An Overview of the Atmospheric Component of the Energy Exascale Earth System Model

P. J. Rasch[1], S. Xie[2], P-L Ma[1], W. Lin[3], H. Wang[1], Q. Tang[2], S. M. Burrows[1], P. Caldwell[2], K. Zhang[1], R. C. Easter[1], P. Cameron-Smith[2], B. Singh[1], H. Wan[1], J.-C. Golaz[2], B. E. Harrop[1], E. Roesler[4], J. Bacmeister[5], V. E. Larson[10], K. J. Evans[6], Y. Qian[1], M. Taylor[4], L.R. Leung[1], Y. Zhang[2], L. Brent[1], M. Branstetter[6], C. Hannay[5], S. Mahajan[6], A. Mametjanov[4], R. Neale[5], J. H. Richter[5], J-H. Yoon[9], C. S. Zender[8], D. Bader[2], M. Flanner[13], J. G. Foucar[4], R. Jacob[11], N.Keen[14], S. A. Klein[2], X. Liu[12], A. G. Salinger[4] , M. Shrivastava[1], Y. Yang[1]

[1]Pacific Northwest National Laboratory, [2]Lawrence Livermore National Laboratory, [3]Brookhaven National Laboratory, [4]Sandia National Laboratory, [5]National Center for Atmospheric Research, [6]Oak Ridge National Laboratory, [7]Los Alamos National Laboratory, [8]University of California Irvine, [9]Gwangju Institute of Science and Technology, [10]U. Wisconsin --- Milwaukee, [11]Argonne National Laboratory, [12]U. Wyoming, [13]U. Michigan, [14]Lawrence Berkeley National Laboratory

Corresponding author: Phil Rasch (philip.rasch@pnnl.gov)

Key Points:

- The atmospheric component of the Department of Energy's Energy Exascale Earth System Model is introduced.
- A brief description and evaluation is provided. Model fidelity has generally improved compared to predecessors and models participating in past international model evaluations.
- Strengths and weaknesses of the model, as well as opportunities for future work are described

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1029/2019MS001629

Abstract

The E3SM Atmosphere Model version 1 (EAMv1), the atmospheric component of the Department of Energy's E3SM (Energy Exascale Earth System Model) is described. The model began as a fork of the well-known Community Atmosphere Model (CAM) but it has evolved in new ways, and coding, performance, resolution, physical processes (primarily cloud and aerosols formulations), testing and development procedures now differ significantly. Vertical resolution was increased (from 30 to 72 layers) and the model top extended to 60 km (~0.1 hPa). A simple ozone photochemistry predicts stratospheric ozone and the model now supports increased and more realistic variability in the upper troposphere and stratosphere. An optional improved treatment of light-absorbing particle deposition to snowpack and ice is available, and stronger connections with Earth system biogeochemistry can be used for some science problems. Satellite and ground-based cloud and aerosol simulators were implemented to facilitate evaluation of clouds, aerosols and aerosol-cloud interactions. Higher horizontal and vertical resolution, increased complexity, and more predicted and transported variables have increased the model computational cost and changed the simulations considerably. These changes required development of alternate strategies for tuning and evaluation as it was not feasible to "brute force" tune the high-resolution configurations; so short-term hindcasts, perturbed parameter ensemble simulations, and regionally refined simulations provided guidance on tuning and parameterization sensitivity to higher resolution. A brief overview of the model and model climate is provided. Model fidelity has generally improved compared to its predecessors and the CMIP5 generation of climate models.

Plain Language Summary

This study provides an overview of a new computer model of the Earth's atmosphere that is used as one component of the Department of Energy's latest Earth System Model. The model can be used to help understand past, present, and future changes in Earth's behavior as the system responds to changes in atmospheric composition (like pollution and greenhouse gases), land and water use, and to explore how the atmosphere interacts with other components of the Earth system (ocean, land, biology, etc). Physical, chemical, and biogeochemical processes treated within the atmospheric model are described, and pointers to previous and recent work are listed to provide additional information. The model is compared to present day observations, and evaluated for some important tests that provide information about what could happen to clouds and the environment as changes occur. Strengths and weaknesses of the model are listed, as well as opportunities for future work.

1 Introduction

This paper describes EAMv1 (the E3SM [Energy Exascale Earth System Model] Atmosphere Model version 1). EAMv1 is a descendant of the Community Atmosphere Model (CAM, Neale et al, 2012), the atmosphere component of the Community Earth System Model (CESM, Hurrell et al, 2013). As part of the development of the E3SM project (E3SM Project, 2018), EAM was forked from the CAM development path at version 5.3, a version that contained many of the parameterizations intended for CAM6, and independent development of EAM began around June 2014 with the goal of producing a tool optimized

for DOE scientific and computational needs. It continues to share many attributes with CAM6 but as the model has evolved it is becoming increasingly distinct, and a number of different design and parameterization choices have been made to satisfy project goals.

The E3SM atmosphere major developmental thrusts were chosen to explore several key emerging questions in climate and climate change that were not easily addressed with previous versions. The E3SM project is focused on modeling and assessment of environmental change occurring over the last 40 years and next 40 years in three research areas: 1) the water cycle; 2) sea level rise; and 3) exploratory studies of the nitrogen and phosphorus on land biogeochemistry (BGC). These research areas motivated attention to specific model features to improve EAM for those science problems. E3SM has targeted the capability of scaling to large processor counts sufficient to run multi-decadal or even century and longer simulations at higher (horizontal and vertical) resolution to better capture critical atmospheric features of the water cycle in the presence of strong anthropogenic forcing, and the surface exchanges that affect other Earth system components (cryosphere, ocean, land) and introduce transient responses over century and longer timescales. E3SM has also focused on the improved treatment of some diabatic processes (primarily on boundary layer turbulence, clouds and aerosols, with less, but some attention to stratospheric ozone) that are important to tropospheric climate variability and predictability, New strategies for model tuning and evaluation (through short simulations, hindcasts and regionally refined configurations) were used during the development process to allow more complex and costly model configurations to be tuned.

Climate models are generally released at relatively low horizontal and vertical resolution to allow a broad variety of scientific problems to be studied at a relatively low computational cost, but the community recognizes that this strategy is not without problems. Orlanski (2008) showed that organized mesoscale motions are needed to produce reasonable fidelity of some atmospheric features. Jung et al (2012) showed model improvements with resolution to tropical precipitation, the tropical atmospheric circulation, and the frequency of occurrence of Euro-Atlantic blocking, and both better representation of sub- synoptic features and better resolution of topography were found to be important. Bauer and Del Genio (2006) showed that a climate model's underprediction of cloudiness and humidity in the subpolar region results from the lack of moisture transport by extratropical cyclones. Ma et al (2015) showed that cloud susceptibility to aerosol changes decreases as horizontal resolution is increased.

The EAM configurations discussed here include a lower- (100 km, LR) and high- (25 km, HR) horizontal resolution configuration. These configurations provide a moderate resolution useful for model development and capable of resolving many synoptic scale features, and a higher resolution configuration better able to address issues of 1) water cycle extremes important to our project (e.g., hurricanes, midlatitude cyclones, atmospheric rivers), water variability and availability (storms, storm tracks, blocking, snowmelt) and Sea Level Rise (SLR) issues, where resolution of the topography associated with the Greenland and Antarctic Ice sheets (GIS and AIS respectively) may become important. A variable resolution horizontal grid (termed a regionally refined model, or RRM), is also sometimes used to reduce the computational cost of high-resolution simulations.

Many climate modeling centers have focused on high horizontal resolution, but it is also clear that vertical resolution matters. Water vapor mean gradients between the surface and tropopause can vary over 3 orders of magnitude, and very large changes are often observed over a few hundred meters near the top of the atmospheric boundary layer. Clouds often occur in layers less than a hundred meters thick near the surface. Lindzen and Fox-Rabinowitz (1989) found that virtually all large-scale models had inadequate vertical resolutions and the situation has not improved much with time for many climate models. An informal look at climate models being used today indicate many put 22-28 layers between the surface and 18 km (e.g., Roeckner et al (2006); Donner et al (2011); Martin et al (2011)). Butchart et al (2018) evaluated 14 modern GCMs, and all but one appeared to have grid spacing in this range. CAM lineage models have historically found it particularly difficult to change vertical resolution near the surface. The CAM lineage used 18 layers in CCM2 (Hack et al ,1993), 27 in CAM3 (Collins et al, 2006), 30 in CAM4 (Neale et al, 2013) and CAM5 (Neale et al, 2012), and CAM6 has added two extra model layers at the tropopause, but the surface layer has remained constant at approximately 100 m over all these generations. Recent changes in the model physics described below (with an extensive discussion of consequences in Xie et al (2018)) have reduced the model sensitivity to vertical resolution, although there is still some dependence. EAMv1 uses a traditional hybridized sigma pressure vertical coordinate. The transition between terrain following and constant pressure coordinate is made at ~200 hPa (~11km). A no flux upper boundary condition is applied at the top (finite constant pressure) layer interface in the adiabatic equations of motion. The EAMv1 vertical resolution was increased from 30 layers with a top at ~2hPa (~40 km) in CAM5 to 72 layers with a top at ~0.1 hPa (64 km). The distribution of vertical layers and thickness are displayed in figure 1 of Xie et al (2018)). The surface layer is now 20 m thick, with 15 layers between the surface and 850 hPa to better capture thin clouds, sharp gradients at top of BL, rapid changes in process rates (autoconversion, accretion, evaporation, radiative heating rates) and cloud properties (drop size, rain rates). Between 850 hPa and 500 hPa vertical grid spacing is slowly increased from 100 m to 500 m because water vapor gradients are still very strong compared to important clouds features and variations in water vapor features and aerosol plumes that are frequently observed to occur at vertical scales of 500 m and less. Resolution from the free troposphere (above 500 hPa) up to the lower stratosphere (70 hPa) was increased from 1200 m to 600 m to allow for adequate representation of upward propagating large scale tropical waves such as Kelvin and mixed-Rossby gravity. This permits simulation of stratospheric features like the Quasi-Biennial-Oscillation (QBO, Richter et al. 2014; Geller et al. 2016), which can also influence tropospheric variability, including the modulation of the amplitude of the Madden-Julian Oscillation (Anstey and Shepherd, 2014, Yoo and Son 2016, Son et al. 2017), and high latitude features in the lower stratosphere like sudden stratospheric warmings.

Higher horizontal and vertical resolution have increased the computational cost and changed the simulations substantially. Parameterization complexity, and more predicted and transported variables (aerosol species, prognostic snow and rain, ozone) also contribute to computational cost increases. The increases in operation count required to perform a simulation motivated the development of some alternate strategies for tuning and evaluation of the HR model configuration. It was not feasible to "brute force" tune the

high-resolution configurations, so short-term hindcasts and perturbed parameter ensemble simulations provided guidance on tuning, and regionally refined simulations were also used to explore and evaluate parameterization sensitivity to high resolution. Details of the impact of specific parameter changes are discussed in more detail in Section 2.2 of Xie et al (2018), and background information and citations discussing the procedures used for the high resolution tuning are described in section 3.

The Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package [COSP; Bodas-Salcedo et al. 2011] version 1.4 was also implemented in EAM to facilitate the comparison of modeled clouds with satellite observations by accounting for observational limitations and observing features of the instruments (Zhang et al. 2019). COSP was further enhanced with an aerosol lidar simulator (P. Ma et al, 2018) that facilitates the evaluation of aerosol cloud interactions and comparison with satellite retrievals of aerosol properties; that study showed that simulator output in combination with COSP cloud diagnostics could be used to better understand and explain discrepancies between estimates of cloud susceptibility to aerosols from satellite products, and estimates typically produced within Earth System Models. A ground-based cloud radar simulator developed from the DOE Atmospheric Radiation Measurement (ARM) program (Y. Zhang et al. 2018) was also integrated into COSP to utilize detailed cloud observations at ARM research sites.

2 Model Description

2.1 Dynamical Formulation and its connection to transport and physical processes

EAMv1 uses a continuous Galerkin spectral finite-element method to solve the primitive equations on a cubed-sphere grid (Taylor et al. 1997; Thomas and Loft 2005; Taylor and Fournier 2010; Dennis et al. 2012). Tracer transport is handled using a variant of the semi-Lagrangian vertical coordinate system from Lin (2004) adapted for the cubed sphere, designed to locally conserve air and trace constituent mass as well as moist total energy (Taylor 2011). This treatment has attractive computational scaling properties and has been shown to scale nearly linearly to thousands of cores (Dennis et al. 2012; Evans et al. 2013). The dynamics, tracer transport, and physical parameterizations are operator split, and sub-cycled in time (see Figure 2. in K. Zhang et al. (2018)). Tendencies for each process (dynamics, physics, and constituent transport) are calculated separately and used to update the model state. The sub-cycling frequency, and strategy for communication of information between updates to the model state by physical parameterizations, trace constituent transport and fluid dynamics by the dynamical core is resolution dependent. These choices make the discretization formally fourth order accurate in the horizontal dimension, and first order accurate in time. A hybrid physics-dynamics coupling that uses different methods for fluid dynamics and tracer transport is used in EAMv1 to ensure water conservation (K. Zhang et al., 2018).

Because the horizontal discretization is localized on individual elements, variable resolution can be introduced through refined meshes. This setup allows variable resolution grids to maintain the key conservation and scalability aspects that makes the dynamical core a desirable model choice for climate simulations. A detailed evaluation of the impacts

of variable resolution on the climate for E3SMv0 is described in Roesler et al. (2018), and for EAMv1 in Tang et al. (2019). Two examples of a horizontal discretization in EAM can be seen in Figure 1, which shows the basic "cubed sphere" grid with an embedded regional refinement over the Continental United States Region (CONUS), and over the Tropical Western Pacific (TWP). Two additional RRM configurations positioned over the ARM Eastern North Atlantic (ENA) site have also been introduced into E3SMv1.

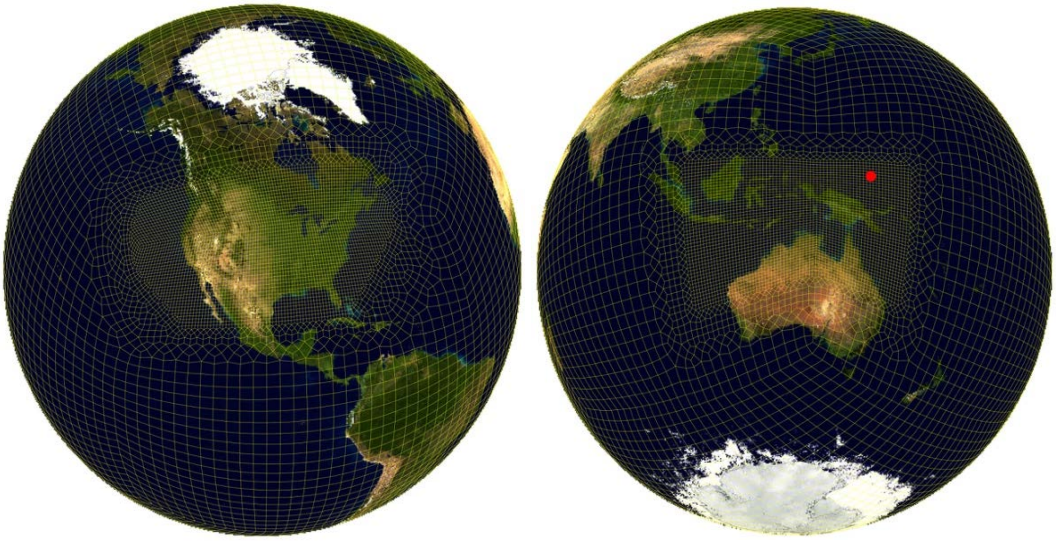

Figure 1: The regionally refined grid configuration for the atmosphere and land over the Continental United States (CONUS) and Tropical West Pacific in EAMv1. The high resolution area has an effective spatial resolution between grid points of 25 km, while the lower resolution area has an effective spatial resolution between grid points of 110 km. A transition region bridges these resolutions. The red dot indicates the location of DOE ARM site in the tropics.

2.2 Physical Processes

This section provides a brief description of the parameterizations used in EAMv1, with citations to papers with more detail, and a discussion of the tunable parameters used in calibrating the model from climate fidelity. Table A.1 of Appendix A lists the values for the tunable parameters used in the calibration process.

2.2.1 Turbulence/Shallow Convection/and Stratiform Clouds

Turbulence, shallow cumulus and stratocumulus clouds are parameterized with a single equation set describing the evolution of a set of probability density functions (PDFs) of fields, using the Cloud Layers Unified By Binormals (CLUBB) parameterization (Golaz et al. 2002; Larson and Golaz 2005; Larson 2017). CLUBB is based on a higher-order closure to the prognostic equations for higher order moments, and it prognoses nine subgrid higher-order moments: the vertical turbulent fluxes of total water (vapor plus cloud liquid) and liquid water potential temperature ($\theta_{\square}$); the variances of total water, $\theta_{\square}$, vertical velocity, zonal velocity, and meridional velocity; the covariance of total water and $\theta_{\square}$; and the third-

order moment of vertical velocity. CLUBB does not yet incorporate a treatment for ice. Within these higher-order equations, turbulent advection and buoyancy terms are closed by analytic integration over the subgrid PDF. The PDF is assumed to have a normal mixture (i.e., double Gaussian) shape using the Analytic Double Gaussian 1 shape of Larson et al. (2002). The PDF in each grid box is chosen to be consistent with the subgrid moments predicted in that grid box. The scalar dissipation and perturbation pressure terms are closed using standard closures from the turbulence literature. Momentum fluxes are parameterized by downgradient eddy diffusion. EAMv1 uses CLUBB revision 7416 of the CLUBB svn repository

2.2.2 Deep Convection

Deep Convection is based upon the formulation originally described in Zhang and McFarlane (1995, hereafter ZM), with modifications by Neale et al (2008), and Richter and Rasch (2008). The parameterization employs a bulk updraft and downdraft intended to represent an ensemble of updrafts (undergoing entrainment, detrainment, condensation and precipitation) and downdrafts (undergoing entrainment and evaporation). Cloud-base mass flux is determined by a closure designed to reduce Convective Available Potential Energy (CAPE) at a prescribed (tunable) timescale. Updraft condensate mass is generated via a saturation adjustment (to the liquid saturation vapor pressure along a moist adiabat) followed by conversion to precipitation with a (tunable) adjustment timescale proportionate to the convective updraft mass flux (differing timescales can be used over land and ocean). Similarly, a tunable downdraft coefficient is used to control the mass flux at the downdraft top and the evaporation rate of the rainfall driving the downdraft. As with CLUBB, ice thermodynamics are generally neglected, but when liquid condensate is detrained from the updraft it is partitioned into liquid and ice over a (tunable) temperature range, and the energy associated with a phase change is applied locally to the temperature tendency to conserve energy. This simple treatment of the thermodynamics of water is not optimal and alternate (more realistic) formulations are planned for future versions of EAM.

Convection parameterizations can be quite sensitive to vertical and horizontal resolution, and significant modifications were made to accommodate EAM's resolution changes. A comprehensive evaluation of the ZM parameterization sensitivity to resolution in EAM is provided in Xie et al (2018), and only a few of the implications are mentioned here. CAM5 used a 100-m (200-m) thick surface layer (and layer above, respectively). EAMv1 has a much thinner (20-m) surface layer and much higher vertical resolution, particularly in the boundary layer supporting much stronger gradients in temperature and water vapor, with implications to many aspects of the convective parameterization. Heat and moisture fluxes are more easily confined to the surface and boundary layer, affecting the buoyancy and lifting condensation level (LCL) of parcels driving the convective updrafts. Additional model layers also allow more variation in layer stability, relative humidity, and buoyancy (i.e. updraft plumes may penetrate more layers before reaching their level of neutral buoyancy). Because water vapor is generally more confined to the boundary layer, the free troposphere can also be drier, in the absence of processes that mix it to higher altitudes. To avoid a significant sensitivity in ZM to surface layer properties the lowest possible launch level was changed from the surface layer to a (tunable) layer at

about 100m. CAM5 ZM formulations also allowed an air parcel to traverse up to five local regions of stability (negative buoyancy) in the calculation of CAPE and convection top before the parcel buoyancy calculation determining the level of neutral buoyancy (LNB) was terminated. Because the parameterization used no quantitative information about the vertical extent or strength of stable regions (e.g. "convective inhibition"), this made the LNB calculation extremely resolution dependent. EAMv1 reduced this dependence by terminating the CAPE calculation when the parcel encounters a single negatively buoyant layer in the plume buoyancy calculation. Impacts of these changes are discussed in Xie et al. (2018). CAPE, and cloud top (LNB) calculations also depend strongly on the environment air properties entrained in the updraft (following Neale et al, 2008), and so the convection top is quite sensitive to free troposphere humidity. Since the higher resolution can support stronger gradients (and a drier free troposphere), convection in a high vertical resolution model is typically shallower than found in a coarser resolution simulation, influencing many meteorological features (radiative heating, altitude of detraining heat and water substances) and the moisture modes that play a significant role in equatorial wave variability. The (tunable) entrainment parameter "dmpdz" was reduced to control the depth of convection. An optional parameterization of convective gustiness was also implemented (off by default for EAMv1) that is designed to represent the increase in surface winds produced in the presence of evaporating precipitation. Implications of this variation of the convection parameterization are discussed in Harrop et al. (2018).

Convective cloud fraction is assumed proportional to the log of the updraft mass flux following Hack et al (1993). Detrained convective condensate is partitioned into liquid and ice. Detraining convective condensate in the liquid phase inherits the particle size estimate from any existing stratiform cloud at the same location, and the ice particle size is specified as a tunable parameter (discussed again below).

2.2.3 Aerosol and Cloud Microphysics

Aerosol microphysics and interactions with stratiform clouds are treated with an updated and improved version of the four-mode version of the Modal Aerosol Module (MAM4, Liu et al., 2016) that predicts the number and mass concentrations of major aerosol species [sulfate, black carbon (BC), primary organic matter (POM), marine organic aerosol (MOA), secondary organic aerosol (SOA), mineral dust, and sea spray] in one coarse and three fine-particle aerosol modes. A new representation of MOA was introduced (Burrows et al., 2014; Burrows et al., 2018) as an important step towards linking ocean biogeochemistry processes with the ocean surface chemistry that determines sea spray aerosol formation. The simplified treatment of SOA-precursor sources in MAM's previous single-lumped-species treatment of SOA was revised by adapting SOA production results from Shrivastava et. al (2015), in which an explicit treatment of multigenerational gas-phase chemistry of SOA precursor gases and particle-phase transformation of SOA gave better agreement with global organic aerosol measurements.

Aging of the primary carbon uses the 8-monolayer (slow ageing) criterion of Liu et al. (2016). A parallel time-split treatment of H2SO4 production by gas-phase chemistry and loss by condensation was implemented that provides more accurate H_2SO_4 concentrations to the new particle formation process and increases small particle (defined as having a diameter < 100 nm) number concentrations, in better agreement with

observations. Most of the changes to aerosol wet removal described in Wang et al. (2013) were also added to EAMv1, notably including a unified treatment for convective transport and scavenging of aerosols (with secondary activation in convective updrafts above cloud base) that more realistically represents wet removal and vertical transport of aerosols by convective clouds. These changes make important improvements to the aerosol concentrations in regions remote from major sources (i.e., the upper tropospheric and polar regions). The resuspension of aerosol matter from evaporating raindrops was revised to return particles to the coarse mode (as a few large particles), rather than to the originating modes (mostly the accumulation mode), and this causes modest reductions to aerosol mass and Cloud Condensation Nuclei (CCN) concentrations.

EAMv1 uses the "Morrison and Gettelman version 2" (MG2) two-moment bulk microphysics parameterization for stratiform clouds described in Gettelman et al. (2015). Eight prognostic variables (representing the mass and number of small and large liquid and ice particles) are used to describe the evolution of condensate. Table A.1 of Appendix A lists the values for MG2 and other tunable parameters. It is noteworthy that the default rate at which liquid is transferred from liquid to ice via the Wegener-Bergeron-Findeison (WBF) process, inherited from a developmental version of the MG2, was set unrealistically low. When coupled with CLUBB and MAM4, MG2 provides a more consistent and coherent set of aerosol-cloud interactions in all stratiform and shallow convective clouds than CAM5, which neglected aerosol interactions in both shallow and deep convective clouds. Aerosol–cloud interactions are still neglected in deep convective clouds in EAMv1, because of the very simple microphysics assumed by the ZM parameterization. Stratiform cloud liquid and ice nucleation processes use a characteristic in-cloud updraft velocity that is diagnosed as a function of the Turbulent Kinetic Energy (TKE) predicted by CLUBB. CLUBB's TKE (and corresponding cloud updraft velocities) does not currently depend directly on radiative cooling rates near cloud top, and often produce weak estimated subgrid vertical velocities so a (tunable) lower limit of 0.2 m/s was chosen for the subgrid velocity used to activate liquid and ice clouds. The lower bound for IN calculation is triggered very frequently, but the impact on ice cloud formation is small with ~0.3 W/m2 change in global annual mean long and shortwave cloud radiative effects that counter each other to produce a near zero impact on the global annual mean TOA net radiative flux.

Liquid cloud drop activation follows Abdul-Razzak and Ghan (2000). EAMv1 uses a Classical-Nucleation-Theory (CNT)-based ice nucleation parameterization for the heterogeneous ice formation in mixed phase clouds, which depends on both interstitial and cloud-borne BC and dust aerosols (Wang et al., 2014). The threshold size for auto-conversion of ice crystals to snow particles was changed from the constant value used in CAM5 to a function of temperature to improve agreement with observed ice distributions.

EAMv1 uses a variant of the popular autoconversion parameterization developed by Khairoutdinov and Kogan (2000, hereafter KK2000) to represent the self-collection of small liquid drops to form precipitation size particles, but as noted by Wood (2005), and Kogan (2013), the three parameters developed for that parameterization (and usually treated as fixed inviolate constants) are actually subject to significant uncertainty, depending upon cloud regime. Those three studies compared very accurate calculations of the precipitation formation process in differing cloud regimes to demonstrate that there is

actually a big range of free parameters in the KK2000 formulation that provide acceptable fits to more accurate calculations, so the parameter values usually adopted in models are actually subject to considerable uncertainty, and the parameter choices have a very strong impact on aerosol indirect radiative effects (e.g., Wang et al., 2012). As shown in Wood (2005) and Kogan (2013) the fitting process used as central values for autoconversion (and accretion) were designed to produce the good agreement in fully developed clouds with substantial condensate mass. Other choices within the range of uncertainty identified in those studies also produce reasonable cloud properties that are much less sensitive to cloud formation and precipitation onset in low aerosol conditions. Our choices for those parameter values reduce radiative biases in regions with low cloud water and drop number in regimes other than stratocumulus regions (e.g. trade cumuli), produce less precipitation through autoconversion than the standard KK2000 treatment, and are very consistent with K2013 and Beheng 1994 rates in low aerosol (cloud droplet) regimes. Autoconversion rates determined by the EAMv1 parameter choices (documented in Appendix A.1) compared to the standard KK2000 values (and Kogan 2013 and Beheng 1994) are shown in figure 2 over a range of cloud droplet numbers and condensate loadings. Results remain very similar to KK2000 for higher droplet number and condensate values. Although autoconversion is somewhat higher at very high (300+ cm^{-3}) Nc (fig 2d-f), this is a situation where precipitation is low anyway and autoconversion is most important at low aerosol concentrations. If default values are used instead of EAMv1 settings the amplitude of the net Cloud Radiative response from changing aerosol emissions between pre-industrial and present-day values changes by about 0.3W/m^2 in amplitude. More details of the model sensitivity to autoconversion formulation will be described elsewhere in a manuscript in preparation.

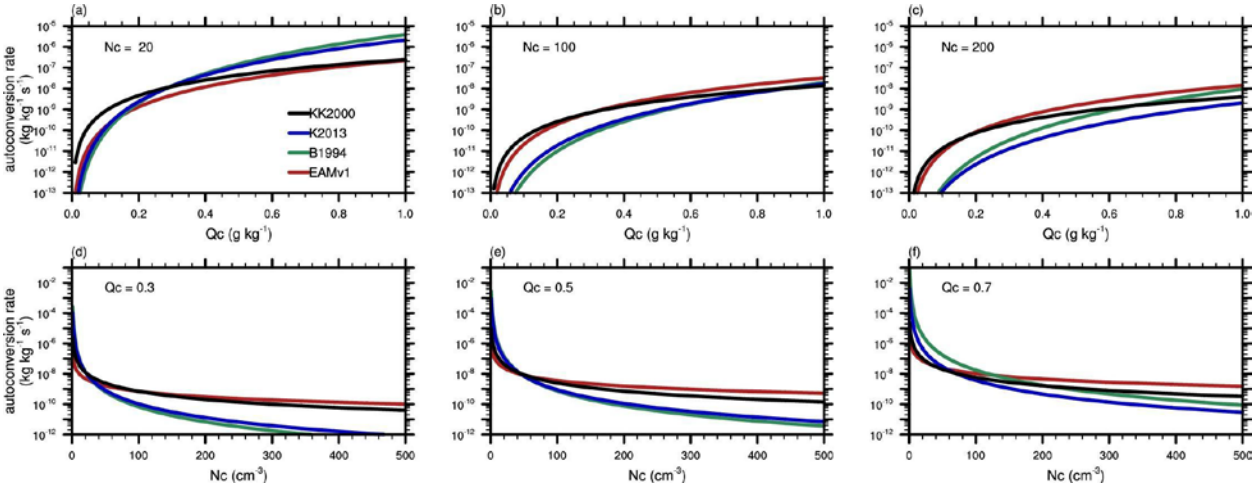

Figure 2*:* Autoconversion rate (**a-c**) as a function of cloud water mixing ratio (Qc) at cloud droplet number concentration (Nc) = 20, 100, and 200 cm^{-3} and (**d-f**) as a function of Nc at Qc = 0.3, 0.5, and 0.7 g kg^{-1}.

Aerosols and light-absorbing particles (e.g., BC and dust) can be deposited on snow and ice surfaces over land and sea-ice, and these treatments were harmonized to treat both external mixing and internal mixing (within-hydrometeor) of BC and snow grains following Flanner et al. (2013). This harmonization shows similar stronger in-snow BC radiative

forcing to the Flanner study compared to the previous external mixing treatment (e.g. Qian et al., 2014).

Hygroscopicity characteristics are specified for soluble species to make particle size depend upon relative humidity. Aerosol mass and number concentrations predicted by MAM4 are ultimately merged with a historical inventory of volcanic aerosols with an assumed mean radius in the stratosphere. Aerosol optical properties for each aerosol mode are then calculated following the treatment described by Ghan and Zaveri (2007), where optical properties of aerosols are combined using a "homogeneous mixing assumption" by volume to produce estimates of single scattering albedo, extinction, and backscatter by wavelength and then provided to the radiative transfer parameterization described below.

2.2.4 Chemistry

A simple chemical mechanism calculates stratospheric ozone concentrations following Hsu & Prather (2009, LINOZ2). Linearized production and loss coefficients characterizing sensitivity to local ozone concentration, local temperature, and the ozone column above, are used to calculate stratospheric production and loss terms for ozone for each decade of the 1850-2000 time period between 12 and 58 km, monthly-mean zonal-mean estimates. A simple parameterization based on Cariolle et al. (1990) is included to represent ozone depletion in the presence of polar stratospheric clouds (PSCs) at cold temperatures in the stratosphere. The rate of exponential decay is a function of the prescribed chlorine loading, so ozone holes will only form when chlorine loadings are high, and the temperature is cold enough to form PSCs. Chlorine loading can be changed to produce or suppress ozone hole formation over Antarctic. These source sink terms are then used to update an ozone-like trace constituent in the stratosphere, which is advected by the EAMv1 dynamics. Advected stratospheric ozone mixing ratios are then merged at the tropopause with prescribed monthly mean tropospheric ozone archived from more comprehensive chemical simulations performed with the MOZART chemical transport model as discussed in Neale et al (2012). This merged ozone distribution is then used in the radiative transfer calculation, and for oxidation of aerosol precursors, as described above.

2.2.5 Radiative Transfer

EAMv1 uses the Rapid Radiative Transfer Model for general circulation models (RRTMG) (Iacono et al., 2008; Mlawer et al., 1997). RRTMG uses a modified correlated-k method to calculate radiative fluxes and heating rates in the clear sky and for condensed phase species. Extinction optical depth, single scattering albedo and asymmetry properties are specified for 14 short-wave bands (extending from 0.2 μm to 12.2 μm), and a mass specific absorption is specified for 16 long-wave bands (extending from 3.1 μm to 1000 μm) representing Rayleigh scattering and sources of extinction for H2O, O3, CO2, O2, CH4, N2O. The longwave calculation uses molecular sources of absorption for the same species, plus CFC-11 (representing multiple CFC species) and CFC-12. Solar irradiance is now specified for the short-wave bands from the Lean dataset (Wang et al., 2005).

A single set of in-cloud properties, and a single cloud fraction for each model cell are calculated and provided to RRTMG. These composite cloud properties are derived as a complicated function of individual estimates of convective and stratiform (liquid and ice) cloud fraction and condensate properties (mass and number mixing ratios), and

assumptions of ice crystal shape using the following recipe. The MG2 stratiform cloud microphysics and CLUBB higher order turbulence parameterization provide explicit values for condensate (liquid and ice [both small and large (snow)] mass and number, and an estimate of stratiform (liquid and ice) cloud fractions. As previously described, convective cloud fraction, condensate partitioned into liquid and ice phase, and convectively detrained ice crystal size are treated as tunable parameters and are chosen to produce reasonable estimates of cloud radiative forcing and condensate mixing ratios. These estimates of condensate mass, size, and fraction are merged into a single set of characteristics prior to the calculation of extinction optical depth, single scattering albedo and asymmetry parameter. The cloud fraction compositing is described in Park et al (2014). Vertical cloud overlap follows the algorithm for CAM5 as described in Neale et al, 2012. Maximum vertical overlap is assumed within each of the 3 regimes representing low ($p > 700$ hPa), middle (400 hPa $< p < 700$ hPa), and high ($p < 400$ hPa) clouds, and a random vertical overlap is assumed between these 3 regimes. Liquid-cloud drop optics are calculated following Wiscombe [1996] and ice-cloud optics are calculated following Mitchell [2002]. Ice-cloud size optics are extended to allow for radiatively active falling snow. A subcolumn decomposition of cloud overlap and fraction generates a set of sub-columns in which cloud fraction is either 1 or 0 in each layer. These composited cloud and aerosol properties for each subcolumn are in turn passed to RRTMG along with concentrations of radiatively active gases. The Monte-Carlo Independent Column Approximation (McICA, Pincus and Morcrette [2003]) that represents sub-grid scale cloud variability is used to calculate the heating rates and fluxes in each subcolumn for specific spectral bands; averaging each sub-column's radiative heating rate over time and subcolumn provides an estimate of grid-mean radiative heating rate.

2.2.6 Gravity Waves

The gravity wave (GW) parameterization in EAMv1 is the same as that described in Richter et al (2010). Three sources of gravity waves are considered – orographic (mountain waves as in McFarlane (1987), convective (Beres et al. 2004), and frontal (Charon and Manzini 2002). The geographical distribution of the frontal GW source is based on a frontogenesis function (Hoskins 1982). When the frontogenesis function is found to exceed a specified threshold at 600 hPa, a spectrum of frontal GWs are launched at 500 hPa with a pre-specified momentum flux amplitude. The spectrum is aligned parallel to the wind at the source level, which is typically the steering level of fronts. 40 individual waves with phase speeds from U-100 to U+100 m/s are launched. The spectrum is Gaussian in shape and centered on the source level wind. Studies in models with tops that extend through the mesosphere indicate that parameterized frontal GWs are critical in reversing high-latitude upper-stratospheric and mesospheric jets and are ultimately responsible for producing the mesopause temperature reversal in models. The top boundary condition for GWs is set so the momentum flux goes out at the model top. While it is theoretically more correct to deposit GW momentum flux within the model domain, experience in the Whole Atmosphere Community Climate Model (a CAM variant with a higher model top), produced several undesirable issues (too much momentum deposition and heating). Hence after deliberation, we chose to accept a (very) small momentum flux source/sink from the column, rather than support spurious tendencies in the stratosphere.

Convective GWs in EAM are specified according to the source spectrum described in Beres et al. 2004. GWs are assumed to be produced by oscillating convective motion with the deep convective cells and are launched when deep convection produced by the ZM parameterization is detected. The width and shape of the GW spectrum as a function of phase speed is determined by the depth of deep-convective heating, which is estimated from the heating profile from the ZM parameterization. The mean wind within the heating region is also taken into account, creating an asymmetry between the eastward and westward propagating waves. The amplitude of GW momentum fluxes is proportional to the square of the maximum heating in ZM scaled by a factor that converts the convective grid-box averaged heating rate to a heating rate representative of heating within a convective cell. The "obstacle effect" of convection is not considered. Orographic GWs are launched from the layer nearest the surface in E3SM. Obstacle heights are determined from RMS deviations of unresolved topography. The procedure to generate the resolved and unresolved topography is described in Lauritzen et al (2015). An ultra high resolution (30') source file is first downscaled to the appropriate (ne30 or ne120) cubed-sphere grid. 16 iterations of the discrete HOMME 2nd order Laplacian operator are then applied by running the dynamical core in a special mode. The unresolved topography is computed to be consistent with the resolved topography. The spectral element core uses a smoother topography than CAM-FV because CAM-SE doesn't have a good pressure gradient fixer. This is quantified in the energy spectra in the Lauritzen et al (2015) paper. Average flow properties are estimated in a "source layer" extending from the model surface to an obstacle peak height based on the RMS deviations of unresolved topography. An isotropic orographic source is assumed, i.e., wave amplitudes are independent of source-layer wind direction, but orographic GW wave orientations are assumed to be normal to the source-layer wind. Near-surface flow nonlinearities arising from flow splitting (blocking) of downslope winds is not represented (e.g. Lott and Miller (1997); Scinocca and McFarlane (2004)). Frontal GW drag plays a significant role in the momentum budget of EAMv1 above 40 km in the NH, and above 50 km in the SH. Convective GWs in EAM are responsible for producing QBO-like fluctuations of the equatorial zonal wind, but the amplitude and period were not optimized in the current EAM and the oscillation is currently too strong. Better tuning with much improved QBOs have been demonstrated in experimental versions of EAMv1 and are planned for implementation in the next model version.

3. Model configurations, tuning, and computational performance

EAMv1 is computationally more expensive than its predecessors. In additional to a significant increase in parameterization complexity, it employs 40 trace constituents (water vapor plus mass and number for 4 categories of condensed water, five gaseous aerosol precursors and oxidants (including ozone), number for the four aerosol modes, and (mass fractions of aerosol components in each mode), compared to the 25 used in CAM5, increasing the cost of the trace constituent advection by a factor of ~1.6. Since the atmospheric model operation count increases roughly in (linear) proportion to the number of model layers, the model cost is thus roughly doubled over previous costs because the vertical resolution has more than doubled. The increase in vertical resolution also has consequences to model numerical stability. In principal, the upper limit on timestep for advection and many physical processes involving transport should decrease in proportion

to the model layer thickness, and since the thinnest layers in EAMv1 are approximately five times thinner than those of CAM5 it might be expected that the model timestep might need to be decreased by a factor of 5. In practice CLUBB and MG2 are both much more complex than the CAM5 parameterizations they replace, and both components are already sub-stepped together to avoid numerical instabilities and increase the numerical accuracy. Sub-stepping adds to the computational cost, but the frequency of communication between physics and dynamics modules required to maintain stability is relatively insensitive to the changes in vertical resolution (although timestep length does have an impact on the model climate). HR model configurations support stronger velocities generating higher Courant numbers and more rapid temporal changes, and therefore use a coupling frequency of 15 minutes, while the LR configurations exchange information at 30 minute intervals.

Approximate costs for a few useful configurations of EAMv1 and its predecessor (CAM5/EAMv0) are shown as a function of throughput and cost on two DOE computers (Cori-KNL and Edison) in Table 1a and b. The configurations reported on are by no means optimal, but they are representative, and provide some insight into model performance and cost. Configurations were generally selected to produce throughput around 1-10 Simulated Years per Day (SYPD) without requiring a large core count (higher throughput is easily attained with more nodes, and the model still scales well to much larger node/processor counts). The cost is reported after assuming all cores on each node are charged for (the typical strategy used at most supercomputer centers. Cori-KNL configurations actually utilize only half the available cores per node, so costs would be lower if the charge were by core-hour actually used). On Cori, the 1degree (ne30) configuration of EAMv1 is approximately 3 times more expensive than the spectral element ne30 configuration of CAM5 while the difference is larger on Edison. The factor of 4 increase in cost on Edison is closer to the anticipated difference explained by vertical resolution, tracer number and parameterization complexity. The global quarter degree (ne120) model configuration is approximately 25 times more expensive than the 1degree (ne30) version on Cori, and 40 times more expensive on Edison (associated with a factor of 16 increase in horizontal elements, a reduction in dynamics timestep, and a factor of two increase in calling frequency of the physics parameterizations). Using a regionally refined CONUS configuration reduces the cost by a factor of 5-10 compared to the globally uniform HR model, and is ~4 times more expensive than the uniform LR configuration on both machines.

Table 1a: Information related to model resolution and cost on Cori (KNL nodes). Note that CAM5 by default uses a finite volume dynamical core and the numbers reported here are for the spectral element dynamical core, and the discussion regarding the cost calculation on Cori.

Resolution identifier	Horizontal Resolution	Vertical Resolution	Number of elements	Throughput: Sim Years Per Day (# Cori-KNL nodes)	Cost in core-hours per sim year (68 cores/node)
CAM5 Ne30	1^o	30	5400	9 SYPD (40N)	7,000
EAM Ne30	1^o	72	5400	6 SYPD (81N)	22,000
EAM Ne120	$1/4^o$	72	86400	2 SYPD (675N)	551,000
EAM CONUS	$1/4^o$ to 1^o	72	9905	1.7 SYPD (88N)	84,000

Table 1b: Information related to model resolution and cost on Edison.

Resolution identifier	Horizontal Resolution	Vertical Resolution	Number of elements	Throughput: Sim Years Per Day (# Edison nodes)	Cost in core-hours per sim year (24 core/node)
CAM5 Ne30	1 ^{o}	30	5400	41 SYPD (113N)	1,600
EAM Ne30	1^{o}	72	5400	10 SYPD (113N)	6,300
EAM Ne120	1/4^{o}	72	86400	0.5 SYPD (200N)	265,000
EAM CONUS	1/4^{o} to 1^{o}	72	9905	1.3 SYPD (113N)	25,000

As previously mentioned, the change in vertical resolution introduced profound changes in the model climate because some atmospheric features are better resolved, and some EAMv1 parameterizations remain sensitive to both horizontal and vertical resolution, so re-tuning was needed to produce good model fidelity at high and low horizontal resolution (for details, see Xie et al, 2018). LR tuning was done employing traditional approaches (2-10 year simulations and subjective evaluations of dozens of climate diagnostics). Since brute force strategies using many multi-year simulations were not feasible at high resolution, other strategies were employed. Xie et al. (2012) and H. Ma et al. (2012) showed that model biases related to fast physical processes in short-term hindcasts resemble those from long-term climate simulations, so two frameworks using short simulations were used as an integral part of model development. A hindcasting methodology identified with Transpose-AMIP and Cloud-Associated Parameterizations Testbed (CAPT) (Philips et al. 2004, Williams et al. 2013) was used to assess candidate physical parameterizations and tune the computationally expensive high-resolution EAM configurations, following protocols described in previous studies (Xie et al. 2004; Boyle and Klein 2010; Liu et al. 2011; Lin et al. 2012; van Weverberg et al. 2015, Qian et al., 2015; H. Ma et al. 2018). Short, few-day simulations are also capable of revealing parametric sensitivities, and Wan et al. (2014) and Qian et al. (2015) developed a protocol for using Perturbed Parameters Ensembles (PPEs) of 3-day simulations to reveal model sensitivities to parameter changes that provide insight similar to that found in longer, multi-year simulations with a substantial saving in computing resources and dramatic reduction of simulation turnaround time. Qian et al. (2018) described the use of this strategy during EAM model development to identify simulation sensitivity to 18 uncertain parameters related to convection, turbulence, cloud microphysics, and orographic gravity wave drag. The resolution dependent tuning parameters are listed in Appendix A. Although it would be desirable to have modified the model to support parameters that could adapt to resolution as the grid varies, the current code does not support that capability. Our initial choice was to select HR tunings over the whole globe in the regionally refined CONUS model configurations.

Although the hindcasting methodology and PPEs were used to understand model sensitivity to changes in vertical resolution, physics parameterizations, and model parameters at various model development stages, our final EAMv1 tuning strategy was simply to produce simulations with reasonable model fidelity similar to or better than that of CMIP5 models based on comparison with recent (about the last 3 decades) observations. We did not strive for the best simulation possible, but rather to produce a good simulation

for a range of resolutions and forcings --- indeed with subsequent efforts after the release of E3SMv1, we are now aware that better simulations with the same model are possible with other tuning choices, which will be documented and considered for adoption in version 2. For version 1, a lot of emphasis was placed on producing reasonable radiative signatures in monthly mean radiative fluxes (driven by state variables), hydrological fields (precipitation, water vapor, clouds), and dominant modes of variability (including the seasonal cycle, ENSO relevant signatures, the Madden Julian Oscillation), etc. Lack of space precludes a comprehensive description of our metrics and evaluation strategies, but some important features are described in the next section.

4. Model Evaluation

Coupled simulations are evaluated in Golaz et al (2019) with EAMv1 connected to interactive ocean and sea ice components, and model configurations using prescribed SST and sea ice extent are evaluated in this section. Since Xie et al. (2018) and Zhang et al (2019) performed an in-depth analysis of clouds and precipitation features, less emphasis is placed on evaluating those fields here. Prescribed mid-monthly values of the fields used as input (boundary conditions and external forcing specifications) for the model were "diddled" as outlined in Taylor et al (2000) to assure that monthly averages of fields interpolated between mid-month values match original prescribed monthly averages. The input fields specify values of:

- Land-Use (LU) and Land-Cover (LC)
- Greenhouse gas concentrations: (CO_2, CH_4, N_2O, CFC-12, CFC-11, other halogenated compounds added as radiatively equivalent amounts of CFC-11) follow CMIP5 specification.
- Aerosols: Most aerosols and pre-cursors used CMIP5 or CMIP6 emissions (http://www.globalchange.umd.edu/ceds/ceds-cmip6-data/, Hoesly et al. (2018)) with the following exceptions. DiMethylSulfide (DMS) was based on Elliott (2009) and Wang, et al (2015). Ocean macromolecule emissions follow Burrows et al., (2014;2018); Ogunro et al. (2015) and Wang, et al (2015). Secondary organic precursors included a vertical distribution of emissions following Shrivastava et al (2015). Open fire emissions were added for SO_2, POM, and BC (van Marle et al., 2017). Aerosol oxidants (O_3, OH, NO_3, HO_2) were specified as discussed in Neale et al (2012). Some simulations also used estimates of historical stratospheric distributions of volcanic aerosol. Finally, dust and sea salt aerosol emission factors were tuned to constrain the global dust and total AOD (see Table A.1).
- Insolation
- Prescribed SST and sea-ice datasets were constructed as a blended product, using the global HadISST OI dataset prior to 1981 and the Smith/Reynolds EOF dataset post-1981 as described in Hurrell et al (2008)

Sea Ice surface temperature, ice thickness and snow depth on ice are prognostic and respond to the forcing.

Three main model configurations were analyzed: Most of the evaluation uses a configuration (identified as FAMIP) with SSTs, sea-ice concentrations, emissions and oxidant distributions that vary month by month and year by year. The second important

configuration uses a climatology of monthly estimates of the surface temperatures and ice concentrations characteristic of a temporal (20 year) window surrounding year 2000 with appropriate solar insolation, GHG and oxidant concentrations, and aerosol emissions characteristic of Present Day (PD) conditions (labelled F2000). The last configuration composited a similar window of years surrounding 1850 to provide climatological conditions relevant to Pre-Industrial (PI) times (SST and sea-ice concentrations used HADISST data as described in Hurrell et al (2008)). 30 year averages of solar insolation were used for the F1850 and F2000 simulations to avoid choosing a specific phase of the solar cycle (1834-1867 and 1976-2007, respectively) and volcanic aerosols were neglected. The F1850 and F2000 model configurations were used to evaluate the model response to anthropogenic forcing agents in the presence of SST change. The FAMIP simulations provide information about the model sensitivity to interannual patterned changes in surface temperature like those seen with ENSO. Three additional model configurations were analyzed to expose the model response to all forcing agents (often termed "adjusted forcing" or AF, F2000LR/AF), and to anthropogenic aerosols (e.g. the Aerosol Effective Radiative Forcing from Aerosols (labelled F2000LR/Aero)), and the modeled response to uniform surface temperature changes, including cloud feedbacks, (suffixed with +4K label). All LR model configurations used default parameter settings specified in Appendix Table A.1. Similarly, HR and CONUS settings used the common high-resolution parameter settings of Table A.1. These model configurations are summarized in Table 2.

Table 2: Description of model simulations performed for evaluation:

Simulation Name	Resolution, Analysis period	Description
F1850LR	100km, Years 2-11	Year 1850 SST, GHG, LU, LC, Aerosol sources, no volcanoes
F2000LR	100km, Years 2-11	Year 2000 SST. GHG, LU, LC, Aerosol sources, no volcanoes
F1850LR+4K	100km, Years 2-11	As in F1850LR, SST + 4K to assess cloud feedback
F2000LR/AF	100km, Years 2-11	As in F2000LR but SST, Sea-ice, Solar Cycle replaced by 1850 values
F2000LR/Aero	100km, Years 2-11	As in F2000LR but Aerosol emissions replaced by 1850 values
FAMIPLR	100km, 1980-2005	Time varying SST, aerosol sources, volcanoes, LU/LCC, GHG
FAMIPLR/AllF	100km, 1980-2005	As in FAMIPLR except aerosol sources, volcanoes, LU/LCC, and GHG fixed at 1850
FAMIPLR/AeroF	100km, 1980-2005	As in FAMIPLR except aerosol sources fixed at 1850
CAM5AMIP	100km, 1980-2005	CAM5 Control for comparison
F2000HR	25km, Years 2-11.	As in F2000LR except high resolution
F1850CONUS	100km + 25km, Years 2-5	CONUS simulation, Free-running
F2000CONUS	100km + 25km, Years 2-5	CONUS simulation, Free-running (nudged and CAPT special features)

Our evaluation uses the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2, Gelaro et al, 2017), the Japanese Meteorological Agency's 25 year Reanalysis (JRA-25, Onogi et al, 2007), and the ECMWF Reanalysis Interim (ERA-I, Dee et al (2011)) baselines for comparison with many state variables (temperatures, winds, etc). Differences between model and observational estimates are generally large enough that the choice of analysis product is not critical, but we occasionally note some features where the choice makes a difference. MERRA-2 has made a strong effort in producing accurate products in the lower stratosphere, an area where EAMv1 and CAM5 simulations differ substantially, so we use that product when analysis extends into the stratosphere. Other observational estimates for top of atmosphere and surface fluxes based on satellite retrievals are noted as they are used.

4.1 Global Average Characteristics

A quick assessment of some of the model characteristics relative to many models that participated in the CMIP5 is provided in Figure 3 in the form of a heat map following Gleckler et al (2008, 2016) for global uncentered root-mean-square errors relative to observations. Rows indicate performance for a variety of meteorological fields described in the figure caption, and columns show various models participating in CMIP5, with EAM LR and HR and CAM5 models highlighted near the left side of the figure. Each square region is divided into quartiles denoting the performance in a particular season. Orange/Red colors indicate worse performance and Blue Colors indicate better performance compared to the median model. Both EAM model configurations generally rank quite highly compared to the ensemble of CMIP5 models. The EAM model tends to rank particularly well in radiative signatures at the top of atmosphere, clouds and radiation, and are somewhat less accurate in terms of winds and the 500 hPa height field.

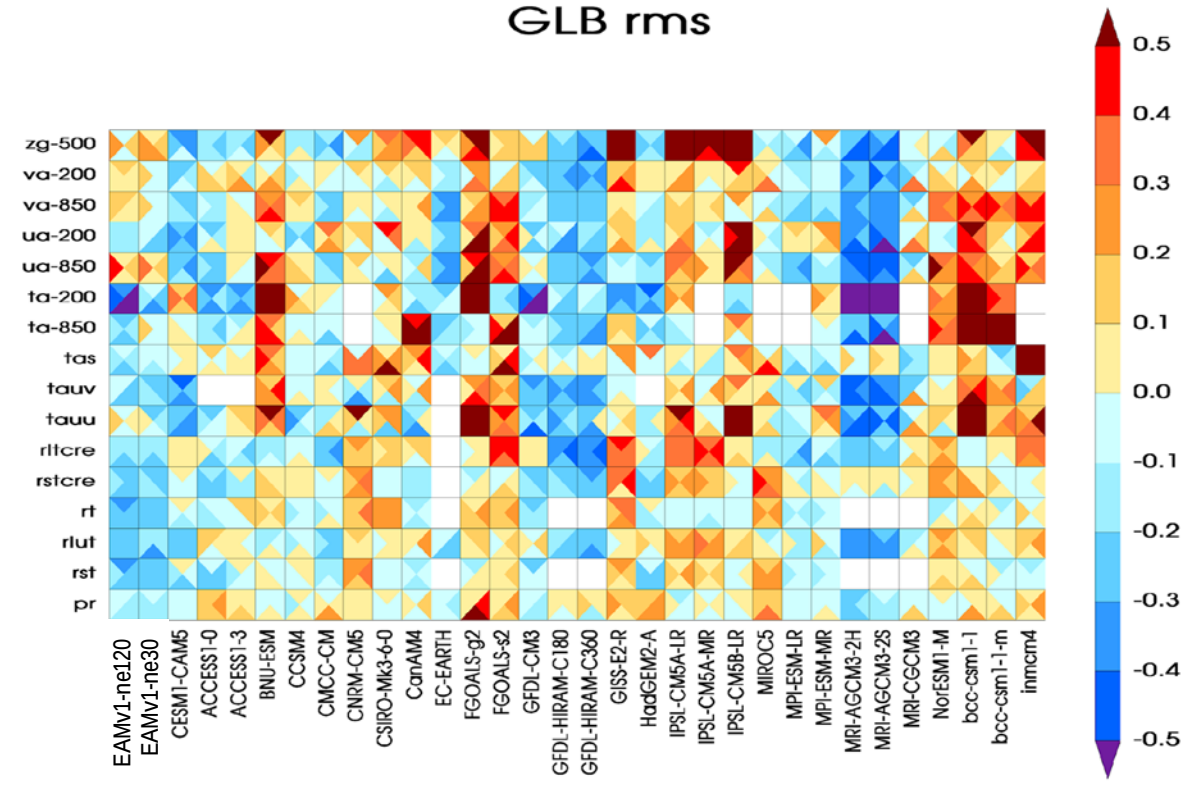

Figure 3: Heatmap diagram evaluating EAMv1 against CMIP5 AMIP simulations following Gleckler et al (2008). The fields listed by row are (from top to bottom), 500 hPa geopotential height, 200 and 850 hPa meridional and zonal wind, and air temperature, surface air temperature, meridional and zonal surface stress, top of atmosphere long and shortwave cloud radiative effect, net, long and shortwave TOA fluxes, and precipitation.

Table 3: Global averages for some important climate quantities for EAMv1 AMIP simulations compared to CAM5 and (central estimates for) observations. Meteorological fields were compared to the Japan Meteorological Agency 25 year reanalysis. Top of Atmosphere Radiative flux estimates are drawn from the CERES-EBAF 2.8 dataset (Loeb et al, 2009). Precipitation from the Global Precipitation Climate Project (GPCP, Adler et al., 2003). Surface Radiative budgets from CERES-EBAF (Kato et al, 2012), AOD from Ma et al (2018). Observational estimates are not entered for liquid and ice water paths because model and observational estimates are sampled so differently that comparison is not very reliable.

Variable	Obs Estimate	CAM5	EAMv1LR
T10 (Ref Height Temp, K)	288.3 (JRA25)	287.7	288.2
LWCRE (Longwave Cloud Radiative Effect, W/m2)	26.1 (CERES-EBAF)	23.6	24.5
SWCRE (Shortwave Cloud Radiative Effect, W/m2)	-47.1 (CERES-EBAF)	-48.6	-49.3
All-sky Longwave (Top of Model vs Top of Atmosphere for obs, W/m2)	239.8 (CERES-EBAF)	234.1	239.4
All-sky Shortwave (Top of Model vs Top of Atmosphere for obs, W/m2)	240.5 (CERES-EBAF)	233.4	239.0
NET Flux at Top of Model (W/m2)	+0.85 (CERES-EBAF)	-0.7	-0.5
Surface Longwave Down (W/m2)	345.2 (EBAFS)	342.9	344.8
Surface Shortwave Down (W/m2)	186.6 (EBAFS)	182.7	184.6
Water Vapor Path	24.4 (JRA25)	25.3	24.4
Precipitation (mm/day)	2.7 (GPCP)	3.0	3.1
Liquid Water Path (g m^{-2})		43.8	53.4
Ice Water Path (g m^{-2})		17.2	11.3
Sensible Heat Flux (W/m2)	19.3 (JRA25)	17.9	19.2
Latent Heat Flux (W/m2)	87.9 (JRA25)	88.2	89.8
U@200hPa (m/s)	15.6 (JRA25)	16.3	15.7
Aerosol Optical Depth@512nm (unitless)	0.12 (GOCAP)	0.11	0.13

More quantitative characteristics of the EAMv1 AMIP simulation compared to present-day observations and a similar CAM5 simulation are provided in Table 3. A few notable biases are apparent in the table. Although top of model (or top of atmosphere) estimates correspond reasonably with observational estimates, EAMv1 estimates of the Longwave Cloud Radiative Effect are, like CAM5, somewhat low, and shortwave Cloud Radiative Effects are too strong compared to the CERES-EBAF estimates --- that is, high clouds do not trap sufficient outgoing longwave energy, and clouds are slightly more reflective than the corresponding observational estimate. EAM and CAM5 estimates of all sky fluxes are reported at the top of model (60 and 40km) respectively while observational estimates from CERES-EBAF are reported at the top of the atmosphere. The atmosphere above the model top is not accounted for in the model fields reported in the table. Since the CAM5 model top is much lower and the layer above holds much more mass, the standard algorithm to account for it requires a large adjustment (1.4 W/m2 in the longwave, and 3.4 W/m2 in the shortwave). The adjustment is much smaller in EAMv1 (0.2 W/m2 in both spectral regions). We chose to report only the top of model results for both models to simplify the discussion. Both models have too much absorption in the clear-sky outgoing longwave energy, necessitating a choice (through model tuning) of a somewhat too weak longwave cloud radiative effect to achieve a reasonable net radiative balance. The water vapor path appears similar to the JRA estimate, which is higher than some other observational estimates, although a little lower than CAM5. The EAMv1 downward flux of both long and shortwave radiation is also closer to observational central estimates than

CAM5. Both models also have a more active hydrologic cycle than indicated by the GPCP observational estimate, although that estimate may be low by as much as 10% (see e.g., Stephens et al, 2012). EAMv1 has higher water paths than CAM5, and lower ice water paths, but these fields are not compared to observational estimates which are judged currently too unreliable to provide a strong constraint. The detailed analysis of Zhang et al (2019) and Xie et al (2018) using observational estimates of cloud amount as a function of latitude, altitude, phase and temperature suggest that the changes in the water amounts and phase partitioning are largely a result of the new treatments of ice nucleation and parameterization of the Wegener-Bergeron-Findeisen process in the model.

4.2 Zonal Average Characteristics

Annual zonal mean biases of AMIP runs compared to present-day observational estimates of temperature, water vapor and winds as a function of height and latitude are shown in Figures 4-7. While temperature biases are quite small in both models, the EAMv1

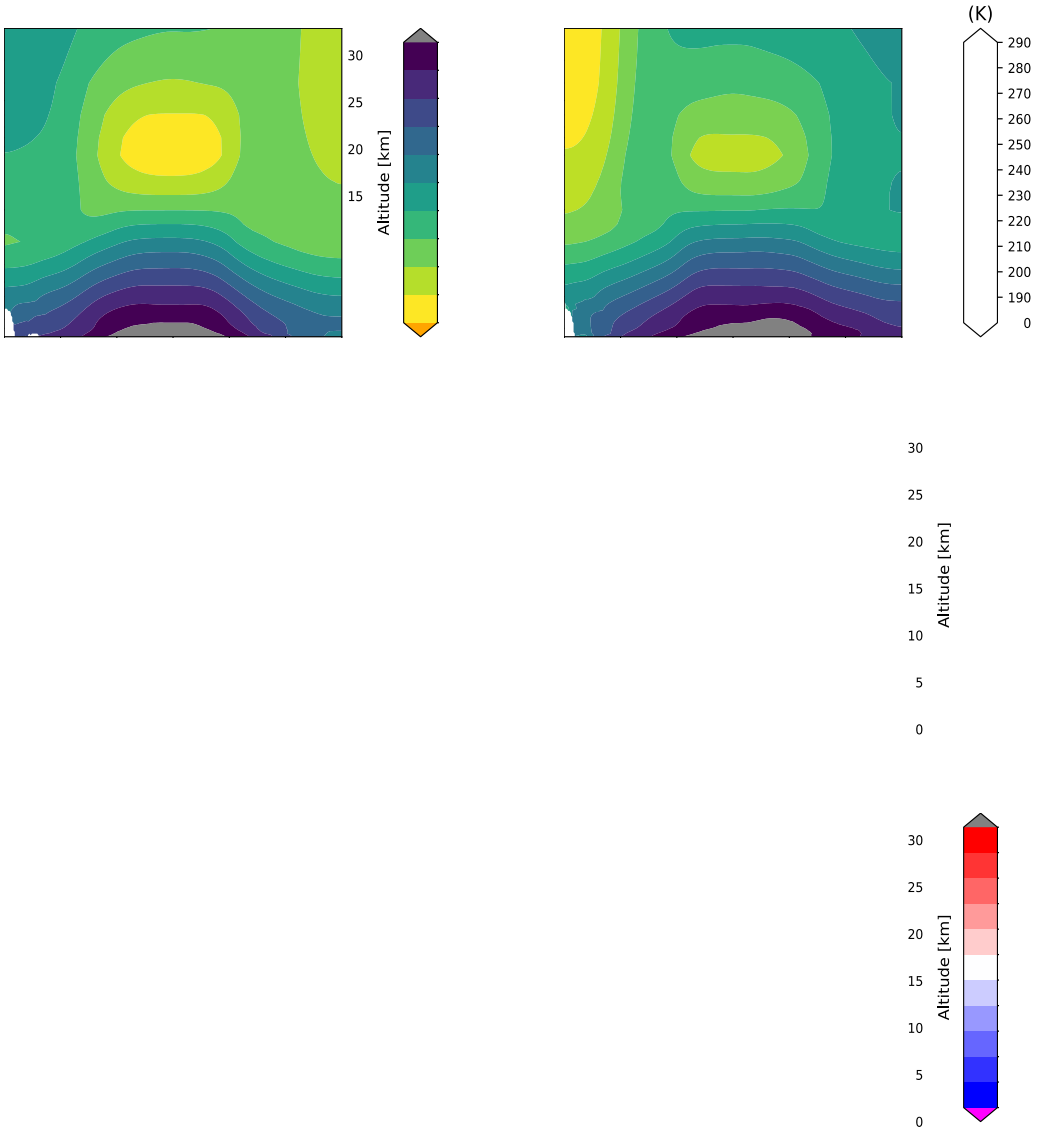

Figure 4: Zonally Averaged model temperature biases (for AMIP simulations) compared to MERRA2 Reanalysis. Left column shows December-January-February Average. The EAM and CAM difference fields share the same color bar within a column. Right column shows June-July-August. Mass and Area Weighted RMS and correlation coefficients are shown.

temperature bias (Figure 4, and see also Figure 15) has generally been substantially reduced compared to CAM5 in the upper troposphere and lower stratosphere, similar to results reported in Richter et al (2014b) when resolution is increased in the upper

troposphere and stratosphere. The EAMv1 climate is generally warmer, and the strong cold bias at the polar tropopause in CAM5 that has persisted over many generations of CAM (particularly in wintertime) has been replaced by a warm bias in the polar lower

stratosphere of about half the amplitude. EAM is a little cooler in the tropical upper troposphere than CAM5, and a little warmer over the Antarctic continent in summer. These differences are associated with the higher vertical resolution and changes in physical

parameterizations (orographic GW drag efficiency has been increased in EAMv1 to 0.25 (from 0.125 in CAM). Zhang et al (2019) have shown that EAMv1probably overestimates the supercooled liquid cloud fraction at high latitudes (unlike most models participating in

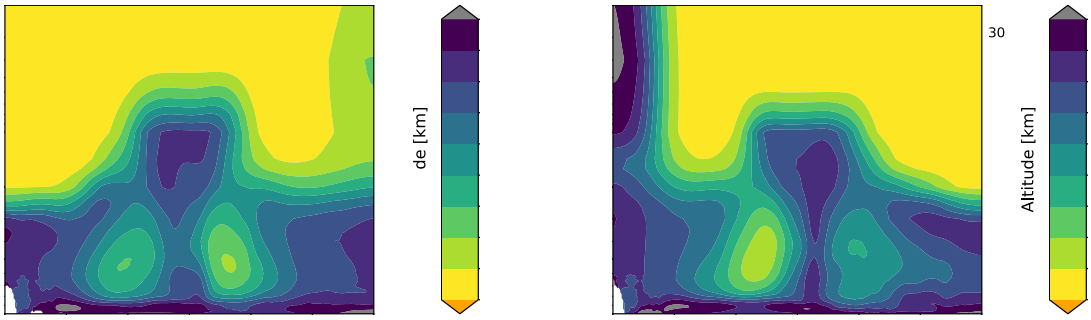

Figure 5: Zonally Averaged difference between model mean relative humidity (over liquid water) and the MERRA re-analysis. Layout details as in figure 4.

recent model intercomparisons), and underestimates pure ice clouds at most temperatures, except near -40C.

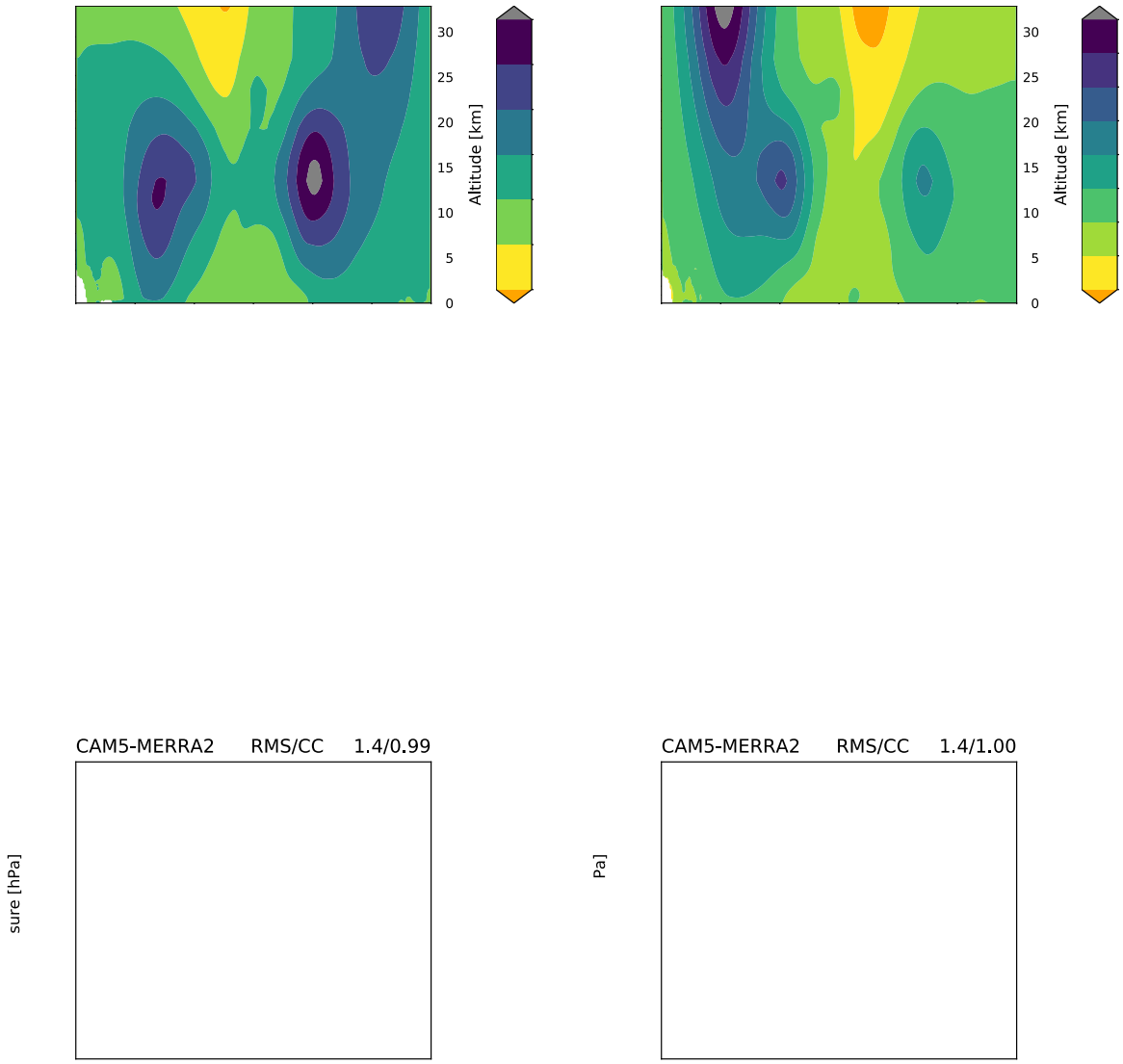

Figure 6: Zonally Averaged difference between zonal mean wind and the MERRA re-analysis. Layout details as in figure 4.

While use of relative humidity (over liquid) as a measure of simulation quality can complicate interpretation of model fidelity of water vapor because of its temperature dependence, the variation of specific humidity over many orders of magnitude can also make differences in that field difficult to interpret, but where temperature biases are relatively small (most of the troposphere) relative humidity is a useful measure. EAM, like CAM5 shows a persistent moist bias through much of the troposphere as indicated in Figure 5 (annual zonal mean relative humidity, see also Figure 15), but EAM is moister than CAM5 through much of the troposphere.

The large differences in the two simulations in the lower stratosphere and at the tropopause arise from a number of contributing factors: 1) The big cold biases in CAM5 at the polar tropopause amplify perceived moist relative humidity biases (a 10 degree change in temperature produces a factor of ~2 reduction in the saturation vapor pressure appearing in the denominator of the relative humidity calculation), but those positive biases are also driven by the larger computational mixing associated with the lower vertical resolution of CAM5. Higher vertical resolution in EAMv1 decreases the computational cross-tropopause water vapor transport, supporting lower relative humidity there. 2) Increased static stability in the lower stratosphere associated with decreased mid-latitude stratosphere-troposphere exchange particularly during northern hemisphere winter (DJF) also decreases mixing between the moist troposphere and the drier stratosphere in EAM. 3) The better resolved stratosphere and higher model top provides an opportunity for a more realistic Brewer Dobson circulation, with upwelling in the tropics, and subsidence at midlatitudes that is driven by "Downward Control", which can result in more realistic equator to pole transport of tracers. The larger moist bias in the CAM5 lower tropical stratosphere during DJF when orographic wave driving is strongest, produces a persistent winter pole moist bias in the lower stratosphere characteristic of "low top models" (see, e.g., Haynes et al (1991); Mote et al (1993)). 4) The generally low relative humidity evident in the EAM stratosphere is also consistent with, and probably explained by an important missing water vapor source there (methane oxidation, see e.g., Mote et al (1993)).

Zonal mean wind biases above 200 hPa are very different in the two models (Figure 6). The general easterly bias in the tropics and subtropics present in CAM5 has been replaced with a stronger westerly bias, but midlatitude biases are somewhat reduced. The westerly bias in the tropical lower stratosphere in EAMv1 is due to the previously mentioned oscillation in the tropical zonal wind field that is similar to a quasi-biennial oscillation, but the amplitude is too strong, and the period too short (not shown). Improvements to these features are being addressed in developmental versions of the model.

4.3 Latitude/Longitude Climate characteristics

There are climatological changes in the near surface (reference height) temperature due to the new parameterizations and much higher vertical resolution near the surface (Figure 7). Since these model simulations were performed with prescribed sea surface temperatures, the reference height temperature over oceans is strongly constrained, and small differences there reflect changes in the near surface temperature gradients due to the change in turbulence parameterization, and the much higher vertical resolution of EAM,

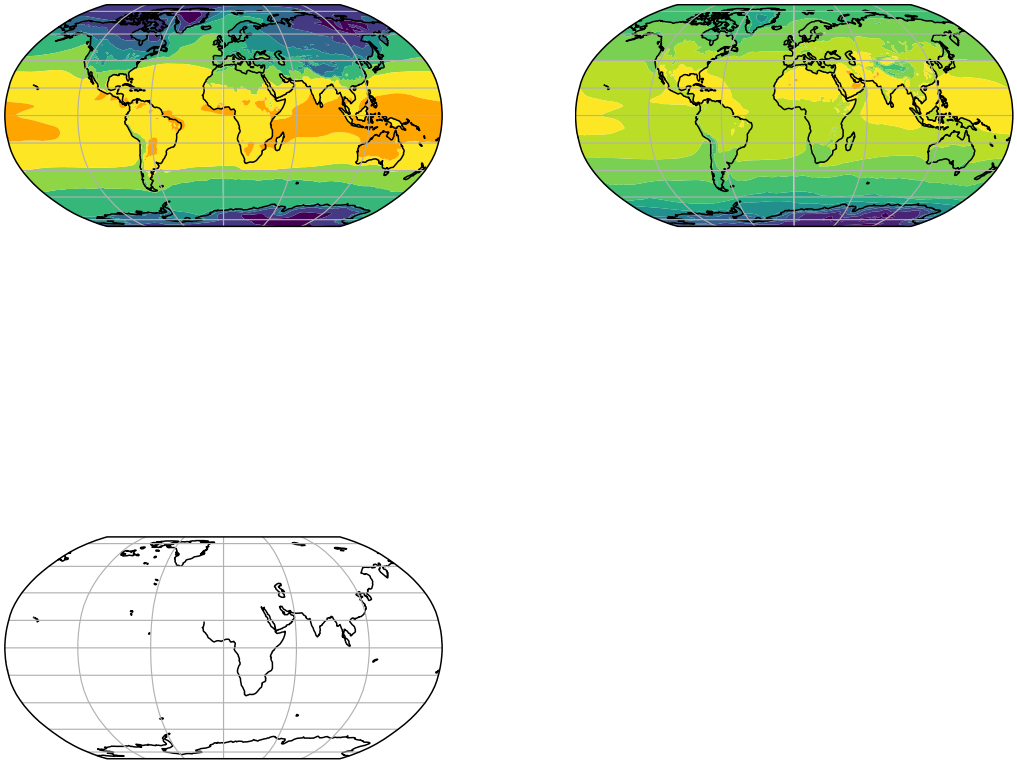

Figure 7: Difference between model reference height temperature and the MERRA2 Reanalysis. Layout details as in figure 4.

which can easily support much stronger gradients near the surface. Temperature responses over land and sea ice are also driven by these same processes but can also evolve due to interactions with the other model components, and so respond to energy and water fluxes driven by clouds, and radiatively active constituents in the atmosphere (aerosols, water vapor, clouds, ozone). EAM temperatures are much warmer over the NH continents in winter poleward of 40N due at least in part to the tuning choices used in the Bergeron process, although the amplitude of the temperature bias is about the same magnitude. Temperature biases in the same regions in NH summer are much reduced compared to CAM5, and many other climate models (Ma et al. 2018). EAM is also generally warmer over sea ice and ice sheets, particularly in winter, and this is due to the substantial changes in the surface energy budget due to differences in clouds and the surface energy budget discussed below.

Annual mean biases in precipitation (left column, Figure 8) show similar patterns in EAMv1 and CAM5, but the amplitude of many biases have been reduced (e.g., RMS average

error in precipitation is 0.94 and 1.12 mm/day in EAMv1 and CAM5 respectively). Excessive precipitation seen in CAM5 over the Arabian Sea, Bay of Bengal, Pacific Warm pool and central Pacific are reduced, and the underpredicted precipitation in that model over S. America is also more realistic. The middle column of Figure 8 shows the difference between summer and winter precipitation as an index of monsoon activity (Wang and Ding, 2008). Biases present in CAM5 in the Asian (S.E. Asia, India, Bay of Bengal), S. American and African Monsoons have all been reduced. More detail about CAM5 and EAMv1 biases in monsoon features and sensitivity to the optional gustiness parameterization mentioned in section 2.1.3 are discussed in Harrop et al (2018).

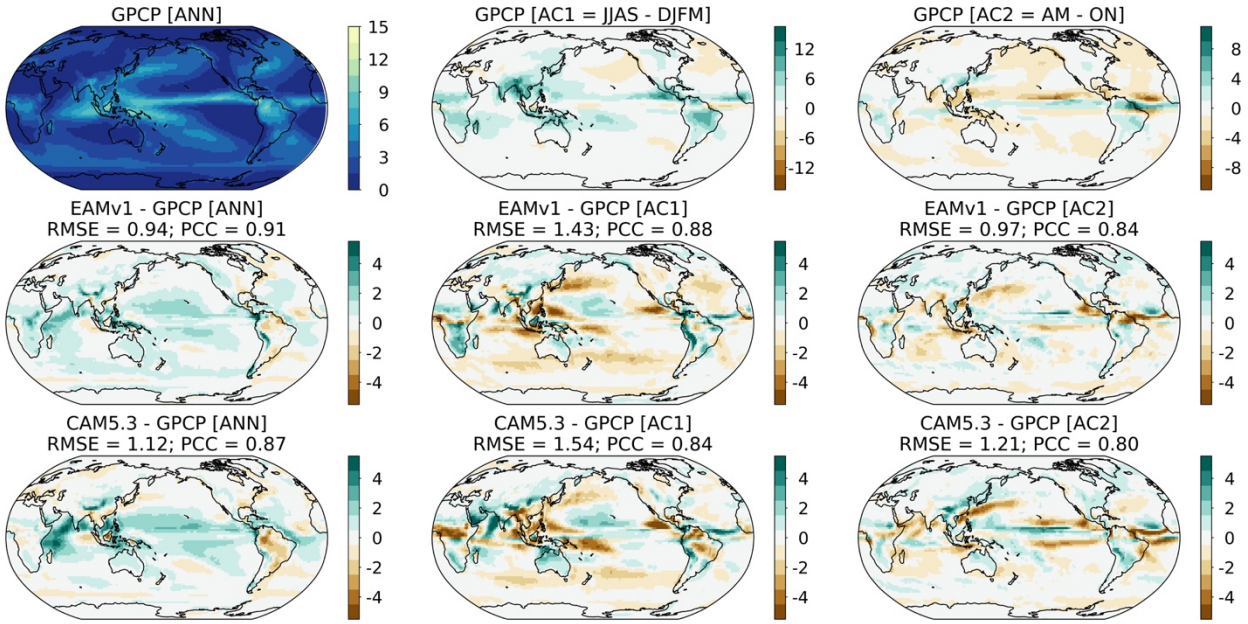

Figure 8: GPCP estimates compared to AMIP model simulations (EAM center row, CAM5.3 bottom row). Left column shows Annual Mean precipitation (mm/day). Center column shows Summer-Winter precipitation as a measure of monsoon strength. Right column shows the precipitation differences across equinox months (April/May – Oct/Nov). The lower right four values (AC1 and AC2) are multiplied by -1 in the S. Hem. to account for the change in seasonality. A white color in the difference fields indicates small biases compared to observational estimates.

Patterns of model biases in aerosol concentrations (Figure 9) are also similar in both models compared to an observational estimate of aerosol optical depth (the GCM-oriented CALIPSO aerosol product, (GOCAP), P. Ma et al (2018)), but EAMv1 biases are substantially lower. Biases over remote regions (oceans and high latitudes) are much lower. Aerosol mass and composition over oceanic regions have changed due to differences in physical parameterizations, and the new and improved aerosol sources. Many distribution changes are also associated with the improved treatments of aerosol wet removal (Wang et al., 2013), ageing of carbonaceous aerosols (Liu et al., 2016), and resuspension of aerosol particles from evaporated raindrops to the coarse mode. There is also a notable reduction in the low AOD bias over East and Southeast Asia, which is likely

due to improved CMIP6 aerosol and precursor emissions and the better treatment of SOA. Burrows et al (2018) noted that the new formulation for marine organic aerosols contributed an additional source of cloud condensation nuclei, strengthening shortwave radiative cooling by clouds, and changing summer cloud radiative effects in the Southern Ocean. AOD in regions dominated by dust emissions remain much higher than observations.

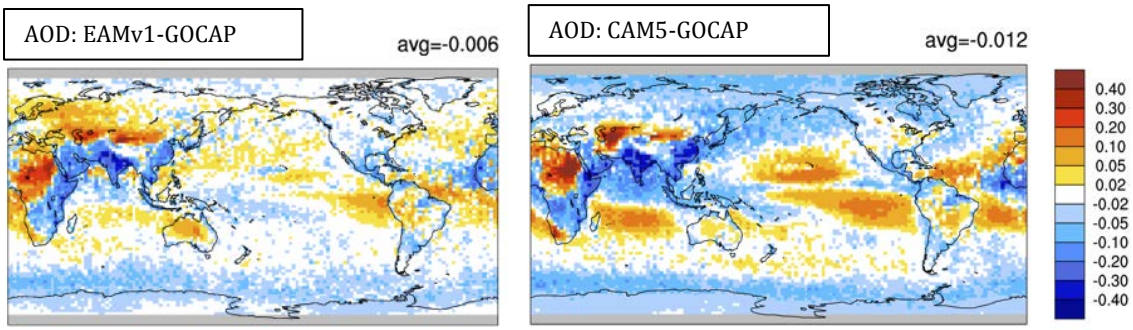

Figure 9: Difference between model (AMIP) and observational estimates (GOCAP) of total Aerosol Optical Depth. Left panel EAMv1, right panel CAM5.

EAMv1 biases in shortwave (SW) Cloud Radiative Effect (CRE, Figure 10) are generally smaller than CAM5 over Southern Hemisphere continents and in the Southern Ocean. The weak amplitude of Southern Ocean SWCRE in CAM5 noted by Kay et al (2014) and common to many models (Tan et al, 2016) is frequently associated with mixed-phase or supercooled liquid clouds. Zhang et al (2019) showed that supercooled liquid water clouds are much more prevalent in EAMv1 than many other GCMs due to our choice of the rate at which water is transferred from liquid to the ice phase via the WBF process (discussed in section 2.2.3), More realistic partitioning occurs when the parameter is increased, and this is a change planned for the next generation of EAM). In addition, the application of Classical nucleation theory (CNT) into EAMv1 (Wang et al. 2014) to replace the Meyers ice nucleation scheme used in CAM5 also leads to an increase of supercooled liquid cloud. This is because the CNT scheme links the number concentration of ice nucleating particles (INPs) to the number concentration of aerosols, which usually leads to a smaller number concentration of INPs compared to what is estimated from the Meyers ice nucleation scheme in which the INP number concentration is often overestimated over the high latitudes (Liu et al. 2011; Xie et al. 2008 and 2013). The reduced INPs result in a slow-down of the WBF process and therefore an increase of supercooled liquid in the mixed-phase clouds. In the Northern Hemisphere biases in both models are generally small, except over the Arctic during JJA where EAM clouds reflects sunlight too strongly, and CAM5 too weakly. Remaining biases in these high latitude cloud systems are still associated with treatments of heterogeneous ice nucleation and the WBF process discussed by Tan et al (2016), and Zhang et al (2019). CAM5 tropical CRE biases appear centered on regions of deep convection, while EAM low latitude biases are located primarily in regions of trade cumuli and summertime stratocumulus (Xie et al, 2018). Experimental versions of EAM now exist with substantial improvements for these features in EAM.

Biases in Longwave CRE (Figure 11) are smaller than CAM5 over most regions and seasons, except:

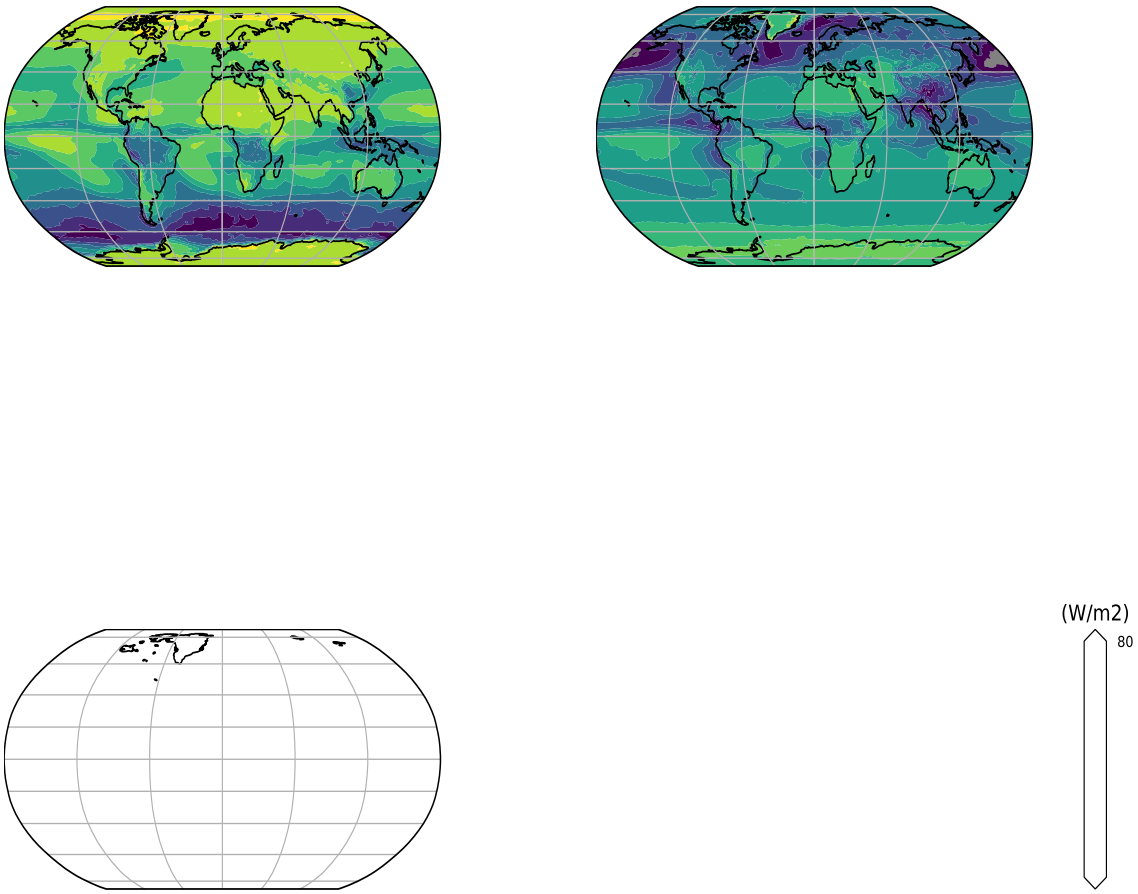

Figure 10. Shortwave Cloud Radiative Effect: Top row from EBAF 4..0 observational estimates. Left column shows December-January-February Average. The EAM and CAM difference fields (for AMIP simulations) share the same color bar within a column. Right column shows June-July-August.

- in the tropical warm pool where EAM high clouds are too optically thick and not extensive enough. The analysis of Zhang et al. (2019) indicate there is a significant underestimate of high clouds of intermediate optical depth compared to observational estimates in the low resolution model, with significant improvements in LWCRE in the high resolution configuration, but the smaller error appears to be a result of error compensation between its underestimated optically intermediate high clouds and overestimated optically thick high clouds.
- The error is also large in the Arctic during JJA where clouds also appear optically too thick, and this is likely due to the revisions to aerosols and clouds. Wang et al (2013) showed similar changes to liquid and ice clouds and associated CRE when similar modifications were introduced in CAM5.

Although figure 12 might formally belong to the zonal plots of section 4.2, it is convenient to concisely summarize here all of the all-sky radiative fluxes (sensitive to clouds, aerosols, and surface albedo) differences between model and CERES-EBAF observational estimates at the top of atmosphere (Loeb et al., 2009) and surface (Kato et al., 2013) using zonally averaged line plots. EAM annually averaged net fluxes (right bottom panel) are generally closer to observations at both the top of atmosphere and surface than CAM5. The differences between the models are more evident when examined by season and partitioned into long and shortwave radiative components (two left columns, two upper rows). It is no surprise that TOA fluxes generally agree more closely with the observational estimates than surface fluxes for both models because 1) the models were

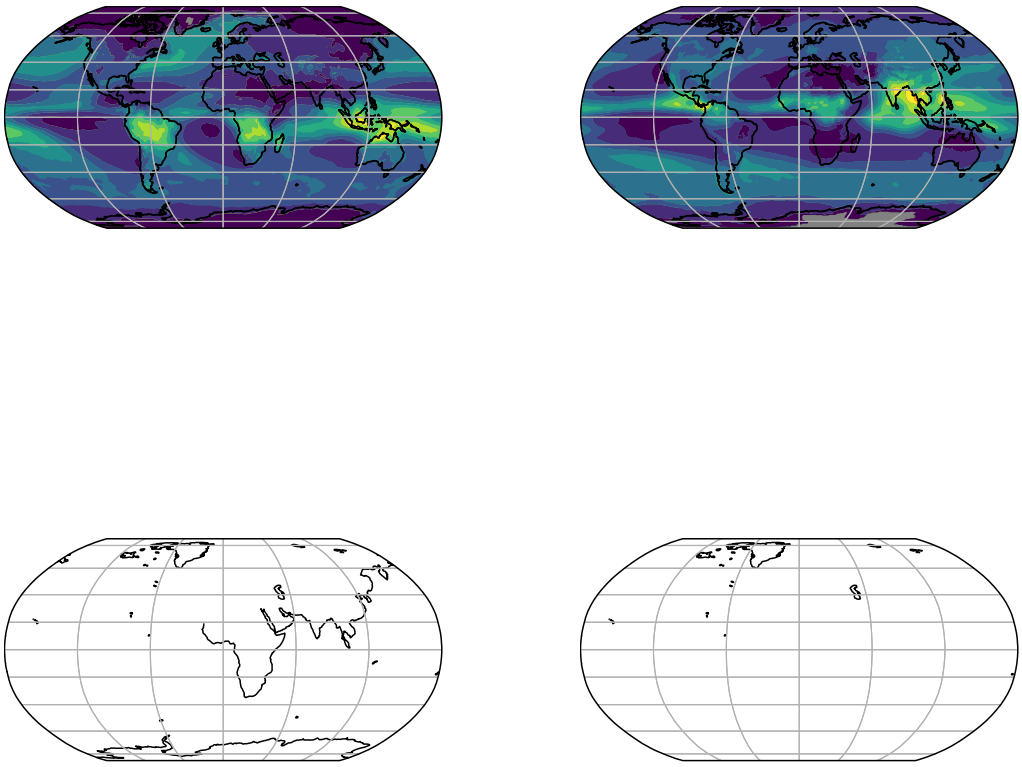

Figure 11. Longwave Cloud Radiative Effect: Top row from EBAF 4.0 observational estimates. Left column shows December-January-February Average. The EAM and CAM (AMIP) difference fields share the same color bar within a column. Right column shows June-July-August.

tuned to optimize fidelity at the TOA (and not at the surface), and 2) the CERES-EBAF surface products require use of a radiative transfer model and atmospheric state to produce surface flux estimates, so the estimates are less strongly constrained by measurements. SW TOA and surface biases for each model are very well correlated because atmosphere is dominated by scattering. EAM SW Flux biases are generally smaller for all seasons at TOA and the surface south of about 40N, particularly in the high southern latitudes (because of the improved characterization of cloud forcing in the Southern Oceans

(40-70S) and in the tropics. EAM SW errors are significantly larger than CAM5 poleward of 60N.

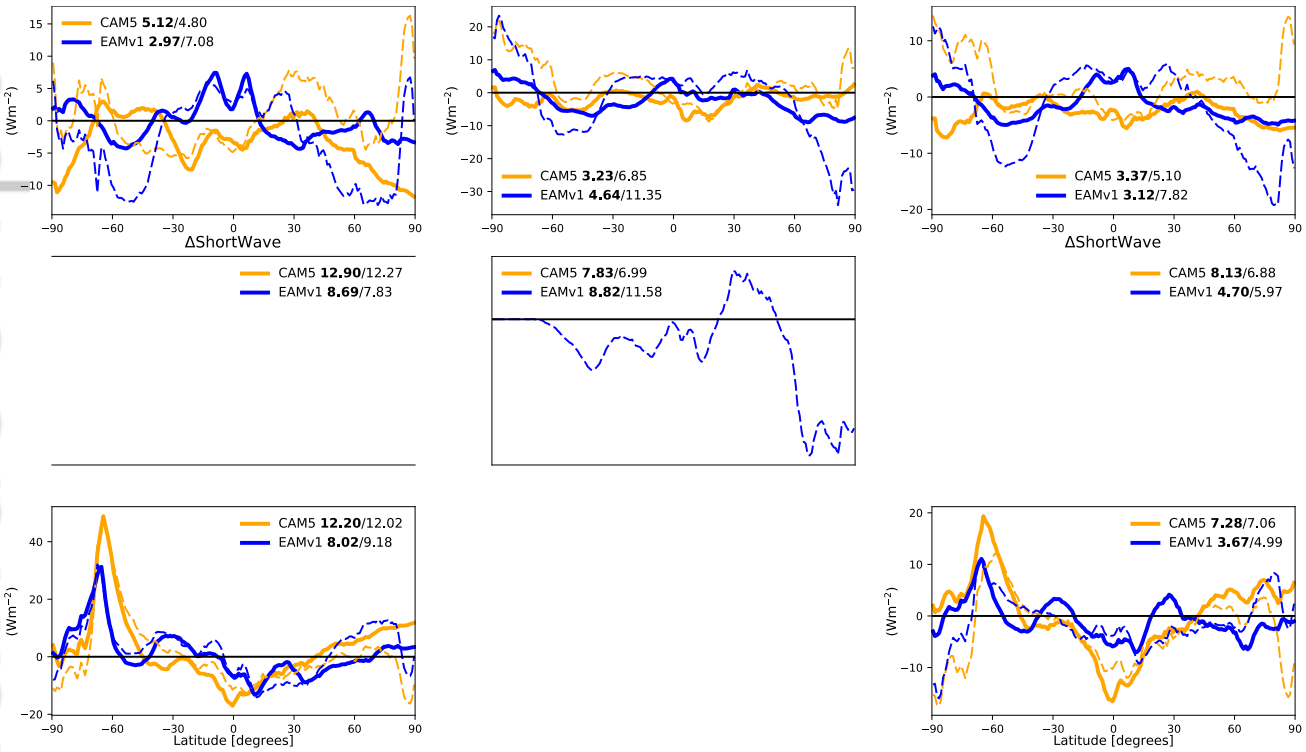

Figure 12: Differences between model AMIP simulations (EAMv1 in blue, CAM5 in orange) and CERES-EBAF observationally based estimates of all sky fluxes using the sign conventions adopted by the CESM community (longwave flux is positive out of surface, shortwave flux positive into surface, net flux is longwave minus shortwave). Heavy Solid lines/numbers show top of atmosphere differences and area weighted RMS Lighter lines/numbers show differences in surface fluxes.

4.4 Sensitivity to Horizontal Resolution

As pointed out by Bacmeister et al (2013), it is not inevitable that the climate of a high-resolution simulation be dramatically improved over a low-resolution counterpart. Explorations with previous generations in the lineage of this model using CAM4, and the simplified versions and variants of CAM5 reported in Bacmeister et al (2013), Wehner et al (2014), and O'Brien et al (2016), have reported a variety of common features, including an increase in the ratio of stratiform to convective precipitation, a decrease in cloud forcing, and after retuning modest improvements to monsoon systems, tropical wind fields, frequency of occurrence of tropical cyclones, and reduction in precipitation biases during winter in the southeast United States, along with a degradation (or no improvement) to the pattern of the intertropical convergence zone, with significant changes (both good and bad) to the precipitation intensity statistics. These differences are due both to the resolution

changes themselves, and to some parameterization's sensitivity to the time step, which must be decreased as resolution increases to avoid computational instability. Some of these same characteristics and experiences occurred during development of EAM. Xie et al (2018) have described the model response to resolution changes without parameter changes and outlined the required changes in parameter settings needed to produce a reasonably tuned model. Although a thorough evaluation of the high-resolution model

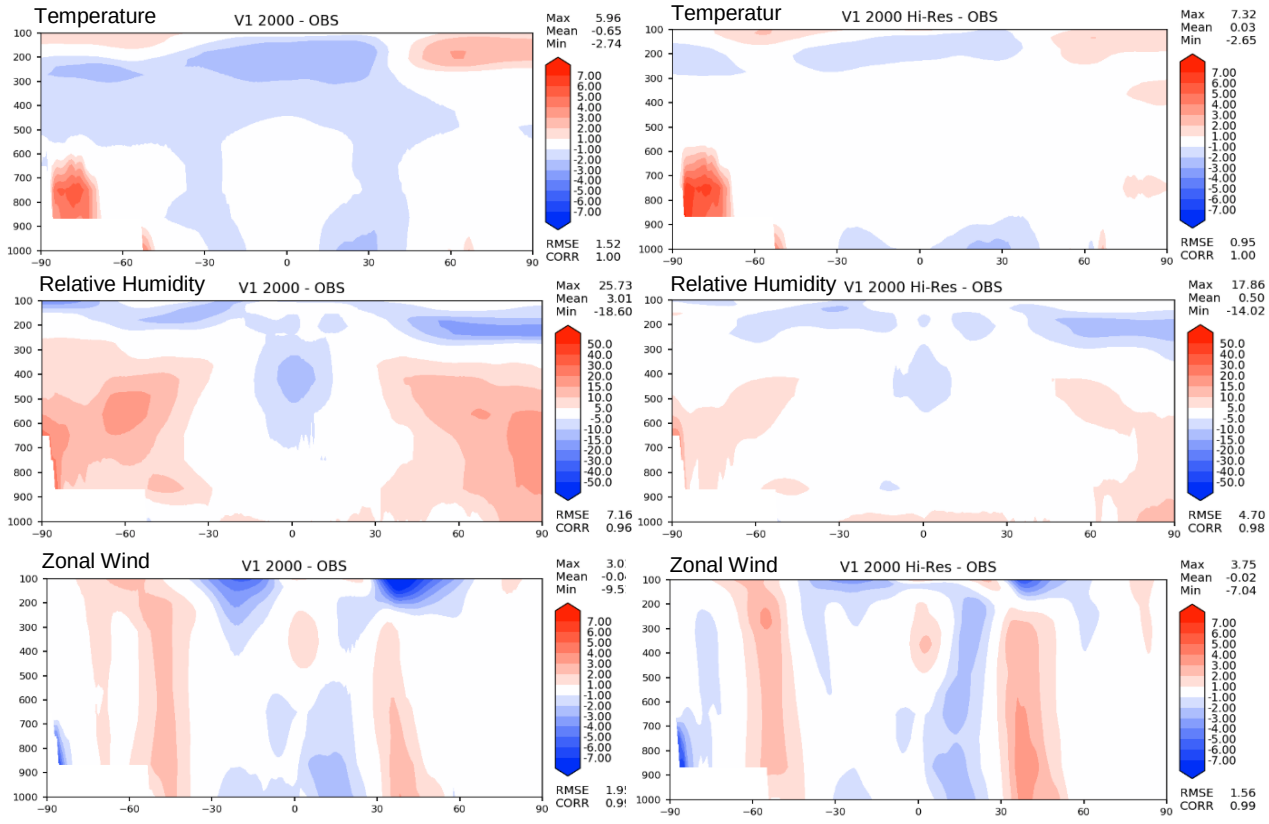

Figure 13: A comparison of model biases in F2000LR Low (left column) and F2000HR High Resolution (Right Column) simulations compared to the ERA-Interim (1980-2004) reanalysis. Top row: Temperature(K); Middle row: Relative Humidity(%); Bottom row: Zonal Wind (m/s).

behavior is beyond the scope of this study (other works are in preparation) we present here a few hints about model behavior that can be gleaned from monthly mean fields.

Figure 13 shows EAM differences in Temperature, Relative Humidity and Zonal Wind as a function of resolution for the F2000 configuration, displaying biases with respect to the ERA reanalysis product at low (left column) and high (right column) horizontal resolution. Since the model tuning has also changed, these differences reflect change in both resolution, and parameterization changes. The high-resolution model configuration seasonal and annual biases are notably smaller for both temperature and relative humidity throughout the troposphere, and there are hints that the stratospheric jet structure in the subtropics is also more realistic. Near surface biases in most variables, including those shown in figure 13 and many other variables appear insensitive to horizontal resolution,

suggesting that those deficiencies are often more strongly driven by inadequacies in physical parameterizations.

Figure 14 shows EAM differences between modeled and GPCP (Huffman et al, 2009) total precipitation estimates for F2000 HR and LR configurations. Error patterns are very similar at both resolutions, but the amplitude of the largest errors (indicated by the magenta color interval) is smaller at high resolution in the tropics (ITCZ and Warm Pool), and over steep topography (Andes and Himalayas). There are also notable improvements during the summer/wet season over southern Africa, Northwest Canada, Alaskan North America (with a degradation in the SE US), and the ITCZ bias that extends east of the dateline at low resolution (bottom right panel) is much reduced at high resolution (bottom left panel).

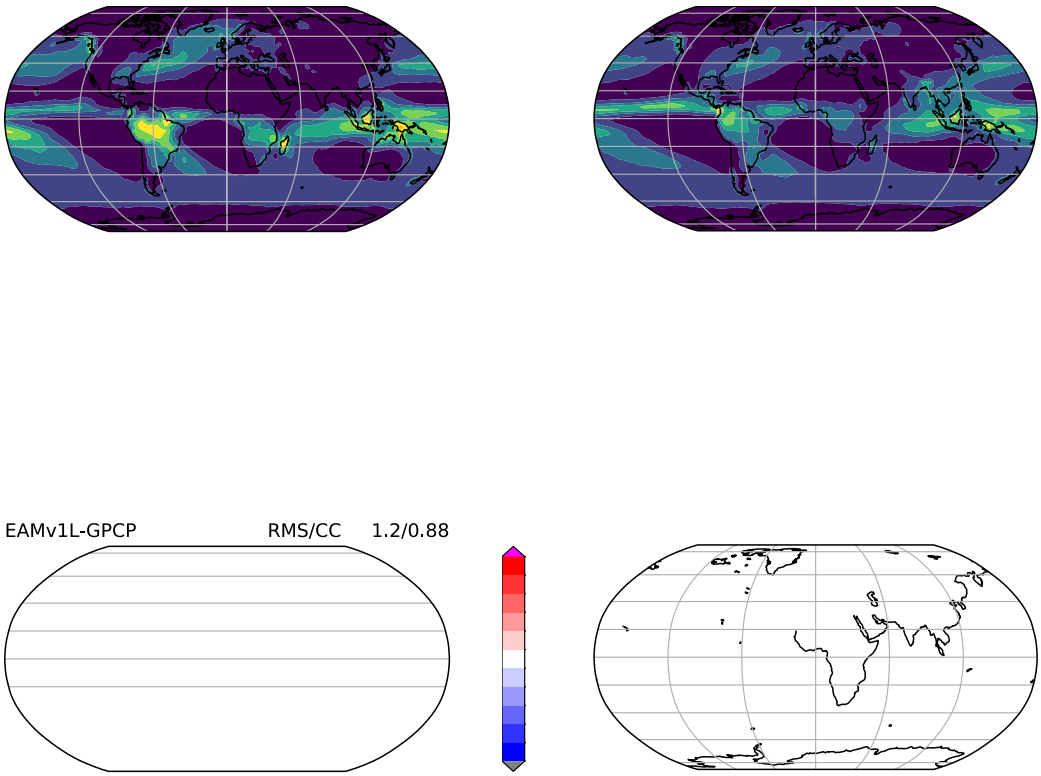

Figure 14: Differences between DJF and JJA mean modeled (F2000) and GPCP precipitation (top row, Huffman et al, 2009) estimates for 4 year averages (years 2-5) of the high (middle row) and standard low resolution (bottom row) model simulations.

Figure 15 shows mean JJA differences between modeled total precipitation and GPCP1DD estimates (Huffman et al, 2001) over the Continental United States (CONUS) at LR and HR standard resolutions and using EAM with regional refinement. Differences between the CONUS RRM and the high-res model are generally statistically insignificant. Simulated precipitation is improved (i.e., smaller model-observation differences)

with increased resolution, especially over mountain regions, such as over the Western United States. These differences are present in Figure 14 but are not apparent because of the larger contour interval in the global plots. Largest departures between low and high-resolution simulations are seen along the eastern seaboard. These results highlight the utility of RRM as a useful tool for the high-resolution model development by mimicking the behavior of the globally uniform high-resolution simulations with substantially reduced computational cost. More discussion of the RRM configurations can be found in Roesler et al. (2018) and Tang et al. (2019).

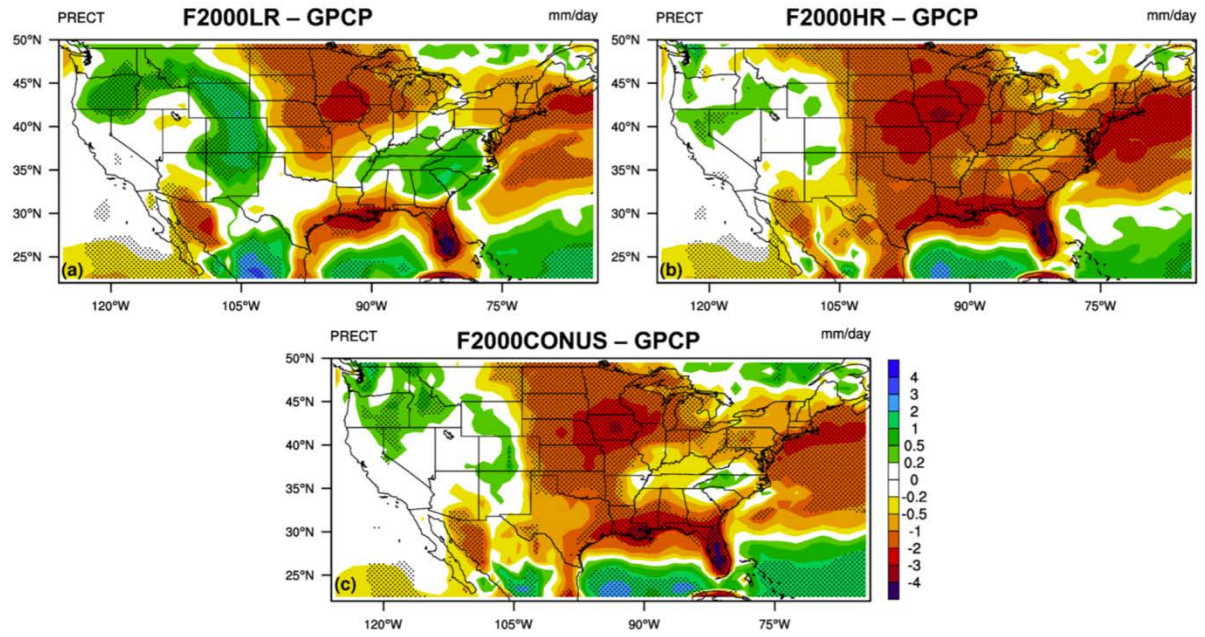

Figure 15: Simulations of JJA mean total precipitation against GPCP1DD precipitation (Huffman et al, 2001) estimates for 4 year averages (years 2-5) of the standard low resolution (left, ne30) and regionally refined model configuration over CONUS (right, ne30->ne120) Dotted areas indicate where the differences are statistically significant at the 95% confidence level with a two-tailed Student's t test.

4.4 Tropical Variability

Although there isn't room for much discussion of EAMv1's major modes of variability, two important tropical features are displayed to provide a little insight into aspects of the simulation where signatures are similar to CAM5, and where they are different.

Figure 16 shows that significant biases remain in the diurnal signal of precipitation in the timing phase, amplitude and coastal coherence (continuity of phase and amplitude as precipitation propagates away from island coasts). Despite the significant changes in turbulence and cloud physics, EAMv1 maintains many of the persistent biases present in

CAM5, including on average a timing that is at least 6 hours too early over land and 4 hours too early over ocean. An increase in diurnal amplitude is in better agreement with observations. Given the similar biases to CAM5 it is unsurprising but reasonable to conclude that the common ZM deep convection in both models exerts the dominant influence on the diurnal variability. Changes in the vertical grid and modifications in the remaining physics schemes would therefore seem to have a secondary role.

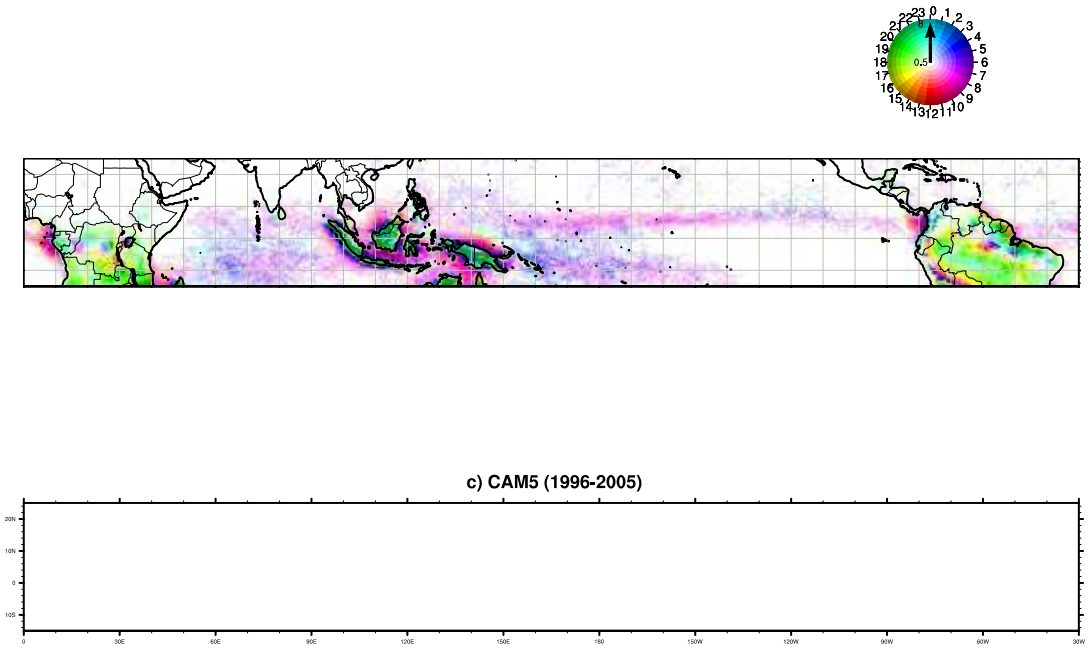

Figure 16: Timing phase (color) and amplitude (color density) of the first diurnal harmonic of total precipitation (mm/day) from 10 years of 3-hourly averaged data for (a) Observed (TRMM), (b) EAMv1 (AMIP) and (c) CAM5 (AMIP).

Figure 17 highlights the fidelity of some of the major sub-seasonal organized modes of variability in the deep tropics by displaying observational and model estimates of the power spectrum for OLR. The observed field shows phase space peaks corresponding to the ubiquitous, equatorially trapped propagating wave modes responsible for much of the regional variability of clouds and precipitation within a season. EAMv1 significantly under-represents the strength of many of these modes, particularly for the key eastward propagating low-frequency Kelvin modes, and this represents a degradation compared to CAM5. However, there is marked improvement in the MJO strength in EAMv1 which appears as a much more coherent propagating signal through the Maritime Continent and into the West Pacific (not shown). Sensitivity experiments performed with prescribed SSTs in EAMv1 indicate that a combination of factors are responsible for the improvement, including the inclusion of CLUBB and the additional vertical levels. Unlike CAM5, optimization (tuning) choices in EAMv1 also appear to strongly affect sub-seasonal variability characteristics. Another more fundamental transformation compared to CAM5 is

that atmosphere ocean coupling plays a very strong role in modulating MJO activity. Coupling was a small influence on the presence and strength of the MJO in CAM5 and CESM1 (not shown, but Fig 17(d) differs minimally from CESM1). EAMv1 produces an altogether different response to coupling, as the strength of the MJO and the wavenumbers over which it is active increases in coupled (E3SM) versus prescribed SST model configurations (panel b vs c). The underlying causes are somewhat unclear, but preliminary study indicates that coupling appears to dramatically reduce the barrier effect of the Maritime Continent. Furthermore, zonal precipitation and surface stress biases in the Indian Ocean, although somewhat degraded from CAM5/CESM1, appear to generate initial MJO-type disturbances that propagate more robustly and consistently toward the Maritime continent. The resolved equatorial wave spectrum is important (among other reasons) in forcing of QBO, and the low amplitude in EAMv1 and E3SMV1 has implications for the the gravity wave drag parameterization tuning. Improving the capability of the model to excite a more realistic tropical spectrum will be an important goal for the next model version.

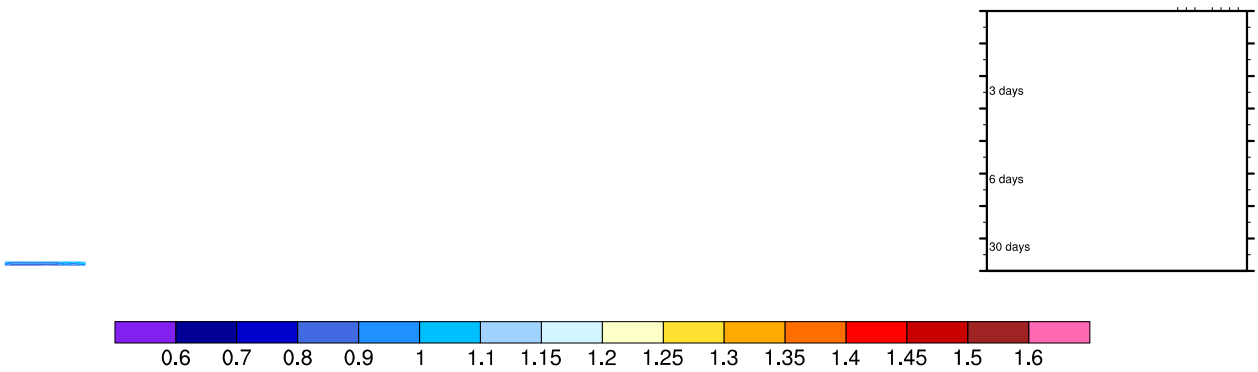

Figure 17: Wavenumber frequency spectra (following Wheeler and Kiladis, 1999) showing the ratio of unfiltered background spectra of daily interpolated (symmetric about the equator) outgoing longwave radiation (1986-2005) for (a) Observed (NOAA), (b) EAMv1 (AMIP), (c) E3SM (coupled), and CAM5 (AMIP).

4.5 Response to Forcing agents and Cloud Feedbacks

Realistic implications of the response to forcing changes must be done in a coupled modeling framework, but simpler simulations can provide insight into the forcing and feedbacks in the model (Ringer et al, 2014; Forster et al, 2016). Effective Radiative Forcing (ERF, Boucher et al, 2013) provides an estimate of the change in fluxes due to a change in forcing agent after "rapid responses" in the atmosphere (clouds and stratosphere) and land take place, but before longer timescale adjustments occur. Forster et al (2016) noted that ERF estimates can be quite sensitive to the calculation method and indicated that robust estimates can be made with fixed SST simulations. Table 4 provides estimates of changes in top of atmosphere flux produced by differences in paired runs designed to expose differences between simulations as climate forcing agents are varied. Some of the runs used to estimate the model response to a forcing agent change in table 4 are drawn from Golaz (2018). Simulations from that study allow estimates of

the model response to be made by comparing paired simulations with different forcing agents when surface temperature and volcanic stratospheric aerosols are varied by year (AMIP SSTs and a historical record of volcanoes were used).

Total ERF characterizes the model response to changes in GHGs, aerosols, land use, and land cover in the absence of SST changes. Column 2 (C2) shows flux changes obtained by differencing a run using CMIP5 emissions, LU/LC, GHGs and (climatological, repeating) SSTs representative of a time window centered on 1850 (F1850), with a run having the same SSTs but forcing agents representative of a 20 year window centered on the year 2000 (F2000AllF). The total ERF or Adjust Forcing, inferred from the FTOA field for these configurations is estimated to be ~1.26 W/m2, somewhat higher than the EAMv1 estimate of ~1.1 W/m2 reported by Golaz (2019) which used an average of three simulations contrasting CMIP6 forcing agents between (1995-2014) and 1850 forcing agents over varying AMIP SSTs for (1995-2014). A much lower estimate of 0.36 W/m^2 (C3, FTOA) is present for that same model configuration using only the third realization (identified as A3 in Golaz et al (2018)) when the estimate is evaluated from a 25-year segment started 15 years earlier (1980-2004). The lower forcing estimate, consistent with figure 24 of Golaz et al (2019), is sensitive to the presence of aerosols from stratospheric volcanoes (absent during the period Golaz et al focused on), which explains most of the difference in the estimates, but emission sources of anthropogenic aerosols were also significantly higher during the earlier (1980-2004) period. The differences compared to the F2000 emissions used in C2 may also be partially explained by the SSTs differences used in columns C2 and C3 --- clouds and cloud responses to aerosols and GHG can be sensitive to surface temperature and the 1980-2014 period is characterized by very large El Nino events not present in the SST climatologies used in C2, or the later time period analyzed in Golaz et al (2019). The differences in emissions, and small differences in methodology compared to that used in Golaz (2019) and the Forster et al (2016) result showed that forcing estimates can depend significantly on the length of analysis interval indicated the sensitivity of the calculation. At any rate, the EAMv1 Total ERF is smaller for all EAMv1 estimates than the Forster et al (2013) mean value 1.7 (±0.9) W/m^2 but within the range identified for 1850 and 2001-2005 from CMIP5 coupled runs. Since Total Adjusted Forcing is made more positive by GHGs and more negative by aerosol forcings, the low values indicate the aerosol forcing is strong compared to other models.

Aerosol Forcing (decomposed into an ERF due to Aerosol Radiation Interactions (ERF_{ARI}) and Aerosol Cloud Interactions (ERF_{ACI}) can be estimated by differencing runs with anthropogenic sources for aerosols changed between PI and PD values with other forcing agents held constant. Unfortunately, diagnostic fields needed to decompose the model response into unbiased estimates of ERF_{ARI} and ERF_{ACI} using the technique in Ghan et al (2013) were not archived for the runs of table 2, but biased estimates that provide rough estimates of the cloud and clear sky impacts of aerosol can be produced by differencing long and shortwave clear sky fluxes (FLNTC and FSNTC), and Cloud Radiative Effects (-LWCRE and SWCRE). The biased estimate based on the runs described in Golaz et al (2018) when CMIP6 aerosol sources vary by year and month over contemporaneously varying SSTs for 1980-2004 is shown in C4, with clear sky aerosol effects in clear (-0.49/-0.15/-0.64 W/m2) and strong cloudy (-1.37/0.57/-1.94 W/m2) flux changes producing an estimated total aerosol ERF about -1.86 W/m^2. Golaz et al (2019) noted a lower amplitude estimate of -1.65 W/m2 for the 1995-2015 time period. In this case, both simulations

included the same volcanic aerosols in the stratosphere so the differences must be associated with the higher anthropogenic aerosol sources or higher SSTs compared to the period analyzed by Golaz et al (2019). Column C5 shows CAM5 values reported in Gettelman et al (2015) with a total Aerosol ERF of -1.6 W/m2.

Table 4 Changes in globally averaged top of atmosphere fluxes (longwave allsky (FLNT), clear sky (FLNTC) and CRE (LWCRE), Shortwave allsky (FSNT), clearsky (FSNTC) and CRE(SWCRE), surface temperature (TS, total/land), Aerosol Optical Depth (AOD). C2, C3: Response to Total forcing (GHG, aerosols, and land use/land cover) C4, C5: Response to aerosol and aerosol precursor emissions. C6: response to SST changes. Estimates for CAM5 (C5) from Gettelman et al (2015) used the Ghan (2013) "clear clean" methodology that can produce differences O(0.1 W/m2) compared to the "clear" CRE estimates (A. Gettelman, private communication). Fluxes use typical sign convention for CESM models; both long and shortwave net fluxes are defined to be positive, the net flux is defined to be shortwave-longwave, and cloud forcing is defined positive for longwave and negative for shortwave.

C1	C2	C3	C4	C5	C6
Description of forcing estimnate	EAMv1 Responses to a change in all forcing agents in presence of 1850 SSTs and no volcanic aerosols	EAMv1 Response to a change in all forcing agents in presence of time varying SSTs, aerosols, and volcanic aerosols (1980-2004)	EAMv1 Response to aerosol emission changes in presence of time varying SSTs, aerosols, and volcanic aerosols (1980-2004)	CAM5.3 Response to changing aerosol emission	EAMv1 response to a +4K increase in SSTs
Cases or reference for calculation → / Change in field ↓	See Table 2 (cases (F2000LR/AF-F1850)	See table 2 of Golaz et al (2018) for description (cases amip_A3 - amip_1850allF_A3)	See table 2 of Golaz et al (2018) for description (cases amip_A3 - amip_1850aerof_A3)	CAM5.3 (AE w MG1)	See Table 2 (cases F1850LR+4k – F1850LE)
ΔFLNT	-3.03	-2.86	-0.72	~0.44	10.1
ΔFLNTC	-3.17	-2.87	-0.15		9.4
ΔLWCRE	-0.14	-0.01	0.57	0.44	-0.7
ΔFSNT	-1.77	-2.5	-2.58	-2.	3.7
ΔFSNTC	-0.63	-1.	-0.64		1.9
ΔSWCRE	-1.13	-1.51	-1.94	-2.02	1.8
ΔFTOA	1.26	0.36	-1.86	-1.59	-6.4
ΔTS/TSLAND	0.12/0.25	0.1/0.18	0.01/-0.11		4.49
ΔAOD	0.03	0.03	0.02		0.01

A few more results using a reproducible decomposition of Aerosol Forcing following the protocols suggested for the AeroCom Indirect Effect Experiments (Ghan et al., 2016, and see https://wiki.met.no/aerocom/phase3-experiments) are noted for future comparison to other models. Two two-year (2006-2007) EAMv1 simulations were made with pre-industrial and present-day emissions. Runs used a 6-hour relaxation time nudging to the MERRA (Rienecker et al., 2011) reanalysis winds, with CMIP6 representative monthly climatological values of aerosol sources composited over years 2000-2014 (Yang et al., 2018). The anthropogenic aerosol ERF using this configuration is -1.75 W m^{-2}, similar to the -1.85 W m^{-2} CAM5 forcing estimate described in Zhang et al (2016) using the CLUBB and MG2 schemes. The aerosol ERF for both models is dominated by the contribution from aerosol-cloud interactions. Contributions from aerosol-radiation interactions and aerosol-surface interactions (Ghan, 2013) are smaller than 0.1 W m^{-2} for both models. However, the contribution of aerosol-cloud interactions from warm clouds (i.e., cloud top temperature warmer than 263 K) is much smaller for EAMv1 than for CAM5. Factorization (Ghan et al., 2016) of the cloud and aerosol changes driving the difference in the response of cloud radiative forcing to anthropogenic emissions (not shown) suggests that the ERF is not driven by different sensitivities between cloud condensation nuclei, droplet number, or even the cloud radiative forcing, but primarily by differences in the control warm cloud radiative forcing, which is far stronger for CAM5 than for EAMv1. Because the global mean cloud radiative forcing and effective radiative forcing are both similar for the two models, some compensation between differences in warm and cold clouds must occur. It is noteworthy that differences in cloud radiative forcing are related to differences in the sensitivity of cloud radiative forcing to aerosol source changes. It is clear that net Aerosol Forcing depends strongly on the details of the evaluation (precise period used to define fluxes, inclusion of volcanoes, and perhaps SST pattern). Under some scenarios EAMv1 has a lower sensitivity to aerosol sources than CAM5, and in others a higher susceptibility. These subtleties are being evaluated and will be reported on in a separate study.

Ringer et al (2014) showed that reasonable estimates of model feedbacks may be estimated by calculating the top of atmosphere flux response to uniform surface temperature changes compared to fully coupled simulations. In that study, long and shortwave clear sky feedbacks calculated using perturbed surface temperatures were found to generally produce a somewhat weaker positive and stronger negative feedback compared to the fully coupled simulations. They also show that CRE feedbacks agree quite well --- longwave cloud feedbacks using prescribed SST changes are very strongly correlated with coupled estimates; shortwave cloud feeds differ more, and in most models this was because of a missing contribution from sea ice reductions in the prescribed SST runs. Differences between the F1850LR and F1850+4K simulations produce EAMv1 (Table 2, column C6) estimates of a total feedback of -1.42. Ringer et al (2014) list the total feedback from 10 +4K CMIP5 runs using AMIP SSTs as having a mean of λ = -1.6 and a range of -1.05 to -1.95, so EAMv1 has a negative total feedback near the central value. Ringer et al indicate that most CMIP5 models show a longwave $\lambda_{lw,CRE}$ feedback around 0 (±0.5) W/m2/K, with a slight tendency for fixed SST calculations to produce a more negative feedback than coupled calculations. EAMv1's value of +0.16W/m2/K is in the inner quartile range of those models. The EAMv1 clearsky longwave feedback $\lambda_{LW,Clear}$ is -

2.09 W/m2/K, again near the center of the CMIP5 range. $\lambda_{SW,CRE}$ is 0.4 W/m2/K, in the upper quartile of the CMIP5 models (with a range about -0.2 – 0.9 W/m2/K). $\lambda_{SW,Clear}$ is 0.42 within the central (inner quartile) range of CMIP5 models. Golaz et al (2019) perform a more accurate calculation of feedbacks using coupled control and abrupt 4xCO2 simulations with a radiative kernel calculation capable of breaking down the feedbacks into the Planck, lapse rate, water vapor feedback etc and conclude that E3SMv1 has a larger cloud feedback than all CMIP5 models evaluated. The stronger response may be due in part to the cloud changes in the vicinity of sea ice, to cloud masking issues, and to the patterned SST change compared to the uniform change evaluated here.

5. Summary

A new version of the E3SM Atmospheric Model (EAMv1), has been developed and released to the community. EAMv1 is a fork of the CAM5 lineage of atmospheric models. In addition to improved treatment of model physical processes, and model numerics, it includes a substantial increase to the vertical resolution, optional increased horizontal resolution, and optional regionally refined configurations that support increased resolution in a region of interest in order to optimize the balance between computational cost and resolution. Vertical resolution and extent was increased (from 30 to 72 layers, with the surface layer resolution increased by up to 5 times and the model top extended to 60 km [~0.1 hPa]). A tuned low (~100km), and high (~25 km) resolution configuration have been provided, along with a regionally refined configuration with high resolution over the continental US and low resolution elsewhere.

The new atmospheric model includes significant changes to turbulence, aerosols, cloud microphysical, convective, and precipitation processes, and chemistry. The model uses the MG2 cloud microphysics parameterization with modifications to the autoconversion process, and ice nucleation. A 4 mode version of the Modal Aerosol Module (MAM4) was included with additions designed to improve the representation of particle aging, secondary organic aerosol, and cloud interactions, and an additional aerosol type was included to represent important missing species (Marine Organic Aerosol). A simple ozone photochemistry was added to represent stratospheric ozone sensitivity to atmosphere dynamics, temperature, and depletion by prescribed halogen; the model now supports increased and more realistic variability in the upper troposphere and stratosphere. An optional improved treatment of light-absorbing particle deposition to snowpack and ice is available, and stronger connections with Earth system biogeochemistry can be used for some science problems. The changes made to turbulence and vertical resolution required modifications to the convection parameterization near the surface to make the parameterization less sensitive to boundary layer structure.

Model workflow also changed: EAMv1 used the E3SM software engineering code development process as part of its development strategy, with strict use of regression testing to prevent answer-changing bugs from entering the code, and modifications to the verification methodology capable of detecting some implementation errors in the development versions of EAM. EAMv1 is computationally 3-4 times more expensive than its predecessors due to a significant increase in parameterization complexity, more tracer constituents (40 compared to the 25 used previously), a doubling of vertical layers, and modifications to maintain numerical stability with increased resolution. Since brute force strategies using many multi-year simulations were not feasible at high resolution, other strategies were also employed during the tuning phase. Short simulations (hindcasts and ensembles of short simulations) were used during

the initial phase of tuning to converge on appropriate model configurations and better understand sensitive to parameter variations.

A brief assessment of some simulation characteristics was performed to assess changes compared to CAM5, and fidelity compared to observations. A lightweight comparison of EAMv1 to models participating in the previous (CMIP5) evaluation indicates that the model climate ranks in the better scoring half in terms of a series of common climatological metrics, and its climate fidelity is generally higher than CAM5 (its predecessor). For version 1, a lot of emphasis was placed on producing reasonable radiative signatures in monthly mean radiative fluxes (driven by state variables), hydrological fields (precipitation, water vapor, clouds), and dominant modes of variability (including the seasonal cycle, ENSO relevant signatures, the Madden Julian Oscillation), etc. Longstanding biases in the lower stratosphere and polar tropopause were much reduced, and the model has more realistic variability there. Top of atmosphere radiative flux biases have been reduced significantly. High clouds still do not trap sufficient outgoing longwave energy, and clouds remain slightly more reflective than the corresponding observational estimate. Temperature biases in Northern Hemisphere summer are much reduced. Monsoon features are improved, and the largest biases in tropic precipitation are reduced compared to CAM5, but they are still quite large. The model cloud susceptibility to aerosols is quite high.

There are many remaining areas in which EAM can be improved; wind biases in the Upper Troposphere and Stratosphere still affect major modes of stratospheric variability (e.g., the QBO, and stratospheric warmings) that are potentially important to tropospheric variability; coastal stratocumulus clouds are underestimated and are responsible for strong and persistent biases in shortwave cloud forcing; significant errors in convective precipitation in the tropical warm pool and associated radiative effects still contribute to errors in major circulation features; the phase of diurnal precipitation indicates that critical remaining problems exist in triggering of convection, which is probably related to insufficient energy in major modes of tropical variability (for example, Kelvin wave with periods from 3-30 days, and inertial gravity wave with periods less than 3 days). We are now aware that more realistic simulations for some of these features with the same model are possible with other tuning choices, and some modifications are being documented in upcoming studies, and being considered for adoption in the next generation of EAM. Other much more extensive modifications are also planned (for example to cloud microphysics, and the cloud extent, phase, and overlap decompositions provided to the radiation). There are many fascinating and important opportunities for improving this tool, and using it for better understanding of the Earth System. We invite use of this new tool and participation by on our science project by members of the research community. More information on this opportunity is provided in the next paragraph.

Code and data availability

The E3SM project, code, simulation configurations, model output, and tools to work with the output are described at https://e3sm.org. Instructions on how to get started running E3SM and its components are available at https://e3sm.org/model/running-e3sm/e3sm-quick-start. All model codes may be accessed on the GitHub repository at https://github.com/E3SM-Project/E3SM. CLUBB is available following free registration at

https://carson.math.uwm.edu/larson-group/clubb_site/signup/, The specific CLUBB variant use here code can be cloned from https://github.com/larson-group/clubb_release/tree/20141204_Rasch_et_al_2019_paper. Model output for original model simulation data are accessible through the DOE Earth System Grid Federation at https://esgf-node.llnl.gov/projects/e3sm. Specific climatologies used in this study can be obtained at http://portal.nersc.gov/archive/home/ACME/www/pjr/EAMv1_overview_data.

Acknowledgments

Phil Rasch and Shaocheng Xie co-led the model development team of over most of its development cycle, with help from Steve Klein and Peter Caldwell during early stages of the project. Many members of the DOE leadership and scientific community influenced the model evolution through formal and informal conversations about model characteristics, strengths, deficiencies, and opportunities for improvement. We are also grateful to many members of the CESM team (particularly the Atmospheric Model Working Group, the Software Engineering Working Group, and Bill Large) for their collaboration as E3SM made the transition to an independent modeling activity, and we continue to appreciate and benefit from discussions with that team. Special thanks to Michael Brunke, Bill Collins, Andrew Gettelman, Michael Prather, Steve Smith, Guang Zhang, Minghua Zhang, and Xubin Zeng for their insight about the model, parameterizations, simulations, input fields that drive the model and discussions of opportunities for improvement.

This research was supported as part of the Energy Exascale Earth System Model (E3SM) project (doi:10.11578/E3SM/dc.20180418.36), funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. The E3SM model code and input data is available at https://e3sm.org. This research used resources of:

- The National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231.
- 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
- The Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.
- A high-performance computing cluster provided by the BER Earth System Modeling program and operated by the Laboratory Computing Resource Center at Argonne National Laboratory
- The Pacific Northwest National Laboratory Institutional Computing (PIC) program.

Work at LLNL was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. Sandia National Laboratories is a multi-mission laboratory managed and operated by NTESS, 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. The Pacific Northwest National Laboratory is operated for the U.S. DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830. J.-H. Yoon was supported by National Research Foundation grant NRF_2017R1A2b4007480.

References

Abdul-Razzak, H., and S. J. Ghan. A Parameterization of Aerosol Activation: 2. Multiple Aerosol Types. *Journal of Geophysical Research* 105, no. D5 (March 1, 2000): 6837–44. https://doi.org/10.1029/1999JD901161.

Adler, R.F., G.J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, and P. Arkin, 2003: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hydrometeor.*, 4,1147-1167

Anstey, J. A. and T.G. Shepherd, T. G. 2014: High-latitude influence of the quasi-biennial oscillation. Q.J.R. Meteorol. Soc., 140, 1–21. doi:10.1002/qj.2132.

Bacmeister, J. T., Wehner, M. F., Neale, R. B., Gettelman, A., Hannay, C., Lauritzen, P. H., et al. (2013). Exploratory High-Resolution Climate Simulations using the Community Atmosphere Model (CAM). *Journal of Climate*, 131217115439008. https://doi.org/10.1175/JCLI-D-13-00387.1

Bauer, M., & Del Genio, A. D. (2006). Composite Analysis of Winter Cyclones in a GCM: Influence on Climatological Humidity. *Journal of Climate*, *19*(9), 1652–1672. https://doi.org/10.1175/JCLI3690.1

Beres, J. H., 2004: Gravity wave generation by a three-dimensional thermal forcing. *J. Atmos. Sci*., 61, 1805–1815.

Bodas-Salcedo, A., et al., 2011: COSP: Satellite simulation software for model assessment, *Bull. Amer. Meteor. Soc*., 92, 1023–1043.

Bogenschutz, P. A., Gettelman, A., Morrison, H., Larson, V. E., Craig, C., and Schanen, D. P., 2013: Higher-order turbulence closure and its impact on climate simulations in the Community Atmosphere Model, *Journal of Climate*, 26, 9655–9676.

Boyle, J., and S. A. Klein, 2010: Impact of horizontal resolution on climate model forecasts of tropical convection and diabatic heating for the TWP-ICE period. *J. Geophys. Res*., 115, D23113, doi: 10.1029/2010JD014262.

Boucher, O., et al. (2013), Clouds and Aerosols, in Climate Change 2013: *The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker, et al., pp. 571–658, Cambridge Univ. Press, Cambridge, and New York, doi:10.1017/CBO9781107415324.016.

Burrows, S. M., Ogunro, O., Frossard, A. A., Russell, L. M., Rasch, P. J., and Elliott, S. M.: A physically based framework for modeling the organic fractionation of sea spray aerosol from bubble film Langmuir equilibria, *Atmos. Chem. Phys.,* 14, 13601-13629, doi:10.5194/acp-14-13601-2014, 2014.

Burrows, S. M., Easter, R., Liu, X., Ma, P.-L., Wang, H., Elliott, S. M., et al. (2018). OCEANFILMS sea-spray organic aerosol emissions; Part 1: implementation and impacts on clouds. *Atmospheric Chemistry and Physics Discussions*, 1–27. https://doi.org/10.5194/acp-2018-70

Butchart, N., Anstey, J. A., Hamilton, K., Osprey, S., McLandress, C., Bushell, A. C., et al. (2018). Overview of experiment design and comparison of models participating in phase 1 of the SPARC Quasi-Biennial Oscillation initiative (QBOi). *Geoscientific Model Development*, *11*(3), 1009–1032. https://doi.org/10.5194/gmd-11-1009-2018

Charron, M., and E. Manzini, 2002: Gravity waves from fronts: Parameterization and middle atmosphere response in a general circulation model. *J. Atmos. Sci.,* 59, 923–941.

Collins, W. D., P. J. Rasch, B. A. Boville, J. J. Hack, J. R. McCaa, D. L. Williamson, B. P. Briegleb, C. M. Bitz, S-J. Lin, and M. Zhang, 2006. "The Formulation and Atmospheric Simulation of the Community Atmosphere Model Version 3 (CAM3)." *Journal of Climate* 19, no. 11, 2144–2161.

Cariolle, D., A. Lasserrebigorry, J. F. Royer, and J. F. Geleyn (1990), A general circulation model simulation of the springtime Antarctic Ozone decrease and its impact on midlatitudes, *Journal of Geophysical Research-Atmospheres*, 95(D2), 1883-1898, doi:10.1029/*JD095iD02p01883*.

Dennis, J. Edwards, K. J. Evans, O. Guba, P.H. Lauritzen, A.A. Mirin, A. St-Cyr, M. A. Taylor, P. H. Worley. CAM-SE: A scalable spectral element dynamical core for the Community Atmosphere Model. *International Journal of High Performance Computing Applications*, 26(1): 74-89, 2012.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. Quarterly Journal of the Royal Meteorological Society, 137(656), 553–597. https://doi.org/10.1002/qj.82

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., et al. (2011). The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. *Journal of Climate*, *24*(13), 3484–3519. https://doi.org/10.1175/2011JCLI3955.1

E3SM Project, DOE. Energy Exascale Earth System Model. Computer Software. https://github.com/E3SM-Project/E3SM.git. 23 Apr. 2018. Web. doi:10.11578/E3SM/dc.20180418.36.

Elliott, S. "Dependence of DMS Global Sea-air Flux Distribution on Transfer Velocity and Concentration Field Type." *Journal of Geophysical Research* 114, no. G2 (April 2, 2009). https://doi.org/10.1029/2008JG000710.

Evans, K. J., P. H. Lauritzen, S. K. Mishra, R. Neale, M.A. Taylor, J.J. Tribbia. AMIP Simulation with the CAM4 Spectral Element Dynamical Core. J. Climate, 26:689-709, 2013.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor. "Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization." *Geoscientific Model Development* 9, no. 5 (May 26, 2016): 1937–58. https://doi.org/10.5194/gmd-9-1937-2016.

Flanner, M. G., Liu, X., Zhou, C., Penner, J. E., and Jiao, C.: Enhanced solar energy absorption by internally-mixed black carbon in snow grains, *Atmos. Chem. Phys.*, 12, 4699-4721, doi:10.5194/acp-12-4699-2012, 2012.

Flanner, M. G. "Arctic Climate Sensitivity to Local Black Carbon." *Journal of Geophysical Research: Atmospheres* 118, no. 4 (February 27, 2013): 1840–51. https://doi.org/10.1002/jgrd.50176.

Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka, 2013: Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models, *J. Geophys. Res. Atmos.*, 118, 1139–1150, doi:10.1002/jgrd.50174.

Forster, P. M., Richardson, T., Maycock, A. C., Smith, C. J., Samset, B. H., Myhre, G., et al. (2016). Recommendations for diagnosing effective radiative forcing from climate models for CMIP6. *Journal of Geophysical Research: Atmospheres*, *121*(20), 12,460-12,475. https://doi.org/10.1002/2016JD025320

Gelaro, R., McCarty, W., Suárez, M. J., Todling, R., Molod, A., Takacs, L., et al. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419–5454. https://doi.org/10.1175/JCLI-D-16-0758.1

Geller, M. A., T. Zhou, D. Shindell, R. Ruedy, I. Aleinov, L. Nazarenko, N. L. Tausnev, M. Kelley, S. Sun, Y. Cheng, R. D. Field and G. Faluvegi (2016): Modeling the QBO—Improvements resulting from higher-model vertical resolution, *Journal of Advances in Modeling Earth Systems*, 8 (3), 1092-1105. https://doi.org/10.1002/2016MS000699

Gettelman, A. and Morrison, H.: Advanced two-moment bulk microphysics for global models. Part I: Off-line tests and comparison with other schemes, *Journal of Climate*, 28, 1268–1287, 2015.

Ghan, S. J., 2013: Technical note: Estimating aerosol effects on cloud radiative forcing. *Atmos. Chem Phys.*, 13, 9971–9974, doi:10.5194/acp-13-9971-2013.

Ghan, S., M. Wang, S. Zhang, S. Ferrachat, A. Gettelman, J. Griesfeller, Z. Kipling, U. Lohmann, H. Morrison, D. Neubauer, D. G. Partridge, P. Stier, T. Takemura, H. Wang, and K. Zhang, 2016: Challenges in constraining anthropogenic aerosol effects on cloud radiative forcing using present-day spatiotemporal variability. *Proc. Nat. Acad. Sci.,* 113, 5804–5811, www.pnas.org/cgi/doi/10.1073/pnas.1514036113.

Gleckler, P., C. Doutriaux, P. Durack, K. Taylor, Y. Zhang, D. Williams, E. Mason, and J. Servonnat (2016), A more powerful reality test for climate models, EOS, 97, doi:10.1029/2016eo051663.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux (2008), Performance metrics for climate models, Journal of Geophysical Research, 113 (D6), doi:10.1029/1488 2007jd008972.

Golaz, J.-C., Larson, V. E., and Cotton, W. R.: A PDF-Based Model for Boundary Layer Clouds. Part I: Method and Model Description., *Journal of Atmospheric Sciences*, 59, 3540–3551, doi:10.1175/1520-0469(2002)059<3540:APBMFB>2.0.CO;2, 2002.

Golaz, J.-C., et al, 2018: The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *JAMES*, submitted.

Guba, O, M. A. Taylor, P. A. Ullrich, J. R. Overfelt, M. N. Levy The spectral element method (SEM) on variable resolution grids: evaluating grid sensitivity and resolution-aware numerical viscosity, *Geoscientific Model Development*, 7, 2803-2816, 2014.

Guba, O., M. A. Taylor, A. St-Cyr. Optimal limiters for the spectral element method. *Journal of Computational Physics* 267 (2014) 176-195.

Guba, O., J. Lorenz, D. Sulsky. On well-posedness of the viscous-plastic sea-ice model. *Journal of Physical Oceanography*. Vol. 43, No. 10., 2194-2209, 2013.

Hack, J., B. Boville, B. Briegleb, J. Kiehl, and D. Williamson. "Description of the NCAR Community Climate Model (CCM2)." *UCAR/NCAR Technical Note*, 1993. https://doi.org/10.5065/d6qz27xv.

Harrop, B. E., Ma, P.-L., Rasch, P. J., Neale, R. B., & Hannay, C. (2018). The Role of Convective Gustiness in Reducing Seasonal Precipitation Biases in the Tropical West Pacific. *Journal of Advances in Modeling Earth Systems*, *10*(4), 961–970. https://doi.org/10.1002/2017MS001157

Haynes, P. H., McIntyre, M. E., Shepherd, T. G., Marks, C. J., & Shine, K. P. (1991). On the "Downward Control" of Extratropical Diabatic Circulations by Eddy-Induced Mean Zonal Forces. *Journal of the Atmospheric Sciences*, *48*(4), 651–678. https://doi.org/10.1175/1520-0469(1991)048<0651:OTCOED>2.0.CO;2

Hoesly, R. M., S. J. Smith, L. Feng, Z. Klimont, G. Janssens-Maenhout, T. Pitkanen, J. J. Seibert, et al. "Historical (1750–2014) Anthropogenic Emissions of Reactive Gases and Aerosols

from the Community Emissions Data System (CEDS)." *Geosci. Model Dev.* 11, no. 1 (January 29, 2018): 369–408. https://doi.org/10.5194/gmd-11-369-2018.

Hoskins, B. J., 1982: The mathematical theory of frontogenesis. *Annu. Rev. Fluid Mech*., 14, 131–151.

Hsu, J., and M. J. Prather (2009), Stratospheric variability and tropospheric ozone, *Journal of Geophysical Research-Atmospheres*, 114, 15, doi:10.1029/2008jd010942.

Huffman, G. J., Adler, R. F., Morrissey, M. M., Bolvin, D. T., Curtis, S., Joyce, R., McGavock, B. and Susskind, J.: Global Precipitation at One-Degree Daily Resolution from Multisatellite Observations, J. Hydrometeorol., 2(1), 36–50, doi:10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2, 2001.

Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu. "Improving the Global Precipitation Record: GPCP Version 2.1." *Geophysical Research Letters* 36, no. 17 (2009).

Hurrell, J. W., J. J. Hack, D. Shea, J. M. Caron, and J. Rosinski (2008), A New Sea Surface Temperature and Sea Ice Boundary Dataset for the Community Atmosphere Model, *J. Climate*, 21(19), 5145–5153, doi:10.1175/2008JCLI2292.1.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *J. Geophys. Res.-Atmos*., 113, D13103, doi:10.1029/2008jd009944, 2008.

Jung, T., Miller, M. J., Palmer, T. N., Towers, P., Wedi, N., Achuthavarier, D., et al. (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. https://doi.org/10.1175/JCLI-D-11-00265.1

Kato, S., Loeb, N. G., Rose, F. G., Doelling, D. R., Rutan, D. A., Caldwell, T. E., et al. (2013). Surface Irradiances Consistent with CERES-Derived Top-of-Atmosphere Shortwave and Longwave Irradiances. *Journal of Climate*, *26*(9), 2719–2740. https://doi.org/10.1175/JCLI-D-12-00436.1

Khairoutdinov M, and Y. Kogan, 2000: A new cloud physics parameterization in a large-eddy simulation model of marine stratocumulus. *Mon. Weather Rev*. 128(1) 229–243, doi:10.1175/1520-0493.

Larson, V. E., & Golaz, J. C. (2005). Using probability density functions to derive consistent closure relationships among higher-order moments. *Monthly Weather Review*, 133, 1023-1042.

Larson, V. E., Golaz, J. C., & Cotton, W. R. (2002). Small-scale and mesoscale variability in cloudy boundary layers: Joint probability density functions. *Journal of the Atmospheric Sciences*, 59, 3519-3539.

Larson, V. E. (2017). CLUBB-SILHS: A parameterization of subgrid variability in the atmosphere. *arXiv*:1711.03675.

Lauritzen, P. H., J. T. Bacmeister, P. F. Callaghan, and M. A. Taylor. "NCAR_Topo (v1.0): NCAR Global Model Topography Generation Software for Unstructured Grids." *Geoscientific Model Development* 8, no. 12 (2015): 3975–86. https://doi.org/10.5194/gmd-8-3975-2015.

Lauritzen, P., P. A. Ullrich, C. Jablonowski, P. A. Bosler, D. Calhoun, A. J. Conley, T. Enomoto, L. Dong, S. Dubey, O. Guba, A. B. Hansen, E. Kaas, J. Kent, J.-F. Lamarque, M. J. Prather, D. Reinert, V. V. Shashkin, W. C. Skamarock, B. Sørensen, M. A. Taylor, and M.A. Tolstykh, 2014. A standard test case suite for two-dimensional linear transport on the sphere: results from a collection of state-of-the-art schemes. *Geosci. Model Dev.*, 7, 105-145.

Lin, Y., L. J. Donner, J. Petch, P. Bechtold, J. Boyle, S. A. Klein, T. Komori, K. Wapler, M. Willett, X. Xie, M. Zhao, S. Xie, S. A. MaFarlane, C. Schumacher, 2012: TWP-ICE global atmospheric model intercomparison: Convection responsiveness and resolution impact. *J. Geophys. Res.*, 117, D09111, doi: 10.1029/2011JD017018.

Liu, X., S. Xie, J. Boyle, S. A. Klein, X. Shi, Z. Wang, W. Lin, S. J. Ghan, M. Earle, P. S. K. Liu, and A. Zelenyuk1, 2011: Testing Cloud Microphysics Parameterizations in NCAR CAM5 with ISD AC and M-PACE Observations. *J. Geophys. Res.,* 116, D00T11, doi: 10.1029/2011JD015889.

Liu, X., P.-L. Ma, H. Wang, S. Tilmes, B. Singh, R.C. Easter, S.J. Ghan, and P.J. Rasch, 2015: Description and evaluation of a new 4-mode version of Modal Aerosol Module (MAM4) within version 5.3 of the Community Atmosphere Model, *Geosci. Model Dev*., 9, 505–522, doi:10.5194/gmd-9-505-2016.

Lindzen, R. S., & Fox-Rabinovitz, M. (1989). Consistent Vertical and Horizontal Resolution. *Monthly Weather Review*, *117*(11), 2575–2583. https://doi.org/10.1175/1520-0493(1989)117<2575:CVAHR>2.0.CO;2

Loeb, N. G., B. A. Wielicki, D. R. Doelling, G. L. Smith, D. F. Keyes, S. Kato, N. Manalo-Smith, and T. Wong, 2009: Toward optimal closure of the Earth's top-of-atmosphere radiation budget. *J. Climate*, 22, 748–766, doi:10.1175/2008JCLI2637.1.

Lott, F., & Miller, M. J. (1997). A new subgrid-scale orographic drag parametrization: Its formulation and testing. *Quarterly Journal of the Royal Meteorological Society*, *123*(537), 101–127. https://doi.org/10.1002/qj.49712353704

McFarlane, N. A., 1987: The effect of orographically excited wave drag on the general circulation of the lower stratosphere and troposphere. *J. Atmos. Sci*., 44, 1775–1800.

Ma, P.-L., Rasch, P. J., Wang, M., Wang, H., Ghan, S. J., Easter, R. C., et al. (2015). How does increasing horizontal resolution in a global climate model improve the simulation of aerosol-cloud interactions? *Geophysical Research Letters*, *42*(12), 5058–5065. https://doi.org/10.1002/2015GL064183

Ma, P.-L., Rasch, P. J., Chepfer, H., Winker, D. M., & Ghan, S. J. (2018). Observational constraint on cloud susceptibility weakened by aerosol retrieval limitations. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-05028-4

Ma, H.-Y., S. Xie, S. A. Klein, K. D. Williams, J. S. Boyle, S. Bony, H. Douville, S. Fermepin, B. Medeiros, S. Tyteca, M. Watanabe, and D. Williamson, 2014: On the correspondence between mean forecast errors and climate errors in CMIP5 models. *J. Clim.*, 27, 1781-1798, doi:10.1175/JCLI-D-13-00474.1.

Ma, H.-Y., S. A. Klein, S. Xie, C. Zhang, S. Tang, Q. Tang, C. J. Morcrette, K. van Weverberg, J. Petch, M. Ahlgrimm, L. K. Berg, F. Cheruy, J. Cole, R. Forbes, W. I. Gustafson, M. Huang, Y. Liu, W. Merryfield, Y. Qian, R. Roehrig, Y.-C. Wang, 2018, CAUSES: On the Role of Surface Energy Budget Errors to the Warm Surface Air Temperature Error Over the Central United States, *J. Geophys. Res.*, https://doi.org/10.1002/2017JD027194.

Martin, G. M., Bellouin, N., Collins, W. J., Culverwell, I. D., Halloran, P. R., Hardiman, S. C., et al. (2011). The HadGEM2 family of Met Office Unified Model climate configurations. *Geoscientific Model Development*, *4*(3), 723–757. https://doi.org/10.5194/gmd-4-723-2011

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough. "Radiative Transfer for Inhomogeneous Atmospheres: RRTM, a Validated Correlated-k Model for the Longwave." *Journal of Geophysical Research: Atmospheres* 102, no. D14 (July 27, 1997): 16663–82. https://doi.org/10.1029/97JD00237.

Mote, P. W., Holton, J. R., Russell, J. M., & Boville, B. A. (1993). A comparison of observed (Haloe) and modeled (CCM2) methane and stratospheric water vapor. *Geophysical Research Letters*, *20*(14), 1419–1422. https://doi.org/10.1029/93GL01764

Neale, R. B., J. Richter, S. Park, P. H. Lauritzen, S. J. Vavrus, P. J. Rasch, and M. Zhang. "The Mean Climate of the Community Atmosphere Model (CAM4) in Forced SST and Fully Coupled Experiments." *Journal of Climate*, no. 2013 (2013). http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00236.1.

Neale, R. B., C. C. Chen, A. Gettelman, P. H. Lauritzen, S. Park, D. L. Williamson, A. J. Conley, et al. "Description of the NCAR Community Atmosphere Model: CAM5.0." NCAR Technical Note, Boulder, Colorado, USA: National Center for Atmospheric Research, 2012.

O'Brien, T. A., Collins, W. D., Kashinath, K., Rübel, O., Byna, S., Gu, J., et al. (2016). Resolution dependence of precipitation statistical fidelity in hindcast simulations. *Journal of Advances in Modeling Earth Systems*, *8*(2), 976–990. https://doi.org/10.1002/2016MS000671

Ogunro, O., S. M. Burrows, S. Elliott, A. A. Frossard, F. Hoffman, R. T. Letscher, J. K. Moore, L. M. Russell, S. Wang, and O. W. Wingenter. "Global Distribution and Surface Activity of Macromolecules in Offline Simulations of Marine Organic Chemistry." *Biogeochemistry* 126, no. 1 (November 1, 2015): 25–56. https://doi.org/10.1007/s10533-015-0136-x.

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., et al. (2007). The JRA-25 Reanalysis. *Journal of the Meteorological Society of Japan. Ser. II*, *85*(3), 369–432. https://doi.org/10.2151/jmsj.85.369

Orlanski, I. (2008). Mesoscale Dynamics. *Eos, Transactions American Geophysical Union*, *89*(42), 408–408. https://doi.org/10.1029/2008EO420004

Park, S., C. S. Bretherton, and P. J. Rasch. "Integrating Cloud Processes in the Community Atmosphere Model, Version 5." *Journal of Climate*, August 11, 2014, 140811131149007. https://doi.org/10.1175/JCLI-D-14-00087.1.

Phillips, T. J., and Coauthors (2004). Evaluating parameterizations in general circulation models: Climate simulation meets weather prediction. *Bull. Amer. Meteor. Soc*., 85, 1903–1915.

Qian Y., H. Wang, R. Zhang, M.G. Flanner, and P.J. Rasch. 2014. "A Sensitivity Study on Modeling Black Carbon in Snow and its Radiative Forcing over the Arctic and Northern China." *Environmental Research Letters* 9, no. 6:Article No. 064001. doi:10.1088/1748-9326/9/6/064001.

Qian, Y., Yan, H., Hou, Z., Johannesson, G., Klein, S., Lucas, D., et al. (2015). Parametric sensitivity analysis of precipitation at global and local scales in the community atmosphere model CAM5. *Journal of Advances in Modeling Earth Systems,* 7, 382–411. https://doi.org/10.1002/2014ms000354.

Qian Y, H Wan, B Yang, J-C Golaz, B Harrop, Z Hou, VE Larson, LR Leung, G Lin, W Lin, P-L Ma, H-Y Ma, P Rasch, B Singh, H Wang, S Xie, and K Zhang. 2018. "Parametric Sensitivity and Uncertainty Quantification in the Version 1 of E3SM Atmosphere Model Based on Short Perturbed Parameter Ensemble Simulations." *Journal of Geophysical Research: Atmospheres* 123. https://doi.org/10.1029/2018JD028927.

Roeckner, E., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., et al. (2006). Sensitivity of Simulated Climate to Horizontal and Vertical Resolution in the ECHAM5 Atmosphere Model. *Journal of Climate*, *19*(16), 3771–3791. https://doi.org/10.1175/JCLI3824.1

Richter., J.H., F. Sassi, and R. R. Garcia, 2010: Toward a Physically Based Gravity Wave Source Parameterization in a General Circulation Model, *J. Atmos. Sci*., **67**, 136 – 156.

Richter, J. H., and P. J. Rasch. "Effects of Convective Momentum Transport on the Atmospheric Circulation in the Community Atmosphere Model, Version 3." *Journal of Climate* 21, no. 7 (April 2008): 1487–99. https://doi.org/10.1175/2007JCLI1789.1.

Richter, J. H., A. Solomon, and J. T. Bacmeister (2014), On the simulation of the quasi-biennial oscillation in the Community Atmosphere Model, version 5, *J. Geophys. Res. Atmos.*, 119, 3045–3062, doi:10.1002/2013JD021122.

Richter, J. H., Solomon, A., & Bacmeister, J. T. (2014b). Effects of vertical resolution and nonorographic gravity wave drag on the simulated climate in the Community Atmosphere Model, version 5. *Journal of Advances in Modeling Earth Systems*, *6*(2), 357–383. https://doi.org/10.1002/2013MS000303

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, R., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G-K, Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications, *J. Climate,* 24, 3624–3648, 2011.

Ringer, M. A., Andrews, T., & Webb, M. J. (2014). Global-mean radiative feedbacks and forcing in atmosphere-only and coupled atmosphere-ocean climate change experiments. *Geophysical Research Letters*, *41*(11), 4035–4042. https://doi.org/10.1002/2014GL060347

Scinocca, J. F., & McFarlane, N. A. (2004). The variability of modeled tropical precipitation. *Journal of the Atmospheric Sciences*, *61*(16), 1993–2015.

Shrivastava, M., et al., 2015: Global transformation and fate of SOA: Implications of low-volatility SOA and gas-phase fragmentation reactions, *J. Geophys. Res.-Atmos.,* 120, 4169-4195, doi:10.1002/2014jd022563.

Son, S.-W., Y. Lim, C. Yoo, H. H. Hendon, and J. Kim, 2017: Stratospheric control of Madden–Julian oscillation. J. Climate, 30, 1909–1922, doi:10.1175/JCLI-D-16-0620.1.

Tang, Q., S. A. Klein, S. Xie, W. Lin, J.-C. Golaz, E. L. Roesler, M. A. Taylor, P. J. Rasch, D. C. Bader, L. K. Berg, P. Caldwell, S. Giangrande, R. Neale, Y. Qian, L. D. Riihimaki, C. S. Zender, Y. Zhang, and X. Zheng, 2019: Regionally refined capability in E3SM Atmosphere Model Version 1 (EAMv1) and applications for high-resolution modelling, *Geosci. Model Dev*., submitted.

Taylor, M. A. (2011), Conservation of mass and energy for the moist atmospheric primitive equations on unstructured grids, in Numerical Techniques for Global Atmospheric Models, *Lecture Notes Comput. Sci. Eng.,* vol. 80, edited by P. H. Lauritzen et al., pp. 357–380, Springer, Heidelberg, Germany.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl. "An Overview of CMIP5 and the Experiment Design." *Bulletin of the American Meteorological Society* 93, no. 4 (April 2012): 485–98. https://doi.org/10.1175/BAMS-D-11-00094.1.

Taylor, K. E., D. Williamson, and F. Zwiers, 2000: The sea surface temperature and sea-ice concentration boundary conditions of AMIP II simulations. *PCMDI Rep. 60,* 20 pp. pdf file available at http://www-pcmdi.llnl.gov/publications/pdf/60.pdf

van Marle, M. J. E., Kloster, S., Magi, B. I., Marlon, J. R., Daniau, A.-L., Field, R. D., Arneth, A., Forrest, M., Hantson, S., Kehrwald, N. M., Knorr, W., Lasslop, G., Li, F., Mangeon, S., Yue, C., Kaiser, J. W., and van der Werf, G. R., 2017: Historic global biomass burning emissions for CMIP6 (BB4CMIP) based on merging satellite observations with proxies and fire models (1750–2015), *Geosci. Model Dev.*, 10, 3329-3357, https://doi.org/10.5194/gmd-10-3329-2017.

Van Weverberg, K., C. J. Morcrette, H.-Y. Ma, S. A. Klein, and J. C. Petch, 2015: Using regime analysis to identify the contribution of clouds to surface temperature errors in weather and climate models. *Quart. J. Roy. Met. Soc.*, 141, 3190-3206, doi: 10.1002/qj.2603.

Wang, H., Easter, R. C., Rasch, P. J., Wang, M., Liu, X., Ghan, S. J., Qian, Y., Yoon, J.-H., Ma, P.-L., and Vinoj, V., 2013: Sensitivity of remote aerosol distributions to representation of cloud–aerosol interactions in a global climate model, *Geosci. Model Dev.*, 6, 765-782, doi:10.5194/gmd-6-765-2013.

Wang M, et al., 2012: Constraining cloud lifetime effects of aerosol using A-Train satellite observations. *Geophys. Res. Lett.* 39 15, L15709, doi:10.1029/2012GL052204.

Wang, B., & Ding, Q. (2008). Global monsoon: Dominant mode of annual variation in the tropics. Dynamics of Atmospheres and Oceans, 44, 165–183. https://doi.org/10.1016/j.dynatmoce.2007.05.002

Wang, Y., Liu, X., Hoose, C., and Wang, B., 2014: Different contact angle distributions for heterogeneous ice nucleation in the Community Atmospheric Model version 5, *Atmospheric Chemistry and Physics*, 14, 10 411–10 430, https://doi.org/10.5194/acp-14-10411-2014.

Wang, S., S. Elliott, M. Maltrud, and P. Cameron-Smith. "Influence of Explicit *Phaeocystis* Parameterizations on the Global Distribution of Marine Dimethyl Sulfide: Explicit *Phaeocystis* in a DMS Model." *Journal of Geophysical Research: Biogeosciences* 120, no. 11 (November 2015): 2158–77. https://doi.org/10.1002/2015JG003017.

Wan, H., Rasch, P. J., Zhang, K., Qian, Y., Yan, H., and Zhao, C.: Short ensembles: an efficient method for discerning climate-relevant sensitivities in atmospheric general circulation models, *Geosci. Model Dev.,* 7, 1961-1977, https://doi.org/10.5194/gmd-7-1961-2014, 2014.

Wehner, M. F., Reed, K. A., Li, F., Prabhat, Bacmeister, J., Chen, C.-T., et al. (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. https://doi.org/10.1002/2013MS000276

Wheeler, M., G.N. Kiladis. Convectively Coupled Equatorial Waves: Analysis of Clouds and Temperature in the Wavenumber-Frequency Domain, *J. Atmos. Sci.*, 1999, 56: 374-399. NOAA interpolated OLR

Williams, K. D., Bodas-Salcedo, A., Deque, M., Fermepin, S., Medeiros, B., Watanabe, M., Jakob, C., Klein, S.A., Senior, C.A., and Williamson, D. L. (2013). The Transpose-AMIP II experiment and its application to the understanding of Southern Ocean cloud biases in climate models. *Journal of Climate*, 26(10), 3258-3274

Xie, S., M. Zhang, J. S. Boyle, R. T. Cederwall, G. L. Potter, and W. Lin, 2004: Impact of a revised convective triggering mechanism on Community Atmosphere Model, Version 2, simulations: Results from short-range weather forecasts. *J. Geophys. Res.*, 109, D14102, doi: 10.1029/2004JD004692.

Xie, S., J. Boyle, S. A. Klein, X. Liu, and S. Ghan (2008), Simulations of Arctic mixed-phase clouds in forecasts with CAM3 and AM2 for M-PACE, J. Geophys. Res., 113, D04211, doi: 10.1029/2007JD009225.

Xie, S., H.-Y. Ma, J. S. Boyle, S. A. Klein, and Y. Zhang, 2012: On the correspondence between short- and long-time-scale systematic errors in CAM4/CAM5 for the Year of Tropical Convection., *J. Clim.*, 25, 7937-7955, doi: 10.1175/JCLI-D-12-00134.1.

Xie, S., X. Liu, C. Zhao, and Y. Zhang (2013). Sensitivity of CAM5 Simulated Arctic Clouds and Radiation to Ice Nucleation, J. Climate. 26, 5981-5999, doi:10.1175/JCLI-D-12-00517.1.

Xie, S., Lin, W., Rasch, P. J., Ma, P.-L., Neale, R., Larson, V. E., et al. (2018). Understanding Cloud and Convective Characteristics in Version 1 of the E3SM Atmosphere Model. *Journal of Advances in Modeling Earth Systems*, *10*(10), 2618–2644. https://doi.org/10.1029/2018MS001350

Yang, Y., Wang, H., Smith, S. J., Zhang, R., Lou, S., Yun, Q., Ma, P.-L., and Rasch, P. J., 2018: Recent intensification of winter haze in China linked to foreign emissions and meteorology, *Sci. Rep.*, 8, 2107, https://doi.org/10.1038/s41598-018-20437-7.

Yoo, C., and S.-W. Son (2016), Modulation of the boreal wintertime Madden-Julian oscillation by the stratospheric quasi-biennial oscillation, Geophys. Res. Lett., 43, 1392–1398, doi:10.1002/2016GL067762.

Zarzycki, C. M., C. Jablonowski, D. R. Thatcher, and M. A. Taylor, 2015. "Effects of Localized Grid Refinement on the General Circulation and Climatology in the Community Atmosphere

Model." *Journal of Climate* 28, no. 7, 2777–2803. https://doi.org/10.1175/JCLI-D-14-00599.1.

Zhang, G. J., and N. A. McFarlane, 1995. "Sensitivity of Climate Simulations to the Parameterization of Cumulus Convection in the Canadian Climate Centre General Circulation Model." *Atmosphere-Ocean* 33, no. 3: 407–446.

Zhang, S., Wang, M., Ghan, S.J., Ding, A., Wang, H., Zhang, K., Neubauer, D., Lohmann, U., Ferrachat, S., Takeamura, T. and Gettelman, A., 2016. On the characteristics of aerosol indirect effect based on dynamic regimes in global climate models. *Atmospheric Chemistry and Physics*, *16*(5), pp.2765-2783.

Zhang, K., Rasch, P. J., Taylor, M. A., Wan, H., Leung, R., Ma, P.-L., Golaz, J.-C., Wolfe, J., Lin, W., Singh, B., Burrows, S., Yoon, J.-H., Wang, H., Qian, Y., Tang, Q., Caldwell, P., and Xie, S., 2018: Impact of numerical choices on water conservation in the E3SM Atmosphere Model version 1 (EAMv1), *Geosci. Model Dev.*, 11, 1971-1988, https://doi.org/10.5194/gmd-11-1971-2018.

Zhang, Y., Xie, S., Lin, W., Klein, S. A., Zelinka, M., Ma, P.-L., et al (2019). Evaluation of clouds in version 1 of the E3SM atmosphere model with satellite simulators. Journal of Advances in Modeling Earth Systems, 11. https://doi.org/10.1029/2018MS001562.

Zhang, Y., S. Xie, S.A. Klein, R. Marchand, P. Kollias, E.E. Clothiaux, W. Lin, K. Johnson, D. Swales, A. Bodas-Salcedo, S. Tang, J.M. Haynes, S. Collis, M. Jensen, N. Bharadwaj, J. Hardin, and B. Isom, 2018: The ARM Cloud Radar Simulator for Global Climate Models: Bridging Field Data and Climate Models. *Bull. Amer. Meteor. Soc.,***99**, 21-26, https://doi.org/10.1175/BAMS-D-16-0258.1

Appendix A

A.1 Tunable Parameters used by EAMv1

Table of tunable parameters that that have been changed compared to the value described in the original papers describing the basic parameterization. When the parameter value is resolution dependent the both Low (LR) and High (HR) resolutions values are listed.

Tunable parameter	Description	LR value	HR value	Comments/nominal value
dtime	physics timestep	1800	900	
seasalt_emis_scale	sea salt emission scale factor	0.85		reduced from 1.35, tuned for global AOD
raytau0	Rayleigh friction	5.0		increased from 0.
cldfrc2m_rhmaxi	Maximum RH for ice cloud fraction	1.05		Reduced from 1.2.
cldfrc_dp1	Parameter for deep convective cloud fraction	0.045	0.039	Reduced from 0.1.
dust_emis_fact	Dust emission scale factor (denominator)	2.05	2.50	Increased from 0.35 to reduce dust emission

clubb_c1	Damping coefficient for w'2	1.335	1.5	Reduce from 2.5 to increase w'2, favoring low cloud formation
clubb_c14	Damping coefficient for u'2 and v'2	1.06	1.75	Increase from 1.0 to reduce TKE and low clouds
clubb_c2rt	Damping coefficient for total water variance	1.75		Increase from 1.0
clubb_c8	Damping coefficient for w'3	4.3	4.73	Increased from 3.0 to reduce w'3, favoring symmetric convection and low cloud formation.
clubb_c_k10	ratio of eddy diffusivity of wind to heat	0.3		0.6
clubb_ice_deep	assumed ice condensate radius detrained from ZM	16	12	25
clubb_liq_deep	assumed liquid condensate radius detrained from ZM	8		10
effgw_beres	frontal gravity wave drag coefficient	0.4		0.1
effgw_oro	orographic gravity wave drag coefficient	0.25		0.1
taubgnd	background source strength parameter (for gravity wave drag)	2.5e-3		1.5e-3
do_tms	turbulent mountain stress parameterization	false		true
ice_sed_ai	ice particle fall speed parameter	500		700
micro_mg_berg_eff_factor	Scale factor for WBF process	0.1		Nominal setting of 1, in MG
so4_sz_thresh_icenuc	Aitken mode SO4 size threshold used for homogeneous ice nucleation	0.1e-6		0.05e-6
micro_mg_accre_enhan_fac	enhancement factor for accretion	1.5		1
zmconv_alfa	Maximum downdraft fraction	0.1	0.2	0.1
zmconv_c0_lnd	Conversion from cloud water to rain water over land	0.007	0.0035	0.0059
zmconv_c0_ocn	Conversion from cloud water to rain water over ocean	0.007	0.0043	0.045
zmconv_cape_cin	# of layers allowed for negative CAPE	1		5
zmconv_dmpdz	Parcel fractional mass entrainment rate	-0.7e-3	-0.2e-3	-1.e-3
zmconv_ke	coefficient for evaporation of convective precipitation	5.e-6	6e-6	1.e-6
zmconv_mx_bot_lyr_adj	Lowest launching level	2		0
zmconv_tiedke_add	initial parcel	0.8		0.5

	buoyancy			
k	coefficient for autoconversion	30500		1350
a	exponent for Qc	3.19		2.47
b	exponent for Nc	-1.2		-1.79
Note 2: clubb_rainevap_turb and clubb_cloudtop_cooling are set to false.				
Note 3: dcs parameter has been replaced by a temperature-dependent parameterization				

Tables:

Table 1a: Information related to model resolution and cost on Cori (KNL nodes). Note that CAM5 by default uses a finite volume dynamical core and the numbers reported here are for the spectral element dynamical core, and the discussion regarding the cost calculation on Cori.

Resolution identifier	Horizontal Resolution	Vertical Resolution	Number of elements	Throughput: Sim Years Per Day (# Cori-KNL nodes)	Cost in core-hours per sim year (68 cores/node)
CAM5 Ne30	1^{o}	30	5400	9 SYPD (40N)	7,000
EAM Ne30	1^{o}	72	5400	6 SYPD (81N)	22,000
EAM Ne120	$1/4^{o}$	72	86400	2 SYPD (675N)	551,000
EAM CONUS	$1/4^{o}$ to 1^{o}	72	9905	1.7 SYPD (88N)	84,000

Table 1b: Information related to model resolution and cost on Edison.

Resolution identifier	Horizontal Resolution	Vertical Resolution	Number of elements	Throughput: Sim Years Per Day (# Edison nodes)	Cost in core-hours per sim year (24 core/node)
CAM5 Ne30	1^{o}	30	5400	41 SYPD (113N)	1,600
EAM Ne30	1^{o}	72	5400	10 SYPD (113N)	6,300
EAM Ne120	$1/4^{o}$	72	86400	0.5 SYPD (200N)	265,000
EAM CONUS	$1/4^{o}$ to 1^{o}	72	9905	1.3 SYPD (113N)	25,000

Table 2: Description of model simulations performed for evaluation:

Simulation Name	Resolution, Analysis period	Description
F1850LR	100km, Years 2-11	Year 1850 SST, GHG, LU, LC, Aerosol sources, no volcanoes
F2000LR	100km, Years 2-11	Year 2000 SST. GHG, LU, LC, Aerosol sources, no volcanoes
F1850LR+4K	100km, Years 2-11	As in F1850LR, SST + 4K to assess cloud feedback
F2000LR/AF	100km, Years 2-11	As in F2000LR but SST, Sea-ice, Solar Cycle replaced by 1850 values
F2000LR/Aero	100km, Years 2-11	As in F2000LR but Aerosol emissions replaced by 1850 values
FAMIPLR	100km, 1980-2005	Time varying SST, aerosol sources, volcanoes, LU/LCC, GHG
FAMIPLR/AllF	100km, 1980-2005	As in FAMIPLR except aerosol sources, volcanoes, LU/LCC, and GHG fixed at 1850
FAMIPLR/AeroF	100km, 1980-2005	As in FAMIPLR except aerosol sources fixed at 1850

CAM5AMIP	100km, 1980-2005	CAM5 Control for comparison
F2000HR	25km, Years 2-11.	As in F2000LR except high resolution
F1850CONUS	100km + 25km, Years 2-5	CONUS simulation, Free-running
F2000CONUS	100km + 25km, Years 2-5	CONUS simulation, Free-running (nudged and CAPT special features)

Table 3: Global averages for some important climate quantities for EAMv1 AMIP simulations compared to CAM5 and (central estimates for) observations. Meteorological fields were compared to the Japan Meteorological Agency 25 year reanalysis. Top of Atmosphere Radiative flux estimates are drawn from the CERES-EBAF 2.8 dataset (Loeb et al, 2009). Precipitation from the Global Precipitation Climate Project (GPCP, Adler et al., 2003). Surface Radiative budgets from CERES-EBAF (Kato et al, 2012), AOD from Ma et al (2018). Observational estimates are not entered for liquid and ice water paths because model and observational estimates are sampled so differently that comparison is not very reliable.

Variable	Obs Estimate	CAM5	EAMv1LR
T10 (Ref Height Temp, K)	288.3 (JRA25)	287.7	288.2
LWCRE (Longwave Cloud Radiative Effect, W/m2)	26.1 (CERES-EBAF)	23.6	24.5
SWCRE (Shortwave Cloud Radiative Effect, W/m2)	-47.1 (CERES-EBAF)	-48.6	-49.3
All-sky Longwave (Top of Model vs Top of Atmosphere for obs, W/m2)	239.8 (CERES-EBAF)	234.1	239.4
All-sky Shortwave (Top of Model vs Top of Atmosphere for obs, W/m2)	240.5 (CERES-EBAF)	233.4	239.0
NET Flux at Top of Model (W/m2)	+0.85 (CERES-EBAF)	-0.7	-0.5
Surface Longwave Down (W/m2)	345.2 (EBAFS)	342.9	344.8
Surface Shortwave Down (W/m2)	186.6 (EBAFS)	182.7	184.6
Water Vapor Path	24.4 (JRA25)	25.3	24.4
Precipitation (mm/day)	2.7 (GPCP)	3.0	3.1
Liquid Water Path (g m^{-2})		43.8	53.4
Ice Water Path (g m^{-2})		17.2	11.3
Sensible Heat Flux (W/m2)	19.3 (JRA25)	17.9	19.2
Latent Heat Flux (W/m2)	87.9 (JRA25)	88.2	89.8
U@200hPa (m/s)	15.6 (JRA25)	16.3	15.7
Aerosol Optical Depth@512nm (unitless)	0.12 (GOCAP)	0.11	0.13

Table 4 Changes in globally averaged top of atmosphere fluxes (longwave allsky (FLNT), clear sky (FLNTC) and CRE (LWCRE), Shortwave allsky (FSNT), clearsky (FSNTC) and CRE(SWCRE), surface temperature (TS, total/land), Aerosol Optical Depth (AOD). C2, C3: Response to Total forcing (GHG, aerosols, and land use/land cover) C4, C5: Response to aerosol and aerosol precursor emissions. C6: response to SST changes. Estimates for CAM5 (C5) from Gettelman et al (2015) used the Ghan (2013) "clear clean" methodology that can produce differences O(0.1 W/m2) compared to the "clear" CRE estimates (A. Gettelman, private communication). Fluxes use typical sign convention for CESM models; both long and shortwave net fluxes are defined to be positive, the net flux is defined to be shortwave-longwave, and cloud forcing is defined positive for longwave and negative for shortwave.

C1	C2	C3	C4	C5	C6
Description of forcing estimnate	EAMv1 Responses to a change in all forcing agents in presence of 1850 SSTs and no volcanic aerosols	EAMv1 Response to a change in all forcing agents in presence of time varying SSTs, aerosols, and volcanic aerosols (1980-2004)	EAMv1 Response to aerosol emission changes in presence of time varying SSTs, aerosols, and volcanic aerosols (1980-2004)	CAM5.3 Response to changing aerosol emission	EAMv1 response to a +4K increase in SSTs
Cases or reference for calculation → / Change in field ↓	See Table 2 (cases (F2000LR/AF-F1850)	See table 2 of Golaz et al (2018) for description (cases amip_A3 - amip_1850allF_A3)	See table 2 of Golaz et al (2018) for description (cases amip_A3 - amip_1850aerof_A3)	CAM5.3 (AE w MG1)	See Table 2 (cases F1850LR+4k – F1850LE)
ΔFLNT	-3.03	-2.86	-0.72	~0.44	10.1
ΔFLNTC	-3.17	-2.87	-0.15		9.4
ΔLWCRE	-0.14	-0.01	0.57	0.44	-0.7
ΔFSNT	-1.77	-2.5	-2.58	-2.	3.7
ΔFSNTC	-0.63	-1.	-0.64		1.9
ΔSWCRE	-1.13	-1.51	-1.94	-2.02	1.8
ΔFTOA	1.26	0.36	-1.86	-1.59	-6.4
ΔTS/TSLAND	0.12/0.25	0.1/0.18	0.01/-0.11		4.49
ΔAOD	0.03	0.03	0.02		0.01

Figure captions

Figure 1: The regionally refined grid configuration for the atmosphere and land over the Continental United States (CONUS) and Tropical West Pacific in EAMv1. The high resolution area has an effective spatial resolution between grid points of 25 km, while the lower resolution area has an effective spatial resolution between grid points of 110 km. A transition region bridges these resolutions. The red dot indicates the location of DOE ARM site in the tropics.

Figure 2*:* Autoconversion rate (**a-c**) as a function of cloud water mixing ratio (Qc) at cloud droplet number concentration (Nc) = 20, 100, and 200 cm^{-3} and (**d-f**) as a function of Nc at Qc = 0.3, 0.5, and 0.7 g kg^{-1}.

Figure 3: Heatmap diagram evaluating EAMv1 against CMIP5 AMIP simulations following Gleckler et al (2008). The fields listed by row are (from top to bottom), 500 hPa geopotential height, 200 and 850 hPa meridional and zonal wind, and air temperature, surface air temperature, meridional and zonal surface stress, top of atmosphere long and shortwave cloud radiative effect, net, long and shortwave TOA fluxes, and precipitation.

Figure 4: Zonally Averaged model temperature biases (for AMIP simulations) compared to MERRA2 Reanalysis. Left column shows December-January-February Average. The EAM and CAM difference fields share the same color bar within a column. Right column shows June-July-August. Mass and Area Weighted RMS and correlation coefficients are shown.

Figure 5: Zonally Averaged difference between model mean relative humidity (over liquid water) and the MERRA re-analysis. Layout details as in figure 4.

Figure 6: Zonally Averaged difference between zonal mean wind and the MERRA re-analysis. Layout details as in figure 4.

Figure 7: Difference between model reference height temperature and the MERRA2 Reanalysis. Layout details as in figure 4.

Figure 8: GPCP estimates compared to AMIP model simulations (EAM center row, CAM5.3 bottom row). Left column shows Annual Mean precipitation (mm/day). Center column shows Summer-Winter precipitation as a measure of monsoon strength. Right column shows the precipitation differences across equinox months (April/May – Oct/Nov). The lower right four values (AC1 and AC2) are multiplied by -1 in the S. Hem. to account for the change in seasonality. A white color in the difference fields indicates small biases compared to observational estimates.

Figure 9: Difference between model (AMIP) and observational estimates (GOCAP) of total Aerosol Optical Depth. Left panel EAMv1, right panel CAM5.

Figure 10. Shortwave Cloud Radiative Effect: Top row from EBAF 4..0 observational estimates. Left column shows December-January-February Average. The EAM and CAM difference fields (for AMIP simulations) share the same color bar within a column. Right column shows June-July-August.

Figure 11. Longwave Cloud Radiative Effect: Top row from EBAF 4.0 observational estimates. Left column shows December-January-February Average. The EAM and CAM (AMIP) difference fields share the same color bar within a column. Right column shows June-July-August.

Figure 12: Differences between model AMIP simulations (EAMv1 in blue, CAM5 in orange) and CERES-EBAF observationally based estimates of all sky fluxes using the sign conventions adopted by the CESM community (longwave flux is positive out of surface, shortwave flux positive into surface, net flux is longwave minus shortwave). Heavy Solid lines/numbers show top of atmosphere differences and area weighted RMS Lighter lines/numbers show differences in surface fluxes.

Figure 13: A comparison of model biases in F2000LR Low (left column) and F2000HR High Resolution (Right Column) simulations compared to the ERA-Interim (1980-2004) reanalysis. Top row: Temperature(K); Middle row: Relative Humidity(%); Bottom row: Zonal Wind (m/s).

Figure 14: Differences between DJF and JJA mean modeled (F2000) and GPCP precipitation (top row, Huffman et al, 2009) estimates for 4 year averages (years 2-5) of the high (middle row) and standard low resolution (bottom row) model simulations.

Figure 15: Simulations of JJA mean total precipitation against GPCP1DD precipitation (Huffman et al, 2001) estimates for 4 year averages (years 2-5) of the standard low resolution (left, ne30) and regionally refined model configuration over CONUS (right, ne30->ne120) Dotted areas indicate where the differences are statistically significant at the 95% confidence level with a two-tailed Student's t test.

Figure 16: Timing phase (color) and amplitude (color density) of the first diurnal harmonic of total precipitation (mm/day) from 10 years of 3-hourly averaged data for (a) Observed (TRMM), (b) EAMv1 (AMIP) and (c) CAM5 (AMIP).

Figure 17: Wavenumber frequency spectra (following Wheeler and Kiladis, 1999) showing the ratio of unfiltered background spectra of daily interpolated (symmetric about the equator) outgoing longwave radiation (1986-2005) for (a) Observed (NOAA), (b) EAMv1 (AMIP), (c) E3SM (coupled), and CAM5 (AMIP).