes insignificant for the period 1950-2014: 0.9 $\pm$ 0.5 Gt yr^{-2} (p-value = 0.090). In addition, the positive precipitation trend disappears when extending time series using a CESM2-based SSP8.5 scenario (not shown), demonstrating that the latter trend originates from internal decadal variability.

In line with Van den Broeke et al. (2016), Fig. 3c shows that $\sim$60% of the recent mass loss acceleration in CESM2-forced RACMO2.3p2 is caused by decreased SMB (6.6 $\pm$ 3.3 Gt yr^{-2}) resulting from enhanced meltwater runoff; the remaining $\sim$40% is ascribed to increased glacial discharge (4.7 $\pm$ 0.5 Gt yr^{-2}). As a result, Greenland mass balance decreased by an estimated rate of 11.3 $\pm$ 3.2 Gt yr^{-2} (or 9.4 $\pm$ 1.6 Gt yr^{-2} for the GrIS only) in good agreement with GRACE (9.4 $\pm$ 1.2 Gt yr^{-2} for 2003-2014; Fig. 3c). This is meaningful for two reasons: for the first time, an ESM, assimilating no observational climatic data except for atmospheric greenhouse gas and aerosol emissions, can 1) reliably reproduce the historical average and variability of SMB and its individual components; 2) accurately represent the recent Greenland mass loss acceleration. These results are essential for forthcoming attribution studies investigating post-1990 GrIS mass loss.

5 Conclusions

Historical output (1950-2014) of the Earth System Model CESM2 ($\sim$111 km) is dynamically downscaled using the regional climate model RACMO2.3p2 ($\sim$11 km) over the GrIS. The resulting SMB components are further statistically downscaled to 1 km spatial resolution to resolve the narrow ablation zones and marginal outlet glaciers of Greenland. Model evaluation using in-situ and remotely sensed measurements demonstrates the ability of CESM2-forced RACMO2.3p2 to accurately reproduce SMB as well as the rapid post-1991 melt and runoff increase. Combining modelled SMB with observed glacial discharge, our new ESM-based SMB product reflects an ice sheet in approximate mass balance before 1991, followed by a rapid mass loss acceleration resulting from enhanced meltwater runoff: two key features that, until now, exclusively showed up in reanalysis-based estimates. This means that, for the first time, an Earth System Model (CESM2), that does not assimilate climatic observations, can be used to force a regional climate model (RACMO2) to accurately reproduce historical GrIS SMB average and variability. Furthermore, our results suggest that CESM2 climate forcing can be used without bias corrections to simulate the climate and SMB of the GrIS for different warming scenario projections to quantify the GrIS contribution to future eustatic sea level rise.

Data availability. Data sets presented in this study are available from the authors upon request and without conditions.

Author contributions. BN prepared the manuscript, conducted the CESM2-forced RACMO2.3p2 simulation and analysed the data. LK and JL provided the historical CESM2 forcing. BW processed the GRACE mass anomalies time series. MB and WJB helped interpreting the results. All authors commented the manuscript.

Competing interests. The authors declare no competing interests.

Acknowledgements. BN, LvK, WJvdB and MRvdB acknowledge funding from the Netherlands Organisation for Scientific Research (NWO) and the Netherlands Earth System Science Centre (NESSC). The Community Earth System Model (CESM) project is supported primarily by the National Science Foundation (NSF). This material is based upon work supported by the National Center for Atmospheric Research (NCAR), which is a major facility sponsored by the NSF under Cooperative Agreement No. 1852977.

References

Bailey, D., Hunke, E., DuVivier, A., Lipscomb, B., Bitz, C., Holland, M., Briegleb, B., and Schramm, J.: CESM CICE5 Users Guide, Release CESM CICE5, Documentation and Software User's Manual from Los Alamos National Laboratory, pp. 1 – 41, http://www.cesm.ucar.edu/models/cesm2/sea-ice/, 2018.

Bales, R. C., Guo, Q., Shen, D., McConnell, J. R., Du, G., Burkhart, J. F., Spikes, V. B., Hanna, E., and Cappelen, J.: Annual accumulation for Greenland updated using ice core data developed during 2000-2006 and analysis of daily coastal meteorological data, Journal of Geophysical Research, 114, D06 116, https://doi.org/10.1029/2008JD011208, 2009.

Bamber, J. L., Westaway, R. M., Marzeion, B., and Wouters, B.: The land ice contribution to sea level during the satellite era, Environmental Research Letters, 13, 063 008, https://doi.org/10.1088/1748-9326/aac2f0, 2018.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hòlm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553 – 597, https://doi.org/10.1002/qj.828, 2011.

Delhasse, A., Fettweis, X., Kittel, C., Amory, C., and Agosta, C.: Brief communication: Impact of the recent atmospheric circulation change in summer on the future surface mass balance of the Greenland ice sheet, The Cryosphere, 12, 3409–3418, https://doi.org/10.5194/tc-12-3409-2018, 2018.

Doherty, S. J., Warren, S. G., Grenfell, T. C., Clarke, A. D., , and Brandt, R. E.: Light-absorbing impurities in Arctic snow, Atmos. Chem. Phys., 10, 11 647 – 11 680, https://doi.org/10.5194/acp-10-11647-2010, 2010.

ECMWF-IFS: Part IV : PHYSICAL PROCESSES (CY33R1), Technical Report, 2008.

Ettema, J., van den Broeke, M. R., van Meijgaard, E., van de Berg, W. J., Box, J. E., and Steffen, K.: Climate of the Greenland ice sheet using a high-resolution climate model - Part 1: Evaluation, The Cryosphere, 4, 511 – 527, https://doi.org/10.5194/tc-4-511-2010, 2010.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geosci. Model Dev., 9, 1937 – 1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Fettweis, X., Franco, B., Tedesco, M., van Angelen, J. H., Lenaerts, J. T. M., van den Broeke, M. R., and Gallée, H.: Estimating the Greenland ice sheet surface mass balance contribution to future sea level rise using the regional atmospheric climate model MAR, The Cryosphere, 7, 469 – 489, https://doi.org/10.5194/tc-7-469-2013, 2013.

Fettweis, X., Box, J. E., Agosta, C., Amory, C., Kittel, C., Lang, C., van As, D., Machguth, H., and Gallée, H.: Reconstructions of the 1900-2015 Greenland ice sheet surface mass balance using the regional climate MAR model, The Cryosphere, 11, 1015 – 1033, https://doi.org/10.5194/tc-11-1015-2017, 2017.

Gettelman, A., Truesdale, J. E., Bacmeister, J. T., Caldwell, P. M., Neale, R. B., Bogenschutz, P. A., and Simpson, I. R.: The Single Column Atmosphere Model Version 6 (SCAM6): Not a Scam but a Tool for Model Evaluation and Development, Journal of Advances in Modeling Earth Systems, 11, 1381 – 1401, https://doi.org/10.1029/2018MS001578, 2019.

Hanna, E., Fettweis, X., and Hall, R. J.: Brief communication: Recent changes in summer Greenland blocking captured by none of the CMIP5 models, The Cryosphere, 12, 3287 – 3292, https://doi.org/10.5194/tc-12-3287-2018, 2018.

Howat, I. M., Negrete, A., and Smith, B. E.: The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets, The Cryosphere, 8, 1509 – 1518, https://doi.org/10.5194/tc-8-1509-2014, 2014.

Kuipers Munneke, P., van den Broeke, M. R., Lenaerts, J. T. M., Flanner, M. G., Gardner, A. S., and van de Berg, W. J.: A new albedo parameterization for use in climate models over the Antarctic ice sheet, Journal of Geophysical Research, 116, D05 114, https://doi.org/10.1029/2010JD015113, 2011.

Lawrence, D.: The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty, Journal of Advances in Modeling Earth Systems, in press, 2019.

Lenaerts, J. T. M., van den Broeke, M. R., Angelen, J. H., van Meijgaard, E., and Déry, S. J.: Drifting snow climate of the Greenland ice sheet: a study with a regional climate model, The Cryosphere, 6, 891 – 899, https://doi.org/10.5194/tc-6-891-2012, 2012.

Lenaerts, J. T. M., Medley, B., van den Broeke, M. R., and Wouters, B.: Observing and Modeling Ice Sheet Surface Mass Balance, Reviews of Geophysics, 57, 376 – 420, https://doi.org/10.1029/2018RG000622, 2019.

Ligtenberg, S. R. M., Helsen, M. M., and van den Broeke, M. R.: An improved semi-empirical model for the densification of Antarctic firn, The Cryosphere, 5, 809 – 819, https://doi.org/10.5194/tc-5-809-2011, 2011.

Ligtenberg, S. R. M., Munneke, P. K., Noël, B., and van den Broeke, M. R.: Brief communication: Improved simulation of the present-day Greenland firn layer (1960-2016), The Cryosphere, 12, 1643 – 1649, https://doi.org/10.5194/tc-12-1643-2018, 2018.

Lipscomb, W. H., Price, S. F., Hoffman, M. J., Leguy, G. R., Bennett, A. R., Bradley, S. L., Evans, K. J., Fyke, J. G., Kennedy, J. H., Perego, M., Ranken, D. M., Sacks, W. J., Salinger, A. G., Vargo, L. J., and Worley, P. H.: Description and evaluation of the Community Ice Sheet Model (CISM) v2.1, Geoscientific Model Development, 12, 387 – 424, https://doi.org/10.5194/gmd-12-387-2019, 2019.

Machguth, H., Thomsen, H., Weidick, A., Ahlstrøm, A. P., Abermann, J., Andersen, M. L., Andersen, S., Bjørk, A. A., Box, J. E., Braithwaite, R. J., Bøggild, C. E., Citterio, M., Clement, P., Colgan, W., Fausto, R. S., Gubler, K. G. S., Hasholt, B., Hynek, B., Knudsen, N., Larsen, S., Mernild, S., Oerlemans, J., Oerter, H., Olesen, O., Smeets, C., Steffen, K., Stober, M., Sugiyama, S., van As, D., van den Broeke, M., and van de Wal, R. S.: Greenland surface mass balance observations from the ice sheet ablation area and local glaciers, Journal of Glaciology, pp. 1 – 27, https://doi.org/10.1017/jog.2016.75, 2016.

Mottram, R., Boberg, F., Langen, P., Yang, S., Rodehacke, C., Christensen, J. H., and Madsen, M. S.: Surface Mass balance of the Greenland ice Sheet in the Regional Climate Model HIRHAM5: Present State and Future Prospects, Low Temperature Science, 75, 105 – 115, https://doi.org/10.14943/lowtemsci.75.105, 2017.

Mouginot, J., Rignot, E., Bjørk, A., van den Broeke B., M., Millan, R., Morlighem, M., Noël, B., Scheuchla, and Wood, M.: Forty six years of Greenland Ice Sheet mass balance from the component method: 1972 to 2018, PNAS, in press., 2019.

Niwano, M., Aoki, T., Hashimoto, A., Matoba, S., Yamaguchi, S., Tanikawa, T., Fujita, K., Tsushima, A., Iizuka, Y., Shimada, R., and Hori, M.: NHM-SMAP: spatially and temporally high-resolution nonhydrostatic atmospheric model coupled with detailed snow process model for Greenland Ice Sheet, The Cryosphere, 12, 635 – 655, https://doi.org/10.5194/tc-12-635-2018, 2018.

Noël, B., van de Berg, W. J., van Meijgaard, E., Munneke, P. K., van de Wal, R. S. W., and van den Broeke, M.: Evaluation of the updated regional climate model RACMO2.3: summer snowfall impact on the Greenland Ice Sheet, The Cryosphere, 9, 1831 – 1844, https://doi.org/10.5194/tc-9-1831-2015, 2015.

Noël, B., van de Berg, W. J., Machguth, H., Lhermitte, S., Howat, I., Fettweis, X., and van den Broeke, M. R.: A daily, 1 km resolution data set of downscaled Greenland ice sheet surface mass balance (1958-2015), The Cryosphere, 10, 2361 – 2377, https://doi.org/10.5194/tc-10-2361-2016, 2016.

Noël, B., van de Berg, W. J., Lhermitte, S., Wouters, B., Machguth, H., Howat, I., Citterio, M., Moholdt, G., Lenarets, J. T. M., and van den Broeke, M. R.: A tipping point in refreezing accelerates mass loss of Greenland's glaciers and ice caps, Nature Communications, 8, 14 730, https://doi.org/10.1038/ncomms14730, 2017.

Noël, B., van de Berg, W. J., van Wessem, J. M., van Meijgaard, E., van As, D., Lenaerts, J. T. M., Lhermitte, S., Munneke, P. K., Smeets, C. J. P. P., van Ulft, L. H., van de Wal, R. S. W., and van den Broeke, M. R.: Modelling the climate and surface mass balance of polar ice sheets using RACMO2, Part 1: Greenland (1958-2016), The Cryosphere, 12, 811 – 831, https://doi.org/10.5194/tc-12-811-2018, 2018.

Noël, B., van de Berg, W. J., Lhermitte, S., and van den Broeke, M. R.: Rapid ablation zone expansion amplifies north Greenland mass loss, Science Advances, 5, eaaw0123, https://doi.org/10.1126/sciadv.aaw0123, 2019.

Overly, T. B., Hawley, R. L., Helm, V., Morris, E. M., and Chaudhary, R. N.: Greenland annual accumulation along the EGIG line, 1959-2004, from ASIRAS airborne radar and detailed neutron-probedensity measurements, The Cryosphere, 10, 1679 – 1694, https://doi.org/10.5194/tc-10-1679-2016, 2016.

Rae, J. G. L., Aõalgeirsdóttir, G., Edwards, T. L., Fettweis, X., Gregory, J. M., Hewitt, H. T., Lowe, J. A., Lucas-Picher, P., Mottram, R. H., Payne, A. J., Ridley, J. K., Shannon, S. R., van de Berg, W. J., van de Wal, R. S. W., and van den Broeke, M. R.: Greenland ice sheet surface mass balance: evaluating simulations and making projections with regional climate models, The Cryosphere, 6, 1275 – 1294, https://doi.org/10.5194/tc-6-1275-2012, 2012.

Shepherd, A., Ivins, E., and et al., E. R.: Mass balance of the Greenland Ice Sheet from 1992 to 2018, Nature, p. 25 pp., https://doi.org/10.1038/s41586-019-1855-2, 2019.

Smith, R., Jones, P., Briegleb, B., Bryan, F., Danabasoglu, G., Dennis, J., Dukowicz, J., Eden, C., Fox-Kemper, B., and Gent, P.: The parallel ocean program (POP) reference manual ocean component of the community climate system model (CCSM) and community earth system model (CESM), Rep. LAUR-01853, 141, 1 – 140, 2010.

Undèn, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A., Jones, C., Lenderlink, G., Mcdonald, A., Mcgrath, R., Navascues, B., Nielsen, N. W., Degaard, V., Rodriguez, E., Rummukainen, M., Sattler, K., Sass, B. H., Savijarvi, H., Schreur, B. W., Sigg, R., and The, H.: HIRLAM-5, Scientific Documentation, technical Report, 2002.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., Bechtold, V. D. C., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka, N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., Berg, L. V. D., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hòlm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., Mcnally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Ster, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961 – 3012, 2005.

Van Angelen, J. H., Lenaerts, J. T. M., van den Broeke, M. R., Fettweis, X., and van Meijgaard, E.: Rapid loss of firn pore space accelerates 21st century Greenland mass loss, Geophysical Research Letters, 40, 2109 – 2113, https://doi.org/10.1002/grl.50490, 2013a.

Van Angelen, J. H., van den Broeke, M. R., Wouters, B., and Lenaerts, J. T. M.: Contemporary (1969-2012) evolution of the climate and surface mass balance of the Greenland ice sheet, Surveys in Geophysics, 35, 1155 – 1174, https://doi.org/10.1007/s10712-013-9261-z, 2013b.

Van de Berg, W. J. and Medley, B.: Brief Communication: Upper-air relaxation in RACMO2 significantly improves modelled interannual surface mass balance variability in Antarctica , The Cryosphere, 10, 459 – 463, https://doi.org/https://doi.org/10.5194/tc-10-459-2016, 2016.

Van den Broeke, M. R., Enderlin, E., Howat, I., Munneke, P. K., Noël, B., van de Berg, W. J., van Meijgaard, E., and Wouters, B.: On the recent contribution of the Greenland ice sheet to sea level change, The Cryosphere, 10, 1933 – 1946, https://doi.org/10.5194/tc-10-1933-2016, 2016.

Van Kampenhout, L., Rhoades, A. M., Herrington, A. R., Zarzycki, C. M., Lenaerts, J. T. M., Sacks, W. J., and van den Broeke, M. R.: Improving the Representation of Polar Snow and Firn in the Community Earth System Model, Journal of Advances in Modeling Earth Systems, 9, 2583 – 2600, https://doi.org/10.1002/2017MS000988, 2017.

Van Kampenhout, L., Lenaerts, J. T. M., Lipscomb, W. H., Lhermitte, S., Noël, B., Vizcaíno, M., Sacks, W. J., and van den Broeke, M. R.: Greenland ice sheet climate and surface mass balance in CESM2, Journal of Geophysical Research-Earth Surface, submitted, 2019a.

Van Kampenhout, L., Rhoades, A. M., Herrington, A. R., Zarzycki, C. M., Lenaerts, J. T. M., Sacks, W. J., and van den Broeke, M. R.: Regional grid refinement in an Earth system model: impacts on the simulated Greenland surface mass balance, The Cryosphere, 13, 1547 – 1564, https://doi.org/10.5194/tc-13-1547-2019, 2019b.

Van Wessem, J. M., van de Berg, W. J., Noël, B. P. Y., van Meijgaard, E., Amory, C., Birnbaum, G., Jakobs, C. L., Krüger, K., Lenaerts, J. T. M., Lhermitte, S., Ligtenberg, S. R. M., Medley, B., nd K. van Tricht, C. H. R., 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, 12, 1479 – 1498, https://doi.org/10.5194/tc-12-1479-2018, 2018.

Wouters, B., Bamber, J. L., van den Broeke, M. R., Lenaerts, J. T. M., and Sasgen, I.: Limits in detecting acceleration of ice sheet mass loss due to climate variability, Nature Geoscience, 6, 613 – 616, https://doi.org/10.1038/ngeo1874, 2013.

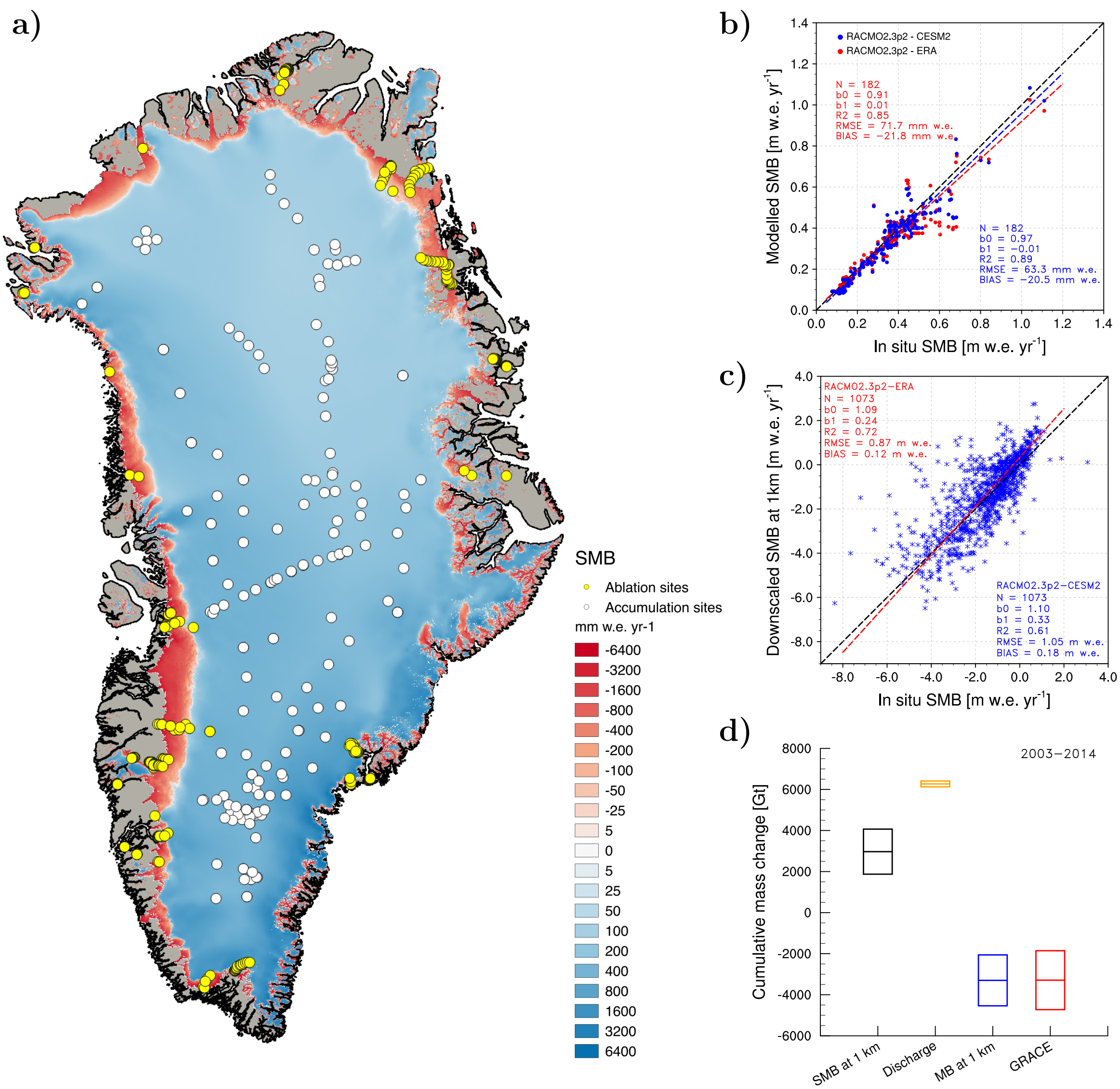


Figure 1. a) Annual mean SMB (1950-2014) as modelled by RACMO2.3p2 forced by CESM2, statistically downscaled to 1 km resolution. b) SMB evaluation in the accumulation zone (182 sites; white dots in Fig. 1 a) and c) in the ablation zone of the GrIS (213 sites; yellow dots in Fig. 1a). Statistics including the number of observations (N), slope (b0) and intercept (b1) of the regressions, determination coefficient (R^2), RMSE and bias are listed for the ERA (red) (Noël et al., 2018) and CESM2-forced RACMO2.3p2 simulation (blue). d) Period (2003-2014) cumulative SMB (black), glacial discharge (orange) (Mouginot et al., 2019), mass balance (MB = SMB - Discharge; blue) and mass loss derived from GRACE (red) (Wouters et al., 2013). To enable a direct comparison with GRACE in d), SMB is integrated over the GrIS, peripheral ice caps and tundra regions of Greenland. Boxes represent one standard deviation around the mean (horizontal bars).

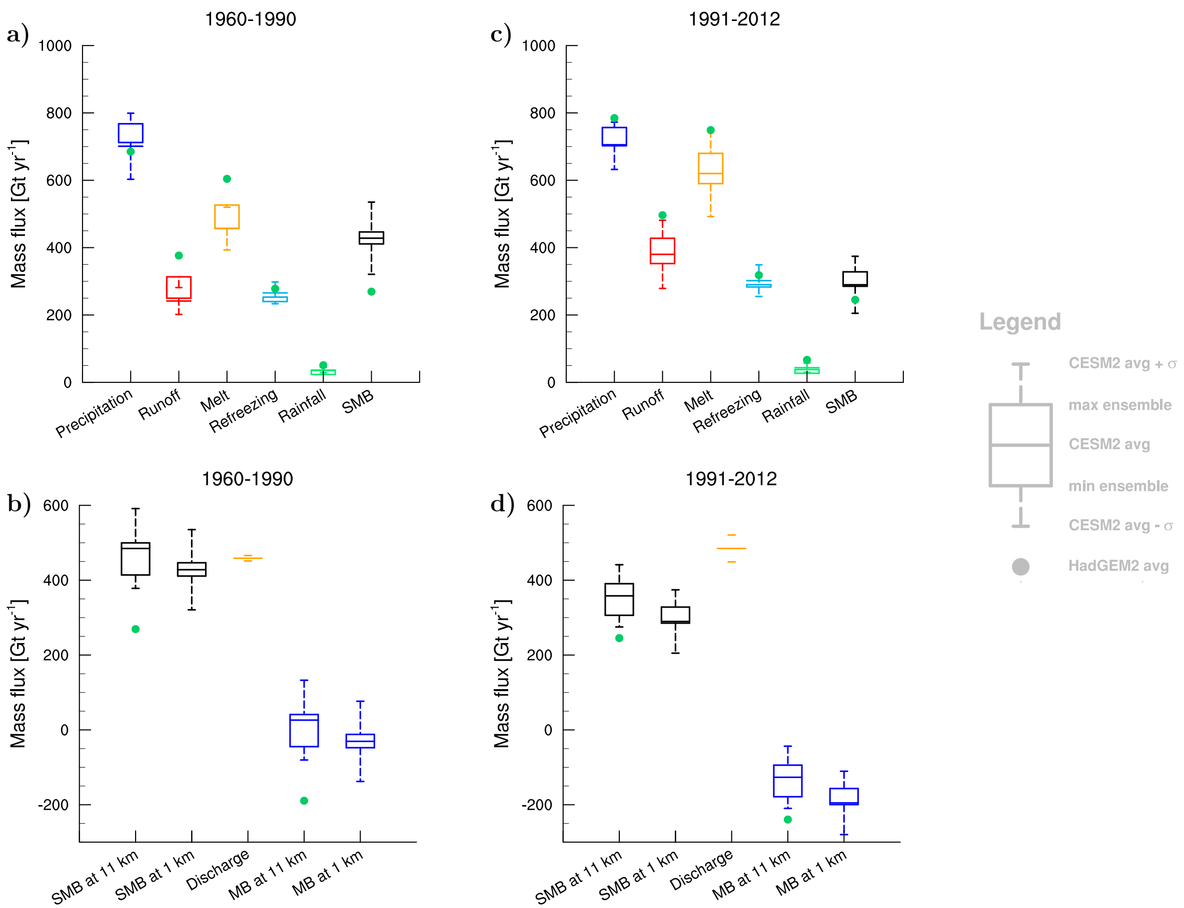


Figure 2. a) Average GrIS SMB (black) and components at 1 km, i.e. total precipitation (blue), runoff (red), melt (orange), refreezing (cyan) and rainfall (green), for the period 1960-1990. b) Annual mean SMB (black), glacial discharge (orange) (Mouginot et al., 2019) and mass balance (MB = SMB - Discharge; blue) at 11 km and 1 km resolution for the period 1960-1990. c) and d) same as a) and b) for the period 1991-2012. The green dots represent values from a previous HadGEM2-forced RACMO2.1 simulation. Boxes around the CESM2-forced RACMO2.3p2 mean (horizontal bars) represent the range of modelled estimates from an ensemble of RACMO2 simulations (five at 11 km, one at 5.5 km, and four at 1 km; Tables A1 and A2). The error bars represent one standard deviation (σ) around the CESM2-forced RACMO2.3p2 mean.

Table A1. Annual mean SMB and components integrated over the GrIS (Gt yr^{-1}) for the period 1960-1990 from an ensemble of RACMO2 simulations using various spatial resolutions and lateral forcing. The uncertainty range corresponds to one standard deviation around the mean. Here mass balance of the GrIS (MB) is estimated as GrIS-integrated SMB minus glacial discharge for the period 1960-1990 (458 Gt yr^{-1}) (Mouginot et al., 2019).

1960-1990	Reference	Forcing	Grid	MB	SMB	PR	RU	ME	RF	RA
RACMO2.1	Van Angelen et al. (2013a)	HadGEM2	11 km	-189±97	269±104	685±82	377±65	604±116	278±66	51±15
RACMO2.1	Van Angelen et al. (2013b)	ERA	11 km	-22±92	437±99	724±77	247±47	424±60	220±27	44±11
RACMO2.3p1	Noël et al. (2015)	ERA	11 km	-45±91	414±98	693±72	237±48	404±62	189±25	22±6
RACMO2.3p2	Noël et al. (2018)	ERA	11 km	28 ±93	487±100	715±73	197±46	345±58	169±20	23±6
RACMO2.3p2	Noël et al. (2019)	ERA	5.5 km	41±96	500 ±104	731±75	197±47	341±59	164±20	23±6
RACMO2.3p2	This study	CESM2	11 km	26±99	485±107	702±98	187±36	332±66	167±35	23±7
RACMO2.3p1	Noël et al. (2016)	ERA	1 km	-48±100	411±108	768±76	313±55	527±66	240±28	24±6
RACMO2.3p2	Noël et al. (2018)	ERA	1 km	-36 ±98	423±106	712±73	257±53	457±56	250±23	35±6
RACMO2.3p2	Noël et al. (2019)	ERA	1 km	-12 ±101	447±109	731±75	250±53	460±59	253±21	33±6
RACMO2.3p2	This study	CESM2	1 km	-31 ±100	428±107	701±98	242±40	457±63	266±32	37±8

Table A2. Annual mean SMB and components integrated over the GrIS (Gt yr^{-1}) for the period 1991-2014 from an ensemble of RACMO2 simulations using various spatial resolutions and lateral forcing. The uncertainty range corresponds to one standard deviation around the mean. Here mass balance of the GrIS (MB) is estimated as GrIS-integrated SMB minus glacial discharge for the period 1991-2012 (485 Gt yr^{-1}) (Mouginot et al., 2019).

1991-2012	Reference	Forcing	Grid	MB	SMB	PR	RU	ME	RF	RA
RACMO2.1	Van Angelen et al. (2013a)	HadGEM2	11 km	-240±80	245±141	785±84	496±114	749±156	318±54	67±17
RACMO2.1	Van Angelen et al. (2013b)	ERA	11 km	-155±80	330±116	731±64	361±100	588±146	279±65	52±14
RACMO2.3p1	Noël et al. (2015)	ERA	11 km	-179±83	306±119	685±64	336±99	543±145	233±60	26±8
RACMO2.3p2	Noël et al. (2018)	ERA	11 km	-95±78	389±114	709±61	286±92	470±133	208±53	26±8
RACMO2.3p2	Noël et al. (2019)	ERA	5.5 km	-94±81	391±117	717±62	291±91	469±131	203±53	26±8
RACMO2.3p2	This study	CESM2	11 km	-127±47	358±83	704±70	314±93	495±131	211±51	32±14
RACMO2.3p1	Noël et al. (2016)	ERA	1 km	-200±95	285±131	757±66	428±109	680±146	283±53	27±9
RACMO2.3p2	Noël et al. (2018)	ERA	1 km	-170±85	315±121	705±61	357±101	590±136	28±52	38±8
RACMO2.3p2	Noël et al. (2019)	ERA	1 km	-157±85	328±121	717±63	353±96	594±131	290±50	36±8
RACMO2.3p2	This study	CESM2	1 km	-195±49	290±85	702±70	380±101	620±128	302±47	44±13

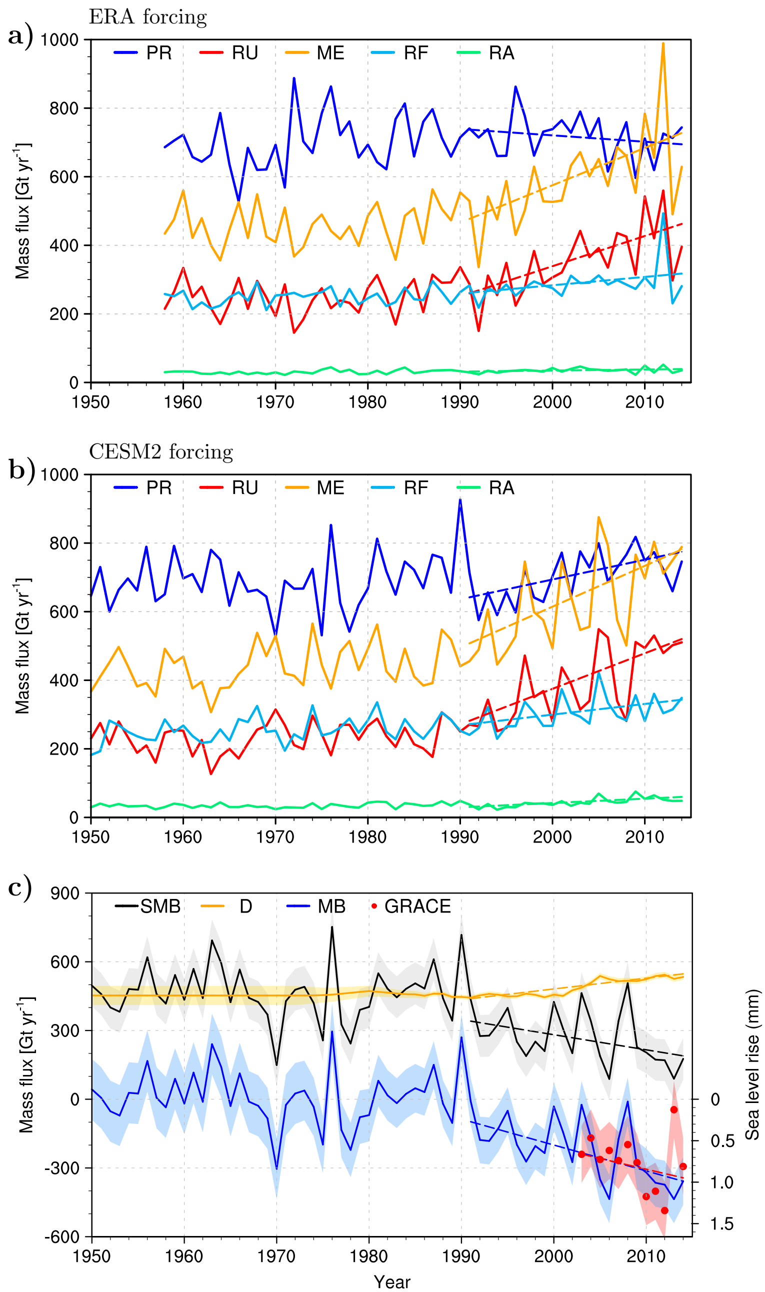


Figure 3. Time series of downscaled (1 km) GrIS-integrated annual SMB components, namely total precipitation (PR; blue), runoff (RU; red), melt (ME; orange), refreezing (RF; cyan) and rainfall (RA; green), as modelled by a) ERA-forced RACMO2.3p2 (1958-2014) (Noël et al., 2019) and b) CESM2-forced RACMO2.3p2 (1950-2014). c) Time series of annual SMB (CESM2-forced RACMO2.3p2 at 1 km), glacial discharge (D) (Mouginot et al., 2019) and mass balance (MB = SMB - D). Mass loss from GRACE (2003-2014) is represented by red dots (Wouters et al., 2013). Dashed lines show the 1991-2014 trends. To enable a direct comparison with GRACE in c), SMB is integrated over the GrIS, peripheral ice caps and surrounding tundra regions of Greenland.

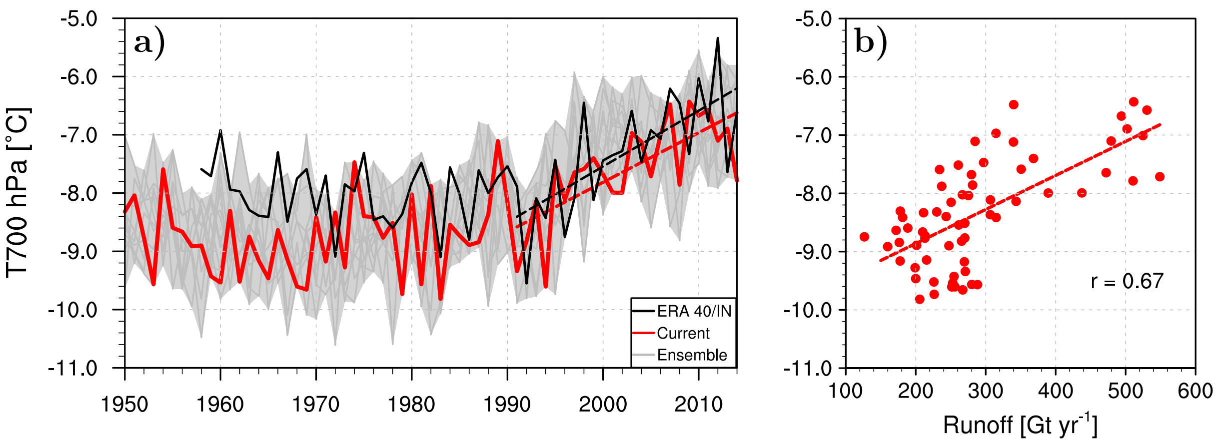


Figure 4. a) Time series of the annual June-July-August (JJA) atmospheric temperature at 700 hPa (T_{700}) averaged over 60-80°N and 20-80°W for the ERA-40 (1958-1978) and ERA-Interim (1979-2014) reanalyses (black), the current CESM2 simulation (red) and 11 additional reference ensemble members (dark grey). The grey belt encompasses annual minimum and maximum values of the whole ensemble. b) Annual GrIS-integrated runoff derived from CESM2-forced RACMO2.3p2 at 1 km resolution as a function of JJA atmospheric temperature at 700 hPa from the current CESM2 simulation (red).