

1      **Impact of grids and dynamical cores in CESM2.2 on**  
2      **the surface mass balance of the Greenland Ice Sheet**

3      **Adam R. Herrington<sup>1</sup>, Peter H. Lauritzen<sup>1</sup>, William H. Lipscomb<sup>1</sup>, Marcus**  
4      **Lofverstrom<sup>2</sup> and Andrew Gettelman<sup>1</sup>**

5      <sup>1</sup>National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, Colorado, USA

6      <sup>2</sup>Department of Geosciences, University of Arizona, 1040 E. 4th Street, Tucson, AZ USA

7      **Key Points:**

- 8      • The transition from latitude-longitude to unstructured grids in CESM2.2 adversely  
9      impacts the surface mass balance of the Greenland Ice Sheet.
- 10     • Two Arctic refined mesh configurations have been developed and made available  
11     in CESM2.2.
- 12     • The Arctic meshes substantially improves the surface mass balance over all  $1^{\circ}$  grids,  
13     including the latitude-longitude grids.

14 **Abstract**

15 Six different configurations, a mixture of grids and atmospheric dynamical cores  
 16 available in the Community Earth System Model, version 2.2 (CESM2.2), are evaluated  
 17 for their skill in representing the climate of the Arctic and the surface mass balance of  
 18 the Greenland Ice Sheet (GrIS). The more conventional  $1^{\circ}$ – $2^{\circ}$  uniform resolution grids  
 19 systematically overestimate both accumulation and ablation over the GrIS. Of these con-  
 20 ventional grids, the latitude-longitude grids outperform the quasi-uniform unstructured  
 21 grids owing to their greater degrees of freedom in representing the GrIS. Two Arctic-refined  
 22 meshes, with  $\frac{1}{4}^{\circ}$  and  $\frac{1}{8}^{\circ}$  refinement over Greenland, are documented as newly supported  
 23 configurations in CESM2.2. The Arctic meshes substantially improve the simulated clouds  
 24 and precipitation rates in the Arctic. Over Greenland, these meshes skillfully represent  
 25 accumulation and ablation processes, leading to a more realistic GrIS surface mass bal-  
 26 ance. As CESM is in the process of transitioning away from latitude-longitude grids, these  
 27 new Arctic refined meshes improve the representation of polar processes in CESM by re-  
 28 recovering resolution lost in the transition to quasi-uniform grids.

29 **Plain Language Summary**

30 [ enter your Plain Language Summary here or delete this section]

31 **1 Introduction**

32 General Circulation Models (GCMs) are powerful tools for understanding the me-  
 33 teorology and climate of the high latitudes, which are among the most sensitive regions  
 34 on Earth to global and environmental change. GCMs differ vastly in their numerical treat-  
 35 ment of polar regions because of the so-called *pole-problem* (? , ?). The pole problem refers  
 36 to numerical instability arising from the convergence of meridian lines into polar singu-  
 37 larities on latitude-longitude grids (e.g., Figure ??a, hereafter referred to as *lat-lon* grids).  
 38 Depending on the numerics, methods exist to suppress this instability, and lat-lon grids  
 39 may be advantageous for polar processes by representing structures with finer resolution  
 40 than elsewhere in the computational domain. With the recent trend towards globally uni-  
 41 form unstructured grids, the benefits of lat-lon grids in polar regions may be lost. (WHL:  
 42 Would it be worth mentioning that quasi-uniform grids also have benefits in polar re-  
 43 gions, e.g. by removing elongated grid cells?) (ARH: I don't think it's worth getting into  
 44 that here. You'll see later in the manuscript that b/c the cells are narrower in the E-W  
 45 direction, more cells can fit into the big ablation zone on the western margin.) (PHL: an-  
 46 other advantage of lat-lon grids is that the coordinate lines are orthogonal!) In this study  
 47 we evaluate a number of grids and dynamical cores (hereafter referred to as *dycores*) avail-  
 48 able in the Community Earth System Model, version 2.2 (CESM2.2; ?, ?), including new  
 49 variable-resolution grids, to understand their impacts on the simulated Arctic climate.  
 50 We focus on the climate and surface mass balance of the Greenland Ice Sheet.

51 In the 1970s the pole problem was largely defeated through the adoption of effi-  
 52 cient spectral transform methods in GCMs. (WHL: Add refs) These methods transform  
 53 grid point fields into a global, isotropic representation in wave space, where linear op-  
 54 erators (e.g., horizontal derivatives) in the equation set can be solved exactly. While spec-  
 55 tral transform methods are still used today, local numerical methods have become de-  
 56 sirable for their ability to run efficiently on massively parallel systems. The pole prob-  
 57 lem has thus re-emerged in contemporary climate models that use lat-lon grids, and some  
 58 combination of reduced grids (define 'reduced grids') and polar filters are necessary to  
 59 ameliorate this instability (? , ?). Polar filters subdue the growth of unstable modes by  
 60 applying additional damping to the solution over polar regions. This damping reduces  
 61 the effective resolution in polar regions such that the resolved scales are *approximately*  
 62 the same everywhere on the grid. We emphasize *approximately* since it's at least con-

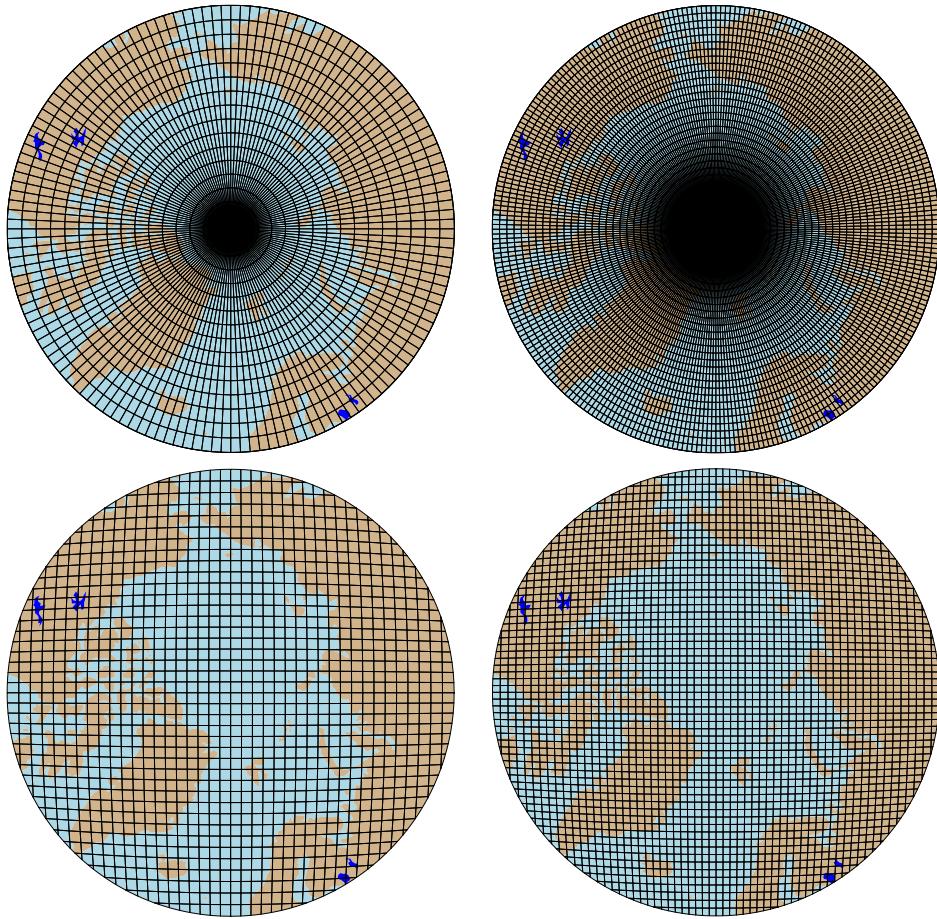
63 ceivable that marginal increases in effective resolution occur over polar region in lat-lon  
64 grids, despite polar filtering.

65 (WHL: This raises the question: How do lat-lon grids better resolve polar regions  
66 at a given nominal global resolution, if the filtering makes them effectively coarser? Does  
67 the damping degrade some processes more than others?)(ARH - The phrase "approx-  
68 imately the same" doing a lot of heavy lifting here b/c I frankly don't know. The results  
69 from this paper suggest that the effective resolution is finer over polar regions, but this  
70 is a bit hand-wavy, despite my attempts to quantify it more exactly, and based on the  
71 precipitation rates having smaller biases. Also, to be clear, effective resolution is specific  
72 to atmospheric features; the land model is free to take advantage of these additional grids  
73 cells representing the ablation zone, and they certainly do. Perhaps I can add something  
74 to the effect of "approximately the same everywhere on the grid, although marginal in-  
75 creases in effective resolution over the polar regions cannot be ruled out...")

76 PHL: These are good points! I think it is useful to separate dynamics from physics  
77 in this discussion. Physics is called at more columns on the lat-lon grid as one approaches  
78 the poles. There is no "polar filtering" in physics so effectively the physics resolution is  
79 much higher than on a the cubed-sphere grid as one approaches the mid-latitudes and  
80 poles. Dynamics, on the other hand, filters out smaller scales so effectively it should only  
81 retain the 1 degree scale. That said, you will get a more accurate solution since you have  
82 more grid-points to represent waves (even though smaller scales are filtered) compared  
83 to the cubed-sphere grid. I think that is pretty clear in flow over orography. One can  
84 compare the lat-lon grid setup to a spectral model where one keeps viscosity coefficients  
85 fixed as one increased horizontal resolution - in such a setup one will get convergence of  
86 the dynamics!

87 An alternative approach is to use unstructured grids, which permit quasi-uniform  
88 grid spacing globally and eliminates the pole-problem (e.g., Figure ??c). (WHL: Add  
89 1 or 2 refs) This grid flexibility also allows variable-resolution or regional grid refinement  
90 (e.g., Figure ??, hereafter abbreviated as VR). In principle, grid refinement over polar  
91 regions can make up for any loss of resolution in transitioning away from lat-lon grids  
92 (e.g., Figure ??). However, this refinement comes at the cost of a smaller CFL-limited  
93 time step in the refined region. (The CFL-condition — short for Courant–Friedrichs–Lewy  
94 condition — is a necessary condition for numerical stability when using discrete data in  
95 time and space.) Unstructured grids also scale more efficiently on parallel systems than  
96 lat-lon grids, another advantage as computing power increases over time.

97 The meteorology and climate of the Arctic are characterized by a range of processes  
98 and scales that are difficult to represent in GCMs (? , ? , ? , ?). For example, while syn-  
99 optic scale storms are generally well represented at typical GCM resolutions of 1 to 2  
100 degrees (? , ? , ?), mesoscale polar lows are not well resolved at these resolutions. These  
101 mesoscale systems are prevalent during winter and produce gale-force winds that can in-  
102 duce large heat and moisture fluxes through the underlying sea-ice/ocean interface. The  
103 Arctic also contains the Greenland Ice Sheet (GrIS), which blankets the world's largest  
104 island. Many of the processes that control the GrIS annual surface mass balance (SMB;  
105 the integrated sum of precipitation and runoff) are only partially resolved at typical GCM  
106 resolutions. For example, GrIS precipitation is concentrated at the ice-sheet margins,  
107 where steep topographic slopes drive orographic precipitation. GrIS ablation areas (marginal  
108 regions where seasonal melting exceeds the annual mass input from precipitation) are  
109 typically less than 100 km wide and are confined to low-lying areas or regions with low  
110 precipitation. GCMs struggle to resolve the magnitude and extent of these features (? ,  
111 ? , ?), which can lead to unrealistic ice sheet growth in models with an interactive ice sheet  
112 component (e.g., ? , ?). (WHL: I suggest adding some language explaining why advance  
113 is more prevalent than retreat. E.g.: "... of these features. For example, CESM2 typ-  
114 ically underestimates ablation in the northern GrIS, leading to unrealistic ice advance  
115 when run with an interactive ice sheet component." Could cite Lofverstrom et al. (2020.))(ARH:



**Figure 1.** Computational grids for the uniform  $1^\circ - 2^\circ$  grids in this study.

I may delete the reference to this anecdote of advancing in northern GrIS altogether b/c I've been unable to show whether it's related to resolution. I don't have confidence defining a direction of the SMB bias due to resolution because resolution inflates both source and sink terms; the net effect can go either way.)

This study compares the representation of high-latitude regions using the spectral-element and finite-volume dycores in CESM2.2, as these two dycores treat high latitudes (i.e., the pole problem) in different ways. It also serves to document the Arctic-refined grids developed by the authors for CESM2.2. Section ?? documents the grids, dycores, and physical parameterizations used in this study, and also describes the experiments, datasets, and evaluation methods. Section ?? analyzes the results of the experiments, and Section ?? provides a general discussion and conclusions.

## 2 Methods

### 2.1 Dynamical cores

The atmospheric component of CESM2.2, the Community Atmosphere Model, version 6.3 (CAM; ?, ?), supports several different atmospheric dynamical cores. These include dycores on lat-lon grids, such as finite-volume (FV; ?, ?) and Eulerian spectral transform (EUL; ?, ?) models, and dycores built on unstructured grids, including spectral-element (SE; ?, ?) and finite-volume 3 (FV3; ?, ?) models. The EUL dycore is the old-



**Figure 2.** Spectral-element grid for the VR Arctic grids in this study. Note that this is not the computational grid; each element has  $3 \times 3$  independent grid points.

est and least supported dycore in CAM. FV3 is the newest dycore, but was not fully incorporated when this work begin. This study compares the performance of the SE and FV dycores, omitting the EUL and FV3 dycores.

[phl: maybe mention that FV was used for IPCC simulations; SE in iHESP]

### 2.1.1 Finite-volume (FV) dynamical core

The FV dycore is a hydrostatic model that integrates the equations of motion using a finite-volume discretization on a spherical lat-lon grid (?, ?). The 2D dynamics evolve in floating Lagrangian layers that are periodically mapped to an Eulerian reference grid in the vertical (?, ?), using a hybrid-pressure vertical coordinate. Hyperviscous damping is applied to the divergent modes, while Laplacian damping is applied to momentum in the top few layers, referred to as a *sponge layer* (?, ?) [phl: as far as I know the Laplacian-like damping is not invoked in lower resolution setups - was implemented for 1/4 applications where the polar night jet became excessively strong]. A polar filter damps computational instability due to the convergence of meridians, allowing a longer time step. It takes the form of a Fourier filter in the zonal direction, with the damping coefficients increasing monotonically in the poleward direction (?, ?).

### 2.1.2 Spectral-element (SE) dynamical core

The SE dycore is a hydrostatic model that integrates the equations of motion using a high-order continuous Galerkin method (?, ?, ?). The computational domain is a cubed-sphere grid tiled with quadrilateral elements (see Figure ??). Each element contains a fourth-order basis set in each horizontal direction, with the solution defined at the roots of the basis functions, the Gauss-Lobatto-Legendre (GLL) quadrature points. This results in 16 GLL nodal points per element, with 12 of the points lying on the (shared) element boundary. Communication between elements uses the direct stiffness summation (?, ?), which applies a numerical flux to the element boundaries to reconcile overlapping nodal values and produce a continuous global basis set.

As with the FV dycore, the dynamics evolve in floating Lagrangian layers that are subsequently mapped to an Eulerian reference grid. A dry mass vertical coordinate was recently implemented for thermodynamic consistency with condensates (?, ?). The 2D

163 dynamics have no implicit dissipation, and so hyperviscosity operators are applied to all  
 164 prognostic variables to remove spurious numerical errors (?, ?). Laplacian damping is  
 165 applied in the sponge layer.

166 The SE dycore supports regional grid refinement via its VR configuration, requiring  
 167 two enhancements to uniform-resolution grids: (1) As the numerical viscosity increases  
 168 with resolution, explicit hyperviscosity relaxes according to the local element size, re-  
 169 ducing in strength by an order of magnitude per halving of the grid spacing. A tensor-  
 170 hyperviscosity formulation is used (?, ?), which adjusts the coefficients in two orthog-  
 171 onal directions to more accurately target highly distorted quadrilateral elements. (2) The  
 172 topography boundary conditions need to be smoothed in a way that does not excite grid  
 173 scale modes, and so the NCAR topography software (?, ?) has been modified to scale  
 174 the smoothing radius by the local element size.

175 For spectral-element grids with quasi-uniform grid spacing, we use a variant in which  
 176 tracer advection is computed using the Conservative Semi-Lagrangian Multi-tracer trans-  
 177 port scheme (CSLAM) (?, ?). (WHL: I'm not sure what CSLAM is a variant of, since  
 178 I don't see a tracer transport scheme described above. How is tracer transport handled  
 179 on VR grids? If it's not too much detail, why does CSLAM not work on VR grids?) (ARH:  
 180 fair point. the spectral-element section is lacking in general, so keeping this here as a re-  
 181 minder to explain the spectral-element method is used for tracer transport in "vanilla"  
 182 cam-se, and we only support vanilla var-res. PHL is planning on adding a description  
 183 of enhancements to cam-se introduced in CESM2.2, since this turns out to be the only  
 184 paper documenting CESM2.2.) CSLAM has improved tracer property preservation and  
 185 accelerated multi-tracer transport. It uses a separate grid from the spectral-element dy-  
 186 namics, dividing each element into  $3 \times 3$  control volumes with quasi-equal area. The phys-  
 187 ical parameterizations are computed from the state on the CSLAM grid, which has clear  
 188 advantages over the default SE dycore in which the physics are evaluated at the GLL  
 189 nodal points (?, ?).

## 190 2.2 Grids

191 We evaluate six grids in this study (Table ??). The FV dycore is run with  $1^\circ$  and  
 192  $2^\circ$  grid spacing, referred to as *f09* and *f19*, respectively (Figure ??a,b). We also run the  
 193  $1^\circ$  equivalent of the CAM-SE-CSLAM grid, referred to as *ne30pg3* (Figure ??c), where  
 194 *ne* refers to a grid with  $ne \times ne$  elements per cubed-sphere face, and *pg* denotes that  
 195 there are  $pg \times pg$  control volumes per element for computing the physics. We run an  
 196 additional  $1^\circ$  CAM-SE-CSLAM grid with the physical parameterizations computed on  
 197 a grid with  $2 \times 2$  control volumes per element, *ne30pg2* (Figure ??d; ?, ?).

198 Two VR meshes were developed for the CESM2.2 release to support grid refine-  
 199 ment over the Arctic (Figure ??). This paper serves as the official documentation of these  
 200 grids. The Arctic meshes were developed using the software package SQuadgen (<https://github.com/ClimateGlobalChange/squadgen>). The *Arctic* grid is a  $1^\circ$  grid with  $\frac{1}{4}^\circ$   
 201 regional refinement over the broader Arctic region. The *Arctic – GrIS* grid is identi-  
 202 cal to the *Arctic* grid, but with an additional patch covering the island of Greenland with  
 203  $\frac{1}{8}^\circ$  resolution.

205 The physics time step should depend on grid resolution. Increased horizontal res-  
 206 olution permits faster vertical velocities that reduce characteristic time scales, so the physics  
 207 time step is reduced to avoid large time truncation errors (?, ?). The *ARCTIC* and *Arctic–*  
 208 *GrIS* grids are run with a  $4 \times$  and  $8 \times$  reduction in physics time step relative to the de-  
 209 fault 1800 s time step for the standard  $1^\circ$  and  $2^\circ$  grids (Table ??).

210 All grids and dycores in this study use 32 levels in the vertical, with a model top  
 211 of about 1 hPa or 40 km. However, any grid or dycore can in principle be run with a higher  
 212 top or finer vertical resolution.

grid name	dycore	$\Delta x_{\text{eq}}$ (km)	$\Delta x_{\text{refine}}$ (km)	$\Delta t_{\text{phys}}$ (s)
<i>f19</i>	FV	278	-	1800
<i>f09</i>	FV	139	-	1800
<i>ne30pg2</i>	SE-CSLAM	167	-	1800
<i>ne30pg3</i>	SE-CSLAM	111	-	1800
<i>ne30pg3*</i>	SE-CSLAM	111	-	450
<i>Arctic</i>	SE	111	28	450
<i>Arctic – GrIS</i>	SE	111	14	225

**Table 1.** Grids and dycores used in this study.  $\Delta x_{\text{eq}}$  is the average equatorial grid spacing,  $\Delta x_{\text{refine}}$  is the grid spacing in the refined region (if applicable), and  $\Delta t_{\text{phys}}$  is the physics time step. FV refers to the finite-volume dycore, SE the spectral-element dycore, and SE-CSLAM the spectral-element dycore with CSLAM tracer advection. We use the *ne30pg3* grid for two runs with different values of  $\Delta t_{\text{phys}}$ .

### 2.3 Physical parameterizations

All simulations in this study use the CAM6.3 physical parameterization package (hereafter referred to as the *physics*; ?, ?). The physics in CAM6 differs from its predecessors through the incorporation of high-order turbulence closure, Cloud Layers Unified by Binormals (CLUBB; ?, ?, ?), which jointly acts as a planetary boundary layer (PBL), shallow convection, and cloud macrophysics scheme. CLUBB is coupled with the MG2 microphysics scheme (? , ?), with prognostic precipitation and classical nucleation theory to represent cloud ice for improved cloud-aerosol interactions. Deep convection is parameterized using a convective quasi-equilibrium, bulk plume mass flux scheme (? , ?, ?) and includes convective momentum transport (? , ?). PBL form drag is modeled after ? (?). Orographic gravity wave drag is represented with an anisotropic method informed by the orientation of topographic ridges at the sub-grid scale.

Initial simulations with the *ne30pg3* SE grid produced weaker shortwave cloud forcing relative to the tuned finite-volume dycore. All runs with the SE dycore therefore have two CLUBB parameter changes to provide more realistic cloud forcing and top-of-atmosphere radiation balance. These are CLUBB’s *gamma* parameter, reduced from 0.308 to 0.270, and *c14*, reduced from 2.2 to 1.6. Briefly, *gamma* scales the width of the sub-grid distribution of vertical velocity, and *c14* controls the strength of the damping term in the equation for the horizontal component of turbulent kinetic energy. For a thorough explanation of how CLUBB parameters impact the simulated climate, see (? , ?).

### 2.4 Experimental design

We run all grids and dycores using an identical transient 1979-1998 Atmospheric Model Inter-comparison Project configuration (AMIP), with prescribed monthly SST and sea ice following ? (?). In CESM terminology, AMIP simulations use the *FHIST* computational set and run out of the box in CESM2.2.

All grids and dycores simulate the GrIS surface mass balance (SMB). The SMB is the sum of the mass accumulation term, i.e., precipitation, and the mass sink term, ablation. Ablation is the sum of evaporation/sublimation and liquid runoff, with runoff being a combination of liquid precipitation and snow and ice melt. Not all liquid precipitation runs off the ice sheet; rain can penetrate pore spaces in the snowpack/firn layer and freeze, increasing the snow mass. These processes are represented by different CESM components, but it is the Community Land Model, version 5 (CLM; ?, ?), that aggregates these processes and computes the SMB.

data product	years used in this study	resolution	citation
ERA5	1979-1998	$\frac{1}{4}^\circ$	? (?)
CERES-EBAF ED4.1	2003-2020	$\frac{1}{4}^\circ$	? (?)
CALIPSO-GOCCP	2006-2017	$1^\circ$	? (?)
RACMO2.3	1979-1998	11 km	? (?)
RACMO2.3p2	1979-1998	5.5 km	? (?)

**Table 2.** Description of observational datasets used in this study.

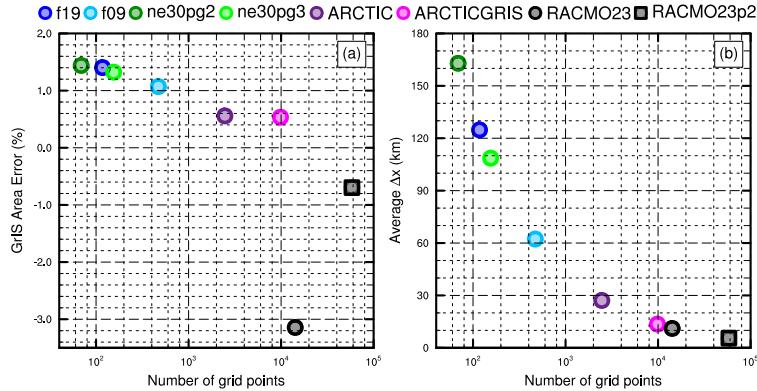
CLM runs on the same grid as the atmosphere, using a downscaling technique to account for sub-grid variability in SMB. In short, the ice sheet patch in a CLM grid cell is subdivided into 10 elevation classes (ECs), each with a distinct surface energy balance and SMB. The area fraction of each EC is derived from a high-resolution GrIS elevation dataset. The near-surface air temperature, humidity, and air density are calculated for each EC using an assumed lapse rate and the elevation difference from the grid-cell mean. Precipitation from CAM is repartitioned into solid or liquid based on the surface temperature of the EC; precipitation falls as snow for  $T < -2^\circ \text{ C}$ , as rain for  $T > 0^\circ \text{ C}$ , and as a rain/snow mix between  $-2^\circ \text{ C}$  and  $0^\circ \text{ C}$ . Snow accumulation in each EC is limited to a depth of 10 m liquid water equivalent. Any snow above the 10-m cap is routed to the ocean as solid runoff. Refreezing of liquid within the snowpack is an additional source of ice. Integrating over all ECs, weighting by the area fractions, provides a more accurate SMB than would be found using the grid-cell mean elevation. For a more detailed description of how the SMB is computed in CESM, we refer the reader to ?, ?, ?, ?.

Changes in ice depth, but not snow depth, count toward the SMB. That is, snow accumulation above the 10-m cap contributes a positive SMB, and surface ice melting (after melting of the overlying snow) yields a negative SMB. Since snow in the accumulation zone must reach the cap to simulate a positive SMB, the snow depths on the VR grids were spun up by forcing CLM in standalone mode, cycling over a 20-year *Arctic FHIST* run for about 500 years. The uniform-resolution grids are initialized with the SMB from an existing *f09* spun-up initial condition.

## 2.5 Observational Datasets

We use several observational datasets (Table ??) to gauge the performance of the simulations. Several products are near-global gridded datasets commonly used to evaluate GCMs. We gathered SMB datasets from several sources. RACMO2.3 11km and RACMO2.3p2 5.5km are regional model simulations targeting Greenland, forced by ERA40, ERA interim and ERA5 reanalysis products at the domain boundaries. The RACMO simulations have been shown to perform skillfully against observations and are often used as modeling targets ?, ?, ?.

In-situ SMB (snow pit and ice cores) and radar accumulation datasets (e.g., IceBridge) are maintained in The Land Ice Verification and Validation toolkit (LIVVkit), version 2.1 ?, ?. However, these point-wise measurements are difficult to compare to model output spanning several different grids, especially since the SMB from each EC is not available from the model output. We used a nearest-neighbor technique for an initial analysis, which showed that the model biases are similar to those computed using the RACMO datasets. Because of the uncertainty of comparing gridded fields to point-wise measurements, and the lack of information added with regard to model biases, we omitted these datasets from our analysis.



**Figure 3.** The spatial properties of the GrIS as represented by different grids in this study. (Left) GrIS area error, computed as the relative differences from a 4km dataset used to create the CESM ice masks, (right) approximate average grid spacing over GrIS.

285

## 2.6 SMB Analysis ARH - this section is under construction

286  
287  
288  
289  
290  
291  
292  
293  
294

We seek to integrate SMB components over a GrIS ice mask and to diagnose their contributions to the GrIS mass budget. However, the ice masks vary across the grids, especially in comparison to the RACMO3.2 ice mask, whose total area is about 3% less than that of the reference dataset (Figure ??). The area errors in Figure ?? may not seem large, but even 1–2% area differences can lead to large differences in integrated SMB (?, ?). CLM’s dataset creation tool generates the model ice mask by mapping a high-resolution dataset to the target grid using the Earth System Modeling Framework first-order conservative remapping algorithm (ESMF; ?, ?). The figure suggests that the mapping errors are less than 1.5% across the CESM2 grids.

295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305

We have taken a common-ice-mask approach by mapping all model fields to the lowest-resolution grids, i.e., the *f19* and *ne30pg2* grids, and integrating over these low-resolution ice masks. The use of low-resolution common ice masks is a conservative decision and is largely because we seek to use first-order remapping algorithms to map fields to the common ice mask, which is not reliable in the downscaling direction, i.e., mapping to a higher-resolution grid than the source grid. We use two remapping algorithms: ESMF first-order conservative and the TempestRemap (?, ?) high-order monotone algorithm. Since mapping errors are sensitive to grid type, we evaluate all quantities on both common ice masks, the *f19* and *ne30pg2* masks. Thus, we evaluate an integrated quantity on a given grid up to four times to estimate the uncertainty due to differences in grid type and remapping algorithms.

306  
307  
308  
309

The SMB is expressed in a form that is agnostic of water phase, a total water mass balance, to facilitate comparisons across different grids with different ice masks and to increase consistency with the variables available in the RACMO datasets. The SMB for total water can be expressed as:

$$SMB = precipitation + runoff + evaporation/sublimation, \quad (1)$$

310  
311  
312  
313  
314  
315

where all terms have consistent sign conventions (positive values contribute mass, negative values reduce mass). The precipitation source term refers to the combined solid and liquid precipitation, runoff refers to the liquid water sink, and evaporation/sublimation is the vapor sink. Since the runoff term aggregates many processes, we isolate the melting contribution by also tracking the combined snow plus ice melt. Note that this SMB expression is different from the internally computed SMB in CLM (?, ?).

316 We consider two approaches for mapping and integrating the SMB components over  
 317 the common ice masks:

- 318 1. Map the grid-cell mean quantities to the common grid, and integrate the mapped  
 319 fields over the common ice masks.
- 320 2. Map the patch-level quantities (i.e., the state over the ice fractional component  
 321 of the grid cell) to the common grid, and integrate the mapped fields over the com-  
 322 mon ice masks.

323 Note that we are mapping to low-resolution grids that have larger GrIS areas (Figure ??). Since the components of equation ?? are not confined to the ice mask, method ??  
 324 reconstructs the SMB over the portion of the common ice mask that is outside the ice  
 325 mask on the source grid. While this may be a an acceptable way to reconstruct the mass  
 326 source terms over different ice masks, ice melt is zero outside the source ice mask, and  
 327 so method ?? will underestimate the mass sink term. Method ??, in contrast, does not  
 328 map values outside the source ice mask to the common ice mask. This tends to be dis-  
 329 sipative, since mapping to the low-resolution grid will average a field of non-zero values  
 330 over the ice mask with a field of zeros outside the ice mask. However, patch-level val-  
 331 ues for processes exclusive to the ice mask (e.g., ice melt) will have larger magnitudes  
 332 than the the grid-mean quantities used in method ??.

333 The different error characteristics of the two methods are used to diversify the en-  
 334 semble. Each of the four regridding combinations (with conservative and high-order remap-  
 335 ping to the *f09* and *ne30pg2* grids) are repeated with each method, resulting in (up to)  
 336 eight values for each integrated quantity. Unfortunately, the patch-level values of evap-  
 337 oration/sublimation are not available from the model output, and we estimate their con-  
 338 tribution by zeroing out the field for grid cells that have no ice, prior to mapping to the  
 339 common ice mask. This will degrade the SMB estimates using method ??, but note that  
 340 we are more interested in the behavior of the processes across grids and dycores, expressed  
 341 as the components of the SMB, rather than the SMB itself.

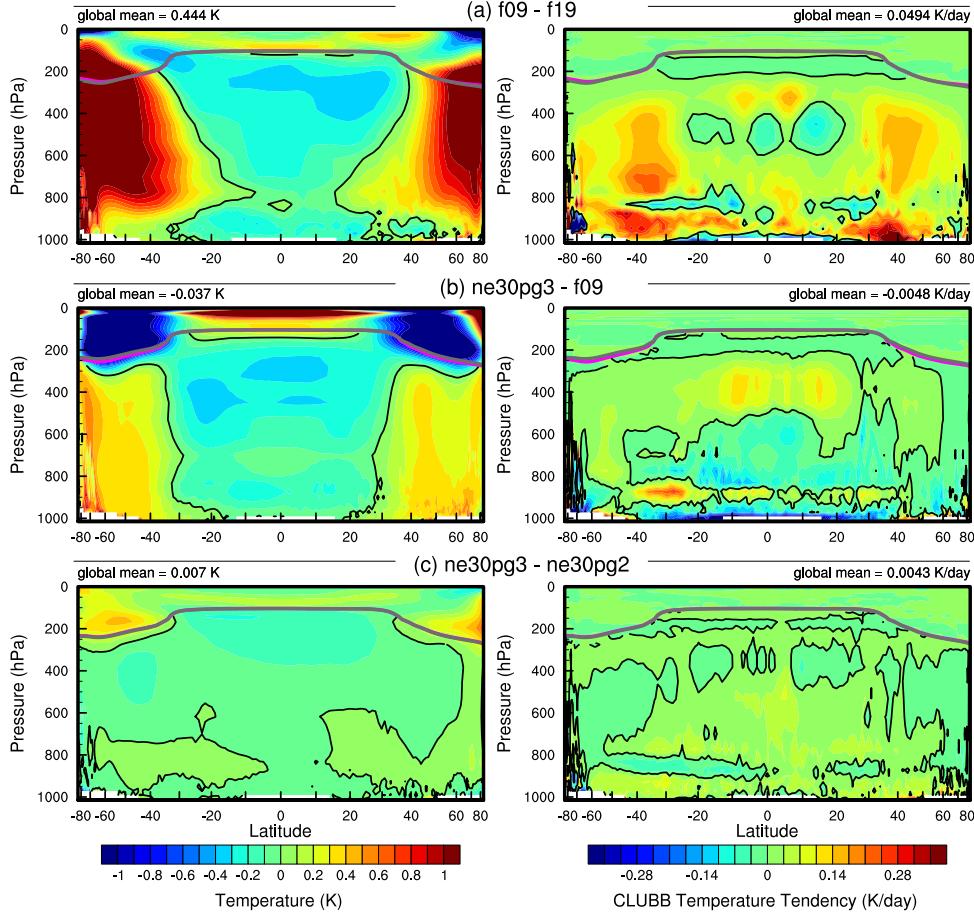
### 343 3 Results

#### 344 3.1 Tropospheric temperatures

345 Before delving into the simulated Arctic, we assess the global mean differences be-  
 346 tween the various grids and dycores. Figure ?? shows 1979-1998 annual mean, zonal mean  
 347 height plots expressed as differences between uniform-resolution grids and dycores. The  
 348 *f09* grid is warmer than the *f19* grid, primarily in the mid-to-high latitudes through-  
 349 out the depth of the troposphere. This is a common response to increasing horizontal  
 350 resolution in GCMs (? , ? , ?). ? (?) have shown that this occurs in CAM due to greater  
 351 resolved vertical velocities which, in turn, drive more condensational heating in the CLUBB  
 352 macrophyics. The right columns in Figure ?? support this interpretation, showing an  
 353 increase in the climatological CLUBB heating in low ([looks more like subtropics](#)) and  
 354 mid-latitudes on the *f09* grid.

355 As the SE dycore is less diffusive than the FV dycore, the resolved vertical veloc-  
 356 ities are larger in the SE dycore, and so the *ne30pg3* troposphere is modestly warmer  
 357 than *f09* (Figure ??). The stratosphere responds differently, with *ne30pg3* much cooler  
 358 than *f09* in the mid-to-high latitudes. Figure ?? also shows small temperature differ-  
 359 ences between *ne30pg3* and *ne30pg2*, with *ne30pg3* slightly warmer near the tropopause  
 360 at high latitudes. This is consistent with the similar climates found for these two grids  
 361 by ? (?).

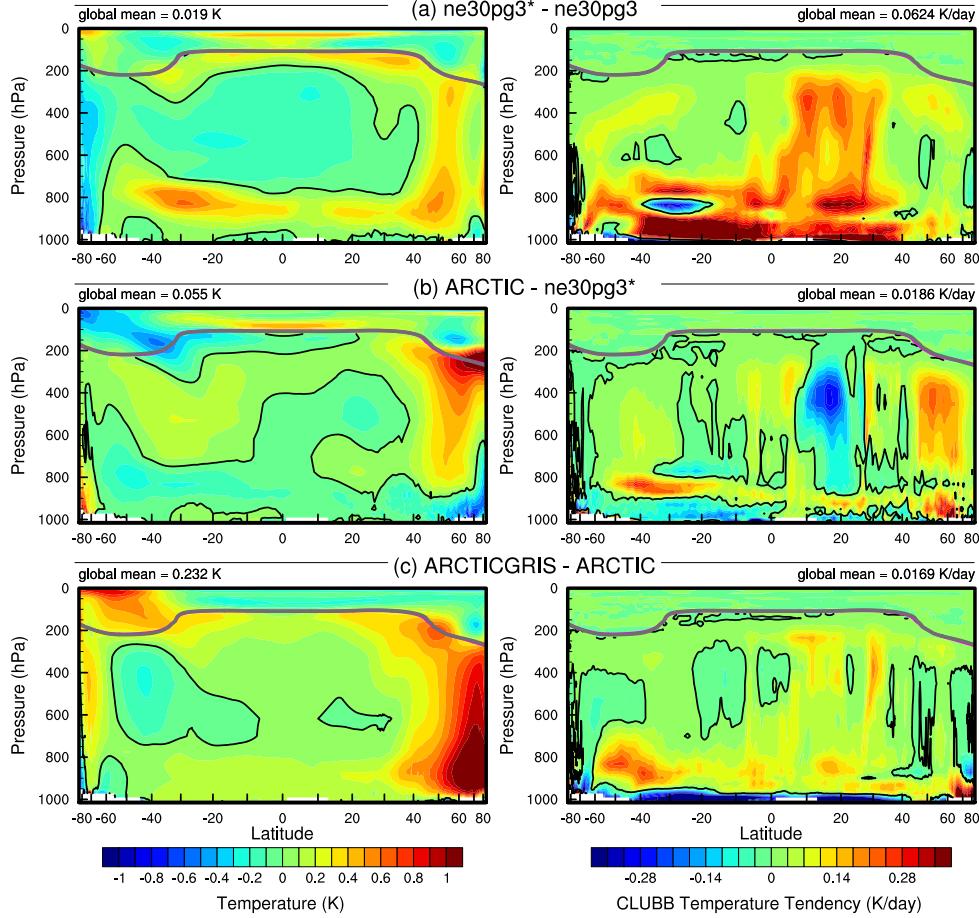
362 Comparing the VR grids to the uniform-resolution grids is complicated because we  
 363 simultaneously increase the resolution and reduce the physics time-step, both of which



**Figure 4.** 1979–1998 annual mean temperature (left column) and CLUBB temperature tendencies (right column) in zonal mean height space, expressed as differences between the various  $1^{\circ} - 2^{\circ}$  grids. The thick grey and magenta lines are the tropopause for the control run and the test run, respectively.

364 impact the solution (?). We therefore run an additional *ne30pg3* simulation with the  
 365 shorter physics time step used in the *Arctic* grid (450 s), referred to as *ne30pg3\**. Figure ?? shows the difference between *ne30pg3\** and *ne30pg3* for climatological summer  
 366 temperatures in zonal-mean height space. Much of the troposphere is warmer with the  
 367 reduced time step, and the mechanism is similar in that the shorter time step increases  
 368 condensational heating by CLUBB. Figure ?? shows the difference in climatological sum-  
 369 mer temperature between the *ARCTIC* grid and the *ne30pg3\** grid. With the same physics  
 370 time step on each grid, the greater condensational heating and warmer temperatures are  
 371 confined to the refined Arctic region.  
 372

373 Figure ?? shows that the *Arctic-GrIS* grid is much warmer than the *Arctic* grid  
 374 in the Arctic summer. This may be due, in part, to the shorter physics time step, but  
 375 the temperature response is too large to be explained by the CLUBB changes alone. This  
 376 summer warming appears to be a result of variations in the stationary wave pattern, with  
 377 anomalous southerly winds to the west of Greenland (not shown). This dynamic response  
 378 is interesting, because other than the physics time step, the only difference between the  
 379 *Arctic-GrIS* and *Arctic* runs is the doubling of resolution over Greenland. This be-  
 380 havior will be explored further in a future study.



**Figure 5.** As in Figure ?? but for the short-time-step experiment and the VR grids. The fields plotted are for the climatological northern hemisphere summer. We focus on summer because that is when the resolution response is largest, and the refined regions are located in the northern hemisphere.

It is useful to understand summer temperature biases due to their control on ice and snow melt over the Greenland Ice Sheet (GrIS; ?, ?). Figure ?? shows the 1979–1998 lower troposphere summer temperature bias relative to ERA5. It is computed from the 500–1000 hPa geopotential thickness, solving for the layer-mean virtual temperature using the hypsometric equation (Might be good to write this down, for those of us who studied atmospheric science a really long time ago). The results are consistent with the zonal mean height plots; increasing resolution from  $f19$  to  $f09$  warms the climate, and the  $1^\circ$  SE grids are warmer than the FV grids. The FV summer temperatures are persistently colder than ERA5, whereas the  $1^\circ$  SE grids are not as cold, and are actually warmer than ERA5 at high-latitudes, north of  $75^\circ$ . All grids show a north-south gradient in bias over Greenland, with the summer temperature bias more positive for the northern part of the ice sheet. This pattern is also evident in the 2m summer temperature bias over Greenland (not shown).

The *Arctic* grid has summer temperatures similar to the  $1^\circ$  SE grids, but is slightly warmer over northern Eurasia and the North Pole. An anomalous cooling patch forms to the west of Greenland, centered over Baffin Island. The *Arctic-GrIS* grid is warmer

397 than the *Arctic* grid over most of the Arctic, but with a similar spatial pattern of sum-  
398 mer temperature bias.

399 Some of these temperature differences may be related to different summer short-  
400 wave cloud forcing for the various grids and dycores. Figure ?? shows the summer short-  
401 wave cloud forcing bias in six runs, using the CERES-EBAF product. A negative bias  
402 corresponds to excessive reflection and cooling. The uniform grids have similar biases,  
403 with the clouds reflecting 20–40 W/m<sup>2</sup> too much shortwave radiation over a wide swath  
404 of the Arctic, primarily the land masses ([please explain what SWCF in caption.](#)). There  
405 is also a halo of positive bias (clouds not reflective enough) around the ocean perime-  
406 ter of Greenland. The *Arctic* grid has much smaller cloud forcing biases over the Arc-  
407 tic land masses, but is still too reflective over Alaska, the Canadian Archipelago, and parts  
408 of Eurasia. Compared to the *Arctic* grid, the *ARCTCGRIS* grid vastly reduces the cloud  
409 forcing bias over Eurasia, and also improves the bias over North America. In both VR  
410 grids, the halo of positive forcing bias around the perimeter of Greenland is absent.

411 While the summer cloud forcing biases are consistent with the summer tempera-  
412 ture biases in Figure ?? – regions where clouds are too reflective coincide with regions  
413 that are too cold – it is not clear whether the cold biases are caused by the cloud biases,  
414 or whether the cold biases amplify the cloud forcing bias. ([This ends abruptly. Anything](#)  
415 [else to add?](#))

### 416 3.2 Shortwave radiation over Greenland

417 In addition to summer temperatures, shortwave radiation is an important deter-  
418 minant of snow and ice melt. Figure ?? shows the summer incident shortwave radiation  
419 bias at the surface, zoomed in over Greenland. The top panel shows the bias relative to  
420 the CERES-EBAF dataset, and the bottom panel relative to the RACMO2.3p2 dataset.  
421 The halo of excessive incident shortwave radiation around the coasts of Greenland is ap-  
422 parent for both datasets in relation to the coarser grids, consistent with the shortwave  
423 cloud forcing biases in Figure ??.

424 The ice sheet interior receives too little shortwave radiation on the coarser grids.  
425 On the VR grids, both the interior shortwave deficit and the excessive shortwave around  
426 the ocean perimeter are improved. This suggests that the coarse-grid clouds are too thick  
427 in the Greenland interior and too thin around the perimeter, and that increasing hor-  
428 izontal resolution reduces these biases. This is consistent with the total summer cloud  
429 fraction bias, computed from the CALIPSO-GOCCP cloud dataset and shown in Fig-  
430 ure ???. Note that total cloud fraction characterizes the cloud field at all vertical levels,  
431 but attenuates the changes arising from any single layer due to the maximum overlap  
432 assumption used to compute this quantity. Despite the attenuated signal, the total cloud  
433 fraction for the VR grids does indicate reduced cloud coverage in the interior and increased  
434 cloudiness around the ocean perimeter.

435 The agreement of the cloud biases in and around Greenland from multiple inde-  
436 pendent datasets shows that the biases are a robust feature of the coarser grids. The re-  
437 duced biases on the VR grids suggest that the coarse-grid biases are a result of insuf-  
438 ficient horizontal resolution.

### 439 3.3 Greenland surface mass balance

440 The accuracy of the simulated SMB is expected to be sensitive to grid resolution.  
441 Figure ??b shows the average grid spacing over the Greenland Ice Sheet (GrIS) in all  
442 six grids in this study. The *ne30pg2* grid has the coarsest representation with an aver-  
443 age  $\Delta x = 160 \text{ km}$ , and the *Arctic-GrIS* grid has the highest resolution with an av-  
444 erage  $\Delta x = 14.6 \text{ km}$ , similar to the grid spacing of the  $11 \text{ km}$  RACMO2.3 grid. The  
445 *ne30pg3* grid has an average  $\Delta x = 111.2 \text{ km}$ , substantially coarser than the *f09* grid,

grid name	accumulation	total melt	runoff	sublimation	SMB
<i>RACMO</i>	768.5 (733.5)	-347.2 (-436.4)	-221.7 (-258.5)	-36.5 (-38.8)	510.3 (436.2)
<i>f19</i>	882.5 (913.5)	-440.3 (-546.5)	-283.7 (-284.3)	-36.6 (-37.5)	562.2 (591.7)
<i>f09</i>	874.8 (882.1)	-418.4 (-482.3)	-255.0 (-212.3)	-38.1 (-37.4)	581.7 (632.4)
<i>ne30pg2</i>	1000. (973.4)	-549.4 (-647.3)	-383.9 (-347.0)	-33.4 (-32.1)	582.7 (594.3)
<i>ne30pg3</i>	934.9 (909.3)	-568.8 (-686.7)	-356.2 (-330.1)	-34.4 (-32.6)	544.3 (546.6)
<i>Arctic</i>	795.9 (818.6)	-367.3 (-436.8)	-208.9 (-194.2)	-44.1 (-43.9)	542.9 (580.5)
<i>Arctic – GrIS</i>	708.7 (747.3)	-471.6 (-610.4)	-261.1 (-307.8)	-50.7 (-51.8)	396.9 (387.7)

**Table 3.** 1979-1998 surface mass balance of the Greenland Ice Sheet in Gt/yr. Values shown are using the common ice mask approach described in the methods section, whereas values in parentheses are from integrating over the native grid and ice mask.

with an average  $\Delta x = 60 \text{ km}$ . Although *ne30pg3* and *f09* have similar average grid spacing over the entire globe, and comparable computational costs, the convergence of meridians on the FV grid enhances the resolution over the GrIS. The *Arctic* grid has an average grid spacing of  $\Delta x = 27.8 \text{ km}$ , and is about 10 times more expensive than the  $1^\circ$  models. The *Arctic–GrIS* grid is about twice as expensive as the *Arctic* grid. (Wondering if this paragraph would go better in an earlier section.)

The lower panels of Figure ?? show the summer climatological mean precipitation bias over the GrIS, expressed as the fractional difference from the RACMO2.3p2 solution. The coarse  $1^\circ$ – $2^\circ$  grids have large, positive biases centered over the southern dome. The *Arctic* grid reduces this bias substantially, and the *Arctic–GrIS* grid reduces it further. This suggests that the southern dome bias arises from inadequate horizontal resolution, consistent with the original GrIS VR experiments in ? (?).

Large GrIS accumulation rates result from synoptic systems arriving from the south. These systems are orographically lifted at the ice sheet margin, especially over the steep slopes in southeast Greenland, concentrating heavy precipitation near the ice margin. At lower resolutions, the topography is too smooth and moisture penetrates inland, erroneously dumping precipitation onto the southern dome. The ability of the VR grids to more accurately simulate orographic precipitation is consistent with the cloud results above. As the precipitation centers move from the interior toward the margins, and even out over the ocean with increasing resolution, the cloud decks move accordingly. Figures ??, ??, and ?? clearly illustrate a shift in clouds from the interior to the ocean perimeter with increasing resolution.

Table ?? shows the 1979-1998 climatological SMB components for each grid, compared with RACMO. The CESM values are averages over the ensemble of common ice masks and regridding methods described in section ??, and the RACMO values are averages over both RACMO datasets (Table ??) using the same common-ice-mask approach. Table ?? also contains (in parentheses) the SMB components derived from evaluating the integrals on each model’s native grid and ice mask. Of note is the large reduction in melt rates using the common-ice-mask approach compared to the native grid, illustrating the dissipation discussed in section ???. The errors are greatest in partially ice-covered grid cells straddling the ice sheet margins, in the ablation zone where melt rate are large. For integrated precipitation, the differences between the native and common-ice-mask approaches are much smaller, since the combined solid/liquid precipitation rates are not directly tied to the ice mask.

Figure ?? shows time series of annually integrated precipitation and snow/ice melt over the GrIS for the various different grids and dycores, with both versions of RACMO shown in black. The 1979-1998 climatological mean values, listed in Table ??, are shown as circles on the right side of the panels. The uniform  $1^\circ$ – $2^\circ$  grids have positive pre-

484 precipitation biases in the interior, whereas the VR grids have the smallest biases, with pre-  
 485 precipitation comparable to RACMO. The *f*19 and *f*09 grids perform similarly, with +110  
 486 Gt/yr bias, whereas *ne30pg3* is biased by about +165 Gt/yr and *ne30pg2* by +230 Gt/yr.  
 487 The larger biases on the uniform-resolution SE grids relative to the FV grids are con-  
 488 sistent with the coarser GrIS resolution on the SE grids (Figure ??).

489 The combined annual snow/ice melt shown in the bottom panel of Figure ?? in-  
 490 dicates that the *Arctic* grid simulates the most realistic melt rates, with the other grids  
 491 having more melt than RACMO. The *Arctic–GrIS* grid overpredicts melting by about  
 492 125 Gt/yr. This is likely due to an anomalously warm lower troposphere during the sum-  
 493 mer, relative to the *Arctic* run (Figure ??). The *f*19 and *f*09 melting rates are improved  
 494 over *Arctic–GrIS*, overestimating melt by only 70–90 Gt/yr. The SE grids have the  
 495 largest positive melt bias, between 200–220 Gt/yr. It is more difficult to attribute these  
 496 differences to resolution alone, since the FV grids have colder summer temperatures than  
 497 the uniform-resolution SE grids. (Not sure I understand this sentence. FV is both cooler  
 498 and higher-resolution than SE, so one might suspect that cooler T is, in fact, connected  
 499 to higher resolution.)

500 To illustrate the regional behavior of the SMB components, Figure ?? shows the  
 501 precipitation and combined snow/ice melt integrated over the basins defined by ? (?).  
 502 The uncertainty due to differences in basin area is larger than for GrIS-wide integrals,  
 503 owing to the differences in basin boundaries as represented by the common ice masks,  
 504 which are shown in the *f*19 and *ne30pg2* panels of Figure ???. Nonetheless, the regional  
 505 totals in Figure ?? correctly show the southeast and southwest basins have the most ac-  
 506 cumulation. In all basins, accumulations drops monotonically with increasing grid res-  
 507 olution, with some exceptions. The *Arctic–GrIS* grid simulates less precipitation than  
 508 RACMO in the central-east and southeast basins, and is closest of all grids to the RACMO  
 509 precipitation in the large southwest basin.

510 The basin-integrated melt rates in Figure ?? depend on the dycore. The uniform-  
 511 resolution SE grids have the largest positive biases in all basins. The *Arctic–GrIS* grid  
 512 is a close second, while the FV grids have systematically smaller melt-rates. The “second-  
 513 place” standing of *Arctic–GrIS* is somewhat unexpected, as this grid has the warmest  
 514 lower-troposphere summer temperatures (Figure ??) and greatest incident shortwave ra-  
 515 diation (Figure ??), yet it has less melting than the uniform-resolution SE grids.

516 Lower troposphere temperature is not a strict proxy for melting; e.g., it may not  
 517 capture microclimate effects as a result of a better representation of the low-elevation  
 518 ablation zones. Positive degree-days (PDD; ?, ?), which accumulate the near-surface tem-  
 519 perature in °C for days with temperature above freezing, are a more accurate proxy. PDD  
 520 is nonlinear in mean monthly temperature (? , ?). We compute it from monthly mean  
 521 2-meter temperature using the method of ? (?), assuming a fixed monthly mean stan-  
 522 dard deviation of 3°C and a degree-day factor of 5 mm d<sup>-1</sup> °C<sup>-1</sup>.

523 Figure ??c shows the basin-integrated PDD melt estimate. In the large southeast  
 524 and southwest basins (and all the other western basins), the *ne30pg3* grid has larger PDD-  
 525 based melt than the *ARCTCGRIS* grid. The FV grids also have large PDD-based melt  
 526 in the southwest basin, relative to *Arctic–GrIS*. The PDD plots indicate that the near-  
 527 surface temperatures which contribute to melt are not well approximated by the sum-  
 528 mer lower-troposphere temperatures in Figure ??.

529 Figure ?? presents the biases in the combined ice/snow melt as map plots. These  
 530 plots show that the largest melt biases are on the southeast and northwest coasts, where  
 531 large coarse-grid cells overlap with the ocean. One possibility is that these problematic  
 532 grid cells are situated at lower elevations than the true ice sheet surface, leading to a warm  
 533 bias and too much melt. Figure ?? shows the representation of the ice sheet surface along  
 534 two transects on the different grids, compared to the high-resolution dataset used to gen-

erate CAM topographic boundary conditions (?, ?, ?). The two transects are shown in Figure ??: the east-west “K-transect” in southwest Greenland and a transect extending from the central dome down to the Kangerlussuaq glacier on the southeast coast. The  $1^{\circ} - 2^{\circ}$  grids are noticeably coarse, with only a few grid cells populating the transect. The *f09* grid is a bit of an exception for the K-transect, with grid cells becoming narrow in the meridional direction at high latitudes. The VR grids are more skillful at reproducing the steep margins of the ice sheet, capturing the parabolic shape of the GrIS margins.

The transects in Figure ?? show that the ice sheet surface on the coarse grids is not systematically lower than the true surface in ablation zones. Rather, the smoothing of the raw topography, necessary to prevent the model from instigating grid-scale modes, flattens the ice sheet, causing the lower-elevation ablation zones to extend beyond the true ice sheet margin, where they lie above the actual ice surface. The *f19* grid has the smoothest topography since its dynamics are coarsest (whereas *f09*, *ne30pg2* and *ne30pg3* use identical smoothing), and has the flattest ice sheet. This suggests that if anything, coarser models will tend to elevate the ablation zones and depress melt rates.

Figure ?? also shows the ice margin boundary, illustrating that the ablation zone lies in a narrow horizontal band where the ice sheet rapidly plunges to sea-level. Due to this abrupt transition, coarse grids will commonly represent the ablation zone with grid cells containing mixtures of ice-covered and ice-free regions. We hypothesize that coarser models have larger melt biases because summer melting is confined to these mixed ice/land/ocean grid cells. CLM deals with land heterogeneity in a complex and sophisticated manner, but CAM only sees a homogenized state due to volume averaging over the sub-grid mixture. Thus, warm ice-free land patches in a grid cell may unduly influence the climate over the entire grid cell, causing a warm bias over the ice-covered patch. ([This is an interesting conclusion pointing to the need for better treatment of surface inhomogeneity in CAM. This might be a way to compensate for coarse resolution in future CESM versions?](#))

Figure ?? shows mean melt bias, relative to both RACMO datasets, conditionally sampled based on grid cell ice fraction in the GrIS region. Errors are computed using the common-ice-mask approach, meaning that all fields are mapped to the common masks, which define the grid cell ice fraction. The figure shows ([Any idea why the errors are smaller in the cells with intermediate ice fraction? I wouldn't have expected this.](#)) that coarser grids generally have two peaks in ice fraction space; a bump in positive melting errors in the 0-20% range, and another in the fully ice-covered cells. Also shown are the  $\pm 1$  standard deviation of the biases for each bin. They indicate that the biases in 0-20% bins are mostly contained in the positive bias region (a fractional bias greater than 0), whereas the fully-covered ice cells have a wider distribution, with many grid cells also containing negative melting biases. The excessive melting in the 0-20% ice fraction bins supports our hypothesis that the prevalence of mixed-grid cells in the ablation zone on coarse grids is responsible for their large melt bias.

### 576 3.4 Precipitation extremes

Synoptic storms are tracked using TempestExtremes atmospheric feature detection software (?, ?). As the *Arctic* grid contains  $\frac{1}{4}^{\circ}$  refinement north of about  $45^{\circ}$  latitude, the storm tracker is applied to this region for the *Arctic* and *ne30pg3* runs to identify differences in storm characteristics due to horizontal resolution. The composite mean precipitation maps are similar between the two grids, and exhibit the iconic comma structure of synoptic cyclones (not shown).

Figure ?? shows monthly PDFs of the precipitation rates associated with storms. The PDFs are constructed by sampling all the precipitation rates within  $30^{\circ}$  of the storm center, for each point on the storm track and for all storms. The PDFs are evaluated on

586 an identical composite grid for all runs, and so storm statistics are not impacted by dif-  
 587 fferences in output resolution. The *Arctic* run has larger extreme precipitation rates com-  
 588 pared to *ne30pg3* in every month, but the increase is greatest in the summer months,  
 589 which coincides with the most extreme events of the year. This is primarily due to in-  
 590 creased resolution and not the reduced physics times-step; the *ne30pg3\** run only marginally  
 591 increases the extreme precipitation rates compared with *ne30pg3* (Figure ??).

592 The extreme precipitation rates in the *Arctic* run are closer than *ne30pg3* to the  
 593 ERA5 reanalysis (Figure ??). It is difficult to know how much the extreme precipita-  
 594 tion rates in ERA5 are constrained by data assimilation, or whether these precipitation  
 595 rates are due to using a similar  $\frac{1}{4}^{\circ}$  model as the *Arctic* grid. However, it is well docu-  
 596 mented that  $\frac{1}{4}^{\circ}$  models are more skillful at simulating extreme events (?, ?, ?). This is  
 597 an additional benefit of the VR grids.

## 598 4 Conclusions

599 Running CESM2.2 in an AMIP-style configuration, we have evaluated six grids from  
 600 two dynamical cores for their performance over the Arctic and in simulating the Green-  
 601 land Ice Sheet (GrIS) surface mass balance (SMB). The  $1-2^{\circ}$  finite-volume grids have  
 602 enhanced resolution over polar regions due to the convergence of meridian lines, although  
 603 a polar filter is used to prevent spurious atmospheric features from forming at this higher  
 604 resolution. Spectral-element grids comparable to the resolution of the finite-volume grids  
 605 have an isotropic grid structure, meaning the grid resolution is similar over the entire  
 606 domain. We developed two VR grids and introduced them into CESM2.2 as part of this  
 607 work. Both use the spectral-element dycore; the *Arctic* grid has  $\frac{1}{4}^{\circ}$  refinement over the  
 608 broader Arctic, whereas the *Arctic–GrIS* grid is identical except for a  $\frac{1}{8}^{\circ}$  patch of re-  
 609 finement over Greenland.

610 In general, the FV grids have colder summer temperatures over the Arctic com-  
 611 pared with the SE grids (including the VR grids). The cloud biases in all the uniform-  
 612 resolution grids, whether FV or SE, are similar, in general being too cloudy over Arc-  
 613 tic land masses. The VR grids reduce the cloud biases. It should be emphasized that our  
 614 analysis is specific to the Arctic summer because of its relevance to GrIS melt rates; im-  
 615 proved clouds in the Arctic do not imply improved clouds at lower latitudes.

616 At the regional level, there is a halo of negative cloud bias (/colorblueI got con-  
 617 fused about signs here, because section 3 talks about a halo of positive cloud SW forc-  
 618 ing bias, which corresponds to a negative bias in cloud amounts. Maybe replace 'cloud  
 619 bias' with 'cloudiness bias' or something similar?) around the ocean perimeter of Green-  
 620 land on all  $1-2^{\circ}$  grids, but not the VR grids. This halo bias coincides with a positive  
 621 cloud bias over the ice sheet interior. This pattern has been traced to the inadequacy  
 622 of the coarser grids for resolving orographic precipitation. With overly smooth topog-  
 623 raphy on the  $1-2^{\circ}$  grids, synoptic systems moving into Greenland are not sufficiently  
 624 lifted when encountering the steep ice margins. As a result, moisture penetrates and dumps  
 625 excess precipitation into the GrIS interior, instead of being concentrated near the coastal  
 626 margins as shown by observations. This results in a positive precipitation and cloud bias  
 627 in the ice sheet interior, and a halo of low cloud bias about the perimeter. The agree-  
 628 ment of different observational data products on this bias lends confidence in the attri-  
 629 bution of causes. The VR grids compare better to the observations and show that oro-  
 630 graphic precipitation in Greenland is largely resolved.

631 We integrated the primary source and sink terms of the SMB equation over the GrIS  
 632 for each of the six grids. The uniform  $1-2^{\circ}$  grids have large positive accumulation bi-  
 633 ases because they fail to resolve orographic precipitation. The uniform SE grids have larger  
 634 accumulation biases, suggesting that the FV grids are more skillful for precipitation due

635 to finer resolution over Greenland, despite a polar filter. The VR grids have the most  
636 accurate accumulation rates of all the grids.

637 The primary mass sink term of the GrIS, ice/snow melt, has similar biases. The  
638 uniform resolution SE grids have too much melt, while the FV grids have smaller biases.  
639 It is difficult to attribute these biases to grid resolution alone. The FV grids have colder  
640 summers, consistent with their lower melt bias. However, the *Arctic–GrIS* grid has  
641 the warmest summer temperatures of all grids, yet it has less melting than the uniform-  
642 resolution SE grids. This suggests that grid resolution is responsible for a large fraction  
643 of the melt biases. We propose a mechanism: Coarse grids represent ablation zones us-  
644 ing grid cells with mixed surface types, ice-covered and ice-free. The warmer ice-free patches  
645 may largely determine the mean state, leading to a warm bias over the ice-covered patches  
646 of the grid cell. This mechanism is supported by analysis of melt biases binned by grid-  
647 cell ice fraction.

648 The *Arctic* grid substantially improves the simulated Arctic climate, including pre-  
649 cipitation extremes and the Greenland SMB, compared to the uniform  $1^\circ - 2^\circ$  grids.  
650 The *Arctic–GrIS* grid has the most realistic cloud and precipitation fields, but its sum-  
651 mer temperatures are too warm. The  $1^\circ$  FV model gives a surprisingly realistic SMB,  
652 likely due to the relatively fine resolution of Greenland on lat-lon grids. In particular,  
653 a greater number of grid cells in the ablation zone reduces the influence of mixed ice-  
654 covered/ice-free grid cells that represent ablation poorly on the other uniform-resolution  
655 grids.

656 As modeling systems move away from lat-lon grids towards quasi-uniform unstruc-  
657 tured grids, it is worth taking stock of whether this will degrade the simulated polar cli-  
658 mate. We have found that the  $1^\circ$  FV model has clear advantages over the  $1^\circ$  SE model  
659 in simulating the surface mass balance of the GrIS. This finding will not interrupt the  
660 ongoing transition towards unstructured grids in CESM, largely driven by gains in com-  
661 putational efficiency, but it has inspired us to develop alternative configurations that re-  
662 cover or improve on the fidelity of polar climate. We have shown here (and in a prior  
663 companion study (? , ?)) that for CESM, Arctic-refined meshes can substantially improve  
664 the simulated mass balances of the GrIS, even compared to the  $1^\circ$  grid. This should re-  
665 assure the CESM modeling community that the ongoing transition away from lat-lon grids  
666 will not adversely impact CESM’s usefulness as a state-of-the-art tool for simulating and  
667 understanding polar processes. (WHL: This last sentence may be too sanguine. Yes, we  
668 can recover the fidelity of Arctic simulations using VR grids, but (so far) only at a con-  
669 siderable cost in cpu-hours. This points to the need for an intermediate resolution that  
670 is more affordable, and/or model development or tuning that reduces the biases on coarse  
671 grids.)

672 We are working to develop a configuration of the *Arctic* grid that is fully-coupled  
673 with the CESM ocean and sea ice components and the Community Ice Sheet Model (CISM),  
674 to provide multi-century projections of the state of the GrIS and its contribution to sea-  
675 level rise. A fully-coupled pre-industrial control configuration has already been devel-  
676 oped and vetted. We have also developed a visualization of the *ARCTCGRIS* run, now  
677 available on youtube<sup>1</sup>. Figure ?? shows a snapshot of this visualization, illustrating mesoscale  
678 katabatic winds descending the southeastern slopes of GrIS. These new grids and con-  
679 figurations will provide new opportunities for CESM polar science. (WHL: I replaced  
680 the previous last sentence because it seemed too much like an advertisement. However,  
681 this new ending seems weak. I wonder if we should say something about future work mo-  
682 tivated by this study, for instance investigating grids and parameterizations that pro-  
683 vide some of the same benefits as these VR grids but at lower cost.)

---

<sup>1</sup> [https://www.youtube.com/watch?v=YwHgqDu75s8&t=4s&ab\\_channel=NCARVisLab](https://www.youtube.com/watch?v=YwHgqDu75s8&t=4s&ab_channel=NCARVisLab)

## 684 Appendix A Details on spectra-element dynamical core improvements 685 since the CESM2.0 release

686 **PHL working on this!**

687 Since the CESM2.0 release of the spectral-element dynamical core documented in  
688 ? (?) some important algorithmic improvements have been implemented and released  
689 with CESM2.2. These pertain mainly to the flow over orography that, for the spectral-  
690 element dynamical core, can lead to noise aligned with the element boundaries (? , ?).

### 691 A1 Reference profiles

692 Significant improvement in removing noise for flow over orography can be achieved  
693 by using reference profiles for temperature and pressure

$$T^{(ref)} = T_0 + T_1 \Pi^{(ref)}, \quad (\text{A1})$$

$$p_s^{(ref)} = p_0 \exp\left(-\frac{\Phi_s}{R^{(d)} T_0}\right) \quad (\text{A2})$$

694 (?, ?) where  $g$  gravity,  $T_0 \equiv 288K$ ,  $T_1 = \Gamma_0 T_0 c_p^{(d)} / g$  with standard lapse rate  $\Gamma_0 \equiv$   
695  $6.5K/km$  ( $c_p^{(d)}$  specific heat of dry air at constant pressure;  $R^{(d)}$  gas constant for dry air),  
696 and  $\Phi_s$  surface geopotential. The reference Exner function is

$$\Pi^{(ref)} = \left(\frac{p^{(ref)}}{p_0}\right)^\kappa \quad (\text{A3})$$

697 where  $\kappa = \frac{R_d}{c_p^{(d)}}$  (the reference surface pressure  $p_0 = 1000hPa$  and at each model level  
698 the reference pressure  $p^{(ref)}$  is computed from  $p_s^{(ref)}$  and the standard hybrid coefficients

$$p^{(ref)}(\eta) = A(\eta) + B(\eta)p_s^{(ref)}, \quad (\text{A4})$$

699 where  $A$  and  $B$  are the standard hybrid coefficients (using a dry-mass generalized ver-  
700 tical mass coordinate  $\eta$ ). These reference profiles are subtracted from the prognostic tem-  
701 perature and pressure-level-thickness states before applying hyperviscosity:

CESM2.0 → CESM2.2

$$\begin{aligned} \nabla_\eta T &\rightarrow \nabla_\eta \left(T - T^{(ref)}\right), \\ \nabla_\eta \delta p^{(d)} &\rightarrow \nabla_\eta \left(\delta p^{(d)} - \delta p^{(ref)}\right). \end{aligned}$$

702 This reduces spurious transport of temperature and mass up/down-slope due to hyper-  
703 viscosity.

### 704 A2 Rewriting the pressure gradient force (PGF)

705 In the CESM2.0 we used to used the following (standard) form of the pressure gra-  
706 dient term:

$$\nabla_\eta \Phi + \frac{1}{\rho} \nabla_\eta p, \quad (\text{A5})$$

707 where  $\Phi$  is geopotential and  $\rho = \frac{R_d T_v}{p}$  is density and  $\eta$  is the dry-mass vertical coor-  
708 dinate (for details see ?, ?). To alleviate noise for flow over orography we switched to  
709 an Exner pressure formulation, that is, we use that (??) can be written in terms of the  
710 Exner pressure

$$\nabla_\eta \Phi + c_p^{(d)} \theta_v \nabla_\eta \Pi, \quad (\text{A6})$$

711 where the Exner pressure is

$$\Pi \equiv \left(\frac{p}{p_0}\right)^\kappa. \quad (\text{A7})$$

712 The derivation showing that (??) and (??) are equivalent is shown here:

$$\begin{aligned}
 c_p^{(d)} \theta_v \nabla_\eta \Pi &= c_p^{(d)} \theta_v \nabla_\eta \left( \frac{p}{p_0} \right)^\kappa, \\
 &= c_p^{(d)} \theta_v \kappa \left( \frac{p}{p_0} \right)^{\kappa-1} \nabla_\eta \left( \frac{p}{p_0} \right), \\
 &= c_p^{(d)} \theta_v \kappa \Pi \left( \frac{p_0}{p} \right) \nabla_\eta \left( \frac{p}{p_0} \right), \\
 &= \frac{c_p^{(d)} \theta_v \kappa \Pi}{p} \nabla_\eta p, \\
 &= \frac{R_d \theta_v \Pi}{p} \nabla_\eta p, \\
 &= \frac{R_d T_v}{p} \nabla_\eta p, \\
 &= \frac{1}{\rho} \nabla_\eta p,
 \end{aligned}$$

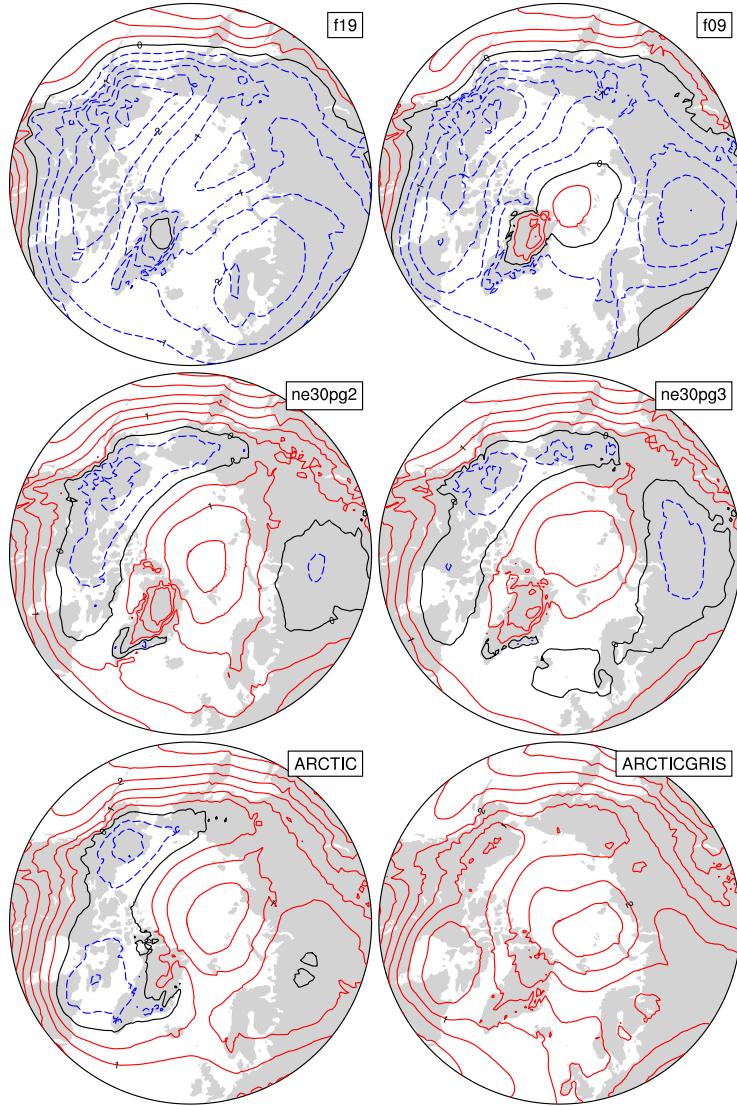
### 713 A3 Results

714 One year averages of *OMEGA*500 (vertical pressure velocity at 500hPa) has been  
 715 found to be a useful quantity to detect spurious up or down-drafts induced by steep orog-  
 716 raphy.

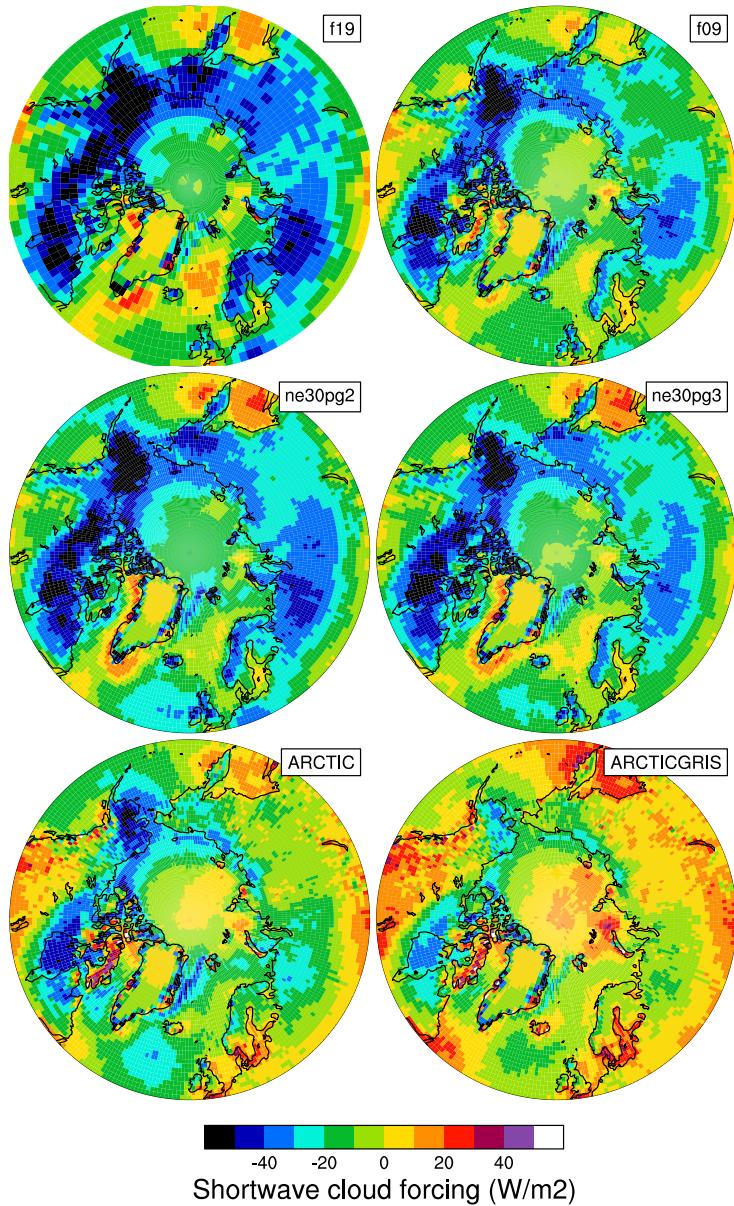
### 717 Acknowledgments

718 This material is based upon work supported by the National Center for Atmospheric Re-  
 719 search (NCAR), which is a major facility sponsored by the NSF under Cooperative Agree-  
 720 ment 1852977. Computing and data storage resources, including the Cheyenne super-  
 721 computer (doi:10.5065/D6RX99HX), were provided by the Computational and Infor-  
 722 mation Systems Laboratory (CISL) at NCAR. A. Herrington thanks Matt Rehme (NCAR/CISL)  
 723 for his role in generating the *Arctic–GrIS* visualization available on youtube ([https://www.youtube.com/watch?v=YwHgqDu75s8&t=4s&ab\\_channel=NCARVisLab](https://www.youtube.com/watch?v=YwHgqDu75s8&t=4s&ab_channel=NCARVisLab)).

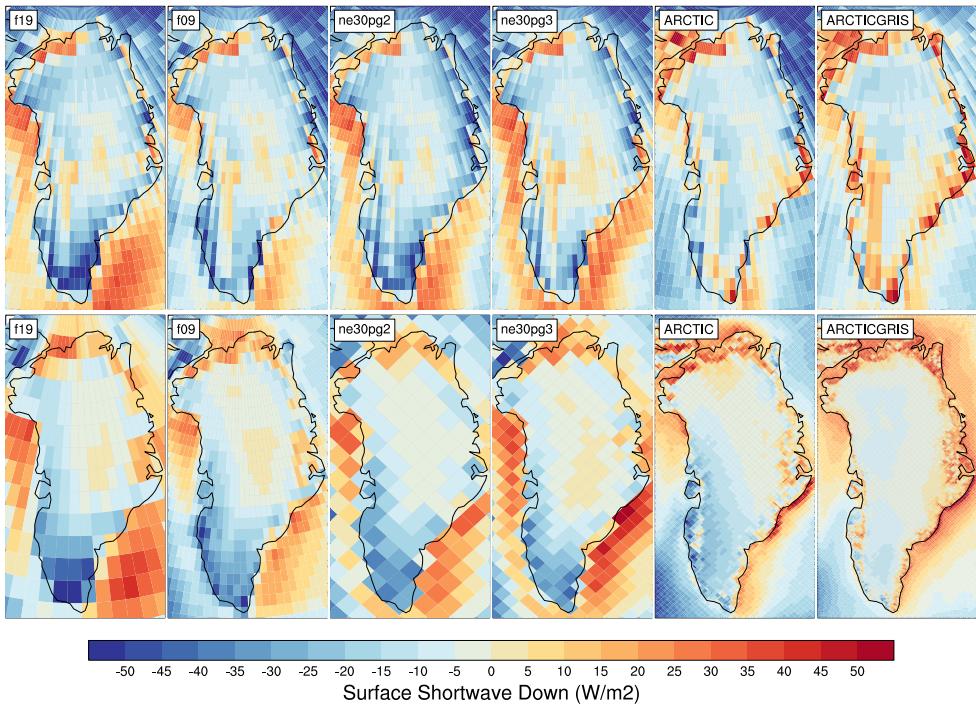
724 The data presented in this manuscript is available at <https://github.com/adamrher/2020-arcticgrids>.



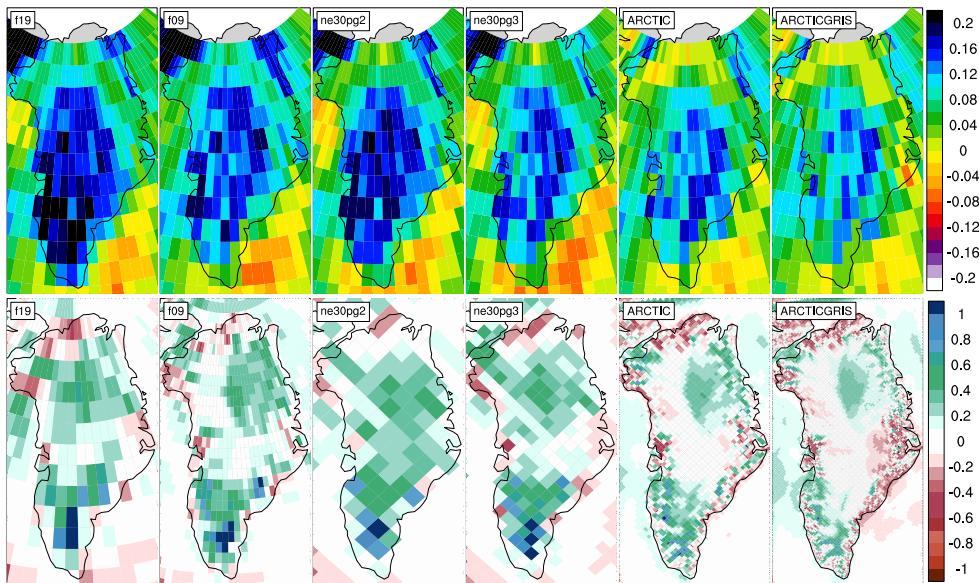
**Figure 6.** 1979-1998 lower troposphere, northern hemisphere summer virtual temperature biases, computed as the difference from ERA5. Lower troposphere layer mean virtual temperature is derived from the 1000 hPa - 500h Pa geopotential thickness, using the hypsometric equation. Differences are computed after mapping the ERA5 data to the finite-volume grids since the geopotential field is only available on the output tapes in the spectral-element runs that have been interpolated to the *f09* grid, inline.



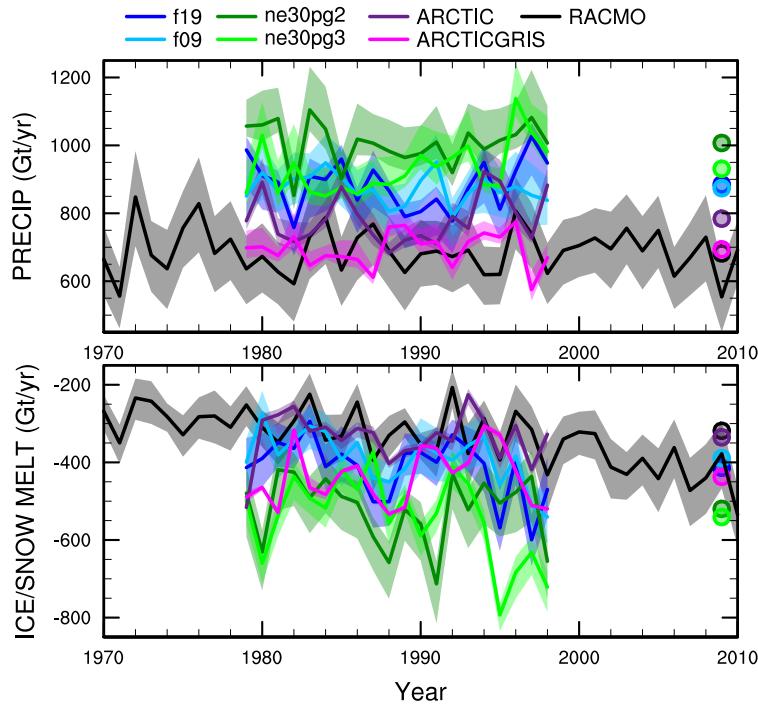
**Figure 7.** 1979-1998 northern hemisphere summer shortwave cloud forcing bias, relative to the CERES-EBAF gridded dataset. Differences are computed after mapping all model output to the  $1^{\circ}$  CERES-EBAF grid.



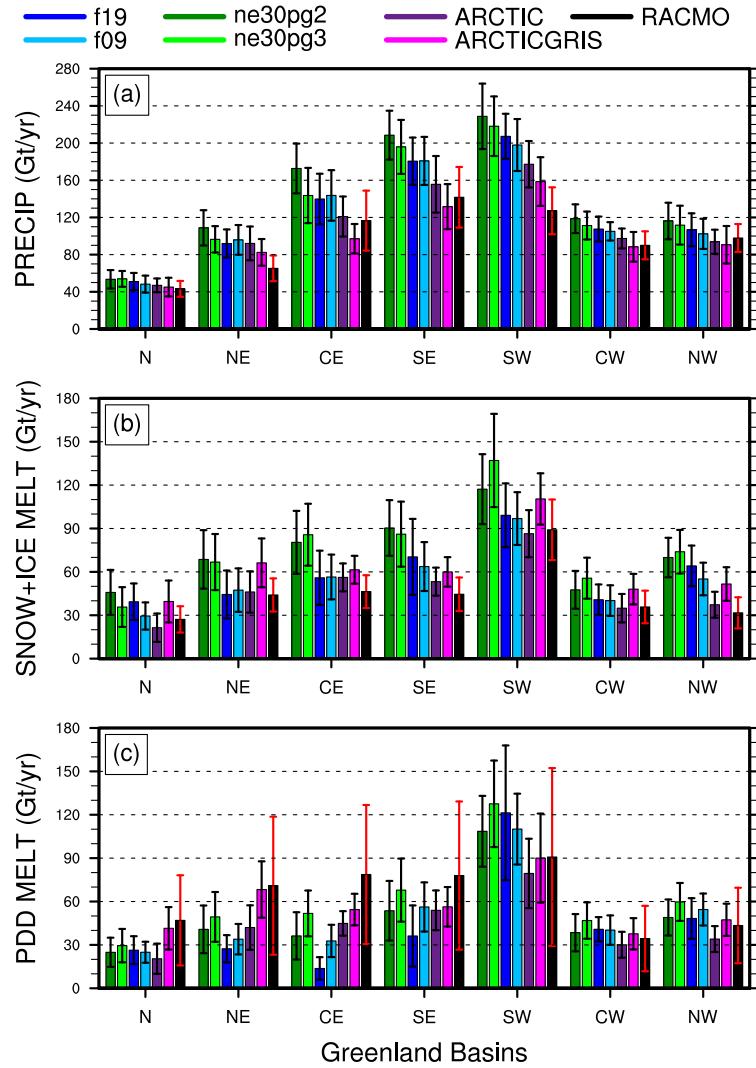
**Figure 8.** 1979-1998 northern hemisphere summer, incident shortwave radiation bias, computed as the difference (top) from CERES-EBAF, and (bottom) RACMO2.3p2 dataset. The differences in the top panel are found by mapping the model output to the 1° CERES-EBAF grid, and differences on the bottom panel are computed after mapping the RACMO2.3p2 dataset to the individual model grids. Note that the averaging period for the CERES-EBAF panels, 2003-2020, is different from the averaging period for the model results.



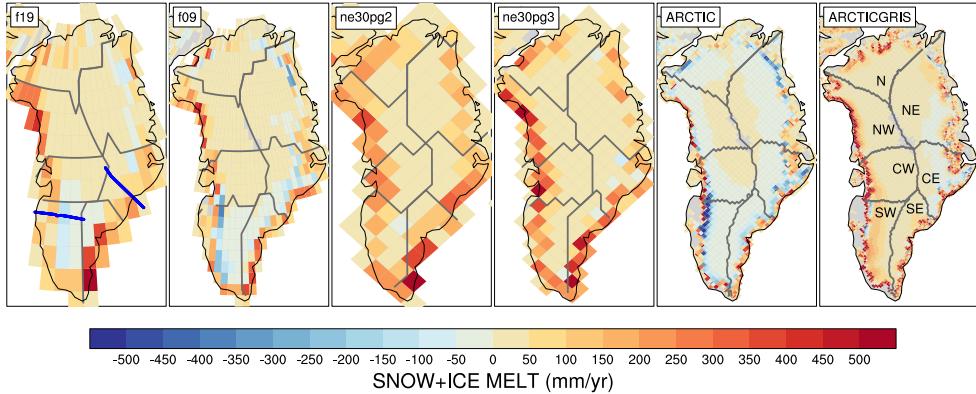
**Figure 9.** 1979-1998 northern hemisphere summer (top) total cloud fraction bias, relative to the CALIPSO-GOCCP dataset, and (bottom) precipitation rate bias, relative to the RACMO2.3p2 dataset. The CALIPSO-GOCCP differences are computed after mapping all model output to the  $1^{\circ}$  grid, whereas the RACMO differences are computed after mapping the RACMO dataset to the individual model grids. Note that the averaging period for the CALIPSO-GOCCP panels, 2006-2017, is different than the model averaging period. **ARH** - still trying to fix the layout of this figure so the label bars aren't on top of the panels.



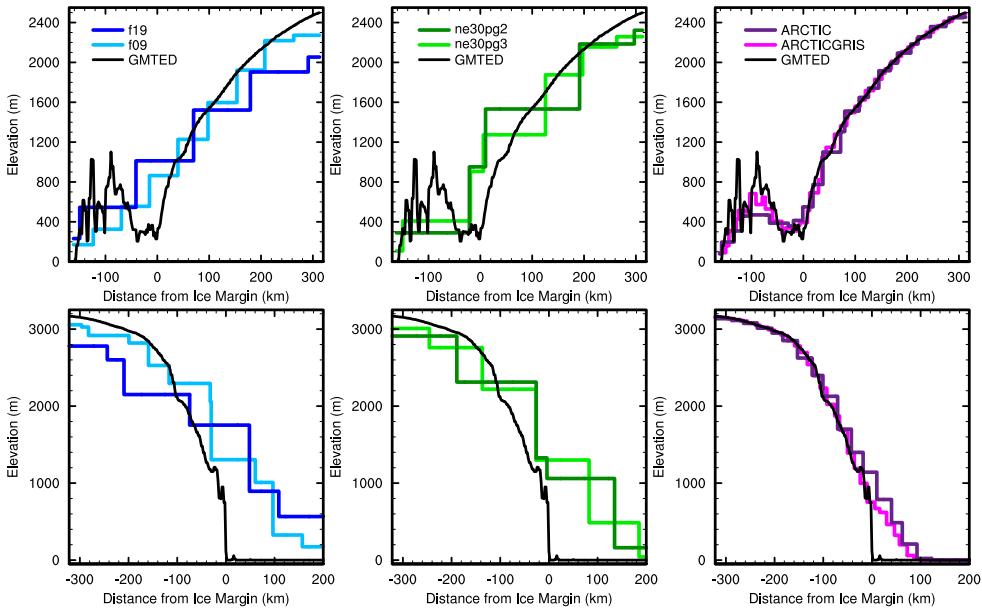
**Figure 10.** Time-series of annual (solid+liquid) precipitation (top) and annual runoff (bottom) integrated over the Greenland Ice Sheet for all six simulations and compared to the RACMO datasets. The time-series were generated using the common ice mask approach, which results in up to 4 ensembles, with the mean value given by the solid line and shading spanning the extent of the ensemble members.



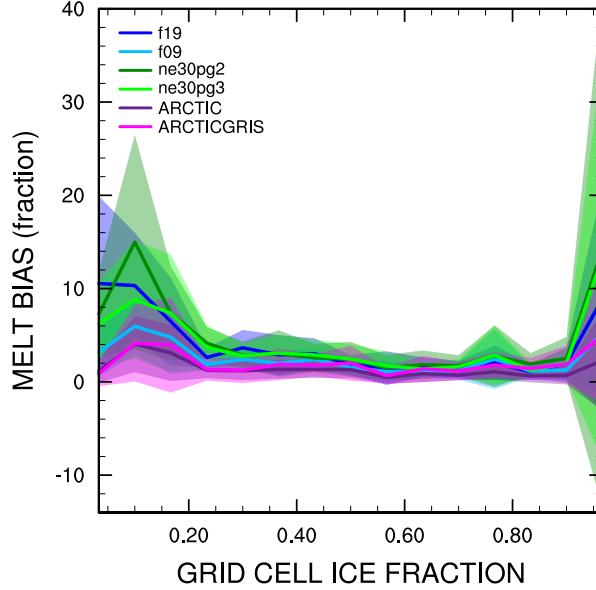
**Figure 11.** 1979–1998 basin integrated components of the SMB; (top) precipitation, (middle) ice/snow melt and (bottom) ice/snow melt estimated from the PDD method. Whiskers span the max/min of the four ensemble members generated from the common-ice-mask approach. Basin definitions are after ?, and are found on the common ice masks using a nearest neighbor approach, and shown in Figure ??.



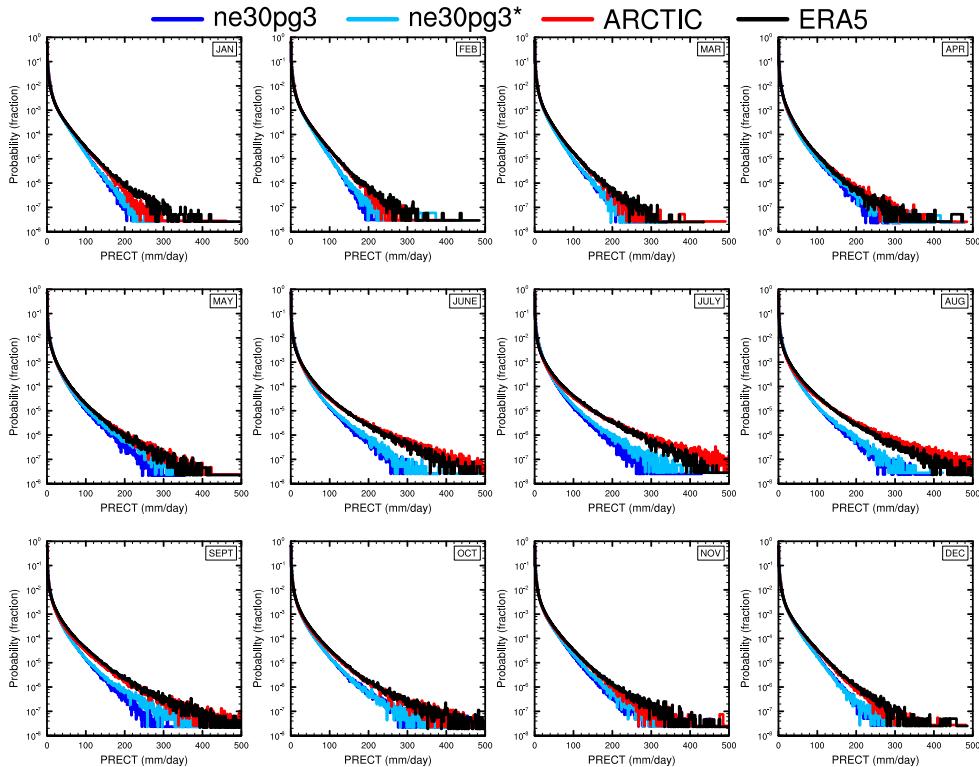
**Figure 12.** 1979–1998 ice/snow melt biases (in mm/yr) relative to RACMO2.3p2, evaluated on the native model grids. The ? (?) basin boundaries are shown in grey for each model grid. Note that Figure ?? uses the basin boundaries for the two common ice masks, shown in the *f19* and *ne30pg2* panels, in computing the basin-scale integrals. Blue lines in the *f19* panel show the location of the two transects plotted in Figure ??.



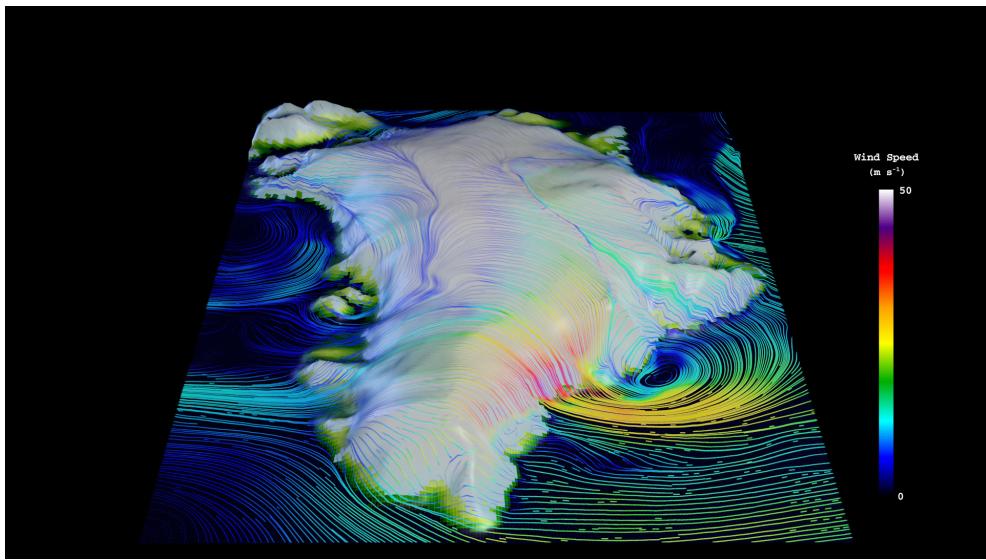
**Figure 13.** Model surface elevation along the (top) K-transect, and (bottom) a transect spanning the central dome down to the Kangerlussuaq glacier in southeast Greenland, for all model grids. The reference surface (GMTED) is a 1 km surface elevation dataset used for generating the CAM topographic boundary conditions.



**Figure 14.** Fractional melt bias over the GrIS, computed relative to the RACMO datasets using the common ice mask approach, and conditionally sampled by grid cell ice fraction provided by the common ice masks. Solid lines are the mean of the distribution with  $\pm$  one standard deviation expressed by shading.



**Figure 15.** PDFs of the total precipitation rate associated with tracked storms, by month, in the *ne30pg3*, *ne30pg3\** and *Arctic* runs, and compared with the ERA5 dataset.



**Figure 16.** Snapshot of the lowest model level streamlines from the *Arctic – GrIS* visualization, with color shading denoting the wind magnitude.