

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

Adam R. Herrington¹, Peter H. Lauritzen¹, Marcus Lofverstrom², William H. Lipscomb¹, Andrew Gettelman¹ and Mark A. Taylor³

¹National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder, Colorado, USA

²Department of Geosciences, University of Arizona, 1040 E. 4th Street, Tucson, Arizona USA

³Sandia National Laboratories, Albuquerque, New Mexico, USA

Key Points:

- The CESM2.2 release includes several enhancements to the spectral-element dynamical core, including two new Arctic refined mesh configurations.
- Uniform resolution grids degrade the surface mass balance of the Greenland Ice Sheet compared with equivalent low resolution latitude-longitude grids.
- The refined Arctic meshes substantially improve the surface mass balance over all low resolution grids.

15 **Abstract**

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

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* (Williamson, 2007). The pole
 36 problem refers to numerical instability arising from the convergence of meridian lines into
 37 polar singularities on latitude-longitude grids (e.g., Figure 1a, hereafter referred to as
 38 *lat-lon* grids). Depending on the numerics, methods exist to suppress this instability, and
 39 lat-lon grids may be advantageous for polar processes by representing structures with
 40 finer resolution than elsewhere in the computational domain. With the recent trend to-
 41 wards globally uniform unstructured grids, any potential benefits of lat-lon grids in po-
 42 lar regions may be lost. In this study we evaluate a number of grids and dynamical cores
 43 (hereafter referred to as *dycores*) available in the Community Earth System Model, ver-
 44 sion 2.2 (CESM2.2; Danabasoglu et al., 2020), including new variable-resolution grids,
 45 to understand their impacts on the simulated Arctic climate. We focus specifically on
 46 the climate and surface mass balance of the Greenland Ice Sheet.

47 In the 1970s, the pole problem was largely defeated through the adoption of effi-
 48 cient spectral transform methods in GCMs (see Williamson, 2007, and references therein).
 49 These methods transform grid point fields into a global, isotropic representation in wave
 50 space, where linear operators (e.g., horizontal derivatives) in the (truncated) equation
 51 set can be solved exactly. While spectral transform methods are still used today, local
 52 numerical methods have become desirable for their ability to run efficiently on massively
 53 parallel systems. The pole problem has thus re-emerged in contemporary climate mod-
 54 els that use lat-lon grids, and some combination of reduced grids (modified lat-lon grids,
 55 with cells elongated in the longitudinal direction over the polar regions) and polar fil-
 56 ters are necessary to ameliorate this instability (Jablonowski & Williamson, 2011). Po-
 57 lar filters subdue the growth of unstable modes by applying additional damping to the
 58 solution over polar regions. This damping reduces the effective resolution in polar re-
 59 gions such that the resolved scales are *approximately* the same everywhere on the grid.
 60 We emphasize *approximately* since it's at least conceivable that marginal increases in ef-
 61 fective resolution occur over polar region in lat-lon grids, despite polar filtering, since
 62 resolved waves can be represented with more grid points than at lower latitudes.

63 Dycores built on lat-lon grids have some advantages over unstructured grids. Lat-
 64 lon coordinate lines are orthogonal, and aligned with zonally symmetric circulations that
 65 characterize many large-scale features of Earth's atmosphere. Lauritzen et al. (2010) has

66 experimented with rotating lat-lon models such that their coordinate lines no longer align
 67 with an idealized, zonally balanced circulation. For the finite-volume lat-lon dycore con-
 68 sidered in this paper (hereafter *FV*), numerical errors were shown to be largest when
 69 the polar singularity is rotated into the baroclinic zone (45°N latitude), generating spu-
 70 rious wave growth much earlier on in the simulation than for other rotation angles. This
 71 illustrates the advantages of coordinate surfaces aligned with latitude bands, albeit an
 72 extreme example where the polar singularity and the polar filter are also contributing
 73 to the spurious wave growth. The unstructured grids all generate spurious baroclinic waves
 74 earlier on in the simulations than the (unrotated) lat-lon models, although the unstruc-
 75 tured model considered in this paper, the spectral-element dycore (hereafter *SE*), holds
 76 a balanced zonal flow without wave growth appreciably longer than the rotated FV ex-
 77 periments, regardless of how the SE grid is rotated (Lauritzen et al., 2010).

78 The polar filter in the FV model impedes efficiency at large processor (CPU) counts
 79 because it requires a spectral transform, which have large communication overhead (Suarez
 80 & Takacs, 1995; Dennis et al., 2012). Unstructured grids support quasi-uniform grid spac-
 81 ing globally, and there is no pole-problem (e.g., Figure 1c). Unstructured grids are be-
 82 coming increasingly common due to their improved performance on massively parallel
 83 systems and lack of constraints on grid structure (Putman & Lin, 2007; Dennis et al.,
 84 2012; Wan et al., 2013). This grid flexibility allows for variable-resolution grids (e.g., Fig-
 85 ure 2; hereafter abbreviated as *VR*), sometimes referred to as regional grid refinement.
 86 In principle, grid refinement over polar regions can make up for any loss of resolution
 87 in transitioning away from lat-lon grids (e.g., Figure 2). However, refinement comes at
 88 the cost of a smaller CFL-limited time step in the refined region. (The CFL-condition
 89 — short for Courant–Friedrichs–Lewy condition — is a necessary condition for numer-
 90 ical stability when using discrete data in time and space.)

91 It is important to emphasize that the pole-problem is a feature of the dycore. Po-
 92 lar filters do not directly interfere with the physical parameterizations, nor do they have
 93 any bearing on the surface models; e.g., the land model can take full advantage of the
 94 greater number of grid cells in polar regions on lat-lon grids. This is particularly rele-
 95 vant for the surface mass balance of the Greenland Ice Sheet (*SMB*; the integrated sum
 96 of precipitation and runoff), which relies on hydrological processes handled by the land
 97 model.

98 The SMB of the Greenland Ice Sheet (hereafter *GrIS*) is determined by processes
 99 occurring over a range of scales that are difficult to represent in GCMs (Pollard, 2010).
 100 GrIS precipitation is concentrated at the ice-sheet margins, where steep topographic slopes
 101 drive orographic precipitation. The truncated topography used by low resolution GCMs
 102 enables moisture to penetrate well into the GrIS interior, manifesting as a positive pre-
 103 cipitation bias (Pollard & Groups, 2000; van Kampenhout et al., 2018). GrIS ablation
 104 areas (marginal regions where seasonal melting exceeds the annual mass input from pre-
 105 cipitation) are typically less than 100 km wide and are confined to low-lying areas or re-
 106 gions with low precipitation. These narrow ablation zones are not fully resolved in low
 107 resolution GCMs, and may further degrade the simulated SMB. For example, CESM,
 108 version 2.0 (CESM2) underestimates ablation in the northern GrIS, leading to unreal-
 109 istic ice advance when run with an interactive ice sheet component (Lofverstrom et al.,
 110 2020).

111 Regional climate models (RCMs) are commonly relied upon to provide more ac-
 112 curate SMB estimates. The limited area domain used by RCMs permits the use of high
 113 resolution grids able to resolve the SMB processes, and can skillfully simulate the GrIS
 114 SMB (Box et al., 2004; Rae et al., 2012; Van Angelen et al., 2012; Fettweis et al., 2013;
 115 Mottram et al., 2017; Noël et al., 2018). However, unlike GCMs, RCMs are not a freely
 116 evolving system and boundary conditions must be prescribed at the lateral boundaries.
 117 The inability of the RCM solution to influence larger-scale dynamics outside the RCM
 118 domain (due to the prescribed boundary conditions) severely limits this approach from

properly representing the role of the GrIS in the climate system. In addition, the boundary conditions are derived from a separate host model, which introduces inconsistencies due to differences in model design between the host model and the RCM.

In order to retain the benefits of RCMs in a GCM, van Kampenhout et al. (2018) utilized the VR capabilities of the SE dycore in CESM, generating a grid where Greenland is represented with $1/4^\circ$ resolution, and elsewhere with the more conventional 1° resolution. The simulated SMB compared favorably to the SMB from RCMs and observations. The VR approach is therefore emerging as a powerful tool for simulating and understanding the GrIS and its response to different forcing scenarios, in the freely evolving GCM framework.

The SE dycore has been an available option since CESM, version 1, but has been under active development ever since. This includes the switch to a dry-mass vertical coordinate (Lauritzen et al., 2018) and incorporation of an accelerated multi-tracer transport scheme (Lauritzen et al., 2017), made available in CESM2. This paper documents several additional enhancements to the SE dycore as part of the release of CESM2.2. These include three new VR configurations (Figure 2), two Arctic meshes and a Contiguous United-States mesh (CONUS; featured in Pfister et al. (2020)). While there are dozens of published studies using VR in CESM (e.g., Zarzycki et al., 2014; Rhoades et al., 2016; Gettelman et al., 2017; Burakowski et al., 2019; Bambach et al., 2021), these studies either used development code or collaborated closely with model developers. CESM2.2 is the first code release that contains out of the box VR functionality in CESM.

This study compares the representation of Arctic regions using the SE and FV dycores in CESM2.2 (see description below), as these two dycores treat high latitudes (i.e., the pole problem) in different ways. Section 2 documents the grids, dycores, and physical parameterizations used in this study, and also describes the experiments, datasets, and evaluation methods. Section 3 analyzes the results of the experiments, and Section 4 provides a general discussion and conclusions.

2 Methods

2.1 Dynamical cores

The atmospheric component of CESM2.2 (Danabasoglu et al., 2020), the Community Atmosphere Model, version 6.3 (CAM; Gettelman et al., 2019; Craig et al., 2021), supports several different atmospheric dynamical cores. These include dycores on lat-lon grids, such as finite-volume (FV; Lin, 2004) and Eulerian spectral transform (EUL; Collins et al., 2006) models, and dycores built on unstructured grids, including spectral-element (SE; Lauritzen et al., 2018) and finite-volume 3 (FV3; Putman & Lin, 2007) models. This study compares the performance of the SE and FV dycores, omitting the EUL and FV3 dycores. CESM2 runs submitted to the Coupled Model Intercomparison Project Phase 6 (Eyring et al., 2016) are based on the FV dycore, whereas the SE dycore is often used for global high-resolution simulations (e.g., Small et al., 2014; Bacmeister & Coauthors, 2018; Chang et al., 2020) due to its higher throughput on massively parallel systems (Dennis et al., 2012).

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 (Lin & Rood, 1997). The 2D dynamics evolve in floating Lagrangian layers that are periodically mapped to an Eulerian reference grid in the vertical (Lin, 2004), 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* (Lauritzen

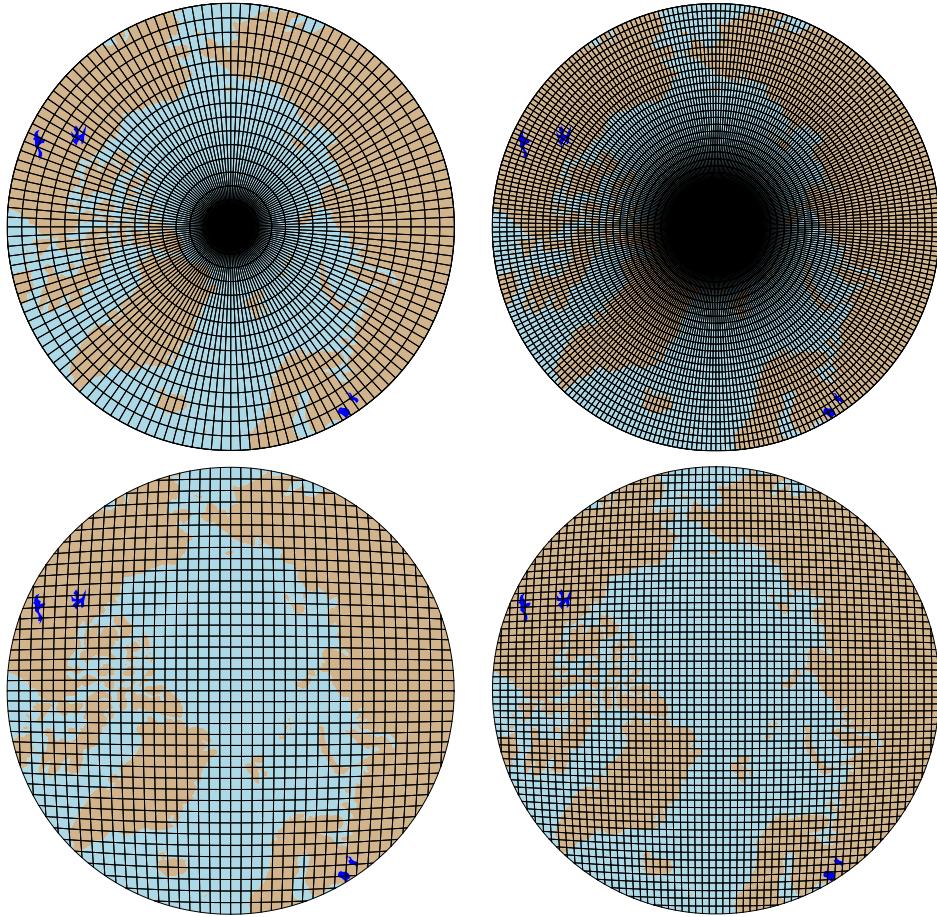


Figure 1. Computational grids for the uniform $1^\circ - 2^\circ$ grids in this study. Grids names after Table 1, (a) f19, (b) f09, (c) ne30pg2 and (d) ne30pg3.

et al., 2011) [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, permitting a more practical time step. It takes the form of a Fourier filter in the zonal direction, with the damping coefficients increasing monotonically in the meridional direction (Suarez & Takacs, 1995). The form of the filter is designed to slow down the propagation of large zonal wave-numbers to satisfy the CFL condition of the shortest resolved wave at some reference latitude.

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 (Taylor et al., 1997; Taylor & Fournier, 2010). The computational domain is a cubed-sphere grid tiled with quadrilateral elements (see Figure 2). 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 quadrature points. This results in 16 nodal points per element, with 12 of the points lying on the (shared) element boundary. Communication between elements uses the direct stiffness summation (Canuto et al., 2007), which applies a numerical flux to the el-

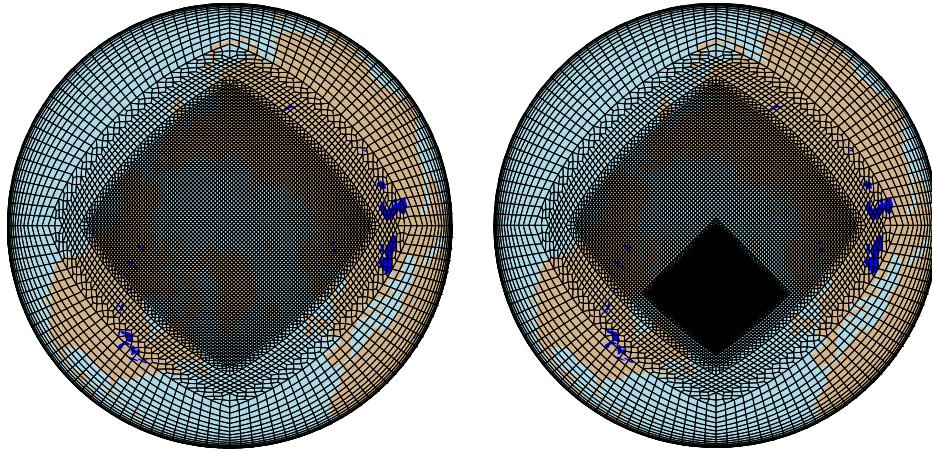


Figure 2. Spectral-element grid for the VR grids in this study, (a) Arctic, (b) Arctic – GrIS and (a) CONUS. Note that this is not the computational grid; each element has 3×3 independent grid points. [add CONUS grid](#)

ement 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 (Lauritzen et al., 2018). The 2D dynamics have no implicit dissipation, and so hyperviscosity operators are applied to all prognostic variables to remove spurious numerical errors (Dennis et al., 2012). Laplacian damping is applied in the sponge layer.

The SE dycore supports regional grid refinement via its VR configuration, requiring two enhancements over uniform-resolution setups. (1) As the numerical viscosity increases with resolution, explicit hyperviscosity relaxes according to the local element size, reducing in strength by an order of magnitude per halving of the grid spacing. A tensor-hyperviscosity formulation is used (Guba et al., 2014), which adjusts the coefficients in two orthogonal directions to more accurately target highly distorted quadrilateral elements. (2) The topography boundary conditions are smoothed in a way that does not excite grid scale modes, and so the NCAR topography software (Lauritzen et al., 2015) has been modified to scale the smoothing radius by the local element size, resulting in rougher topography in the refinement zone.

For SE grids with quasi-uniform grid spacing, The SE tracer transport scheme is replaced with the Conservative Semi-Lagrangian Multi-tracer transport scheme (CSLAM) (Lauritzen et al., 2017). CSLAM has improved tracer property preservation and accelerated multi-tracer transport. It uses a separate grid from the spectral-element dynamics, dividing each element into 3×3 control volumes with quasi-equal area. The physical parameterizations are computed from the state on the CSLAM grid, which has clear advantages over the default SE dycore in which the physics are evaluated Gauss-Lobatto-Legendre points (Herrington et al., 2018). [Rene/Andrew - maybe elaborate further on the advantages of CSLAM?](#)

grid name	dycore	Δx_{eq} (km)	Δx_{refine} (km)	Δt_{phys} (s)
f19	FV	278	-	1800
f09	FV	139	-	1800
ne30pg2	SE-CSLAM	167	-	1800
ne30pg3	SE-CSLAM	111	-	1800
ne30pg3*	SE-CSLAM	111	-	450
Arctic	SE	111	28	450
Arctic – GrIS	SE	111	14	225

Table 1. Grids and dycores used in this study. Δx_{eq} is the average equatorial grid spacing, Δx_{refine} is the grid spacing in the refined region (if applicable), and Δt_{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 Δt_{phys} . *Rene - maybe include model cost?*

211

2.2 Grids

212

We evaluate model simulations on six different grids in this study (Table 1). The
 213 FV dycore is run with nominal 1° and 2° grid spacing, referred to as f09 and f19, re-
 214 spectively (Figure 1a,b). We also run the 1° equivalent of the SE-CSLAM grid, referred
 215 to as ne30pg3 (Figure 1c), where ne refers to a grid with $ne \times ne$ elements per cubed-
 216 sphere face, and pg denotes that there are $pg \times pg$ control volumes per element for com-
 217 puting the physics. We run an additional 1° SE-CSLAM simulation with the physical
 218 parameterizations computed on a grid with 2×2 control volumes per element, ne30pg2
 219 (Figure 1d; Herrington et al., 2019, note CSLAM is still run on the 3×3 control vol-
 220 ume grid).

221

Three VR meshes were developed for the CESM2.2 release to support grid refinement
 222 over the Arctic and CONUS (Figure 2). This paper serves as the official documen-
 223 tation of these grids. The Arctic meshes were developed using the software package SQuad-
 224 gen (<https://github.com/ClimateGlobalChange/squadgen>). The Arctic grid is a 1°
 225 grid with $1/4^{\circ}$ regional refinement over the broader Arctic region. The Arctic–GrIS
 226 grid is identical to the Arctic grid, but with an additional patch covering the island of
 227 Greenland with $1/8^{\circ}$ resolution. The CONUS grid contains $1/8^{\circ}$ refinement over the United
 228 States, and 1° everywhere else. The CONUS grid is not discussed any further in this pa-
 229 per; see Pfister et al. (2020) for simulations using the CONUS grid.

230

The physics time step depends on the grid resolution. Increased horizontal reso-
 231 lution permits faster vertical velocities that reduce characteristic time scales, so the physics
 232 time step is reduced to avoid large time truncation errors (Herrington & Reed, 2018).
 233 The Arctic and Arctic – GrIS grids are run with a $4 \times$ and $8 \times$ reduction in physics
 234 time step relative to the default 1800 s time step used in the 1° and 2° grids (Table 1).

235

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

238

2.3 Physical parameterizations

239

All simulations in this study use the CAM6.3 physical parameterization package
 240 (hereafter referred to as the *physics*; Gettelman et al., 2019). The physics in CAM6 dif-
 241 fers from its predecessors through the incorporation of high-order turbulence closure, Cloud
 242 Layers Unified by Binormals (CLUBB; Golaz et al., 2002; Bogenschutz et al., 2013), which
 243 jointly acts as a planetary boundary layer, shallow convection, and cloud macrophysics
 244 scheme. CLUBB is coupled with the MG2 microphysics scheme (Gettelman & Morri-

son, 2015; Gettelman et al., 2015), 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 (Zhang & McFarlane, 1995; Neale et al., 2008) and includes convective momentum transport (Richter et al., 2010). Boundary layer form drag is modeled after Beljaars et al. (2004), and orographic gravity wave drag is represented with an anisotropic method informed by the orientation of topographic ridges at the sub-grid scale (the ridge orientation is derived from a high-resolution, global topography dataset (J. J. Danielson & Gesch, 2011)).

Initial simulations with the `ne30pg3` SE grid produced weaker shortwave cloud forcing relative to the tuned finite-volume dycore. The SE dycore in CESM2.2 therefore has two CLUBB parameter changes to provide more realistic cloud forcing and top-of-atmosphere radiation balance. We reduced the width of the sub-grid distribution of vertical velocity (`clubb_gamma` = 0.308 → 0.270) and also reduced the strength of the damping for horizontal component of turbulent energy (`clubb_c14` = 2.2 → 1.6) to increase cloudiness. For a thorough explanation of how CLUBB parameters impact the simulated climate, see (Guo et al., 2015).

2.4 Modeled SMB

All grids and dycores simulate the GrIS Surface Mass Balance (SMB). The SMB is the sum of mass accumulation of precipitation and mass loss from 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 or snow/ice melt runs off the ice sheet; this water can penetrate pore spaces in the snowpack/firn layer and freeze, increasing the ice mass. These processes are represented by different CESM components, but it is the Community Land Model, version 5 (CLM; Lawrence et al., 2019), that aggregates these processes and computes the SMB.

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 temperatures between $T < -2^{\circ}$ C, as rain for $T > 0^{\circ}$ C, and as a linear combination of rain and snow for temperatures between -2° C and 0° 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 water 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 Lipscomb et al. (2013); Sellevold et al. (2019); van Kampenhout et al. (2020); Muntjewerf et al. (2021). **Marcus - mention you're not running CISIM.**

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.

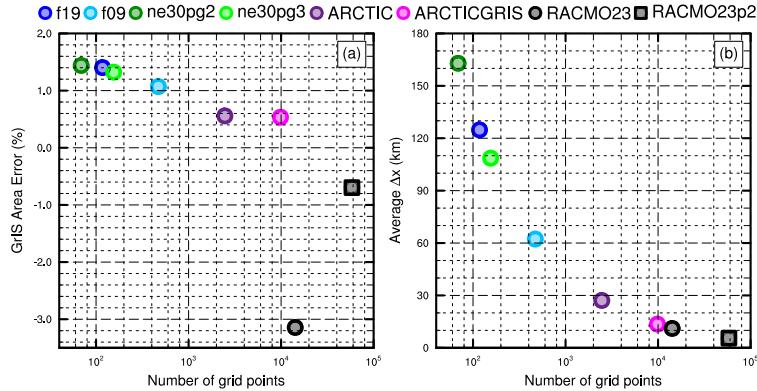


Figure 3. The spatial properties of the GrIS as represented by different grids in this study. (Left) GrIS area error, computed as the relative differences from a 4km dataset used to create the CESM ice masks, (right) approximate average grid spacing over GrIS.

294

2.5 SMB Analysis

295
296
297
298
299
300
301
302
303

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 3). The area errors in Figure 3 may not seem large, but even 1–2% area differences can lead to large differences in integrated SMB (Hansen et al., 2022). 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 (ESMF) first-order conservative remapping algorithm (Team et al., 2021). The figure suggests that the mapping errors are less than 1.5% across the CESM2 grids.

304
305
306
307
308
309
310
311
312
313
314

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 (Ullrich & Taylor, 2015) 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.

315
316
317
318

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:

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

319
320
321
322
323

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

324 expression is different from the internally computed SMB in CLM (van Kampenhout et
 325 al., 2020).

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

- 328 1. Map the grid-cell mean quantities to the common grid, and integrate the mapped
 329 fields over the common ice masks.
- 330 2. Map the patch-level quantities (i.e., the state over the ice fractional component
 331 of the grid cell) to the common grid, and integrate the mapped fields over the com-
 332 mon ice masks.

333 Note that we are mapping to low-resolution grids that have larger GrIS areas (Figure
 334 3). Since the components of equation 1 are not confined to the ice mask, method 1
 335 reconstructs the SMB over the portion of the common ice mask that is outside the ice
 336 mask on the source grid. While this may be an acceptable way to reconstruct the mass
 337 source terms over different ice masks, ice melt is zero outside the source ice mask, and
 338 so method 1 will underestimate the mass sink term. This underestimation is systematic
 339 in method 2; all variables are exclusive to the ice mask. Method 2 is more dissipative,
 340 since mapping to the low-resolution grid will average a field of non-zero values over the
 341 ice mask with a field of zeros outside the ice mask. However, patch-level values for pro-
 342 cesses exclusive to the ice mask (e.g., ice melt) will on average have larger magnitudes
 343 than the the grid-mean quantities used in method 1.

344 The different error characteristics of the two methods are used to diversify the en-
 345 semble. Each of the four regridding combinations (with conservative and high-order remap-
 346 ping to the f09 and ne30pg2 grids) are repeated with each method, resulting in (up to)
 347 eight values for each integrated quantity. Unfortunately, the patch-level values of evap-
 348 oration/sublimation are not available from the model output, and we estimate their con-
 349 tribution by zeroing out the field for grid cells that have no ice, prior to mapping to the
 350 common ice mask. This will degrade the SMB estimates using method 2, but note that
 351 we are more interested in the behavior of the processes across grids and dycores, expressed
 352 as the components of the SMB, rather than the SMB itself.

353 2.6 Experimental design

354 All simulations described here use an identical transient 1979-1998 Atmospheric
 355 Model Inter-comparison Project (AMIP) configuration, with prescribed monthly sea-surface
 356 temperature and sea ice following Hurrell et al. (2008). In CESM terminology, AMIP
 357 simulations use the FHIST computational set and run out of the box in CESM2.2.

358 2.7 Observational Datasets

359 We use several observational datasets (Table 2) to assess the performance of the
 360 simulations. We gathered SMB datasets from several sources. Regional Atmospheric Cli-
 361 mate Model, version 2.3 11km (RACMO23; Noël et al., 2015) and version 2.3p2 5.5km
 362 (RACMO2.3p2; Noël et al., 2018, 2019) are RCM simulations targeting Greenland, forced
 363 by ERA40, ERA interim, and ERA5 renalysis products at the domain's lateral bound-
 364aries. The RACMO simulations have been shown to perform skillfully against observa-
 365 tions and are often used as modeling targets (e.g., Evans et al., 2019; van Kampenhout
 366 et al., 2020).

367 In-situ SMB (snow pit and ice cores) and radar accumulation datasets (e.g., Ice-
 368 Bridge) are maintained in The Land Ice Verification and Validation toolkit (LIVVkit),
 369 version 2.1 (Evans et al., 2019). However, these point-wise measurements are difficult
 370 to compare to model output spanning several different grids, especially since the SMB

data product	years used in this study	resolution	citation
ERA5	1979-1998	1/4°	Copernicus (2019)
CERES-EBAF ED4.1	2003-2020	1°	Loeb et al. (2018)
CALIPSO-GOCCP	2006-2017	1°	Chepfer et al. (2010)
RACMO2.3	1979-1998	11 km	Noël et al. (2015)
RACMO2.3p2	1979-1998	5.5 km	Noël et al. (2019)

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

from each elevation class 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 grid-ded fields to point-wise measurements, and the lack of information added with regard to model biases, we omitted these datasets from our analysis. **Andrew - maybe do show them, briefly?**

3 Results

3.1 Tropospheric temperatures

Before delving into the simulated Arctic, we assess the global mean differences between the various grids and dycores. Figure 4 shows 1979-1998 annual mean, zonal mean height plots expressed as differences between uniform-resolution grids and dycores. The f09 grid is warmer than the f19 grid, primarily in the mid-to-high latitudes throughout the depth of the troposphere. This is a common response to increasing horizontal resolution in GCMs (Pope & Stratton, 2002; Roeckner et al., 2006). Herrington and Reed (2020) have shown that this occurs in CAM due to greater resolved vertical velocities which, in turn, drive more condensational heating in the CLUBB macrophyiscs. The right columns in Figure 4 support this interpretation, showing an increase in the climatological CLUBB heating in low ([looks more like subtropics](#)) and mid-latitudes on the f09 grid.

As the SE dycore is less diffusive than the FV dycore, the resolved vertical velocities are larger in the SE dycore, and so the ne30pg3 troposphere is modestly warmer than f09 (Figure 4). The stratosphere responds differently, with ne30pg3 much cooler than f09 in the mid-to-high latitudes. Figure 4 also shows small temperature differences between ne30pg3 and ne30pg2, with ne30pg3 slightly warmer near the tropopause at high latitudes. This is consistent with the similar climates found for these two grids by Herrington et al. (2019).

Comparing the VR grids to the uniform-resolution grids is complicated because we simultaneously increase the resolution and reduce the physics time-step, both of which impact the solution (Williamson, 2008). We therefore run an additional ne30pg3 simulation with the shorter physics time step used in the **Arctic** grid (450 s), referred to as ne30pg3*. Figure 5a shows the difference between ne30pg3* and ne30pg3 for climatological summer temperatures in zonal-mean height space. Much of the troposphere is warmer with the reduced time step, and the mechanism is similar in that the shorter time step increases condensational heating by CLUBB. Figure 5b shows the difference in climatological summer temperature between the **Arctic** grid and the ne30pg3* grid. With the same physics time step on each grid, the greater condensational heating and warmer temperatures are confined to the refined Arctic region.

Figure 5c shows that the **Arctic-GrIS** grid is much warmer than the **Arctic** grid in the Arctic summer. This may be due, in part, to the shorter physics time step, but the temperature response is too large to be explained by the CLUBB changes alone. This

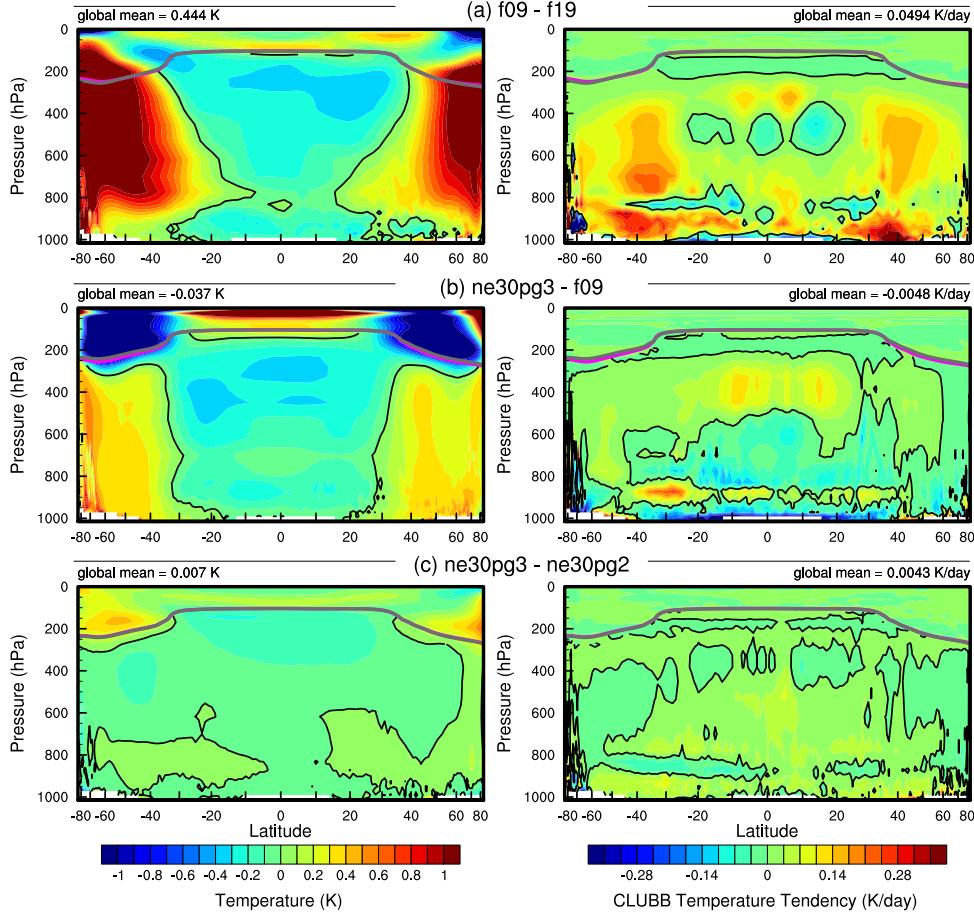


Figure 4. 1979–1998 annual mean temperature (left column) and CLUBB temperature tendencies (right column) in zonal mean height space, expressed as differences between the various 1° – 2° grids. The thick grey and magenta lines are the tropopause for the control run and the test run, respectively.

summer warming appears to be a result of variations in the stationary wave pattern, with anomalous southerly winds to the west of Greenland (not shown). This dynamic response is interesting, because other than the physics time step, the only difference between the Arctic–GrIS and Arctic runs is the doubling of resolution over Greenland. This behavior will be explored further in a future study.

It is useful to understand summer temperature biases due to their control on ice and snow melt over the Greenland Ice Sheet (GrIS; Ohmura, 2001). Figure 6 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° SE grids are warmer than the FV grids. The FV summer temperatures are persistently colder than ERA5, whereas the 1° SE grids are not as cold, and are actually warmer than ERA5 at high-latitudes, north of 75° . All grids show a north-south gradient in bias over Greenland, with the summer temperature bias more positive

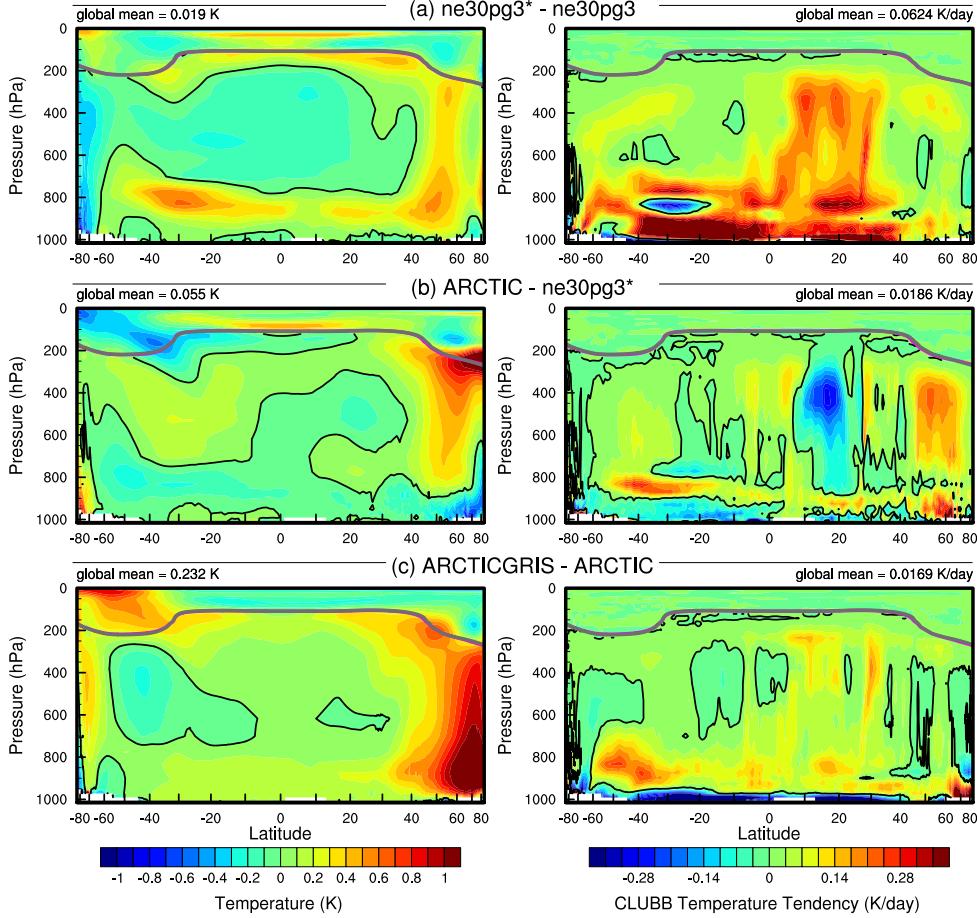


Figure 5. As in Figure 4 but for the short-time-step experiment and the VR grids. The fields plotted are for the climatological northern hemisphere summer. We focus on summer because that is when the resolution response is largest, and the refined regions are located in the northern hemisphere.

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° 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 than the **Arctic** grid over most of the Arctic, but with a similar spatial pattern of summer temperature bias.

Some of these temperature differences may be related to different summer shortwave cloud forcing for the various grids and dycores. Figure 7 shows the summer shortwave cloud forcing bias in six runs, using the CERES-EBAF product. A negative bias corresponds to excessive reflection and cooling. The uniform grids have similar biases, with the clouds reflecting 20–40 W/m² too much shortwave radiation over a wide swath of the Arctic, primarily the land masses (please explain what SWCF in caption.). There is also a halo of positive bias (clouds not reflective enough) around the ocean perimeter of Greenland. The **Arctic** grid has much smaller cloud forcing biases over the Arctic land masses, but is still too reflective over Alaska, the Canadian Archipelago, and parts

443 of Eurasia. Compared to the **Arctic** grid, the **Arctic – GrIS** grid vastly reduces the
 444 cloud forcing bias over Eurasia, and also improves the bias over North America. In both
 445 VR grids, the halo of positive forcing bias around the perimeter of Greenland is absent.

446 While the summer cloud forcing biases are consistent with the summer tempera-
 447 ture biases in Figure 6 – regions where clouds are too reflective coincide with regions that
 448 are too cold – it is not clear whether the cold biases are caused by the cloud biases, or
 449 whether the cold biases amplify the cloud forcing bias. (This ends abruptly. Anything
 450 else to add? Causation I agree is hard to prove due to energy transport, but in summer
 451 cloud fraction is probably playing a role. The likely other thing that could be playing
 452 a role is surface albedo (absorbed solar). Could note that.)

453 3.2 Shortwave radiation over Greenland

454 In addition to summer temperatures, shortwave radiation is an important deter-
 455 minant of snow and ice melt. Figure 8 shows the summer incident shortwave radiation
 456 bias at the surface, zoomed in over Greenland. The top panel shows the bias relative to
 457 the CERES-EBAF dataset, and the bottom panel relative to the RACMO2.3p2 dataset.
 458 The halo of excessive incident shortwave radiation around the coasts of Greenland is ap-
 459 parent for both datasets in relation to the coarser grids, consistent with the shortwave
 460 cloud forcing biases in Figure 7.

461 The ice sheet interior receives too little shortwave radiation on the coarser grids.
 462 On the VR grids, both the interior shortwave deficit and the excessive shortwave around
 463 the ocean perimeter are improved. This suggests that the coarse-grid clouds are too thick
 464 in the Greenland interior and too thin around the perimeter, and that increasing hor-
 465 izontal resolution reduces these biases. This is consistent with the total summer cloud
 466 fraction bias, computed from the CALIPSO-GOCCP cloud dataset and shown in Fig-
 467 ure 9. Note that total cloud fraction characterizes the cloud field at all vertical levels,
 468 but attenuates the changes arising from any single layer due to the maximum overlap
 469 assumption used to compute this quantity. Despite the attenuated signal, the total cloud
 470 fraction for the VR grids does indicate reduced cloud coverage in the interior and increased
 471 cloudiness around the ocean perimeter.

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

476 3.3 Greenland surface mass balance

477 The accuracy of the simulated SMB is expected to be sensitive to grid resolution.
 478 Figure ??b shows the average grid spacing over the Greenland Ice Sheet (GrIS) in all
 479 six grids in this study. The **ne30pg2** grid has the coarsest representation with an aver-
 480 age $\Delta x = 160 \text{ km}$, and the **Arctic – GrIS** grid has the highest resolution with an av-
 481 erage $\Delta x = 14.6 \text{ km}$, similar to the grid spacing of the 11 km RACMO2.3 grid. The
 482 **ne30pg3** grid has an average $\Delta x = 111.2 \text{ km}$, substantially coarser than the **f09** grid,
 483 with an average $\Delta x = 60 \text{ km}$. Although **ne30pg3** and **f09** have similar average grid
 484 spacing over the entire globe, and comparable computational costs, the convergence of
 485 meridians on the FV grid enhances the resolution over the GrIS. The **Arctic** grid has
 486 an average grid spacing of $\Delta x = 27.8 \text{ km}$, and is about 10 times more expensive than
 487 the 1° models. The **Arctic–GrIS** grid is about twice as expensive as the **Arctic** grid.
 488 (Wondering if this paragraph would go better in an earlier section. Rene agrees)

489 The lower panels of Figure 9 show the summer climatological mean precipitation
 490 bias over the GrIS, expressed as the fractional difference from the RACMO2.3p2 solu-
 491 tion. The coarse 1° – 2° grids have large, positive biases centered over the southern dome.

grid name	accumulation	total melt	runoff	sublimation	SMB
RACMO	768.5 (733.5)	-347.2 (-436.4)	-221.7 (-258.5)	-36.5 (-38.8)	510.3 (436.2)
f19	882.5 (913.5)	-440.3 (-546.5)	-283.7 (-284.3)	-36.6 (-37.5)	562.2 (591.7)
f09	874.8 (882.1)	-418.4 (-482.3)	-255.0 (-212.3)	-38.1 (-37.4)	581.7 (632.4)
ne30pg2	1000. (973.4)	-549.4 (-647.3)	-383.9 (-347.0)	-33.4 (-32.1)	582.7 (594.3)
ne30pg3	934.9 (909.3)	-568.8 (-686.7)	-356.2 (-330.1)	-34.4 (-32.6)	544.3 (546.6)
Arctic	795.9 (818.6)	-367.3 (-436.8)	-208.9 (-194.2)	-44.1 (-43.9)	542.9 (580.5)
Arctic – GrIS	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. ARH - these numbers need updating ... they are only for method 1.

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 van Kampenhout et al. (2018).

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 7, 8, and 9 clearly illustrate a shift in clouds from the interior to the ocean perimeter with increasing resolution.

Table 3 shows the 1979-1998 climatological SMB components for each grid, compared with RACMO Andrew - do you want to talk about the values at all?. The CESM values are averages over the ensemble of common ice masks and regridding methods described in section 2.5, and the RACMO values are averages over both RACMO datasets (Table 2) using the same common-ice-mask approach. Table 3 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 2.5. 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 10 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 3, are shown as circles on the right side of the panels. The uniform 1° – 2° grids have positive precipitation biases in the interior, whereas the VR grids have the smallest biases, with precipitation comparable to RACMO. The f19 and f09 grids perform similarly, with +110 Gt/yr bias, whereas ne30pg3 is biased by about +165 Gt/yr and ne30pg2 by +230 Gt/yr. The larger biases on the uniform-resolution SE grids relative to the FV grids are consistent with the coarser GrIS resolution on the SE grids (Figure 3).

The combined annual snow/ice melt shown in the bottom panel of Figure 10 indicates that the Arctic grid simulates the most realistic melt rates, with the other grids having more melt than RACMO. The Arctic–GrIS grid overpredicts melting by about

125 Gt/yr. This is likely due to an anomalously warm lower troposphere during the summer, relative to the **Arctic** run (Figure 6). The **f19** and **f09** melting rates are improved over **Arctic–GrIS**, overestimating melt by only 70–90 Gt/yr. The SE grids have the largest positive melt bias, between 200–220 Gt/yr. It is more difficult to attribute these differences to resolution alone, since the FV grids have colder summer temperatures than the uniform-resolution SE grids. (Not sure I understand this sentence. FV is both cooler and higher-resolution than SE, so one might suspect that cooler T is, in fact, connected to higher resolution.)

To illustrate the regional behavior of the SMB components, Figure 11 shows the precipitation and combined snow/ice melt integrated over the basins defined by Rignot and Mouginot (2012). The uncertainty due to differences in basin area is larger than for GrIS-wide integrals, owing to the differences in basin boundaries as represented by the common ice masks, which are shown in the **f19** and **ne30pg2** panels of Figure 12. Nonetheless, the regional totals in Figure 11 correctly show the southeast and southwest basins have the most accumulation. In all basins, accumulations drops monotonically with increasing grid resolution, with some exceptions. The **Arctic–GrIS** grid simulates less precipitation than RACMO in the central-east and southeast basins, and is closest of all grids to the RACMO precipitation in the large southwest basin.

The basin-integrated melt rates in Figure 11 depend on the dycore. The uniform-resolution SE grids have the largest positive biases in all basins. The **Arctic–GrIS** grid is a close second, while the FV grids have systematically smaller melt-rates. The “second-place” standing of **Arctic–GrIS** is somewhat unexpected, as this grid has the warmest lower-troposphere summer temperatures (Figure 6) and greatest incident shortwave radiation (Figure 8), yet it has less melting than the uniform-resolution SE grids.

Lower troposphere temperature is not a strict proxy for melting; e.g., it may not capture microclimate effects as a result of a better representation of the low-elevation ablation zones. Positive degree-days (PDD; Braithwaite, 1984), which accumulate the near-surface temperature in °C for days with temperature above freezing, are a more accurate proxy. PDD is nonlinear in mean monthly temperature (Reeh, 1991). We compute it from monthly mean 2-meter temperature using the method of Calov and Greve (2005), assuming a fixed monthly mean standard deviation of 3°C and a degree-day factor of 5 mm d⁻¹ °C⁻¹.

Figure 11c shows the basin-integrated PDD melt estimate. In the large southeast and southwest basins (and all the other western basins), the **ne30pg3** grid has larger PDD-based melt than the **ARCTCGRIS** grid. The FV grids also have large PDD-based melt in the southwest basin, relative to **Arctic–GrIS**. The PDD plots indicate that the near-surface temperatures which contribute to melt are not well approximated by the summer lower-troposphere temperatures in Figure 6.

Figure 12 presents the biases in the combined ice/snow melt as map plots. These plots show that the largest melt biases are on the southeast and northwest coasts, where large coarse-grid cells overlap with the ocean. One possibility is that these problematic grid cells are situated at lower elevations than the true ice sheet surface, leading to a warm bias and too much melt. Figure 13 shows the representation of the ice sheet surface along two transects on the different grids, compared to the high-resolution dataset used to generate CAM topographic boundary conditions (J. Danielson & Gesch, 2011; Lauritzen et al., 2015). The two transects are shown in Figure 12: 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°–2° 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.

581 The transects in Figure 13 show that the ice sheet surface on the coarse grids is
 582 not systematically lower than the true surface in ablation zones. Rather, the smoothing
 583 of the raw topography, necessary to prevent the model from instigating grid-scale modes,
 584 flattens the ice sheet, causing the lower-elevation ablation zones to extend beyond the
 585 true ice sheet margin, where they lie above the actual ice surface. The f19 grid has the
 586 smoothest topography since its dynamics are coarsest (whereas f09, ne30pg2 and ne30pg3
 587 use identical smoothing), and has the flattest ice sheet. This suggests that if anything,
 588 coarser models will tend to elevate the ablation zones and depress melt rates.

589 Figure 13 also shows the ice margin boundary, illustrating that the ablation zone
 590 lies in a narrow horizontal band where the ice sheet rapidly plunges to sea-level. Due to
 591 this abrupt transition, coarse grids will commonly represent the ablation zone with grid
 592 cells containing mixtures of ice-covered and ice-free regions. We hypothesize that coarser
 593 models have larger melt biases because summer melting is confined to these mixed ice/land/ocean
 594 grid cells. CLM deals with land heterogeneity in a complex and sophisticated manner,
 595 but CAM only sees a homogenized state due to volume averaging over the sub-grid mix-
 596 ture. Thus, warm ice-free land patches in a grid cell may unduly influence the climate
 597 over the entire grid cell, causing a warm bias over the ice-covered patch. (This is an in-
 598 teresting conclusion pointing to the need for better treatment of surface inhomogeneity
 599 in CAM. This might be a way to compensate for coarse resolution in future CESM ver-
 600 sions?)

601 Figure 14 shows mean melt bias, relative to both RACMO datasets, conditionally
 602 sampled based on grid cell ice fraction in the GrIS region. Errors are computed using
 603 the common-ice-mask approach, meaning that all fields are mapped to the common masks,
 604 which define the grid cell ice fraction. The figure shows (Any idea why the errors are smaller
 605 in the cells with intermediate ice fraction? I wouldn't have expected this.) that coarser
 606 grids generally have two peaks in ice fraction space; a bump in positive melting errors
 607 in the 0-20% range, and another in the fully ice-covered cells. Also shown are the ± 1 stan-
 608 dard deviation of the biases for each bin. They indicate that the biases in 0-20% bins
 609 are mostly contained in the positive bias region (a fractional bias greater than 0), whereas
 610 the fully-covered ice cells have a wider distribution, with many grid cells also contain-
 611 ing negative melting biases. The excessive melting in the 0-20% ice fraction bins sup-
 612 ports our hypothesis that the prevalence of mixed-grid cells in the ablation zone on coarse
 613 grids is responsible for their large melt bias.

614 Rene - is the melt map consistent with the bin figure? Smallest errors are interior
 615 points which correspond to glacier fraction bin of 1, which the largest errors. ARH - I
 616 can check. The bin figure is fractional change so its a different metric, which could ex-
 617 plain the apparent inconsistency.

618 3.4 Precipitation extremes

619 Synoptic storms are tracked using TempestExtremes atmospheric feature detec-
 620 tion software (Ullrich et al., 2021). As the Arctic grid contains $1/4^\circ$ refinement north
 621 of about 45° latitude, the storm tracker is applied to this region for the Arctic and ne30pg3
 622 runs to identify differences in storm characteristics due to horizontal resolution. The com-
 623 posite mean precipitation maps are similar between the two grids, and exhibit the iconic
 624 comma structure of synoptic cyclones (not shown).

625 Figure 15 shows monthly PDFs of the precipitation rates associated with storms.
 626 The PDFs are constructed by sampling all the precipitation rates within 30° of the storm
 627 center, for each point on the storm track and for all storms. The PDFs are evaluated on
 628 an identical composite grid for all runs, and so storm statistics are not impacted by dif-
 629 ferences in output resolution. The Arctic run has larger extreme precipitation rates com-
 630 pared to ne30pg3 in every month, but the increase is greatest in the summer months,
 631 which coincides with the most extreme events of the year. This is primarily due to in-

632 creased resolution and not the reduced physics times-step; the `ne30pg3*` run only marginally
 633 increases the extreme precipitation rates compared with `ne30pg3` (Figure 15).

634 The extreme precipitation rates in the `Arctic` run are closer than `ne30pg3` to the
 635 ERA5 reanalysis (Figure 15). It is difficult to know how much the extreme precipitation
 636 rates in ERA5 are constrained by data assimilation, or whether these precipitation rates
 637 are due to using a similar $1/4^\circ$ model as the `Arctic` grid. However, it is well documented
 638 that $1/4^\circ$ models are more skillful at simulating extreme events (Bacmeister et al., 2013;
 639 Obrien et al., 2016). This is an additional benefit of the VR grids.

640 4 Conclusions

641 Running CESM2.2 in an AMIP-style configuration, we have evaluated six grids from
 642 two dynamical cores for their performance over the Arctic and in simulating the Green-
 643 land Ice Sheet (GrIS) surface mass balance (SMB). The $1-2^\circ$ finite-volume grids have
 644 enhanced resolution over polar regions due to the convergence of meridian lines, although
 645 a polar filter is used to prevent spurious atmospheric features from forming at this higher
 646 resolution. Spectral-element grids comparable to the resolution of the finite-volume grids
 647 have an isotropic grid structure, meaning the grid resolution is similar over the entire
 648 domain. We developed two VR grids and introduced them into CESM2.2 as part of this
 649 work. Both use the spectral-element dycore; the `Arctic` grid has $1/4^\circ$ refinement over
 650 the broader Arctic, whereas the `Arctic–GrIS` grid is identical except for a $1/8^\circ$ patch
 651 of refinement over Greenland.

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

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

673 We integrated the primary source and sink terms of the SMB equation over the GrIS
 674 for each of the six grids. The uniform $1-2^\circ$ grids have large positive accumulation bi-
 675 ases because they fail to resolve orographic precipitation. The uniform SE grids have larger
 676 accumulation biases, suggesting that the FV grids are more skillful for precipitation due
 677 to finer resolution over Greenland, despite a polar filter. The VR grids have the most
 678 accurate accumulation rates of all the grids.

679 The primary mass sink term of the GrIS, ice/snow melt, has similar biases. The
 680 uniform resolution SE grids have too much melt, while the FV grids have smaller biases.
 681 It is difficult to attribute these biases to grid resolution alone. The FV grids have colder

682 summers, consistent with their lower melt bias. However, the **Arctic**–**GrIS** grid has
 683 the warmest summer temperatures of all grids, yet it has less melting than the uniform-
 684 resolution SE grids. This suggests that grid resolution is responsible for a large fraction
 685 of the melt biases. We propose a mechanism: Coarse grids represent ablation zones us-
 686 ing grid cells with mixed surface types, ice-covered and ice-free. The warmer ice-free patches
 687 may largely determine the mean state, leading to a warm bias over the ice-covered patches
 688 of the grid cell. This mechanism is supported by analysis of melt biases binned by grid-
 689 cell ice fraction.

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

698 As modeling systems move away from lat-lon grids towards quasi-uniform unstruc-
 699 tured grids, it is worth taking stock of whether this will degrade the simulated polar cli-
 700 mate. We have found that the 1° FV model has clear advantages over the 1° SE model
 701 in simulating the surface mass balance of the GrIS. This finding will not interrupt the
 702 ongoing transition towards unstructured grids in CESM, largely driven by gains in com-
 703 putational efficiency, but it has inspired us to develop alternative configurations that re-
 704 cover or improve on the fidelity of polar climate. We have shown here (and in a prior
 705 companion study (van Kampenhout et al., 2018)) that for CESM, Arctic-refined meshes
 706 can substantially improve the simulated mass balances of the GrIS, even compared to
 707 the 1° grid. This should reassure the CESM modeling community that the ongoing trans-
 708 sition away from lat-lon grids will not adversely impact CESM’s usefulness as a state-of-
 709 the-art tool for simulating and understanding polar processes. (WHL: This last sentence
 710 may be too sanguine. Yes, we can recover the fidelity of Arctic simulations using VR grids,
 711 but (so far) only at a considerable cost in cpu-hours. This points to the need for an in-
 712 termediate resolution that is more affordable, and/or model development or tuning that
 713 reduces the biases on coarse grids.) Andrew - Maybe better to just state that higher res-
 714 olution is better: 1deg better than 2deg FV, and ne30 is coarser than 1deg at Greenland
 715 latitudes so it’s worse. But VR has higher resolution so SR-VR is better. It’s all about
 716 resolution.

717 We are working to develop a configuration of the **Arctic** grid that is fully-coupled
 718 with the CESM ocean and sea ice components and the Community Ice Sheet Model (CISM),
 719 to provide multi-century projections of the state of the GrIS and its contribution to sea-
 720 level rise. We have also developed a visualization of the *ARCTCGRIS* run, now avail-
 721 able on youtube¹. Figure 16 shows a snapshot of this visualization, illustrating mesoscale
 722 katabatic winds descending the southeastern slopes of GrIS. These new grids and con-
 723 figurations will provide new opportunities for CESM polar science and aims to contribute
 724 to an improved understanding of the polar environment. (WHL: I replaced the previ-
 725 ous last sentence because it seemed too much like an advertisement. However, this new
 726 ending seems weak. I wonder if we should say something about future work motivated
 727 by this study, for instance investigating grids and parameterizations that provide some
 728 of the same benefits as these VR grids but at lower cost.)

¹ https://www.youtube.com/watch?v=YwHgqDu75s8&t=4s&ab_channel=NCARVisLab

729 **Appendix A Details on spectra-element dynamical core improvements
730 since the CESM2.0 release**

731 Since the CESM2.0 release of the spectral-element dynamical core documented in
732 Lauritzen et al. (2018) some important algorithmic improvements have been implemented
733 and released with CESM2.2. These pertain mainly to the flow over orography that, for
734 the spectral-element dynamical core, can lead to noise aligned with the element bound-
735aries (Herrington et al., 2018).

736 **A1 Reference profiles**

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

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

$$p_s^{(ref)} = p_0 \exp\left(-\frac{\Phi_s}{R^{(d)}(T_0 + T_1)}\right), \quad (\text{A2})$$

739 **Mark - Note: new definition of T0 means this equation should also be changed** (Simmons
740 & Jiabin, 1991) where g gravity, $T_1 = \Gamma_0 T_0 c_p^{(d)} / g \approx 192K$ with standard lapse rate
741 $\Gamma_0 \equiv 6.5K/km$ and $T_0 \equiv 288K - T_1 \approx 97K$ ($c_p^{(d)}$ specific heat of dry air at constant
742 pressure; $R^{(d)}$ gas constant for dry air), and Φ_s surface geopotential. The reference Exner
743 function is

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

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

$$p^{(ref)}(\eta) = A(\eta)p_0 + B(\eta)p_s^{(ref)}, \quad (\text{A4})$$

746 where A and B are the standard hybrid coefficients (using a dry-mass generalized ver-
747 tical mass coordinate η). These reference profiles are subtracted from the prognostic tem-
748 perature and pressure-level-thickness states before applying hyperviscosity:

$$\text{CESM2.0} \rightarrow \text{CESM2.2} \quad (\text{A5})$$

$$(\text{A6})$$

$$\nabla_\eta^4 T \rightarrow \nabla_\eta^4 \left(T - T^{(ref)}\right), \quad (\text{A7})$$

$$\nabla_\eta^4 \delta p^{(d)} \rightarrow \nabla_\eta^4 \left(\delta p^{(d)} - \delta p^{(ref)}\right). \quad (\text{A8})$$

749 This reduces spurious transport of temperature and mass up/down-slope due to the hy-
750 perviscosity operator.

751 **A2 Rewriting the pressure gradient force (PGF)**

752 In the CESM2.0 the following (standard) form of the pressure gradient term was
753 used:

$$\nabla_\eta \Phi + \frac{1}{\rho} \nabla_\eta p, \quad (\text{A9})$$

754 where Φ is geopotential and $\rho = \frac{R^{(d)} T_v}{p}$ is density (for details see Lauritzen et al., 2018).
755 To alleviate noise for flow over orography, we switched to an Exner pressure formulation
756 following Taylor et al. (2020), which uses that (A9) can be written in terms of the Exner
757 pressure

$$\nabla_\eta \Phi + c_p^{(d)} \theta_v \nabla_\eta \Pi, \quad (\text{A10})$$

758 where the Exner pressure is

$$\Pi \equiv \left(\frac{p}{p_0} \right)^\kappa. \quad (\text{A11})$$

759 The derivation showing that (A9) and (A10) are equivalent is shown here:

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

760 Using the reference states from (Simmons & Jiabin, 1991),

$$\bar{T} = T_0 + T_1 \Pi, \quad (\text{A12})$$

$$\bar{\theta} = T_0 / \Pi + T_1, \quad (\text{A13})$$

761 we can define a geopotential as a function of Exner pressure

$$\bar{\Phi} = -c_p^{(d)} (T_0 \log \Pi + T_1 \Pi - T_1). \quad (\text{A14})$$

762 This "balanced" geopotential obeys

$$c_p^{(d)} \bar{\theta} \nabla \Pi + \nabla \bar{\Phi} = 0 \quad (\text{A15})$$

763 for any Exner pressure. Subtracting this "reference" profile from the PGF yields

$$\begin{aligned} \nabla_\eta \Phi + c_p^{(d)} \theta_v \nabla_\eta \Pi &= \nabla_\eta (\Phi - \bar{\Phi}) + c_p^{(d)} (\theta_v - \bar{\theta}) \nabla_\eta \Pi, \\ &= \nabla_\eta \Phi + c_p^{(d)} \theta_v \nabla_\eta \Pi + c_p^{(d)} T_0 \left[\nabla_\eta \log \Pi - \frac{1}{\Pi} \nabla_\eta \Pi \right]. \end{aligned} \quad (\text{A16})$$

764 In the continuum, the two formulations (left and right-hand side of (A16)) are identi-
765 cal. But under discretization, the second formulation can have much less truncation er-
766 ror.

767 A3 Results

768 [Adam: have you defined ne30np4 in the main text?]

769 One year averages of vertical pressure velocity at 500hPa (`OMEGA500`) have been
770 found to be a useful quantity to detect spurious up or down-drafts induced by steep orog-
771 raphy (Figure A1). While the true solution is not known, strong vertical velocities aligned
772 with element edges that are not found in the CAM-FV reference solution (Figure A1(a))
773 are likely not physical (spurious). The older CESM2.0 version of SE (Figure A1(d)) us-
774 ing the "traditional" discretization of the PGF, (A16), exhibits significant spurious noise
775 patters around steep orography compared to CAM-FV (e.g., around Himalayas and An-
776 des). This is strongly alleviated by switching to the Exner formulation of the PGF (A10;
777 Figure A1(c)). By also subtracting reference profiles from pressure-level thickness and

778 temperature, equations (A7) and (A8) respectively, reduces strong up-down drafts fur-
 779 ther (Figure A1(d)). Switching to the CAM-SE-CSLAM version where physics ten-
 780 dencies are computed on an quasi-equal area physics grid and using the CSLAM transport
 781 scheme, marginal improvements are observed in terms of a smoother vertical velocity field
 782 (Figure A1(e,f)). The configuration shown in Figure A1(d) is used for the simulations
 783 shown in the main text of this paper.

784 It is interesting to note that the noise issues and algorithmic remedies found in the
 785 real-world simulations discussed above, can be investigated by replacing all of physics
 786 with a modified version of the Held-Suarez forcing (Held & Suarez, 1994). The original
 787 formulation of the Held-Suarez idealized test case used a flat Earth ($\Phi_s = 0$) and a dry
 788 atmosphere. By simply adding the surface topography used in ‘real-world’ simulations
 789 and removing the temperature relaxation in the lower part of domain ($\sigma > 0.7$; see Held
 790 and Suarez (1994) for details), surprisingly realistic vertical velocity fields (in terms of
 791 structure) result (see Figure A2). Since this was a very useful development tool it is shared
 792 in this manuscript.

793 Acknowledgments

794 This material is based upon work supported by the National Center for Atmospheric Re-
 795 search (NCAR), which is a major facility sponsored by the NSF under Cooperative Agree-
 796 ment 1852977. Computing and data storage resources, including the Cheyenne super-
 797 computer (doi:10.5065/D6RX99HX), were provided by the Computational and Informa-
 798 tion Systems Laboratory (CISL) at NCAR. A. Herrington thanks Matt Rehme (NCAR/CISL)
 799 for his role in generating the Arctic–GrIS visualization available on youtube (https://www.youtube.com/watch?v=YwHgqDu75s8&t=4s&ab_channel=NCARVisLab).
 800

801 The data presented in main part of this manuscript is available at <https://github.com/adamrher/2020-arcticgrids>. The source code and data for the Appendix is avail-
 802 able at <https://github.com/PeterHjortLauritzen/CAM/tree/topo-mods>.
 803

804 References

- 805 Bacmeister, J. T., & Coauthors. (2018). Projected changes in tropical cyclone activ-
 806 ity under future warming scenarios using a high-resolution climate model. *Cli-*
 807 *matic Change*, 146, 547–560. Retrieved from <http://dx.doi.org/10.1007/s10584-016-1750-x> doi: 10.1007/s10584-016-1750-x
 808
 809 Bacmeister, J. T., Wehner, M. F., Neale, R. B., Gettelman, A., Hannay, C., Lau-
 810 ritzen, P. H., ... Truesdale, J. E. (2013). Exploratory high-resolution climate
 811 simulations using the community atmosphere model (cam). *J. Climate*, 27(9),
 812 3073–3099. doi: 10.1175/JCLI-D-13-00387.1
 813 Bambach, N. E., Rhoades, A. M., Hatchett, B. J., Jones, A. D., Ullrich, P. A., &
 814 Zarzycki, C. M. (2021). Projecting climate change in south america using
 815 variable-resolution community earth system model: An application to chile.
International Journal of Climatology.
 816 Beljaars, A., Brown, A., & Wood, N. (2004). A new parametrization of turbulent
 817 orographic form drag. *Quart. J. Roy. Meteor. Soc.*, 130(599), 1327–1347. doi:
 818 10.1256/qj.03.73
 819
 820 Bogenschutz, P. A., Gettelman, A., Morrison, H., Larson, V. E., Craig, C., & Scha-
 821 nen, D. P. (2013). Higher-order turbulence closure and its impact on climate
 822 simulations in the community atmosphere model. *Journal of Climate*, 26(23),
 823 9655–9676.
 824 Box, J. E., Bromwich, D. H., & Bai, L.-S. (2004). Greenland ice sheet surface mass
 825 balance 1991–2000: Application of polar mm5 mesoscale model and in situ
 826 data. *Journal of Geophysical Research: Atmospheres*, 109(D16).
 827 Braithwaite, R. J. (1984). Calculation of degree-days for glacier-climate research.

- 828 *Zeitschrift für Gletscherkunde und Glazialgeologie*, 20(1984), 1–8.
- 829 Burakowski, E. A., Tawfik, A., Ouimette, A., Lepine, L., Zarzycki, C., Novick, K.,
 830 ... Bonan, G. (2019). Simulating surface energy fluxes using the variable-
 831 resolution community earth system model (vr-cesm). *Theoretical and Applied
 832 Climatology*, 138(1), 115–133.
- 833 Calov, R., & Greve, R. (2005). A semi-analytical solution for the positive degree-day
 834 model with stochastic temperature variations. *Journal of Glaciology*, 51(172),
 835 173–175.
- 836 Canuto, C., Hussaini, M. Y., Quarteroni, A., & Zang, T. (2007). *Spectral methods:
 837 Evolution to complex geometries and applications to fluid dynamics* (1st ed.).
 838 Springer.
- 839 Chang, P., Zhang, S., Danabasoglu, G., Yeager, S. G., Fu, H., Wang, H., ... oth-
 840 ers (2020). An unprecedented set of high-resolution earth system simulations
 841 for understanding multiscale interactions in climate variability and change.
 842 *Journal of Advances in Modeling Earth Systems*, 12(12), e2020MS002298.
- 843 Chepfer, H., Bony, S., Winker, D., Cesana, G., Dufresne, J., Minnis, P., ... Zeng, S.
 844 (2010). The gcm-oriented calipso cloud product (calipso-goccp). *Journal of
 845 Geophysical Research: Atmospheres*, 115(D4).
- 846 Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., McCaa, J. R., Williamson,
 847 D. L., ... Zhang, M. (2006). The formulation and atmospheric simulation
 848 of the community atmosphere model version 3 (cam3). *Journal of Climate*,
 849 19(11), 2144–2161.
- 850 Copernicus, C. (2019). Era5 monthly averaged data on pressure levels from
 851 1979 to present. URL: <https://cds.climate.copernicus.eu>. doi: 10.24381/
 852 cds.6860a573
- 853 Craig, C., Bacmeister, J., Callaghan, P., Eaton, B., Gettelman, A., Goldhaber, S. N.,
 854 ... Vitt, F. M. (2021). *Cam6.3 user's guide* (Tech. Rep.). NCAR/TN-
 855 571+EDD. doi: 10.5065/Z953-ZC95
- 856 Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D., DuVivier, A., Ed-
 857 wards, J., ... others (2020). The community earth system model ver-
 858 sion 2 (cesm2). *Journal of Advances in Modeling Earth Systems*, 12(2),
 859 e2019MS001916.
- 860 Danielson, J., & Gesch, D. (2011). *Global multi-resolution terrain elevation data
 861 2010 (GMTED2010)* (Open-File Report 2011-1073). U.S. Geological Survey.
 862 (<http://pubs.usgs.gov/of/2011/1073/pdf/of2011-1073.pdf>)
- 863 Danielson, J. J., & Gesch, D. B. (2011). *Global multi-resolution terrain elevation
 864 data 2010 (GMTED2010)* (Open File Rep. No. 2011-1073). US Geological Sur-
 865 vey. doi: <https://doi.org/10.3133/ofr20111073>
- 866 Dennis, J. M., Edwards, J., Evans, K. J., Guba, O., Lauritzen, P. H., Mirin, A. A.,
 867 ... Worley, P. H. (2012). CAM-SE: A scalable spectral element dynamical
 868 core for the Community Atmosphere Model. *Int. J. High. Perform. C.*, 26(1),
 869 74–89. Retrieved from <http://hpc.sagepub.com/content/26/1/74.abstract>
 870 doi: 10.1177/1094342011428142
- 871 Evans, K. J., Kennedy, J. H., Lu, D., Forrester, M. M., Price, S., Fyke, J., ... oth-
 872 ers (2019). Livvkit 2.1: automated and extensible ice sheet model validation.
 873 *Geoscientific Model Development*, 12(3), 1067–1086.
- 874 Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., &
 875 Taylor, K. E. (2016). Overview of the coupled model intercomparison project
 876 phase 6 (cmip6) experimental design and organization. *Geoscientific Model
 877 Development*, 9(5), 1937–1958.
- 878 Fettweis, X., Franco, B., Tedesco, M., Van Angelen, J., Lenaerts, J. T., van den
 879 Broeke, M. R., & Gallée, H. (2013). Estimating the greenland ice sheet
 880 surface mass balance contribution to future sea level rise using the regional
 881 atmospheric climate model mar. *The Cryosphere*, 7(2), 469–489.
- 882 Gettelman, A., Callaghan, P., Larson, V., Zarzycki, C., Bacmeister, J., Lauritzen, P.,

- 883 ... Neale, R. (2017). Regional climate simulations with the community earth
 884 system model. *J. Adv. Model. Earth Syst.*. (submitted)
- 885 Gettelman, A., Hannay, C., Bacmeister, J. T., Neale, R. B., Pendergrass, A., Dan-
 886 abasoglu, G., ... others (2019). High climate sensitivity in the community
 887 earth system model version 2 (cesm2). *Geophysical Research Letters*, 46(14),
 888 8329–8337.
- 889 Gettelman, A., & Morrison, H. (2015). Advanced two-moment bulk microphysics for
 890 global models. part i: Off-line tests and comparison with other schemes. *Journal*
 891 of *Climate*, 28(3), 1268–1287.
- 892 Gettelman, A., Morrison, H., Santos, S., Bogenschutz, P., & Caldwell, P. (2015).
 893 Advanced two-moment bulk microphysics for global models. part ii: Global
 894 model solutions and aerosol–cloud interactions. *Journal of Climate*, 28(3),
 895 1288–1307.
- 896 Golaz, J.-C., Larson, V. E., & Cotton, W. R. (2002). A pdf-based model for bound-
 897 ary layer clouds. part i: Method and model description. *Journal of the Atmo-*
 898 *spheric Sciences*, 59(24), 3540–3551. doi: 10.1175/1520-0469(2002)059<3540:
 899 apbmbf>2.0.co;2
- 900 Guba, O., Taylor, M. A., Ullrich, P. A., Overfelt, J. R., & Levy, M. N. (2014). The
 901 spectral element method (sem) on variable-resolution grids: evaluating grid
 902 sensitivity and resolution-aware numerical viscosity. *Geosci. Model Dev.*, 7(6),
 903 2803–2816. doi: 10.5194/gmd-7-2803-2014
- 904 Guo, Z., Wang, M., Qian, Y., Larson, V. E., Ghan, S., Ovchinnikov, M., ... Zhou,
 905 T. (2015). Parametric behaviors of clubb in simulations of low clouds in the c
 906 ommunity atmosphere model (cam). *Journal of Advances in Modeling Earth*
 907 *Systems*, 7(3), 1005–1025.
- 908 Hansen, N., Simonsen, S. B., Boberg, F., Kittel, C., Orr, A., Souverijns, N., ... Mot-
 909 tram, R. (2022). Brief communication: Impact of common ice mask in surface
 910 mass balance estimates over the antarctic ice sheet. *The Cryosphere*, 16(2),
 911 711–718.
- 912 Held, I. M., & Suarez, M. J. (1994). A proposal for the intercomparison of the dy-
 913 namical cores of atmospheric general circulation models. *Bull. Am. Meteorol.*
 914 *Soc.*, 73, 1825–1830.
- 915 Herrington, A. R., Lauritzen, P., Taylor, M. A., Goldhaber, S., Eaton, B. E.,
 916 Bacmeister, J., ... Ullrich, P. (2018). Physics-dynamics coupling with element-
 917 based high-order galerkin methods: quasi equal-area physics grid. *Mon. Wea.*
 918 *Rev.*, 47, 69–84. doi: 10.1175/MWR-D-18-0136.1
- 919 Herrington, A. R., Lauritzen, P. H., Reed, K. A., Goldhaber, S., & Eaton, B. E.
 920 (2019). Exploring a lower resolution physics grid in cam-se-cslam. *Journal of*
 921 *Advances in Modeling Earth Systems*, 11.
- 922 Herrington, A. R., & Reed, K. A. (2018). An idealized test of the response of the
 923 community atmosphere model to near-grid-scale forcing across hydrostatic
 924 resolutions. *J. Adv. Model. Earth Syst.*, 10(2), 560–575.
- 925 Herrington, A. R., & Reed, K. A. (2020). On resolution sensitivity in the commu-
 926 nity atmosphere model. *Quarterly Journal of the Royal Meteorological Society*,
 927 146(733), 3789–3807.
- 928 Hurrell, J. W., Hack, J. J., Shea, D., Caron, J. M., & Rosinski, J. (2008). A new
 929 sea surface temperature and sea ice boundary dataset for the community at-
 930 mosphere model. *Journal of Climate*, 21(19), 5145–5153.
- 931 Jablonowski, C., & Williamson, D. L. (2011). The pros and cons of diffusion, fil-
 932 ters and fixers in atmospheric general circulation models. In P. H. Lauritzen,
 933 C. Jablonowski, M. Taylor, & R. Nair (Eds.), *Numerical techniques for global*
 934 *atmospheric models* (pp. 381–493). Berlin, Heidelberg: Springer Berlin Heidel-
 935 berg. doi: 10.1007/978-3-642-11640-7_13
- 936 Lauritzen, P. H., Bacmeister, J. T., Callaghan, P. F., & Taylor, M. A. (2015). Ncar
 937 global model topography generation software for unstructured grids. *Geosci-*

- entific Model Development Discussions*, 8(6), 4623–4651. doi: 10.5194/gmdd-8-4623-2015

Lauritzen, P. H., Jablonowski, C., Taylor, M., & Nair, R. D. (2010). Rotated versions of the jablonowski steady-state and baroclinic wave test cases: A dynamical core intercomparison. *J. Adv. Model. Earth Syst.*, 2(15), 34 pp.

Lauritzen, P. H., Mirin, A., Truesdale, J., Raeder, K., Anderson, J., Bacmeister, J., & Neale, R. B. (2011). Implementation of new diffusion/filtering operators in the CAM-FV dynamical core. *Int. J. High Perform. Comput. Appl.*. doi: 10.1177/1094342011410088

Lauritzen, P. H., Nair, R., Herrington, A., Callaghan, P., Goldhaber, S., Dennis, J., ... Dubos, T. (2018). 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. *J. Adv. Model. Earth Syst.*, 10(7), 1537–1570. doi: 10.1029/2017MS001257

Lauritzen, P. H., Taylor, M. A., Overfelt, J., Ullrich, P. A., Nair, R. D., Goldhaber, S., & Kelly, R. (2017). CAM-SE-CSLAM: Consistent coupling of a conservative semi-lagrangian finite-volume method with spectral element dynamics. *Mon. Wea. Rev.*, 145(3), 833–855. doi: 10.1175/MWR-D-16-0258.1

Lawrence, D. M., Fisher, R. A., Koven, C. D., Oleson, K. W., Swenson, S. C., Bonan, G., ... others (2019). The community land model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *Journal of Advances in Modeling Earth Systems*, 11(12), 4245–4287.

Lin, S.-J. (2004). A 'vertically Lagrangian' finite-volume dynamical core for global models. *Mon. Wea. Rev.*, 132, 2293–2307.

Lin, S.-J., & Rood, R. B. (1997). An explicit flux-form semi-Lagrangian shallow-water model on the sphere. *Q.J.R.Meteorol.Soc.*, 123, 2477–2498.

Lipscomb, W. H., Fyke, J. G., Vizcaíno, M., Sacks, W. J., Wolfe, J., Vertenstein, M., ... Lawrence, D. M. (2013). Implementation and initial evaluation of the glimmer community ice sheet model in the community earth system model. *Journal of Climate*, 26(19), 7352–7371.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., ... Kato, S. (2018). Clouds and the earth's radiant energy system (ceres) energy balanced and filled (ebaf) top-of-atmosphere (toa) edition-4.0 data product. *Journal of Climate*, 31(2), 895–918.

Lofverstrom, M., Fyke, J. G., Thayer-Calder, K., Muntjewerf, L., Vizcaino, M., Sacks, W. J., ... Bradley, S. L. (2020). An efficient ice sheet/earth system model spin-up procedure for cesm2-cism2: Description, evaluation, and broader applicability. *Journal of Advances in Modeling Earth Systems*, 12(8), e2019MS001984.

Mottram, R., Boberg, F., Langen, P., Yang, S., Rodehacke, C., Christensen, J. H., & Madsen, M. S. (2017). Surface mass balance of the greenland ice sheet in the regional climate model hirham5: Present state and future prospects. *Low Temp. Sci.*, 75, 105–115.

Muntjewerf, L., Sacks, W. J., Lofverstrom, M., Fyke, J., Lipscomb, W. H., Ernani da Silva, C., ... Sellevold, R. (2021). Description and Demonstration of the Coupled Community Earth System Model v2–Community Ice Sheet Model v2 (CESM2-CISM2). *Journal of Advances in Modeling Earth Systems*, 13(6), e2020MS002356.

Neale, R. B., Richter, J. H., & Jochum, M. (2008). The impact of convection on ENSO: From a delayed oscillator to a series of events. *J. Climate*, 21, 5904–5924.

Noël, B., Van De Berg, W., Van Meijgaard, E., Kuipers Munneke, P., Van De Wal, R., & Van Den Broeke, M. (2015). Evaluation of the updated regional climate model racmo2. 3: summer snowfall impact on the greenland ice sheet. *The Cryosphere*, 9(5), 1831–1844.

- Noël, B., van de Berg, W. J., Lhermitte, S., & van den Broeke, M. R. (2019). Rapid ablation zone expansion amplifies north greenland mass loss. *Science advances*, 5(9), eaaw0123.
- Noël, B., van de Berg, W. J., Van Wessem, J. M., Van Meijgaard, E., Van As, D., Lenaerts, J., ... others (2018). Modelling the climate and surface mass balance of polar ice sheets using racmo2–part 1: Greenland (1958–2016). *The Cryosphere*, 12(3), 811–831.
- Obrien, T. A., Collins, W. D., Kashinath, K., Rübel, O., Byna, S., Gu, J., ... Ullrich, P. A. (2016). Resolution dependence of precipitation statistical fidelity in hindcast simulations. *J. Adv. Model. Earth Syst.*, 8(2), 976–990. Retrieved from <http://dx.doi.org/10.1002/2016ms000671> doi: 10.1002/2016ms000671
- Ohmura, A. (2001). Physical basis for the temperature-based melt-index method. *Journal of applied Meteorology*, 40(4), 753–761.
- Pfister, G. G., Eastham, S. D., Arellano, A. F., Aumont, B., Barsanti, K. C., Barth, M. C., ... others (2020). The multi-scale infrastructure for chemistry and aerosols (musica). *Bulletin of the American Meteorological Society*, 101(10), E1743–E1760.
- Pollard, D. (2010). A retrospective look at coupled ice sheet–climate modeling. *Climatic Change*, 100(1), 173–194.
- Pollard, D., & Groups, P. P. (2000). Comparisons of ice-sheet surface mass budgets from paleoclimate modeling intercomparison project (pmip) simulations. *Global and Planetary Change*, 24(2), 79–106.
- Pope, V., & Stratton, R. (2002). The processes governing horizontal resolution sensitivity in a climate model. *Climate Dynamics*, 19(3-4), 211–236.
- Putman, W. M., & Lin, S.-J. (2007). Finite-volume transport on various cubed-sphere grids. *J. Comput. Phys.*, 227(1), 55–78.
- Rae, J., Adalgeirsdóttir, G., Edwards, T. L., Fettweis, X., Gregory, J., Hewitt, H., ... others (2012). Greenland ice sheet surface mass balance: evaluating simulations and making projections with regional climate models. *The Cryosphere*, 6(6), 1275–1294.
- Reeh, N. (1991). Parameterization of melt rate and surface temperature in the greenland ice sheet. *Polarforschung*, 59(3), 113–128.
- Rhoades, A. M., Huang, X., Ullrich, P. A., & Zarzycki, C. M. (2016). Characterizing sierra nevada snowpack using variable-resolution cesm. *Journal of Applied Meteorology and Climatology*, 55(1), 173–196. Retrieved from <https://doi.org/10.1175/JAMC-D-15-0156.1> doi: 10.1175/JAMC-D-15-0156.1
- Richter, J. H., Sassi, F., & Garcia, R. R. (2010). Toward a physically based gravity wave source parameterization in a general circulation model. *J. Atmos. Sci.*, 67, 136–156. doi: dx.doi.org/10.1175/2009JAS3112.1
- Rignot, E., & Mouginot, J. (2012). Ice flow in greenland for the international polar year 2008–2009. *Geophysical Research Letters*, 39(11).
- Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., ... Schulzweida, U. (2006). Sensitivity of simulated climate to horizontal and vertical resolution in the echam5 atmosphere model. *Journal of Climate*, 19(16), 3771–3791.
- Sellevold, R., Van Kampenhout, L., Lenaerts, J., Noël, B., Lipscomb, W. H., & Vizcaino, M. (2019). Surface mass balance downscaling through elevation classes in an earth system model: Application to the greenland ice sheet. *The Cryosphere*, 13(12), 3193–3208.
- Simmons, A. J., & Jiabin, C. (1991). The calculation of geopotential and the pressure gradient in the ECMWF atmospheric model: Influence on the simulation of the polar atmosphere and on temperature analyses. *Quart. J. Roy. Meteor. Soc.*, 117(497), 29–58. Retrieved from <https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.49711749703> doi:

- 1048 https://doi.org/10.1002/qj.49711749703
- 1049 Small, R. J., Bacmeister, J., Bailey, D., Baker, A., Bishop, S., Bryan, F., ... Verten-
1050 stein, M. (2014). A new synoptic scale resolving global climate simulation
1051 using the community earth system model. *J. Adv. Model. Earth Syst.*, 6(4),
1052 1065–1094. doi: 10.1002/2014MS000363
- 1053 Suarez, M. J., & Takacs, L. L. (1995). Volume 5 documentation of the aries/geos dy-
1054 namical core: Version 2.
- 1055 Taylor, M. A., & Fournier, A. (2010). A compatible and conservative spectral el-
1056 ement method on unstructured grids. *J. Comput. Phys.*, 229(17), 5879 - 5895.
1057 doi: 10.1016/j.jcp.2010.04.008
- 1058 Taylor, M. A., Guba, O., Steyer, A., Ullrich, P. A., Hall, D. M., & Eldred, C.
1059 (2020). An energy consistent discretization of the nonhydrostatic equations
1060 in primitive variables. *Journal of Advances in Modeling Earth Systems*, 12(1).
1061 doi: 10.1029/2019MS001783
- 1062 Taylor, M. A., Tribbia, J., & Iskandarani, M. (1997). The spectral element method
1063 for the shallow water equations on the sphere. *J. Comput. Phys.*, 130, 92-108.
- 1064 Team, E. J. S., Balaji, V., Boville, B., Collins, N., Craig, T., Cruz, C., ... others
1065 (2021). *Esmf user guide* (Tech. Rep.).
- 1066 Ullrich, P. A., & Taylor, M. A. (2015). Arbitrary-order conservative and consis-
1067 tent remapping and a theory of linear maps: Part i. *Monthly Weather Review*,
1068 143(6), 2419–2440.
- 1069 Ullrich, P. A., Zarzycki, C. M., McClenney, E. E., Pinheiro, M. C., Stansfield, A. M.,
1070 & Reed, K. A. (2021). Tempestextremes v2. 1: a community framework for
1071 feature detection, tracking and analysis in large datasets. *Geoscientific Model
1072 Development Discussions*, 1–37.
- 1073 Van Angelen, J., Lenaerts, J., Lhermitte, S., Fettweis, X., Kuipers Munneke, P.,
1074 Van den Broeke, M., ... Smeets, C. (2012). Sensitivity of greenland ice
1075 sheet surface mass balance to surface albedo parameterization: a study with a
1076 regional climate model. *The Cryosphere*, 6(5), 1175–1186.
- 1077 van Kampenhout, L., Lenaerts, J. T., Lipscomb, W. H., Lhermitte, S., Noël, B.,
1078 Vizcaíno, M., ... van den Broeke, M. R. (2020). Present-day greenland ice
1079 sheet climate and surface mass balance in csm2. *Journal of Geophysical
1080 Research: Earth Surface*, 125(2).
- 1081 van Kampenhout, L., Rhoades, A. M., Herrington, A. R., Zarzycki, C. M., Lenaerts,
1082 J. T. M., Sacks, W. J., & van den Broeke, M. R. (2018). Regional grid refine-
1083 ment in an earth system model: Impacts on the simulated greenland surface
1084 mass balance. *The Cryosphere Discuss.*. doi: 10.5194/tc-2018-257
- 1085 Wan, H., Giorgetta, M. A., Zängl, G., Restelli, M., Majewski, D., Bonaventura, L.,
1086 ... others (2013). "the icon-1.2 hydrostatic atmospheric dynamical core on
1087 triangular grids, part i: formulation and performance of the baseline version".
1088 *Geosci. Model Dev.*, 6, 735–763.
- 1089 Williamson, D. (2007). The evolution of dynamical cores for global atmospheric
1090 models. *J. Meteor. Soc. Japan*, 85, 241-269.
- 1091 Williamson, D. (2008). Convergence of aqua-planet simulations with increasing res-
1092 olution in the community atmospheric model, version 3. *Tellus A*, 60(5), 848–
1093 862. doi: 10.1111/j.1600-0870.2008.00339.x
- 1094 Zarzycki, C. M., Jablonowski, C., & Taylor, M. A. (2014). Using variable-resolution
1095 meshes to model tropical cyclones in the community atmosphere model. *Mon.
1096 Wea. Rev.*, 142(3), 1221-1239. doi: 10.1175/MWR-D-13-00179.1
- 1097 Zhang, G., & McFarlane, N. (1995). Sensitivity of climate simulations to the
1098 parameterization of cumulus convection in the canadian climate center general-
1099 circulation model. *ATMOSPHERE-OCEAN*, 33(3), 407-446.

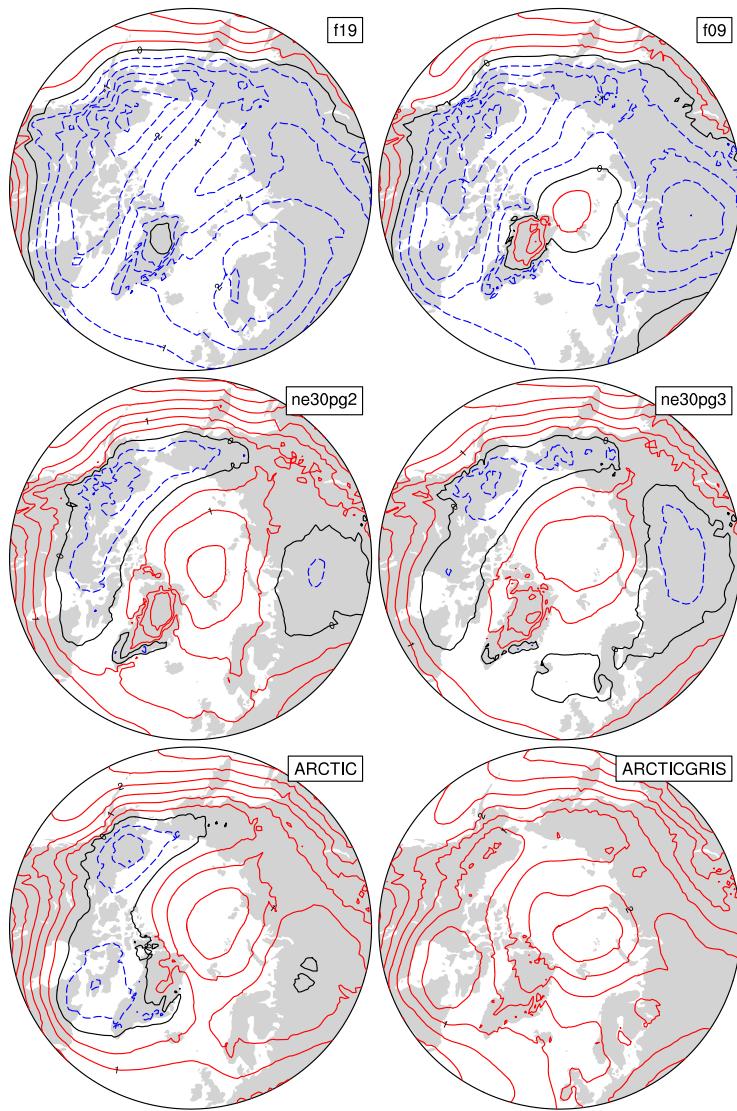


Figure 6. 1979-1998 lower troposphere, northern hemisphere summer virtual temperature biases, computed as the difference from ERA5. Lower troposphere layer mean virtual temperature is derived from the 1000 hPa - 500h Pa geopotential thickness, using the hypsometric equation. Differences are computed 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.

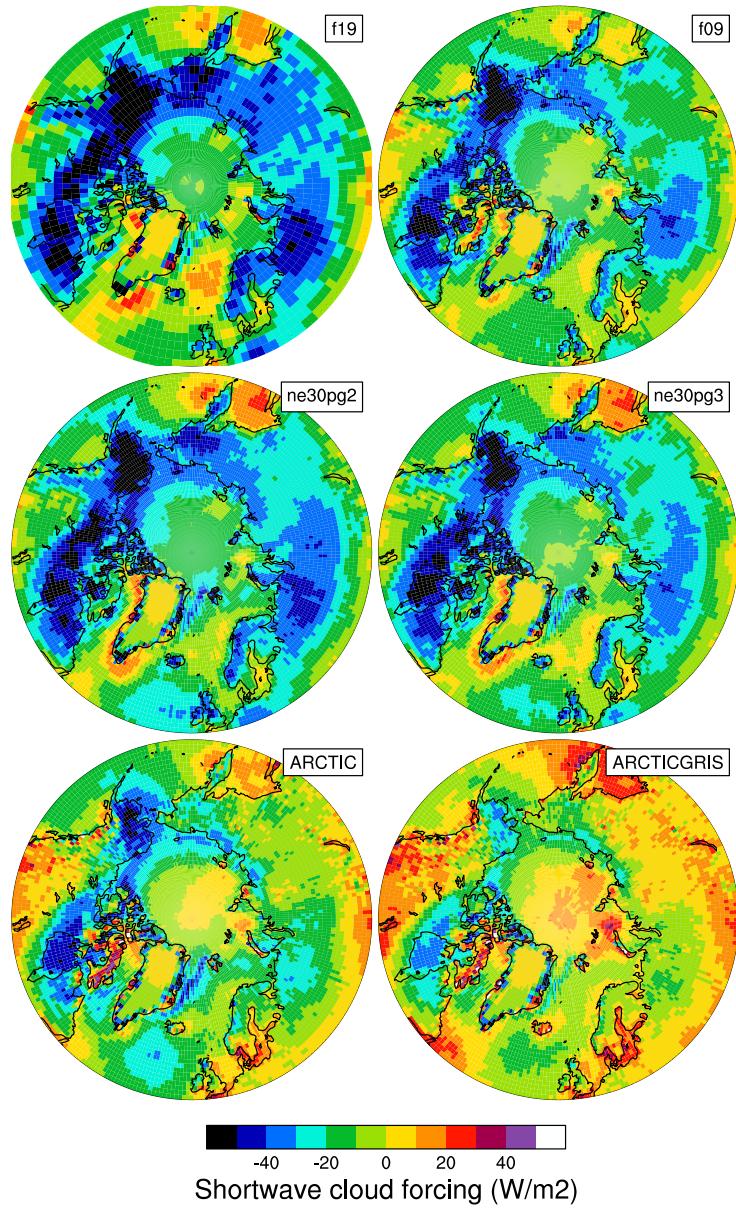


Figure 7. 1979-1998 Northern Hemisphere summer shortwave cloud forcing bias, relative to the CERES-EBAF gridded dataset. Differences are computed after mapping all model output to the 1° CERES-EBAF grid.

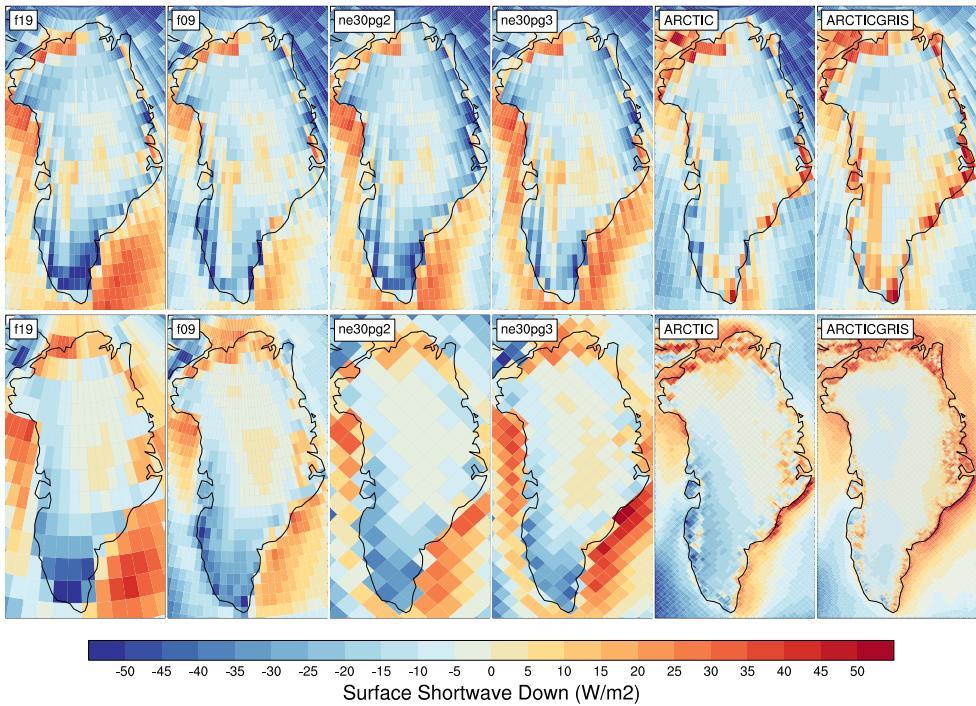


Figure 8. 1979-1998 northern hemisphere summer, incident shortwave radiation bias, computed as the difference (top) from CERES-EBAF, and (bottom) RACMO2.3p2 dataset. The differences in the top panel are found by mapping the model output to the 1° CERES-EBAF grid, and differences on the bottom panel are 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.

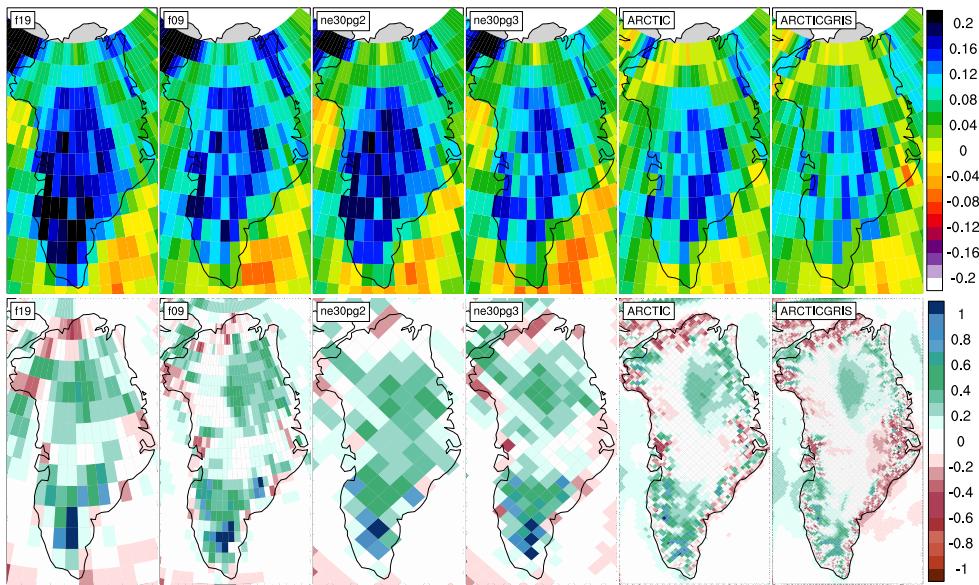


Figure 9. 1979-1998 northern hemisphere summer (top) total cloud fraction bias, relative to the CALIPSO-GOCCP dataset, and (bottom) precipitation rate bias, relative to the RACMO2.3p2 dataset. The CALIPSO-GOCCP differences are computed after mapping all model output to the 1° grid, whereas the RACMO 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.

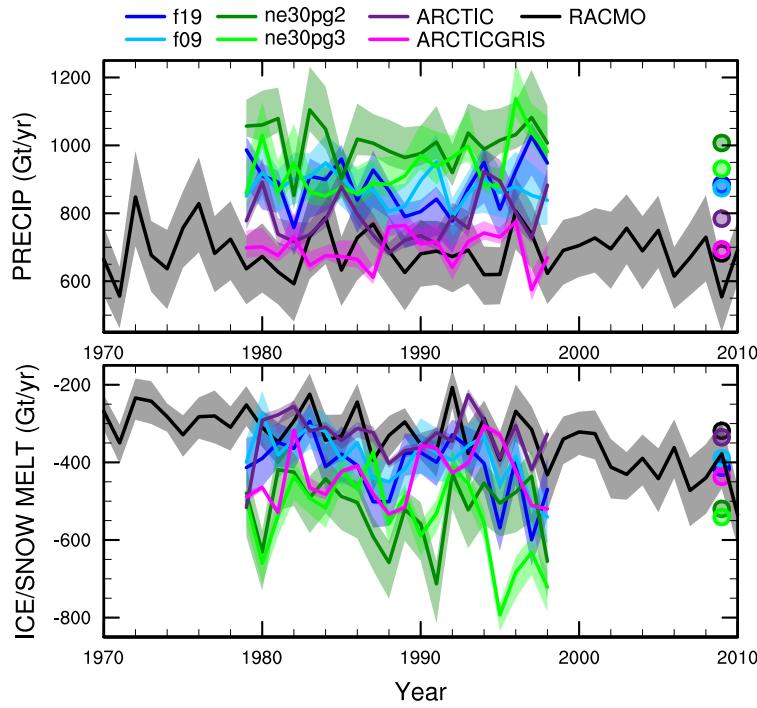


Figure 10. Time-series of annual (solid+liquid) precipitation (top) and annual runoff (bottom) integrated over the Greenland Ice Sheet for all six simulations and compared to the RACMO datasets. The time-series were generated using the common ice mask approach, which results in up to 4 ensembles, with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

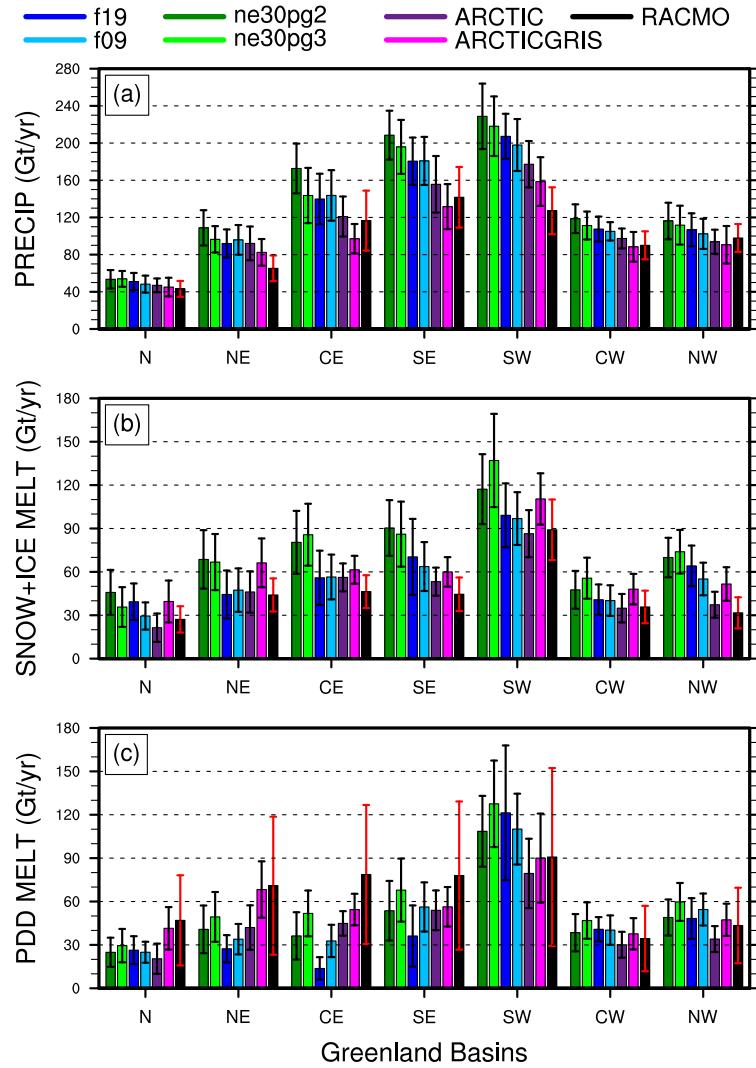


Figure 11. 1979–1998 basin integrated components of the SMB; (top) precipitation, (middle) ice/snow melt and (bottom) ice/snow melt estimated from the PDD method. Whiskers span the max/min of the four ensemble members generated from the common-ice-mask approach. Basin definitions are after Rignot and Mouginot (2012), and are found on the common ice masks using a nearest neighbor approach, and shown in Figure 12.

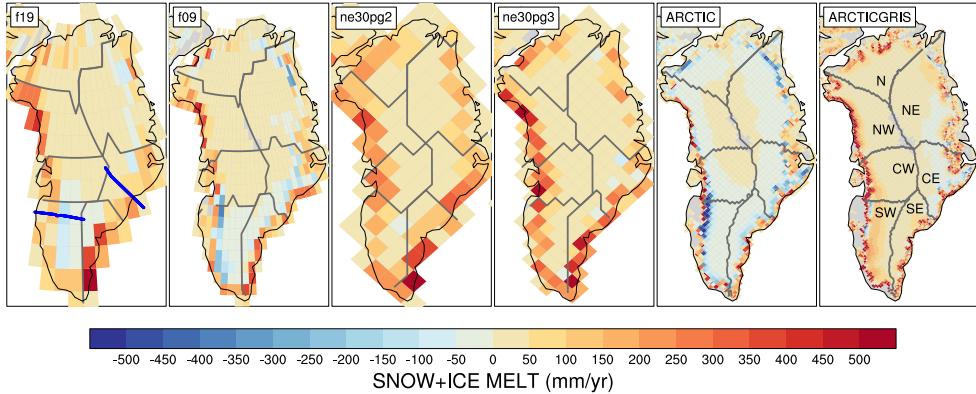


Figure 12. 1979–1998 ice/snow melt biases (in mm/yr) relative to RACMO2.3p2, evaluated on the native model grids. The Rignot and Mouginot (2012) basin boundaries are shown in grey for each model grid. Note that Figure 11 uses the basin boundaries for the two common ice masks, shown in the f19 and ne30pg2 panels, in computing the basin-scale integrals. Blue lines in the f19 panel show the location of the two transects plotted in Figure 13.

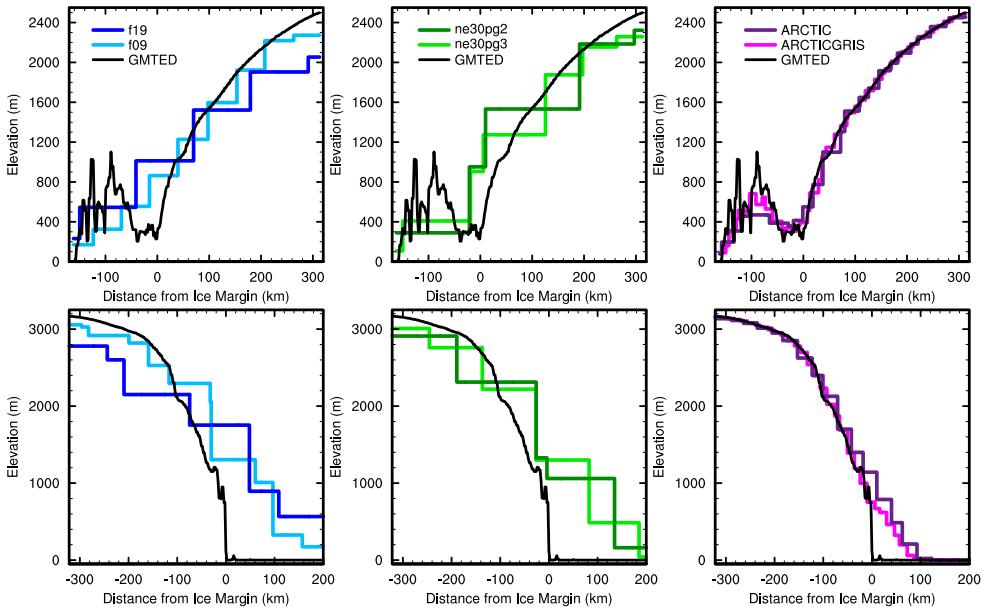


Figure 13. Model surface elevation along the (top) K-transect, and (bottom) a transect spanning the central dome down to the Kangerlussuaq glacier in southeast Greenland, for all model grids. The reference surface (GMTED) is a 1 km surface elevation dataset used for generating the CAM topographic boundary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

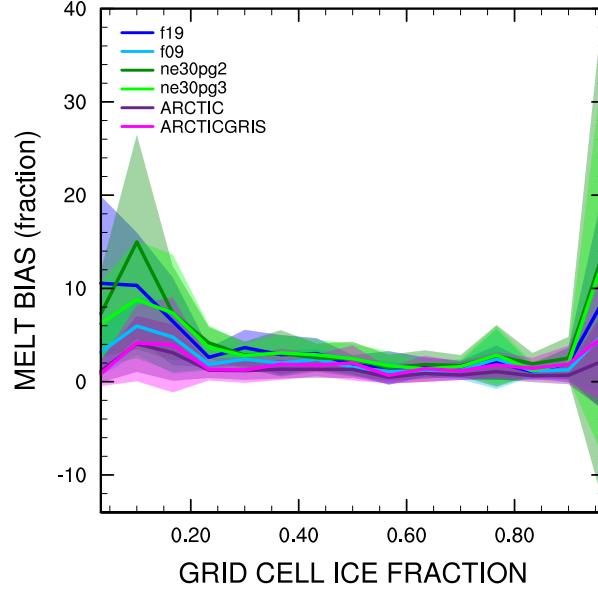


Figure 14. Fractional melt bias over the GrIS, computed relative to the RACMO datasets using the common ice mask approach, and conditionally sampled by grid cell ice fraction provided by the common ice masks. Solid lines are the mean of the distribution with \pm one standard deviation expressed by shading.

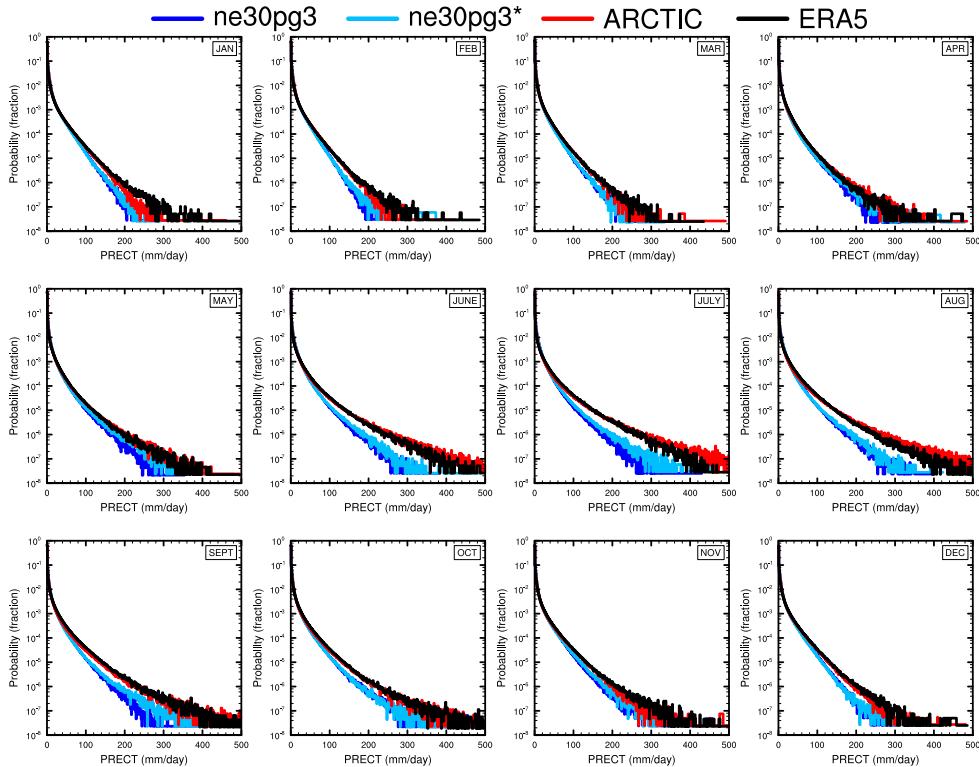


Figure 15. PDFs of the total precipitation rate associated with tracked storms, by month, in the ne30pg3, ne30pg3* and Arctic runs, and compared with the ERA5 dataset.

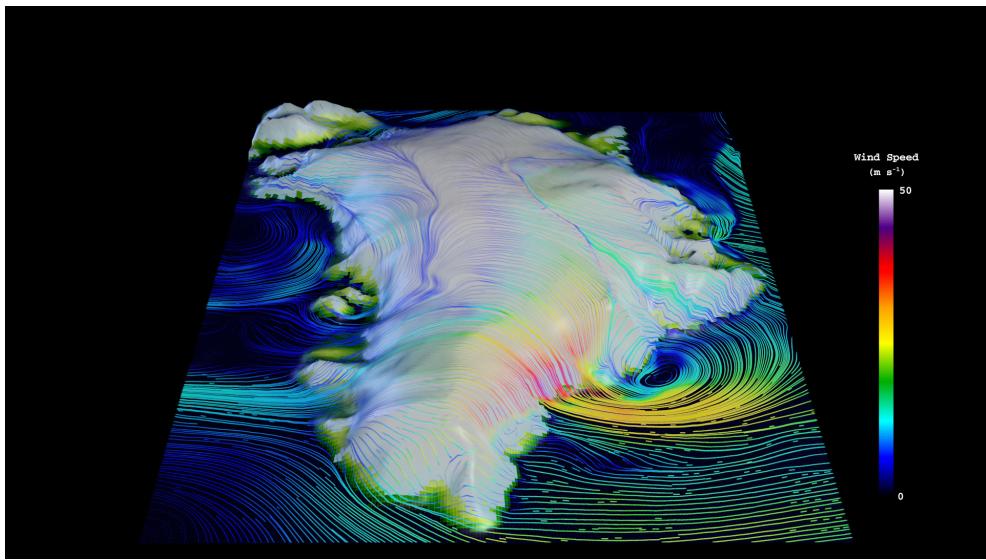


Figure 16. Snapshot of the lowest model level streamlines from the `Arctic - GrIS` visualization, with color shading denoting the wind magnitude.

OMEGA500, 1 year average, F2000climo, 32 levels

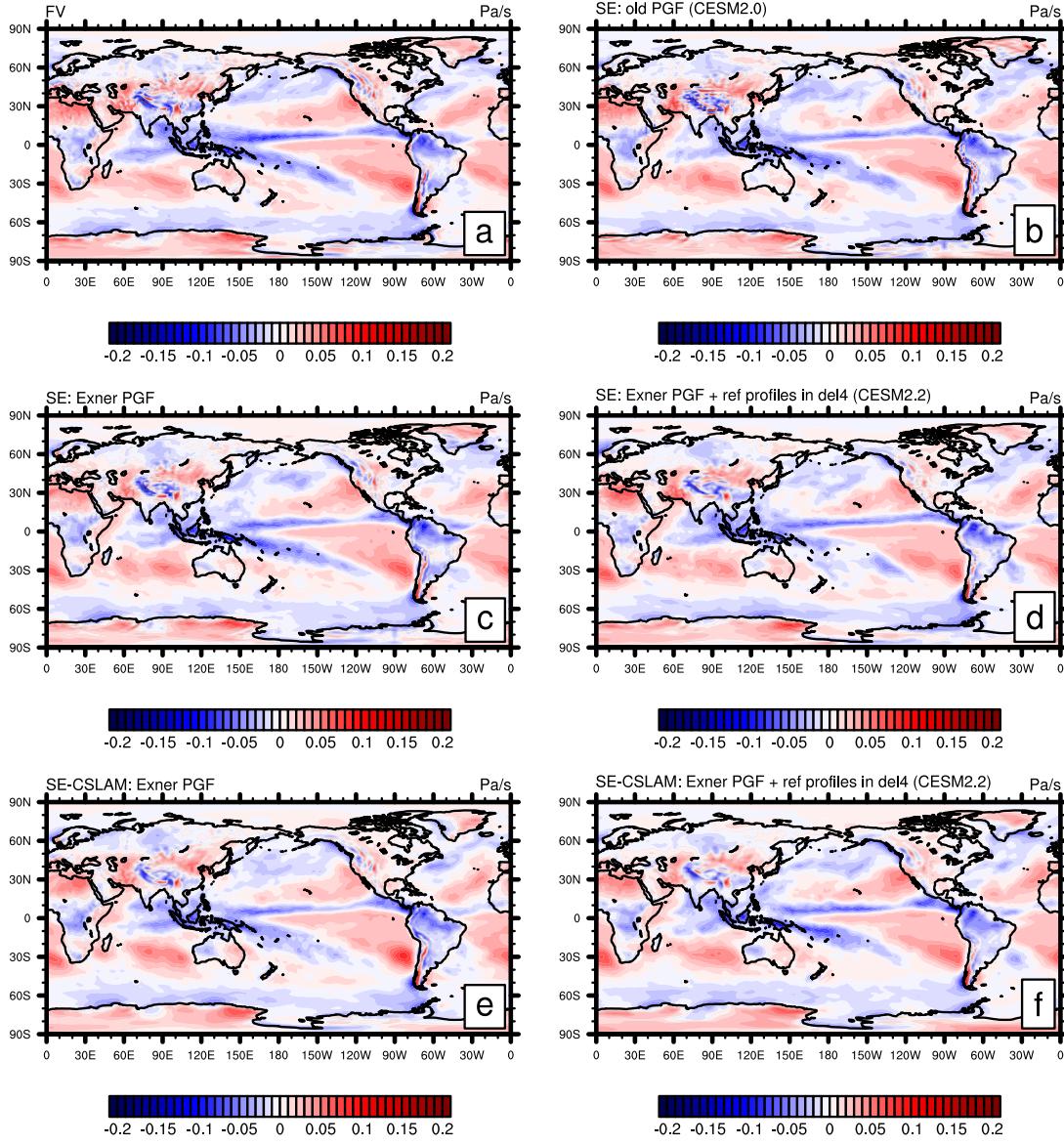


Figure A1. One year averages of vertical pressure velocity at 500hPa (**OMEGA500**) using (a) CAM-FV (Finite-Volume dynamical core) and (b-f) various versions of the spectral-element (SE) dynamical core at approximately 1° horizontal resolution and using 32 levels. (b) is equivalent to the CESM2.0 version of the SE dynamical core using the "traditional"/"old" discretization of the pressure-gradient force (PGF). Plot (c) is equivalent to configuration (b) but using the Exner form of the PGF. Plot (d) is the same as configuration (c) but also subtracting reference profiles from pressure and temperature before applying hyperviscosity operators (which is equivalent to the CESM2.2 version of SE in terms of the dynamical core). Plots (e) and (f) are equivalent to (c) and (d), respectively, by using the SE-CSLAM (`ne30pg3`) version of the SE dynamical core (i.e. separate quasi-uniform physics grid and CSLAM transport scheme).

OMEGA500, 18 months average, FHS94 forcing, 32 levels

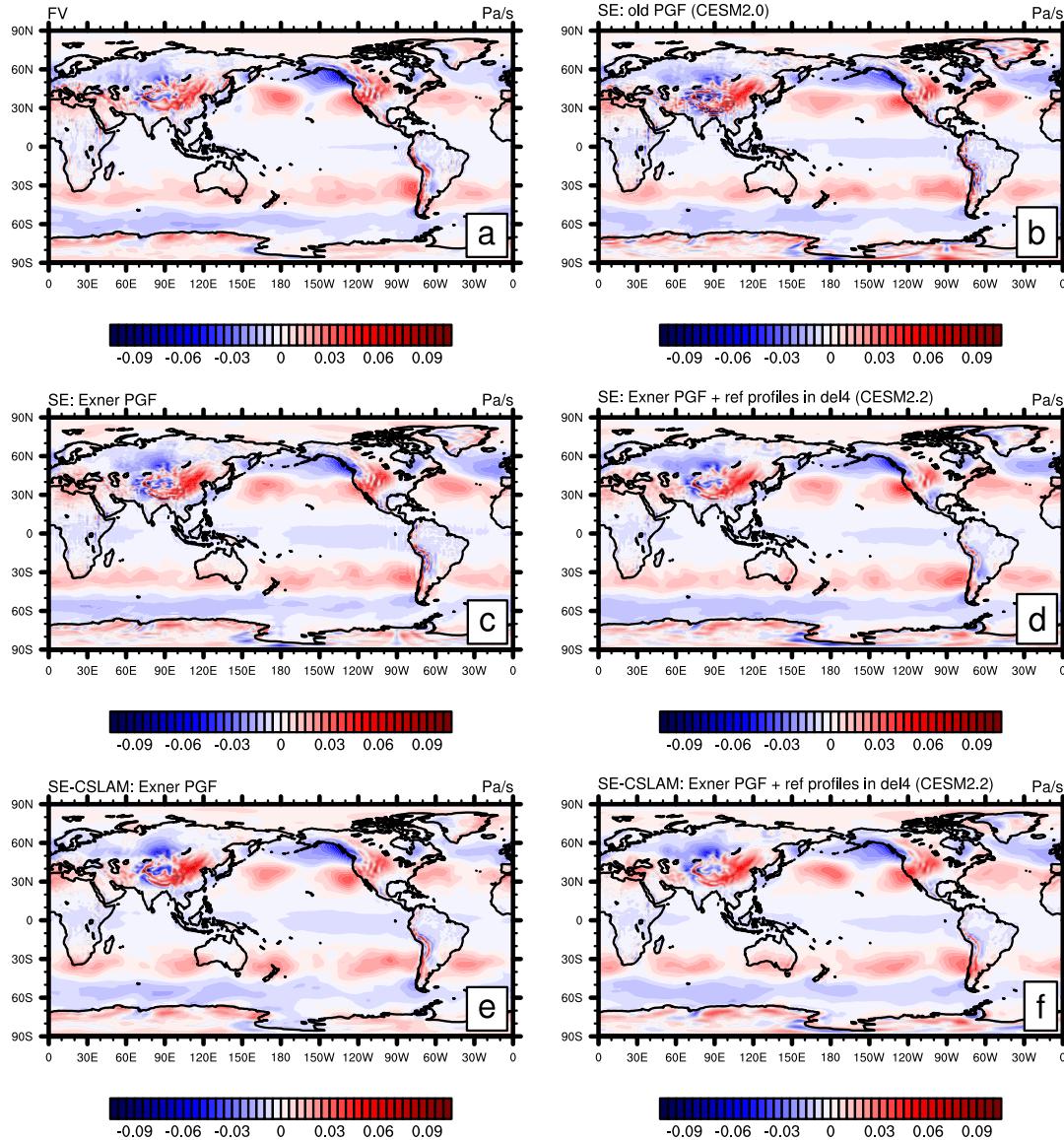


Figure A2. Same as Figure A1 but using modified Held-Suarez forcing and the average is over 18 months (excl. spin-up).