Convection-Permitting Simulations with the E3SM Global Atmosphere Model

Peter Martin Caldwell[1,1], Christopher Ryutaro Terai[2,2], Benjamin R Hillman[3,3], Noel D. Keen[4,4], Peter A Bogenschutz[5,5], Wuyin Lin[6,6], Hassan Beydoun[7,7], Mark A Taylor[3,3], Luca Bertagna[8,8], Andrew Bradley[9,9], Thomas C Clevenger[10,10], Aaron Sheffield Donahue[5,5], Christopher Eldred[10,11], James G Foucar[12,12], Jean-Christophe Golaz[1,1], Oksana Guba[3,3], Robert L Jacob[13,13], Jeff Johnson[14,14], Jagadish Krishna[8], Weiran Liu[15,15], Kyle G Pressel[16,16], Andrew G. Salinger[9,9], Balwinder Singh[17,17], Andrew Steyer[9,9], Paul Ullrich[18,18], Danqing Wu[19,19], Xingqiu Yuan[19,19], Jacob Shpund[8,8], Hsi-Yen Ma[20,20], Charles Sutton Zender[21,21], and Jayesh Krishna[19]

[1]Lawrence Livermore National Laboratory (DOE)
[2]University of California - Irvine
[3]Sandia National Laboratories
[4]Lawrence Berkeley National Laboratory (DOE)
[5]Lawrence Livermore National Laboratory
[6]Brookhaven National Laboratory
[7]Karlsruhe Institute of Technology
[8]Unknown
[9]Sandia National Laboratory
[10]Sandia National Lab
[11]LAGA, University of Paris
[12]Sandia National Laboratory (DOE)
[13]Argonne Notional Laboratory
[14]Cohere LLC
[15]UC Davis
[16]Pacific Northwest National Laboratory
[17]Pacific Northwest National Laboratory (DOE)
[18]University of California Davis
[19]Argonne National Lab
[20]LLNL
[21]University of California, Irvine

November 30, 2022

Abstract

This paper describes the first implementation of the d x=3.25 km version of the Energy Exascale Earth System Model (E3SM) global atmosphere model and its behavior in a 40 day prescribed-sea-surface-temperature simulation (Jan 20-Feb 28, 2020). This simulation was performed as part of the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) phase 2 model intercomparison. Effective resolution is found to be $\sim 6x the horizontal grid resolution despite using a coarser grid for physical parameterizations. Despite this new model being in an immature and untuned state, moving to 3.25 km grid spacing solves several long-standing problems with the E3SM model. In particular, Amazon precipitation is much more realistic, the frequency of light and heavy precipitation is improved, agreement between the simulated and observed diurnal cycle of tropical precipitation is excellent, and the vertical structure of tropical convection and coastal stratocumulus look good. In addition, the new model is able to capture the frequency and structure of important weather events (e.g. hurricanes, midlatitude storms including atmospheric rivers, and cold air outbreaks). Interestingly, this model does not get rid of the erroneous southern branch of the intertropical convergence zone nor the tendency for strongest convection to occur over the Maritime Continent rather than the West Pacific, both of which are classic climate model biases. Several other problems with the simulation are identified, underscoring the fact that this model is a work in progress.

Convection-Permitting Simulations with the E3SM Global Atmosphere Model

P. M. Caldwell[1], C. R. Terai[1], B. Hillman[2], N. D. Keen[3], P. Bogenschutz[1], W. Lin[4], H. Beydoun[1], M. Taylor[2], L. Bertagna[2], A. M. Bradley[2], T. C. Clevenger[2], A. S. Donahue[1], C. Eldred[2], J. Foucar[2], J.-C. Golaz[1], O. Guba[2], R. Jacob[5], J. Johnson[6], J. Krishna[5], W. Liu[7], K. Pressel[8], A. G. Salinger[2], B. Singh[8], A. Steyer[2], P. Ullrich[7], D. Wu[5], X. Yuan[5], J. Shpund[8], H.-Y. Ma[1], C. S. Zender[9]

[1]Lawrence Livermore National Lab, Livermore, CA, USA
[2]Sandia National Lab, Albuquerque, NM, USA
[3]Lawrence Berkeley National Lab, Berkeley, CA, USA
[4]Brookhaven National Lab, Upton, NY, USA
[5]Argonne National Lab, Lemont, IL, USA
[6]Cohere Consulting LLC, Seattle, WA, USA
[7]Department of Land, Air, and Water Resources, University of California-Davis, Davis, CA, USA
[8]Pacific Northwest National Lab, Richland, WA, USA
[9]Departments of Earth System Science and Computer Science, University of California, Irvine, USA

Key Points:

- Describes the Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM)
- SCREAM performs well in a 40 day boreal winter simulation at 3.25 km Δx
- Resolving deep convection solves many long-standing climate model biases

---

Corresponding author: Peter Caldwell, `caldwell19@llnl.gov`

Abstract

This paper describes the first implementation of the $\Delta x = 3.25$ km version of the Energy Exascale Earth System Model (E3SM) global atmosphere model and its behavior in a 40 day prescribed-sea-surface-temperature simulation (Jan 20 through Feb 28, 2020). This simulation was performed as part of the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains (DYAMOND) phase 2 model intercomparison. Effective resolution is found to be $\sim 6\times$ the horizontal dynamics grid resolution despite using a coarser grid for physical parameterizations. Despite this new model being in an immature and untuned state, moving to 3.25 km grid spacing solves several long-standing problems with the E3SM model. In particular, Amazon precipitation is much more realistic, the frequency of light and heavy precipitation is improved, agreement between the simulated and observed diurnal cycle of tropical precipitation is excellent, and the vertical structure of tropical convection and coastal stratocumulus look good. In addition, the new model is able to capture the frequency and structure of important weather events (e.g. tropical cyclones, extratropical cyclones including atmospheric rivers, and cold air outbreaks). Interestingly, this model does *not* get rid of the erroneous southern branch of the intertropical convergence zone nor the tendency for strongest convection to occur over the Maritime Continent rather than the West Pacific, both of which are classic climate model biases. Several other problems with the simulation are identified, underscoring the fact that this model is a work in progress.

Plain Language Summary

This paper describes the new global 3.25 km version of the Energy Exascale Earth System Model (E3SM) atmosphere model and its behavior in a 40-day northern-hemisphere wintertime simulation. In exchange for huge computational expense, this high-resolution model avoids many but not all biases common in lower-resolution models. It also captures several types of extreme weather that would simply not be resolved in lower-resolution models. Several opportunities for further development are identified.

1 Introduction

Because the processes controlling Earth's weather and its climatology are complex and inter-related, numerical models are a critical tool for predicting future conditions. Global coverage is necessary because local behavior propagates rapidly to distant areas of the globe. Simulating the whole planet imposes severe computational challenges, however. In the past, this has typically been handled by coarsening model grid spacing until simulations became affordable on the machines of the time. As of 2020, this translated to horizontal grid spacing of ∼10 km for weather models (which simulate days to weeks at a time) and ∼100 km for climate models (which are typically run for centuries). These grid spacings are too coarse to capture many important atmospheric processes.

The impacts of sub-grid scale processes on model climate are instead *parameterized* based on available grid-scale quantities. Typical parameterized processes include turbulent transport and mixing, gravity-wave motions, greenhouse gas and aerosol chemistry and physics, radiative transfer, and cloud physics. Cloud parameterizations are in particular complicated yet important for accurate predictions. Vapor transport, collisions, and other physics involving micron-scale water drops or ice crystals (collectively called microphysics) are critical for predicting precipitation and future changes in cloud shading. Condensation and evaporation of clouds and resulting fractional cloudiness within a grid cell (often called macrophysics) involve larger spatial scales but are still important to parameterize in conventional models. Condensational heating in convective clouds causes narrow but intense upward vertical motions which are a primary source of vertical transport of heat, moisture, and momentum in the tropical atmosphere (Riehl & Malkus, 1958). Because the microphysics and macrophysics of these intense updrafts are tightly entwined with their motions, convective parameterizations tend to include their own microphysics and macrophysics treatments. Inconsistency between microphysical treatments for convective- versus resolved-scale motions is a large source of model biases (Song & Zhang, 2011; Storer et al., 2015). Convection in general has proven to be particularly difficult to parameterize from quantities available on the grid scale (Randall et al., 2003; Stevens & Bony, 2013) and has been implicated as a primary source of climate change uncertainty (Sanderson et al., 2008; Sherwood et al., 2014).

Another challenge posed by coarse resolution is interaction with Earth's surface. Topography is not resolved at typical global model grid spacing and in fact must be even further smoothed to avoid model instability (Lauritzen et al., 2015). Because topography can force air upwards until it condenses, smoothing out high mountain peaks causes major problems for cloud and precipitation climatology (Giorgi & Marinucci, 1996). Insufficient surface roughness means wind stresses are also too weak over smoothed topography and must be parameterized. Subgrid-scale surface heterogeneity also poses problems for coarse models (Prein et al., 2015). And while the focus of this paper is on simulations with prescribed sea surface temperature, it is worth noting that ocean eddies on spatial scales <10 km play a critical role in heat transport (Maslowski et al., 2008) and their parameterization has proven as problematic for ocean models as convective clouds are for atmosphere models (Hewitt et al., 2020). Ocean/atmosphere interaction at convection- and ocean-eddy resolving scales has not (to our knowledge) been studied but is also likely to have important impacts on model behavior.

Because so much is lost at coarse resolution, the global atmospheric modeling community has long pushed towards higher resolution. Unsurprisingly, better topographic resolution improves orographic precipitation, snowpack, and stream flow (Pope & Stratton, 2002; Duffy et al., 2003; Delworth et al., 2012; Caldwell et al., 2019). Sea breeze effects become better captured as coastal boundaries are better resolved (Boyle & Klein, 2010; Love et al., 2011). Because finer grid spacing allows smaller spatial and temporal scales to be resolved, higher-resolution GCMs also better capture extreme precipitation events (Iorio et al., 2004; Wehner et al., 2014; Terai et al., 2018). As GCM grid spacing falls to 25 km or less, tropical cyclones begin to be resolved (Atlas et al., 2005; Bacmeister et al., 2014; Wehner et al., 2014; Caldwell et al., 2019), though capturing details of spatial structure requires still finer resolution (Judt et al., 2021). Some classic model problems are, however, relatively unaffected by reducing grid spacing to 25 km. In particular, increased resolution does not get rid of the erroneous southern branch of the Intertropical Convergence Zone (ITCZ) common in climate models (McClean et al., 2011; Bacmeister et al., 2014; Caldwell et al., 2019). Simulation of the Madden-Julian Oscillation (MJO) is likewise unaffected (Jung et al., 2012; Bacmeister et al., 2014). In addition, precipitation improvement has been found primarily in wintertime (Duffy et al., 2003).

It is notable that these remaining deficiencies are related to convective motions which are unresolved even at high GCM resolutions. Given the aforementioned difficulty of parameterizing convection, this situation is perhaps expected. A small number of global models with grid spacing fine enough to explicitly resolve the largest convection events (hereafter called global convection-permitting models or GCPMs) have also been built. The number of these models has exploded recently because recent advances in computing have tended towards allowing more calculations to be performed in parallel rather than making individual calculations faster. Conventional global simulations already exploit all available parallelism, so won't run faster on these new machines. Higher horizontal resolution is a ready source of increased parallelism, so is attractive in this new computing environment. Unfortunately, smaller timesteps are needed to resolve finer spatial scales. Thus even if all columns could be computed in parallel, a given integration at finer resolution requires more timesteps and therefore has a longer time-to-solution.

As a result, GCPM simulations can't be run as routinely nor as long as conventional global models.

The history of GCPM modeling is nicely summarized in Satoh et al. (2019). Briefly, the first GCPM was NICAM, described in Tomita et al. (2005); Satoh et al. (2008, 2014). For several years its only companion was the Multiscale Modeling Framework (MMF) described in Grabowski and Smolarkiewicz (1999), Randall et al. (2003), and Grabowski (2016). The MMF isn't exactly a GCPM, however, as it replaces the physical parameterizations inside each grid cell of a conventional GCM with a limited-area convection-permitting model (CPM). The MMF is much cheaper than a GCPM because embedded CPMs are typically contained within a single computational node, avoiding MPI communication costs. Additionally, the grid of the CPM is decoupled from that of the GCM, so CPMs are typically 2d and have domain size smaller than the GCM grid cell width. The second GCPM was NASA's GEOS model (Putman & Suarez, 2011), which was used as a synthetic laboratory for designing and testing satellite campaigns (Gelaro et al., 2015) in addition to more general analysis. In the last few years, enough new GCPMs have been developed to warrant their own intercomparison. Called DYnamics of the Atmospheric general circulation Modeled On Nonhydrostatic Domains (DYAMOND), the first phase of this intercomparison focused on a 40 day simulation starting Aug 1, 2016 and included 8 models with grid spacing less than 5 km globally. An overview of this intercomparison is presented in Stevens et al. (2019). Stevens' study shows striking agreement in outgoing longwave radiation, precipitation, and precipitable water between participating models. Shortwave radiation differs between models, presumably due to differences in low clouds, which aren't well resolved at GCPM resolutions. Models also tend to predict a spurious peak in precipitation just south of the equator, suggesting that km-scale resolution is not the solution to the double-ITCZ problem endemic to conventional climate models (Li & Xie, 2014). Based on the success of this first intercomparison, a second DYAMOND intercomparison (called DYAMOND2) is now underway. The current paper documents a new contribution to DYAMOND2.

GCPMs can be viewed as a natural and beneficial extension of conventional GCMs to finer resolution, but they can also be seen as the extension towards larger domains of a robust research community focused on limited-area CPMs. Beginning with the explicit simulation of a single convective event (Ogura, 1963), cloud-resolving simulations have steadily grown in duration and domain size. Recently, Bretherton and Khairoutdinov (2015) and Narenpitak et al. (2017) describe multi-month 4 km simulations simulating the entire tropical channel between 45°N and 45°S. CPMs tend to offer more benefit for summertime convection rather than wintertime cyclones (Prein et al., 2015), as may be expected given the spatial scale of these storm types. Limited-area CPM research suggests that resolution finer than $\sim$4 km is needed to resolve convective ensemble statistics (Weisman et al. (1997); also found for GCPMs by Miyamoto et al. (2013)) but resolution finer than that adds relatively little value (Kain et al., 2008; Schwartz et al., 2009; Langhans et al., 2013). Cloud fraction tends to decrease as resolution becomes finer (Prein et al., 2013; Langhans et al., 2013; Fosser et al., 2014), a feature also found in GCPMs (Noda et al., 2010; Hohenegger et al., 2020).

A great deal of CPM research has been organized around the Global Energy and Water Cycle Experiment Cloud Systems Study (GCSS). As described in a review by Krueger et al. (2016), GCSS organized intercomparisons of CPMs and single-column versions of GCMs for intensive observing periods spanning a wide variety of cloud regimes. These intercomparisons clarified processes CPMs could and couldn't handle, often leading to idealized follow-up experiments. These follow-up studies have proven invaluable for providing process insights and subsequent model improvements. DYAMOND is in some ways the reincarnation of GCSS for the next generation of models.

In general, high-resolution regional studies have added value primarily by resolving fine-scale features rather than through upscale effects onto scales resolved by conventional models (Prein et al., 2015; Caldwell, 2010). One potential reason for this is that lateral boundary conditions impose strong constraints on domain-averaged properties (Edman & Romps, 2014). Thus while GCPMs may be overkill for looking at fine-scale features which could be studied via limited-area models, they offer fresh new potential to solve long-standing deficiencies in the general circulation.

The goal of this paper is to introduce the GCPM being developed by the Energy Exascale Earth System Model (E3SM) project and to provide an initial look at its behavior in the DYAMOND2 case study. Details about this model are provided in Section 2. Sections 3-5 describe experimental design, data for evaluation, and computational performance (respectively). Results in Section 6 are broken into an analysis of effective resolution in subsection 6.1, general attributes in subsection 6.2, clouds and radiation in subsection 6.3, precipitation in subsection 6.4, and specific weather phenomena in subsequent subsections. Conclusions follow in Section 7.

2 Model Description

As described in Golaz et al. (2019), the E3SM project was born from the US Department of Energy (DOE)'s need for quantitative information about future climate for use in energy-sector decisions. Given DOE's leadership in high-performance computing, it has been natural for E3SM to focus on compute-intensive frontiers in climate science. One of those efforts has been to develop a new GCPM called the Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM).

Our ultimate goal is to make SCREAM as fast as possible on exascale machines by writing it in C++ using the Kokkos library (Carter-Edwards et al., 2014) for performance portability. See Bertagna et al. (2019, 2020) for a description of our design strategy and initial performance results. We are, however, approaching this goal by first creating a prototype version in Fortran using the existing E3SM atmosphere infrastructure. This initial implementation - which is the focus of the current study - is being used as the template for the C++ implementation as well as giving us an early look at model behavior. The final implementation should be scientifically identical to this prototype version but will be much faster because of its ability to run on GPU-powered computers.

Our strategy has been to make our first implementation as simple as possible and to start using it for science as quickly as possible. This strategy is expected to result in sub-optimal skill in our first implementation, but allows us to more rapidly produce, understand, and improve our model. We believe that it is better to start with an overly-simple model and to add complexity as needed rather than to start with a more sophisticated/accurate model which we don't understand.

Simplicity in particular means that SCREAM consists solely of nonhydrostatic fluid dynamics, a turbulence/cloud fraction scheme, a microphysics scheme, a radiation scheme, an energy fixer, and prescribed-aerosol functionality. These pieces are described in the subsections below. SCREAM does not parameterize sub-grid scale gravity-wave drag or deep convection. This initial implementation uses the E3SM land model described in Golaz et al. (2019). It also uses prescribed-ice mode from CICE4 (Hunke & Lipscomb, 2008) to compute surface fluxes, snow depth, albedos, and surface temperature, resetting sea ice thickness after each timestep to 2 m in the northern hemisphere and 1 m in the southern hemisphere. Sea surface temperature (SST) is prescribed.

2.1 Fluid Dynamics

SCREAM's fluid-dynamics solver (hereafter dycore) solves the nonhydrostatic equations of motion in a rotating reference frame with the shallow atmosphere approximation and a hyperviscosity based turbulence closure. It additionally transports several constituents, including multiple forms of water and various aerosols. It is implemented in the High Order Method Modeling Environment (HOMME) (Dennis et al., 2005, 2012; Evans et al., 2013). HOMME contains several spectral element based dycores, including the hydrostatic dycore used by E3SM (Rasch et al., 2019; Golaz et al., 2019; Caldwell et al., 2019) and the Community Earth System Model (Small et al., 2014; S. Zhang et al., 2020). We refer to the new nonhydrostatic dycore developed for SCREAM as HOMME-NH.

HOMME-NH uses the nonhydrostatic formulation of the equations from Taylor et al. (2020). It solves the equations in a terrain following mass based vertical coordinate (Kasahara, 1974; Laprise, 1992), with prognostic equations for the three components of the velocity field, the mass-coordinate pseudo-density, the geopotential height, and a thermodynamic variable, for which we use virtual potential temperature. The prognostic equations consist of the time-reversible adiabatic terms from Taylor et al. (2020), a ∇^4 hyperviscosity following Dennis et al. (2012) and Guba et al. (2014), and a sponge layer at the model top. For the adiabatic terms, we use a structure preserving formulation in order to preserve the discrete Hamiltonian and produce an energetically consistent model. The horizontal discretization uses the collocated mimetic spectral element method from Taylor and Fournier (2010), with conservative and monotone semi-Lagrangian tracer transport (Bradley et al., 2019). The vertical discretization uses a Lorenz staggered extension of the mimetic centered difference from Simmons and Burridge (1981). With this vertical staggering, prognostic variables are located at level midpoints, with the exception of the vertical velocity and geopotential, which are located at level interfaces. For the vertical transport terms, we use a vertically Lagrangian approach adapted from Lin (2004).

For the temporal discretization, we use a Horizontally Explicit Vertically Implicit (HEVI) approach (Satoh, 2002), discretized with an IMplicit-EXplicit (IMEX) Runge Kutta method (Ascher et al., 1997). The HEVI splitting decomposes the equations into a set of terms which represent vertically propagating acoustic waves (treated implicitly), and the remaining terms which include all horizontal derivatives (treated explicitly). We use a highly efficient IMEX method from Steyer et al. (2019) and Guba et al. (2020), with a 2nd-order accurate coupling of a high-stage high-CFL scheme for the explicit terms and a Diagonally Implicit Runge Kutta (DIRK) scheme for the implicit terms. Due to the use of the Laprise mass coordinate, the vertical acoustic waves are isolated to only two terms in the equations for vertical velocity and geopotential solved at level interfaces, leading to an implicit system for a single variable.

There are several sources of dissipation in the dynamical core. The ∇^4 hyperviscosity is the largest. It is applied to all prognostic variables and on every model layer, with a hyperviscosity coefficient of 2.5×10^{10} m^4 s^{-1} for the 3.25 km grid. Because tuning at 3.25 km is expensive, we chose this value based on a Δx^3 scaling of the hyperviscosity coefficient used by E3SM at lower resolutions. For the model-top sponge-layer, we applied a ∇^2 Laplacian operator to all prognostic variables according to the reference-pressure based ramp function from Lauritzen et al. (2011). This ramp starts at layer 14 ($\sim$ 19hPa) with a coefficient of 0.189×10^{-4} m^2 s^{-1} and ramps up to 6.93×10^{-4} m^2 s^{-1} at the model top. In addition, vertical dissipation is introduced by the monotone vertical remap operator. A smaller amount of dissipation is also generated by the Runge-Kutta timestepping.

2.2 Model Grid

Our horizontal grid for dynamics is a cubed-sphere grid with 1024×1024 spectral elements on each face, denoted ne1024. The total number of elements is therefore 6,291,456. Within each element, fields are represented by degree-3 polynomials, using nodal values

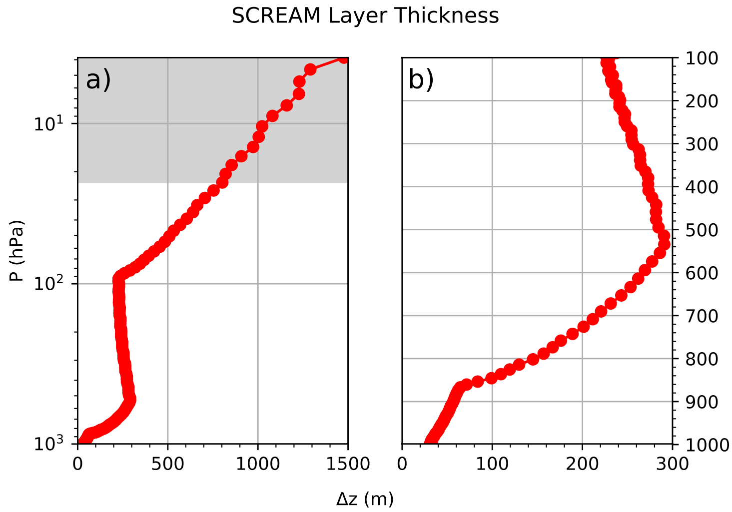

Figure 1. SCREAM grid spacing. Panel a shows the complete vertical grid using logarithmic pressure to emphasize the upper atmosphere. Panel b zooms in on the troposphere using linear pressure spacing to emphasize lower levels. The sponge layer is indicated by gray-shading.

on a 4×4 grid of Gauss-Lobatto-Legendre (GLL) nodes. The edge and corner nodes are shared by adjacent elements, resulting in an average spacing between GLL nodes of ∼3.25 km. The nonuniform spacing of GLL nodes presents some challenges to the physical parameterizations (Herrington et al., 2019), which we avoid by evaluating the parameterizations on a uniformly spaced 2×2 grid within each spectral element. This *physics grid* has 4/9 as many physics columns as would be in a GLL-collocated physics grid. Tests show that the 2×2 physics grid provides very similar results to simulations with physics running on every GLL node (Hannah et al., 2021). Our land model is run on a 1/8° latitude-longitude grid. SST and sea ice are computed on the high-resolution ocean grid used by Caldwell et al. (2019), which tapers from 18 km in the tropics to 6 km near the poles. The ocean and sea ice grids have minor impact since SST and ice extent are interpolated from 0.5° datasets. It would be better to have all surface calculation on the 3.25 km atmosphere grid, but resolution challenges with the E3SM input data toolchain made doing so impractical for this initial simulation.

We use a relatively-fine 128 layer vertical grid with a model top at 40 km (2.25 hPa) and a sponge layer above ∼19 hPa (as as described in the previous subsection). Vertical grid spacing is presented in Fig. 1. Representative grid spacing in the boundary layer is ~ 50 m, in trade Cu is ~ 100 m, and in tropical cirrus anvils is ~ 250 m.

2.3 Topography

To generate the SCREAM v0 surface topography, we use the NCAR topography tool chain (Lauritzen et al., 2015) to first compute the unfiltered height field on the atmosphere grid. We then smooth the height field on the GLL grid using 16 iterations of the spectral element Laplace operator. To quantify the amount of smoothing, we follow Evans et al. (2013) and compare power spectra $E(k)$ from the spherical harmonic representation of the filtered and unfiltered height field, and then compute the lowest wave number $k_{1/2}$ for which the smoothing has reduced $E(k_{1/2})$ by 50%. The SCREAM v0 topography has a $k_{1/2}$ corresponding to wavelength $6.4\Delta x$.

2.4 Clouds and Turbulence

Boundary layer clouds and their associated circulations are still largely unresolved at 3.25 km so a parameterization of interaction between clouds and turbulence is critical. Because GCPMs push the boundary of computational possibility, it is important that these processes are handled efficiently. These goals are accomplished in SCREAM via the Simplified Higher Order Closure (SHOC; Bogenschutz & Krueger, 2013). Similar to other widely used assumed PDF-based schemes (Golaz et al., 2002; Cheng & Xu, 2008), SHOC computes subgrid-scale liquid cloud and turbulence using an assumed double-Gaussian probability density function (PDF). SHOC is more efficient than the aforementioned schemes, however, because it diagnoses rather than prognoses the higher order moments that are needed to close the double Gaussian PDF. Bogenschutz and Krueger (2013) demonstrate that when SHOC is used in limited-area cloud-resolving simulations of boundary layer clouds, the solution is insensitive to the horizontal resolution choice. This is in contrast to a standard 1.5-order TKE closure, which suffers from large horizontal-resolution sensitivity when used in the same cloud-resolving model.

SHOC has undergone several updates since Bogenschutz and Krueger (2013) to improve numerical stability and performance among the wider range of regimes SHOC is subjected to in a global model. Chief among these updates is the implementation of an implicit diffusion solver, a revised formulation of the turbulence length scale to better achieve vertical convergence, and a revised formulation of the eddy diffusivities for the stable boundary layer (similar to those implemented in Bretherton and Park (2009)). The turbulence length scale is now a continuous formulation that avoids the separate definitions of in-cloud vs sub-cloud length scales documented in Bogenschutz and Krueger (2013) and performs scientifically similarly to the original formulation.

In addition to the liquid cloud fraction supplied by SHOC, we require an ice cloud fraction. For simplicity, our initial implementation includes the same ice cloud fraction used by E3SMv1 and inherited from CESM1. This implementation assumes ice cloud starts forming when an ice-modified relative humidity $\mathrm{RH}_i = (q_v + q_i)/q_{\mathrm{sat,i}}$ reaches a user-specified minimum value and reaches 100% at a user-specified maximum value. Unfortunately, these parameters were left at their low-resolution E3SMv1 defaults of 80% and 105% (respectively) in our DYAMOND2 simulation. The impact of this mistake is shown in Section 6.2.

2.5 Microphysics

SCREAM microphysics is based on the Predicted Particle Properties (P3) scheme of Morrison and Milbrandt (2015) taken from version 4.1 of the Weather Research and Forecasting (WRF) model (Skamarock et al., 2019). The novel feature of P3 is that it avoids arbitrary cutoffs between cloud-borne and precipitating ice categories by employing a single ice category which is allowed to evolve naturally from small pristine crystals into large and possibly rimed snowflakes. While the WRF version of P3 allows for multiple simultaneous populations of these ice crystals within a grid cell, SCREAM currently only supports a single population because the modest improvements from multiple ice populations reported in Milbrandt and Morrison (2016) were not deemed worth the additional software engineering time required to support this feature. The liquid phase of the P3 scheme - like most microphysics codes - separates drops into cloud and rain categories.

One feature of this scheme is the clever use of supersaturation to diagnose condensation, evaporation, sublimation, and deposition. This approach works well for Large-Eddy simulations (LES) which explicitly model each updraft, but probably underpredicts condensation for the 3.25 km grid spacing used in SCREAM (Morrison & Grabowski, 2008). The great benefit of this supersaturation approach is that it treats ice growth at the expense of nearby liquid (Wegener, 1911; Bergeron, 1935; Findeisen, 1938, hereafter

WBF process) in a very natural way. Unfortunately, allowing supersaturation in P3 directly conflicts with the instantaneous saturation adjustment assumption which forms the foundation of SHOC's PDF. For consistency, our P3 implementation instead handles vapor deposition, sublimation, and the associated WBF process following Gettelman and Morrison (2015). In particular, maximum overlap between liquid and ice is assumed when liquid and ice coexist, leading to efficient liquid-to-ice transition via the WBF process. If all liquid is removed within a microphysics timestep, vapor deposition onto ice for the remainder of that timestep is computed based on cell-average water vapor content.

Another inconsistency between SHOC and the WRF version of P3 is the use of fractional cloudiness and precipitation. P3 neglected all sub-grid variability such that cloud and precipitation covered the entire grid cell where they exist and otherwise the cell was entirely devoid of condensate. SHOC provides fractional cloudiness, so we modified P3 to only operate in the cloudy or precipitating portion of each cell. Our fractional cloudiness implementation is similar to Jouan et al. (2020), which was implemented in WRF P3 around the same time as we made our modifications. The fraction of each cell containing precipitation is also important. In SCREAM this was taken to be equal to the largest cloud fraction of all cells including and above the layer of interest. This approach is crude (as noted by Zheng et al., 2020) and will be a subject of future research.

SHOC's subgrid assumptions require further modifications. SHOC uses a double-Gaussian PDF to model subgrid-scale variations in liquid water potential temperature, total water mixing ratio, and vertical velocity. Larson and Griffin (2013) provide an analytical formulation for incorporating SHOC's variability into microphysical processes expressed as power functions. We intend to implement this consistent scheme in our version of P3 eventually, but for the moment we have instead implemented the partially-consistent approach from (Morrison & Gettelman, 2008), which instead assumes a gamma distribution for liquid water mixing ratio and ignores subgrid temperature variations. The benefit of the gamma distribution is that the expected value of a power-law-based microphysical process rate can be written as that power law applied to the cell-mean value multiplied by an easily-calculated scaling factor.

Finally, water vapor saturation was changed in our version of P3 to be consistent with the Murphy and Koop (2005) (MK) implementation used in SHOC. MK is more accurate at very low temperatures than the Flatau et al. (1992) implementation originally used in P3, but is more computationally expensive. We found this performance difference, however, to have a negligible impact on total run time.

2.6 Radiation

Gas optical properties and radiative fluxes are computed using the RTE+RRTMGP radiative transfer package (Pincus et al., 2019). Active gases in SCREAM include H_2O, CO_2, O_3, N_2O, CO, CH_4, O_2, and N_2. Cloud and aerosol optical properties are computed as in the Community Atmosphere Model (CAM). The approach is described in detail in Neale et al. (2012). Briefly, condensed phase optical properties (extinction coefficient, single scattering albedo, and asymmetry parameter for shortwave bands and absorption coefficient for longwave bands) are computed per unit mass for liquid, ice, and aerosol, then multiplied by the appropriate mass mixing ratio for use in RTE+RRTMGP.

Liquid cloud optical properties are calculated from a table-lookup after being computed offline using a Mie scattering code (Wiscombe, 1996) based on the assumption (taken from microphysics) that the total number of liquid drops with diameter D follows a gamma histogram

$$n(D) = N_0 D^{\mu} e^{-\lambda D}$$

with intercept parameter N_0, slope parameter λ, and spectral size dispersion μ taken every timestep from P3. In this initial implementation, in-cloud liquid water content is assumed to be homogeneously distributed. This is inconsistent with our implementation of P3, which (as noted above) assumes a gamma distribution for spatial variations in cloud liquid. Fixing this inconsistency is a future goal.

Ice cloud optical properties are computed for each shortwave and longwave band used by the radiation code using a lookup table based on the modified anomalous diffraction approximation (Mitchell, 2002). The only input to these table lookups is ice effective radius, which is computed in P3. Because ice mass-density relationships are different for different size and riming regimes, ice effective radius is calculated via a table lookup described in Morrison and Milbrandt (2015). Because P3 merges the ice and snow categories used by traditional microphysics schemes into a single ice mode, radiation naturally acts on all frozen hydrometeors. Aerosol optical properties are specified in a lookup table as a function of wet refractive index and wet surface mode radius (Ghan & Zaveri, 2007).

Vertical overlap of partially-cloudy cells is accounted for by assuming maximum-random overlap (Geleyn & Hollingsworth, 1979) using the Monte Carlo Independent Column approach (MCICA Pincus et al., 2003).

2.7 Prescribed Aerosol

E3SMv1 uses a 4 Mode Aerosol Model (MAM4 Liu et al., 2016). For computational efficiency, we employ a version where this modal aerosol information is prescribed using monthly-average climatologies interpolated to the model grid from a 1° resolution E3SMv1 simulation. Implementation and use of prescribed-aerosol functionality is described in K. Zhang et al. (2013), Lebassi-Habtezion and Caldwell (2015), and Shi and Liu (2018). The default prescribed-aerosol implementation scales aerosols by different random perturbations every day to improve agreement between prescribed- and prognostic-aerosol simulations at high latitudes. These random daily jumps are confusing for analysis of short timeseries, so we've set the magnitude of random perturbations to zero for DYAMOND2. This might degrade aerosol behavior in polar regions.

Like E3SMv1, cloud condensation nuclei (CCN) concentration is derived from Abdul-Razzak and Ghan (2000). Ice nucleation follows Gettelman et al. (2010) for deposition nucleation and homogeneous freezing of solution droplets but retains the original P3 implementation for cloud and rain drop freezing.

2.8 Energy Fixer

SCREAM inherited its energy fixer from CAM. As described in Lauritzen and Williamson (2019), this energy fixer corrects errors due to pressure work, time integration in the dynamical core, inconsistent formulations of equation of state, and other minor sources of non-conservation. Historically, CAM and the atmospheric component of E3SM had used an incorrect formulation for energy. Williamson et al. (2015) documents this problem and provides a correction, which is used in SCREAM.

2.9 Timesteps

Like most atmosphere models, SCREAM's timestepping is a complex mixture of substepping and superstepping of individual processes. Ideally, model timesteps would be small enough that modest changes wouldn't have a noticeable effect on model behavior. Unfortunately, climate models have not yet reached that goal (Santos et al., 2020). Thus we list the timesteps used for the DYAMOND2 simulation in Table 1.

Main	Dycore	Dycore Remap	Advection	Radiation
75	9.375	18.75	75	300

Table 1. Timesteps used in SCREAM DYAMOND2 simulation (in sec). Processes not listed use Main timestep.

2.10 Tuning

Tuning is important for optimal performance of any weather or climate model, but should become less important at higher resolution where more processes are explicitly resolved and therefore expressed in a more complete and physical way. Because of time constraints and a reticence to tune away problems before understanding their source, the only parameter adjustment we made was to modify the lower limit of the eddy diffusivity damping timescale to get net top-of-atmosphere (TOA) radiation to match observations and to control surface temperatures under stable conditions at high latitudes. Because our tuning was based on short (1 or 2 day) simulations and therefore required comparison against higher-time frequency radiative observations which (as described in Sect. 4) have larger global-average bias than the monthly-average data used to assess the simulation-average radiation, the TOA net bias reported here still ended up being somewhat large. Our crude tuning approach also resulted in clouds which are too stratiform rather than convective (as described in Section 6.3). High latitude land surface temperature biases remain high, indicating that more tuning work is needed.

3 Experimental Design

The focus of this study is a 40 day global simulation (Jan 20 through Feb 28) performed as part of the DYAMOND2 intercomparison. Our implementation follows the guidance at `https://www.esiwace.eu/services/dyamond/winter` as closely as practicable. Atmospheric initial conditions come from the European Center for Medium Range Weather Forecasting (ECMWF) Integrated Forecasting System (IFS) at its native 9 km grid spacing. Whereas some DYAMOND2 entrants are running with interactive ocean models, SCREAM is not yet able to do this. Instead we use SST at 6-hourly resolution as prescribed from IFS output smoothed by a 7 day running mean.

As mentioned in Sect. 2.7, aerosol distributions are prescribed from a 1° E3SMv1 simulation. This simulation was 6 years long with annually-repeating forcings (SST, sea ice extent, land use, solar forcing, aerosol emissions, greenhouse gases, and volcanic aerosols) values typical for the decade surrounding 2010. The last 5 years of this simulation are averaged to create a monthly varying aerosol field.

Soil and snowpack initial conditions were computed in 2 steps. First, the E3SM land model was run from Jan 1, 1979 through Aug 1, 2016 at the target resolution forced by observed atmospheric conditions from Version 7 of the Climatic Research Unit - National Centers for Environmental Prediction (CRUNCEPv7, Viovy (2018)) atmospheric forcing data. This simulation couldn't be extended beyond 2016 because of CRUNCEPv7 data availability. The second step was therefore to run from Aug 1, 2016 to Jan 20, 2020 using EAMv1 at 1° nudged to ERA5 reanalysis with a 6 hr timescale. Prescribed weekly SST and sea ice from OISSTv2 (Reynolds et al. (2002)) is used for this simulation. The machinery for this second step came from the Cloud-Associated Parameterizations Testbed (Phillips et al. (2004); Ma et al. (2015)).

Nodes	8x16 Dycore timing in minutes	8x16 with IO SDPD	8x16 without IO SDPD	16x8 without IO SDPD
1536	100.8	5.1	5.8	OOM
3072	53.9	8.6	10.3	OOM
4096	44.4	not run	not run	14.2
6144	29.2	14.2	19.2	23.1

Table 2. SCREAM timings as a function of KNL node count using either 8x16 MPI tasks vs OpenMP threads or 16x8 MPI tasks vs OpenMP threads per node. All timing runs were 1 day in length. Timings with IO include all standard output for our DYAMOND simulation. OOM means Out of Memory and IO stands for Input/Output.

4 Observations for Evaluation

The short duration of this simulation and our focus on small time and spatial scales limit the range of observational datasets suitable for comparison. We rely heavily on the European Centre for Medium-Range Weather Forecasting's ERA5 reanalysis (Hersbach et al., 2020). This retrospective simulation assimilates a massive array of observations, runs at 31 km horizontal resolution with 137 vertical levels and a top at 0.01 hPa, and is available at hourly resolution. Because model formulation strongly affects cloud and precipitation predictions from reanalysis, we use satellite products for cloud-related variables. In particular, we use half-hourly 0.1° gridded Global Precipitation Measurement (Hou et al., 2014, GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG) product version V06B (G. J. Huffman & coauthors, 2019) for global precipitation. For radiative fluxes, we use CERES-EBAF 1° data averaged over February 2020 (Loeb et al., 2018). To examine the radiative properties of individual storms, we also use CERES-SYN hourly 1° data (Doelling et al., 2013, 2016). Cloud fraction and liquid water content are taken from CloudSat (Austin et al., 2009; Su et al., 2011) and from the CERES–CALIPSO–CloudSat–MO merged product (Kato et al., 2010, C3M). CloudSat and C3M are not available for the 2020 dates simulated and are instead climatological averages.

Where possible, we compute long-term averages using the last 30 days of the simulation (Jan 30th through Feb 28th); we exclude the first 10 days of the run as spinup (though SCREAM fields stabilize after just one day of spinup, see Fig. 5). As noted above, some observational datasets are only available as monthly averages. For corresponding variables, we show results using just days in Feb. Finally, the first week or so of the simulation can be treated as a weather forecast, we use all 40 days of the simulation for some analysis of storm behavior.

5 Performance

The DYAMOND2 simulation was performed as a series of 1536-node job submissions using the Knights Landing (KNL) nodes of Cori at the National Energy Research Supercomputing Center (NERSC). We found that using 8 MPI processes and 16 OpenMP threads per node provided the optimal balance of memory usage and performance for these 1536-node jobs. The overall throughput for the 40-day simulation, including I/O, was about 4-5 simulated days per day (SDPD). Further details about the performance of this 40-day DYAMOND2 simulation can be explored at `https://pace.ornl.gov/search/SCREAMv0.SCREAM-DY2.ne1024pg2.20201127`. As shown in Table 2, the model scales quite well - particularly in the dycore - and can achieve up to 23.1 SDPD without input or output (IO) on 6144 KNL nodes.

The simulation used the Software for Caching Output and Reads for Parallel I/O (SCORPIO) library for reading input data and writing simulation output to the file system. SCORPIO is derived from the Parallel I/O library (Hartnett & Edwards, 2021) and continues to support the same application programming interface. To improve the I/O write performance the library caches and rearranges output data between MPI processes before using low level I/O libraries like the netCDF, Parallel netCDF (PnetCDF) (Latham et al., 2003), and ADIOS (Godoy et al., 2020) libraries to write the data to the file system. On Cori the simulation produced ∼4.5 TB of data per simulated day and achieved an average I/O write throughput of ∼2.5 GB/s using the PnetCDF library.

Unsurprisingly for such a large run, we experienced several node failures during the simulation requiring restarts from the previous day. Because E3SM is bit-for-bit reproducible for identical initial conditions and forcings, these failures should not have any impact on our results. During model development, we had problems with occasional extremely cold temperatures near the surface at wintertime high latitudes. We fixed this problem by increasing turbulent diffusivity in stable atmospheric conditions, but this had the side effect of increasing time-average warm bias in polar regions. The tuning used here balances model stability against bias.

6 Results

6.1 Kinetic Energy Spectrum

At convection permitting resolutions, the simulated atmosphere's kinetic energy spectra recovers many features seen in observations and reveals many aspects of model diffusion, filtering and parameterization behavior (Skamarock et al., 2014). As a first look at this in SCREAM, we plot the horizontal kinetic energy power spectra at 250 hPa and 500 hPa in Fig. 2. The spectra are computed via spherical harmonic transforms of 3-hour flow snapshots from days 22 and 23 of the simulation. We denote by $E(k)$ the power of the spherical harmonics of degree k. We plot compensated spectra, $E(k)k^{5/3}$, to better illustrate the high wave number $k^{-5/3}$ regime. SCREAM reproduces the observed Nastrom-Gage transition from a k^{-3} scaling at low wavenumbers to a $k^{-5/3}$ regime (Nastrom & Gage, 1985; Lindborg, 1999). The $k^{-5/3}$ region extends to $\sim 6\Delta x$ wavelength (wavenumber 2000), where the spectra start to roll off and become dominated by model diffusion. Thus SCREAM's effective resolution is similar to ICON and IFS (Neumann et al., 2019) despite SCREAM's novel use of a coarser grid for physical parameterizations.

6.2 General Features

Global-average model biases are modest in size but are generally larger than the range of observed day-to-day variability within the simulation period (Fig. 3). TOA net shortwave (SW) radiative absorption $\mathrm{SW_{net}}$ and longwave (LW) emission $\mathrm{LW_{net}}$ are both too strong but (as noted in Section 2.10) were tuned to compensate each other such that TOA radiative bias $\mathrm{rad_{net}}$ exhibits only a modest warming tendency. Radiative biases are almost entirely due to clouds rather than clear-sky bias (not shown). Too little $\mathrm{SW_{net}}$ reflection and excessive $\mathrm{LW_{net}}$ emission suggests a lack of clouds, so it is surprising that model calculated vertically-projected cloud fraction is 5% too large. This is an unfortunate result of using a RH-based ice cloud fraction parameterization without retuning for higher resolution. As a result, large cloud fraction occurs in cold regions which don't necessarily have cloud mass (Fig. 4). Fortunately, 'clouds' without condensate are treated like clear-sky air by radiation, so our mistake is mostly cosmetic in nature. In the future we intend to switch to a mass-based all-or-nothing ice cloud fraction scheme to avoid this problem. An offline version of this mass-based approach (shown in Fig. 4) is used in the remainder of this paper wherever upper-level cloud fraction is required.

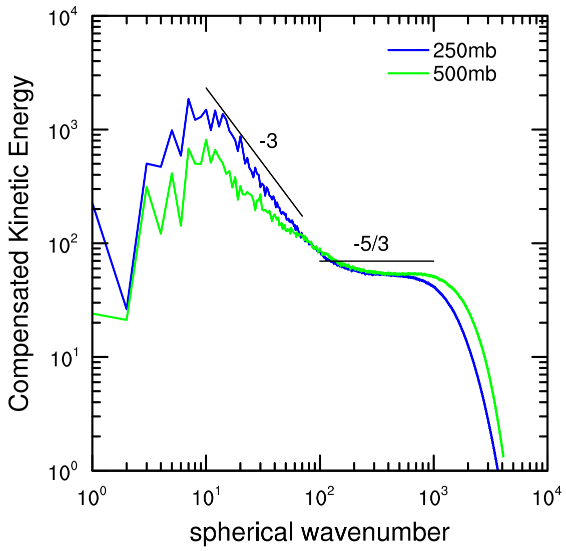

Figure 2. Compensated kinetic energy spectra ($E(k)k^{5/3}$) at 500 hPa and 250 hPa from days 23-24 of the simulation. The black lines show idealized $E(k) \approx k^{-3}$ and $E(k) \approx k^{-5/3}$ scalings. See text for details.

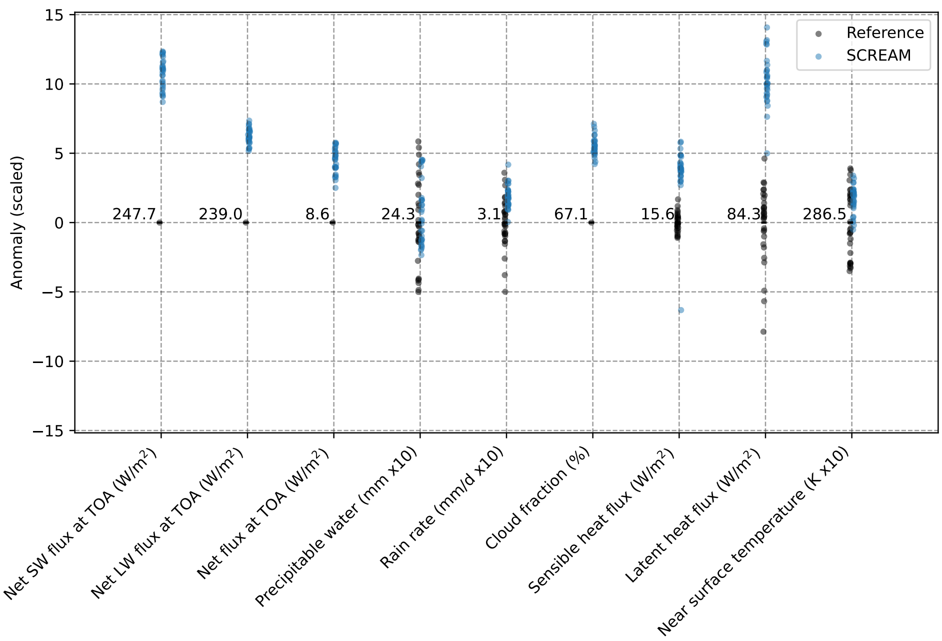

Figure 3. Global-mean anomaly in variables listed along x-axis. Anomalies are calculated relative to the February 2020 average of CERES-EBAF (for radiative fluxes and cloud fraction), and the January 30 2020 through Feb 28 2020 average of ERA5 (for precipitable water, sensible and latent heat fluxes, and near surface temperature), and GPM (for precipitation). Each dot represents a single daily average, so vertical spread gives a sense of temporal variability. There is a dot for each day in Feb 2020. Units for each variable are included in the x-axis labels.

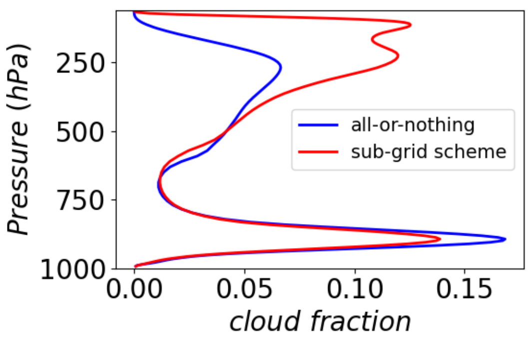

Figure 4. Vertical profile of Feb-mean tropics-averaged (30°S-30°N) cloud fraction computed by SCREAM compared to an offline calculation of cloud fraction based on assuming an entire cell is saturated whenever cloud water content $> 10^{-5}$ kg kg^{-1}.

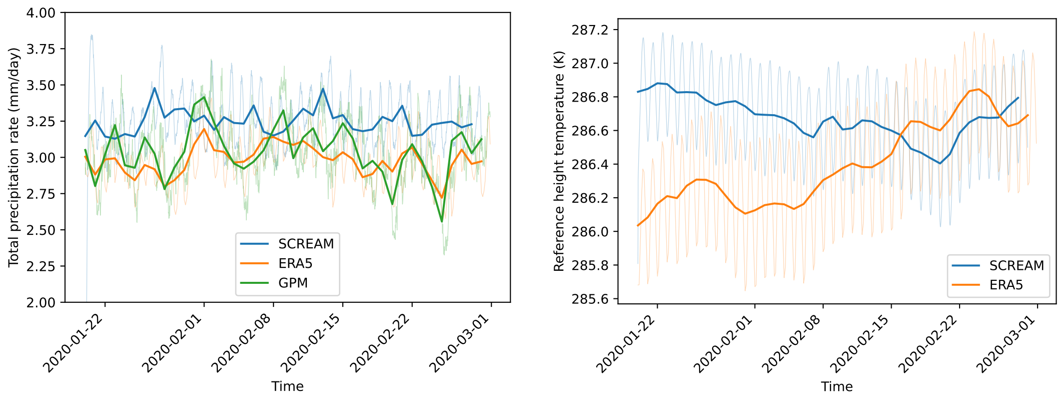

Figure 5. 15 minute (thin curves) and daily-mean (thick curves) time series of global-average precipitation (left) and 2 m temperature (right) for the duration of the DYAMOND2 simulation.

Global-average precipitation is ~0.3 mm day^{-1} larger in SCREAM than GPM, which is consistent with a general tendency for models to have higher precipitation rates than observations (Terai et al., 2018), including in the previous DYAMOND intercomparison (Stevens et al., 2019). Temperature at 2 m height (T2m) and vertically-integrated vapor lie within observed day-to-day variability in the global average, though we show later that this is due in part to compensating errors. Sensible heat flux (SHF) and surface evaporation (a.k.a. latent heat flux; LHF) are larger than observed, probably due to near-surface wind speed biases discussed later.

Fig. 5 demonstrates that our simulation doesn't drift rapidly in time, even in the first few days of the run. Time tendencies in other key variables are likewise small (not shown). The amplitude of global-average diurnal variations is also reasonable. Interestingly, GPM and ERA5 contain periods where global-average precipitation drops, while SCREAM is more temporally invariant. Understanding what causes these global-average drops is an interesting question for future work.

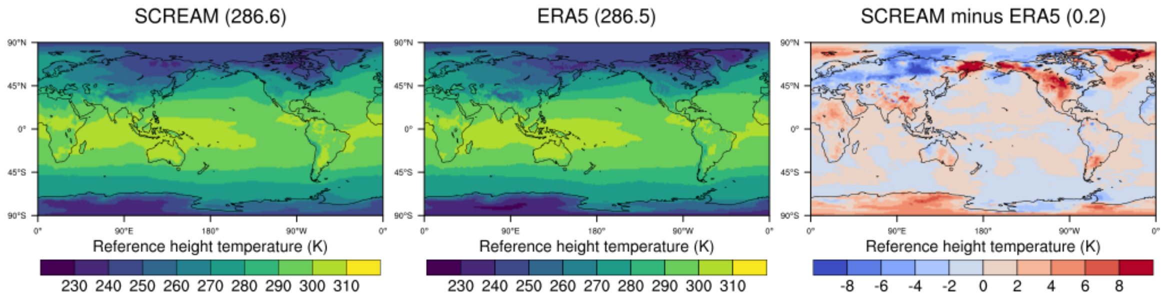

Figure 6. Near-surface temperature averaged over Jan 30 through Feb 28, 2020 from SCREAM and ERA5 reanalysis.

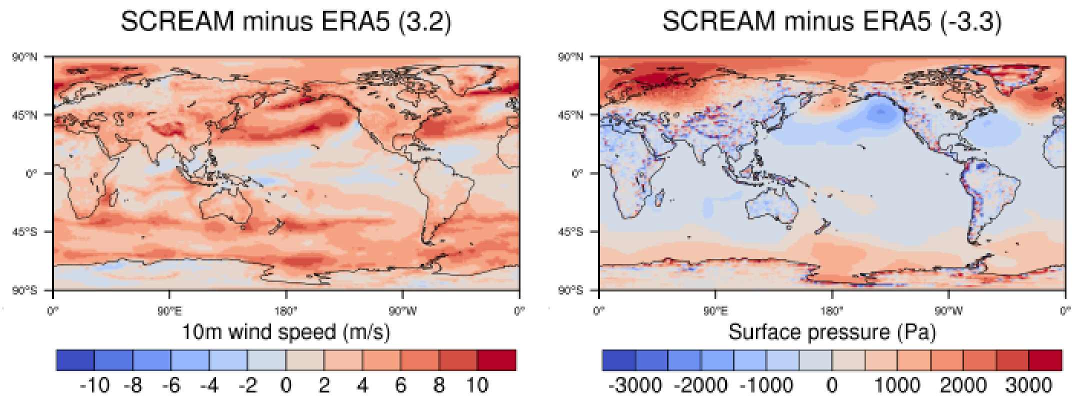

Figure 7. Bias (relative to ERA5) in 10 m wind speed (left) and surface pressure (right) averaged over the last 30 days of the simulation.

Near-surface temperature biases are modest at low latitudes and larger at high latitudes (Fig. 6). In the first few days of our simulation, T2m was uniformly too high at high latitudes (not shown), which we attribute to a land initial condition created by driving our land model with a 1° atmosphere model which one might expect to handle snowpack poorly. We tuned overturning turbulent mixing in stable conditions to compensate the warm biases we saw in our initial short testing runs; it appears in retrospect that we overdid it. Averaged over the last 30 days of the simulation, the US, Greenland, and the far eastern side of Russia retain >6 K warm biases, while north Asia and the Canadian Arctic are ∼5 K too cold. Improving these temperature biases is a future goal. Surface pressure is also too large at high latitudes (right-hand panel of Fig. 7), which will translate (through thermal wind balance) to errors in wind speed.

Near-surface wind speed is too high almost everywhere but particularly over mid-latitude oceans (Fig. 7). Bias is smallest in the tropics. We are still working to understand and fix this deficiency, but note that switching to the Zeng et al. (2002) scheme significantly alleviates excessive wind speeds. Consistently positive wind biass is solely a feature of the surface layer - even at 925 hPa wind biases are much more balanced around zero. Overall, it is surprising that so many aspects of our simulation look quite good in spite of this near-surface wind bias. Overly strong SHF and LHF mentioned earlier are unsurprising given strong near-surface wind speed.

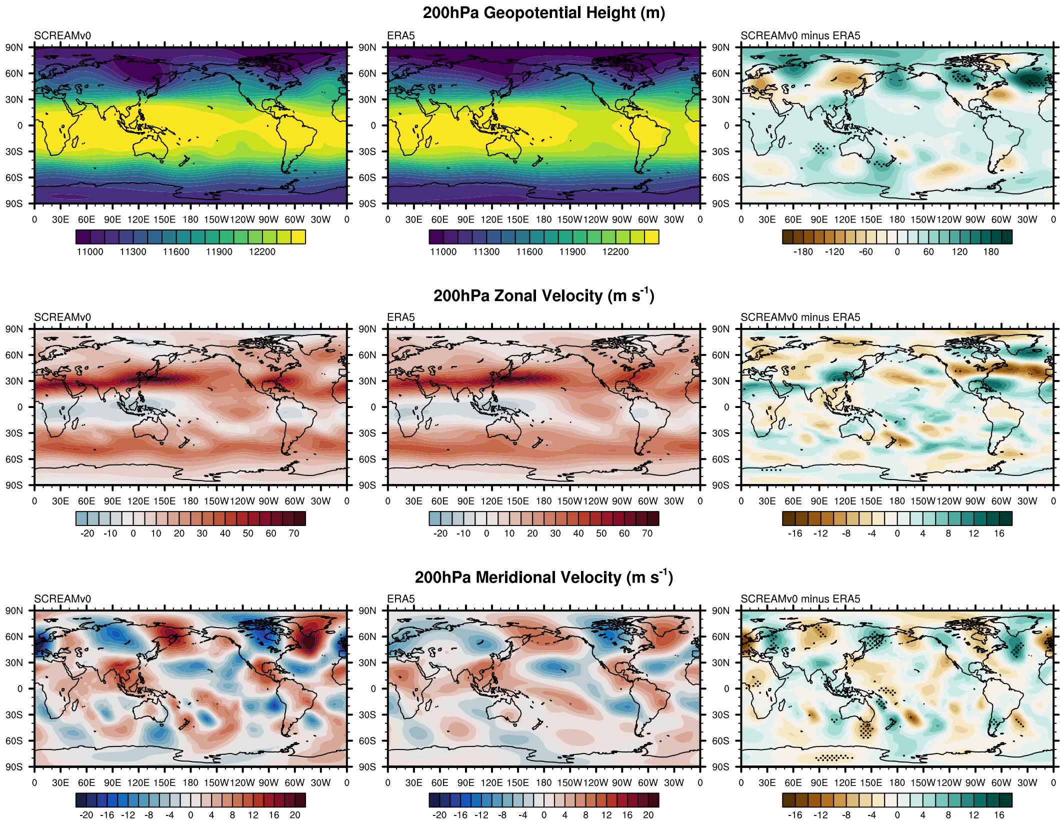

Figure 8. 200hPa geopotential height (top), zonal wind speeds (middle), and meridional wind speeds (bottom) averaged over the last 30 days of the simulation. Stippling in the difference plots (right panels) indicates regions where SCREAM falls outside the range of mean values for all years in ERA5 1979-2020.

Fig. 8 shows geopotential height and wind speeds on the 200hPa pressure surface averaged over the period from January 30th to February 28th. Although there is generally strong agreement between SCREAM and ERA5, two hotspots emerge. First, the wintertime Rossby wave train that reinforces the upper-level trough over Greenland is markedly more intense in SCREAM than in ERA5. The result is southward displacement of the subtropical jet (STJ) over the West Atlantic and anomalously strong poleward flow from the STJ towards Greenland. In fact, this anomaly in the Central Atlantic is largely barotropic, present even at 850hPa with approximately the same magnitude (not shown). A second region of anomalous behavior also exists around the periphery of Australia where the 200hPa geopotential surface is enhanced, producing spurious meridional flow throughout this region. Notably, the bias pattern present in the difference plots suggest an enhancement in wavenumber 4 in both hemispheres centered around the locations of cubed-sphere corners in the dynamics grid. The bias appears slightly stronger in the first 20 days of the simulation than the last 20 days (not shown). The source of this behavior is under investigation.

6.3 Radiation and Clouds

SW_{net} and LW_{net} radiation biases were found in Fig. 3 to somewhat cancel in the global mean; Fig 9 reveals that this cancellation also holds regionally in many places. Cancellation between SW and LW biases is a hallmark of high clouds. Further evidence

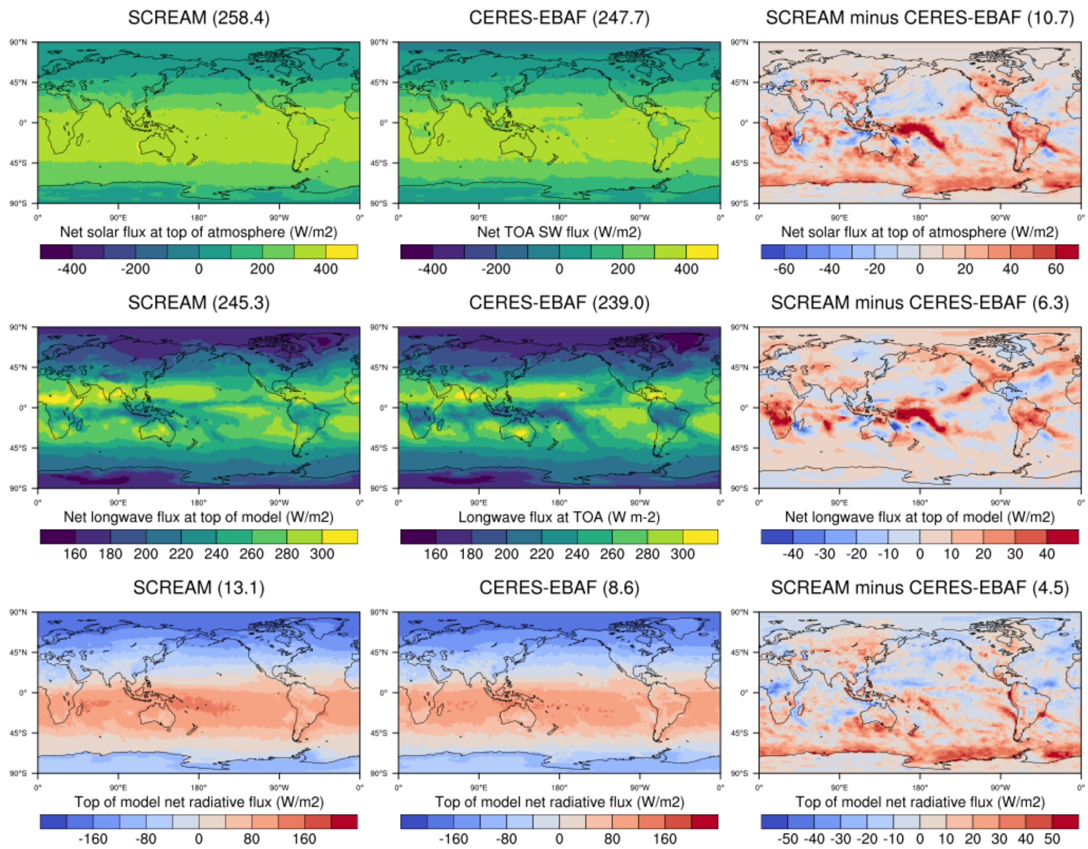

Figure 9. TOA radiation averaged over February 2020. Top is SW (>0 warms the planet), middle is LW (>0 cools the planet), and bottom is net (>0 warms the planet).

of problems with high clouds is the pattern of LW_{net} bias, which is large where deep convective clouds are expected.

Fig. 10 explores the vertical profile of tropical clouds compared to climatological CloudSat measurements. Because SCREAM results are for one month only, detailed comparison is not appropriate. Nonetheless, SCREAM's ability to capture the general features from CloudSat data is very good, particularly compared to the (albeit old) GCMs analyzed in Su et al. (2011). In particular, SCREAM captures the bimodality of deep and shallow clouds and does a reasonable job of matching the quantitative magnitude of each peak. Ability to better capture the structure of tropical convection is perhaps unsurprising given that resolving such convection was a primary motivation for developing a 3.25 km model. Both simulated cloud peaks sit lower in the atmosphere than they do in the measurements. Another notable deficiency in SCREAM is the lack of mid-level clouds, which may be tied to either the absence of significant cloud detrainment at mid-levels, overly efficient sedimentation of cloud particles through mid-layers, or both. Reasonable or even excessive SCREAM anvil condensate in Fig. 10 and erroneously large high cloud fraction in Fig. 4 are at odds with excessive LW emission to space in Fig. 9. We are still working to understand this conundrum.

Net outgoing radiation over the northern hemisphere oceans is found in Fig. 9 to be too strong in general (i.e. the oceans in Fig. 9i are colored blue indicating more ra-

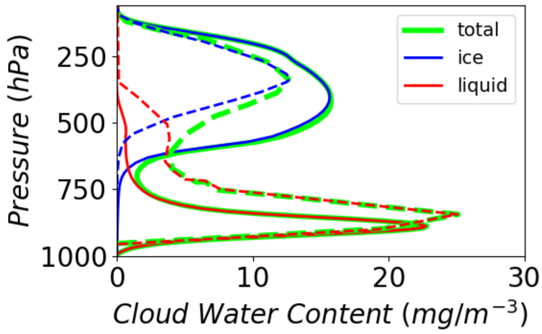

Figure 10. Cloud water content (CWC) profiles from SCREAM (solid) versus CloudSat observations from Su et al. (2011) (dashed). Data are averaged over all longitudes and latitudes between 30°S-30°N and over all 40 simulated days.

diation leaving than entering the atmosphere). This is due mainly to trapping of LW emission; SW_{net} insolation at higher northern latitudes is too small in wintertime to matter.

Away from high-latitude winter regions, the impacts of high clouds on SW_{net} and LW_{net} tend to cancel so rad_{net} is a good indicator of lower-level cloudiness. Fig. 9 reveals a lack of low clouds over the southern ocean, but generally decent low-cloud radiative forcing in the stratocumulus decks off the west coast of the continents. Anemic stratocumulus is a perennial GCM bias (Nam et al., 2012; Jian et al., 2020), so capturing this cloud type in SCREAM is exciting. This is particularly surprising since 3.25 km grid spacing is generally considered insufficient to capture boundary-layer clouds like this. One potential reason for improvement is our higher-order turbulence closure. Increased vertical resolution (~50 m in the boundary layer) in addition to SCREAM's high horizontal resolution also likely helps; Bogenschutz et al. (2021) and Lee et al. (2021) demonstrate that increased vertical resolution helps to ameliorate these biases in E3SM, owing to better representation of the cloud top cooling and turbulence feedback, but both studies hypothesize that concurrent increases in the horizontal and vertical resolution are needed to adequately simulate the coastal Sc. Results with SCREAM support that hypothesis.

Figures 11a-b display the February 2020 average profiles of cloud fraction and cloud liquid water for SCREAM and the February 2006-2010 climatology from C3M. These profiles are averaged over a small domain neighboring the coast of Peru and Chile. This domain was selected as it represents the area of most intense shortwave cloud radiative effect (SWCRE) biases associated with low clouds in the northern-hemisphere winter season for standard-resolution GCMs (e.g. Golaz et al. (2019); Danabasoglu et al. (2020)). Although different averaging periods are used for C3M versus SCREAM data, stratocumulus are a persistent feature in this region so broad comparison is reasonable. SCREAM produces cloud structure quite similar to the observations. Though SCREAM cloud fraction in Fig. 11a may appear to be underrepresented, we note that its deficiencies are small compared to most GCMs (Bogenschutz et al., 2021). In addition, cloud liquid water in Fig. 11b matches observations almost perfectly. Fig. 11c depicts a snapshot of the SWCRE on 01 March, 2020 at 18:00:00 UTC from SCREAM to demonstrate the model's ability to simulate healthy coastal Sc cloud decks and the gradual transition to more broken cloud.

Fig. 12 displays the temporally-averaged curtain of cloud fraction along the 20°S transect across the stratocumulus-to-deep-convection transition for SCREAM February

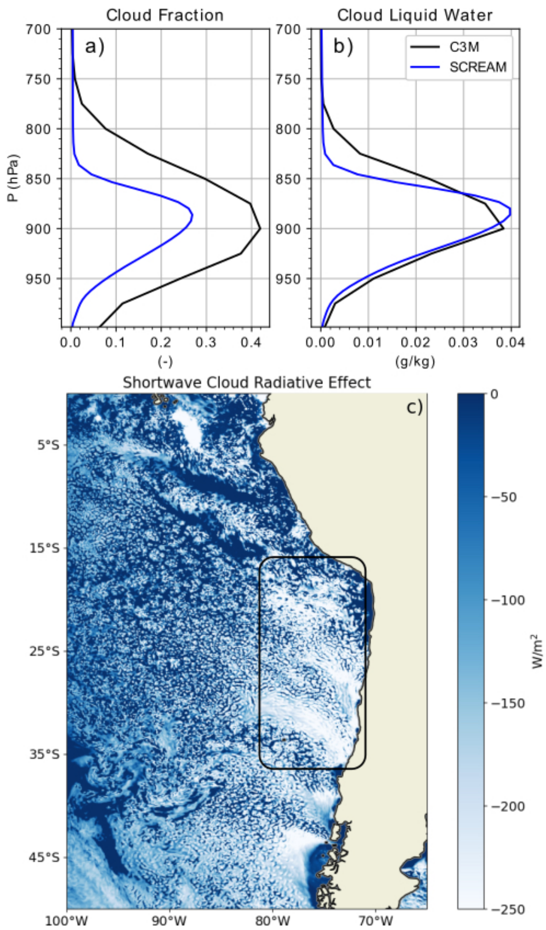

Figure 11. Temporally and spatially averaged profiles of cloud fraction (a) and cloud liquid water (b) for SCREAM and C3M. SCREAM profiles are averaged for the month of February 2020 while the C3M represents the February climatology from 2006-2010. Both SCREAM and C3M profiles represent spatial averages from the southeast Pacific coastal stratocumulus region bounded from 35°S to 15°S and 275°E to 290°E. The area used for spatial averaging is denoted in (c), which represents a snapshot of shortwave cloud radiative effect from SCREAM for 01 March 2020 at 18:00:00 UTC.

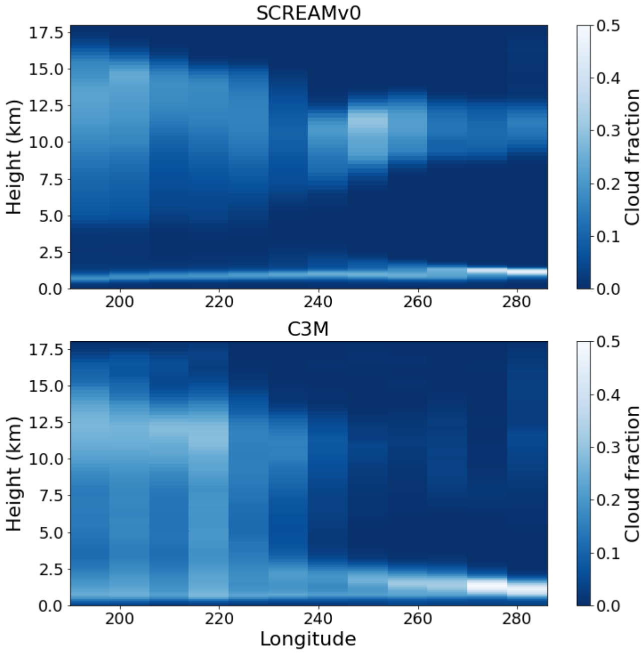

Figure 12. Temporally-averaged curtain of cloud fraction along the 20°S transect across the stratocumulus to deep convection transition. SCREAM clouds are averaged over the month of February 2020 while the C3M represents the February climatology from 2006-2010. Both SCREAM and C3M profiles represent curtains bounded from 24°S to 16°S.

2020 average and C3M February climatology from 2006-2010. When read from right to left (i.e. along the direction of prevailing easterly winds), C3M observations depict a gradual deepening of cloud in the lower troposphere over progressively warmer SSTs. SCREAM looks reasonable near the coast but fails to deepen to the W and is generally too thin in depth and too weak. This was an unintended consequence of tuning choices made in the SHOC parameterization to achieve reasonable radiation balance; further tuning since this simulation has improved the realism of trade cumulus.

6.4 Precipitation

Evaluating the spatial distribution of precipitation from a 40 day simulation is challenging. Forty days is too long for comparison against weather events but too short to average out the effects of individual storms. Zonal-averaging beats down some of this weather noise and large-scale tropical precipitation structure is probably robust, but results should still be taken with caution. In Fig. 13, zonal-average precipitation is found to generally agree well with both GPM and ERA5 except for excessive rainfall on the equatorward side of the northern-hemisphere storm track and at the poleward edges of the tropics. GPM is known to be biased low at higher latitudes due to problems detecting light rain and snow (G. Huffman et al., 2019), which might partially explain storm track and polar biases. Fig. 14 shows that tropical zonal-mean bias is due to a complicated mixture of differences in the meridional structure of precipitation. SCREAM tends to have stronger precipitation on the east side of land masses, in particular over the Maritime Continent (which has been a long-standing bias in E3SM; Golaz et al., 2019) and

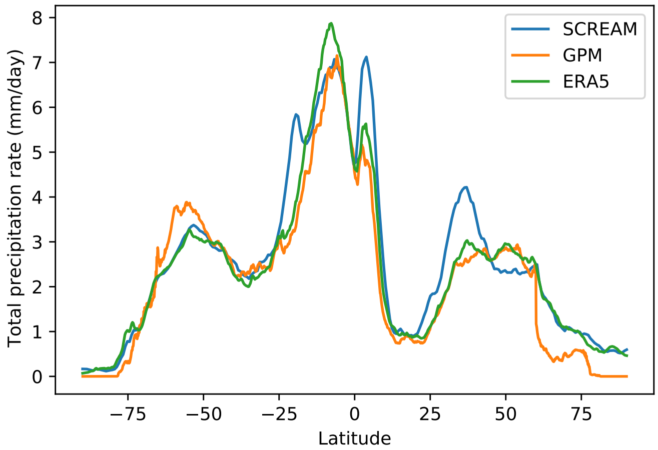

Figure 13. Zonal-average precipitation averaged over the last 30 days of the SCREAM simulation.

west of Madagascar. Heavy precipitation in the ITCZ extends too far east, which is another persistent E3SM bias. Precipitation in the South Pacific Convergence Zone (SPCZ) is, on the other hand, too weak and a bit too zonal. This may indicate that SCREAM (like most climate models) suffers from double-ITCZ problems (Li & Xie, 2014). Precipitation over the Amazon rain forest is slightly too strong, which is the opposite from what is seen in conventional climate models (Yin et al., 2012).

A great success of cloud-resolving models are their ability to simulate the diurnal cycle of precipitation (Khairoutdinov et al., 2005; Sato et al., 2009; Stevens et al., 2019). This is a feature which coarser resolution models struggle with (Covey et al., 2016), though progress has been reported (Xie et al., 2019). As documented in Fig. 15, SCREAM is able to capture the morning-time peak over the oceans and late afternoon peak over land. The diurnal cycle over the Maritime Continent and Madagascar - two areas dominated by sea breezes - is actually stronger than observed in GPM (but is weaker in magnitude than TRMM's observed climatology; not shown). Stronger diurnal amplitude in these areas is perhaps unsurprising given that daily mean precipitation was also noted to be too high in these regions.

Like conventional GCMs (Stephens et al., 2010; Na et al., 2020), SCREAM has a tendency towards having too much drizzle and not enough strong precipitation (Fig. 16). The magnitude of this bias is, however, much smaller than typically found in conventional GCMs (e.g. Caldwell et al. (2019) Fig. 12). Thus we consider simulation of heavy precipitation to be a victory for SCREAM.

Hovmöller diagrams showing precipitation averaged from 5° N to 5° S latitude as a function of longitude and time are useful for evaluating the temporal intermittency and propagation of tropical convection which collectively result in the Madden-Julian Oscillation (MJO; (Madden & Julian, 1971)). Usually MJO analyses filter out signals outside of a 20-90 day window, but our 40 day simulation precludes such processing. A longer simulation is needed for statistical robustness, but it seems clear in Fig. 17 that SCREAM triggers convection too frequently. This feature is also apparent in instantaneous snapshots of precipitation, water vapor, and cloud mass (not shown). We are still investigating the source of this "popcorn convection", which also appears in other convection-permitting regional and global models (Arnold et al., 2020; Kendon et al., 2012). As found for other GCPMs (Miura et al., 2007; Miyakawa et al., 2014), SCREAM does a good job of propagating convective events eastward, though its propragation speed is perhaps slightly fast.

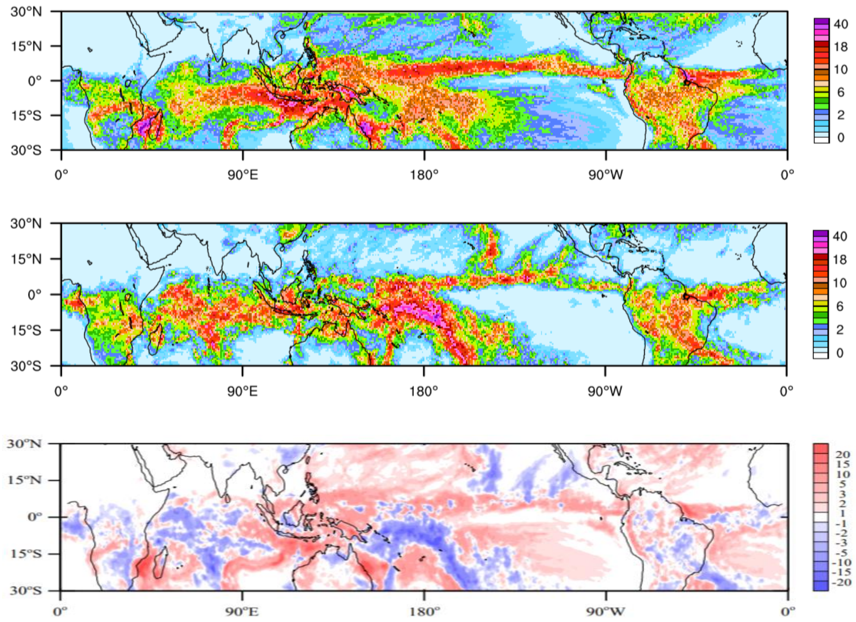

Figure 14. Tropical precipitation over the last 30 days of the SCREAM run (top), GPM observations averaged over the same period (middle), and their difference (SCREAM minus GPM, bottom).

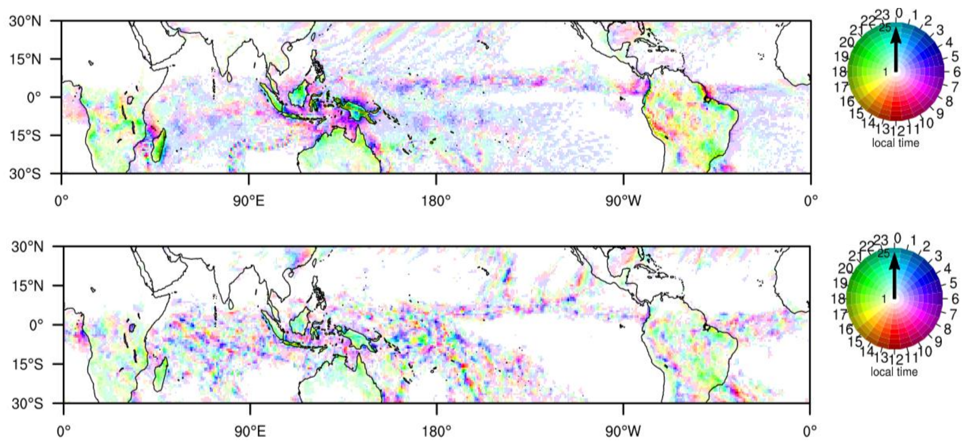

Figure 15. Diurnal cycle of precipitation averaged over the last 30 days of the SCREAM run (top) and GPM observations (bottom). Hue indicates time of peak precipitation and intensity indicates diurnal amplitude. Amplitudes less than 1 mm day^{-1} are colored white and colors saturate at 25 mm day^{-1}.

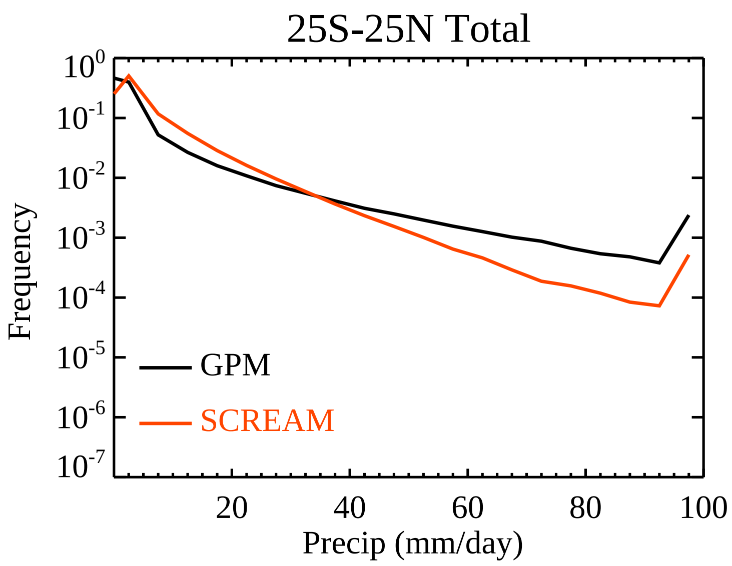

Figure 16. Histogram of tropical precipitation over the ocean for the last 30 days of the SCREAM simulation compared to equivalent days from GPM.

The statistical analysis of precipitation above is important, but it ignores the fact that precipitation comes from storms whose characteristics vary regionally. The next few subsections explore SCREAM's treatment of important storm types.

6.5 Tropical Cyclones

Tropical cyclones (TCs) are some of the most intense storms in the world, combining intense precipitation with winds frequently in excess of 30 m s^{-1}. Although some global models are able to represent TC frequency and intensity well at 0.25° grid spacing, resolving the inner structure of these storms requires much finer resolution (Wehner et al., 2014; Zarzycki & Jablonowski, 2015; Judt et al., 2021). A key advantage of running global convection-permitting models is the ability to represent and study multiscale interactions between the inner structure of tropical cyclones and the large-scale environment (Satoh et al., 2019). In the first phase of the DYAMOND project, models produced a wide range of tropical cyclone counts and intensities with counts as low as 4 to as high as 20, while in reality there were 14 (Stevens et al., 2019; Judt et al., 2021). In this section, we provide a brief and broad overview of the tropical cyclones identified in the SCREAM simulation.

Fig. 18 shows TC tracks during the simulation period from SCREAM, ERA5, and IBTrACS observations (Knapp et al., 2010, 2018). SCREAM and ERA5 tracks are computed using the TempestExtremes (TE) algorithm and the criteria described in appendix A1, while IBTrACS are based on expert judgement. Large discrepancies between IBTrACS and reanalysis datasets highlight the importance of using consistent criteria to classify storms. Note as well that the chaotic nature of weather means that storms later in the SCREAM simulation are not expected to match those found in ERA5 or in observations. Within the days of potential predictability (up to two weeks), one TC exists in both the SCREAM simulation and ERA5 data (Moderate Tropical Storm Diane). Another storm that is present in ERA5 (Moderate Tropical Storm Esami) does not organize in SCREAM, although a weak low pressure region does persist. Over the entire simulation period, we identify five tropical cyclone tracks in SCREAM during the 40 day simulation and six

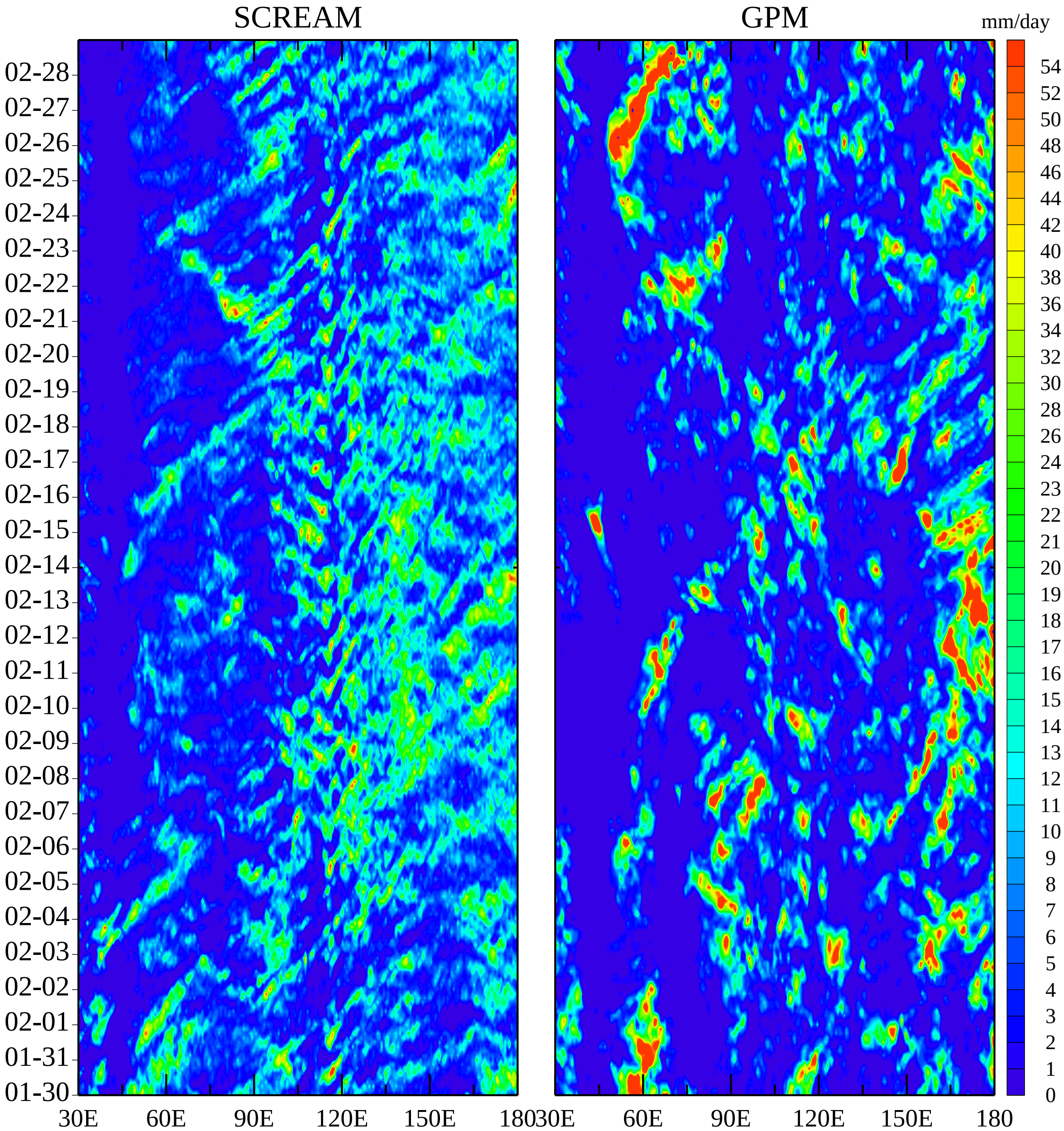

Figure 17. Precipitation averaged from 5° N to 5° S as a function of longitude (x-axis) and time (y axis) from SCREAM (left) and GPM precipitation observations (right).

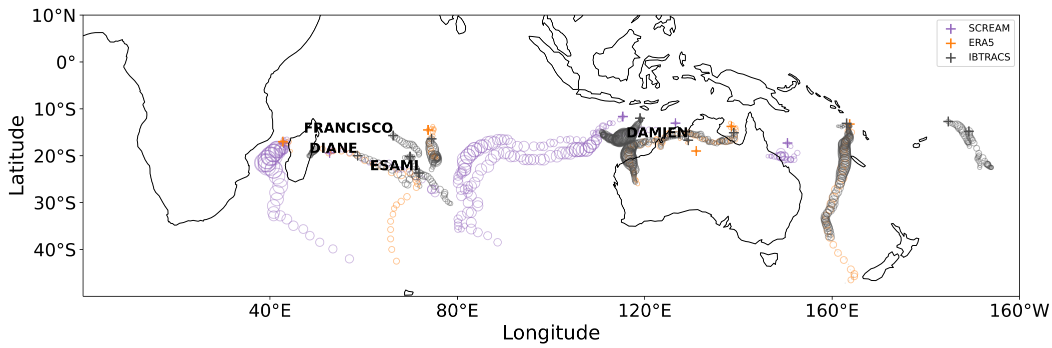

Figure 18. Tracks of tropical storms from IBTrACS (grey) and identified by the Tempest Extremes algorithm in SCREAM (purple) and in ERA5 (orange) between Jan 20 and Feb 28, 2020. Starting location is indicated with a plus (+).

tracks in the ERA5 reanalysis data. All five TCs in SCREAM occur in the Southern Hemisphere, with four over the Indian Ocean and one off the northwestern coast of Australia over the Pacific Ocean, all broadly located where TCs are found in the reanalysis.

In ERA5, Diane starts off as a tropical depression with central pressure of 1020 hPa, but its pressure drops down to 990 hPa by Jan 26 with sustained maximum winds of 25 m/s (49 knots) or more. The simulated storm track from SCREAM closely follows that found in the reanalysis (Fig. 19a), although it forms farther to the east and moves eastward more slowly. The maximum wind speed within a 6°×6° box around the storm is also higher in the model, but this is likely due to the use of native grid data in SCREAM and the coarser regridding of the reanalysis data. Precipitation rates in SCREAM and reanalysis closely follow each other until ERA5 starts tapering off while SCREAM continues growing. GPM precipitation, however, includes a period of much stronger precipitation which isn't captured by either model simulation. Interestingly, SCREAM has a strong and regular diurnal cycle of precipitation which isn't found in the other timeseries. Although the data for this storm from IBTrACS spans a much shorter period in the storm lifetime than identified by Tempest Extremes from either the reanalysis or the SCREAM simulation, the magnitudes of central pressure and maximum 10-m wind speed are in good agreement between SCREAM, ERA5, and IBTrACS for the period that does overlap.

Because Severe Tropical Storm Diane does not fully develop a canonical tropical cyclone structure and exhibits hurricane force winds only for a few hours, we take a more detailed look at a stronger storm in the model which forms on Feb 10 and produces surface wind speeds which classify it as a category 3 hurricane (Fig. 19g). For reference, the storm's maximum intensity (based on minimum surface pressure values) is the median of the five storms tracked in SCREAM (not shown). Fig. 19e shows the cyclone track, which spans sixteen days. The surface pressure rapidly drops from Feb 11 to Feb 14, a minimum pressure of 930 hPa on Feb 16, when maximum 10-m wind speeds are also reached. By that point, the storm has formed a distinctive eye, ringed by strong precipitation rates reaching 100 mm/hr and wind speeds greater than 60 m/s (Fig. 20). The high surface wind speeds drive surface latent heat fluxes greater than 500 W m^{-2}, and a vertical north-south curtain centered on the point of minimum surface pressure shows the boundary layer flow is transporting energy towards the eye, particularly in the southern half of the storm (Fig. 20).

More analysis is necessary for an in depth study of the storm characteristics in SCREAM, as was done by Judt et al. (2021) for the models participating in the first phase of DYA-

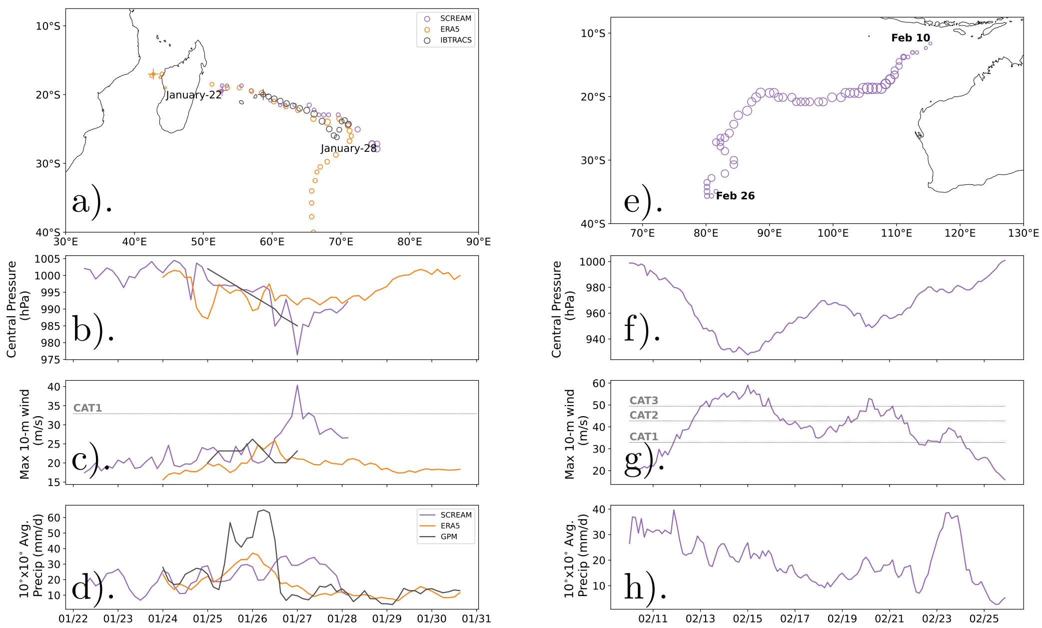

Figure 19. (a) Tracks of the tropical storm Diane from IBTrACS (grey) and as identified by the Tempest Extremes algorithm in SCREAM (purple) and in ERA5 (orange). Starting location is indicated with a + symbol. Shown below the tracks are time evolution of the storm's minimum central pressure (b), maximum 10-m wind speeds within 5° of the storm center (c), and area-averaged precipitation rate (d). (e-h) Same as (a-d) but for Feb 10 tropical cyclone in SCREAM simulation. No observational equivalent is shown, because it is outside the period of predictability.

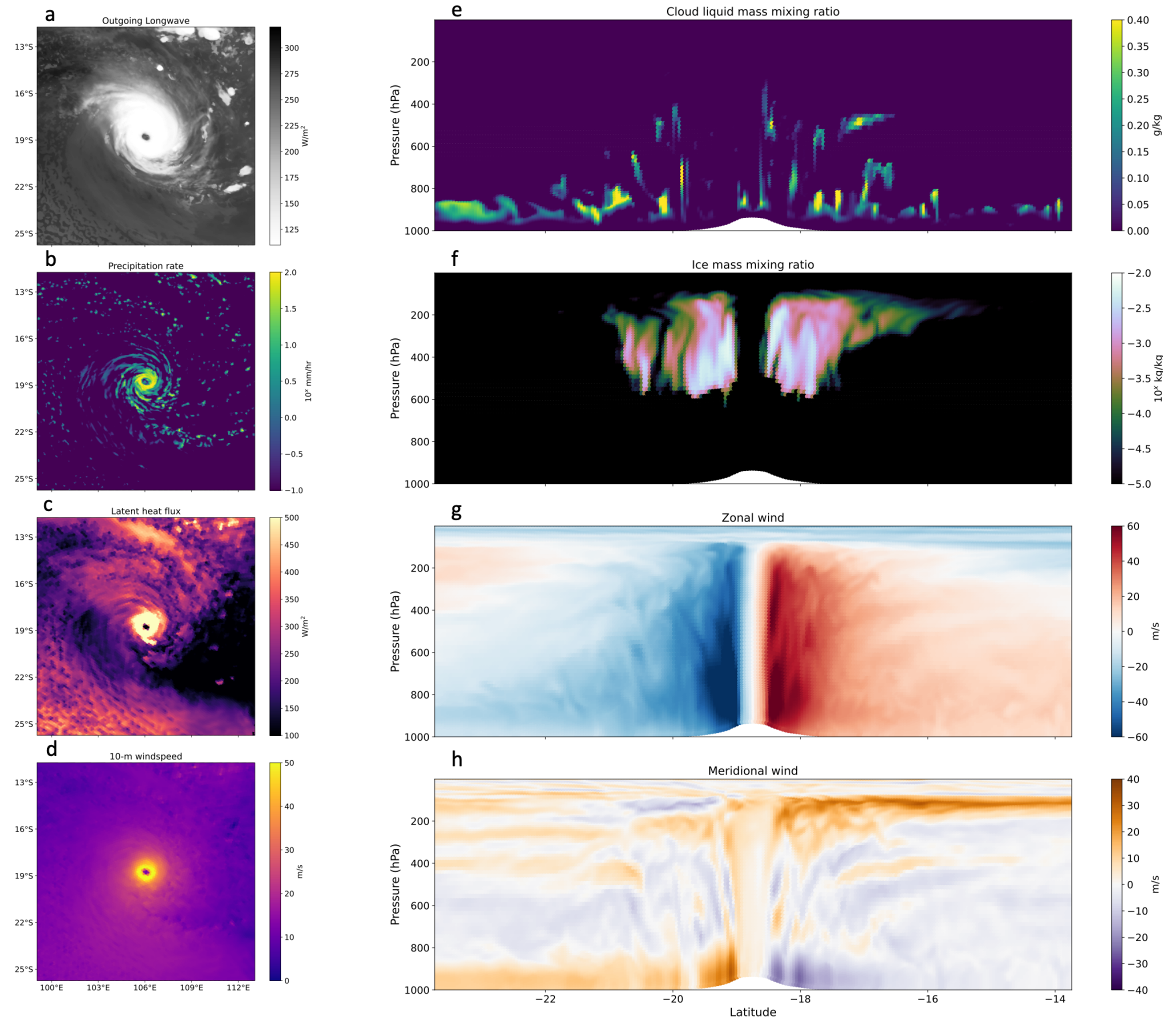

Figure 20. Instantaneous planar and curtain view of Feb 11 tropical cyclone at maximum intensity on Feb 16 0UTC. On the left column are planar views of the outgoing longwave radiation (a), precipitation rate (b), latent heat flux (c), and 10-m wind speed (d). On the right column is a north-south curtain snapshot through the center of the storm of the cloud liquid mass mixing ratio (e), ice mass mixing ratio (f), zonal wind speed (g), and meridional wind speed (h).

MOND. However, as Fig. 19 and 20 indicate, SCREAM produces tropical cyclones with reasonable eye-wall structure and adequate surface wind intensities, which provide promise for future attempts to simulate observed tropical cyclones using the model.

6.6 Extratropical Cyclones

In mid- and high-latitude regions, extratropical cyclones (ECs) are a large source of day-to-day weather variability. ECs are a major pathway for water evaporated from the ocean to precipitate over land; Hawcroft et al. (2012) suggest that as much as 90% of the surface precipitation along midlatitude storm tracks is attributed to ECs. ECs are also behind a majority of extreme precipitation events, particularly in the northeast US where ECs are responsible for more than 80% of winter-time extreme precipitation (Pfahl & Wernli, 2012; Agel et al., 2015). With increasing resolution, ECs are better represented in global models (Jung et al., 2006), and a recent study using a set of global storm-resolving model simulations shows an increase of 7%/K in precipitation rate from the most intense extratropical cyclones with warming, which differs from the 2-3%/K increase expected in the global mean (Kodama et al., 2019).

Over the simulation time period, 87 ECs are identified in SCREAM and 80 are found in ERA5 using the TempestExtremes algorithm (see Appendix A2 for details). Their geographic distributions in the Southern and Northern Hemispheres are shown in Fig. 21a and b. In the Northern Hemisphere, the density of storms in both SCREAM and ERA5 is largest over the Atlantic and Pacific Ocean basins, with many storms originating close to the western boundary currents. This is consistent with observed climatologies of cyclone statistics (Sinclair, 1997). Bomb cyclones (ECs with surface low pressures dropping more than 24 hPa over a 24 hour period (Sanders & Gyakum, 1980)) are present in both SCREAM (11) and in ERA5 (15). While small numbers prevent us from making conclusive statements, spatial distributions in ERA5 and SCREAM seem consistent.

Fig. 21c shows the frequency of ECs by latitude band. ECs are counted separately in each 6 hourly snapshot in this plot, so counts in this plot are much higher than the $\sim$80 storms quoted above for SCREAM and ERA5, which tracked single storms across time. In both hemispheres, SCREAM has a more peaked distribution with maximum frequency at the upper limit of the observed count from the 1979-2020 period. The excessively peaked EC count structure in the northern hemisphere is consistent with zonal precipitation bias shown in Fig. 13. Interestingly, modeled southern hemisphere storm track precipitation in Fig. 13 matches ERA5 almost perfectly despite having excessive EC count around 50°S. Storm composites show that Southern Hemisphere extratropical cyclones in SCREAM are associated with less rain than ERA5, which might explain this apparent paradox (not shown). Peak latitude is roughly consistent with observations in each hemisphere, though is displaced slightly poleward in the northern hemisphere.

We noted earlier that large swaths of the Southern Ocean in SCREAM have too much absorbed shortwave radiation compared to CERES-EBAF retrievals (Fig. 9). Many climate models share biases where the cold sector of storms does not reflect enough incoming shortwave radiation, while the warm sector is less biased (Bodas-Salcedo et al., 2014). To examine whether this is the case in SCREAM, we construct composites of the cyclones tracked in SCREAM between 40°S and 60°S. This latitude band is consistent with those of Bodas-Salcedo et al. (2014), but ignores storms with centers poleward of 60°S (to remove complications due to the reflectivity of sea ice). Fig. 22 shows the composite of the pseudo-cloud albedo for SCREAM and its difference with CERES-SYN-based estimates. The pseudo-cloud albedo is defined here as the shortwave cloud radiative effect divided by the local solar insolation. By using a pseudo-cloud albedo rather than reflected shortwave radiation, we remove the potential impact of biases in the latitudinal distribution of ECs on our assessment of SCREAM's cloud reflectivity. Indeed, like the GCMs studied by Bodas-Salcedo et al. (2014), there is less cloud reflection in the cold sector of SCREAM's storms (-4.9 % in the cold western half of the storm), compared to the storms captured in ERA5. However, the warm-sector of the storm also shows lower cloud albedo (-3.8 % in the warm eastern half of the storm), showing that in SCREAM, there is a general lack of cloud reflection. Fixing this bias is a research priority.

6.7 Atmospheric Rivers

Atmospheric rivers (ARs) are long, narrow, and transient corridors of enhanced vapor transport typically associated with the low-level jet stream ahead of the cold front of an extratropical cyclone (AMS, 2019). As noted by Zhu and Newell (1998), ARs are responsible for approximately 90% of poleward vapor transport. Water resources in the western U.S. are strongly tied to ARs, with landfalling ARs providing approximately $20-50\%$ of total wet season precipitation (Dettinger et al., 2011; Lavers & Villarini, 2015)

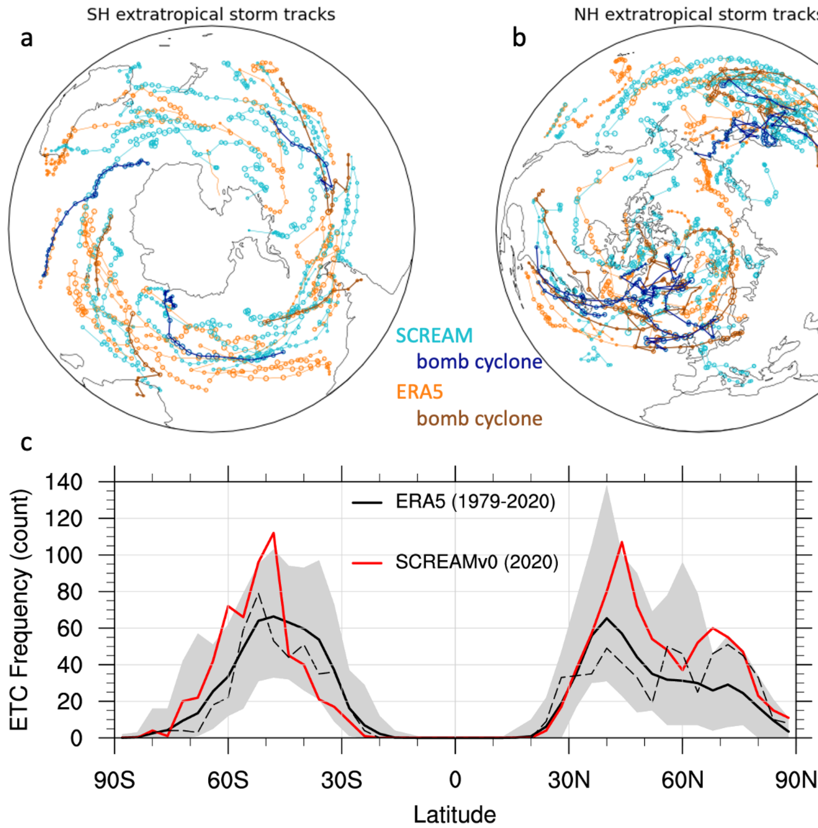

Figure 21. Geographic distribution of extratropical cyclones identified in SCREAM (cyan) and ERA5 (orange) using the TE algorithm (described in Appendix A2) for the Southern Hemisphere (a) and Northern Hemisphere (b). Dark blue tracks indicate bomb cyclones in SCREAM, whereas brown tracks indicate bomb cyclones in ERA5. (c) The latitudinal distribution of 6 hourly snapshots of extratropical cyclones in ERA5 (black) and SCREAM (red). The dashed black line indicates the distribution found in ERA5 for the DYAMOND2 period (Jan 20 through Feb 28, 2020). Solid black line indicates the average distribution for Jan 20 to Feb 28 of 1979 through 2020 in ERA5 with gray shading indicating maximum and minimum ranges for each year.

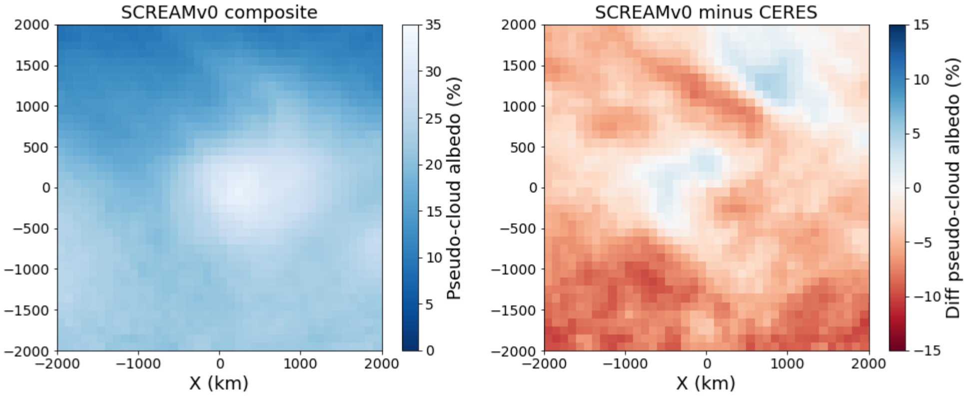

Figure 22. (left) Composite of the pseudo-cloud albedo in extratropical cyclones found between 45°S and 60°S in SCREAM (Jan 20 – Feb 28). Composites average over all 6 hourly snapshots centered on identified ECs and are plotted such that north is oriented upward. (right) Difference in storm composite pseudo-cloud albedo between storms in SCREAM and in reanalysis and satellite data (ERA5 / CERES-SYN) for storms occurring during the same period.

and 30 − 40% of mountain snowpack (Guan et al., 2010). One such landfalling atmospheric river observed in the SCREAM simulation along the west coast of North America is depicted in Fig. 23.

To assess the quality of ARs in the SCREAM simulation, we track ARs over the simulation period using the TempestExtremes atmospheric river detection and tracking algorithm (McClenny et al., 2020; Ullrich & Zarzycki, 2017) as described in Appendix A3. In Fig. 24 the properties of these tracked features are then compared to analogously tracked features from all January 20 through Mar 28 periods in ERA5 data (1979-2020), roughly following the approach discussed in Rutz et al. (2019). In general SCREAM falls well within the climatological range from ERA5 historical simulations, except for a slight underestimation of AR frequency around 35° north and south of 50°S. For 2020, ERA5 predicts abnormally high AR activity while SCREAM is slightly weaker than ERA5's long-term average. Without an ensemble of simulations to compare against, however, such a discrepancy could very easily be attributed to interannual variability.

The underestimation of AR frequency in southern high latitudes is associated with anomalously low eastward integrated vapor transport (IVT), which is in turn due to anomalously low eastward wind speeds compared to ERA5 (not shown). Interestingly, Fig. 21 shows that EC frequency was actually *too high* where we find AR frequency to be too low. Perhaps ECs are spending too much time in this region due to low wind speeds? Nonetheless, the fractional contribution of ARs to poleward transport of moisture is almost identical to the climatological mean performance from ERA5, suggesting consistency of the underlying physical processes. Overall we conclude that SCREAM performs well in its representation of ARs and their associated contribution to poleward transport of vapor.

6.8 Cold-Air Outbreaks

Marine cold air outbreaks (MCAOs) occur when cold air of polar or continental origin flow over warm ocean waters. Because of the strong air-sea temperature differences and typical higher surface wind speeds, cold air outbreaks are regions of strong surface turbulent heat fluxes that can reach 1000 W m^{-2} (Shapiro et al., 1987) and can im-

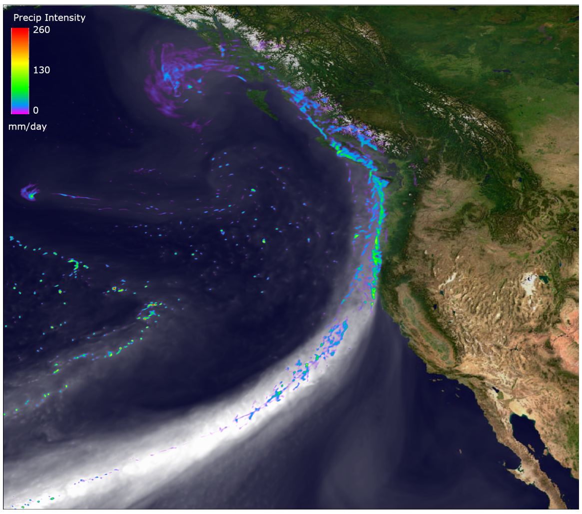

Figure 23. Snapshot of a landfalling atmospheric river along the west coast of North America that occurs on February 11th 23:00:00 UTC. Vertically-integrated water vapor is indicated in transparent grayscale with opaque/white regions having integrated vapor greater than 40 kg m^{-2}. Colors indicate precipitation intensity.

pact frontogenesis (Terpstra et al., 2016). General circulation models (GCMs) have, however, not represented clouds under these conditions very well (Rémillard & Tselioudis, 2015). The models tend to simulate too little stratiform cloud cover in these regions (Field et al., 2014; Bodas-Salcedo et al., 2014). In this section, we describe the frequency and intensity of MCAOs in the SCREAM simulation relative to reanalysis (ERA5) during the same time period and examine the surface flux and cloud properties for a single cold air outbreak event that occurs early in the simulation over the Kuroshio current.

To identify and quantify cold air outbreaks, we use the cold air outbreak index (M) as described by Fletcher et al. (2016), which is quantified as the potential temperature difference between the surface skin and 800hPa. Any oceanic region with a positive value of M denotes a region undergoing a cold air outbreak. If we compare the frequency of cold air outbreaks in SCREAM and in ERA5 over the global oceans, we see general agreement of where and how often cold air outbreaks occur (Fig. 25a and c). Cold air outbreaks tend to occur most prominently in the winter Northern Hemisphere along the eastern edges of continents and southern edges of the sea-ice. In regions where SCREAM produces cold air outbreaks (e.g. over the Kuroshio current, Gulf stream current, and south of Alaska), M frequency tends to be higher. MCAOs are, however, greatly underestimated to the south and east of Greenland. This is unsurprising since 2-m temperature is far too warm over Greenland (Fig. 6), likely due to meridional wind biases discussed in Sect. 6.2. Except for a slight overestimation, SCREAM also tends to capture well the intensity of the strongest of cold air outbreaks (Fig. 25b and d).

To study the cloud fields that form under the simulated cold air outbreaks in SCREAM, we focus on a cold air outbreak event that flows off the Asian continent over the Kuroshio current from Jan 21st to Jan 22nd. We examine the cold air outbreak characteristics over the 24 hour period of Jan 22nd to exclude any impacts of the cold front. The simulated sensible heat flux generally matches ERA5, but is a bit too smooth and too big (Fig. 26a and d). Good spatial agreement may be an artifact of prescribed SST; smooth features are probably due to use of a coarser ($\sim$6 km) ocean grid in this region. Excessive magnitude is unsurprising given surface wind speed biases mentioned in Sect. 6.2 and again

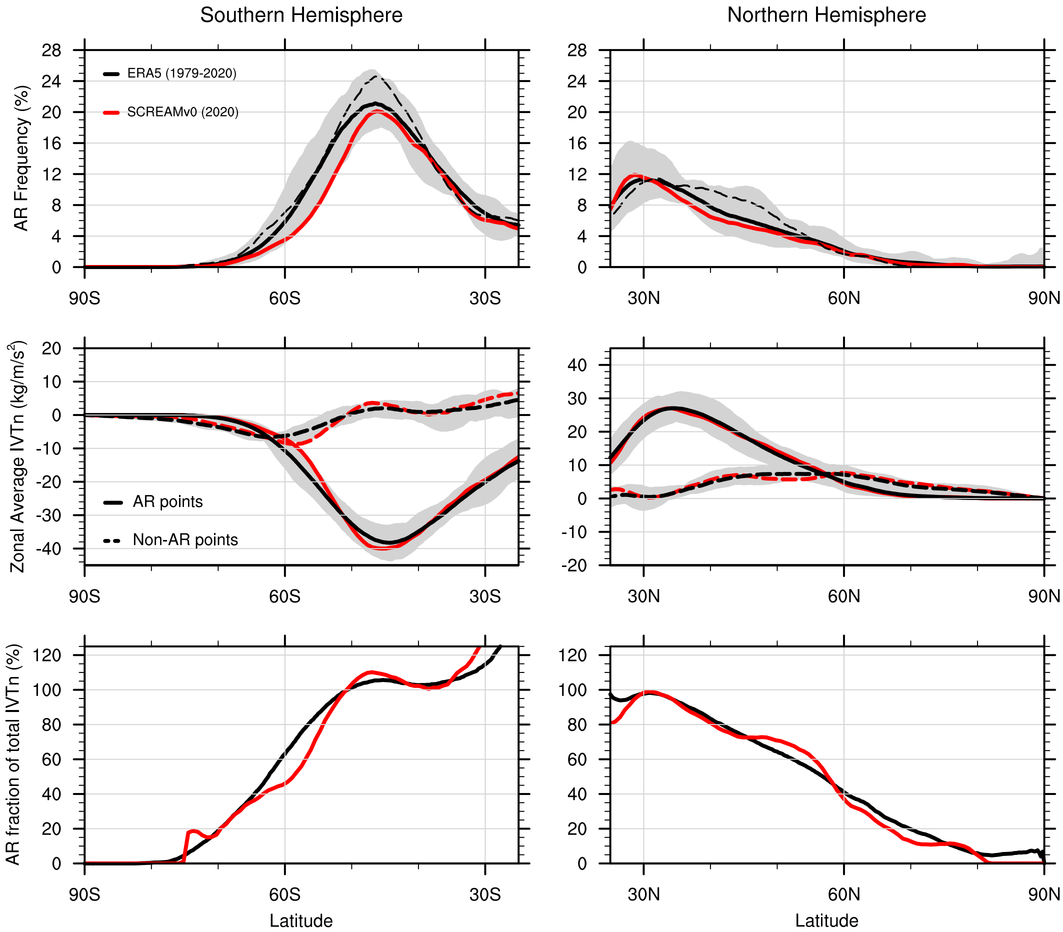

Figure 24. Properties of tracked atmospheric rivers in both hemispheres over the period January 20 through February 28 of each year in (red) the SCREAM DYAMOND2 simulation and (gray shaded region with mean shown with black solid line) 1979-2020 ERA5 reanalysis. Plots refer to (top) average atmospheric river frequency, as a percent of the full longitudinal band, with results from 2020 depicted with a black dashed line; (middle) zonally averaged northward integrated vapor transport (IVTn) at grid points flagged as part of / not part of atmospheric rivers; (bottom) mean fractional contribution of northward vapor transport from atmospheric rivers relative to all northward vapor transport.

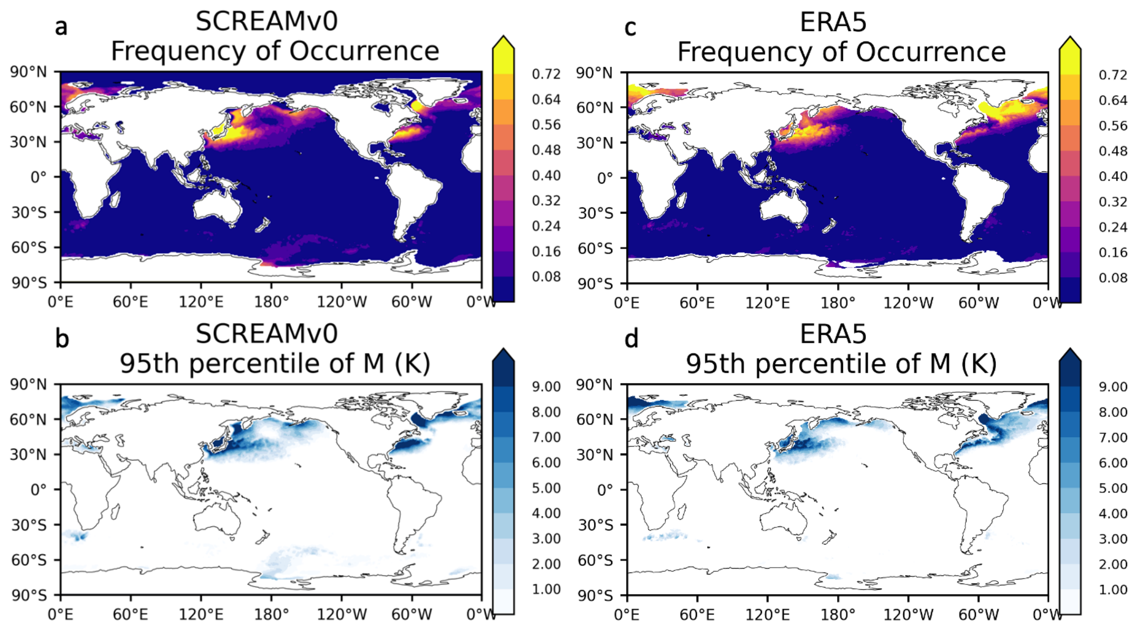

Figure 25. Frequency of cold air outbreaks (based on the M of Fletcher et al., 2016) in SCREAM over the month of February 2020 in SCREAM (a) and in ERA5 (c). Also shown is the 95th percentile value of M (including non-cold air outbreak instances) over the same period in SCREAM (b) and ERA5 (d).

apparent from comparing Fig. 26 panels b versus e. Surface air temperature bias does not contribute to excessive surface fluxes (not shown).

Although GCMs tend to underestimate the occurrence of MCAO clouds and SCREAM itself was shown earlier to suffer from a deficiency in clouds in other regimes, a comparison of the shortwave cloud radiative effect between the model and CERES-SYN suggests good agreement in the MCAO regime (Fig. 26c and f). In Fig. 27 we take a closer look at cloud structure by comparing a snapshot of shortwave cloud radiative effect from SCREAM against a visible satellite image taken at the same time from Himawari-8 (Bessho et al., 2016). Similarity between the observed and simulated cloud structures is striking, particularly since this image is taken 2 days into a free-running simulation. In particular, cloud streets in SCREAM form along the direction of the flow, before transitioning into broken and open-cellular convection further offshore. The model's ability to capture this transition suggests that SCREAM's combination of resolution and boundary layer/cloud parameterizations contains the physics necessary to capture cloud transitions in cold air outbreaks. Further analyses compositing many cold air outbreak events would be necessary to draw more general conclusions.

7 Conclusions

The overall takeaway from this work is that 3.25 km global models solve a lot of the long-standing problems in global climate modeling even without the detailed optimization and tuning which is typically so important for GCM skill. In particular, SCREAM does an excellent job simulating precipitation; its diurnal cycle (Fig. 15) and intensity distribution (Fig. 16) are particularly realistic. Tropical and extratropical storm frequency and structure (Sections 6.5-6.7) are also impressive. The vertical structure of tropical convection (Fig. 10) is also much improved relative to typical GCMs. Coastal stratocu-

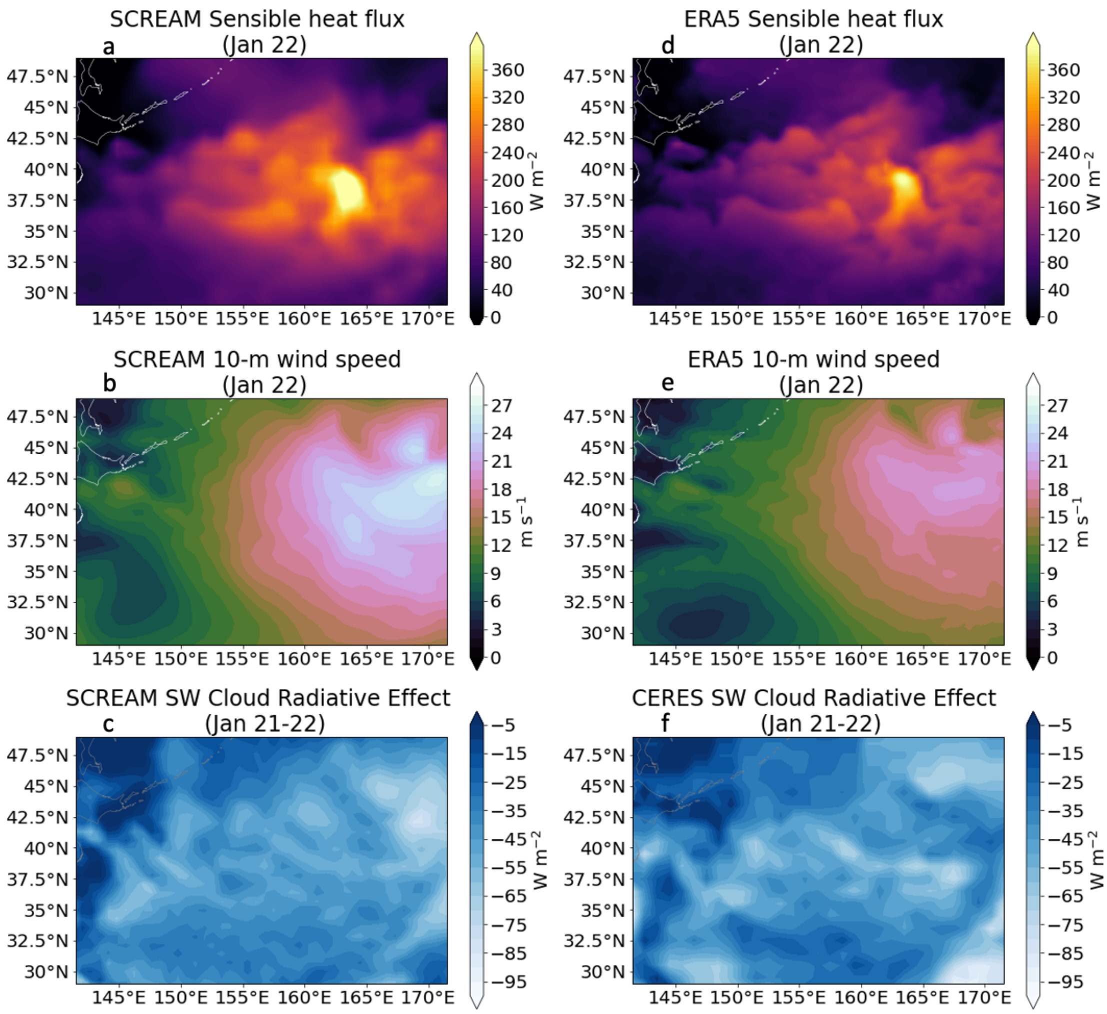

Figure 26. The daily-mean sensible heat flux over the Kuroshio region bounded from 29N to 49N and 141.5E to 171.5E in SCREAM (a) and ERA5 (d) for the cold air outbreak on January 22. Also shown are similar daily mean values of 10-m wind speed (b - SCREAM; e - ERA5) and shortwave cloud radiative effect (c - SCREAM; f - CERES-SYN).

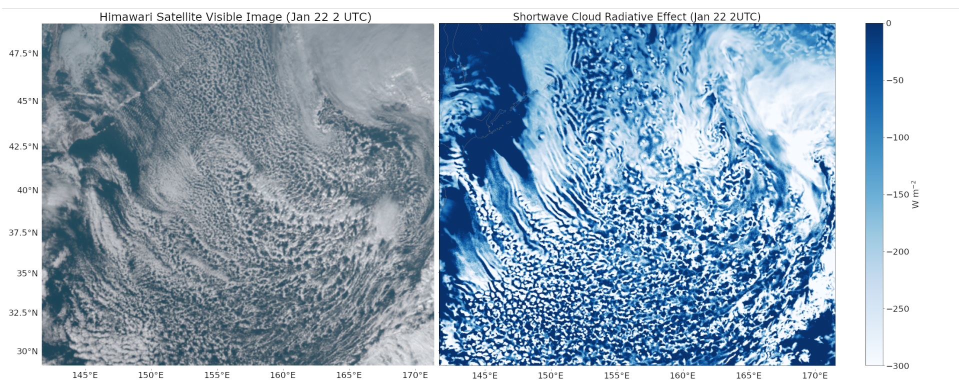

Figure 27. Cold-air outbreak off Siberia on January 22nd 2020 at 2:00:00 UTC (∼local noon) from a Himawari visible satellite image (left) and shortwave cloud radiative effect from SCREAM (right). Visualization is over a region bounded by 29N to 49N and 141.5E to 171.5E.

mulus (Fig. 11) and cold-air outbreaks (25-27), which are perennially difficult to simulate not just in GCMs (Rémillard & Tselioudis, 2015) but also in limited-area CPMs (Klein et al., 2009), are also well captured. We suspect that the SHOC cloud/turbulence parameterization and fine vertical resolution within SCREAM were important for this success.

Several biases in SCREAM are familiar from conventional GCMs. Clarifying whether these biases are caused by processes unresolved at 3.25 km grid spacing would be a large step towards understanding and therefore fixing these perennial problems. One such bias is the tendency for the South Pacific Convergence Zone to be too zonal (Fig. 13-14). This suggests that resolution doesn't resolve the double-ITCZ bias that plagues lower-resolution models. This finding is consistent with the result of Stevens et al. (2019) for other GCPMs. Another bias in lower-resolution versions of E3SM which persists in SCREAM is a tendency for precipitation in the West Pacific to be maximized over the Maritime Continent rather than to the east over the ocean.

Analysis for this paper also revealed several deficiencies which will be fixed in future model versions. First, cloud fraction near the tropopause is corrupted by the use of a relative-humidity based ice cloud fraction scheme tuned for low resolutions (Fig. 4). Because these spurious clouds had no mass, they had little practical impact on the simulation, but users of SCREAM DYAMOND2 data should be careful to use our post-facto-generated cloud-mask-based cloud fraction for future analysis. Overly strong surface wind speed is a second deficiency (Fig. 7). Upper level winds are generally reasonable but have unrealistic poleward transport south of Greenland and around Australia (Fig. 8). Surface temperature at high latitudes is also problematic (Fig. 6). One potential reason for this is a land initial condition with low snowpack in mountainous regions exacerbated by potentially poor tuning of the lower limit on turbulent mixing in stable conditions and aforementioned biases related to heat transport into polar regions. Another issue is a prevalence of frequent, small "popcorn" convective events (Fig. 17). Finally, cloud tuning should be improved. Shortwave reflection and longwave emission are too weak (Fig. 9) and low-level clouds tend too much towards stratus and too little towards shallow convection (Fig. 12). Issues like these are expected for a new model version and many of these issues have an obvious solution. We are releasing this initial model without fixing these problems to match the timing of the DYAMOND2 intercomparison, because there will *always* be something else to fix, and because using a model for science and writing papers is by far the fastest way to find problems.

This simulation is a milestone rather than an endpoint in SCREAM development. In addition to fixing the issues identified above, the major focus of the SCREAM project is on completing the computationally-performant C++ implementation of the model. We hope to perform longer, more realistic simulations soon.

Appendix A Feature tracking with TempestExtremes

For feature tracking in the DYAMOND2 simulation we use TempestExtremes 2.1 (Ullrich & Zarzycki, 2017), available from ZENODO at `http://dx.doi.org/10.5281/zenodo.4385656` and GitHub at `https://github.com/ClimateGlobalChange/tempestextremes`. The exact commands employed in this analysis are provided in this section for reference.

A1 Tropical Cyclones

Tropical cyclone tracking is performed on 6-hourly data following (Zarzycki et al., 2017). The search is performed for local minima in the sea level pressure (PSL) which are accompanied by an increase of 200 Pa over a distance of 5.5 degrees great circle distance (GCD). Tropical cyclones are further defined by the presence of an upper-level warm core which is characterized by anomalous thickness in the geopotential height between

500 hPa and 200 hPa. Here we require that this thickness drop by 6.0 meters over a distance of 6.5 degrees GCD, where the maxima in the layer thickness must be within 1.0 degrees GCD of the pressure minima. Following this only the most intense features within 6.0 degrees GCD are retained. Tracks are then stitched together in time, where sequential features must be within 8.0 degrees GCD, must persist for at least 10 time steps (2.5 days), can have no more than 3 sequential 6-hourly time steps where no detection is found, must have a 10 meter wind speed greater than 10 m s^{-1} for at least 10 steps along the trajectory, and must be within 50°S and 50°N for at least 10 steps along the trajectory. The commands are as follows:

```
$TEMPESTEXTREMESDIR/DetectNodes --in_data_list DYAMOND_TC_files.txt
--out DYAMOND_DN.txt --searchbymin PSL
--closedcontourcmd "PSL,200.0,5.5,0;_DIFF(Z200,Z500),-6.0,6.5,1.0"
--mergedist 6.0 --outputcmd "PSL,min,0;WINDSPD_10M,max,2" --timefilter "6hr"

$TEMPESTEXTREMESDIR/StitchNodes --in DYAMOND_DN.txt
--out DYAMOND_TC_tracks.txt --in_fmt "lon,lat,slp,wind" --range 8.0
--mintime "10" --maxgap "3"
--threshold "wind,>=,10.0,10;lat,<=,50.0,10;lat,>=,-50.0,10"
```

A2 Extratropical Cyclones

As with tropical cyclones, extratropical cyclone tracking is performed on 6-hourly data. Candidates are first detected as minima in the difference between the sea-level pressure (PSL) and the average sea-level pressure over the entire simulation (PSL_climo). We require that this difference increase by 200 Pa within 5.5 degrees GCD of the candidate. We further eliminate points that have an upper-level warm core, as these are likely tropical cyclones, by removing candidates with a drop in the 500-200hPa layer thickness of 6.0 meters within 6.5 degrees GCD of the point of maximum layer thickness within 1.0 degrees of the candidate. Following this only the most intense features within 6.0 degrees GCD are retained. Tracks are then stitched together in time, where sequential features must be within 8.0 degrees GCD, must persist for at least 8 time steps (2.0 days), can have no more than 2 sequential 6-hourly time steps where no detection is found, must have a surface geopotential less than 700.0 for at least 8 time steps, and must have a distance of 6.0 degrees GCD between genesis and termination point. The commands for these operations are as follows:

```
$TEMPESTEXTREMESDIR/bin/DetectNodes --in_data_list DYAMOND_ETC_files.txt
--out DYAMOND_DN_ETCs.txt --searchbymin "_DIFF(PSL,PSL_climo)" --timefilter "6hr"
--closedcontourcmd "_DIFF(PSL,PSL_climo),200.0,5.5,0"
--noclosedcontourcmd "_DIFF(Z200,Z500),-6.0,6.5,1.0" --mergedist 6.0
--outputcmd "PSL,min,0;_DIFF(PSL,PSL_climo),min,0;WINDSPD_10M,max,5;PHIS,min,0"

$TEMPESTEXTREMESDIR/bin/StitchNodes --in DYAMOND_DN_ETCs.txt
--out DYAMOND_ETC_tracks.txt --in_fmt "lon,lat,psl,pslanom,wind,phis" --range 8.0
--mintime "8" --maxgap "2" --min_endpoint_dist 6.0 --threshold "phis,<=,700,8"
```

A3 Atmospheric Rivers

Atmospheric river tracking is performed using the tracker employed in (McClenny et al., 2020). Grid points poleward of 15 degrees N/S are flagged where the Laplacian of the integrated vapor transport (evaluated using 8 points with radius 10 degrees GCD) is less than 20000 kg m^{-1} s^{-1} rad^{-2}. Only contiguous regions with area greater than 4×10^5 km^2 are retained in this operation. Since high IVT blobs can include tropical cyclones, we also remove all points within 10 degrees GCD of TCs detected using the method described in section A1. The commands for these operations are as follows:

```
$TEMPESTEXTREMESDIR/DetectBlobs --in_data CAT_TUQ,TVQ_256x512.eam.nc
--out CAT_ARs_256x512.eam.nc --minabslat 15 --geofiltercmd "area,>=,4e5km2"
--thresholdcmd "_LAPLACIAN{8,10}(_VECMAG(TUQ,TVQ)),<=,-20000,0"

$TEMPESTEXTREMESDIR/NodeFileFilter --in_nodefile DYAMOND_TC_tracks.txt
--in_fmt "lon,lat" --in_data CAT_ARtag_256x512.eam.nc
--out_data CAT_ARtag_TCfiltered_256x512.eam.nc --var "binary_tag"
--bydist 10.0 --invert
```

Acknowledgments

This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. It used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility located at Lawrence Berkeley National Laboratory, operated under Contract No. DE-AC02-05CH11231. This research also used resources from the Argonne Leadership Computing Facility at Argonne National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-06CH11357. This paper was prepared by LLNL under Contract DE-AC52-07NA27344. H.-Y. Ma is funded by the Regional and Global Model Analysis program area and Atmospheric System Research program of the U.S. Department of Energy. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government. The SCREAM output described in this paper is publicly available as part of the DYAMOND2 intercomparison as described at `https://www.esiwace.eu/services/dyamond`. The GPM dataset used in this study is Version 06B Level 3 half-hourly 0.1 degree x 0.1 degree final-run gauge-calibrated data as described in G. Huffman et al. (2019) and accessible via `https://disc.gsfc.nasa.gov/datasets/GPM_3IMERGHH_06/summary`. We acknowledge and thank Jingyu Wang for accessing, formatting, and making the data available to us. The version of CERES SYN1deg data used here is Edition 4.1 with a release date of August 22, 2019 and was accessed from: `https://ceres-tool.larc.nasa.gov/ord-tool/jsp/SYN1degEd41Selection.jsp`. The Himawari-8 visible image used in this paper was supplied by the P-Tree System of the Japan Aerospace Exploration Agency (JAXA). The authors would also like to thank Jiwen Fan and Kai Zhang for their contributions to fixing bugs in our P3 microphysics implementation and 3 anonymous reviewers for excellent suggestions to improve the manuscript.

References

Abdul-Razzak, H., & Ghan, S. J. (2000). A parameterization of aerosol activation: 2. multiple aerosol types. *Journal of Geophysical Research: Atmospheres*, *105*(D5), 6837-6844. doi: https://doi.org/10.1029/1999JD901161

Agel, L., Barlow, M., Qian, J.-H., Colby, F., Douglas, E., & Eichler, T. (2015). Climatology of Daily Precipitation and Extreme Precipitation Events in the Northeast United States. *Journal of Hydrometeorology*, *16*(6), 2537–2557. doi: 10.1175/JHM-D-14-0147.1

AMS. (2019, March). *Atmospheric river.* American Meteorological Society. Retrieved from http://glossary.ametsoc.org/wiki/Atmospheric_river

Arnold, N. P., Putman, W. M., & Freitas, S. R. (2020). Impact of Resolution and Parameterized Convection on the Diurnal Cycle of Precipitation in a Global Nonhydrostatic Model. *Journal of the Meteorological Society of Japan. Ser. II*, *advpub*. doi: 10.2151/jmsj.2020-066

Ascher, U. M., Ruuth, S. J., & Spiteri, R. J. (1997). Implicit-explicit Runge-Kutta methods for time-dependent partial differential equations. *Applied Num. Math.*, *25*(2-3), 151–167. doi: 10.1016/s0168-9274(97)00056-1

Atlas, R., Reale, O., Shen, B.-W., Lin, S.-J., Chern, J.-D., Putman, W., ... Radakovich, J. (2005). Hurricane forecasting with the high-resolution nasa finite volume general circulation model. *Geophysical Research Letters*, *32*(3). doi: 10.1029/2004GL021513

Austin, R. T., Heymsfield, A. J., & Stephens, G. L. (2009). Retrieval of ice cloud microphysical parameters using the cloudsat millimeter-wave radar and temperature. *Journal of Geophysical Research: Atmospheres*, *114*(D8). doi: https://doi.org/10.1029/2008JD010049

Bacmeister, J. T., Wehner, M. F., Neale, R. B., Gettelman, A., Hannay, C., Lauritzen, P. H., ... Truesdale, J. E. (2014). Exploratory high-resolution climate simulations using the community atmosphere model (cam). *Journal of Climate*, *27*(9), 3073-3099. doi: 10.1175/JCLI-D-13-00387.1

Bergeron, T. (1935). *On the physics of clouds and precipitation.* Proces Verbaux de l'Association de Météorologie, International Union of Geodesy and Geophysics.

Bertagna, L., Deakin, M., Guba, O., Sunderland, D., Bradley, A. M., Tezaur, I. K., ... Salinger, A. G. (2019). Hommexx 1.0: a performance-portable atmospheric dynamical core for the energy exascale earth system model. *Geoscientific Model Development*, *12*(4), 1423–1441. doi: 10.5194/gmd-12-1423-2019

Bertagna, L., Guba, O., Taylor, M., Foucar, J., Larkin, J., Bradley, A., ... Salinger, A. (2020). A performance-portable nonhydrostatic atmospheric dycore for the energy exascale earth system model running at cloud-resolving resolutions. In *2020 SC20: International Conference for High Performance Computing, Networking, Storage and Analysis (SC)* (p. 1304-1317). Los Alamitos, CA, USA: IEEE Computer Society. doi: 10.1109/SC41405.2020.00096

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., ... Yoshida, R. (2016). An introduction to himawari-8/9; japan's new-generation geostationary meteorological satellites. *Journal of the Meteorological Society of Japan. Ser. II*, *94*(2), 151-183. doi: 10.2151/jmsj.2016-009

Bodas-Salcedo, A., Williams, K. D., Ringer, M. A., Beau, I., Cole, J. N. S., Dufresne, J.-L., ... Yokohata, T. (2014). Origins of the Solar Radiation Biases over the Southern Ocean in CFMIP2 Models. *Journal of Climate*, *27*(1), 41–56. doi: 10.1175/JCLI-D-13-00169.1

Bogenschutz, P., & Krueger, S. K. (2013). A simplified PDF parameterization of subgrid-scale clouds and turbulence for cloud-resolving models. *Journal of Advances in Modeling Earth Systems*, *5*(2), 195-211. doi: https://doi.org/10.1002/jame.20018

Bogenschutz, P., Yamaguchi, T., & Lee, H.-H. (2021). *E3SM Simulations with high vertical resolution in the lower troposphere.* (Accepted, J. Adv. Modeling Earth Sys.)

Boyle, J., & Klein, S. A. (2010). Impact of horizontal resolution on climate model forecasts of tropical precipitation and diabatic heating for the twp-ice period. *Journal of Geophysical Research: Atmospheres*, *115*(D23). doi: 10.1029/2010JD014262

Bradley, A. M., Bosler, P. A., Guba, O., Taylor, M. A., & Barnett, G. A. (2019). Communication-efficient property preservation in tracer transport. *SIAM Jour-*

nal on Scientific Computing, *41*(3), C161–C193. doi: 10.1137/18m1165414

Bretherton, C. S., & Khairoutdinov, M. F. (2015). Convective self-aggregation feedbacks in near-global cloud-resolving simulations of an aquaplanet. *Journal of Advances in Modeling Earth Systems*, *7*(4), 1765-1787. doi: https://doi.org/10.1002/2015MS000499

Bretherton, C. S., & Park, S. (2009). A new moist turbulence parameterization in the community atmosphere model. *Journal of Climate*, *22*(12), 3422 - 3448. doi: 10.1175/2008JCLI2556.1

Caldwell, P. M. (2010). California wintertime precipitation bias in regional and global climate models. *Journal of Applied Meteorology and Climatology*, *49*(10), 2147 - 2158. doi: 10.1175/2010JAMC2388.1

Caldwell, P. M., Mametjanov, A., Tang, Q., Van Roekel, L. P., Golaz, J.-C., Lin, W., … Zhou, T. (2019). The DOE E3SM Coupled Model Version 1: Description and Results at High Resolution. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4095-4146. doi: https://doi.org/10.1029/2019MS001870

Carter-Edwards, H., Trott, C. R., & Sunderland, D. (2014). Kokkos: Enabling manycore performance portability through polymorphic memory access patterns. *J. Parallel and Dist. Comp.*, *74*(12), 3202–3216.

Cheng, A., & Xu, K.-M. (2008). Simulation of boundary-layer cumulus and stratocumulus clouds using a cloud-resolving model with low-and third-order turbulence closures. *Journal of the Meteorological Society of Japan. Ser. II*, *86A*, 67-86. doi: 10.2151/jmsj.86A.67

Covey, C., Gleckler, P. J., Doutriaux, C., Williams, D. N., Dai, A., Fasullo, J., … Berg, A. (2016). Metrics for the diurnal cycle of precipitation: Toward routine benchmarks for climate models. *Journal of Climate*, *29*(12), 4461 - 4471. doi: 10.1175/JCLI-D-15-0664.1

Danabasoglu, G., Lamarque, J.-F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., … Strand, W. G. (2020). The community earth system model version 2 (cesm2). *Journal of Advances in Modeling Earth Systems*, *12*(2), e2019MS001916. (e2019MS001916 2019MS001916) doi: https://doi.org/10.1029/2019MS001916

Delworth, T. L., Rosati, A., Anderson, W., Adcroft, A. J., Balaji, V., Benson, R., … Zhang, R. (2012). Simulated climate and climate change in the gfdl cm2.5 high-resolution coupled climate model. *Journal of Climate*, *25*(8), 2755-2781. doi: 10.1175/JCLI-D-11-00316.1

Dennis, J., Edwards, J., Evans, K., Guba, O., Lauritzen, P., Mirin, A., … Worley, P. H. (2012). CAM-SE: A scalable spectral element dynamical core for the community atmosphere model. *Int. J. High Perf. Comp. Appl.*, *26*, 74–89.

Dennis, J., Fournier, A., Spotz, W. F., St.-Cyr, A., Taylor, M. A., Thomas, S. J., & Tufo, H. (2005). High resolution mesh convergence properties and parallel efficiency of a spectral element atmospheric dynamical core. *Int. J. High Perf. Comp. Appl.*, *19*, 225–235.

Dettinger, M. D., Ralph, F. M., Das, T., Neiman, P. J., & Cayan, D. R. (2011). Atmospheric Rivers, Floods and the Water Resources of California. *Water*, *3*(2), 445–478. doi: 10.3390/w3020445

Doelling, D. R., Loeb, N. G., Keyes, D. F., Nordeen, M. L., Morstad, D., Nguyen, C., … Sun, M. (2013). Geostationary enhanced temporal interpolation for ceres flux products. *Journal of Atmospheric and Oceanic Technology*, *30*(6), 1072 - 1090. doi: 10.1175/JTECH-D-12-00136.1

Doelling, D. R., Sun, M., Nguyen, L. T., Nordeen, M. L., Haney, C. O., Keyes, D. F., & Mlynczak, P. E. (2016). Advances in geostationary-derived longwave fluxes for the ceres synoptic (syn1deg) product. *Journal of Atmospheric and Oceanic Technology*, *33*(3), 503 - 521. doi: 10.1175/JTECH-D-15-0147.1

Duffy, P. B., Govindasamy, B., Iorio, J. P., Milanovich, J., Sperber, K. R., Taylor, K. E., … Thompson, S. L. (2003). High-resolution simulations of global climate, part 1: present climate. *Clim. Dyn.*, *21*, 371-390.

Edman, J. P., & Romps, D. M. (2014). An improved weak-pressure-gradient scheme for single-column modeling. *Journal of the Atmospheric Sciences*, *71*(7), 2415–2429.

Evans, K., Lauritzen, P., Mishra, S., Neale, R., Taylor, M., & Tribbia, J. (2013). AMIP simulation with the CAM4 spectral element dynamical core. *J. Climate*, *26*(3), 689–709.

Field, P. R., Cotton, R. J., McBeath, K., Lock, A. P., Webster, S., & Allan, R. P. (2014). Improving a convection-permitting model simulation of a cold air outbreak. *Quarterly Journal of the Royal Meteorological Society*, *140*(678), 124–138.

Findeisen, W. (1938). *Kolloid-meteorologische vorgänge bei neiderschlags-bildung.* (Vol. 55).

Flatau, P. J., Walko, R. L., & Cotton, W. R. (1992). Polynomial fits to saturation vapor pressure. *Journal of Applied Meteorology (1988-2005)*, *31*(12), 1507–1513.

Fletcher, J., Mason, S., & Jakob, C. (2016). The Climatology, Meteorology, and Boundary Layer Structure of Marine Cold Air Outbreaks in Both Hemispheres. *Journal of Climate*, *29*(6), 1999–2014. doi: 10.1175/JCLI-D-15-0268.1

Fosser, G., Khodayar Pardo, S., & Berg, P. (2014). Benefit of convection permitting climate model simulations in the representation of convective precipitation. *Climate Dynamics*, *44*, 45-60. doi: 10.1007/s00382-014-2242-1

Gelaro, R., Putman, W. M., Pawson, S., et al. (2015). *Evaluation of the 7-km geos-5 nature run* (Tech. Rep. Series on Glob. Model Data Assim.) Greenbelt MD, USA: NASA.

Geleyn, J., & Hollingsworth, A. (1979). An economical analytical method for the computation of the interaction between scattering and line absorption of radiation. *Beitr. Phys. Atmos.*, *52*(1), 1 – 16.

Gettelman, A., Liu, X., Ghan, S. J., Morrison, H., Park, S., Conley, A. J., ... Li, J.-L. F. (2010). Global simulations of ice nucleation and ice supersaturation with an improved cloud scheme in the community atmosphere model. *Journal of Geophysical Research: Atmospheres*, *115*(D18). doi: https://doi.org/10.1029/2009JD013797

Gettelman, A., & Morrison, H. (2015). Advanced two-moment bulk microphysics for global models. part i: Off-line tests and comparison with other schemes. *Journal of Climate*, *28*(3), 1268 - 1287. doi: 10.1175/JCLI-D-14-00102.1

Ghan, S. J., & Zaveri, R. A. (2007). Parameterization of optical properties for hydrated internally mixed aerosol. *Journal of Geophysical Research: Atmospheres*, *112*(D10). doi: https://doi.org/10.1029/2006JD007927

Giorgi, F., & Marinucci, M. R. (1996). A investigation of the sensitivity of simulated precipitation to model resolution and its implications for climate studies. *Monthly Weather Review*, *124*(1), 148 - 166. doi: 10.1175/1520-0493(1996)124⟨0148:AIOTSO⟩2.0.CO;2

Godoy, W. F., Podhorszki, N., Wang, R., Atkins, C., Eisenhauer, G., Gu, J., ... Klasky, S. (2020). Adios 2: The adaptable input output system. a framework for high-performance data management. *SoftwareX*, *12*, 100561. doi: https://doi.org/10.1016/j.softx.2020.100561

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., ... Zhu, Q. (2019). The DOE E3SM Coupled Model Version 1: Overview and Evaluation at Standard Resolution. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2089-2129. doi: https://doi.org/10.1029/2018MS001603

Golaz, J.-C., Larson, V. E., & Cotton, W. R. (2002). A pdf-based model for boundary layer clouds. part i: Method and model description. *Journal of the Atmospheric Sciences*, *59*(24), 3540 - 3551. doi: 10.1175/1520-0469(2002)059⟨3540:

APBMFB⟩2.0.CO;2
Grabowski, W. W. (2016). Towards global large eddy simulation: Super-parameterization revisited. *Journal of the Meteorological Society of Japan. Ser. II*, *94*(4), 327-344. doi: 10.2151/jmsj.2016-017
Grabowski, W. W., & Smolarkiewicz, P. K. (1999). Crcp: a cloud resolving convection parameterization for modeling the tropical convecting atmosphere. *Physica D: Nonlinear Phenomena*, *133*(1), 171-178. doi: https://doi.org/10.1016/S0167-2789(99)00104-9
Guan, B., Molotch, N. P., Waliser, D. E., Fetzer, E. J., & Neiman, P. J. (2010). Extreme snowfall events linked to atmospheric rivers and surface air temperature via satellite measurements. *Geophysical Research Letters*, *37*(20). doi: 10.1029/2010GL044696
Guba, O., Taylor, M., & St.-Cyr, A. (2014). Optimization based limiters for the spectral element method. *J. Comput. Phys.*, *267*, 176–195. doi: 10.1016/j.jcp.2014.02.029
Guba, O., Taylor, M. A., Bradley, A. M., Bosler, P. A., & Steyer, A. (2020). A framework to evaluate imex schemes for atmospheric models. *Geoscientific Model Development*, *13*(12), 6467–6480. doi: 10.5194/gmd-13-6467-2020
Hannah, W., Bradley, A., Guba, O., Tang, Q., Golaz, J.-C., & Wolfe, J. (2021). *Separating physics and dynamics grids for improved computational efficiency in spectral element earth system models.* (Accepted, J. Adv. Modeling Earth Sys.) doi: 10.1029/2020MS002419
Hartnett, E., & Edwards, J. (2021). The parallelio (pio) c/fortran libraries for scalable hpc performance..
Hawcroft, M. K., Shaffrey, L. C., Hodges, K. I., & Dacre, H. F. (2012). How much northern hemisphere precipitation is associated with extratropical cyclones? *Geophysical Research Letters*, *39*(24). doi: https://doi.org/10.1029/2012GL053866
Herrington, A. R., Lauritzen, P. H., Taylor, M. A., Goldhaber, S., Eaton, B. E., Bacmeister, J. T., ... Ullrich, P. A. (2019). Physics-dynamics coupling with element-based high-order galerkin methods: Quasi-equal-area physics grid. *Mon. Weath. Rev.*, *147*(1), 69 - 84. doi: 10.1175/MWR-D-18-0136.1
Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J.-N. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049. doi: https://doi.org/10.1002/qj.3803
Hewitt, H., Roberts, M., Mathiot, P., Biastoch, A., Blockley, E., Chassignet, E., ... Zhang, Q. (2020). Resolving and parameterising the ocean mesoscale in earth system models. *Current Climate Change Reports*. doi: 10.1007/s40641-020-00164-w
Hohenegger, C., Kornblueh, L., Klocke, D., Becker, T., Cioni, G., Engels, J. F., ... Stevens, B. (2020). Climate statistics in global simulations of the atmosphere, from 80 to 2.5 km grid spacing. *Journal of the Meteorological Society of Japan. Ser. II*, *advpub*. doi: 10.2151/jmsj.2020-005
Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., ... Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), 701 - 722. doi: 10.1175/BAMS-D-13-00164.1
Huffman, G., Stocker, E., Bolvin, D., Nelkin, E., & Tan, J. (2019). *GPM IMERG Final Precipitation L3 Half Hourly 0.1 degree x 0.1 degree V06.* (Available from Goddard Earth Sciences Data and Information Services Center, Accessed 19 March 2020) doi: 10.5067/GPM/IMERG/3B-HH/06
Huffman, G. J., & coauthors. (2019). *Nasa global precipitation measurement (gpm) integrated multi-satellite retrievals for gpm (imerg)* (NASA Algorithm Theoretical Basis Doc. No. version 06). Greenbelt MD, USA: NASA.

Hunke, E. C., & Lipscomb, W. H. (2008). *CICE: The Los Alamos sea ice model, documentation and software, version 4.0* (Tech. Rep. No. LACC-06-012). Los Alamos, NM: Los Alamos National Laboratory.

Iorio, J. P., Duffy, P. B., Govindasamy, B., Thompson, S. L., Khairoutdinov, M., & Randall, D. (2004). Effects of model resolution and subgrid-scale physics on the simulation of precipitation in the continental United States. *Clim. Dyn.*, *23*, 243-258.

Jian, B., Li, J., Yuxin, Z., He, Y., Wang, J., & Huang, J. (2020, 06). Evaluation of the cmip6 planetary albedo climatology using satellite observations. *Climate Dynamics*, *54*. doi: 10.1007/s00382-020-05277-4

Jouan, C., Milbrandt, J., Vaillancourt, P., Chosson, F., & Morrison, H. (2020). Adaptation of the predicted particles properties (p3) microphysics scheme for large-scale numerical weather prediction. *Weather and Forecasting*, *35*, 2541–2565. doi: 10.1175/WAF-D-20-0111.1

Judt, F., Klocke, D., Rios-Berrios, R., Vanniere, B., Ziemen, F., Auger, L., ... Zhou, L. (2021). Tropical Cyclones in Global Storm-Resolving Models. *Journal of the Meteorological Society of Japan. Ser. II*, *advpub*. doi: 10.2151/jmsj.2021-029

Jung, T., Gulev, S. K., Rudeva, I., & Soloviov, V. (2006). Sensitivity of extratropical cyclone characteristics to horizontal resolution in the ECMWF model. *Quarterly Journal of the Royal Meteorological Society*, *132*(619), 1839–1857. doi: https://doi.org/10.1256/qj.05.212

Jung, T., Miller, M. J., Palmer, T. N., Towers, P., Wedi, N., Achuthavarier, D., ... Hodges, K. I. (2012). High-resolution global climate simulations with the ecmwf model in project athena: Experimental design, model climate, and seasonal forecast skill. *Journal of Climate*, *25*(9), 3155-3172. doi: 10.1175/JCLI-D-11-00265.1

Kain, J. S., Weiss, S. J., Bright, D. R., Baldwin, M. E., Levit, J. J., Carbin, G. W., ... Thomas, K. W. (2008). Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing nwp. *Weather and Forecasting*, *23*(5), 931 - 952. doi: 10.1175/WAF2007106.1

Kasahara, A. (1974). Various vertical coordinate systems used for numerical weather prediction. *Mon. Weath. Rev.*, *102*, 509–522.

Kato, S., Sun-Mack, S., Miller, W. F., Rose, F. G., Chen, Y., Minnis, P., & Wielicki, B. A. (2010). Relationships among cloud occurrence frequency, overlap, and effective thickness derived from calipso and cloudsat merged cloud vertical profiles. *Journal of Geophysical Research: Atmospheres*, *115*(D4). doi: https://doi.org/10.1029/2009JD012277

Kendon, E. J., Roberts, N. M., Senior, C. A., & Roberts, M. J. (2012). Realism of Rainfall in a Very High-Resolution Regional Climate Model. *Journal of Climate*, *25*(17), 5791–5806. doi: 10.1175/JCLI-D-11-00562.1

Khairoutdinov, M., Randall, D., & DeMott, C. (2005). Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes. *Journal of the Atmospheric Sciences*, *62*(7), 2136 - 2154. doi: 10.1175/JAS3453.1

Klein, S. A., McCoy, R. B., Morrison, H., Ackerman, A. S., Avramov, A., Boer, G. d., ... Zhang, G. (2009). Intercomparison of model simulations of mixed-phase clouds observed during the arm mixed-phase arctic cloud experiment. i: single-layer cloud. *Quarterly Journal of the Royal Meteorological Society*, *135*(641), 979-1002. doi: https://doi.org/10.1002/qj.416

Knapp, K. R., Diamond, H. J., Kossin, J. P., Kruk, M. C., & Schreck, C. J. (2018). International best track archive for climate stewardship (IBTrACS) project, version 4. *NOAA National Centers for Environmental Information*. (Accessed June 2021) doi: https://doi.org/10.25921/82ty-9e16

Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J., & Neumann, C. J. (2010). The international best track archive for climate stewardship (IB-

TrACS): Unifying tropical cyclone best track data. *Bulletin of the American Meteorological Society*, *91*, 363-376. doi: 10.1175/2009BAMS2755.1

Kodama, C., Stevens, B., Mauritsen, T., Seiki, T., & Satoh, M. (2019). A New Perspective for Future Precipitation Change from Intense Extratropical Cyclones. *Geophysical Research Letters*, *46*(21), 12435–12444. doi: https://doi.org/10.1029/2019GL084001

Krueger, S. K., Morrison, H., & Fridlind, A. M. (2016). Cloud-resolving modeling: Arm and the gcss story. *Meteorological Monographs*, *57*, 25.1 - 25.16. doi: 10.1175/AMSMONOGRAPHS-D-15-0047.1

Langhans, W., Schmidli, J., Fuhrer, O., Bieri, S., & Schar, C. (2013). Long-term simulations of thermally driven flows and orographic convection at convection-parameterizing and cloud-resolving resolutions. *Journal of Applied Meteorology and Climatology*, *52*(6), 1490 - 1510. doi: 10.1175/JAMC-D-12-0167.1

Laprise, R. (1992). The Euler equations of motion with hydrostatic pressure as an independent variable. *Mon. Weath. Rev.*, *120*(1), 197–207.

Larson, V. E., & Griffin, B. M. (2013). Analytic upscaling of a local microphysics scheme. part i: Derivation. *Quarterly Journal of the Royal Meteorological Society*, *139*(670), 46-57. doi: https://doi.org/10.1002/qj.1967

Latham, R., Zingale, M., Thakur, R., Gropp, W., Gallagher, B., Liao, W., ... Li, J. (2003). Parallel netcdf: A high-performance scientific i/o interface. In *Sc conference* (p. 39). Los Alamitos, CA, USA: IEEE Computer Society. doi: 10.1109/SC.2003.10053

Lauritzen, P. H., Bacmeister, J. T., Callaghan, P. F., & Taylor, M. A. (2015). Ncar_topo (v1.0): Ncar global model topography generation software for unstructured grids. *Geoscientific Model Development*, *8*(12), 3975–3986. doi: 10.5194/gmd-8-3975-2015

Lauritzen, P. H., Mirin, A. A., Truesdale, J., Raeder, K., Anderson, J. L., Bacmeister, J., & Neale, R. B. (2011, June). Implementation of new diffusion/filtering operators in the CAM-FV dynamical core. *The International Journal of High Performance Computing Applications*, *26*(1), 63–73. Retrieved from `https://doi.org/10.1177/1094342011410088` doi: 10.1177/1094342011410088

Lauritzen, P. H., & Williamson, D. L. (2019). A total energy error analysis of dynamical cores and physics-dynamics coupling in the community atmosphere model (cam). *Journal of Advances in Modeling Earth Systems*, *11*(5), 1309-1328. doi: https://doi.org/10.1029/2018MS001549

Lavers, D. A., & Villarini, G. (2015). The contribution of atmospheric rivers to precipitation in Europe and the United States. *Journal of Hydrology*, *522*, 382–390. doi: 10.1016/j.jhydrol.2014.12.010

Lebassi-Habtezion, B., & Caldwell, P. M. (2015). Aerosol specification in single-column community atmosphere model version 5. *Geoscientific Model Development*, *8*(3), 817–828. doi: 10.5194/gmd-8-817-2015

Lee, H.-H., Bogenschutz, P., & Yamaguchi, T. (2021). The Implementation of Framework for Improvement by Vertical Enhancement (FIVE) into Energy Exasacle Earth System Model (E3SM). *Journal of Advances in Modeling Earth Systems*, in review.

Li, G., & Xie, S.-P. (2014). Tropical biases in cmip5 multimodel ensemble: The excessive equatorial pacific cold tongue and double itcz problems. *Journal of Climate*, *27*(4), 1765 - 1780. doi: 10.1175/JCLI-D-13-00337.1

Lin, S.-J. (2004). A vertically Lagrangian finite-volume dynamical core for global models. *Mon. Weath. Rev.*, *132*, 2293–2397.

Lindborg, E. (1999). Can the atmospheric kinetic energy spectrum be explained by two-dimensional turbulence? *Fluid Mech.*, *388*, 259–288.

Liu, X., Ma, P.-L., Wang, H., Tilmes, S., Singh, B., Easter, R. C., ... Rasch, P. J. (2016). Description and evaluation of a new four-mode version of the modal aerosol module (mam4) within version 5.3 of the community atmosphere model. *Geoscientific Model Development*, *9*(2), 505–522. doi: 10.5194/gmd-9-505-2016

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. doi: 10.1175/JCLI-D-17-0208.1

Love, B. S., Matthews, A. J., & Lister, G. M. S. (2011). The diurnal cycle of precipitation over the maritime continent in a high-resolution atmospheric model. *Quarterly Journal of the Royal Meteorological Society*, *137*(657), 934-947. doi: 10.1002/qj.809

Ma, H.-Y., Chuang, C. C., Klein, S. A., Lo, M.-H., Zhang, Y., Xie, S., ... Phillips, T. J. (2015). An improved hindcast approach for evaluation and diagnosis of physical processes in global climate models. *J. Adv. Mod. Earth Syst.*, *7*, 1810-1827. doi: https://doi.org/10.1002/2015MS000490

Madden, R. A., & Julian, P. R. (1971). Detection of a 40–50 day oscillation in the zonal wind in the tropical pacific. *Journal of Atmospheric Sciences*, *28*(5), 702 - 708. doi: 10.1175/1520-0469(1971)028⟨0702:DOADOI⟩2.0.CO;2

Maslowski, W., Clement-Kinney, J., Marble, D., & Jakacki, J. (2008). Towards eddy-resolving models of the Arctic Ocean. In M. Hecht & H. Hasumi (Eds.), *Ocean modeling in an eddying regime* (Vol. 177, pp. 241–264). American Geophysical Union.

McClean, J. L., Bader, D. C., Bryan, F. O., Maltrud, M. E., Dennis, J. M., Mirin, A. A., ... Worley, P. H. (2011). A prototype two-decade fully-coupled fine-resolution ccsm simulation. *Ocean Modelling*, *39*(1), 10 - 30. (Modelling and Understanding the Ocean Mesoscale and Submesoscale) doi: https://doi.org/10.1016/j.ocemod.2011.02.011

McClenny, E. E., Ullrich, P. A., & Grotjahn, R. (2020). Sensitivity of atmospheric river vapor transport and precipitation to uniform sea surface temperature increases. *Journal of Geophysical Research: Atmospheres*, *125*(21), e2020JD033421.

Milbrandt, J. A., & Morrison, H. (2016). Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. part iii: Introduction of multiple free categories. *Journal of the Atmospheric Sciences*, *73*(3), 975 - 995. doi: 10.1175/JAS-D-15-0204.1

Mitchell, D. L. (2002). Effective diameter in radiation transfer: General definition, applications, and limitations. *Journal of the Atmospheric Sciences*, *59*(15), 2330–2346.

Miura, H., Satoh, M., Nasuno, T., Noda, A. T., & Oouchi, K. (2007). A madden-julian oscillation event realistically simulated by a global cloud-resolving model. *Science*, *318*(5857), 1763–1765. doi: 10.1126/science.1148443

Miyakawa, T., Satoh, M., Miura, H., Tomita, H., Yashiro, H., Noda, A., ... Yoneyama, K. (2014). Madden–julian oscillation prediction skill of a new-generation global model demonstrated using a supercomputer. *Nature communications*, *5*, 3769. doi: 10.1038/ncomms4769

Miyamoto, Y., Kajikawa, Y., Yoshida, R., Yamaura, T., Yashiro, H., & Tomita, H. (2013). Deep moist atmospheric convection in a subkilometer global simulation. *Geophysical Research Letters*, *40*(18), 4922-4926. doi: https://doi.org/10.1002/grl.50944

Morrison, H., & Gettelman, A. (2008). A new two-moment bulk stratiform cloud microphysics scheme in the community atmosphere model, version 3 (cam3). part i: Description and numerical tests. *Journal of Climate*, *21*(15), 3642 - 3659. doi: 10.1175/2008JCLI2105.1

Morrison, H., & Grabowski, W. W. (2008). Modeling supersaturation and subgrid-scale mixing with two-moment bulk warm microphysics. *Journal of the Atmospheric Sciences*, *65*(3), 792 - 812. doi: 10.1175/2007JAS2374.1

Morrison, H., & Milbrandt, J. A. (2015). Parameterization of cloud microphysics based on the prediction of the bulk ice particle properties. part i: Scheme description and idealized tests. *J. Atmos. Sci.*, *72*, 287–311.

Murphy, D. M., & Koop, T. (2005). Review of the vapour pressures of ice and supercooled water for atmospheric applications. *Quarterly Journal of the Royal Meteorological Society*, *131*(608), 1539-1565. doi: https://doi.org/10.1256/qj.04.94

Na, Y., Fu, Q., & Kodama, C. (2020). Precipitation probability and its future changes from a global cloud-resolving model and cmip6 simulations. *Journal of Geophysical Research: Atmospheres*, *125*(5), e2019JD031926. doi: https://doi.org/10.1029/2019JD031926

Nam, C., Bony, S., Dufresne, J.-L., & Chepfer, H. (2012). The 'too few, too bright' tropical low-cloud problem in cmip5 models. *Geophysical Research Letters*, *39*(21). doi: https://doi.org/10.1029/2012GL053421

Narenpitak, P., Bretherton, C. S., & Khairoutdinov, M. F. (2017). Cloud and circulation feedbacks in a near-global aquaplanet cloud-resolving model. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1069-1090. doi: https://doi.org/10.1002/2016MS000872

Nastrom, G., & Gage, K. S. (1985). A climatology of atmospheric wavenumber spectra of wind and temperature observed by commercial aircraft. *J. Atmos. Sci.*, *42*, 950–960.

Neale, R. B., Gettelman, A., Park, S., Chen, C.-C., Lauritzen, P. H., Williamson, D. L., . . . Taylor, M. A. (2012). *Description of the NCAR Community Atmosphere Model (CAM 5.0)* (NCAR Technical Note Nos. TN-486+STR). Boulder CO USA: NCAR.

Neumann, P., Düben, P., Adamidis, P., Bauer, P., Brueck, M., Kornblueh, L., . . . Biercamp, J. (2019). Assessing the scales in numerical weather and climate predictions: Will exascale be the rescue? *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *377*. doi: 10.1098/rsta.2018.0148

Noda, A. T., Oouchi, K., Satoh, M., Tomita, H., ichi Iga, S., & Tsushima, Y. (2010). Importance of the subgrid-scale turbulent moist process: Cloud distribution in global cloud-resolving simulations. *Atmospheric Research*, *96*(2), 208-217. (15th International Conference on Clouds and Precipitation) doi: https://doi.org/10.1016/j.atmosres.2009.05.007

Ogura, Y. (1963). The evolution of a moist convective element in a shallow, conditionally unstable atmosphere: A numerical calculation. *Journal of Atmospheric Sciences*, *20*(5), 407 - 424. doi: 10.1175/1520-0469(1963)020⟨0407:TEOAMC⟩2.0.CO;2

Pfahl, S., & Wernli, H. (2012). Quantifying the Relevance of Cyclones for Precipitation Extremes. *Journal of Climate*, *25*(19), 6770–6780. doi: 10.1175/JCLI-D-11-00705.1

Phillips, T. J., Potter, G. L., Williamson, D. L., Cederwall, R. T., Boyle, J. S., Fiorino, M., . . . Yio, J. J. (2004). Evaluating parameterizations in general circulation models: Climate simulation meet with weather prediction. *Bull. Amer. Meteor. Soc.*, *85*, 1903-1915. doi: https://doi.org/10.1175/BAMS-85-12-1903

Pincus, R., Barker, H. W., & Morcrette, J.-J. (2003). A fast, flexible, approximate technique for computing radiative transfer in inhomogeneous cloud fields. *Journal of Geophysical Research: Atmospheres*, *108*(D13). doi: https://doi.org/10.1029/2002JD003322

Pincus, R., Mlawer, E. J., & Delamere, J. S. (2019). Balancing accuracy, efficiency, and flexibility in radiation calculations for dynamical models. *Journal of Advances in Modeling Earth Systems*, *11*(10), 3074-3089. doi: https://doi.org/10.1029/2019MS001621

Pope, V. D., & Stratton, R. A. (2002). The processes governing horizontal resolution sensitivity in a climate model. *Clim. Dyn.*, *19*, 211-236.

Prein, A. F., Gobiet, A., Suklitsch, M., Truhetz, H., Awan, N. K., Keuler, K., & Georgievski, G. (2013). Added value of convection permitting seasonal simulations. *Climate Dynamics*, *41*(9-10), 2655-2677. doi: 10.1007/s00382-013-1744-6

Prein, A. F., Langhans, W., Fosser, G., Ferrone, A., Ban, N., Goergen, K., … Leung, R. (2015). A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Reviews of Geophysics*, *53*(2), 323-361. doi: https://doi.org/10.1002/2014RG000475

Putman, W. M., & Suarez, M. (2011). Cloud-system resolving simulations with the nasa goddard earth observing system global atmospheric model (geos-5). *Geophysical Research Letters*, *38*(16). doi: https://doi.org/10.1029/2011GL048438

Randall, D., Khairoutdinov, M., Arakawa, A., & Grabowski, W. (2003). Breaking the cloud parameterization deadlock. *Bulletin of the American Meteorological Society*, *84*(11), 1547 - 1564. doi: 10.1175/BAMS-84-11-1547

Rasch, P. J., Xie, S., Ma, P.-L., Lin, W., Wang, H., Tang, Q., … Shrivastava, M. (2019). An overview of the atmospheric component of the Energy Exascale Earth System Model. *J. Adv. Model. Earth Sys.*, *11*(8), 2377–2411. doi: 10.1029/2019ms001629

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An improved in situ and satellite sst analysis for climate. *J. Climate*, *15*, 1609-1625. doi: https://doi.org/10.1175/1520-0442(2002)015⟨1609:AIISAS⟩2.0.CO;2

Riehl, M., & Malkus, J. (1958). On the heat balance of the equatorial trough. *Geophysica (Helsinki)*, *6*, 503-538.

Rutz, J. J., Shields, C. A., Lora, J. M., Payne, A. E., Guan, B., Ullrich, P., … others (2019). The atmospheric river tracking method intercomparison project (artmip): quantifying uncertainties in atmospheric river climatology. *Journal of Geophysical Research: Atmospheres*, *124*(24), 13777–13802.

Rémillard, J., & Tselioudis, G. (2015). Cloud Regime Variability over the Azores and Its Application to Climate Model Evaluation. *J. Climate*, *28*(24), 9707–9720. doi: 10.1175/JCLI-D-15-0066.1

Sanders, F., & Gyakum, J. R. (1980). Synoptic-Dynamic Climatology of the "Bomb". *Monthly Weather Review*, *108*(10), 1589–1606. doi: 10.1175/1520-0493(1980)108⟨1589:SDCOT⟩2.0.CO;2

Sanderson, B., Piani, C., Ingram, W., Stone, D., & Allen, M. R. (2008). Towards constraining climate sensitivity by linear analysis of feedback patterns in thousands of perturbed-physics gcm simulations. *Clim. Dyn.*, *30*, 175—190. doi: 10.1007/s00382-007-0280-7

Santos, S. P., Bretherton, C., & Caldwell, P. (2020). Cloud Process Coupling and Time Integration in the E3SM Atmosphere Model. *Earth and Space Science Open Archive*, 34. (In review at JAMES) doi: 10.1002/essoar.10504538.2

Sato, T., Miura, H., Satoh, M., Takayabu, Y. N., & Wang, Y. (2009). Diurnal cycle of precipitation in the tropics simulated in a global cloud-resolving model. *Journal of Climate*, *22*(18), 4809 - 4826. Retrieved from `https://journals.ametsoc.org/view/journals/clim/22/18/2009jcli2890.1.xml` doi: 10.1175/2009JCLI2890.1

Satoh, M. (2002). Conservative scheme for the compressible nonhydrostatic models with the horizontally explicit and vertically implicit time integration scheme. *Mon. Weath. Rev.*, *130*, 1227–1245.

Satoh, M., Matsuno, T., Tomita, H., Miura, H., Nasuno, T., & Iga, S. (2008). Nonhydrostatic icosahedral atmospheric model (nicam) for global cloud resolving simulations. *Journal of Computational Physics*, *227*(7), 3486-3514. (Predicting weather, climate and extreme events) doi: https://doi.org/10.1016/

j.jcp.2007.02.006
Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W. M., & Duben, P. (2019). Global cloud-resolving models. *Curr Clim Change Rep*, *5*, 172-184. doi: https://doi.org/10.1007/s40641-019-00131-0
Satoh, M., Tomita, H., Yashiro, H., Miura, H., Kodama, C., Seiki, T., ... Kubokawa, H. (2014). The non-hydrostatic icosahedral atmospheric model: description and development. *Progress in Earth and Planetary Science*, *1*, 18. doi: 10.1186/s40645-014-0018-1
Schwartz, C. S., Kain, J. S., Weiss, S. J., Xue, M., Bright, D. R., Kong, F., ... Coniglio, M. C. (2009). Next-day convection-allowing wrf model guidance: A second look at 2-km versus 4-km grid spacing. *Monthly Weather Review*, *137*(10), 3351 - 3372. doi: 10.1175/2009MWR2924.1
Shapiro, M. A., Fedor, L. S., & Hampel, T. (1987). Research aircraft measurements of a polar low over the Norwegian Sea. *Tellus A*, *39A*(4), 272–306. doi: https://doi.org/10.1111/j.1600-0870.1987.tb00309.x
Sherwood, S., Bony, S., & Dufresne, J. (2014). Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, *505*, 37-42. doi: 10.1038/nature12829
Shi, X., & Liu, X. (2018). Sensitivity study of anthropogenic aerosol indirect forcing through cirrus clouds with cam5 using three ice nucleation parameterizations. *Journal of Meteorological Research*, *32*(qxxbywb-32-5-shixiangjun), 693. doi: 10.1007/s13351-018-8011-z
Simmons, A. J., & Burridge, D. M. (1981). An energy and angular momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Mon. Weath. Rev.*, *109*, 758–766.
Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., ... Huang, X.-Y. (2019). *A Description of the Advanced Research WRF Version 4* (Tech. Rep.). Boulder CO, USA: NCAR Tech. Note NCAR/TN-556+STR.
Skamarock, W. C., Park, S.-H., Klemp, J. B., & Snyder, C. (2014, October). Atmospheric kinetic energy spectra from global high-resolution nonhydrostatic simulations. *Journal of the Atmospheric Sciences*, *71*(11), 4369–4381. Retrieved from `https://doi.org/10.1175/jas-d-14-0114.1` doi: 10.1175/jas-d-14-0114.1
Small, R. J., Bacmeister, J., Bailey, D., Baker, A., Bishop, S., Bryan, F., ... Vertenstein, M. (2014). A new synoptic scale resolving global climate simulation using the Community Earth System Model. *Journal of Advances in Modeling Earth Systems*, *6*(4), 1065–1094. doi: 10.1002/2014ms000363
Song, X., & Zhang, G. J. (2011). Microphysics parameterization for convective clouds in a global climate model: Description and single-column model tests. *Journal of Geophysical Research: Atmospheres*, *116*(D2). doi: https://doi.org/10.1029/2010JD014833
Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettelmen, A., Golaz, J.-C., Bodas-Salcedo, A., ... Haynes, J. (2010). Dreary state of precipitation in global models. *Journal of Geophysical Research: Atmospheres*, *115*(D24). doi: https://doi.org/10.1029/2010JD014532
Stevens, B., & Bony, S. (2013). What are climate models missing? *Science*, *340*(6136), 1053–1054. doi: 10.1126/science.1237554
Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., ... Zhou, L. (2019). DYAMOND: the DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Progress in Earth and Planetary Science*, *6*(1), 61. doi: 10.1186/s40645-019-0304-z
Steyer, A., Vogl, C. J., Taylor, M., & Guba, O. (2019). Efficient IMEX Runge-Kutta methods for nonhydrostatic dynamics. *arXiv e-prints*, arXiv:1906.07219.
Storer, R. L., Zhang, G. J., & Song, X. (2015). Effects of convective microphysics parameterization on large-scale cloud hydrological cycle and radiative budget

in tropical and midlatitude convective regions. *Journal of Climate*, *28*(23), 9277 - 9297. doi: 10.1175/JCLI-D-15-0064.1

Su, H., Jiang, J. H., Teixeira, J., Gettelman, A., Huang, X., Stephens, G., ... Perun, V. S. (2011). Comparison of regime-sorted tropical cloud profiles observed by cloudsat with geos5 analyses and two general circulation model simulations. *Journal of Geophysical Research: Atmospheres*, *116*(D9). doi: https://doi.org/10.1029/2010JD014971

Taylor, M. A., & Fournier, A. (2010). A compatible and conservative spectral element method on unstructured grids. *J. Comput. Phys.*, *229*, 5879– 5895. doi: 10.1016/j.jcp.2010.04.008

Taylor, M. A., Guba, O., Steyer, A., Ullrich, P. A., Hall, D. M., & Eldrid, C. (2020). An energy consistent discretization of the nonhydrostatic equations in primitive variables. *Journal of Advances in Modeling Earth Systems*, *12*(1). doi: 10.1029/2019MS001783

Terai, C. R., Caldwell, P. M., Klein, S. A., Tang, Q., & Branstetter, M. L. (2018). The atmospheric hydrologic cycle in the acme v0.3 model. *Climate Dynamics*, *50*(9), 3251–3279. doi: 10.1007/s00382-017-3803-x

Terpstra, A., Michel, C., & Spengler, T. (2016). Forward and Reverse Shear Environments during Polar Low Genesis over the Northeast Atlantic. *Monthly Weather Review*, *144*(4), 1341–1354. doi: 10.1175/MWR-D-15-0314.1

Tomita, H., Miura, H., Iga, S., Nasuno, T., & Satoh, M. (2005). A global cloud-resolving simulation: Preliminary results from an aqua planet experiment. *Geophysical Research Letters*, *32*(8).

Ullrich, P. A., & Zarzycki, C. M. (2017). Tempestextremes: A framework for scale-insensitive pointwise feature tracking on unstructured grids. *Geoscientific Model Development*, *10*(3), 1069.

Viovy, N. (2018). *Cruncep version 7 - atmospheric forcing data for the community land model.* Boulder CO: Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory.

Wegener, A. (1911). *Thermodynamik der atmosphäre.*

Wehner, M. F., Reed, K. A., Li, F., Prabhat, Bacmeister, J., Chen, C.-T., ... Jablonowski, C. (2014). The effect of horizontal resolution on simulation quality in the Community Atmospheric Model, CAM5.1. *Journal of Advances in Modeling Earth Systems*, *6*(4), 980-997. doi: 10.1002/2013MS000276

Weisman, M. L., Skamarock, W. C., & Klemp, J. B. (1997). The resolution dependence of explicitly modeled convective systems. *Monthly Weather Review*, *125*(4), 527 - 548. doi: 10.1175/1520-0493(1997)125⟨0527:TRDOEM⟩2.0.CO;2

Williamson, D. L., Olson, J. G., Hannay, C., Toniazzo, T., Taylor, M., & Yudin, V. (2015). Energy considerations in the community atmosphere model (cam). *Journal of Advances in Modeling Earth Systems*, *7*(3), 1178-1188. doi: https://doi.org/10.1002/2015MS000448

Wiscombe, W. J. (1996). *Mie scattering calculations: Advances in technique and fast, vector-speed computer codes* (NCAR Technical Note Nos. TN-140+STR). Boulder CO, USA: NCAR.

Xie, S., Wang, Y.-C., Lin, W., Ma, H.-Y., Tang, Q., Tang, S., ... Zhang, M. (2019). Improved Diurnal Cycle of Precipitation in E3SM With a Revised Convective Triggering Function. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2290-2310. doi: https://doi.org/10.1029/2019MS001702

Yin, L., Fu, R., Shevliakova, E., & Dickinson, R. (2012). How well can cmip5 simulate precipitation and its controlling processes over tropical south america? *Climate Dynamics*, *41*. doi: 10.1007/s00382-012-1582-y

Zarzycki, C. M., & Jablonowski, C. (2015). Experimental tropical cyclone forecasts using a variable-resolution global model. *Monthly Weather Review*, *143*(10), 4012–4037.

Zarzycki, C. M., Thatcher, D. R., & Jablonowski, C. (2017). Objective tropical cyclone extratropical transition detection in high-resolution reanalysis and climate model data. *Journal of Advances in Modeling Earth Systems*, *9*(1), 130–148.

Zeng, X., Zhang, Q., Johnson, D., & Tao, W.-K. (2002). Parameterization of wind gustiness for the computation of ocean surface fluxes at different spatial scales. *Monthly Weather Review*, *130*(8), 2125 - 2133. Retrieved from `https://journals.ametsoc.org/view/journals/mwre/130/8/1520-0493_2002_130_2125_powgft_2.0.co_2.xml` doi: 10.1175/1520-0493(2002)130⟨2125:POWGFT⟩2.0.CO;2

Zhang, K., Liu, X., Yoon, J.-H., Wang, M., Comstock, J. M., Barahona, D., & Kooperman, G. (2013). Assessing aerosol indirect effect through ice clouds in cam5. *AIP Conference Proceedings*, *1527*(1), 751-751. doi: 10.1063/1.4803379

Zhang, S., Fu, H., Wu, L., Li, Y., Wang, H., Zeng, Y., ... Guo, Y. (2020). Optimizing high-resolution Community Earth System Model on a Heterogeneous Many-Core Supercomputing Platform (CESM-HR_SW1.0). *Geoscientific Model Development Discussions*, *2020*, 1–38. doi: 10.5194/gmd-2020-18

Zheng, X., Klein, S., Ghate, V., Santos, S., Mcgibbon, J., Caldwell, P., ... Cadeddu, M. (2020). Assessment of Precipitating Marine Stratocumulus Clouds in the E3SMv1 Atmosphere Model: A Case Study from the ARM MAGIC Field Campaign. *Monthly Weather Review*, *148*. doi: 10.1175/MWR-D-19-0349.1

Zhu, Y., & Newell, R. E. (1998). A Proposed Algorithm for Moisture Fluxes from Atmospheric Rivers. *Monthly Weather Review*, *126*(3), 725–735. doi: 10.1175/1520-0493(1998)126⟨0725:APAFMF⟩2.0.CO;2