RC1

In their paper, Kampenhout and collogues investigate the impact of regional grid refinement on the Greenland surface mass balance. They compare three CESM simulations with regionally refined grids to a CESM simulation with a uniform grid. Further, they evaluate the model performance using remote sensing data, in-situ measurements as well as regional model simulations. They find, that a grid refinement improves accumulation patterns, however biases develop in the ablation zones. In order to explain differences in the ablation between the simulations they investigate differences in the atmospheric large-scale circulation and clouds. The manuscript is very interesting and well written and discusses a very relevant topic, which after major revision can provide a step forward in understanding resolution dependency of the surface mass balance.

We thank the reviewer for his / her kind words and critical review. Below, we respond to his / her remarks in blue.

Major comments

Results and Limitations section (P.8-13 and 18):

It should be appreciated that the error in the land cover input file is reported in the manuscript. Given the relatively short simulations in combination with the acceptable performance and throughput per day I was wondering why the authors did not consider to rerun the simulations. Although the error might not change the overall conclusions, a lot of absolute values are presented throughout the results section which are very likely to become different if the error was fixed.

The reviewer raises a fair point. Since the other reviewer also raised this point we decided to re-run both VR simulations that were affected by the error. All tables, figures and numbers in the text have been updated accordingly.

Results and Discussions:

I am very skeptical towards not using the same ice sheet mask for a direct comparison between RACMO and the CESM simulations. The ice sheet extent has a significant influence on the Greenland integrated mass balance and specifically melt. Hence, I do not believe that a “fair” comparison can be made without using the same mask. As the authors focus on the low bias of melt in the VR-CESM simulations the authors should reconsider using an identical mask for the comparison; there is a good chance that the bias is a result of the smaller CESM ice mask. The mass conservation argument, as stated in their response to the editors comments, does not hold in my opinion. If the authors decide not to change the overall analysis in the result section, such an analysis should be included somewhere in the manuscript in order to estimate the effects of the mask on the integrated values.

We feel ambiguous on this point, and believe there is value in both. As a middle road, we have updated Table 2 with extra rows that display the integrated mass values over a common ice mask, but maintain the rows in which the components are integrated over the native mask. The analysis in the text uses primarily the common mask and has been updated accordingly.

Minor comments

Abstract: “The SMB in the accumulation zone is significantly improved compared ...” – It should be added that the refinement leads to such improvement.

Thanks, we now start the sentence as follows: “On the refined grids, the SMB …”

Introduction/Model description: It is not clear to me which version of CESM is used. This should be pointed out more clearly (CESM2 or CESM1) from the beginning.

The version of CESM used is neither exactly CESM1 nor exactly CESM2. Indeed, it is an unsupported configuration of components, which gave the best results at the time. It uses the CAM5.4 atmosphere model (part of CESM1), with MG2 microphysics (part of CESM2). We believe that the main results should be reproducible using CESM2, as soon as CAM6 gives good results with variable resolution. At the start of our study, however, VR in CAM6 was still undergoing heavy testing.

For clarity, we’ve added this line to the introduction (P2 L33):
“The version of CESM used resembles the recently released CESM version 2 (CESM2), of which a more in-depth evaluation is planned in the near future.”

Further, the following line was added to Methods (P3, L15):
“VR capabilities in CAM6, the new atmosphere model in CESM version 2 (CESM2, \url{www.cesm.ucar.edu/models/cesm2.0}) were still under beta testing at the start of our study, which explains the slightly older model version of CAM. “

Page 4, Line 8: Are these chosen constants a result of model tuning or used in other studies? Please add references. See also comment on the lapse rate below.

These are taken from the literature. Two references have been added.

Table 1: The values for RACMO should be included for reference, as it would increase the readability and a direct comparison.

We have done this.

Page 12, Line 29: Please hypothesize which counteractive effects are acting.

This sentence has been removed in the updated manuscript, since we did more analysis.

Page 14, Line 12: “. .” – remove one .

Done.

Page 14, Line 23: Please state that a 0.25o mesh is chosen here.

Done.

Page 17, Section “Latent Heat”: Regionally there are large differences and the response seems to be very dependent on the topography. The authors should be a little more detailed in their analysis.

The analysis of latent heat was removed from the revised manuscript, since it is deemed not to be of first order importance.

Page 17, Section “Subgrid downscaling”: Could this be tested by an additional simulation in which the lapse rate during the downscaling is slightly changed? Or can the authors think of another method to test this? See also the lapse rate comment above.

This section has been removed in the revised manuscript, since it was no longer deemed relevant. Just for completeness, there is currently a paper in preparation by Raymond Sellevold (TU Delft) specifically looking at the impact of different lapse rates on melt.

Page 18, Line 23: "&" to "and"

Done

Page 14, Line 32: Have you investigated the blocking frequencies in the different simulations? Or is this a hypothesis?

This was a hypothesis. Note that the Discussions section has been extensively rewritten in the revised manuscript.

RC2

This paper describes a (currently unsupported) configuration of the CAM+CLM atmosphere/land models in a set of short AMIP-style experiments where regional grid refinements have been made over Greenland. The focus of the analysis is on the impact of the local grid refinements on the simulated surface mass balance of the Greenland ice sheet, by comparison with observations and a benchmark regional model. The paper is timely and generally clear and well written. Although useful already, I think it would benefit significantly from a few key changes, if the authors could manage them.

We thank the reviewer for his / her kind words and critical review. Below, we respond to his / her remarks in blue.

Major comments

Some of what I'd like to see improved is already acknowledged by the authors in section 5, Limitations. To start with, is the model so expensive that even one further run cannot be done to properly quantify the impact of the erroneous input file, even if you can't go further and rerun both the VR simulations with the correct boundary files? Section 2.1.5 suggests this should only take a couple of days.

Following this comment and a similar comment by Reviewer 1, we decided to re-do both simulations that were affected by the erroneous input file. All tables, figures and numbers in the text have been updated accordingly.

Another matter highlighted here that would seem to be possible to address is that all results of CESM vs RACMO in the Tables are presented as area-integrated, absolute quantities, although the area of the ice sheet used in each case is different. The difference in areas considered does make it difficult to know how far the anomalies wrt the RCM are really down to the CESM- VR physics, especially in the crucial, high ablation zones at the edges. I think some common area could be defined for analysis purposes, even without formally separating out the main sheet from the peripheral glaciers in CESM as section 5 states cannot yet be done.

We feel ambiguous on this point, and believe there is value in both. As a middle road, we have updated Table 2 with extra rows that display the integrated mass values over a common ice mask, but maintain the rows in which the components are integrated over the native mask. The analysis in the text uses primarily the common mask and has been updated accordingly.

More generally, from the analysis that has been done the paper doesn't come to any firm conclusion as to why there's apparently a systematic trend towards lower Greenland ablation with higher resolution, nor whether this really represents an improvement or degradation in model physics or overall performance. Going further, one would really like to see how the changes play out when the CISM ice sheet is coupled into the system. On pg18, there is a scope-limiting statement that has been left right to the end of the paper - if this paper is a limited, "exploratory first step" I think it that should be stated earlier to set up the readers expectations appropriately. I don't think these matters should really be out of scope for this study, but we should at least be warned earlier if the authors are going to declare that they are.

We agree with RC2 in that it was disappointing that the paper did not come to any firm conclusions on the trend towards lower ablation. We decided to extend our analysis accordingly and added Section 3.1 on large scale circulation and Section 3.8 on clouds and radiation. Still, there is some room for interpretation, which is left to Section 4. Discussions in the revised document.

Further, we feel that this paper should primarily remain concerned with SMB and the interpretation of atmospheric fields, and have not added any new analysis on the dynamical impact. We have added new scope limiting statements, notably:

P1. L9: "…pilot study.."
P2. L27: "… explore the impacts that the refinement has on GrIS SMB."
P21. L25: "… our case study…"

Minor comments

pg1,line7: "starts developing": clarify that the growing bias is a function of resolution rather than eg time.

Good point, we have added "with resolution".

pg1,line13: I don't think you need "relative"

OK, we removed this word.

pg3,line9: I wasn't sure until it was noted as a Limitation in section 5 that there really had only been a single, twenty year run conducted for each CESM configuration shown here. Could that be made clearer, earlier?

We added the following sentences to Sect. 2.1 (Modelling setup)
"Three AMIP-style CESM simulations are performed over the years 1980-1999, a period prior to the onset of persistent circulation change and a strong decline in GrIS SMB in the 2000s \citep{Fettweis2013a, vandenBroeke2016}. Two CESM simulations are regionally refined, which makes that the resolution of the GrIS is different across all simulations (111, 55 and 28 km)."

pg5,line4: "similar to what is done in two-way coupled setups..." this isn't very helpful for readers unfamiliar with the CISM coupling. Could you rephrase, or add a reference?

We've added the "CESM Land Ice Documentation and User Guide" as a reference, available online at https://escomp.github.io/cism-docs/cism-in-cesm/versions/release-cesm2.0/html/clm-cism-coupling.html

pg14, line 12: two full-stops after "resolution"

Thanks, this is fixed now.

section 4.1: how do the circulation anomalies shown compare with the variability in CESM? Could they be further compared with CESM-ERA biases for the relevant period, to judge whether the VR has improved matters? I think our Editor has previously suggested that 700mb is the best altitude to assess temperature biases for relevance to Greenland climate?

We have added Section 3.1 on large scale circulation, where we compare geopotential height and temperature at 500 hPa to ERA-Interim. Also, we have added Figure S4 in the Supplementary Material which does the same for JJA at 700 hPa.

section 4.2: The above-suggested comparison with ERA could usefully be extended to cloud properties, and the surface fluxes could be compared with those from RACMO?

In the newly added Section 3.8 we look at difference in cloud water path w.r.t. the reference simulation (Uniform CESM). We agree that clouds are an important subject for further study, and highlight clouds as a potential model weakness in Section 4. Discussions of the revised manuscript. Further, an evaluation of all surface energy fluxes using RACMO is a significant amount of work, and should involve the elevation class output from CESM. An evaluation of SEB is currently underway for the recently released CESM2, using CAM6 with the standard Finite Volume dynamical core.

I don't think the content in sections 4.3 and 4.4 is really extensive enough to deserve their own separate headings?

The number of separate headings has been reduced, including these.

pg19,line27: Even though the VR configuration isn't supported yet, is it worth adding a link to where people can get CESM, and/or lodging your configuration description with the paper?

We have added a link to the CESM2 web page in the Methods section. Further, our configuration is the result of many technical steps which were not documented in a manner appropriate for publication. Anyone that wishes to use VR-CESM is invited to contact the CESM Atmosphere Model Working Group liaison person for possibilities: http://www.cesm.ucar.edu/working_groups/Atmosphere/

Regional Grid Refinement in an Earth System Model: Impacts on the Simulated Greenland Surface Mass Balance

Leonardus van Kampenhout[1], Alan M. Rhoades[2], Adam R. Herrington[3], Colin M. Zarzycki[4], Jan T.M. Lenaerts[5], William J. Sacks[4], and Michiel R. van den Broeke[1]

[1]Institute for Marine and Atmospheric Research Utrecht, Utrecht University, The Netherlands
[2]Lawrence Berkeley National Laboratory, Berkeley CA, USA
[3]Stony Brook University, Stony Brook NY, USA
[4]National Center for Atmospheric Research, Boulder CO, USA
[5]Department of Atmospheric and Oceanic Sciences, University of Colorado, Boulder CO, USA

Correspondence: Leo van Kampenhout (L.vankampenhout@uu.nl)

Abstract. In this study, the resolution dependence of the simulated Greenland ~~Ice Sheet~~ ice sheet surface mass balance in the variable-resolution Community Earth System Model (VR-CESM) is investigated. Coupled atmosphere-land simulations are performed on ~~three~~ two regionally refined grids over Greenland at ~~1° (∼111 km),~~ 0.5°(∼55 km) ~~,~~ and 0.25° (∼28 km), maintaining a quasi-uniform resolution of 1°(∼111 km) over the rest of the globe. ~~The~~ On the refined grids, the SMB in the accumulation zone is significantly improved compared to airborne radar and in-situ observations, with a general wetting at the margins and a drying in the interior GrIS. Total precipitation decreases with resolution, which is in line with best-available regional climate model results. In the ablation zone, ~~VR-CESM~~ CESM starts developing a positive SMB bias ~~in some locations. Potential driving mechanisms are proposed, amongst which are diversions in large scale circulation,~~ with increased resolution in some basins, notably in the east and the north. The mismatch in ablation is linked to changes in cloud cover ~~, and changes in summer snowfall~~in VR-CESM, and a cold bias that CESM has in these regions. Overall, our ~~results demonstrate that VR-CESM~~ pilot study demonstrates that variable resolution is a viable new tool in the cryospheric sciences~~and can~~ , and could e.g. be used to dynamically downscale ~~future scenarios and/or be interactively coupled to~~ SMB in scenarios simulations, or to force dynamical ice sheet models through the CESM coupling framework.

1 Introduction

The ~~relative~~ contribution of the Greenland ~~Ice Sheet~~ ice sheet (GrIS) to global sea level rise is increasingly determined through its surface mass balance (SMB) (van den Broeke et al., 2016). Accurate estimates of future GrIS SMB are therefore key in providing projections for sea level rise. Arguably the most realistic SMB projections to date are derived from general circulation model (GCM) scenario output downscaled using regional climate models (RCMs — e.g., Rae et al. (2012); van Angelen et al.

(2013); Fettweis et al. (2013a); Mottram et al. (2017); Noël et al. (2018)). Compared to GCMs, the regional models offer more sophisticated snow models that have improved representation of albedo, melt, firn densification and refreezing, features that are lacking in most current GCMs (Ziemen et al., 2014; Helsen et al., 2017). In addition, RCMs typically run at a horizontal grid resolution of $\mathcal{O}$(10 km) whereas atmospheric GCMs are typically run using 1° or $\mathcal{O}$(100 km) grids. RCMs therefore tend to better resolve topographic gradients~~which lead~~ , which leads to more accurate spatio-temporal distributions in precipitation, wind, cloud cover, and temperature, enabling a detailed comparison to in-situ meteorological data. A fine spatial resolution seems indispensable for resolving narrow ablation zones found around the GrIS margins (Lefebre et al., 2005; Pollard, 2010).

Recently, significant efforts have been invested into making GCMs more suitable for ~~ice sheet SMB modelling (e.g., Punge et al., 2012; C~~ and SMB modelling (e.g., Punge et al., 2012; Cullather et al., 2014; Fischer et al., 2014; Helsen et al., 2017; van Kampenhout et al., 2017; Shannon et al., 2019 In particular, improvements made to the Community Earth System Model (CESM) include a multilayer snow model with a two-way radiative transfer model for albedo (Flanner and Zender, 2005), ~~improved~~ enhanced snow density parameterizations (van Kampenhout et al., 2017), and the introduction of multiple elevation classes for downscaling SMB with height (Lipscomb et al., 2013).

~~Although globally uniform~~ Still, significant biases remain with respect to RCMs, as the spatial resolution is limited (Vizcaíno et al., 2013; Helsen et al., 2017). Although high-resolution ~~climate~~ GCM simulations exist (50 km, Delworth et al. (2011); 25 km, ~~Wehner et al. (2014); Small et al. (2014)~~Wehner et al. (2014); Small et al. (2014); Bacmeister et al. (2014); 80 km, Müller et al. (2018)) ~~,~~a majority of ongoing modelling experiments, notably the forthcoming CMIP6 experiments (Eyring et al., 2016), maintain a $\sim$1° atmosphere grid due to limitations in computational resources.

A middle road may be found in new techniques that apply regional grid refinement within a global climate model. In this approach, a static global mesh is constructed which has increased resolution over a specified region of interest. Over the past five years, ~~rapid~~ progress has been made in developing regional grid refinement in CESM — variable resolution or VR-CESM. ~~Studies have been carried out looking~~ To date, studies looked at the effect of grid refinement on the global circulation and climatology ~~(Zarzycki et al., 2015)~~(Zarzycki et al., 2015; Gettelman et al., 2018), the effect on tropical cyclones (Zarzycki and Jablonowski, 2014), regional climate in the presence of mountains ~~(Rhoades et al., 2017; Huang et al., 2016; Rhoades et al., 2015)~~(Rhoades et al., 2015; Huang et al., 2016; Rhoades et al., 2017), and the scale dependence of the underlying physics ~~(Gettelman et al., 2018; Herrington and Reed, 2018)~~(Gettelman et al., 2018; Herrington and Reed, 2018). Compared to RCM downscaling, Huang et al. (2016) notes several advantages of the variable resolution (VR) approach. First, using a unified modelling framework avoids the inconsistencies between RCM and GCM, in particular the different dynamical core and physics that are used. Second, VR allows for two-way interactions (i.e., downstream / upstream ~~feedbacks~~effects) between the refinement region and the global domain, which an RCM downscaling approach does not. ~~Some~~ Finally, some more practical advantages are the ~~relative simplicity~~ attractiveness of operating a single ~~model~~ modelling framework, and the relatively low computational cost associated with ~~VR compared to a multi-model setup~~VR-CESM.

In this ~~study~~paper, we apply ~~VR-CESM over Greenland and evaluate the impact~~ regional grid refinement over the Greenland area using VR-CESM, and explore the impacts that the refinement has on GrIS SMB. ~~In particular, we aim to understand the dependence of resolution on outstanding GrIS climate biases in CESM. Simulations are carried out at different spatial resolutions (i.e., 111,~~ Two VR meshes are constructed with refined patches centered around the GrIS with 55 ~~,~~ km and 28 km ~~)~~ ~~over the greater Greenland region~~resolution, respectively. A 20-year atmosphere-only simulation spanning the historical period (1980-1999) is carried out over each of those grids and is then compared to a reference simulation that has no refinement. Results are ~~compared to~~ also compared to reanalyses, airborne snow accumulation radar, in-situ SMB measurements, and gridded climate data from an RCM, in an ongoing effort to improve the representation of ice sheets in CESM (Lipscomb et al., 2013; Vizcaíno et al., 2013; Lenaerts et al., 2016; van Kampenhout et al., 2017). ~~A more~~ The version of CESM used resembles the recently released CESM version 2 (CESM2), of which a more in-depth ~~evaluation of CESM over the GrIS, including surface energy balance, is planned for the upcoming release of CESM 2.1.~~

evaluation will be published in the near future. The layout of the manuscript is as follows. In Section 2, the modelling setup and benchmark data are described ~~. Main results~~ in further detail. The main findings are presented in Section 3. ~~The discussion~~ A discussion follows in Section 4 that is concerned with ~~attributing the changes in ablation found across the different resolutions. Some general limitations of our modelling study are reviewed in Section ??. Finally, an overall summary as well as conclusions are provided~~ interpreting an outstanding bias in CESM, not alleviated by the grid refinement, and guiding future research directions. Finally, a summary with conclusions is found in Section 5.

2 Methodology

2.1 ~~Model description~~Modelling setup

The Community Earth System Model (CESM) is a global climate modelling framework comprised of several components, ~~e.g.~~ i.e. atmosphere, ocean, land surface, sea ice, and land ice, that may operate partially or fully coupled. When partially coupled, the missing components can be substituted by external data or even inactive (stub) components. Here, we follow the protocol of the Atmospheric Model Intercomparison Project (AMIP, Gates et al. (1999)) and dynamically couple the atmosphere-land components and prescribe ocean and sea ice data at monthly intervals ~~. The prescribed sea ice and sea surface temperatures are taken from Hurrell et al. (2008) for the years of simulation~~ (Hurrell et al., 2008). Our three AMIP-style CESM simulations are carried out over the years 1980-1999, a period prior to the onset of persistent circulation change and a strong decline in GrIS SMB in the 2000s (Fettweis et al., 2013b; van den Broeke et al., 2016). Aerosol and trace gas emissions are taken as observed.

2.1.1 ~~Atmosphere model~~

The atmosphere component used is the Community Atmosphere Model ~~(CAM)~~ version 5.4 ~~(Neale et al., 2012)~~ (CAM5.4, Neale et al., 2012) with the spectral element ~~(SE) dynamical core (Dennis et al., 2012; Lauritzen et al., 2018)~~ dynamical core (CAM-SE, Dennis et al., 2012; Lauritzen et al., 2018), the only dynamical core currently in CESM supporting

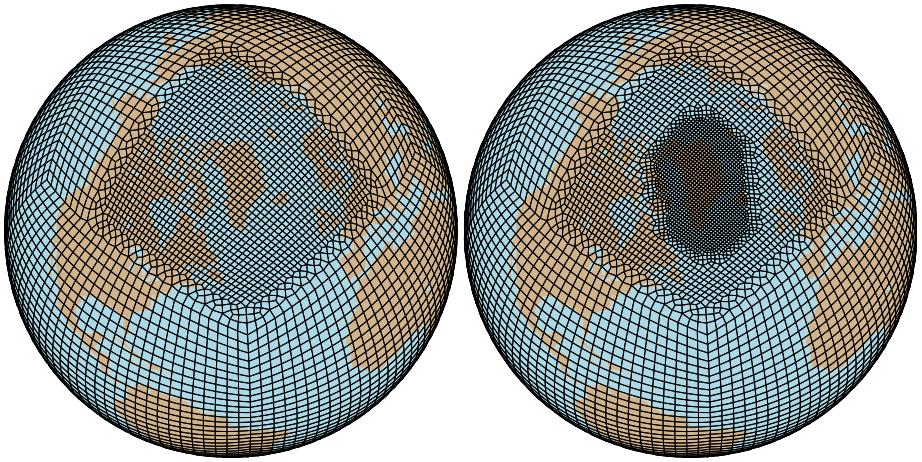

Figure 1. Computational domains of experiments VR-CESM55 (left) and VR-CESM28 (right). Each spectral element visible here contains an additional 3-by-3 grid of points, the exact position of which are determined by the spectral element method (Zarzycki and Jablonowski, 2014).

VR capabilities (Zarzycki et al., 2014). VR capabilities in CAM6, the new atmosphere model in CESM version 2 (CESM2, www.cesm.ucar.edu/models/cesm2.0) were still under beta testing at the start of our study, which explains the slightly older model version of CAM. Our model configuration ~~is almost identical to~~ broadly follows that of Zarzycki and Jablonowski (2014), with a few ~~small differences~~modifications. These include updated rain and snowfall microphysics (MG2, Gettelman and Morrison (2014)), a new dry-mass~~vertical coordinate and slightly reduce~~ , floating Lagrangian, vertical coordinate with 32 levels in the vertical, and slightly reduced horizontal diffusion in the SE dynamical core (Lauritzen et al., 2018)~~, and the use of the~~ . Further, we adopt the Beljaars et al. (2004) orographic drag parameterization ~~, replacing~~ which replaces the turbulent mountain stress (TMS) scheme of CESM1 (Neale et al., 2012) in order to achieve ~~higher and more realistic~~ more realistic (higher) wind speeds over the ice sheets. Physics tuning coefficients were set to default values specified in the supported CAM5 release.

~~Hyperviscosity coefficients are to be scaled by the grid-resolution for~~

2.2 Grids

Our reference simulation, referred to as Uniform CESM, uses a standard cubed sphere grid which is quasi-uniform at 1° (∼111 km) resolution globally (Evans et al., 2012). The first non-uniform VR mesh has a refined patch of 0.5° (∼55 km) over the greater Greenland region and is referred to as VR-CESM55. The patch was constructed such that the boundary of the patch always extends at least six spectral elements away from the Greenland coast (Figure 1). This buffer region is intended to allow incoming "low-resolution" storms to develop finer-scale structures after entering the VR zone and prior to making landfall (Matte et al., 2017). The second VR mesh is constructed off the VR-CESM55 grid, yet features a second level of refinement at 0.25° (∼28 km) inside the first level. This second patch was chosen such that, again, the boundary extends at least six spectral

elements away from the Greenland coast. The simulation on this grid is referred to as VR-CESM28 and both VR grids were constructed using SQuadGen (Ullrich, 2014).

Topographic height over Greenland was interpolated from the 4 km CISM ice sheet domain, which in turn has been derived from the 90 m Greenland Ice Mapping Project product (GIMP, Howat et al. (2014)). Topography is static in time – ice sheet dynamics are not active in this configuration – a reasonable assumption for the decadal length simulations presented in this paper. The new ice topography was spliced into the global topography, similar to what is done in two-way coupled setups where ice sheet dynamics are turned on[1]. Due to the hybrid sigma vertical coordinate system implemented in CAM-SE, a differential smoothing procedure was applied to ensure numerical stability and ~~filter undesirable numerical artifacts. In~~ realistic flow, as described by Zarzycki et al. (2015). Subgrid height variances, used by the orographic drag parameterization, are consistently recomputed as a residual of the smoothed topography.

The resulting GrIS topographies are shown in Figure 2, with a rendering of GIMP for comparison. As one would expect, a more detailed and accurate representation of topography is possible on finer resolutions. The feature most prominently improving is the southern ice dome, that "rises up" from ~2300 m at 111 km to ~2900 m at 28 km. Furthermore, the 28 km resolution seems sufficiently detailed to start resolving some of the fjord structures, especially in the east. The non-zero topographic heights over open ocean in Figure 2 are explained by the differential smoothing procedure.

The CAM physics (dynamics) time steps for Uniform CESM was 1800 (150) seconds. For the VR-CESM ~~, local hyperviscosity~~ runs, the physics time step was set to 450 s and the CAM dynamics time steps scaled with horizontal resolution with VR-CESM55 at 150 s and VR-CESM28 at 75 s. Hyperviscosity coefficients are scaled by the ~~underlying local element dimensions~~grid-resolution (element dimensions) for numerical stability and filter undesirable numerical artifacts (Guba et al., 2014). Here the scaling is such that the hyperviscosity coefficients are reduced by an order of magnitude for each doubling of the resolution (Zarzycki and Jablonowski, 2014).

~~Note that whereas CAM-SE is an integral part of the release of CESM2, the VR framework is not yet officially supported~~Some minor grid imprinting was noted in the grid transition zone over distorted SE elements. It is ~~expected that a few scientifically validated VR-CESM simulations will be publicly available at a later stage during the release cycle of CESM2~~deemed unlikely, however, that these small, local anomalies materially impact the large-scale synoptic flow in the interior of the domain.

2.2.1 ~~Land surface model~~

2.3 Land surface model

CAM is coupled to the Community Land Model (CLM) version 5.0, which incorporates several important bug fixes and snow parameters updates for CESM2. CLM simulates the interaction of the atmosphere with the land surface, notably the surface energy balance and hydrological processes such as interception by canopy, throughfall, infiltration, and runoff (Oleson, 2013). For radiation calculations over snow, the two-way radiative transfer model SNICAR is used (Flanner and Zender, 2005). The

[1] For details, see the CESM Land Ice Documentation and User Guide, https://escomp.github.io/cism-docs/cism-in-cesm/versions/release-cesm2.0/html/clm-cism-coupling.html

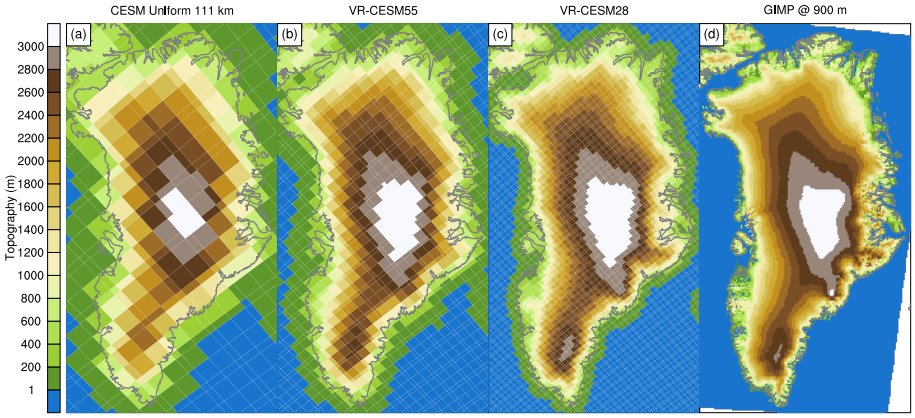

Figure 2. Topographic height in the three CESM simulations. For plotting purposes, spectral element node heights are displayed on control volumes equal to the area that they represent. The control volumes are identical to those used by the CESM coupler to conserve mass and energy. For reference, topographic height according to the Greenland Ice Mapping Project (GIMP, Howat et al. (2014)) is shown, which has been upscaled to 900 m.

snow pack hydrological and thermal evolution is modelled as a one-dimensional column, which can reach depths of up to 10 metres of w.e. (water-equivalent). Several snow model modifications have been implemented specifically for ice sheets, such as wind-dependent fresh snow density and wind driven snow compaction (van Kampenhout et al., 2017) and temperature dependent fresh snow grain size. Bare ice albedo is assumed constant, and is set to 0.50 (0.30) for the visible (near-infrared) spectrum, reflecting the distinction made between these two shortwave bands in CLM (Oleson, 2013).

Over glaciated grid cells, CLM maintains 10 different elevation classes in order to more accurately ~~downscale SMB and temperature to the 4 km ice sheet grid with class weights~~ capture SMB gradients. Each elevation class is linked to an independent CLM column, and is given a weight proportional to the subgrid topography present in the CLM grid cell (Lipscomb et al., 2013). Classes with zero weights are considered "virtual" and do not contribute to the grid cell average. ~~In each elevation class or bin, CLM maintains a independent column for snow and ice (or soil , over tundra). These columns~~ The columns are independent as they maintain their own private snow pack and ice / soil variables. Columns tend to evolve differently depending on their height, as temperature, specific humidity and longwave radiation are downscaled ~~with~~ to the mean class elevation. Lapse rates are used for temperature (6 K km^{-1}, Lipscomb et al. (2013)) and longwave radiation (32 W m^{-2} km^{-1}, based off Figure 6 in Van Tricht et al. (2016a)), whereas specific humidity follows the assumption that relative humidity remains constant with height. Elevation classes have been shown to improve SMB gradients over Greenland (Vizcaíno et al., 2013). In order to activate elevation classes over Greenland, the Community Ice Sheet Model (CISM) must be active as a diagnostic component.

SMB is the main focus of this paper. We deviate from the standard CLM definition of SMB, which does not take into account changes in snow pack height, in favour of the definition that is common to glaciology, in absence of redistribution/erosion by

drifting snow

$$\mathrm{SMB} = \mathrm{Precipitation} - \mathrm{Sublimation} - \mathrm{Runoff}. \quad (1)$$

For the remainder of the paper, modelled accumulation (ablation) is defined as modelled SMB for locations where SMB > 0 (SMB < 0). Unless indicated otherwise, annual CESM SMB data used have been downscaled to the 4 km CISM ice sheet grid, thus taking advantage of the multiple elevation classes in CLM. In this downscaling procedure, elevation classes are used in the vertical direction whereas a bi-linear interpolation is applied in the horizontal direction to prevent artificial jumps between grid cells (Leguy et al., 2018).

Following Rhoades et al. (2018), the distribution of plant functional types in CLM is assumed constant at year 2000 values for all simulations. As the main focus of this work is on precipitation and snow cover in non-vegetated regions, we argue this assumption has a negligible impact on our results.

~~Computational domains of experiments VR-CESM55 (left) and VR-CESM28 (right). Each spectral element visible here contains an additional 3-by-3 grid of points, the exact position of which are determined by the spectral element method (Zarzycki and Jablonowski, 2014).~~

2.3.1 ~~Grids, time stepping, topography~~

~~Three identical atmosphere-land simulations are performed, each at a different grid. The first mesh is a standard, globally quasi-uniform grid at 1° (~111 km) resolution. The simulation on this grid is referred to as Uniform CESM. The second mesh is nonuniform and has a refined patch of 0.5° (~55 km) over the greater Greenland region, constructed using SQuadGen (Ullrich, 2014), and is referred to in this paper as VR-CESM55. The patch was constructed such that the boundary of the patch always extends at least six (nominally, $20\Delta x$) spectral elements away from the Greenland coast (Figure 1). This buffer region is intended to allow incoming "low-resolution" storms to develop finer-scale structures after entering the VR zone, and prior to making landfall. The third mesh is constructed from the VR-CESM55 grid but features a second level of refinement at 0.25° (~28 km) inside the first level. Again, this second patch was chosen such that the boundary extends at least six spectral elements away from the Greenland coast. The simulation on this grid is referred to as VR-CESM28.~~

~~Topographic height over Greenland was interpolated from the 4 km CISM ice sheet domain, which in turn has been derived from the 90 m Greenland Ice Mapping Project product (GIMP, Howat et al. (2014)). Topography is static in time – ice sheet dynamics are not active in this configuration – a reasonable assumption for the decadal length simulations presented in this paper. The new ice topography was spliced into the global topography, similar to what is done in two-way coupled setups where ice sheet dynamics are turned on. Due to the hybrid sigma vertical coordinate system implemented in CAM-SE a differential smoothing procedure was applied to ensure numerical stability and realistic flow, as described in Zarzycki et al. (2015). Subgrid height variances, used by the orographic drag parameterization are consistently computed as a residual of the smoothed topography.~~

~~The resulting topographies are shown in Figure 2, with a rendering of GIMP for comparison. As one would expect, a more detailed and accurate representation of topography is possible on finer resolutions. The feature most prominently improving~~

~~is the southern ice dome, that "rises up" from ~2300 m at 111 km to ~2900 m at 28 km. Furthermore, the 28 km resolution is sufficient to start resolving some of the fjord structures, especially in the east. The non-zero topographic heights over open ocean in Figure 2 are explained by the differential smoothing procedure.~~

~~The CAM physics (dynamics) time steps for Uniform CESM was 1800 (150) seconds. For the VR-CESM runs, the physics time step was set to 450 s and the CAM dynamics time steps scaled with horizontal resolution with VR-CESM55 at 150 s and VR-CESM28 at 75 s.~~

~~Topographic height in the three CESM simulations. For plotting purposes, spectral element node heights are displayed on control volumes equal to the area that they represent. The control volumes are identical to those used by the CESM coupler to conserve mass and energy. For reference, topographic height according to the Greenland Ice Mapping Project (GIMP, Howat et al. (2014)) is shown, which has been upscaled to 900 m.~~

2.3.1 ~~Initialisation~~

2.4 Initialisation

In glaciated regions, the subsurface conductive heat flux at the ice sheet surface is potentially large due to the high thermal conductivity of ice. To avoid unrealistic energy losses or gains from the subsurface, one should start with ice that is in thermal equilibrium with the ambient climate. In our modelling setup, however, a sufficiently long spinup period to achieve such equilibrium was not feasible due to computational constraints. Instead, it was decided to initialise deep ice temperature from values close to observed, in this case 10 m firn temperatures from a firn densification model, forced by RCM-downscaled reanalysis data (Ligtenberg et al., 2018). A nearest neighbour procedure was followed to interpolate ice temperature from the 11 km firn model to the different resolutions used in this study.

Snow height was reset globally to a low value of 100 mm w.e. to avoid snow cover hysteresis arising from errors in the interpolated initial conditions. This reset was limited to grid cells — or rather, CLM columns — below 1774 m, which is an estimate of the maximum present-day GrIS equilibrium line altitude. A spinup simulation was carried out to rebuild snow in any CLM columns residing below the reset altitude (1774 m) where the local climate dictates perennial snow cover. The relevance of this spinup is two-fold: (1) the dependence of fractional snow cover on snow height (Swenson and Lawrence, 2012), (2) the refreezing capacity of the snow pack. For both of these, a period of 5 years was deemed sufficiently long to capture the first-order effect. Nonetheless, it is recognised that the resulting snow depth distribution over the GrIS contains an artificial jump at 1774 m.

2.4.1 ~~Performance~~

2.5 Performance

All simulations have been performed on NCAR's supercomputing facility "Cheyenne" in Wyoming, USA, which is equipped with Intel Broadwell processors. No real load-balancing was needed since the active components (i.e., CAM, CLM, CISM, and

coupler) perform well when sharing all the available cores. On 1800 cores (or 50 compute nodes) the cost of Uniform CESM at 1° (48,602 CAM-SE grid points) amounts to ∼1070 core hours per simulated year. Keeping the number of cores the same, this cost was tripled to 3250 core hours for the VR-CESM55 simulation with the ~~large~~ refined patch of 0.5° ~~,~~ (59,402 CAM-SE grid points), and quadrupled to ∼4300 core hours for the VR-CESM28 simulation with the ~~addition of the smaller~~ additional 0.25° patch (69,887 CAM-SE grid points). By comparison, the computational cost of limited area model RACMO2 at 11 km is ∼6800 core hours per simulated year (Brice Noël, pers. comm.). The throughput was ∼25, ∼13 and ∼10 simulated years per day ~~,~~ for Uniform CESM, VR-CESM55 ~~,~~ and VR-CESM28, respectively.

2.6 Reference data

~~Model results are quantitatively compared to accumulation and ablation observations (remotely-sensed and in-situ) and RCM output. CESM and RCM model data are taken from 1980-1999, a period prior to~~

Output from the three CESM simulations is interpreted using reference data from a variety of sources. The evaluation of the climate at synoptic scales is supported by atmospheric reanalyses, i.e. hindcast climate models that employ data assimilation to match the observed state of atmosphere as close as possible. In particular, temperature and geopotential height from the European Centre for Medium-Range Weather Forecasts Reanalysis (ERA-Interim, Dee et al. (2011)) and the ~~onset of persistent circulation change and a strong decline in GrIS SMB in the 2000s (Fettweis et al., 2013b; van den Broeke et al., 2016). The field~~ Modern-Era Retrospective Analysis for Research and Applications-2 (MERRA2, Molod et al. (2015)) products are used.

For evaluation of GrIS near-surface climate and surface mass balance, data from the Royal Netherlands Meteorological Institute (KNMI) regional atmospheric climate model (RACMO) version 2.3p2 (RACMO2 hereafter) are used. RACMO2 is a state-of-the-art polar climate model that has been extensively evaluated over the GrIS (Noël et al., 2018, 2015) and compares favourably to observations. At its lateral boundaries, RACMO2 was forced using ERA-Interim data and the native spatial resolution of the data is 11 km. When appropriate, however, the statistically downscaled product at 1 km is used, which better resolves narrow ablation zones and low-lying regions (Noël et al., 2016, 2018). We argue that it is fair to compare VR-CESM directly to the downscaled 1 km RACMO2 product as (i) CESM also performs on-line downscaling using the semi-statistical elevation classes (Section 2.3), and (ii) best-estimate data is preferred in order to identify either model improvements or regressions, in line with the purpose of this paper. Still, these best-estimate benchmark data are subject to some uncertainty. Noël et al. (2018) characterise the native spatial resolution of 11 km as a source of model uncertainty, as well as the representation of surface roughness and surface albedo. Two prime uncertainties in the RACMO2 downscaling procedure arise from the bare ice albedo used to correct runoff, and the ice sheet extent (Noël et al., 2016).

Field data analysis has been ~~partly carried out in~~ carried out through the Land Ice Verification & Validation Toolkit (LIVVkit), an open source software package designed for evaluating ice sheet models ~~(Kennedy et al., 2017; Evans et al., 2018). For details on how model coordinates are mapped to the observations, the reader is referred to Evans et al. (2018).~~

~~Airborne accumulation radar data~~ and their forcing (Kennedy et al., 2017; Evans et al., 2018). Three observational SMB datasets available for the GrIS are used: (i) airborne radar, (ii) field accumulation (SMB > 0) measurements, and (iii) field ablation (SMB < 0) measurements (Evans et al., 2018). The airborne radar data stems from NASA's Operation IceBridge

~~are used to evaluate modelled accumulation rates in~~ and covers most of the GrIS interior. The raw data, as described by Lewis et al. (2016), ~~described in detail by Lewis et al. (2016). There are 18, 968 values in this dataset, available as a seasonal average over one or more time periods~~provides seasonal estimates for a given pixel, uniquely determined by its latitude and longitude. Following Evans et al. (2018), a simple time average is applied over all available periods for each record to yield a single accumulation value (in mm w.e. yr^{-1}) per location.

~~In-situ measurements of SMB (in mm w. e. yr^{-1}) are split into accumulation zone (SMB > 0) and ablation zone (SMB < 0) subsets. The accumulation zone collection is made up out of 421 firn core/snow pit/stake measurements taken from various field campaigns (Cogley, 2004; Bales et al., 2009; Evans et al., 2018). The ablation zone collection~~ The resulting number of IceBridge data points is a sizeable 18,968, which means that the spatial density is quite high over the radar transects. During the evaluation, a nearest neighbour method is used to determine the model cell closest to each observation.

The in-situ field accumulation dataset is a compilation of different field campaigns carried out in the GrIS accumulation zone (Cogley, 2004; Bales et al., 2009; Evans et al., 2018). Only records that have been retrieved using firn cores, snow pits or stake measurements are included in the evaluation. Moreover, if there are multiple measurements at one location then the data is averaged in time to yield a climatological SMB estimate for that location. In total, the number of accumulation zone measurements is 421. The in-situ field ablation dataset is a subset of the compilation of GrIS ablation zone SMB measurements by Machguth et al. (2016). ~~Each~~ Again, each record location is averaged in time to yield an annual SMB estimate. Only records that are on the CISM ice mask and have a record length equal or close (i.e., within a 5% difference) to a full year are kept, which brings the total number of records down from 627 to ~~142~~163, spread over 22 rather than 46 glaciers. It is important to mention at this point that the spatial coverage of the ablation zone measurements is quite sparse. ~~This can be seen from~~ Indeed, Figure 1 in Evans et al. (2018) ~~, which shows that all~~ illustrates that all in-situ ablation data stem from ~~only~~ merely 8 transects in total.

~~The regional climate model used is the Royal Netherlands Meteorological Institute (KNMI)regional atmospheric climate model (RACMO) version 2.3p2 (RACMO2 hereafter)~~

3 Results

3.1 Large-scale circulation

We start with a comparison of modelled mid-troposphere climate to reanalyses data, which serves two purposes. First, it is useful to identify any significant climatic biases that CESM possesses, which could aid in interpreting e.g. snow melt rates later on. Second, the VR approach allows for feedbacks between the domain of interest and the global climate system, in contrast to dynamical downscaling using RCMs. One such feedback could be changes to the strength and location of planetary waves both in and outside the VR domain, due to the higher and steeper topography (Figures 1 and 2). If such upstream / downstream dynamical effects are present in our modelling setup they would make an imprint on mid-tropospheric climate on a hemispheric scale.

CESM geopotential height (Z500) at ~~11 km, a state-of-the-art polar climate model that has been extensively evaluated over~~ 500 hPa is compared against ERA-Interim over the period 1980-1999. Note that the choice of ERA-Interim versus MERRA-2 does not impact our results much, so only ERA-Interim is shown. In boreal summer, the season most relevant to GrIS SMB, anomaly maps of Z500 display consistent patterns across all three CESM simulations (Figure 3a-c). A positive height anomaly is found over the Arctic ocean, which is most pronounced in the Uniform CESM 111 km simulation. It is surrounded by a band of negative height anomalies in all three simulations, with one of the maxima approximately centred over Iceland / South Greenland, indicating more cyclonic flow over the GrIS in CESM. At mid-latitudes, positive height anomalies are found instead, indicating more anti-cyclonic flow there. Over the region 55-90N, the ~~GrIS (Noël et al., 2018, 2015) and compares favourably to observations. At its lateral boundaries, RACMO2 was forced using~~ Z500 root mean squared error (RMSE) is decreased from 3.6 dam (Uniform CESM) to 3.5 dam (-2%, VR-CESM55) and 3.3 dam (-8%, VR-CESM28), which could signal minor benefits of the grid refinement on resolving the large-scale circulation. However, no VR-CESM grid point within the domain of interest can be significantly differentiated from Uniform CESM ($p < 0.05$, indicated by hatching in Figure 3b-c). Furthermore, similar decreases in RMSE are not evident in the other three seasons (Figures S1-S3, Supplementary Material) so these changes can be attributed to internal variability.

Similarly, anomalies of 500 hPa air temperature (T500) with respect to ERA-Interim ~~reanalysis data. The 11 km RACMO2 fields, notably runoff, have been statistically downscaled to a resolution of~~ are computed (Figure 3d-f). Major features in T500 are again shared by the three simulations, such as a cold bias exceeding 0.75 °C over Russia, which is most pronounced in VR-CESM55 and VR-CESM28. A slight JJA warm bias of around 0.5 - 1 ~~km in order to better resolve narrow ablation zones and low-lying regions (Noël et al., 2016, 2018). For the purpose of this work, it is considered fair to compare VR-CESM directly to the downscaled RACMO2 product as (i) CESM also performs on-line downscaling using the semi-statistical elevation classes,~~ °C is indicated over the Arctic Ocean and Northern Greenland in all three simulations. Over the region 55-90N, RMSE is decreased from 0.74 °C (Uniform CESM) to 0.68 °C (-8%, VR-CESM55) and ~~(ii)the best-estimate data would be used in order to identify possible model improvements or deficits, which is in line with the purpose of this paper~~0.63 °C (-14 %, VR-CESM28). Similar improvements are found in MAM (Figure S1) but not in the other two seasons (Figures S2 and S3). Some point-by-point significance is found between VR-CESM and Uniform, albeit not over the Greenland area (Figure 3d-f).

4 ~~Results~~

To conclude, heights at 500 hPa seem not substantially affected by the enhanced resolution and topography in VR-CESM, in particular over the area of interest (Greenland). In all three CESM simulations, more cyclonic flow is indicated with respect to ERA-Interim. Temperature at 500 hPa demonstrates a weakly positive bias in CESM, and shows no significant change with refinement over the GrIS. A weak signal cannot be excluded, however, as it may remain undetected by the Student's t-test due to the relatively short sample period of 20 years.

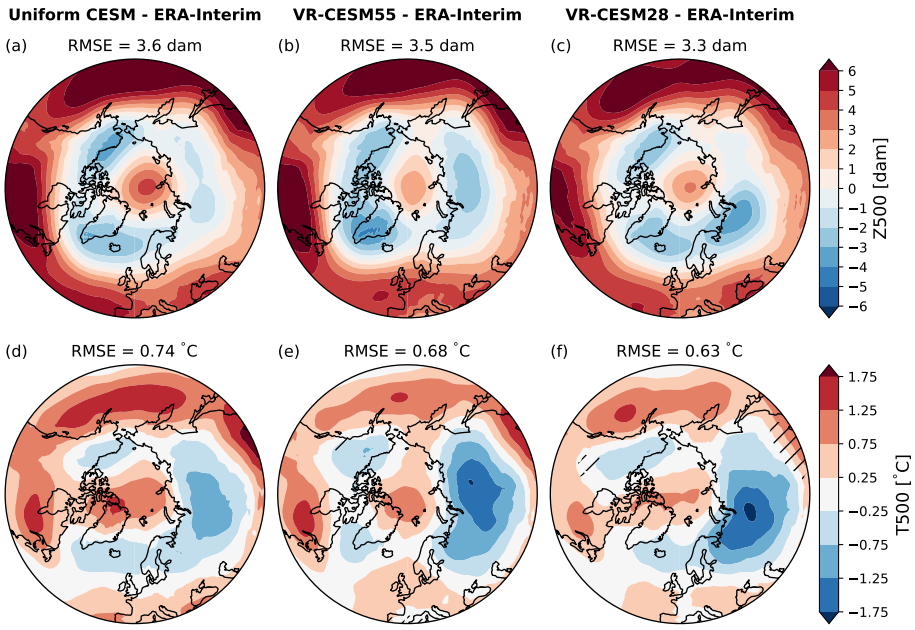

Figure 3. Mean summer (JJA) anomalies of 500 hPa geopotential height (Z500, panels a-c) and air temperature (T500, panels d-f) with respect to ERA-Interim over the period 1980-1999. Shown is 55-90N, the same region over which the area-weighted RMSE was calculated that is listed above each panel. VR-CESM simulations are significantly different ($p < 0.05$) from Uniform CESM in the hatched areas (panels b,c,e,f). Note that no significance is found in panels b, c, and e. Prior to subtraction, all data have been regridded to a common regular mesh of 1° using bi-linear interpolation.

3.1 Precipitation

~~Steep~~ Both the steep edges of ice sheets as well as topographic promontories are effective drivers of orographic precipitation, as is e.g. apparent from the RACMO2 precipitation field (Figure 4d). The largest source of moisture is the North Atlantic basin, which is connected to Greenland by large-scale storm systems (Sodemann et al., 2008). Cyclonic activity associated with the persistent Icelandic Low drives warm and moist air onto land from the south-east, resulting in strong orographic uplift which causes rapid cooling, condensation, and precipitation. By comparison, northern Greenland is much drier with accumulation rates locally below ~~100~~ 150 mm yr^{-1} (Cogley, 2004).

Since orographic precipitation is so dominant over southern Greenland, it is not surprising that we find significant improvements with increasing resolution ~~, compared with RACMO2~~ (Figure 4), compared with RACMO2. At uniform 111 km resolution, CESM correctly predicts a band of high (> 1000 mm yr^{-1}) precipitation rates in the south-east, however it extends too far into the interior (~~panel~~ Figure 4a). This is attributed to the fact that the poorly resolved topography is ~600 m lower in the model than in reality (Figure 2a) and that topographic gradients are smoothed out, which weakens the effect of orographic uplift. The VR-CESM55 result (~~panel~~ Figure 4b) shows that this is mostly a resolution issue as the band of high precipitation rates is more confined to the low-lying areas and slopes, similar to RACMO2. Other effects that can be seen

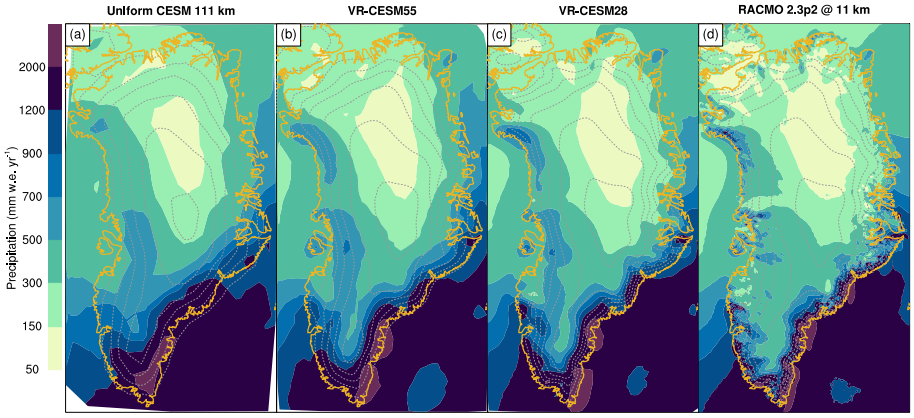

Figure 4. Spatial distribution of mean annual precipitation over Greenland. ~~Note the non-linear colour scale.~~ CESM data are displayed at the native CAM resolution for the period 1980-1999. RACMO2 data are shown at native 11 km resolution for the same period. Coastlines and 500 m elevation contours are overlain in orange and grey, respectively. Note the non-linear colour scale.

in this VR-CESM55 result ~~is~~ are the emergence of orographic precipitation in other locations around the margins, ~~however~~ albeit weak, and a general drying of the northern interior. In VR-CESM28, similar resolution dependent patterns continue to emerge, with even stronger orographic precipitation and more pronounced drying in the north (Figure 4c). Integrated over the entire GrIS, including peripheral glaciers and ice caps, precipitation is reduced from 946 ± 107 Gt yr^{-1} (Uniform CESM) to ~~823~~ 870 $\pm$ ~~55~~ 72 (VR-CESM55) and ~~806~~ 821 $\pm$ ~~61~~ 62 Gt yr^{-1} (VR-CESM28). By comparison, RACMO2 simulates a mean annual precipitation flux of 743 ± 64 Gt yr^{-1} over these glaciated areas. Both the improved patterns (Figure 4) and the more reasonable integrated amount of precipitation over the GrIS are positive results for the application of VR-CESM to this region. ~~Next, we look into accumulation and carry out a more quantitative analysis using the available reference data (Section ??) .~~

3.2 ~~Accumulation (~~IceBridge~~)~~

~~The IceBridge radar data supports~~ Operation IceBridge accumulation data is used to further quantify the effect of the improved precipitation patterns on SMB. As described in Section 2.3, CESM SMB is downscaled to 4 km using the elevation classes method and averaged over the period 1980-1999, prior to comparison to the processed IceBridge SMB samples (Section 2.6). Figure 5 displays the resulting SMB anomalies in mm w.e. yr^{-1}. As can be seen, the IceBridge radar data support the pattern of interior drying with increasing resolution. In Uniform CESM at 111 km resolution, we find a mean wet bias of 81 mm yr^{-1} which is most pronounced in regions near the edges of the IceBridge domain (Figure 5a). The strongest ~~wet~~ bias is found in the south, where absolute precipitation rates are highest (Figure 4) ~~which means that~~ and any relative error will consequently lead to a ~~greater~~ larger absolute error. ~~Distributions of point-by-point deviations are depicted in Figure 6a. Here it can be seen that with~~ With increasing resolution, the mean bias drops from 81 mm to ~~22~~ 37 mm (VR-CESM55) and ~~21~~ 24 mm (VR-CESM28), which suggests that the largest improvement is made going from 111 km to 55 km ~~. At the same time the spread of deviations~~ (Figure 1). The largest SMB differences remain to be found near the margins of the IceBridge domain

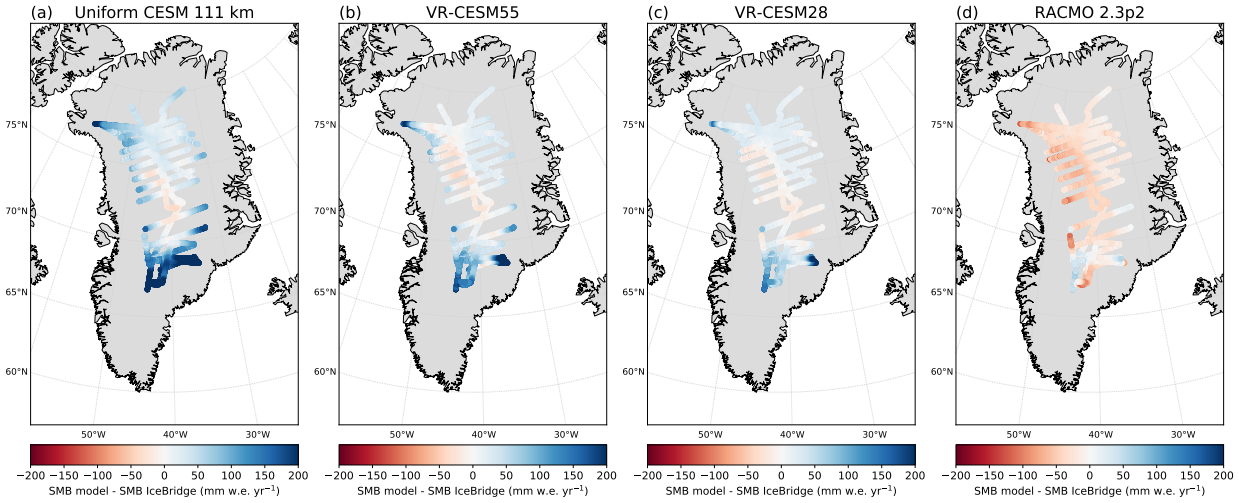

Figure 5. ~~Accumulation difference CESM -~~ SMB differences between IceBridge radar (mean over available period) and model climatology (1980-1999). Blue (red) colours indicate that the model is wetter (dryer) than observations.

(Figure 5b-c). The spread in SMB anomaly also decreases with resolution, ~~i.e. the distribution narrows. For instance, the 95th percentile, as a one-sided~~ which can be visually seen as a narrowing of the SMB anomaly distribution in Figure 6a. As a measure for this spread, ~~falls from 298 mm to 122~~ the difference between the 95th percentile and the 5th percentile falls from 308 mm (Uniform CESM) to 178 mm (VR-CESM55) and ~~96~~ 115 mm (VR-CESM28)~~. IceBridge data suggest that a relatively small dry bias starts to develop in the western GrIS at higher resolutions (Figure 5). The corresponding 5th-percentiles of the point-by-point deviations are -12~~ , respectively. As another measure, the RMSE decreases from 126 mm (Uniform CESM) ~~, -28 mm (~~to 68 mm (-46%, VR-CESM55) and ~~-20 mm (~~46 mm (-64%, VR-CESM28). ~~As a result, the root mean squared error (RMSE) in VR-CESM55 (VR-CESM28) is reduced by 56% (65%) with respect to the uniform resolution, and~~ At the same time, the coefficient of determination (r^2~~is enhanced by 0.1 (0.15))~~ is enhanced substantially (Table 1). The bias and RMSE of the regional RACMO2 are -25 mm and 38 mm, respectively, which suggests a dry bias in RACMO2 (Figure 5d). We conclude that based these statistics, VR-CESM28 performs on-par with RACMO2 (Table 1).

3.3 Accumulation ~~(in-situ measurements)~~sites

~~Next, a~~ A similar analysis is carried out for the in-situ accumulation zone observations. Compared to the airborne radar data, these measurements cover a greater portion of the GrIS~~(cf. Figure 1 in Evans et al., 2018)~~, including the southern dome ~~, which makes it arguably~~ (cf. Figure 1 in Evans et al., 2018), which should make it more representative of the GrIS as a whole~~, although the spatial density is smaller~~. As before, the greatest absolute improvement is found in the doubling of resolution from 111 km to 55 km, with smaller benefits going further to 28 km (Figure 6b and Table 1). The mean bias ~~is halved (-107 mm) going from the uniform 111 km resolution to 55 km, and is further reduced by 30 mm~~ substantially reduces from 187 mm (Uniform CESM) to ~~63 at 28 km (Table 1) . This pattern of marginal returns is also reflected in RMSE (53%vs 62%~~

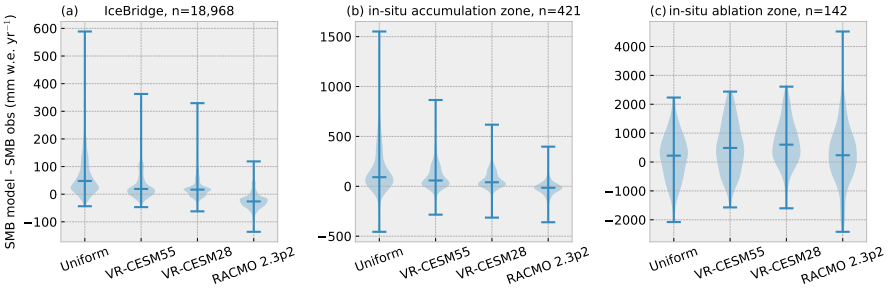

Figure 6. Distribution of point-by-point SMB differences between model ~~(at 4 km)~~ and reference observations. Horizontal ~~lines~~ line segments indicate ~~high/low maxima and~~ maximum, median, ~~respectively~~and minimum value. ~~Note~~ Model climatologies have been computed over the ~~different scales~~period 1980-1999.

Table 1. ~~Statistics~~ Selected statistics of CESM ~~simulations~~ climatological SMB (~~at~~ downscaled to 4 km) and RACMO2 climatological SMB (downscaled to 1 km) with respect to ~~reference data (~~IceBridge radar data, ~~in-situ~~ accumulation zone sites, ~~in-situ~~ and ablation zone ~~)~~sites. Shown are mean bias, coefficient of determination, and root mean square error~~(RMSE)~~. ~~All values are significant in~~ Model climatologies have been computed over the ~~two-sided Student's t-test (p<0.01). On each metric~~period 1980-1999, which not necessarily overlaps with the ~~best ranking CESM simulation is highlighted in bold~~date of each measurement. ~~The reference data are described in the text.~~

	IceBridge ($n = 18,968$)			Acc. sites ($n = 421$)			Abl. sites ($n = 163$)		
	bias (mm)	r^2	RMSE (mm)	bias (mm)	r^2	RMSE (mm)	bias (mm)	r^2	RM
Uniform CESM 1°	~~81.3~~ 81	0.78	126	187	0.61	319	~~**184**~~ 170	~~0.67~~ 0.71	~~****
VR-CESM55 ~~/ 0.5 °~~	~~22.2~~ 37	0.88	~~55.7~~ 68	~~79.6~~ 105	0.74	~~150~~ 172	~~251~~ 462	~~0.66~~ 0.69	~~**
VR-CESM28 ~~/ 0.25 °~~	~~**21.4**~~ 24	~~**0.93**~~ 0.92	~~**44.7**~~ 46	~~**63.4**~~ 71	~~**0.79**~~ 0.79	~~**120**~~ 124	600	0.72	
RACMO 2.3p2	-25	~~574~~0.94	~~**0.70**~~ 38	-13	~~995~~0.71	91	160	0.54	

~~decrease) and r^2 (+0.13 vs +0.18)for the~~ 105 mm (-44%, VR-CESM55~~and~~) and 71 mm (-62%, VR-CESM28) and the RMSE reduces from 319 mm (Uniform CESM) to 172 mm (-46%, VR-CESM55) and 124 (-61%, VR-CESM28). A small positive accumulation bias remains even in the highest resolution run (VR-CESM28~~simulations, respectively. To put the numbers in Table 1 in perspective, Noël et al. (2018) reports a mean bias for~~), a bias that is not apparent in the RACMO2 data (Table 1). For RACMO2, the bias and RMSE values are similar to those mentioned by Noël et al. (2018), who report in their Figure 11a an accumulation zone mean bias of -22 mm (here: -13 mm) and an RMSE of 72 mm ~~for their accumulation zone measurements (version 2.3p2, their Figure 11a), and $r^2 = 0.85$. The bias in VR-CESM28 is still three times that of RACMO2, but the RMSE and r^2 are comparable~~(here: 91 mm). Our r^2 is slightly lower, however, 0.71 against 0.85. These differences can be explained by the different methodology used. Namely, Noël et al. (2018) correlate SMB values based off daily data, thus reflecting the meteorological conditions during which the measurement was made, whereas here we compare climatological averages of the model to each measurement, which introduces additional noise in the comparison.

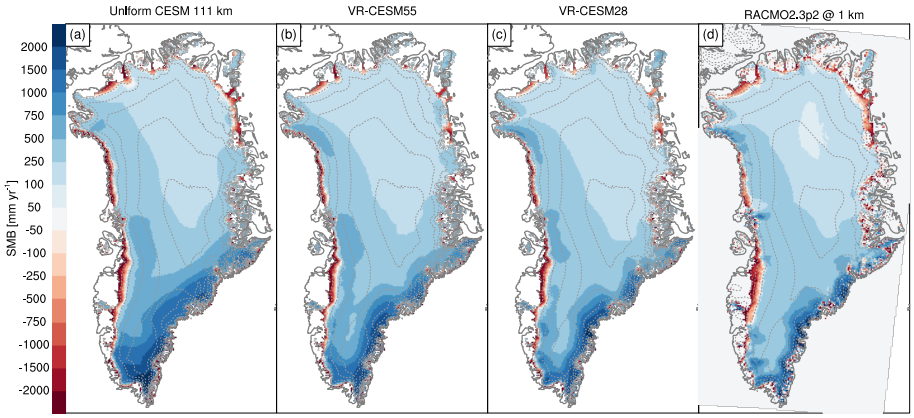

Figure 7. Mean annual SMB in mm yr^{-1}. All CESM data are downscaled to 4 km CISM resolution for the period 1980-1999. RACMO2 data have been statistically downscaled from 11 to 1 km. Note the non-linear colour scale.

3.4 Ablation ~~(in-situ measurements)~~sites

High up on the ice sheet, and thus deep into the accumulation zone, SMB is dominated by snowfall. In the ablation zone, by contrast, there is a delicate balance between different factors — snowfall, sublimation, snowmelt, refreezing, and runoff — that complicates SMB modelling. Furthermore, SMB gradients are typically much stronger in the ablation zone than they are in the accumulation zone, mainly due to steep topography and non-linearity of SMB with height (Figure 7). Therefore, as one ~~therefore~~ expects, model skill in the ablation zone is lower than in the accumulation zone, ~~with errors exceeding 2000 mm~~ signalled by a larger spread and modelling biases exceeding 1000 mm at many locations (Figure 6c)~~at some locations. Still, the coefficient of determination is reasonable ($r^2 > 0.65$ for all simulations, Table 1) and~~ . Nonetheless, ablation zones are mostly predicted in the right locations (Figure 7), owing to the elevation class subgrid parameterization ~~(Section 2.3)~~described in Section 2.3.

~~Different to accumulation zone~~In contrast to the accumulation zone and somewhat surprisingly, model skill in the ablation zone does not improve with resolution (Table 1). The mean bias ~~deepens from 184~~ grows from 170 mm (Uniform CESM) to ~~251~~ 462 mm (VR-CESM55) and ~~574~~ 600 mm (VR-CESM28), which are ~~increases of 36 % and 212~~ substantial increases of +172% and +253%, respectively. The ~~r^2 is similar across all simulations, with the largest value of 0.70 for simulation VR-CESM28. The RMSE is slightly increased in both VR simulations to 970 mm (+15 %)and 995 mm (+18 %) , respectively. To put these numbers in perspective again, Noël et al. (2018) reports a mean bias for~~ model spread is only marginally detoriated, and RMSE ranges 793 - 951 mm for all simulations (Table 1). The ablation statistics of the overall best simulation (Uniform CESM) are comparable to those of RACMO2 which are, analogous to CESM, computed using a 1980-1999 climatology. The bias, r^2, and RMSE of RACMO2 are considerably worse than those reported by Noël et al. (2018), who find a bias of 120 mm ~~and an~~ (here: 160 mm), an r^2 of 0.72 (here: 0.54), and RMSE of 870 mm ~~for~~ (here: 922 mm) in their ablation zone ~~measurements (version 2.3p2 downscaled to 1km,~~ comparison with similar data (their Figure 11c). Again, this is explained by the different

Table 2. Mean GrIS mass fluxes for the period 1980-1999 in gigatonnes per year with standard deviation between brackets. The area of integration is listed in the first column and includes peripheral glaciers and ice caps (GIC). CESM data are integrated at the native resolution with elevation class weighing. The statistically downscaled 1 km RACMO2.3p2 data is averaged over the same period and described in Noël et al. (2018). RACMO2 does not differentiate between snow and ice melt in its output files so only total melt is reported.

Model name	Ice area km^2	Precipitation Gt yr^{-1}	Ice melt Gt yr^{-1}	Total melt Gt yr^{-1}	Refreezing Gt yr^{-1}	Runoff Gt yr^{-1}	Sublimatio Gt yr^{-1}
	native ice sheet extent, including GIC						
Uniform CESM 1°	1,812,467	946 (107)	217 (48)	468 (100)	178 (43)	349 (67)	28 (3)
VR-CESM55 ~~/ 0.5 °~~	1,812,254	~~823 (55~~870 (72)	~~179 (46~~146 (25)	~~447 (108~~387 (70)	~~191 (49~~185 (39)	~~300 (69~~260 (42)	39 (~~43~~)
VR-CESM28 ~~/ 0.25 °~~	1,812,254	~~806 (61)~~ 821 (62)	~~142 (52~~131 (34)	~~384 (104~~377 (73)	~~183 (42~~195 (35)	~~248 (74~~239 (47)	44 (~~52~~)
RACMO2 ~~1 km~~	1,761,475	743 (64)	-	577 (81)	309 (27)	344 (68)	33 (2)
	contiguous GrIS extent						
Uniform CESM 1°	1,705,508	893 (104)	157 (37)	361 (85)	150 (40)	258 (53)	26 (3)
VR-CESM55	1,692,629	796 (69)	115 (20)	314 (62)	159 (37)	203 (34)	36 (3)
VR-CESM28	1,697,054	745 (59)	105 (28)	304 (63)	165 (33)	184 (38)	40 (2)
RACMO2	1,700,772	707 (61)	-	509 (72)	263 (25)	298 (58)	32 (2)

methodology used. In particular, we believe that some extreme ablation events that happened after the year 2000 are not well captured by the climatological mean of the two 20th century decades considered here. When the period of the RACMO2 climatology is changed to 1995-2017, we find a bias of -9 mm, an r^2 of 0.69, and RMSE of 722 mm, which confirms that the time frame used is a crucial factor. Overall, we conclude that both VR-CESM55 and ~~$r^2 = 0.72$. We note that although the bias is larger in all our CESM simulations, RMSE and r^2 are comparable. The bias in the ablation zone of the non-downscaled RACMO2 is 600 mm(Noël et al. (2018), their Figure 11b).~~ VR-CESM28, despite their higher resolution over the GrIS, fail to recover in-situ ablation rates with a skill similar or better than the reference simulation. Instead, a strong positive SMB bias develops in some ablation zone sites, suggesting too little runoff and/or too much precipitation in these locations.

~~Mean annual SMB in mm yr^{-1}. All CESM data are downscaled to 4 km CISM resolution for the period 1980-1999. RACMO2 data have been statistically downscaled from 11 to 1 km. Note the non-linear colour scale.~~

3.5 ~~Integrated SMB (regional climate model data)~~

~~The main SMB components are integrated over their respective ice masks and listed~~

3.5 Integrated SMB

All major surface mass balance components are now integrated over the ice mask native to each model, as well as a common ice mask. Compared to RCMs, which are strongly forced by atmospheric reanalyses, our AMIP-style simulations experience

relatively weak forcing at the ocean boundaries, which renders it unlikely that the actual historical Greenland weather conditions are reasonably resolved. Furthermore, a 20-year model simulation is arguably not long enough to attain a robust mean climate. Hence, the numbers presented in Table 2 ~~. Compared to the 1-km statistically downscaled~~ should be interpreted with some caution, as RACMO2 ~~data,~~ and CESM are not necessarily experiencing the same climate. The common ice mask is constructed based on the contiguous GrIS definition, as laid out by the PROMICE mapping project (Citterio and Ahlstrøm, 2013), which is bilinearly upscaled from the 1 km RACMO domain to the respective CESM grids. In the remainder of this section, we will focus on the results that were obtained on the common ice mask.

GrIS-integrated precipitation is overestimated in all CESM simulations ~~overestimate GrIS-integrated precipitation~~with respect to the RACMO2 regional model (Table 2). The bias in precipitation is largest for Uniform CESM (~~203 ± 125~~ +186 Gt yr^{-1}, or +~~27~~ 26 %) and reduces ~~to 11 % (~~with increasing resolution to +89 Gt (+13%, VR-CESM55) and ~~8 % (~~+38 Gt (+5%, VR-CESM28)~~which~~ . This is in line with our earlier findings of progressive drying with increased resolution ~~. Part of remaining bias may be related to the ~ 3 % larger ice sheet extent in CESM, relative to RACMO2.~~

discussed in Sections 3.1, 3.2, and 3.3. Melt, on the other hand, is consequently underestimated in all CESM simulations (Table 2). The bias ~~of~~ in total melt volume is smallest for coarse-resolution Uniform CESM (~~-19~~ -148 Gt, or -29%) and largest for fine-resolution run VR-CESM28 (~~-34 %), which is consistent with the findings presented in the previous section on ablation zone measurements. Going from Uniform CESMat 111 km to~~ -205 Gt, or -40%). Melt is reduced by 47 Gt in VR-CESM55 ~~the reduction in ice melt(38 Gt ± 66) exceeds the reduction in total melt (21 ± 147), which must mean that snow melt is increasing. This is no longer the case at the next resolution doubling (VR-CESM28)where both ice and snow melt are reduced. The enhanced snow melt in VR-CESM55 helps to explain why total refreezing is largest in this simulation (~~and by 57 Gt in VR-CESM28, with respect to Uniform CESM. The majority of that is due to ice melt, which sees similar reductions of 42 Gt and 52 Gt, respectively (Table 2), with snow melt accounting for the remainder of 5 Gt in both cases. Refreezing volume is comparable across the three different CESM simulations (Table 2)~~, although the differences across the CESM simulations are small.~~

~~In CESM, surface runoff~~ , with only slightly higher numbers at higher resolution. These could be explained, for instance, by lower snow temperatures ("cold content") in these runs, which is consistent with the lower melt rates found. Surface runoff in CESM is the sum of bare ice melt and drainage from the bottom of the snow pack, i.e. liquid water originating from rain or melt that does not refreeze. Due to the reductions in total melt volume, runoff is also significantly reduced at higher resolutions (Table 2), leading to ~~large~~ significant negative biases when compared against the downscaled 1 km RACMO2 data. ~~A small positive bias in Uniform CESM~~ With respect to RACMO2, Uniform CESM underestimates runoff by 40 Gt (~~5 Gt ±~~ -13%), VR-CESM55 by 95 Gt ~~) drops to negative biases in VR-CESM55 (-44~~ (-32%), and VR-CESM28 by 114 Gt (-38%), which agrees with the reduction in ablation found in Section 3.4. Sublimation is enhanced in both VR runs compared to Uniform CESM (Table 2) which we attribute to higher 10 m wind speeds occurring in VR-CESM (not shown). GrIS sublimation in VR-CESM28 is 54 % higher than in Uniform CESM, and exceeds the RACMO2 figure by 8 Gt.

Overall, GrIS integrated SMB exceeds 500 Gt in all CESM simulations (Table 2), which is markedly more than the 376 ± ~~97 Gt)~~ 99 that RACMO2 estimates over the common mask. There appear to be two balancing factors. On one hand, precipitation

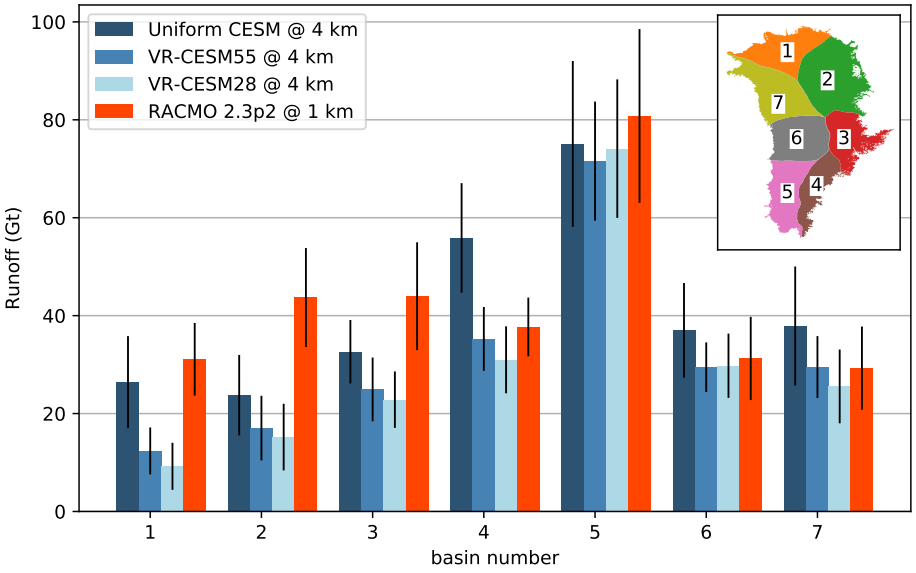

Figure 8. Mean basin-integrated runoff over the period 1980-1999. Error bars represent one standard deviation. CESM data have been manually downscaled down from their native resolution to 4 km using vertical SMB profiles generated by the elevation classes. For reference, RACMO2 downscaled runoff at 1 km resolution is shown. The extent of all basins combined equals the common ice mask in Table 2. Due to the manual interpolation, however, the total runoff for CESM does not match the value reported in Table 2.

is overestimated in all CESM runs, and more so in the runs at low resolution (Uniform CESM and VR-CESM55). On the other hand, runoff is underestimated in all CESM runs, and more so in the runs at high resolution (VR-CESM55 and VR-CESM28~~(-96~~). However, the decrease in precipitation is larger than decrease in runoff, which makes that the lowest integrated SMB value is found in VR-CESM28 (521 $\pm$ ~~100 Gt~~ 77 Gt yr^{-1}).

~~To investigate further the spatial variability~~

3.6 Runoff

The decreased integrated runoff found in both VR-CESM simulations is an unexpected regression with respect to the uniform resolution simulation, and may be linked to a similar regression in reproducing ablation zone measurements (Section 3.4). It is interesting to examing the spatial heterogeneity of runoff, ~~we computed total runoff across~~ and uncover any regional differences. To this end, runoff is aggregated over 7 ~~GrIS~~ major GrIS drainage basins, derived from an ~~updated~~ ice flow mosaic ~~of Rignot and Mouginot (2012). The results can been seen~~ updated from Rignot and Mouginot (2012). Downscaled RACMO2 runoff at 1 km is used as a reference and the results are shown in Figure 8. ~~CESM runoff is monotonically decreasing with increasing resolution in all regions except region 5 and 6. In the northern and eastern parts of the GrIS (basins~~ Clearly, Uniform CESM underestimates mean runoff in basin 1 ~~,~~ (north), basin 2 ~~and 3) this leads to a deeper negative bias with respect to RACMO2. In these basins, the mean runoff as simulated by VR-CESM28 is in fact outside the range of one RACMO2 standard~~

deviation , in contrast to Uniform CESM which is still within that range in basin 1 and 3. In the south-east (north-east), and basin 3 (east). In both VR runs, runoff decreases further in these regions and now falls completely outside of the standard deviation of RACMO2. In basin 4 (south-east), runoff is significantly substantially overestimated in Uniform CESM . This is attributed to (Figure 8). This can be explained through the poorly resolved precipitation field , which in reality exhibits in Uniform CESM. In reality, precipitation has steep gradients over this region (see e.g. basin that are not resolved due to the coarse resolution (Figure 4). With increased precipitation at In both VR runs, precipitation shifts to lower elevations, more meltwater can be buffered and less bare ice reaches the surface, thereby limiting runoff . The runoff simulated by which enhances meltwater buffering / refreezing and prevents bare ice exposure, two mechanisms through which runoff can be limited. Indeed, VR-CESM55 runoff now falls within one standard deviation of the RACMO2 result RACMO2 standard deviation in basin 4, which is a positive result for this simulation, whereas VR-CESM28 runoff is slightly lower than RACMO2. The largest absolute runoff flux is simulated in south-western basin 5. The results of this region are interesting because the response of runoff to resolution increase is not consistently in the same direction, which suggests that different, counter-acting mechanisms are active. Nevertheless, all CESM results found in basin 5 (south-west), which is equally well resolved by all CESM simulations, with numbers that fall within one standard deviation of the RACMO2 due to the high variability in this region. Runoff in basins estimate. Finally, runoff in basin 6 and (west) and basin 7 is consistently overestimated in all CESMsimulations with (north-west) is slightly overestimated in Uniform CESM, a bias that appears to be removed in both VR runs. In summary, our basin analysis indicates that runoff is particularly biased in the north, north-east, and east basins of the GrIS, and that this bias grows stronger with increasing resolution. In the other basins, VR-CESM runoff seems not substantially biased with respect to RACMO2, but less so at higher resolutions.

At this point it is worth reiterating that the comparison to RACMO 2.3p2 downscaled data at 1 km Noël et al. (2018) is a rather stringent test.

3.7 Clouds

What could be causing the observed reduction in runoff, especially pronounced in the northern and eastern basins? Here, we aim to shed some light on that, in a qualitative fashion. Likely, large-scale circulation changes can be excluded as the prime driving mechanism, based on the results presented in Section 3.1. In particular, no statistically significant changes in Z500 and T500 were found over the Greenland area in summer (Figure 3). The same holds true when the lower troposphere is inspected at the 700 hPa pressure level (Figure S4, Supplementary Material), although some cooling is observed over northern Greenland, but this may include a thermodynamic effect of a cooler surface.

This leaves local meteorological conditions as the prime candidate mechanism. Clouds in particular are an important modulator of the surface energy balance, with thick clouds emitting more longwave radiation towards the surface and blocking incoming shortwave radiation. Here we present our model results of cloud presence and surface radiation, with a focus on north and east Greenland during summertime, when runoff rates are typically largest. To inter-compare meteorological and surface conditions across the different CESM simulations, output from the atmosphere model CAM has been bilinearly regridded to a common 0.25° mesh. The anomalies thus computed (Figure 9) should be interpreted with caution since they incorporate

interpolations errors. Nevertheless, ~~these benchmark data are also subject to uncertainty. Noël et al. (2018) characterise the native spatial resolution of 11 km as a source of model uncertainty, as well as the representation of surface roughness and surface albedo. Two prime uncertainties in the RACMO2 downscaling procedure arise from the bare ice albedo used to correct runoff, and the ice sheet extent (Noël et al., 2016)~~these spatial anomaly plots may provide us with useful qualitative leads on some of the first order changes taking place in the VR refined simulations.

~~Finally, sublimation is enhanced in both VR runs compared~~ Figure 9a shows VR-CESM surface elevation anomalies with respect to Uniform CESM~~(Table 2) which is attributed~~ . The patterns are very similar in both VR-CESM55 and VR-CESM28, with lower surface topography over the ocean and near the margins of the island, and with higher surface elevations inland. As discussed previously, these features are explained by the smoothing operator applied by CAM to the topography, the imprint of which is much wider at low resolution than it is at high resolution (cf. Figure 2). Due to the higher ~~wind speeds that are present in the VR-CESM simulations (not shown). In VR-CESM28, the increase relative to Uniform CESM is largest and exceeds 50 %. Still, sublimation is a relatively minor term in the surface mass budget of the GrIS.~~

~~To conclude, the GrIS integrated SMB is consistently overestimated in all CESM simulations relative to the downscaled RACMO2 data. As discussed above, there are different leading terms underpinning this bias in each simulation. In Uniform CESM and VR-CESM55, the SMB component with the largest deviation from the reference data is precipitation, with comparably minor biases in runoff and sublimation. Precipitation is better resolved in VR-CESM28 which therefore compares more favourably to RACMO2 in a GrIS-integrated sense. In this simulation, however, runoff is more strongly biased than in the other simulations~~ and steeper terrain near the margins, orographic uplift and condensation are enhanced leading to increased cloud water path (CWP), ~~such that it is now the SMB component with the largest absolute deviation. The basin analysis indicated that runoff is particularly biased in the north and north-eastern regions of the GrIS.Out of~~ the ~~three CESM simulations, VR-CESM55 has an integrated SMB value most similar to RACMO2, with a positive bias of 120 ± 143 Gt (Table 2)~~vertically integrated mass of liquid water and solid ice contained in clouds. Higher on the ice sheet, CWP decreases, as illustrated by Figure 9b. Some exceptions to this pattern exist, e.g. in north-east Greenland where CWP is locally reduced over the margin and ocean as well. Either changes in meso-scale flow driven by local topography, or increased katabatic surface winds (not shown) are possible explanations for this.

4 ~~Discussion~~

Downwelling longwave radiation is to some extent governed by CWP, which is apparent from Figure 9c, which resembles closely the cloud patterns (Figure 9b). As a result, negative longwave radiation anomalies are found nearly everywhere across the GrIS, with anomalies exceeding -10 W m^{-2} in the north and the east. Some areas where cloud water path has increased strongly (e.g. the patch far north) are not matched by strong positive longwave anomalies. This may be explained as a saturation effect at this particular location, while the slightly reduced lower troposphere temperature cannot be excluded as well (Figure S4). The reduced thermal flux exhibits a broad correlation to the ice sheet skin temperature (Figure 9d), which indicates a general cooling across the GrIS interior and strong negative anomalies in locations with reduced longwave. Some marked

positive temperature anomalies are seen as well, which we interpret as a better resolved tundra region, with fewer grid cells having mixed land surface types of glacier and bare land. This is relevant since the skin temperature of the "glacier" land surface type never exceeds the freezing point (Oleson, 2013), whereas the tundra land surface type does in summer.

Importantly, meteorological conditions in all of the CESM simulation are inadequate to support a snow-free tundra in north Greenland during summer (Figure S5, Supplementary Material), despite the snow height reset to an extremely low value at the time of initialization (see Section 2.4). Snow buildup over the northern tundras is a known model bias in this version of CESM and our results suggest that this bias worsens, rather than improves, on VR refined grids. This is reflected in the albedo difference maps (Figure 9e) which shows that both VR simulations predict higher JJA albedo than the uniform resolution run. Albedo anomalies exceed 0.1 over much of north Greenland. As a result, significantly less shortwave radiation is absorbed which acts as a positive feedback sustaining the snow cover. In effect, the net radiation budget is negative over the entire north Greenland region (Figure 9f); this radiation deficit is partially offset by an enhanced sensible heat flux in both VR runs (not shown), which is however insufficient to prevent a net energy loss at the surface, i.e. the sum of all surface radiative and turbulent fluxes is negative. Note that albedo is also higher over the adjacent sea ice (Figure 9e). In fact, JJA sea ice albedo is found to increase over the entire Arctic basin (not shown), which appears linked to precipitation phase and frequency.

~~One important and perhaps unexpected finding of the previous section is that melt and runoff generally decrease at higher resolutions, deteriorating the agreement with the – albeit sparse – ablation zone measurements. Here, we aim to identify possible driving mechanisms in an attempt to understand what might be causing this behaviour.~~

~~The first candidate for such a mechanism, naturally, is the redistribution of precipitation~~To summarize, a changed cloud distribution is proposed as the key mechanism that could explain the lower ablation rates in VR-CESM, relative to Uniform CESM. In particular, reduced cloud water paths are found across large stretches in the northern and eastern regions of the GrIS leading to less heating of the surface. As a result, VR-CESM simulates a lower skin temperature in those regions, which is at least partly due to the reduction in longwave radiation, as the two variables broadly correlate. At the ~~margins, local increases in snowfall occur due to the steeper terrain at higher resolution (Figure 2), enhancing orographic uplift and condensation. More winter snow means delayed bare-ice exposure in summer and an increased refreezing capacity in spring and autumn. Moreover, summer snowfall is potentially even more important since it can cover dark bare ice with bright fresh snow and interrupt the snow-albedo-melt feedback (Noël et al., 2015). But precipitation does not explain everything.~~ same time, CESM fails to simulate a seasonal snow cover over the northern tundras in all simulations, thereby reinforcing the cold bias that this region experiences, through albedo. The reduced ablation rates in VR-CESM are explained by the negative net radiation anomalies seen in these simulations.

~~In particular, the comparison with the RACMO2 data indicated that runoff in VR-CESM28 is most markedly underestimated in the northern and north-eastern basins of the GrIS (Figure 8), where precipitation does not change substantially at higher resolution.~~

4 Discussion

All of our CESM runs are simulating a perennial snow cover over the North Greenland tundra, or at least partly (Figure S5, Supplementary Material). In reality, these tundras and the adjacent bare ice zones do experience a seasonal snow cover. This is a delicate balance, however, with little snowfall (cf. Figure 4) and a relatively short melt season, that could tip towards the positive side without much difficulty. Any incidental snow excess delays exposure of the bare land or ice underneath, i.e. triggers the albedo feedback. We now present four possible factors or model weaknesses that could underpin this chain of events, leading up to a perennial snow cover in places where it should not exist. The factors / weaknesses are not mutually exclusive and may even be related (e.g. rain and clouds). The aim of this discussion is to improve understanding of our model results on one hand, and to guide future model development and analysis on the other.

First, in regions with little snowfall and melt, erosion and sublimation by blowing snow are two relatively important SMB components. ~~In this region, low precipitation rates in combination with high wind speeds lead to a high contribution of blowing snow processes to SMB .~~ Therefore, even state-of-the-art regional climate models, such as RACMO2, struggle to capture the SMB of the northern GrIS. For example, blowing snow was overestimated in a previous version of RACMO2 (RACMO 2.3p1), causing a too-wide ablation zone in the north (Noël et al., 2018). Here, CESM simulates an ablation zone that is too narrow in all three simulations (Figure 7), which may be explained by the fact that neither drifting snow redistribution, nor drifting snow sublimation are included in ~~CESM (van Kampenhout et al., 2017)~~CLM (see Equation 1).

4.1 ~~Large-scale circulation~~

~~Another important driving mechanism for runoff changes could be a shift/change in the large-scale atmospheric circulation induced by the resolution changes, leading to lower tropospheric temperatures over Greenland and an associated reduced downwelling longwave flux at the surface. To this end, we compared CESM mid-troposphere geopotential height and temperature across the different resolutions. Circulation anomalies are computed by subtracting the Uniform CESM fields from VR-CESM55 and VR-CESM28, respectively, which is done after regridding to a common mesh. The resulting anomaly maps in 500 hPa summer (JJA) geopotential height (Z500) and temperature (T500) are depicted in Figure ??.~~

~~In both VR-CESM simulations, positive Z500 anomalies (increased heights) are apparent over the north-east Atlantic Ocean and the Bering Strait, with negative anomalies elsewhere in the Arctic. The broad similarity between the anomaly patterns in both VR-CESM runs makes it unlikely that internal variability is responsible for generating these patterns, so we interpret these patterns as actual large-scale circulation changes. This claim is further supported by anomaly fields in the other seasons, which also indicate consistent circulation patterns across both VR-CESM simulations (not shown). Therefore we conclude that these changes are~~ Second, along the same reasoning, incidental rainfall could be playing an important role as well. Summer rainfall events add liquid water to the snow pack, thereby speeding up grain growth and thus lowering snow albedo (Oleson, 2013). Moreover, rain water may be refrozen, thereby releasing latent heat into the snow pack and permanently increasing the snow albedo due to the ~~subsequent refinement of topography over Greenland (Figure 2) and elsewhere within the VR domain (Figure 1). The higher (and steeper) topography promotes atmospheric blocking and leads to disturbances in the near-stationary Rossby / planetary waves and is an example of the upstream / downstream impacts that the VR-CESM approach captures, which may not otherwise be present in offline RCM dynamical downscaling.~~

Over Greenland, decreased geopotential height indicates more cyclonic flow, drawing in cold air from the Arctic ocean. By and large, anomalies in summer T500 are fairly well correlated with anomalies in Z500, which suggests that the changes in T500 are predominantly forced by circulation (Figure **??**). Summer T500 anomalies ranging from -0.2 to -0.6 K are found across Greenland. This indicates that, indeed, large scale circulation changes are likely an important driver of the lower melt and runoff rates found at higher resolutions.

Summer (JJA) anomalies of selected variables. Prior to subtraction, all data have been regridded to a common regular mesh of 0.25° using bi-linear interpolation. The anomaly map for 2-m air temperature (not shown) is similar to that of the skin temperature (panel d).

4.1 Clouds and radiation

Next, the role of cloud cover is investigated. Clouds are an important modulator of the surface energy balance and therefore melt and runoff. Anomalies of summer cloud water path of both VR-CESM simulations have been computed and are displayed in Figure 9a. Cloud water path exhibits negative anomalies – implying that VR-CESM clouds contain less water and ice than their counterparts in the Uniform CESM simulation – in places of reduced precipitation (Figure 9b). Positive anomalies are mainly found over the open ocean and close to the margins, where locally the terrain has lowered due to the better resolved height gradients (Figure 2) . More importantly, negative anomalies are also found across most of the northern and eastern parts of the GrIS. Likely, this is a combination of topographic effects and circulation changes, through which less moisture is brought to the region.

As a result of both the reduced cloud content and the colder ambient troposphere (Figure **??**), downwelling longwave radiation is significantly reduced. In VR-CESM28, anomalies below -10 W m^2 are found across large areas in the north and east (Figure 9larger snow grain size of refrozen grains (set to 1000 μg in this version of CLM, Oleson (2013)) with respect to fresh snow. As illustrated in Figure 10, CAM-SE currently does not produce the >50 mm accumulated rainfall values over the northern tundras that RACMO2 simulates (panels a and d), with the caveat that RACMO2 rainfall rates have not been validated with station data in this region. Still, we deem it unlikely that CAM-SE simulates realistic rainfall (both intensity and frequency) in this area, and rather underpredicts summer rain by a factor of 2 or more (Figure 10). Further, there seems no significant change in north Greenland rainfall rates in both VR simulations. (panels b and c). We argue that this must be the leading cause for the ~1 K skin temperature anomaly found across most of the GrIS (Figure 9d). This temperature decrease is most likely to impact melt in areas that are already cold (the north) and see a relatively short melt season. This is in line with other studies that showed that longwave radiation from clouds is a dominant factor for initiation of melt and extreme melt events in cold regions of the ice sheet (e.g. (Bennartz et al., 2013; Van Tricht et al., 2016b; Cullather and Nowicki, 2018)).

Reduced cloud cover in general also increases downwelling shortwave radiation received by the surface, here shown in Figure 9e. The magnitude of the insolation anomaly is comparable to that of the longwave flux, but it is still modulated by the high surface albedo. Therefore, we conclude that the net effect of all incoming radiation on the surface energy balance is negative, except for locations where albedo is low.

4.1 ~~Latent heat~~

~~Higher summer wind speeds in both VR-CESM simulations~~ This ties in with Bacmeister et al. (2014), who remarked that increasing horizontal resolution by itself does not lead to dramatically improved climate simulations, and must be accompanied by new cloud and convection parametrizations. Existing parameterizations in CAM were developed with specific spatial and temporal scales in mind, and contain assumptions that may break down at higher resolutions (Bacmeister et al., 2014). For completeness, it is remarked that in this version of CESM, precipitation is phase repartitioned by the land component CLM based on the atmospheric surface temperature[2]. We find that summer rainfall rates after repartitioning are markedly higher over specific ablation areas / glaciers, but not over the northern tundra region (not shown)~~enhance sublimation along the GrIS margins, which acts as a surface energy sink in the ablation zone (Figure 9f). In response, less energy is available for warming the surface through sensible heat and surface melting~~. The phase repartitioning model mechanic is therefore not sufficient in compensating for a possible rain bias during summertime.

4.1 ~~Subgrid downscaling~~

~~The multiple elevation classes method introduced by Lipscomb et al. (2013) as a means of subgrid downscaling (see Section 2.3) gradually becomes less important when resolution is increased, since the topographic variability within a single grid cell is smaller. We hypothesise that a bias in the lapse rates used in the meteorological downscaling has the potential to bias the CESM runoff results, especially at coarser resolutions.~~

The third factor, cloud physics, is similar in origin to the previous factor of rainfall. As mentioned earlier, current CAM cloud physics have been developed with specific spatial and temporal resolutions in mind, and are not necessarily performing better on a decreased grid spacing and a smaller physics time step. Moreover, a recent study indicated that CAM5 simulates insufficient cloud cover in summer, especially non-opaque liquid-containing clouds that have a strongly positive cloud radiative effect (Lacour et al., 2018). We believe that summer cloud phase, frequency, and opacity are important metrics to consider in future studies that are set out to resolve the north-Greenland permanent snow bias.

4.1 ~~Summary~~

~~Several potential driving mechanisms have been proposed that each may explain the reduced melt and runoff fluxes found at higher CESM resolutions . Since some of these processes are intertwined, e.g. the role of cloud cover and tropospheric temperature on downwelling longwave radiation, based on the available data it is not possible to quantify their individual contributions. Only a full reconstruction of the surface energy balance in monitored sites could shed further light on that. However, this is deemed outside of the scope of the current study, which is intended as an exploratory first stepin applying VR-CESM to Greenland.~~

[2] CLM5 Technical Note, https://escomp.github.io/ctsm-docs/doc/build/html/tech_note/Ecosystem/CLM50_Tech_Note_Ecosystem.html

5 ~~Limitations~~

~~Unfortunately, a mistake was found in the land cover input file that was used in both VR-CESM simulations, but not in the Uniform CESM run. The faulty input files led to the unphysical release of sensible heat into the lower atmosphere during bare ice melt conditions. The additional amount of sensible heat released equals the melt energy that was generated in the same time step. It is noted that the magnitude of the erroneous energy flux is equivalent to a similar flux common over non-glaciated surfaces (i. e. that cannot melt), except that here it is released as sensible heat rather than e.g. outgoing longwave radiation. Since our results suggest a lack of bare ice melt in VR-CESM, rather than an excess, this error does not impact the main conclusions of this study.~~

~~Little attention has been given in this study to the degree in which internal variability affects Greenland SMB. Compared to RCMs, which are both forced at the domain edge and ocean surface, there is relatively weak "weather" forcing in our atmosphere-only CESM simulations. Arguably,~~ Finally, we propose CLM snow physics as a possible factor or model weakness that could be responsible for reinforcing a positive mass balance. Currently, the CLM snow model operates with a ~~model run of 20 years is not quite long enough to attain a robust mean climate. Therefore, the numbers presented in Table 2 should be used with some caution. We plan on doing a follow-up study where the impact of atmospheric variability on Greenland SMB in AMIP simulations is diagnosed.~~

~~The difference in ice masks between~~ tipping bucket model for liquid water, while maintaining an irreducible water content of 3.3 % (Oleson, 2013). Further, the grain size of refrozen snow is set to 1000 μg. A previous study carried out with RACMO2 ~~and CESM remains another source of uncertainty, especially when calculating GrIS integrated numbers like in Table 2. In follow-up studies it is therefore recommended to use overlapping masks, or to distinguish between the contiguous GrIS and peripheral glaciers and ice caps. Unfortunately, neither CLM nor CISM currently provide such a distinction, so a workaround solution may be needed. In the case of runoff , the peripheral contribution in CESM could be estimated by computing the integrated runoff over the different basins (Figure 8). The drawback is that these numbers are calculated using bi-linear interpolation and are not necessarily very accurate~~indicated that both the irreducible water content and the refreezing grain size are sensitive parameters that impact total melt and runoff generated by the model (van Angelen et al., 2012). It could at some point be worthwhile to explore this avenue of model development in future studies that target Greenland, especially since they are simple parameters to adjust.

5 Summary and Conclusions

For the first time, regionally refined GCM simulations using VR-CESM have been performed at 55 ~~&~~ and 28 km over the greater Greenland region to study the impact of spatial resolution on GrIS SMB. Compared to a uniform resolution (1° or ∼111 km) control run, topography is resolved with greater fidelity, leading to improved patterns in orographic precipitation, most notably in southern Greenland and along the western and eastern margins. At the same time, a general drying in the GrIS interior occurs, which ~~significantly~~ substantially improves correlations to IceBridge accumulation radar and in-situ measurements of accumulation. Arguably, ~~VR-CESM28 now~~ VR-CESM performs on-par with RCMs in ~~these regions (Lewis et al., 2016). In~~

transient simulations, the reproducing these observations, especially at 28 km. GrIS integrated precipitation is reduced from 893 to 745 Gt in VR-CESM28, which is within 6% of a best-estimate RCM figure (707 Gt). The improved distribution of snowfall accumulation may prove pivotal as it in transient simulations, as snowfall modulates the timing and strength of the snow-albedo feedback (Picard et al., 2012) and impacts ice advection. GrIS integrated precipitation is reduced to 806 Gt in VR-CESM28, which is more in line with best-estimate RCM figures.

In the ablation zone, the CESM simulations were evaluated using geographically sparse in-situ measurements. Despite its coarse resolution of ∼111 km, we found that Uniform CESM reproduces these measurements to a reasonable degree, which due to the complexity of ablation zone could be a fortuitous combination of compensating biases, but may also indicate the effectiveness of multiple elevation classes in CESM (Lipscomb et al., 2013). By comparison, represents a positive result for CESM at low resolution, suggesting that the subgrid elevation classes are effective (Section 2.3). In both VR-CESM simulationsshow an increasingly large , a positive SMB bias , indicating developed in the ablation zone, which signalled an unexpected regression (too little ablation. This is reflected in their GrIS-wide integrated melt and runoffluxes, which are). This was reflected in GrIS-integrated runoff, which was found to be substantially lower in VR-CESM55 and VR-CESM28 than in compared to Uniform CESM. A basins-by-basin analysis revealed that the largest reductions in runoff are found in the northern and eastern basins.

Several driving mechanisms were explored that could be underpinning these reductions. It was shown that the topographic changes on Greenland and in the rest of the domain causes , with a good agreement in the other basins. Likely, large-scale circulation changes which, as a result, led to lower 500 hPa summer temperatures over the GrIS. Combined with local reductions in cloud cover, this affects can be excluded as the prime driving mechanism, as T500 and T700 difference maps indicated no significant cooling over these areas in summer. Instead, we linked the reduction in runoff to clouds, as clouds impact the radiative budget at the surface , in particular the downwelling longwave flux, leading up to a decrease in skin temperature by 1–2 K. Enhanced sublimation by higher wind speeds at increased resolution are hypothesised to play a role as well, drawing away energy from the surface energy budgetthrough the emission of longwave radiation. We highlighted the fact that downwelling longwave radiation is reduced in both VR-CESM simulations, in particular over the regions where runoff decreased with resolution.

Several possible avenues for future research are suggested; (1) with a focus on large-scale circulation, one could compare tropospheric conditions in VR-CESM to reanalyses and alleviate potential biases by tuning relevant atmospheric parameters; (2) one could explore yet higher VR resolutions, like 14 or 7 km Rhoades et al. (2018), and study the impact on Greenland SMB; (3) drifting snow sublimation and redistribution could be implemented and evaluated over the GrIS using VR-CESM; (4) as of present it is not possible to run VR-CESM with an active ocean model. Still, projections on future GrIS SMB could be generated using VR-CESM when high-frequency output from fully-coupled scenario simulations are used as boundary conditions to VR-CESMAt the same time, we could not ignore the outstanding bias in CESM of permanent snow cover over north Greenland peripheral glaciers and tundras. This bias was found to worsen in VR-CESM, rather than improve, leading to more positive surface albedo and a more negative net radiative budget, compared to Uniform CESM. We highlighted four candidate model weaknesses that could be underpinning the bias. With these four factors — blowing snow, rainfall, clouds,

snow physics — we provided directions for future studies that are concerned with Greenland. We strongly believe that for GrIS SMB studies, addressing these model weaknesses will prove more important than e.g. increasing the horizontal resolution further, as is done in other studies (Rhoades et al., 2018).

To conclude, our ~~study shows~~ case study demonstrates that VR-CESM is a ~~viable~~ promising technique for dynamically downscaling GCM climate simulations over an Arctic region, while maintaining model consistency and allowing for feedbacks between the region of interest and the rest of the globe. A finer resolution leads to better resolved storms that are taking different pathways than their low-resolution counterparts, and therefore change precipitation and cloud cover patterns on a local scale. VR-CESM can serve as a tool for modellers that are interested in the dynamical response of the GrIS to future SMB changes~~since~~ , at a reasonable computational cost. At the time this manuscript was written, it was not possible to run VR-CESM ~~can be coupled to ice sheet models like CISM.~~ in coupled mode with an active ocean model. Still, high-resolution future projections of GrIS SMB could be generated using VR-CESM when high-frequency output from a fully-coupled scenario simulation is used as a boundary conditions at the sea surface.

Data availability. Climate model data used in our analysis are available on Zenodo, https://doi.org/10.5281/zenodo.2579606

Author contributions. LvK and MRB originally conceived the study. The simulations were set up and carried out by LvK and AMR with help from ARH and CMZ. LvK led the analysis and the writing of the manuscript, with contributions from all the other authors.

Competing interests. The authors declare no competing interests.

Acknowledgements. ~~The authors thank~~

The authors would like to thank the editor, Xavier Fettweis, and two anonymous referees for their constructive comments. Further, we had fruitful discussions with Paul Ullrich (UC Davis) and Aarnout van Delden (UU/IMAU)~~for useful discussions and~~ , and the RACMO2 data was kindly provided by Brice Noël (UU/IMAU)~~for sharing the RACMO2 data used~~. This work was carried out under the program of the Netherlands Earth System Science Centre (NESSC), financially supported by the Ministry of Education, Culture and Science (OCW, Grantnr. 024.002.001). Author Alan Rhoades was funded by the U.S. Department of Energy, Office of Science "An Integrated Evaluation of the Simulated Hydroclimate System of the Continental US" project (award no. DE-SC0016605).

References

Bacmeister, J. T., Wehner, M. F., Neale, R. B., Gettelman, A., Hannay, C., Lauritzen, P. H., Caron, J. M., and Truesdale, J. E.: Exploratory High-Resolution Climate Simulations Using the Community Atmosphere Model (CAM), J. Climate, 27, 3073–3099, https://doi.org/10.1175/JCLI-D-13-00387.1, 2014.

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: Atmospheres, 114, https://doi.org/10.1029/2008JD011208, 2009.

Beljaars, A. C. M., Brown, A. R., and Wood, N.: A New Parametrization of Turbulent Orographic Form Drag, Q.J.R. Meteorol. Soc., 130, 1327–1347, https://doi.org/10.1256/qj.03.73, 2004.

Bennartz, R., Shupe, M. D., Turner, D. D., Walden, V. P., Steffen, K., Cox, C. J., Kulie, M. S., Miller, N. B., and Pettersen, C.: July 2012 Greenland Melt Extent Enhanced by Low-Level Liquid Clouds, Nature, 496, 83–86, https://doi.org/10.1038/nature12002, 2013.

Citterio, M. and Ahlstrøm, A. P.: Brief Communication "The Aerophotogrammetric Map of Greenland Ice Masses", The Cryosphere, 7, 445–449, https://doi.org/10.5194/tc-7-445-2013, 2013.

Cogley, J. G.: Greenland Accumulation: An Error Model, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/10.1029/2003JD004449, 2004.

Cullather, R. I. and Nowicki, S. M. J.: Greenland Ice Sheet Surface Melt and Its Relation to Daily Atmospheric Conditions, Journal of Climate, 31, 1897–1919, https://doi.org/10.1175/JCLI-D-17-0447.1, 2018.

Cullather, R. I., Nowicki, S. M. J., Zhao, B., and Suarez, M. J.: Evaluation of the Surface Representation of the Greenland Ice Sheet in a General Circulation Model, Journal of Climate, 27, 4835–4856, https://doi.org/10.1175/JCLI-D-13-00635.1, 2014.

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., Hólm, 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, Q.J.R. Meteorol. Soc., 137, 553–597, https://doi.org/10.1002/qj.828, 2011.

Delworth, T. L., Rosati, A., Anderson, W., Adcroft, A. J., Balaji, V., Benson, R., Dixon, K., Griffies, S. M., Lee, H.-C., Pacanowski, R. C., Vecchi, G. A., Wittenberg, A. T., Zeng, F., and Zhang, R.: Simulated Climate and Climate Change in the GFDL CM2.5 High-Resolution Coupled Climate Model, J. Climate, 25, 2755–2781, https://doi.org/10.1175/JCLI-D-11-00316.1, 2011.

Dennis, J. M., Edwards, J., Evans, K. J., Guba, O., Lauritzen, P. H., Mirin, A. A., St-Cyr, A., Taylor, M. A., and Worley, P. H.: CAM-SE: A Scalable Spectral Element Dynamical Core for the Community Atmosphere Model, The International Journal of High Performance Computing Applications, 26, 74–89, https://doi.org/10.1177/1094342011428142, 2012.

Evans, K. J., Lauritzen, P. H., Mishra, S. K., Neale, R. B., Taylor, M. A., and Tribbia, J. J.: AMIP Simulation with the CAM4 Spectral Element Dynamical Core, J. Climate, 26, 689–709, https://doi.org/10.1175/JCLI-D-11-00448.1, 2012.

Evans, K. J., Kennedy, J. H., Lu, D., Forrester, M. M., Price, S., Fyke, J., Bennett, A. R., Hoffman, M. J., Tezaur, I., Zender, C. S., and Vizcaíno, M.: LIVVkit 2.1: Automated and Extensible Ice Sheet Model Validation, Geoscientific Model Development Discussions, pp. 1–31, https://doi.org/10.5194/gmd-2018-70, 2018.

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, Geoscientific Model Development, 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, 2013a.

Fettweis, X., Hanna, E., Lang, C., Belleflamme, A., Erpicum, M., and Gallée, H.: Brief Communication "Important Role of the Mid-Tropospheric Atmospheric Circulation in the Recent Surface Melt Increase over the Greenland Ice Sheet", The Cryosphere, 7, 241–248, https://doi.org/10.5194/tc-7-241-2013, 2013b.

Fischer, R., Nowicki, S., Kelley, M., and Schmidt, G.: A System of Conservative Regridding for Ice-Atmosphere Coupling in a General Circulation Model (GCM), Geoscientific Model Development, 7, p. 883–907, https://doi.org/10.5194/gmd-7-883-2014, 2014.

Flanner, M. G. and Zender, C. S.: Snowpack Radiative Heating: Influence on Tibetan Plateau Climate, Geophysical Research Letters, 32, https://doi.org/10.1029/2004GL022076, 2005.

Gates, W. L., Boyle, J. S., Covey, C., Dease, C. G., Doutriaux, C. M., Drach, R. S., Fiorino, M., Gleckler, P. J., Hnilo, J. J., Marlais, S. M., Phillips, T. J., Potter, G. L., Santer, B. D., Sperber, K. R., Taylor, K. E., and Williams, D. N.: An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP I), Bull. Amer. Meteor. Soc., 80, 29–56, https://doi.org/10.1175/1520-0477(1999)080<0029:AOOTRO>2.0.CO;2, 1999.

Gettelman, A. and Morrison, H.: Advanced Two-Moment Bulk Microphysics for Global Models. Part I: Off-Line Tests and Comparison with Other Schemes, J. Climate, 28, 1268–1287, https://doi.org/10.1175/JCLI-D-14-00102.1, 2014.

Gettelman, A., Callaghan, P., Larson, V. E., Zarzycki, C. M., Bacmeister, J., Lauritzen, P. H., Bogenschutz, P. A., and Neale, R.: Regional Climate Simulations With the Community Earth System Model, Journal of Advances in Modeling Earth Systems, https://doi.org/10.1002/2017MS001227, 2018.

Guba, O., Taylor, M. A., Ullrich, P. A., Overfelt, J. R., and Levy, M. N.: The Spectral Element Method (SEM) on Variable-Resolution Grids: Evaluating Grid Sensitivity and Resolution-Aware Numerical Viscosity, Geosci. Model Dev., 7, 2803–2816, https://doi.org/10.5194/gmd-7-2803-2014, 2014.

Helsen, M. M., van de Wal, R. S. W., Reerink, T. J., Bintanja, R., Madsen, M. S., Yang, S., Li, Q., and Zhang, Q.: On the Importance of the Albedo Parameterization for the Mass Balance of the Greenland Ice Sheet in EC-Earth, The Cryosphere, 11, 1949–1965, https://doi.org/10.5194/tc-11-1949-2017, 2017.

Herrington, A. and Reed, K.: An Idealized Test of the Response of the Community Atmosphere Model to Near-Grid-Scale Forcing Across Hydrostatic Resolutions, Journal of Advances in Modeling Earth Systems, 10, 560–575, https://doi.org/10.1002/2017MS001078, 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.

Huang, X., Rhoades, A. M., Ullrich, P. A., and Zarzycki, C. M.: An Evaluation of the Variable-Resolution CESM for Modeling California's Climate, J. Adv. Model. Earth Syst., 8, 345–369, https://doi.org/10.1002/2015MS000559, 2016.

Hurrell, J. W., Hack, J. J., Shea, D., Caron, J. M., and Rosinski, J.: A New Sea Surface Temperature and Sea Ice Boundary Dataset for the Community Atmosphere Model, J. Climate, 21, 5145–5153, https://doi.org/10.1175/2008JCLI2292.1, 2008.

Kennedy, J. H., Bennett, A. R., Evans, K. J., Price, S., Hoffman, M., Lipscomb, W. H., Fyke, J., Vargo, L., Boghozian, A., Norman, M., and Worley, P. H.: LIVVkit: An Extensible, Python-Based, Land Ice Verification and Validation Toolkit for Ice Sheet Models, Journal of Advances in Modeling Earth Systems, 9, 854–869, https://doi.org/10.1002/2017MS000916, 2017.

Lacour, A., Chepfer, H., Miller, N. B., Shupe, M. D., Noel, V., Fettweis, X., Gallee, H., Kay, J. E., Guzman, R., and Cole, J.: How Well Are Clouds Simulated over Greenland in Climate Models? Consequences for the Surface Cloud Radiative Effect over the Ice Sheet, J. Climate, 31, 9293–9312, https://doi.org/10.1175/JCLI-D-18-0023.1, 2018.

Lauritzen, P. H., Nair, R. D., Herrington, A. R., Callaghan, P., Goldhaber, S., Dennis, J. M., Bacmeister, J. T., Eaton, B. E., Zarzycki, C. M., Taylor, M. A., Ullrich, P. A., Dubos, T., Gettelman, A., Neale, R. B., Dobbins, B., Reed, K. A., Hannay, C., Medeiros, B., Benedict, J. J., and Tribbia, J. J.: NCAR Release of CAM-SE in CESM2.0: A Reformulation of the Spectral-Element Dynamical Core in Dry-Mass Vertical Coordinates with Comprehensive Treatment of Condensates and Energy, Journal of Advances in Modeling Earth Systems, 0, https://doi.org/10.1029/2017MS001257, 2018.

Lefebre, F., Fettweis, X., Gallée, H., Ypersele, J.-P. V., Marbaix, P., Greuell, W., and Calanca, P.: Evaluation of a High-Resolution Regional Climate Simulation over Greenland, Climate Dynamics, 25, 99–116, https://doi.org/10.1007/s00382-005-0005-8, 2005.

Leguy, G., Lipscomb, W. H., and Sacks, W. J.: CESM Land Ice Documentation and User Guide, https://escomp.github.io/cism-docs/cism-in-cesm/doc_work/html/, 2018.

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, 2016.

Lewis, G., Osterberg, E., Hawley, R., Whitmore, B., and Marshall, H. P.: Regional Greenland Accumulation Variability from Operation IceBridge Airborne Accumulation Radar, The Cryosphere Discussions, pp. 1–29, https://doi.org/10.5194/tc-2016-248, 2016.

Ligtenberg, S. R. M., Kuipers Munneke, P., Noël, B. P. Y., 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., Fyke, J. G., Vizcaíno, M., Sacks, W. J., Wolfe, J., Vertenstein, M., Craig, A., Kluzek, E., and Lawrence, D. M.: Implementation and Initial Evaluation of the Glimmer Community Ice Sheet Model in the Community Earth System Model, Journal of Climate, 26, 7352–7371, https://doi.org/10.1175/JCLI-D-12-00557.1, 2013.

Machguth, H., Thomsen, H. H., Weidick, A., Ahlstrøm, A. P., Abermann, J., Andersen, M. L., Andersen, S. B., Bjørk, A. A., Box, J. E., Braithwaite, R. J., Bøggild, C. E., Citterio, M., Clement, P., Colgan, W., Fausto, R. S., Gleie, K., Gubler, S., Hasholt, B., Hynek, B., Knudsen, N. T., Larsen, S. H., Mernild, S. H., Oerlemans, J., Oerter, H., Olesen, O. B., Smeets, C. J. P. P., Steffen, K., Stober, M., Sugiyama, S., As, D. V., Broeke, M. R. V. D., and Wal, R. S. W. V. D.: Greenland Surface Mass-Balance Observations from the Ice-Sheet Ablation Area and Local Glaciers, Journal of Glaciology, 62, 861–887, https://doi.org/10.1017/jog.2016.75, 2016.

Matte, D., Laprise, R., Thériault, J. M., and Lucas-Picher, P.: Spatial Spin-up of Fine Scales in a Regional Climate Model Simulation Driven by Low-Resolution Boundary Conditions, Clim Dyn, 49, 563–574, https://doi.org/10.1007/s00382-016-3358-2, 2017.

Molod, A., Takacs, L., Suarez, M., and Bacmeister, J.: Development of the GEOS-5 Atmospheric General Circulation Model: Evolution from MERRA to MERRA2, Geosci. Model Dev., 8, 1339–1356, https://doi.org/10.5194/gmd-8-1339-2015, 2015.

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, p. 12, 2017.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F., and Marotzke, J.: A Higher-

Resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR), Journal of Advances in Modeling Earth Systems, 0, https://doi.org/10.1029/2017MS001217, 2018.

Neale, R. B., Chen, C.-C., Gettelman, A., Lauritzen, P. H., Park, S., Williamson, D. L., Conley, A. J., Garcia, R., Kinnison, D., and Lamarque, J.-F.: Description of the NCAR Community Atmosphere Model (CAM 5.0), NCAR Tech. Note NCAR/TN-486+ STR, 2012.

Noël, B., van de Berg, W. J., van Meijgaard, E., Kuipers Munneke, P., van de Wal, R. S. W., and van den Broeke, M. R.: 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., van Wessem, J. M., van Meijgaard, E., van As, D., Lenaerts, J. T. M., Lhermitte, S., Kuipers Munneke, P., 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.

Oleson, K. W.: Technical Description of Version 4.5 of the Community Land Model (CLM), NCAR Technical Note NCAR/TN-503+ STR, National Center for Atmospheric Research, Boulder, CO, 2013.

Picard, G., Domine, F., Krinner, G., Arnaud, L., and Lefebvre, E.: Inhibition of the Positive Snow-Albedo Feedback by Precipitation in Interior Antarctica, Nature Climate Change, 2, 795–798, https://doi.org/10.1038/nclimate1590, 2012.

Pollard, D.: A Retrospective Look at Coupled Ice Sheet–Climate Modeling, Climatic Change, 100, 173–194, https://doi.org/10.1007/s10584-010-9830-9, 2010.

Punge, H. J., Gallée, H., Kageyama, M., and Krinner, G.: Modelling Snow Accumulation on Greenland in Eemian, Glacial Inception, and Modern Climates in a GCM, Climate of the Past, 8, 1801– 1819, https://doi.org/10.5194/cp-8-1801-2012, 2012.

Rae, J. G. L., Aðalgeirsdóttir, G., Edwards, T. L., Fettweis, X., Gregory, J. M., Hewitt, H. T., Lowe, J. A., Lucas-Picher, b., 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.

Rhoades, A. M., Huang, X., Ullrich, P. A., and Zarzycki, C. M.: Characterizing Sierra Nevada Snowpack Using Variable-Resolution CESM, J. Appl. Meteor. Climatol., 55, 173–196, https://doi.org/10.1175/JAMC-D-15-0156.1, 2015.

Rhoades, A. M., Ullrich, P. A., and Zarzycki, C. M.: Projecting 21st Century Snowpack Trends in Western USA Mountains Using Variable-Resolution CESM, Clim Dyn, pp. 1–28, https://doi.org/10.1007/s00382-017-3606-0, 2017.

Rhoades, A. M., Ullrich, P. A., Zarzycki, C. M., Johansen, H., Margulis, S. A., Morrison, H., Xu, Z., and Collins, W.: Sensitivity of Mountain Hydroclimate Simulations in Variable-Resolution CESM to Microphysics and Horizontal Resolution, Journal of Advances in Modeling Earth Systems, https://doi.org/10.1029/2018MS001326, 2018.

Rignot, E. and Mouginot, J.: Ice Flow in Greenland for the International Polar Year 2008–2009, Geophysical Research Letters, 39, https://doi.org/10.1029/2012GL051634, 2012.

Shannon, S., Smith, R., Wiltshire, A., Payne, T., Huss, M., Betts, R., Caesar, J., Koutroulis, A., Jones, D., and Harrison, S.: Global Glacier Volume Projections under High-End Climate Change Scenarios, The Cryosphere, 13, 325–350, https://doi.org/https://doi.org/10.5194/tc-13-325-2019, 2019.

Small, R. J., Bacmeister, J., Bailey, D., Baker, A., Bishop, S., Bryan, F., Caron, J., Dennis, J., Gent, P., Hsu, H.-m., Jochum, M., Lawrence, D., Muñoz, E., diNezio, P., Scheitlin, T., Tomas, R., Tribbia, J., Tseng, Y.-h., and Vertenstein, M.: A New Synoptic Scale Resolving Global Climate Simulation Using the Community Earth System Model, Journal of Advances in Modeling Earth Systems, 6, 1065–1094, https://doi.org/10.1002/2014MS000363, 2014.

Sodemann, H., Schwierz, C., and Wernli, H.: Interannual Variability of Greenland Winter Precipitation Sources: Lagrangian Moisture Diagnostic and North Atlantic Oscillation Influence, Journal of Geophysical Research: Atmospheres, 113, https://doi.org/10.1029/2007JD008503, 2008.

Swenson, S. C. and Lawrence, D. M.: A New Fractional Snow-Covered Area Parameterization for the Community Land Model and Its Effect on the Surface Energy Balance, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2012JD018178, 2012.

Ullrich, P. A.: SQuadGen: Spherical Quadrilateral Grid Generator, 2014.

van Angelen, J. H., Lenaerts, J. T. M., Lhermitte, S., Fettweis, X., Kuipers Munneke, P., van den Broeke, M. R., van Meijgaard, E., and Smeets, C. J. P. P.: Sensitivity of Greenland Ice Sheet Surface Mass Balance to Surface Albedo Parameterization: A Study with a Regional Climate Model, The Cryosphere, 6, 1175–1186, https://doi.org/10.5194/tc-6-1175-2012, 2012.

van Angelen, J. H., M. Lenaerts, J. T., van den Broeke, M. R., Fettweis, X., and van Meijgaard, E.: Rapid Loss of Firn Pore Space Accelerates 21st Century Greenland Mass Loss, Geophys. Res. Lett., 40, 2109–2113, https://doi.org/10.1002/grl.50490, 2013.

van den Broeke, M. R., Enderlin, E. M., Howat, I. M., Kuipers Munneke, P., Noël, B. P. Y., 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., Lenaerts, J. T. M., Lipscomb, W. H., Sacks, W. J., Lawrence, D. M., Slater, A. G., and van den Broeke, M. R.: Improving the Representation of Polar Snow and Firn in the Community Earth System Model, J. Adv. Model. Earth Syst., 9, 2583–2600, https://doi.org/10.1002/2017MS000988, 2017.

Van Tricht, K., Lhermitte, S., Gorodetskaya, I. V., and van Lipzig, N. P. M.: Improving Satellite-Retrieved Surface Radiative Fluxes in Polar Regions Using a Smart Sampling Approach, The Cryosphere, 10, 2379–2397, https://doi.org/10.5194/tc-10-2379-2016, 2016a.

Van Tricht, K., Lhermitte, S., Lenaerts, J. T. M., Gorodetskaya, I. V., L'Ecuyer, T. S., Noël, B., van den Broeke, M. R., Turner, D. D., and van Lipzig, N. P. M.: Clouds Enhance Greenland Ice Sheet Meltwater Runoff, Nature Communications, 7, 10 266, https://doi.org/10.1038/ncomms10266, 2016b.

Vizcaíno, M., Lipscomb, W. H., Sacks, W. J., van Angelen, J. H., Wouters, B., and van den Broeke, M. R.: Greenland Surface Mass Balance as Simulated by the Community Earth System Model. Part I: Model Evaluation and 1850–2005 Results, Journal of Climate, 26, 7793–7812, https://doi.org/10.1175/JCLI-D-12-00615.1, 2013.

Wehner, M. F., Reed, K. A., Li, F., Prabhat, Bacmeister, J., Chen, C.-T., Paciorek, C., Gleckler, P. J., Sperber, K. R., Collins, W. D., Gettelman, A., and Jablonowski, C.: The Effect of Horizontal Resolution on Simulation Quality in the Community Atmospheric Model, CAM5.1, J. Adv. Model. Earth Syst., 6, 980–997, https://doi.org/10.1002/2013MS000276, 2014.

Zarzycki, C. M. and Jablonowski, C.: A Multidecadal Simulation of Atlantic Tropical Cyclones Using a Variable-Resolution Global Atmospheric General Circulation Model, J. Adv. Model. Earth Syst., 6, 805–828, https://doi.org/10.1002/2014MS000352, 2014.

Zarzycki, C. M., Levy, M. N., Jablonowski, C., Overfelt, J. R., Taylor, M. A., and Ullrich, P. A.: Aquaplanet Experiments Using CAM's Variable-Resolution Dynamical Core, Journal of Climate, 27, 5481–5503, https://doi.org/10.1175/JCLI-D-14-00004.1, 2014.

Zarzycki, C. M., Jablonowski, C., Thatcher, D. R., and Taylor, M. A.: Effects of Localized Grid Refinement on the General Circulation and Climatology in the Community Atmosphere Model, J. Climate, 28, 2777–2803, https://doi.org/10.1175/JCLI-D-14-00599.1, 2015.

Ziemen, F., Rodehacke, C., and Mikolajewicz, U.: Coupled Ice Sheet–Climate Modeling under Glacial and Pre-Industrial Boundary Conditions, Climate of the Past, 10, 1817–1836, https://doi.org/10.5194/cp-10-1817-2014, 2014.

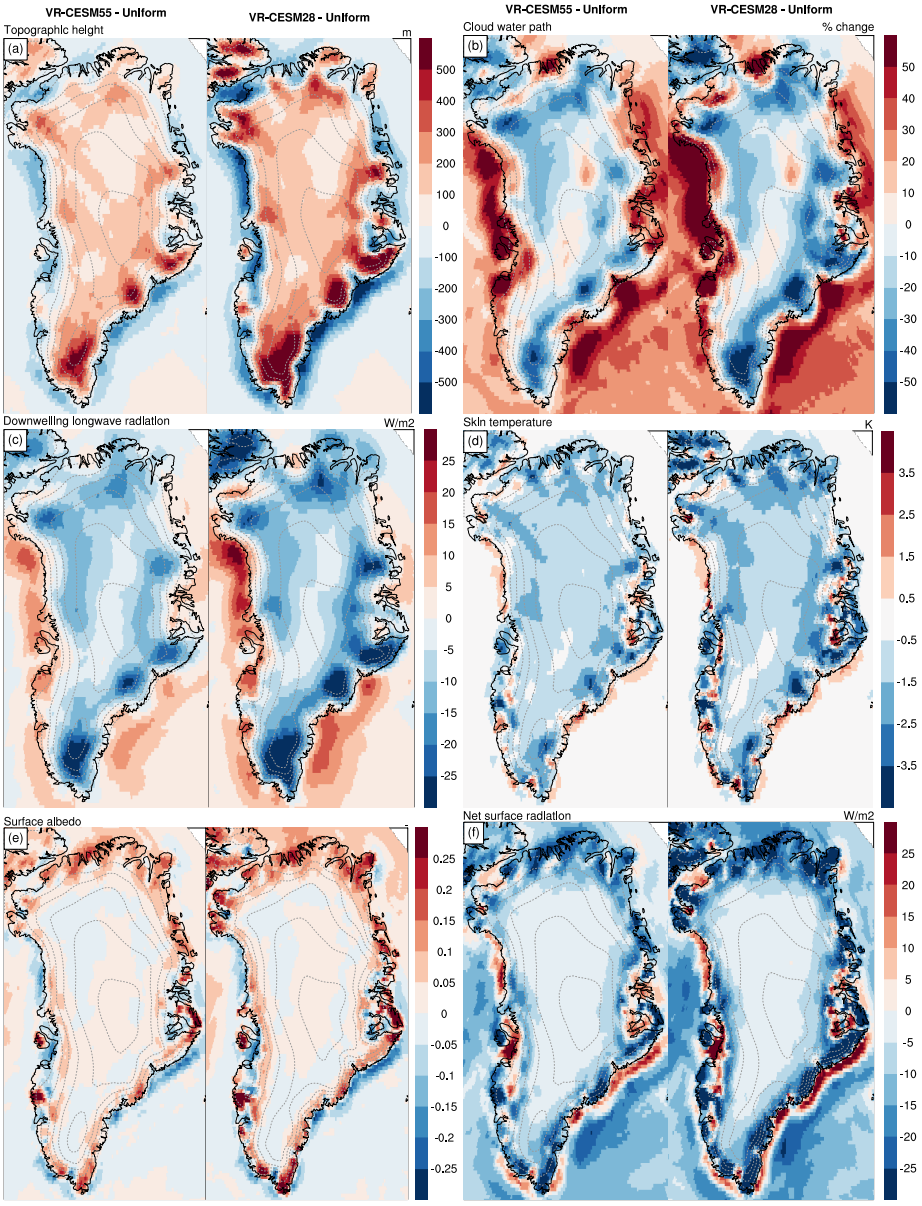

Figure 9. ~~Basin-wide runoff for~~ Summer (JJA) anomalies of atmospheric variables over the period 1980-1999~~. Error bars represent one standard deviation. CESM data have been bi-linearly interpolated down from their native resolution~~ , relative to ~~4 km while making using~~ the ~~vertical elevation class output~~coarse resolution reference simulation (Uniform CESM). ~~The~~ Panel (~~already downscaled~~a) ~~RACMO2 data have been processed~~ topographic height [m], (b) cloud water path [% change], (c) downwelling longwave radiation at ~~their native 1 km resolution. The total runoff across all basins differs from~~ the ~~value reported in Table 2 due to differences in ice extent~~surface [W m^{-2}], ~~since the basins only cover the contiguous ice sheet (total area 1~~d) skin temperature [K], ~~700~~(e) surface albedo [fraction], ~~772 km^2~~(f) net radiation at the surface [W m^{-2}]. Prior to subtraction, all data have been regridded to a common regular mesh of 0.25° using bi-linear interpolation.

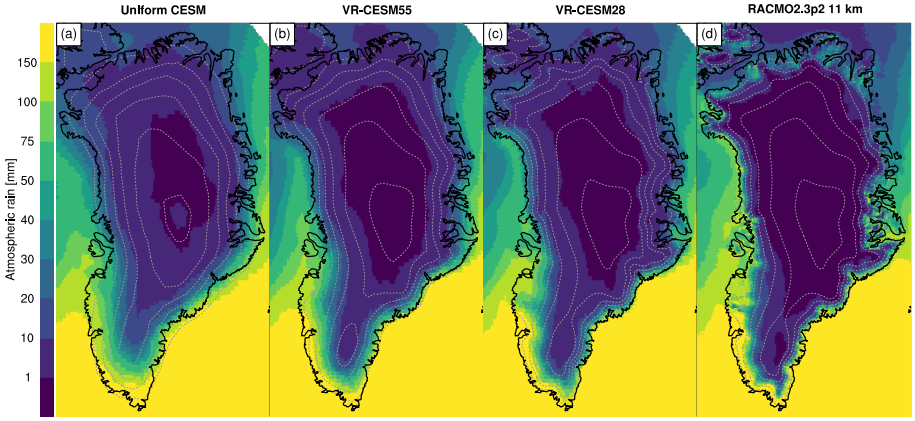

Figure 10. ~~Changes~~ Mean JJA rainfall over the period 1980-1999 in ~~large-scale circulation and mid-troposphere temperature between VR-CESM and uniform resolution CESM~~mm as simulated by CAM-SE. ~~Solid lines represent Z500 from the Uniform~~ RACMO2 data is shown for reference. All three CESM ~~run, dashed lines represent VR-CESM Z500, both drawn at 20 m intervals~~simulations underpredict rain across North Greenland. ~~Prior to subtraction, all~~ CESM data have been regridded to a common regular mesh of 0.25° using bi-linear interpolation. Note the non-linear colour scale.