UC Davis
UC Davis Previously Published Works

Title

The DOE E3SM Coupled Model Version 1: Description and Results at High Resolution

Permalink

https://escholarship.org/uc/item/2qx7r30d

Journal

Journal of Advances in Modeling Earth Systems, 11(12)

ISSN

1942-2466

Authors

Caldwell, Peter M
Mametjanov, Azamat
Tang, Qi
et al.

Publication Date

2019-12-01

DOI

10.1029/2019ms001870

Peer reviewed

Key Points:
- The high-resolution E3SMv1 model was run for 50 years using 1950 forcing data according to the HighResMIP protocol
- Higher resolution and associated retuning improved bias relative to coarser versions of E3SMv1, particularly in ocean and sea ice metrics
- Aerosol and climate sensitivity were relatively unaffected by resolution change; resolution-related tuning had a larger impact

Correspondence to:
P. Caldwell,
caldwell19@llnl.gov

The DOE E3SM Coupled Model Version 1: Description and Results at High Resolution

Peter M. Caldwell[1], Azamat Mametjanov[2], Qi Tang[1], Luke P. Van Roekel[3], Jean-Christophe Golaz[1], Wuyin Lin[4], David C. Bader[1], Noel D. Keen[5], Yan Feng[2], Robert Jacob[2], Mathew E. Maltrud[4], Andrew F. Roberts[3], Mark A. Taylor[6], Milena Veneziani[3], Hailong Wang[7], Jonathan D. Wolfe[3], Karthik Balaguru[7], Philip Cameron-Smith[1], Lu Dong[7], Stephen A. Klein[1], L. Ruby Leung[7], Hong-Yi Li[8], Qing Li[3], Xiaohong Liu[9], Richard B. Neale[10], Marielle Pinheiro[11], Yun Qian[7], Paul A. Ullrich[11], Shaocheng Xie[1], Yang Yang[7], Yuying Zhang[1], Kai Zhang[7], and Tian Zhou[7]

[1]Lawrence Livermore National Lab, Livermore, CA, USA, [2]Argonne National Laboratory, Lemont, IL, USA, [3]Los Alamos National Laboratory, Los Alamos, NM, USA, [4]Brookhaven National Laboratory, Upton, NY, USA, [5]Lawrence Berkeley National Laboratory, Berkeley, CA, USA, [6]Sandia National Laboratories, Albuquerque, NM, USA, [7]Pacific Northwest National Laboratory, Richland, WA, USA, [8]Department of Civil and Environmental Engineering, University of Houston, Houston, TX, USA, [9]Department of Atmospheric Science, University of Wyoming, Laramie, WY, USA, [10]National Center for Atmospheric Research, Boulder, CO, USA, [11]Department of Land, Air, and Water Resources, University of California-Davis, Davis, CA, USA

Abstract This study provides an overview of the coupled high-resolution Version 1 of the Energy Exascale Earth System Model (E3SMv1) and documents the characteristics of a 50-year-long high-resolution control simulation with time-invariant 1950 forcings following the HighResMIP protocol. In terms of global root-mean-squared error metrics, this high-resolution simulation is generally superior to results from the low-resolution configuration of E3SMv1 (due to resolution, tuning changes, and possibly initialization procedure) and compares favorably to models in the CMIP5 ensemble. Ocean and sea ice simulation is particularly improved, due to better resolution of bathymetry, the ability to capture more variability and extremes in winds and currents, and the ability to resolve mesoscale ocean eddies. The largest improvement in this regard is an ice-free Labrador Sea, which is a major problem at low resolution. Interestingly, several features found to improve with resolution in previous studies are insensitive to resolution or even degrade in E3SMv1. Most notable in this regard are warm bias and associated stratocumulus deficiency in eastern subtropical oceans and lack of improvement in El Niño. Another major finding of this study is that resolution increase had negligible impact on climate sensitivity (measured by net feedback determined through uniform +4K prescribed sea surface temperature increase) and aerosol sensitivity. Cloud response to resolution increase consisted of very minor decrease at all levels. Large-scale patterns of precipitation bias were also relatively unaffected by grid spacing.

Plain Language Summary The Energy Exascale Earth System Model (E3SM) is a relatively new fully coupled Earth system and climate model used in major international model simulation projects and mission-defined efforts for the U.S. Department of Energy. This paper describes the first simulation of the model in its high-resolution configuration. This higher-resolution version is able to capture the most energetic motions in the ocean, which are poorly represented in standard resolution coupled climate models, as well as the largest of storms in the atmosphere. Evaluation of this simulation confirms the benefits of high resolution found by other models with a few notable exceptions. These discrepancies with other studies are interesting because they provide a richer understanding of how and why resolution affects model bias. Another key finding is that climate and aerosol sensitivity in E3SM is unaffected by resolution change. This affirms the usefulness of coarser-resolution models for understanding global-scale climate change. This study also confirms the benefits of increased resolution for studying fine-scale features such as hurricanes and orographic precipitation. Finally, the high-resolution version of E3SM is shown to compare favorably to its low-resolution counterpart and to the models participating in Phase 5 of the Coupled Model Intercomparison Project.

1. Introduction

Earth System Models (ESMs) are one of our most important tools available for predicting, reconstructing, and understanding climate change. They can be used both to provide projections of future climate and to test proposed physical explanations for climate system behavior. ESMs are created by coupling global models of the physical processes governing the temporal evolution of the atmosphere, ocean, land, and cryosphere. These components are typically modeled by breaking the planet into a large number of interdependent grid cells from the bottom of the ocean to nearly the top of the atmosphere.

Global coverage, combined with the need to run for centuries at a time, puts severe restrictions on ESM grid resolution. Coarse resolution is a major problem for ESMs because a lot of important climate processes occur on scales smaller than a typical ESM grid cell. Subgrid scale variations must be *parameterized* in terms of grid-cell averages. Clouds, for example, are often smaller than the O(100 km) grid spacing typical for GCMs in 2019 (Wood & Field, 2011). Ocean mesoscale eddies, a critical mechanism for heat advection, occur at scales of less than 10 km in high latitudes (Maslowski et al., 2008). This is far smaller than the ∼50-km resolution often used at high latitudes to generate ensembles of 21st century Arctic sea ice loss (Jahn et al., 2016; Snape & Forster, 2013). These are just a few examples of resolution-dependent mechanisms that remain problematic in most fully coupled simulations of Earth's climate.

Recognition of the uncertainties caused by coarse resolution has prompted many studies evaluating the impact of increased spatial resolution. For the atmosphere, resolution sensitivity of many aspects of climate seems to be model dependent. Some commonalities do emerge, however. Precipitation, snow-pack, and streamflow in hilly or mountainous terrain improves with increasing horizontal resolution (Delworth et al., 2012; Duffy et al., 2003; Iorio et al., 2004; Pope & Stratton, 2002) due to better resolution of the terrain itself. Extreme precipitation also improves with increasing horizontal resolution (Iorio et al., 2004; Kiehl & Williamson, 1991; Terai et al., 2018; Wehner et al., 2010, 2014) because smaller grid cells support more concentrated vertical motions and therefore larger condensation and fallout. The ratio of stratiform-to-convective precipitation also increases with horizontal resolution as more scales are explicitly resolved (Bacmeister et al., 2014; Demory et al., 2014; Hagemann et al., 2006; Hertwig et al., 2015; Pope & Stratton, 2002; Terai et al., 2018). In general, precipitation seems to improve more in winter than in summer because summertime rainfall tends to occur in the form of convective events that are smaller than even the high-resolution grid cells commonly used for GCMs (Duffy et al., 2003). The impact of model resolution on clouds is more complicated. High resolution has been shown to reduce bias in tropical high cloud but increase bias in low cloud (Wang et al., 2018). Blocking of flow with resulting air stagnation problems is also better resolved as mountains and valleys come into focus (Berckmans et al., 2013; Jung et al., 2012; Matsueda et al., 2009; Schiemann et al., 2017). Sea breeze effects, which are important near the coasts in warm areas, also begin to be resolved as horizontal resolution improves (Boyle & Klein, 2010; Love et al., 2011). Because average eyewall radius for strong Atlantic hurricanes (Cat 2 or higher) is about 23 km (Knaff et al., 2003), resolution at least this fine is expected to be a necessary condition for accurately resolving tropical cyclones. Shaevitz et al. (2014) confirmed this expectation in an analysis of 11 simulations from eight distinct climate models ranging in resolution from 130 to 28 km. They found that tropical cyclone intensity was only well captured (relative to observations) by the model with 28-km resolution. Several other studies have affirmed the ability of ∼25-km atmosphere models to reproduce the frequency and intensity of tropical storms (Atlas et al., 2005; Bacmeister et al., 2014; Wehner et al., 2014).

Ocean and sea ice model fidelity are particularly affected by the ability of a model to resolve eddies and deformation, respectively. The impact of permitting oceanic mesoscale eddies in fully coupled models is dramatic, affecting, among many things, the spread of near-inertial energy (Zhai et al., 2007), water mass modification (Gnanadesikan & Hallberg, 2006), and modulation of sea ice extent and thickness (McPhee, 1992; McGeehan & Maslowski, 2011). Resolving the Gulf Stream improves the simulation of meridional heat transport in the Atlantic (Joyce & Zhang, 2010; Volkov et al., 2008), may increase the strength of the Atlantic Meridional Overturning Circulation (AMOC), and may improve simulated AMOC variability (e.g., Cheng et al., 2013; Hirschi et al., 2013; Marti et al., 2010; Stepanov & Haines, 2014). High resolution also helps to resolve fine-scale sea ice deformation, which is important for simulating coupled fluxes against a cold atmosphere. Until recently, ocean models were most often coupled daily to atmospheric, land and sea ice components, filtering out the semidiurnal ice-ocean response to storms, especially in the

Southern Ocean (Roberts et al., 2015), and adversely affecting mixed layer depth in the high north (Holdsworth & Myers, 2015; Jin et al., 2018). Even with subinertial (~hourly) coupling switched on, regions of intense ocean-atmosphere flux exchange are adversely affected in coarse-resolution models by horizontal diffusion in the wind stress curl (DuVivier & Cassano, 2013) and spreading of freshwater fluxes from land models (GarciaQuintana et al., 2019). Increased horizontal resolution improves upwelling of ocean water off the west coast of continents (Delworth et al., 2012; McClean et al., 2011; Small, Bacmeister, et al., 2014) due to improved surface wind stress from better-resolved coastal topography (Gent et al., 2010). This increased upwelling is particularly important because it is expected to increase stratocumulus, the lack of which is a common feature in climate models. Improved atmospheric resolution enhances Antarctic katabatic winds (Lenaerts et al., 2012) and their forcing of a properly resolved slope current surrounding that continent, which require a minimum oceanic resolution of 0.5° (Mathiot et al., 2011). The benefits of high-resolution bathymetry are most keenly felt where narrow passages are critical to water mass exchanges, such as through Nares Strait between Greenland and Ellesmere Island, which can have a significant sea ice flux (Kwok et al., 2010) and is often too wide or completely closed off in coarse-resolution ESMs. El Niño Southern Oscillation often also improves with resolution (Delworth et al., 2012; Sakamoto et al., 2012). However, we note that a key mechanism for model improvement of ENSO, tropical instability waves (Roberts et al., 2009), are already well simulated in the standard resolution configuration of E3SMv1 (Golaz et al., 2019).

While many aspects of climate do improve with increasing horizontal resolution, some seem to be robustly insensitive to resolution. Resolution does not seem to be a panacea for double ITCZ problems (Bacmeister et al., 2014; McClean et al., 2011), except perhaps in the Eastern Pacific where water advected by the trade winds has more reasonable SST due to improved upwelling (Delworth et al., 2012). Resolutions of 25 km in the atmosphere also do not seem to improve simulation of the Madden-Julian Oscillation (Bacmeister et al., 2014; Jung et al., 2012).

It is worth noting that the above discussion focuses on global atmosphere models running in the hydrostatic regime and ocean models able to resolve only the largest eddies. A new class of global atmospheric models that explicitly resolve cloud systems is gaining popularity (Satoh et al., 2014). Because these models capture such different scales, it is likely that their resolution sensitivity is quite different from that described above. Although E3SM has plans to develop such a model, running these cloud-resolving models is unfeasible currently for decadal to centennial climate scales and is therefore outside of the scope of the current analysis.

While past studies of the impact of changing horizontal resolution are plentiful, relatively little analysis has been made of the impact of vertical resolution. Parametric assumptions usually are entangled with the vertical grid, so changing vertical resolution convolves discretization error with parametric errors, oftentimes resulting in little payback in model skill from additional computational expense. This is why even though Lindzen and Fox-Rabinovitz (1989) provide a sound theoretical argument for high vertical resolution in GCMs, we know of no GCMs that actually use the theoretically appropriate resolution. Richter et al. (2014) study the effect of increased vertical resolution in CAM5, finding that an increased model top provides better quasi-biennial oscillations and improved seasonal cycle of temperature in midlatitude. Partly based on this work, E3SM has switched from 30 to 72 vertical layers and pushed the model top up to 60 km for all model resolutions. The impact of increased vertical resolution on E3SM is discussed in Rasch et al. (2019) and Xie et al. (2018).

The E3SM project was launched in large part to tackle the challenge of high-resolution (Bader et al., 2014). The high-resolution coupled model configuration described here is a central part of that strategy and is a standard configuration of the v1 release. This paper is to serves as a reference for the coupled high-resolution configuration to complement the standard resolution configuration of E3SMv1 described in Golaz et al. (2019). In addition, this paper assesses the impact of horizontal resolution on E3SMv1. While there have already been many papers on this latter topic, the conclusions for E3SMv1 are somewhat novel and therefore of broad interest. Section 2 briefly describes the E3SMv1 model and all differences between its low- and high-resolution configurations, as well as the simulations used in this study. This is followed by an in-depth evaluation of the model's 1950 control simulation in section 3. Modes of temporal variability are discussed in section 4, followed by an analysis of the climatology of extreme weather events in section 5. This is followed by a short discussion of the high-resolution model's sensitivity to anthropogenic aerosol and greenhouse gas changes in section 6. Finally, conclusions are presented in section 7.

2. Experimental Design

2.1. Model Description

The E3SMv1 Earth System Model is described in Golaz et al. (2019). The atmospheric component is based on the spectral-element atmospheric dynamical core (Dennis et al., 2012) with 72 vertical levels. It features the Zhang-McFarlane (ZM; Neale et al., 2008; Richter & Rasch, 2008; Zhang & McFarlane, 1995) deep convection scheme and the Cloud-Layers Unified by Binormals (CLUBB; Golaz et al., 2002; Larson, 2017; Larson & Golaz, 2005) for macrophysics, turbulence, and shallow convection. Morrison-Gettelman Version 2 (MG2; Gettelman & Morrison, 2015; Gettelman et al., 2015) microphysics, four-mode Modal Aerosol Model (MAM4; Liu et al., 2016), and Rapid Radiative Transfer Model for general circulation models (RRTMG; Iacono et al., 2008; Mlawer et al., 1997) are also used. Additional details are provided in Xie et al. (2018), Rasch et al. (2019), Golaz et al. (2019).

Ocean and sea ice components in E3SMv1 are based on the Model for Prediction Across Scales (MPAS; Petersen et al., 2019; Ringler et al., 2013). For the high-resolution simulations presented here, the Gent-McWilliams (GM) mesoscale eddy parameterization is not needed and is disabled. For the low-resolution simulation with high-resolution tunings, the GM bolus diffusivity is set to a constant 1,800 m^2/s (consistent with Golaz et al., 2019). Neither the high-resolution or low-resolution simulations include a parameterization of submesoscale eddy transport. The land model is a slightly revised version of that found in CESM1 (Hurrell et al., 2013), as described in Golaz et al. (2019).

The river transport component is the newly developed Model for Scale Adaptive River Transport (MOSART; Li et al., 2013, 2015). The core of MOSART is the simplified form of the one-dimensional Saint-Venant equation, a kinematic routing method that provides explicit simulation of riverine hydrologic variables such as channel velocity, channel water depth, and water surface area. Simulating these variables makes incorporating water management, heat, and biogeochemical processes straightforward. MOSART was developed to work with both grid- and vector-based representations (Li et al., 2013; Tesfa et al., 2014). In E3SMv1, the grid-based representation of MOSART is adapted.

As noted in Golaz et al. (2019), the ocean model in E3SMv1 accounts for changes in heat content of water with temperature, but the atmosphere does not. Keeping track of the heat content of water in the atmosphere is challenging, so E3SMv1 instead applies an ad hoc correction where the global-average difference between the temperature at which water evaporates from the ocean and returns back as rain or stream outflow is applied as a globally uniform sensible heat flux correction. This correction was found to have negligible effect on model climate and is described in more detail in Golaz et al. (2019) Appendix A.

2.2. Model Resolution

In this section, we describe the grids used for the high-resolution (HR) and low-resolution (LR) configurations of E3SMv1. Details of the HR and LR grid configurations are described below and summarized in Table 1. To isolate the effect of resolution, we will focus particularly on a LR run using the same parameterization constants (tuning parameters) as HR; we will call this run LRtunedHR to differentiate it from the E3SMv1 standard resolution configuration (hereafter LRv1) described by Golaz et al. (2019).

At all resolutions, E3SMv1 uses one grid for atmosphere and land components, another grid for ocean and sea ice, and a third grid for streamflow. The atmosphere/land grids are based on a cubed-sphere topology. Atmospheric parameterizations and the land model are called on each of 16 Gauss-Lobatto nodes for each spectral element. After accounting for cells on element edges being shared between elements, this results in atmospheric and land processes being called on approximately nine times more columns than used for dynamics. The high-resolution configuration of E3SMv1 contains 120 quadrilateral spectral elements in both x and y directions of each face of the cube sphere for a total of ∼800,000 columns corresponding to an approximate grid spacing of 25 km. The LR atmosphere, in comparison, has 30^2 total elements on each side of its cube for a total of ∼50,000 columns (110-km nominal grid spacing). At all resolutions, E3SMv1 contains 72 vertical layers on a sigma-pressure coordinate system. As described in Xie et al. (2018), layers are unevenly spaced with finer resolution near the Earth's surface. Atmosphere and land resolution of 25 km was chosen as a compromise between computational feasibility and improved representation of topography; it is not expected to explicitly capture turbulence or convection.

The MPAS ocean and sea ice models share the same horizontal grid, which at high resolution contains 3.7 million horizontal cells. In the HR configuration, ocean resolution is coarsest in the tropics and becomes

Table 1
Information About Grids Used in This Paper

Grid	~ Δx	# of columns	vertical levels
HR atm	25 km	777,602	72
HR land	25 km	777,602	15
HR ocean	8–16 km	3,693,225	80
HR sea ice	8–16 km	3,693,225	5
HR river	0.125°	4,147,200	1
LR atm	110 km	48,602	72
LR land	110 km	48,602	15
LR ocean	30–60 km	235,160	60
LR sea ice	30–60 km	235,160	5
LR river	0.5°	259,200	1

Note. Grid spacing is given in km for near equal-distance grids and in degrees for equal-angle grids. For the sea ice model, the number of vertical levels refers to the number of thickness distribution categories.

finer near the poles. The grid is defined to well resolve the first Rossby Radius of deformation (Chelton et al., 1998) across most of the globe. This grid resolution is broadly similar with other high-resolution ocean configurations (e.g., Delworth et al., 2012; McClean et al., 2011; Small, Bacmeister, et al., 2014). The LR MPAS mesh was designed such that the horizontal resolution has a similar meridional distribution (on average) as the standard CESM1 ocean grid, varying from 60 km in the midlatitudes to 30 km near the equator and poles. The high-resolution version of the E3SMv1 ocean model uses 80 levels in the vertical, with spacings following Stewart et al. (2017). This vertical grid is designed to resolve the vertical structure of mesoscale eddies. The LR ocean model vertical discretization is very similar to CESM1, with 60 layers varying from 10 m spacing at the surface to 250 m at depth. The sea ice model includes five thickness categories in each grid cell divided at 0.65, 1.39, 2.47, and 4.56 m.

For coupled simulations, MOSART employs a uniform latitude/longitude grid with grid spacing listed in Table 1. Only grid cells over land are active; there are 1.5 million active cells at high resolution and 0.1 million at low resolution. MOSART is not discretized in the vertical.

Table 2
Time Steps (in Min) Used in High-Resolution and Low-Resolution Configurations

Time step	HR	LR
atm radiation	60	60
atm phys/dyn coupling	15	30
atm CLUBB+microphysics	5	5
atm dyn remap	2.5	15
atm advection and dyn	1.25	5
atm hyperviscosity	0.417	1.67
ocn	6	10
ocn barotropic	0.2	0.67
ice thermodynamics	15	15
ice dynamics	7.5	15
river	60	60
river coupling	180	180
atm/ice/lnd coupling	15	30
ocn coupling	30	30

Finer grids typically require shorter time steps for stability. This is certainly true for E3SMv1, as evident from the time step information in Table 2. Simultaneous modification of grid spacing and time step can result in model sensitivity to time step being erroneously attributed to spatial resolution sensitivity (Jung et al., 2012). Due to time pressures, we do not decouple time step and resolution sensitivity in this study and instead use LR time steps for the LRtunedHR simulation. As a result, what we describe as resolution sensitivity should be interpreted as including all changes required to run at that resolution.

2.3. Tuning

The exact values for many parameters in climate models are not precisely known. Values for these parameters are generally chosen to minimize model bias in reproducing current climate or to achieve global top-of-atmosphere (TOA) radiation balance for preindustrial conditions. Because empirically chosen tuning parameters can compensate for imperfect process representations, they play a critical role in model development (Hourdin et al., 2017; Schmidt et al., 2017). Tuning involves testing a variety of parameter settings and choosing the best configuration. Thorough tuning often requires hundreds of tests, which quickly becomes prohibitively expensive at high resolution. While parameters chosen for LR simulations often behave reasonably well at higher resolution, optimal tuning choices are generally resolution dependent. This is because parameters are used to compensate for unresolved physics, and more physical processes become resolved at higher resolution.

Because of computational expense, tuning of high-resolution configurations is often accomplished by changing only one or two parameters to remove the most egregious model biases. As a result, it is sometimes hard to disentangle whether lack of improvement at high resolution is due to unimportance of resolution or poor tuning (Bacmeister et al., 2014). One feature of the current study is that substantial effort was put into tuning the high-resolution atmosphere component, with the result (shown later) that the high-resolution version of E3SMv1 has generally lower bias than its low-resolution counterpart. In order to tune the high-resolution model in a computationally efficient manner, the Cloud-Associated Parameterization Testbed (CAPT; Ma, Chuang, et al., 2015) was extensively employed. CAPT takes advantage of the fact that many atmospheric biases show up in 48 hr forecasts, which are cheaper to run and can be compared against observed weather. Tuning started from the LR parameter settings and targeted mainly the global-mean TOA radiative budget and surface precipitation. Interestingly, the biases that showed up as a result of increasing horizontal resolution were similar to the biases resulting from increasing the number of vertical layers in the LR configuration. This allowed us to use similar tuning strategies for both problems. These strategies are described in more detail in Xie et al. (2018).

Another major difference between HR and LRv1 versions of E3SMv1 is in their treatment of heterogeneous ice nucleation. The LRv1 version of E3SMv1 employs a new parameterization of this process based on Classical Nucleation Theory (CNT; Wang et al., 2014). CNT depends explicitly on aerosol properties (e.g., nucleation efficiency and surface area) and computes ice nucleation separately for immersion, deposition, condensation, and contact modes. CNT replaces the Meyers et al. (1992) parameterization in CESM1. The Meyers scheme only depends on ice supersaturation for liquid-saturated conditions (and therefore is uniquely specified by temperature). The Meyers scheme was derived from Northern Hemisphere midlatitude measurements. Because aerosol concentrations are lower in polar regions than in Northern Hemisphere midlatitudes, the Meyers scheme is expected to overestimate number concentrations of ice crystals at high latitudes. In contrast, CNT is likely to underestimate ice crystal number at high latitudes because E3SMv1 underestimates aerosol transport to high latitudes (shown later). At high resolution, ice number concentration using the CNT scheme became extremely low, resulting in an unrealistic blanket of supercooled liquid over polar regions during winter. This blanket warmed the wintertime polar regions by more than 10 K, wrecking the model's sea ice distribution. To avoid this problem, we reverted the heterogeneous ice nucleation scheme back to the Meyers scheme as part of the HR tuning. We later realized, however, that CNT was behaving poorly because the rate at which liquid was being converted to ice by the Bergeron process was being scaled by a factor of 0.1 to compensate for the Meyers scheme's propensity to create too many ice crystals and therefore too much ice. Thus, CNT behaved poorly not because of a design flaw, but because it disturbed a balance between compensating errors. In retrospect, a better approach would have been to keep CNT and scale Bergeron to a reasonable value. We will pursue this option in E3SM Version 2.

A downside to our strategy of intensive tuning at high resolution is that our LR and HR configurations differ due to a mixture of resolution and tuning choices. Differences in tuning parameters are listed in Table 3.

Table 3
List of Tuning Parameters That Differ Between Low- and High-Resolution Versions of E3SMv1

Parameter	HR	LRv1	Parameterization	Meaning (units)
alfa	0.2	0.1	ZM	ratio of upward versus downward mass flux at downdraft initiation level (-)
c0_lnd	3.5e–3	7e–3	ZM	Precip efficiency over land (m^{-1})
c0_ocn	4.3e–3	7e–3	ZM	Precip efficiency over ocean (m^{-1})
dmpdz	−0.2e–3	−0.7e–3	ZM	parcel fractional mass entrainment rate (m^{-1})
ke	6e–6	5e–6	ZM	Precip evaporation efficiency ($(kg\ m^{-2}\ s^{-1})^{-1/2}\ s^{-1}$)
dp1	3.9e–2	4.5e–2	ZM	Deep convective cloud fraction (−)
c1	1.5	1.335	CLUBB	$\overline{w'^2}$ dissipation strength (−)
c8	4.73	4.3	CLUBB	$\overline{w'^3}$ damping strength (−)
c14	1.75	1.06	CLUBB	Strength of $\overline{u'^2}$ and $\overline{v'^2}$ dissipation (−)
dust_emis_fact	2.5	2.05	Dust emission	Inverse of dust emission strength (−)

It is also worth noting that a few minor bugs were found and final coupled tuning was done between the time the atmosphere-component papers from Zhang et al. (2019), Tang et al. (2019), Xie et al. (2018), and Rasch et al. (2019) were published and the coupled simulations documented in Golaz et al. (2019) were completed. As a result, LRv1 values of ke, c14, and so4_sz_thresh_icenuc published here differ from those in Xie et al. (2018). For the most part, these changes did not impact model biases. One exception is an unrealistic lack of high clouds in the Tropical West Pacific in earlier atmosphere-component papers, which is greatly improved here and in the fixed-SST simulations shown by Golaz et al. (2019).

Several more bugs were found and corrected after LRv1 was completed and before these HR simulations started. In particular, a bug was found and corrected in precomputed values of the rrtmg_sw_reftra table lookup (https://github.com/E3SM-Project/E3SM/pull/2202). Another bug related to energy errors in land-surface water-phase change and lake/snow interactions (https://github.com/ESCOMP/ctsm/pull/307) was also fixed. Both of these bugs had little effect on model climate.

Unlike the low-resolution configuration, the coupled version of the HR model uses identical tunings to the fixed-SST HR simulations described in Zhang et al. (2019), Tang et al. (2019), Xie et al. (2018), and Rasch et al. (2019). This decision was made because the coupled model seemed to be in relatively good radiative balance without further coupled tuning and time pressures required us to start our simulation campaign. In addition, the efficacy of our tuning parameters seemed to saturate at the values currently used for HR such that further tuning provided little benefit. Another tool for tuning E3SM at high resolution is the ability to use a variable-resolution mesh in the atmosphere (Tang et al., 2019). While we did not use this capability much for v1 tuning, we expect this technology to play an important role in the development of future model versions.

2.4. Description of Simulations

Aside from LRv1, all coupled E3SMv1 simulations described in this paper use time-invariant 1950 forcings and follow the HighResMIP protocol of Haarsma et al. (2016) as closely as possible. Using 1950 conditions instead of the preindustrial conditions typically used for low-resolution control runs is necessary in order to extend a control simulation into a transient 20th century/future climate simulation—simulating the period between 1850 and 1950 would exhaust our computing allocation, preventing the eventual simulation of the more interesting 1950–2050 period. Although the planet was not strictly in radiative equilibrium in 1950 (as assumed in a control simulation), most anthropogenic greenhouse gas (GHG) emissions and warming have occurred since 1950 (e.g. IPCC, 2014, Fig. SPM.1), so this approach is a reasonable compromise. Using anthropocene-era control simulations for high-resolution control runs is a common practice (Delworth et al., 2012; Sakamoto et al., 2012).

Forcings for the 1950 coupled simulations are identical to the forcings used for year 1950 of the LRv1 historical simulation (Golaz et al., 2019) with a few exceptions. The mean of solar insolation from 1939 through 1961 was used in order to average over two periods of the 11-year solar cycle. Tropospheric aerosol emissions

were handled similarly in HR to LR but use higher resolution 0.5° data from input4MIPS. Stratospheric volcanic aerosols are averaged over 1850–2014 to capture the conditions that would be observed if simulating a long period of time. Tropospheric oxidants are specified using 1955 values because data for 1950 are not available and 1955 values should be negligibly different. Dimethyl sulfide is computed by interpolating 1849 and 2000 values to the period of interest; this treatment is identical to Golaz et al. (2019) but was not mentioned in the previous paper. Anthropogenic and fire emissions for aerosols and precursor gases are taken from the CMIP6 emissions data set. One exception to this is secondary organic aerosol (SOA) gas emissions, which are handled specially as described by Rasch et al. (2019) in order to improve the vertical distribution of SOA.

While the focus of this paper is on the high-resolution configuration of E3SM, it is occasionally useful to compare against the E3SMv1 LR release (LRv1) configuration. In particular, we will occasionally examine the preindustrial (PI) control and the 20th century historical simulations described Golaz et al. (2019).

To provide crude estimates of model sensitivity to anthropogenic GHG and aerosol changes, we also perform short simulations with prescribed sea surface temperature (SST) and sea ice extent. To estimate the net climate feedback parameter, we perform a pair of 5-year current-climate simulations with annually repeating forcings. SST is uniformly raised by 4 K in one simulation and feedback strength is inferred from the difference in resulting TOA radiative imbalance between runs. Aerosol sensitivity is computed from a pair of 15-month simulations nudged to observations with a 6-hr relaxation timescale. In one simulation, anthropogenic aerosols are omitted. In the other, they are included. Detailed explanation of these sensitivity studies and their results are provided in section 6.

For all prescribed-SST runs, forcings representative of 2010 are used unless explicitly mentioned. As recommended for HighResMIP, these runs use daily-resolution prescribed SST and sea ice extent at 25-km resolution from HadISST2 (Haarsma et al., 2016). GHG and volcanic forcings are taken from year 2010 of the CMIP6 Historical experiment for these runs. Solar forcing is the average of CMIP6 Historical data for the 22 years centered on 2010.

2.5. Initialization

The initial ocean temperature and salinity distributions were interpolated from the polar hydrographic climatology (Steele et al., 2001). To generate currents and propagate initial gravity waves, MPAS-Ocean was run in standalone mode for three months, restoring temperature and salinity to the initial climatology and imposing an annually averaged wind stress. The results were used to drive a simulation with interactive ocean and ice components and atmospheric forcing prescribed from the COREv2 interannually varying (Large & Yeager, 2009) dataset. Ice was originally set to a uniform thickness of 1 m at all locations poleward of 70° for this calculation. Only years 1947–1950 were simulated in this mode. This 3-year spin-up period is a compromise between minimizing the amount of model drift inherited by the fully coupled simulation and preventing large shocks upon coupling. The resulting ocean and ice state was combined with an atmosphere initial condition from an earlier experimental HR simulation with fixed SST and a land initial condition interpolated from year 1950 of a LR precursor to the E3SMv1 CMIP LRv1. This configuration was run for 5 years before realizing that polar wintertime surface temperatures were unacceptable and deciding to revert from the CNT mixed-phase nucleation scheme to the old Meyers et al. (1992) scheme to fix the problem (as described in section 2.3). Instead of discarding those 5 years of simulation with the CNT scheme, we decided to begin our production off the end of the 5-year CNT run in order to start the model from a slightly more equilibrated state. For consistency, our LR simulation with HR tuning follows an identical spin-up protocol, including 5 years of coupled simulation with the CNT mixed-phase nucleation scheme. For both simulations, we denote year 1 as the first year with the Meyers rather than CNT nucleation scheme.

2.6. Performance

The v1 HR simulation was performed at DOE's Argonne Leadership Computing Facility (ALCF) Theta supercomputer, which has 4,400 nodes of Intel KNL ("Knights Landing") with 64 1.3 GHz cores and 192 GB memory per node. The coupled E3SM model consists of the five components described in section 2.1, which exchange fluxes through a top-level coupling driver (CPL). Overall model throughput is optimized by laying out components across processing cores such that none of the components is a throughput limiter, while at the same time none of the processing cores are idle. Figure 1 shows the layout used for most of our v1-HR simulation. E3SM's component synchronization approach (RASM_OPTION1) prevents the atmosphere model from running in parallel with sea ice, land, or river runoff models, but allows the ocean to run
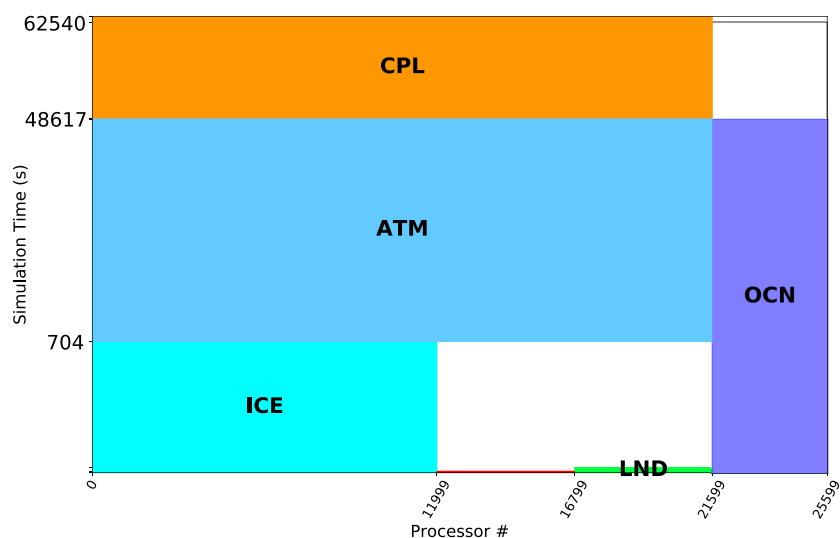
Figure 1. Typical processor layout and throughput for a 6-month simulation on Theta. This layout uses 32 MPI processors and 4 OpenMP threads per process for a total of 128 hardware threads per node. In this figure, the atmosphere models is called ATM, the ocean model is OCN, the sea ice model is ICE, the land model is LND, and the river model is the small red bar at the center-bottom of the plot.

in parallel on its own computational cores. As evident in this figure, the atmosphere and sea ice models are the main determinants of total run time, with the number of ocean-model cores chosen to fit within the time taken by these other processes. Inability of the sea ice model to scale to large process counts, combined with its inability to run in parallel with the atmosphere model, results in unavoidable idle time (shown as white space at the center-bottom of Figure 1). CPL sequences, interpolates, and exchanges data among all components and consumes about 20% of overall run time. CPL scaling limitations result in the idle time indicated by white space in the top-right corner of Figure 1. Overall throughput for this layout was ~0.7 simulated years per wall-clock day. Various performance optimizations (reduced synchronization and improved threading) are being added to ATM and ICE to raise throughput in v2. Using more processor cores would have improved simulation speed but would have been more expensive (because of increased communication lag time) and would have resulted in more time waiting in queue. The 800-node layout shown was chosen because it is the smallest layout that qualified for the largest production queue and its associated 24-hr wall-clock limit.

LRtunedHR was run using 54 nodes of Argonne's Anvil cluster, which has 36 64-GB Intel Xeon Broadwell cores per nodes. It achieved approximately five simulated years per day (SYPD) and consumed ~9,000 core hours per simulated year. LR runs with prescribed SST were run on 169 nodes of NERSC Cori Haswell nodes, which have two 2.3-GHz 16-core Haswell processors per node. It ran at ~13 SYPD and cost ~10,000 core hours per simulated year. Prescribed-SST HR simulations completed ~0.80 SYPD using ~640,000 core hours per simulated year on 323 NERSC Cori KNL nodes, which have one 1.4-GHz 68-core Intel Xeon Phi Processor 7250 per node.

The HR configuration contains ∼ 16× more grid points than LR (Table 1) and requires a time step which is overall ∼ 2× shorter (Table 2). As a result, we expect (and find) HR runs to be ~32 times more expensive than LR on a given machine. Resolution increases have less effect on time to solution because of greater opportunities for parallelism.

Source code for the simulations has been taken from a long-term maintenance branch `maint-1.0` (https://github.com/E3SM-Project/E3SM/tree/maint-1.0) that will be maintained to ensure that runs can be exactly (bit for bit) reproduced if run on the same model with the same software environment.

3. Model Climate
3.1. Global-Average Evaluation
Figure 2 displays the time series over the simulation for several key large-scale climate parameters. While the LRv1 and LRtunedHR runs are close to TOA radiation balance, the HR run loses energy at the rate of ~0.5 W/^{2}. This slow energy loss rate results in a slight cooling trend to bring the globally averaged surface air
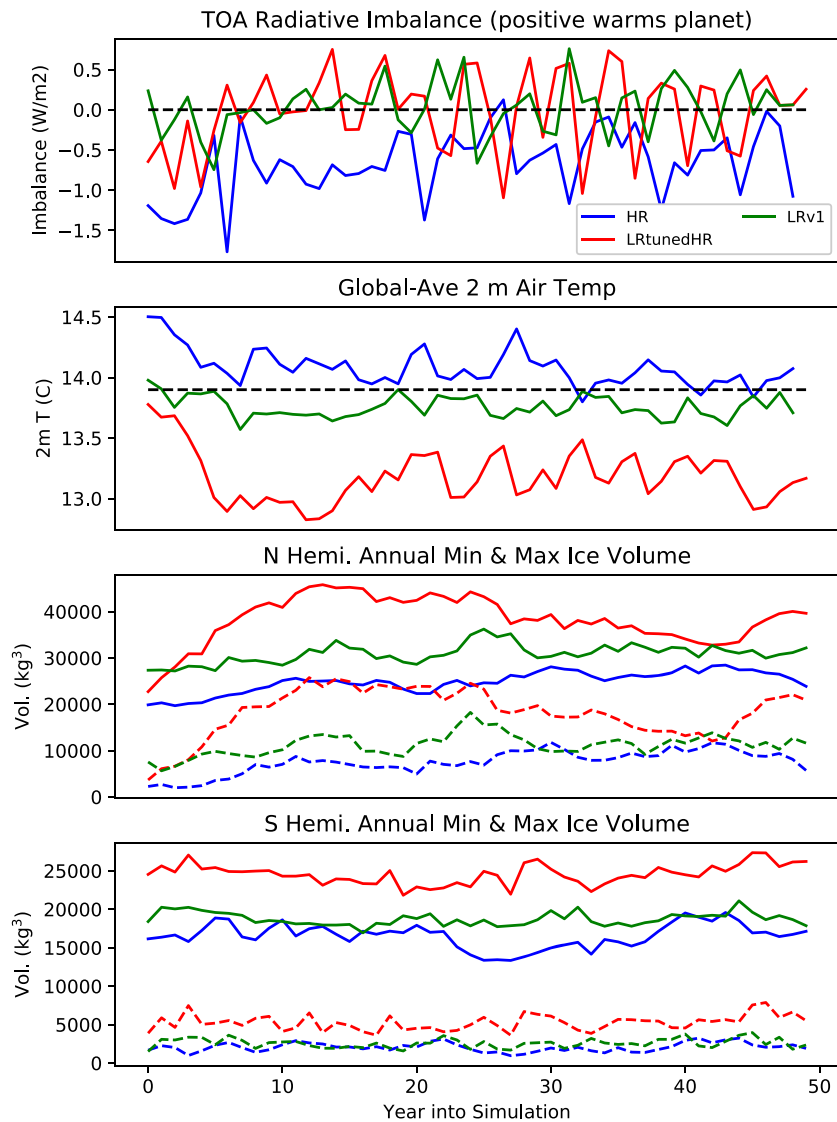
Figure 2. Time series of key global-average quantities. Zero TOA net energy balance is depicted as a dashed line in the top panel. In the second panel, horizontal dashed line at 13.9°C is an estimate observed global average temperature over the 20th century (NOAA Global Climate Report, 2018; https://www.ncdc.noaa.gov/sotc/global/201813).

temperature closer to the 20th century observed value, whereas the LRtunedHR run is over 0.5° too cold. The small imbalance and slight global trend in our HR run is similar in magnitude to the imbalances found in the first several decades of many fully coupled high-resolution climate simulations. HR simulations tend to have larger radiative imbalances and corresponding temperature trends than their low-resolution counterparts because they are too expensive to optimize in a fine-tuned way. It is recognized in the HighResMip protocol that the 50-year spin-up period is a compromise between computer expense and a drift-free model. The near-zero energy imbalance in the LRtunedHR simulation was a pleasant surprise, as it took many iterations to achieve an acceptable, drift-free climate for the standard resolution LRv1 run (Golaz et al., 2019). Trends in sea ice volume are shown in the lower panels of 2. In both hemispheres, ice volume is relatively steady across the simulation, indicating that evaluating climatological averages is appropriate.

Equilibration of the ocean is shown in Figure 3. Consistent with the overall loss of energy of the planet, the global ocean is cooling in HR over most of the water column, but especially in the upper 1,000 m. The deeper ocean remains fairly isolated and cools with time. HR is accumulating a small fresh bias in the upper ocean, but this bias is insignificant compared to the fresh bias in the upper 800 m of the water column
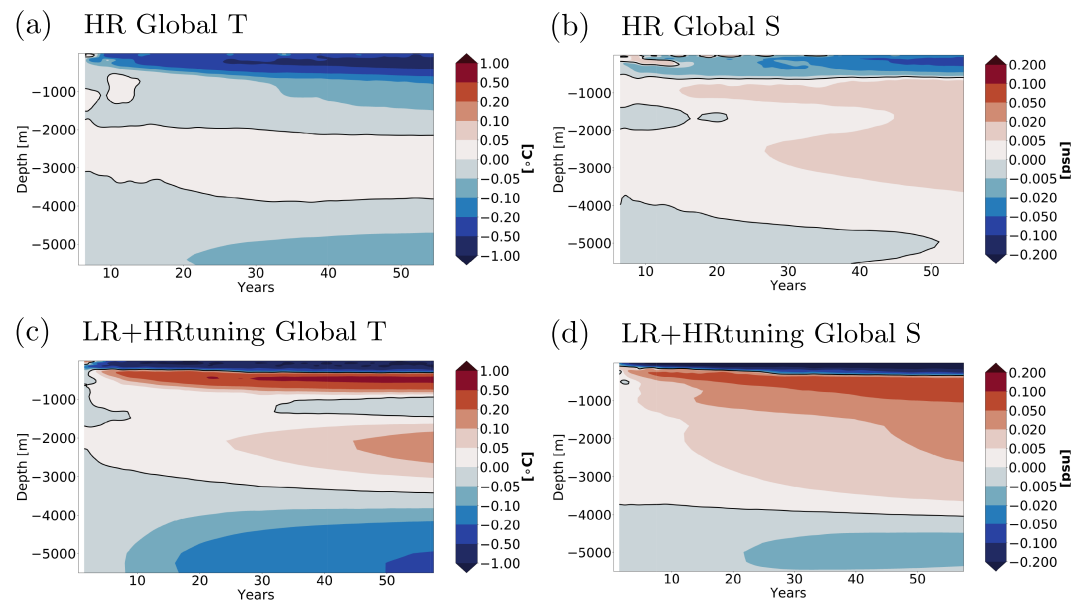
Figure 3. Trends of globally averaged ocean temperature (left panels) and salinity (right panels) versus depth for (a and b) HR and (c and d) LRtunedHR. Trends are computed with respect to annual mean of the first simulated year.

of LRtunedHR. A separate study is now underway to investigate this accumulating fresh bias at the ocean surface in v1 low-resolution simulations.

Because of the better simulation of scale interactions, increased resolution is widely assumed to result in better accuracy at scales resolved by lower-resolution models. We test that assumption by comparing large-scale measures of model bias among the LRv1, LRtunedHR, and HR model runs.

To get a crude sense of HR bias relative to LRtunedHR, LRv1, and models developed by other groups, Figure 4 provides uncentered global root-mean-squared error (RMSE) values for a variety of climatically important variables for all 3 E3SMv1 models and all available CMIP5 models. Prior to analysis, output from all models was spatially filtered through area-average regridding to 2.5°. We chose uncentered RMSE to account for global-mean bias in the simulations. As was seen in similar comparisons of the ECHAM atmospheric model by Hertwig et al. (2014) and CESM by Small, Tomas, et al. (2014), HR has lower error than LRv1 for most variables and seasons. In particular, HR is never in the lowest 25% of CMIP5 models and is among the best CMIP5 models for most variables. Because HR differs from LRv1 due to tuning as well as resolution differences while LRtunedHR has not been tuned to reduce bias at all, it is inappropriate to conclude that HR's superior climatology is due solely to resolution changes. In addition, CMIP5 runs had completed many hundreds of years of simulation by the time we sampled their errors, while HR's spin-up period was only a few decades. While short spin-up is typical (and perhaps unavoidable) for high-resolution simulations, this discrepancy may give high-resolution simulations an unfair advantage in evaluations of this type. Assessing the impact of initialization approaches on HR evaluation is the subject of important future work.

3.2. Clouds and Radiation

Spatial gradients in the amount of radiation absorbed or emitted by the Earth system drive the circulations in both the ocean and the atmosphere, control local temperature and humidity structure, and power the hydrological cycle. As noted above, HR is losing heat at a globally average rate of ~0.5 W/m^2 with significant interannual variability. In nature and models, the globally averaged net energy imbalance is the small residual of the highly variable regional TOA net energy fluxes typically 1–2 orders of magnitude larger. As shown in Figure 5, regional TOA net energy biases are fairly small except for positive bias (too much absorption) in the eastern subtropical ocean basins, over Eastern China, in the Atlantic sector near Antarctica, and (to a lesser extent) over Africa. These are compensated by milder but more widespread negative biases along the west coasts of continents, over the East/Central tropical Pacific (probably associated with intertropical convergence zone biases described later), over the Indian Ocean, and at high northern latitudes. The regional-scale high- and low-resolution radiative imbalances in both resolutions of E3SMv1 look extremely
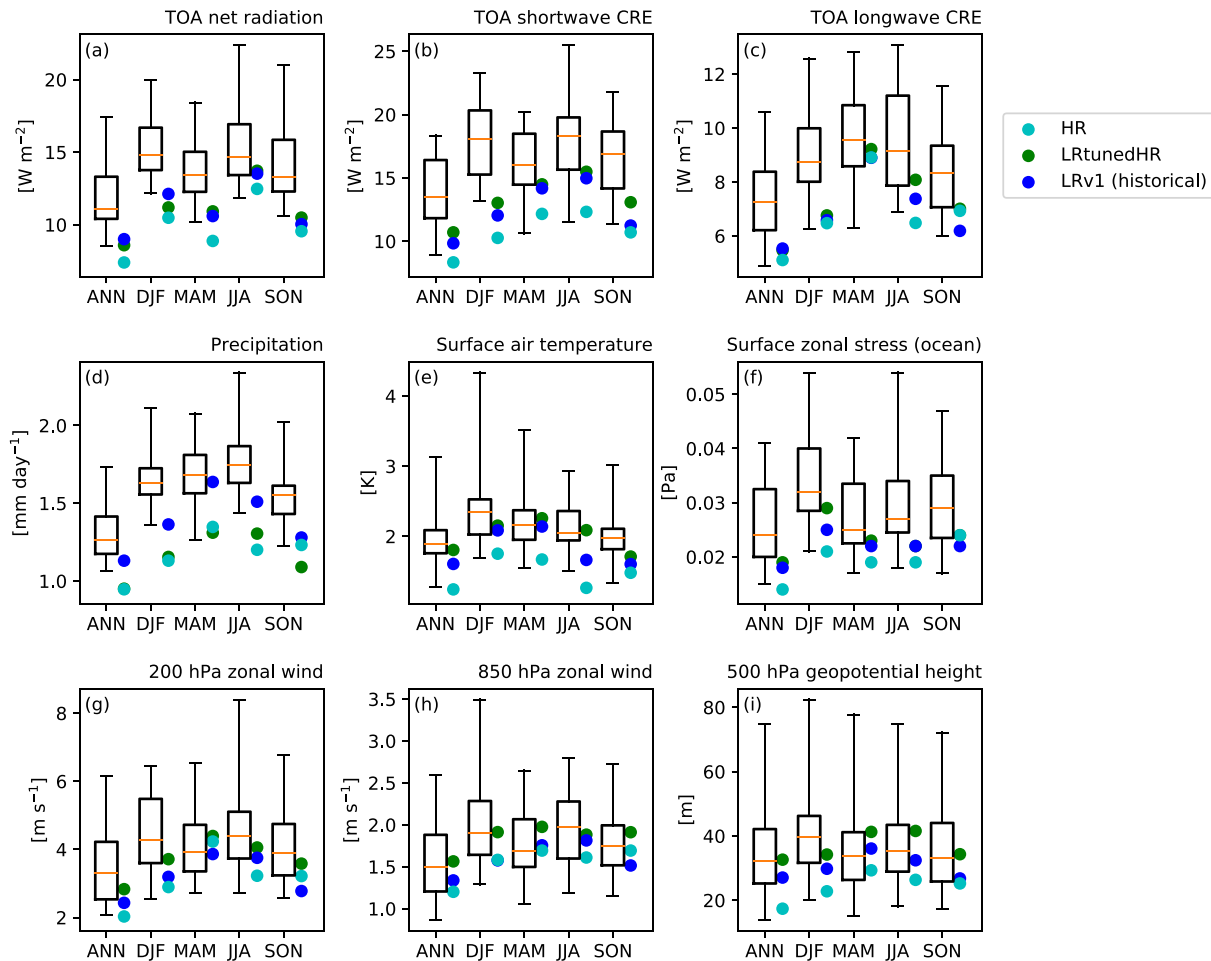
Figure 4. Comparison of uncentered spatial RMSE (computed following Gleckler et al., 2016) from years 1981–2005 of 45 CMIP5 historical simulations (box and whiskers with boxes showing 25th–75th percentile values and whiskers showing the span of the data) versus values from years 21–50 of 1950 control runs for HR (cyan) and LRtunedHR (green). Dark blue dots show RMSE values for 1981–2005 of the first member of the E3SMv1 LRv1 historical ensemble. Error for each variable is calculated relative to the relevant observations: CERES-EBAF for radiation (Loeb et al., 2009), GPCP for precipitation (Adler et al., 2003), and ERA-Interim for the remaining variables (Dee et al., 2011).

similar (compare Fig. 5 with Figure 4 from Golaz et al., 2019). LRtunedHR also shows very similar pattern (not shown), indicating that this net radiative bias structure is a robust feature of E3SMv1 physics.

Clouds are the main source of bias in net radiation. This is evident from the fact that bias in shortwave (SW) cloud forcing (TOA SW radiation minus clear-sky SW radiation, hereafter SWCF) for v1 HR in Figure 6a has a very similar bias pattern to net radiation in Figure 5b. From these figures, we conclude that excessive absorption in the eastern subtropical oceans is due to a lack of SW cloud reflection. Longwave (LW) cloud forcing (LWCF) in v1 HR is too weak almost everywhere except at high latitudes. Eliminating the LW bias through tuning proved difficult, forcing us to compensate with a smaller globally averaged SWCF to maintain global-average net radiative balance. This is yet another example of the limitations of tuning. The problem here is an inability to capture transitions between cloud types such that increasing stratocumulus in the eastern subtropical oceans also increases low clouds in convective regions, exacerbating negative SWCF biases in those regions.

While net radiation looked very similar between HR and LRtunedHR, more marked differences appear in SWCF and LWCF. In particular, Figures 6e and 6f show that finer resolution generally results in SWCF becoming more positive and LWCF becoming less positive. This is fairly uniform except at high latitudes where cloudiness seems to increase. Both of these tendencies are consistent with a reduction in cloudiness with increasing resolution, which is apparent in plots of total cloud fraction (Figure 7). This reduction of cloudiness with increasing resolution was also noted in runs with fixed SST and sea ice (Xie et al., 2018;
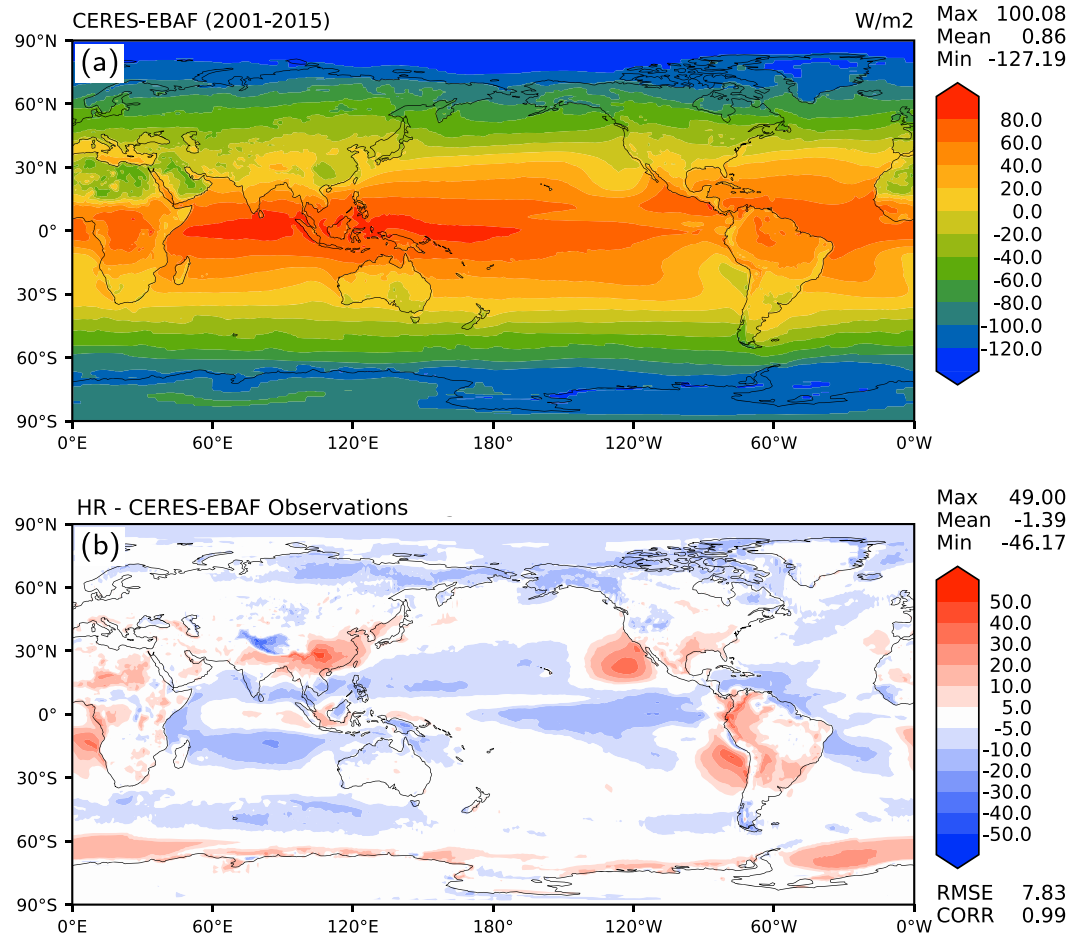
Annual net TOA radiation

Figure 5. TOA radiative imbalance (in W/m^2, positive warms planet). Panel a shows the annual-average climatological distribution from CERES-EBAFv4.0, while panel b shows bias in HR.

Zhang et al., 2019) as well as in other models (Bacmeister et al., 2014). Reductions in both SWCF and LWCF indicate a reduction of high clouds with resolution, while the ~1 W m^{-2} larger reduction in SWCF versus LWCF indicates reduction from low clouds as well. It is worth noting that although clouds are sensitive to resolution in E3SMv1, resolution sensitivity is greatly decreased relative to many earlier generation models (Bacmeister et al., 2014).

Figure 7 reveals that E3SM underestimates cloudiness at low latitudes regardless of resolution. This finding is particularly true in the stratocumulus areas of the eastern subtropical oceans, but also along the equator and over low-latitude land masses. The marine stratocumulus errors are collocated with large SST errors (Figure 16), particularly to the west of Baja California and Namibia/Angola. E3SM atmosphere-only simulations with observed SSTs (not shown) have very similar but somewhat smaller errors, demonstrating that the stratocumulus error, while originating in the atmosphere model, is amplified by the SST error in the coupled model, similar to the E3SM low-resolution LRv1 simulations (Golaz et al., 2019). Other cloud characteristics such as the cloud altitude and optical depth distributions of this E3SM high-resolution model are quite similar to those found in atmosphere-only E3SM simulations including those analyzed in Zhang et al. (2019). An exception is a moderate reduction in the amount of super-cooled liquid in this high-resolution model (not shown) due to the use of CNT rather than Meyers ice nucleation at high resolution (see section 2.3). While we have emphasized model errors, it is worth bearing in mind that E3SM's simulated clouds compare favorably to those of other models. Scalar metrics of model performance from Klein et al. (2013) indicate that E3SM would be in the top rank of pre-CMIP6 models (Zhang et al., 2019).
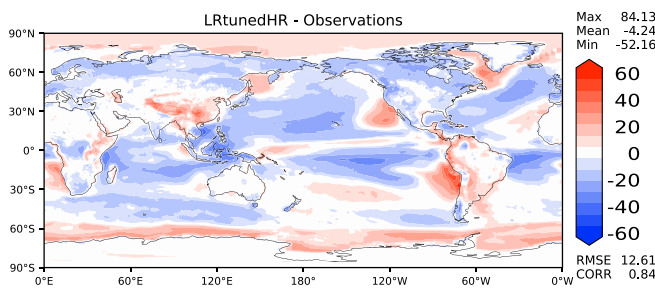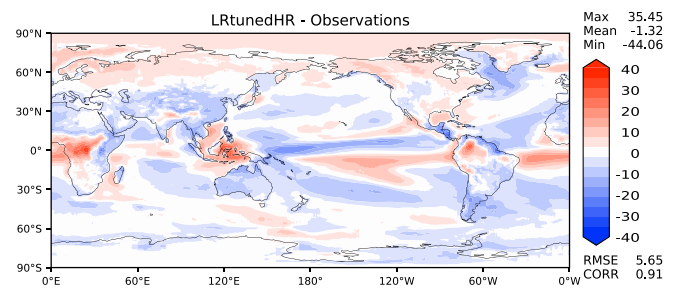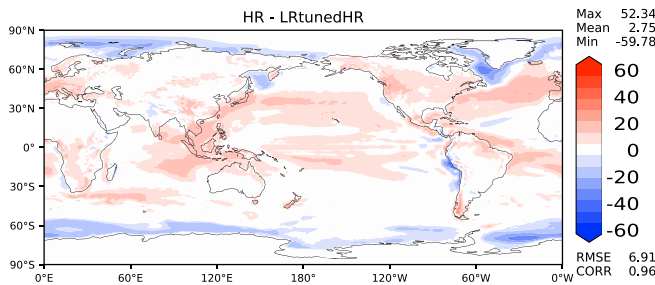
(a) HR SWCF Bias
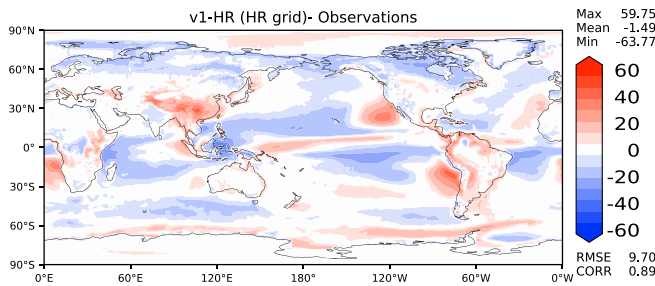
(b) HR LWCF Bias
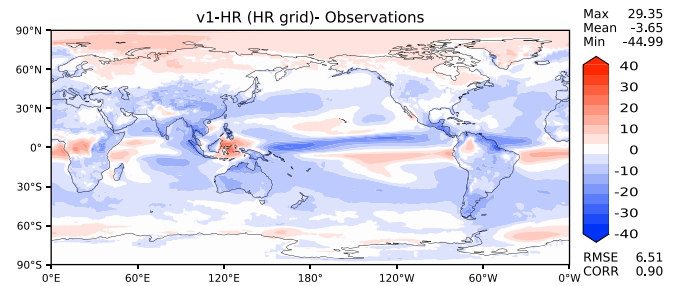
(c) LRtunedHR SWCF Bias

(d) LRtunedHR LWCF Bias

(e) SWCF: HR - LRtunedHR

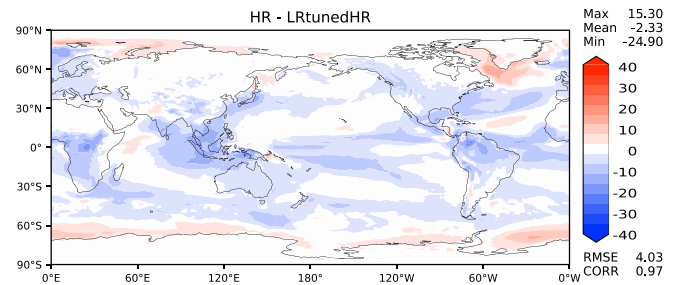## (f) LWCF: HR - LRtunedHR

Figure 6. Panels a–d show annual-average SWCF and LWCF bias in SWCF (left) and LWCF (right) relative to CERES-EBAFv4.0. Panels e–f show differences between HR and LRtunedHR simulations. Positive SWCF values mean less reflection by clouds, while positive LWCF values mean clouds are trapping more LW radiation. Units for all panels are W/m^2.
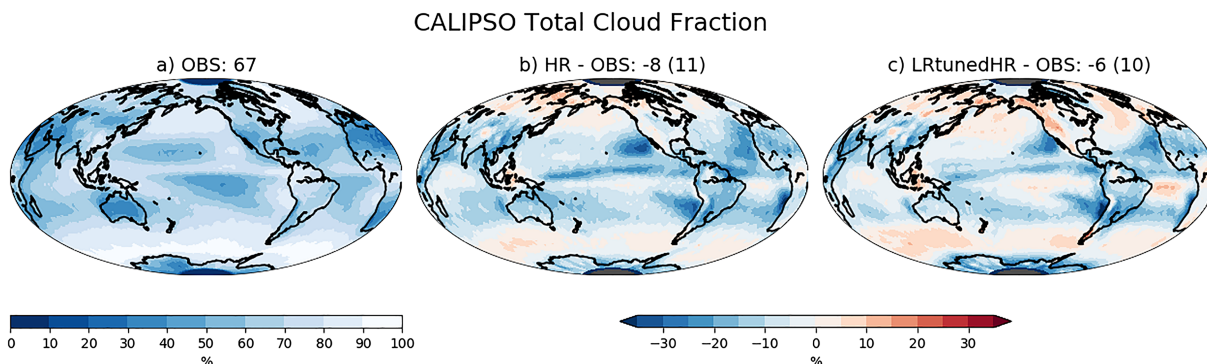
CALIPSO Total Cloud Fraction

Figure 7. Comparison of annual mean total cloud fraction between E3SM and Calipso observations. (a) Observed (OBS) total cloud fraction as retrieved from the Calipso satellite with the GOCCP v3.1.2 algorithms (Chepfer et al., 2008) for the years 2006–2017. (b) Difference between HR and Calipso observations (model minus observed) and (c) the difference between LRtunedHR and observations. Model fields are those after interpretation with the use of a lidar simulator (Chepfer et al., 2008) and are taken from simulated years 41–50. Number above a indicates the global mean total cloud fraction, whereas numbers above b and c indicate the mean bias (model minus observed) and the root-mean-square error (in parenthesis).
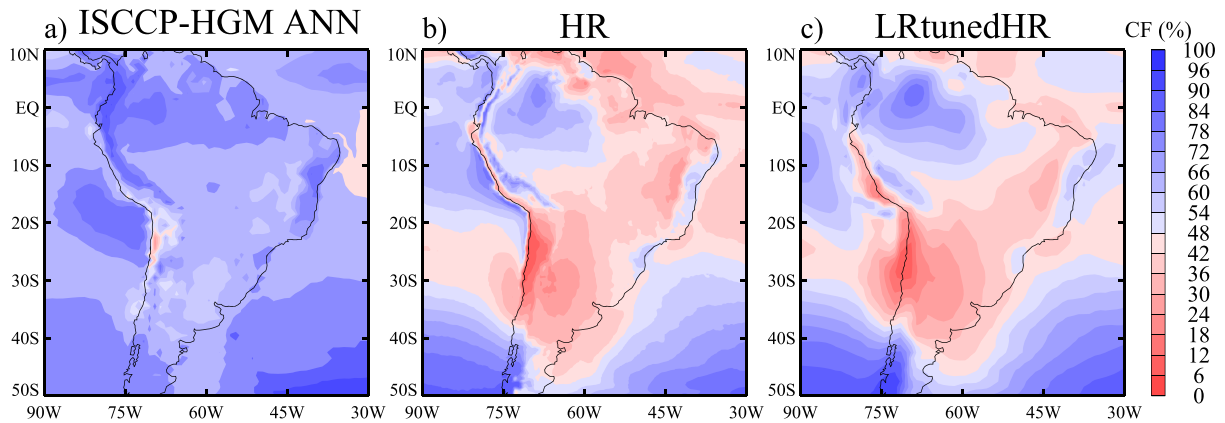
Figure 8. Comparison of annual mean total cloud fraction between E3SM and observations over South America. Total cloud fraction is shown from (a) ISCCP HGM data (Rossow et al., 2016) for the years 1983–2012 at 1° resolution, (b) the HR at 0.25° resolution, and (c) the LRtunedHR at 1° resolution. Model output is for simulated years 41–50, and modeled total cloud fraction is computed from the ISCCP satellite simulator (Klein and Jakob, 1999) using optical depths >0.3.

To appreciate the value of higher resolution in the climatological distribution of clouds requires closer examination of the fine-scale structure of simulated fields. Figure 8 shows this structure over South America with model fields plotted close to their native resolution after interpolation to a latitude-longitude grid. Despite the large-scale biases present (such as the general underestimate of cloud over South America and in the marine stratocumulus region west of the continent), much of the regional structure—particularly near mountain ranges and coastal boundaries—is improved. In particular, v1 HR is able to capture the sharp gradient between cloudy and cloud-free conditions off the west coast. The Atacama desert in particular shows up as an area with extremely low cloud fraction, which jumps discontinuously to much larger cloudiness just offshore. Another improvement is the more narrow cross-width to the northwest-to-southeast feature immediately northeast of the Cordillera Oriental portion of the Andes in Peru and Bolivia.

3.3. Aerosols

Because aerosol observations are lacking for the 1950s and aerosol concentrations have changed greatly since then, this section focuses mainly on sensitivity of simulated aerosol life cycle to horizontal resolution. This is one of the first studies that investigates both the individual and total aerosol responses to resolution change, as previous studies have focused on either dust (Ridley et al., 2013) or carbonaceous aerosols (Ma et al., 2015; Liu et al., 2016). We also examine the impact of model tuning (LRtunedHR vs. LRv1) on aerosol simulations.

Table 4 summarizes the global estimates of aerosol budgets. The annual-mean aerosol optical depth (AOD) is estimated to be 0.123 for the HR run, nearly two thirds of which is contributed by naturally emitted dust and sea-spray aerosols. In the LRtunedHR simulations, coarser model resolution leads to smaller dust and sea-spray aerosol emissions, a consequence of the dependence of these emissions on grid-mean surface wind speeds. The HR model produces more frequent occurrences of high winds due to resolution of fine-scale topography and dynamics (e.g., Ridley et al., 2013). Zhang et al. (2016) have developed a parameterization that accounts for the spatial variability in the subgrid winds for dust and sea-spray aerosol emissions. The impact of applying such a scheme will be explored in future E3SM model development.

Increasing resolution also modifies aerosol removal (i.e., deposition) and long-range transport. To isolate the resolution effect on individual aerosol dry or wet removal pathways, aerosol removal rate, defined as the inverse of the aerosol lifetime (in unit of day^{-1}) for a particular removal pathway (Textor et al., 2006), is calculated and included in Table 4 (numbers in parentheses). By comparing removal rates rather than the actual values of deposition fluxes, complications due to the different aerosol emissions and burdens in the two simulations are avoided. As a result of better-resolved subgrid turbulent transport in HR, the dry deposition rates for dust and sea-spray aerosols are smaller in HR than in LRtunedHR. This result is consistent with previous findings for carbonaceous aerosol (Liu et al., 2016). If aerosols are removed predominantly by dry deposition, this reduction would result in an increase of aerosol burden and residence time at higher resolution.

Table 4
Global Aerosol Budgets

| Characteristic	HR	LRtunedHR	LRv1	Expected PD value
Global AOD	0.123	0.116	0.14	
Dust AOD	0.032	0.025	0.026	
Emission (Tg/year)	4,577.	3,881.	4,702.	1,840. ±49%
Dry depo (Tg/year)	3,309. (0.33)	3,053. (0.42)	3,653. (0.43)	
Wet depo (Tg/year)	1,268. (0.13)	983. (0.12)	1,048. (0.12)	
Burden (Tg)	27.2	19.8	23.5	19.2 ±40%
Lifetime (days)	2.17	1.86	1.82	4.14 ±43%
Sea spray AOD	0.046	0.042	0.049	
Emission (Tg/year)	3,728.	3,434.	3,601.	16,600. ±199%
Dry depo (Tg/year)	2,320. (1.15)	2,451. (1.46)	2,551. (1.39)	
Wet depo (Tg/year)	1,404. (0.70)	983. (0.59)	1,050. (0.57)	
Burden (Tg)	5.52	4.60	5.04	7.52 ±54%
Lifetime (days)	0.54	0.49	0.51	0.48 ±58%
Sulfate AOD	0.013	0.014	0.023	
Dry depo (Tg/year)	15.9 (0.035)	17.1 (0.036)	27.0 (0.035)	
Wet depo (Tg/year)	78.0 (0.173)	78.9 (0.165)	117.2 (0.151)	
Burden (Tg)	1.24	1.31	2.06	2.02 ±25%
Lifetime (days)	4.8	5.0	5.39	4.12 ±18%
BC AOD	0.002	0.003	0.005	
Dry depo (Tg/year)	1.47 (0.059)	1.69 (0.064)	2.95 (0.056)	
Wet depo (Tg/year)	2.84 (0.114)	2.60 (0.099)	4.78 (0.090)	
Burden (Tg)	0.068	0.072	0.145	0.24 ±42%
Lifetime (days)	5.8	6.13	6.85	7.12 ±33%
POM AOD	0.005	0.005	0.01	
Dry depo (Tg/year)	9.41 (0.047)	10.5 (0.049)	16.3 (0.038)	
Wet depo (Tg/year)	21.4 (0.107)	20.2 (0.095)	33.7 (0.079)	
Burden (Tg)	0.55	0.58	1.17	1.7 ±27%
Lifetime (days)	6.51	6.9	8.5	6.54 ±27%
SOA AOD	0.025	0.027	0.025	
Dry depo (Tg/year)	7.32 (0.009)	7.72 (0.009)	9.83 (0.011)	
Wet depo (Tg/year)	61.0 (0.073)	60.5 (0.068)	65.6 (0.076)	
Burden (Tg)	2.3	2.46	2.36	0.57 ±117%
Lifetime (days)	12.3	13.2	11.4	6.70 ±115%

Note. The numbers in parentheses are calculated dry (or wet) removal (or deposition rates) in units of day^{-1}, defined as dry (or wet) deposition/(burden*365.) in Textor et al. (2006). Also shown are the model outputs from LRv1 for simulations with year 2000 aerosol forcing. The means and normalized standard deviations of the Expected PD values are taken from other modeling studies as collected by Liu et al. (2012).

In contrast, aerosol wet removal is enhanced in HR due to improved representation of spatial inhomogeneity in clouds and precipitation (Ma, Rasch, et al., 2015). This enhancement is larger for sea-spray aerosols, which are more susceptible to wet scavenging than dust. Increased wet removal counteracts but does not overcome increased emissions and decreased dry deposition, with the net effect that increasing resolution leads to increased sea spray and dust burdens, AODs, and lifetime. Dust is more sensitive to increased resolution than sea-spray aerosols, with the resulting larger enhancement in dust AOD (+28%) versus sea spray (+10%).
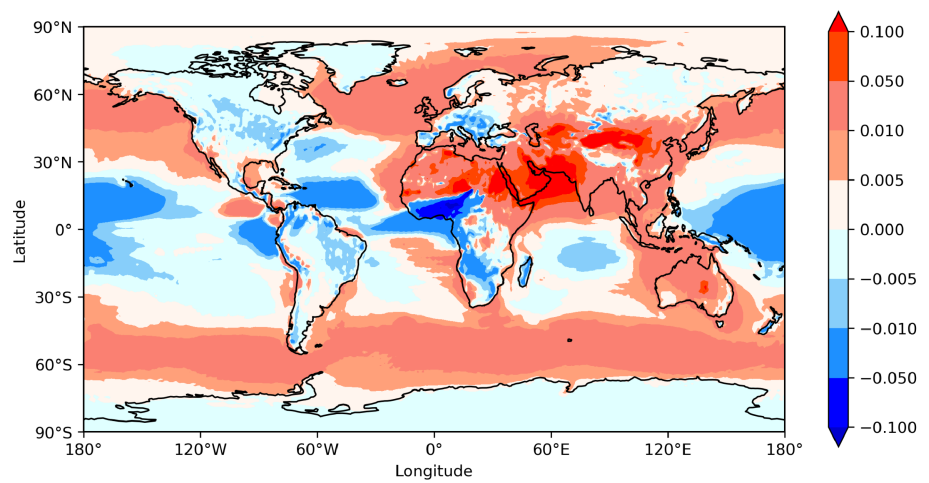
In comparison with natural aerosols, the resolution effects on anthropogenic aerosols (sulfate, black carbon (BC), particulate organic matter (POM), and SOA) are less pronounced because the emissions of the anthropogenic aerosol species or their precursors are inventory based and independent of resolution.
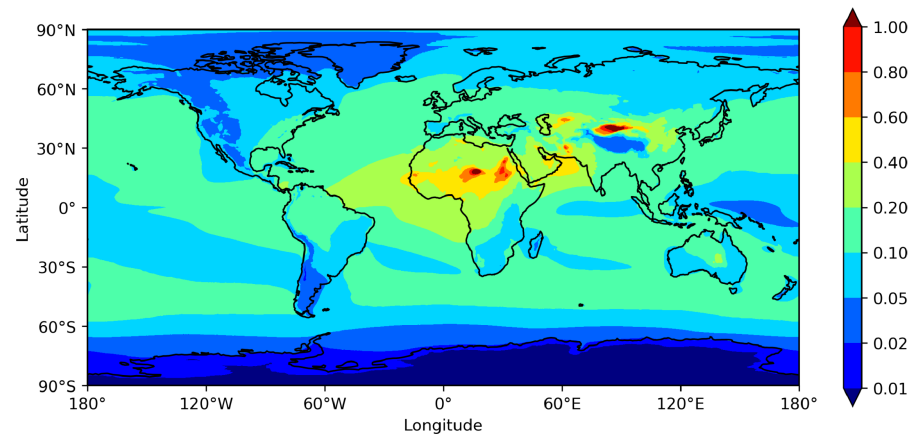
(a) HR
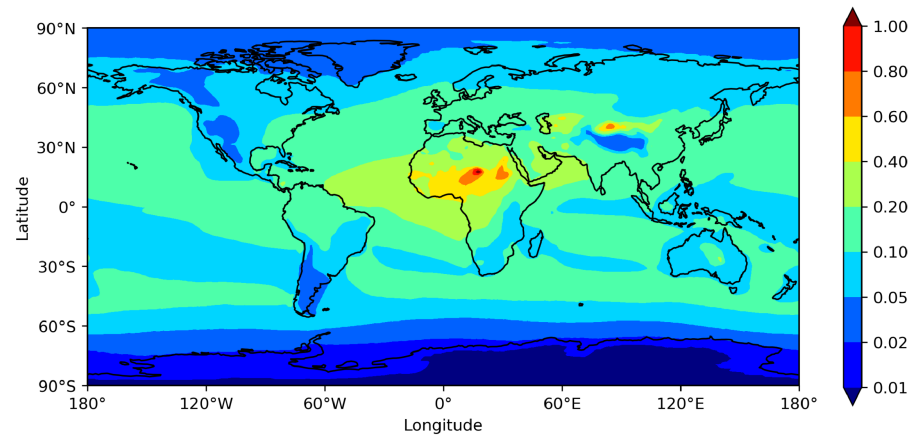
(b) LRtunedHR

(c) HR - LRtunedHR

Figure 9. Annual mean aerosol optical depth (AOD) at 550 nm from the v1: (a) HR and (b) LRtunedHR. Panel (c) shows ΔAOD (HR-LRtunedHR).

Anthropogenic aerosols are, however, subject to the same resolution sensitivity of removal mechanisms as described previously for natural aerosols. This is visible in Table 4 in the form of decreased dry deposition and increased wet deposition rates for anthropogenic aerosols. Note, however, that anthropogenic aerosols are removed preferentially by wet deposition (in terms of total mass) rather than dry deposition due to their finer particle sizes. As a result, the net effect of increasing resolution leads to smaller aerosol burden and AOD, and shorter lifetime for sulfate, BC, POM, and SOA.

In our simulations, enhancement in the natural aerosol AOD dominates the reduction in anthropogenic AOD as model resolution increases, resulting in a net increase in total AOD. The estimated global and annual mean total AOD is about 6% higher in HR than in LRtunedHR.

As expected from industrialization over the last 70 years, global AOD in LRv1 is higher than in both HR and LRtunedHR. As mentioned in the beginning of this section, it is not meaningful to compare total and anthropogenic AOD or burdens in LRv1 with HR or LRtunedHR. Aerosol removal rates and lifetime, however, are less dependent on the emission scenarios and therefore worthy of comparison. Differences between LRtunedHR and LRv1 reflect the impact of different tunings since both runs use the same grid spacing. It is interesting to note that the simulated dust and sea-spray aerosol removal rates and lifetime are very similar between LRtunedHR and LRv1, suggesting that our tuning differences have little impact on aerosol removal and lifetime. Dust and sea-spray emissions, on the other hand, were purposefully decreased during the HR tuning in order for their AODs to match expected real-world conditions. This explains why the global emissions of dust and sea-spray aerosols are similar between HR and the LRv1 despite the resolution differences. Because of the resolution effect on aerosol removal, however, lifetime of natural aerosols is longer in HR than LRv1. For anthropogenic aerosols, the HR simulation results in shorter aerosol lifetime than found in LRv1 or in LRtunedHR. This is the expected result of enhanced aerosol wet removal at higher resolution. Resolution-induced changes in individual aerosol lifetime may lead to further changes in corresponding aerosol distributions and radiative forcings. These changes are beyond the scope of this overview paper but are explored for dust in Feng et al. (2019). Comparison with the expected present-day values in the last column of Table 4 indicates that the simulated aerosol lifetime of natural and anthropogenic aerosols in HR are within the uncertainty range of the AeroCom models (Textor et al., 2006) as summarized in Liu et al. (2012).

Resolution effects on aerosol simulations are much more remarkable on regional scales. In particular, the sign of the regional resolution effect changes depending on the predominant aerosol types. The geographical distributions of the annual mean AOD simulated by v1 HR and LRtunedHR are shown in Figures 9a and 9b. The aerosol climatology in both figures look as expected, with the lowest AOD in polar regions and highest AOD over the major dust source regions such as the Taklamakan and Gobi deserts in Asia and the Sahara desert in Africa. Compared to the LRtunedHR, HR predicts larger AOD over deserts and over midlatitude oceans, consistent with the stronger dust and sea-salt emissions at higher resolution noted in Table 4 (Figure 9c). The maximum increase in regional AOD is much larger than the 6% increase in global AOD, that is, more than a factor of 2 over the Arabian peninsula. Regional reductions in AOD occur mainly over the tropical oceans, associated with the enhanced aerosol wet removal by precipitation and clouds. Changes in sulfate and SOA (not shown) also contribute to the decrease of AODs over some polluted areas, including the Eastern United States and the adjacent North Atlantic, Southern Europe, and Southern Africa. A benefit of HR is the ability to capture highly resolved regional-scale features, especially over the complex terrains of central Asia and south America. These resolution effects on AOD are particularly clear in Figure 9c, which shows AOD differences between HR and LRtunedHR simulations.

In addition, increased horizontal resolution also results in higher aerosol concentrations near the surface and lower concentrations aloft (not shown). This is consistent with increased sea spray and dust emissions and decreased turbulence removal of aerosols in the boundary layer. Weakened vertical transport by convection is also probably playing a role here; previous studies (Wang et al., 2018; Xie et al., 2018) have found that convective transport decreases at higher resolution because more grid-scale lifting is resolved at a higher resolution. The resultant changes in aerosol vertical distributions are likely to have important implications on aerosol radiative forcing and cloud droplet activation—especially on the regional scales—and will be investigated in further studies.

Table 5
Land and Ocean Water Budget

	HR	LRtunedHR	DECK	R15 Obs
Global-ave Pr	2.96	2.82	3.07	2.79±0.15
Ocn Pr	78.6	79.4	77.4	77.7
Ocn Evap	85.6	85.2	84.8	86.5
Land Pr	21.4	20.6	22.6	22.3
Land Evap	14.4	13.9	15.2	13.6
Ocn→Lnd Transport	7.0	5.7	7.4	8.9

Note. Global-ave Pr values are in mm/day, and all other terms are expressed as percents of global-average precipitation. Observational estimates ("R15 Obs") are taken from Rodell et al. (2015), which provides estimates based mainly on satellite measurements with help from data-integrating models and combined in a variational framework that enforces several water and energy budget constraints.

3.4. Hydroclimate

Global-average precipitation is slightly lower in HR than in LRv1 (Table 5). As discussed in Golaz et al. (2019), deriving observational uncertainty for global-average precipitation is challenging. Thus, while E3SMv1 global-average precipitation is larger than the Rodell et al. (2015) estimate used in Table 5 and the GPCP estimate used in Figure 10, it is within uncertainty bounds for Stephens et al. (2012) and Wild et al. (2012). LRtunedHR has lower precipitation, as expected because it is generally colder so has less moisture in the air. The fraction of precipitation or evaporation coming from ocean versus land closely matches observations, and hence moisture flux convergence over land is reasonable, but slightly weaker than observed (Table 5).

Because precipitation depends strongly on topography and small-scale baroclinic instabilities, improved precipitation is one of the main motivations for finer resolution. Thus, it is with keen interest that we examine Figure 10, which provides global maps of precipitation bias. One clear benefit of higher resolution is removal of dry bias to the south of Greenland. This dry bias is caused by erroneous freezing of the Labrador Sea at lower resolutions, which shuts off moisture flux from the ocean surface and therefore starves South Greenland clouds of moisture. See section 3.6 for further discussion of improved Labrador Sea ice distribution at high resolution. Further corroboration of the importance of sea ice extent comes from Figure 6 of Golaz et al. (2019) that shows that Greenland precipitation bias in LRv1 disappears when SST and sea ice extent are prescribed. Wet biases over mountainous terrain (as described in Golaz et al., 2019) are slightly improved at large scales, but by no means alleviated. Unfortunately, increased resolution does not appear to correct E3SMv1's other major precipitation biases. The double ITCZ, dry Amazon, and wet Maritime Continent problems are virtually unchanged with changing resolution.

Even though increased resolution does not improve large-scale features, it has clear benefit at regional scales. This is demonstrated in Figure 11, which zooms in on the west coast of the United States. Because snowpack is the main source of this region's drinking water, capturing the spatial distribution of precipitation over the U.S. West Coast is critical. While LRtunedHR is unable to capture the maxima of orographic precipitation over the region's mountain ranges, HR is better able to capture these features. Nevertheless, HR overpredicts the maximum value in the Sierras, a problem seen in other high-resolution simulations of Sierra precipitation (e.g., Caldwell, 2010; Caldwell et al., 2009). The benefit of high resolution on stream flow is examined later in this section.

In addition to mean precipitation rate, the intensity of precipitation is important for climate impacts. To show the impact of grid resolution on precipitation intensity, the annual mean daily precipitation intensity distributions in the tropics (25°S to 25°N) are examined in Figure 12 with a focus on heavy precipitation. For such histograms, it is more useful to focus on a particular region than on global averages in order to avoid the conflation of disparate changes in different regions. The tropics is a useful region because it dominates global precipitation, it involves significant contributions from both large-scale and convective precipitation at all seasons, and it displays the largest sensitivity to resolution change. Histograms over the continental United States, however, showed similar but muted responses to resolution change (not shown).
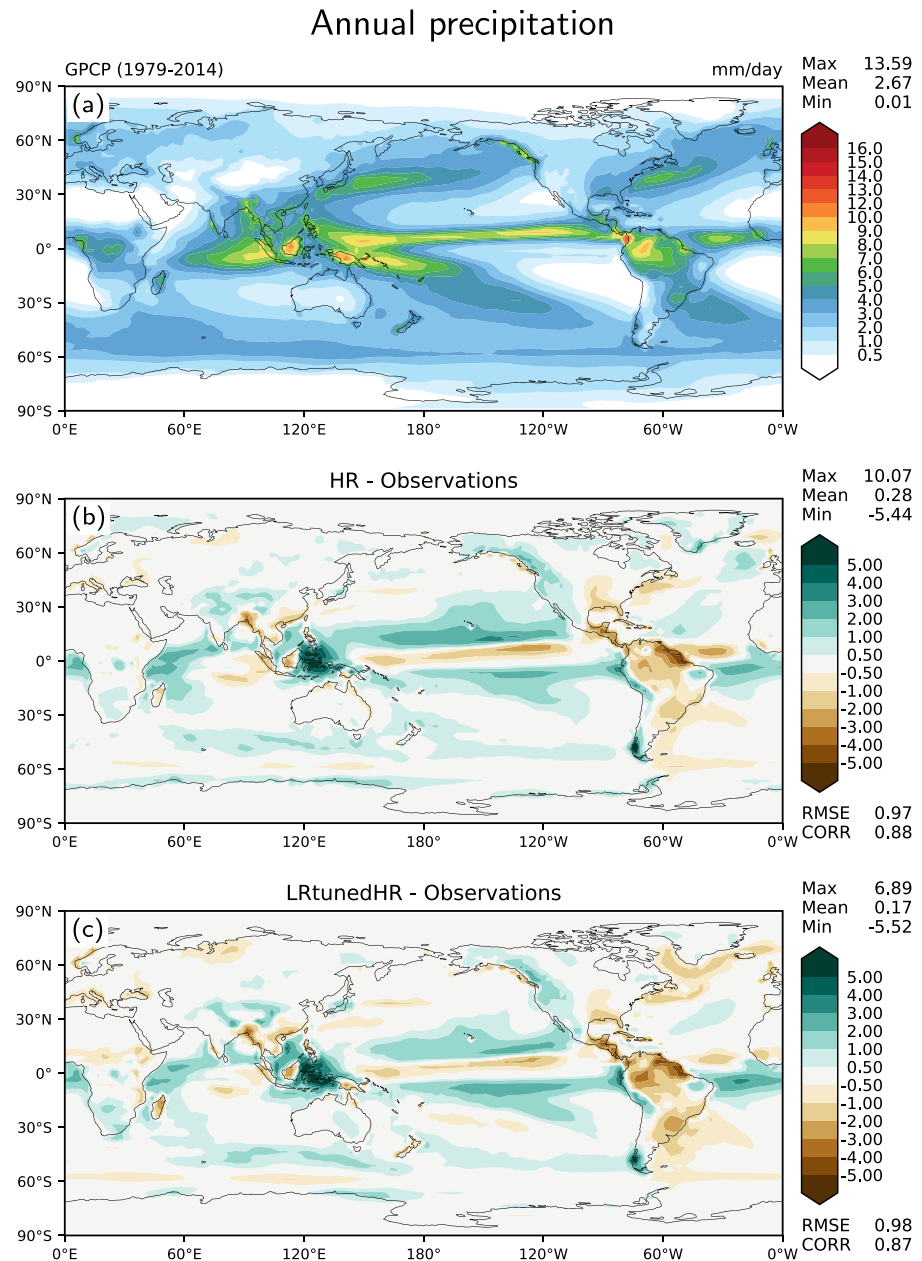
Figure 10. Annual surface precipitation (in mm/day) from (a) Global Precipitation Project (GPCP; Adler et al., 2003) v2.2 observations at 2.5° × 2.5° resolution and model biases for (b) HR and (c) LRtunedHR.

The observations are from the GPCP 1° daily (GPCP1DD) data (Huffman et al., 2001), and TRMM (Huffman & Bolvin, 2018) 3B42 and CPC MORPHing technique (CMORPH; Joyce et al., 2004) at 0.25° × 0.25°. The differences between the observational data sets provide a crude estimate of potential observational uncertainty. We use years from 1998 to 2013. Before aggregating the distribution, CMORPH/TRMM and modeled precipitation rates are remapped to the same 1° × 1° grids as the GPCP1DD data. For each day, each tropical grid cell is assigned to a precipitation bin using 1 mm/day bin widths. Similar to Wehner et al. (2014), we placed all precipitation rates larger than 100 mm/day in a single bin. This results in a change in slope at the right side of the plot.

Both LRtunedHR and HR models generally share the same "too frequent, too weak" problem as shown in other climate models (e.g., Stephens et al., 2010; Trenberth et al., 2003). They largely overpredict the frequency of precipitation occurrence for rainfall rates less than 15 mm/day and underestimate rainfall between
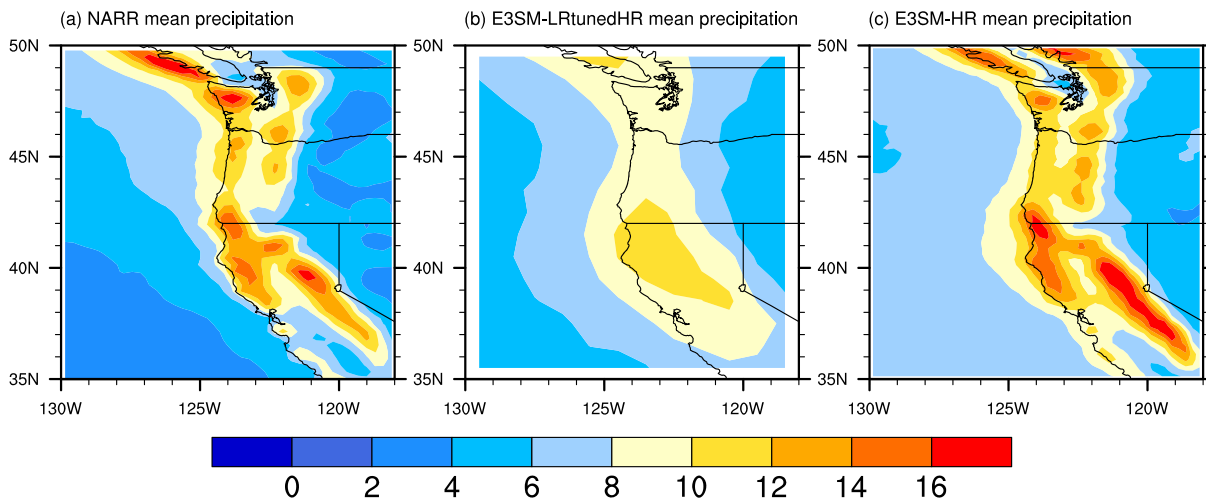
Figure 11. Climatalogical pattern of precipitation (in mm/day) over the west coast of North America as simulated by The North American Regional Reanalysis (NARR; Mesinger et al., 2006) and E3SM v1 model versions.

20 and 70 mm/day (Figure 12a). The high-resolution model clearly shows an increase in the frequency of heavy precipitation rates (>50 mm/day) compared to its low-resolution counterpart. For extreme precipitation rates in the tropics (>70 mm/day), the v1 HR simulation is within the range of the observations while the LRtunedHR simulation is considerably lower than the observed frequency.

The increase of heavy precipitation as model horizontal resolution increases is shown over both land and ocean (Figures 12b and 12c), though the increase over land is more significant, bringing the simulated distribution of intense precipitation closer to the observations. This increase is due to resolved-scale (stratiform) precipitation (Figure 12e), which increases with increasing resolution for all bins greater than <0.2 mm/day. Convective precipitation, on the other hand, actually decreases with increasing resolution for all bins greater than 10 mm/day.

The diurnal cycle of precipitation is another long-standing model bias, which one might hope is ameliorated by higher resolution. There is a slight tendency for precipitation to move later in the day over larger islands of the Maritime Continent, but in general, increased resolution does little to improve the model's tendency for precipitation to peak too early in the day (not shown).

In Figure 13, the land-model-simulated mean annual runoff averaged across the year 21–50 simulation period for both HR and LRtunedHR setups is compared with the half degree University of New Hampshire (UNH)/Global Runoff Data Centre (GRDC) Composite Monthly Runoff data (hereafter "GRDC runoff"; Fekete & Vorosmarty, 2011). The 10-year-long (1986-1995) runoff data are essentially a data assimilation product that preserves the spatial specificity simulated by a large-scale hydrologic model (Water Balance Model, WBM) while constrained by the in situ streamflow discharge measurements from the Global Runoff Data Centre (GRDC). Therefore, it is useful for spatial pattern comparisons at river basin scale level but may not be suitable as a benchmark product for bias identifications at the gridcell level. Our comparisons (Figures 13a, 13b, and 13c) suggest that both HR and LRtunedHR capture the general spatial pattern of the GRDC runoff. In arid regions such as the western United States, southern Africa, central Asia, and Australia, the model has a wet bias, while in relatively humid areas such as the Amazon basin and central Africa, dry bias is observed. These findings are generally consistent with the precipitation results (Figure 10). The difference between HR and LRtunedHR (Figure 13d) indicates that the HR tends to better represent runoff heterogeneity in mountainous regions (e.g., Andes Mountains and Tibetan Plateau). In addition, the regions noted as having dry bias in LR are noticeably wetter in HR (e.g., central Africa) and regions that are too wet tend to be drier in HR (e.g., western United States).

To evaluate the performance of the E3SM river model (MOSART), we compared the MOSART-simulated streamflow (years 21 to 50) with the Global Streamflow Indices and Metadata Archive (GSIM; Do et al., 2018) at 4,515 locations (Figure 14). These gauges were selected to accommodate the need for global spatial
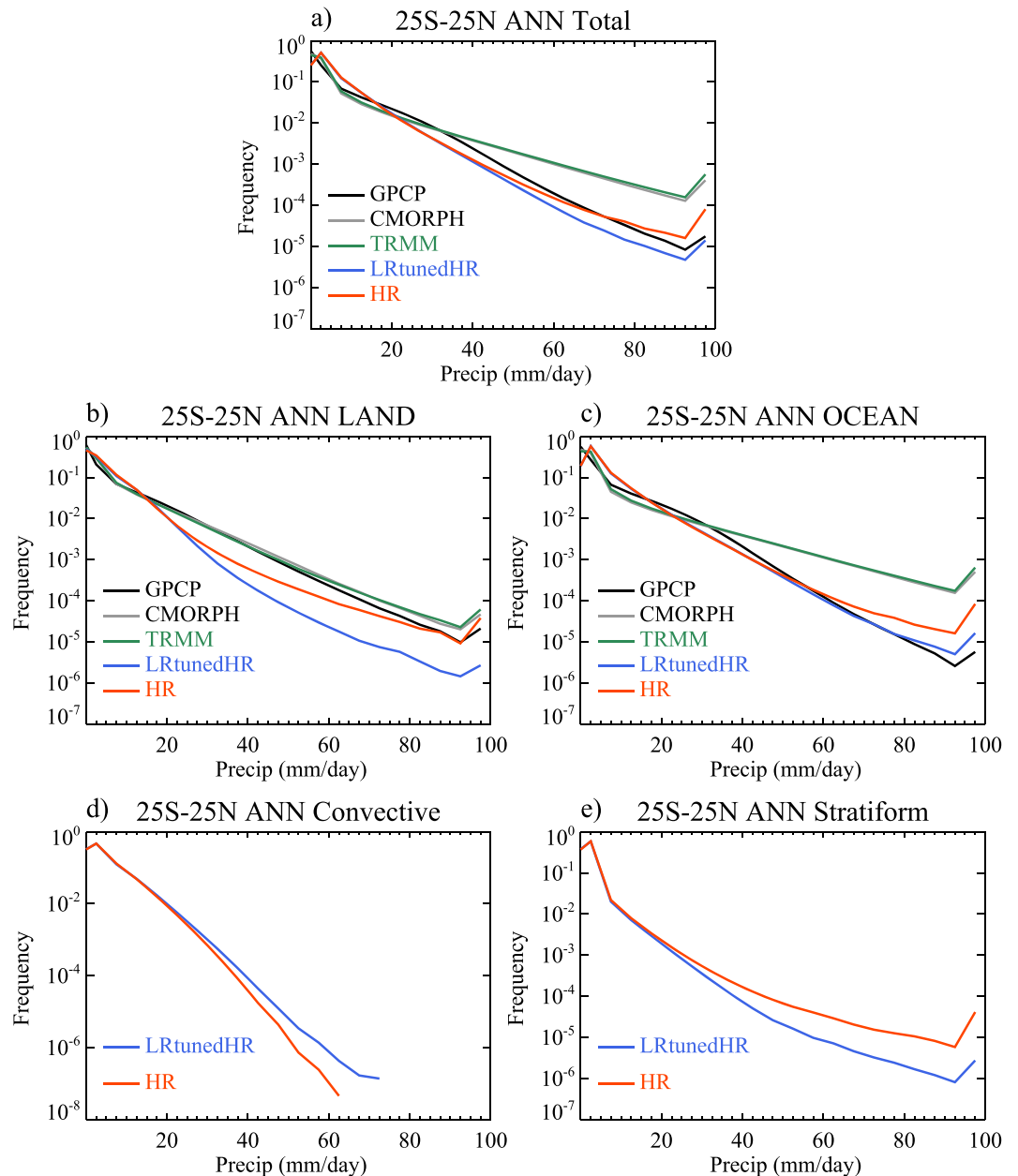
Figure 12. Annual mean precipitation frequency of daily precipitation as a function of precipitation intensity between the models and observations over 25°S to 25° N; black: GPCP1DD; gray: CMORPH; green: TRMM; blue: LRtunedHR; and red: HR. Daily mean precipitation rates (unit: mm/day) gridded at 1° × 1° are used to derive the distribution. Any precipitation rates larger than 100 mm/day are assigned to the last bin that results in an uptick at the end of the plot. (a) Total precipitation, (b) total precipitation over land, (c) total precipitation over ocean, (d) convective precipitation, and (e) stratiform precipitation. Note that a different scale for *Y* axis is used for convective precipitation.

coverage and for having river gauge drainage areas accurately represented in both HR and LRtunedHR models. Unsurprisingly, Figure 14 shows that biases in runoff carry through into streamflow. Nonetheless, the overall performance of both simulations is very good, with R^2 over 0.8. In sites with large discharge (e.g., $>10^4$ m^3 s^{-1}), HR's performance is generally better than LRtunedHR.

In addition to the mean annual streamflow, we also examined the seasonality of the streamflow (Figure 15) with color indicating the peak month of the hydrograph and the size of each dot indicating the seasonality index (SI) of the streamflow, which ranges from 1 to 12 and quantifies the level of seasonal variations of the
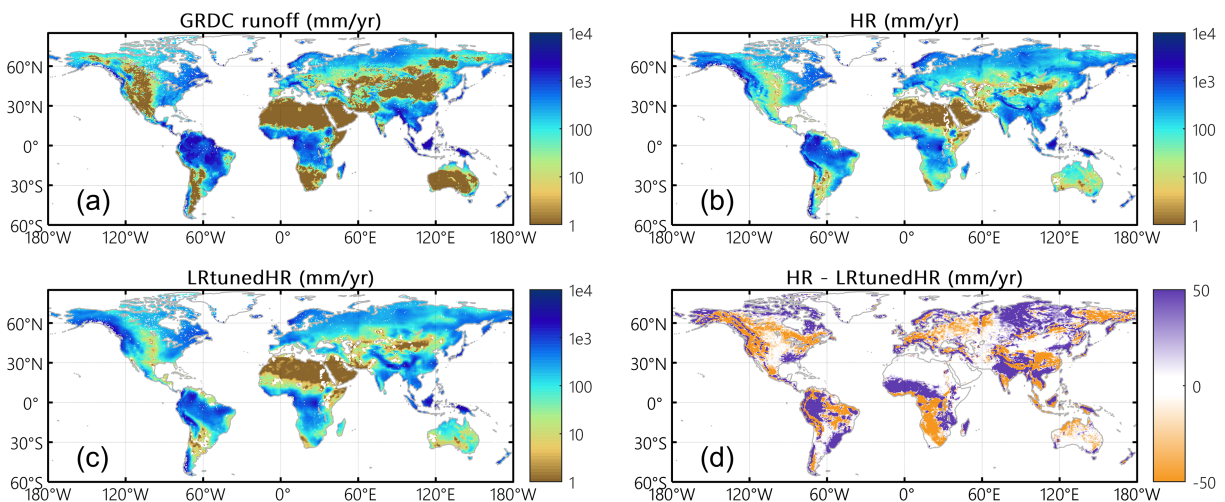
Figure 13. Mean annual runoff (mm/year) from (a) GRDC, (b) HR, (c) LRtunedHR, and (d) the difference between HR and LRtunedHR.

hydrograph. SI equal to one indicates uniformly distributed hydrograph across the year, that is, no seasonal variation; SI equal to 12 indicates that peak streamflow occurs in a single month while the rest months are equal, that is, strongest seasonal variability. The equations for SI calculations can be found in Golaz et al. (2019). The results suggest that simulated streamflow is generally a good match with observations in terms of both seasonality and peak timing, though simulated SI is slightly overestimated for most rivers (Figure 15d). This is not surprising because the model did not represent dam operations, which usually lower the seasonality of the hydrographs. We do notice that SI is underrepresented in northern Eurasia in HR for unclear reasons (Figure 15b). There otherwise does not appear to be any systematic change in SI as a function of resolution.
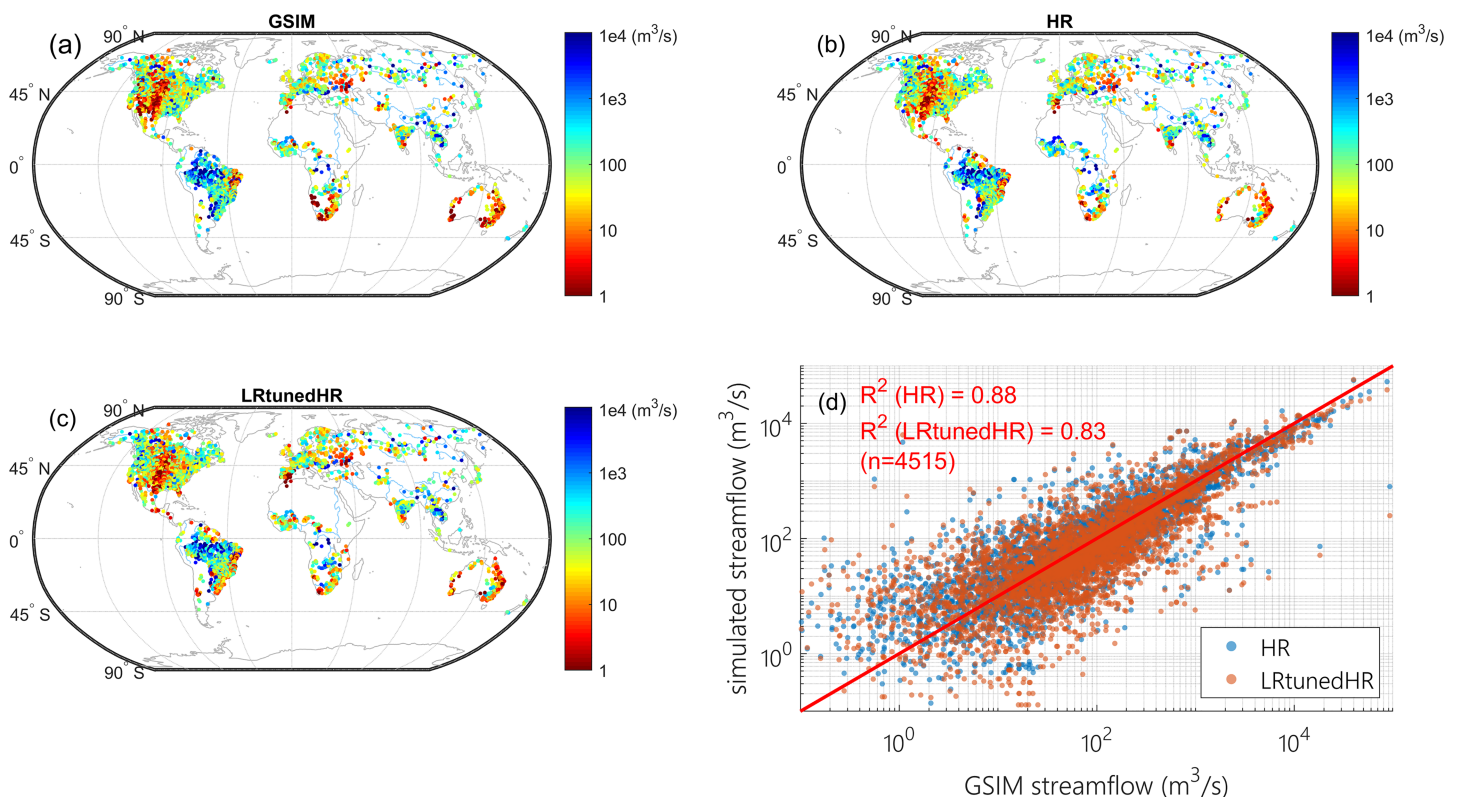
Figure 14. Mean annual streamflow at 4,515 gauge locations from (a) GSIM, (b) HR, (c) LRtunedHR, and (d) the scatter plot between GSIM observations and simulated streamflow.
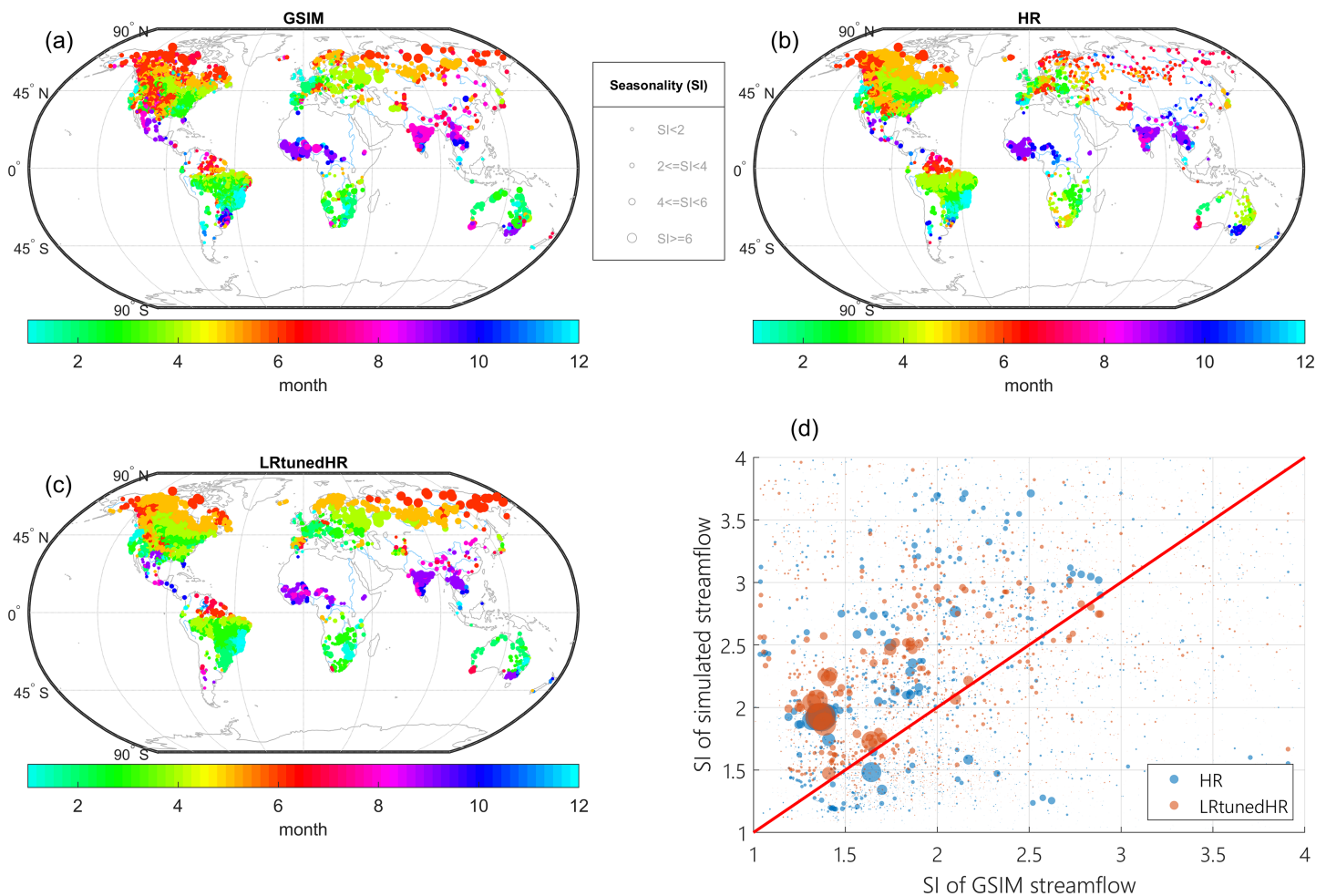
Figure 15. Seasonality index (SI, size of the dots) (a) for GSIM gauge observations, (b) HR, and (c) LRtunedHR. Panel d provides a scatter plot of observed SI against simulated SI with size of the dots indicating magnitude of the discharge.
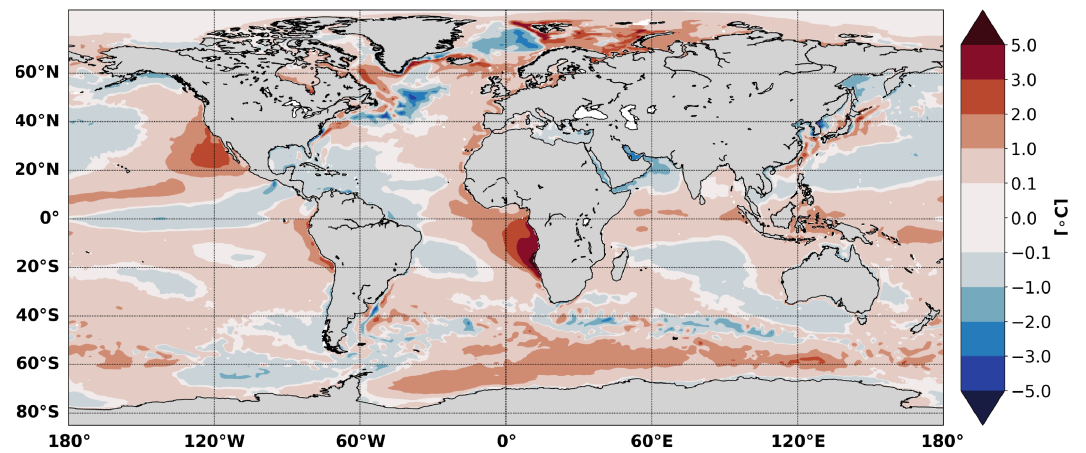
3.5. Ocean

A representation of the global upper ocean is shown in Figures 16–18 as annual averages of Sea Surface Temperature (SST), Sea Surface Salinity (SSS), and Mixed Layer Depth (MLD) computed using the 0.03 kg/m^3 density threshold criterion. Moving from low to high horizontal resolution has generally beneficial effects. In particular, the cold bias in the subtropical gyres is greatly reduced at HR, and the extremely cold and fresh biases in the Labrador Sea disappears. This latter bias is most likely related to sea ice, as we shall see in section 3.6. Also noticeable is the general improvement in SSS between LRtunedHR and HR, with LRtunedHR showing a fairly strong fresh bias in most of the global ocean. The origins of this fresh bias are being investigated in a separate study but that are believed to be at least partially due to ocean mixing. Biases in LRtunedHR SST and SSS look very similar to those shown for LRv1 in Golaz et al. (2019), implying that the aforementioned issues are intrinsic to the low-resolution version of the model rather than the result of poor atmosphere tuning.

One problem that increased resolution does not fix is the warm SST bias in the Eastern Boundary Current regions (e.g., west coast of North America, off the coast of Peru, and in the Buenguela region of South East Africa). Warm bias is caused by an unrealistic lack of stratocumulus clouds in this region (as shown in Figure 7) combined with overly weak alongshore winds. Lack of improvement in eastern boundary current SST with increased resolution is surprising; all other high-resolution modeling studies we know of found robust improvements in this feature with increasing resolution. Improvement is expected due to improved fidelity of alongshore winds, which is likely the result of a better representation of orography (Gent et al., 2010). Upwelling does strengthen in HR, stratocumulus do increase, and SST biases do improve

(a)   SST observations (Hadley/OI)
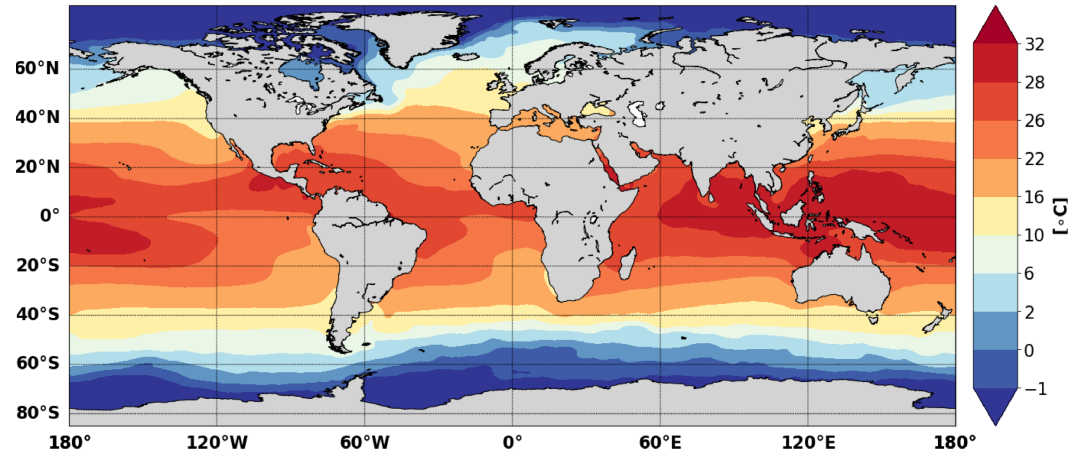
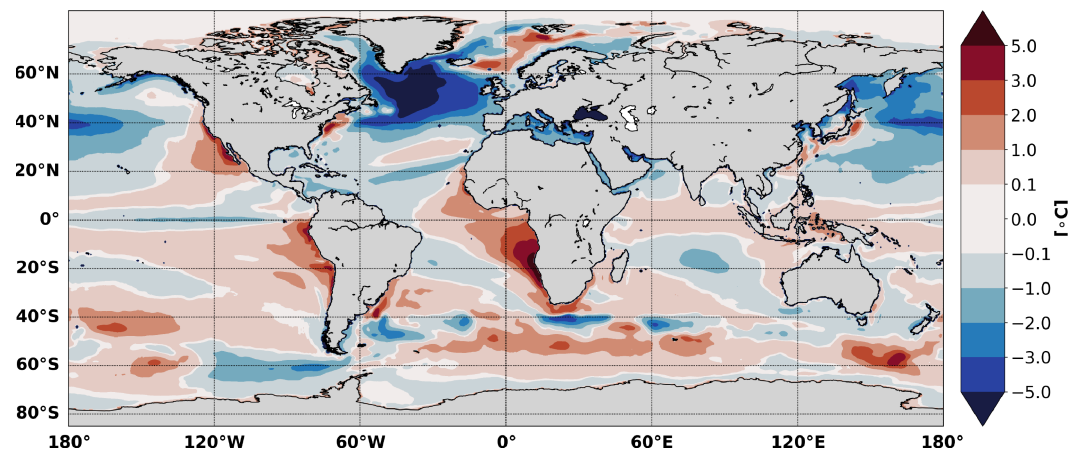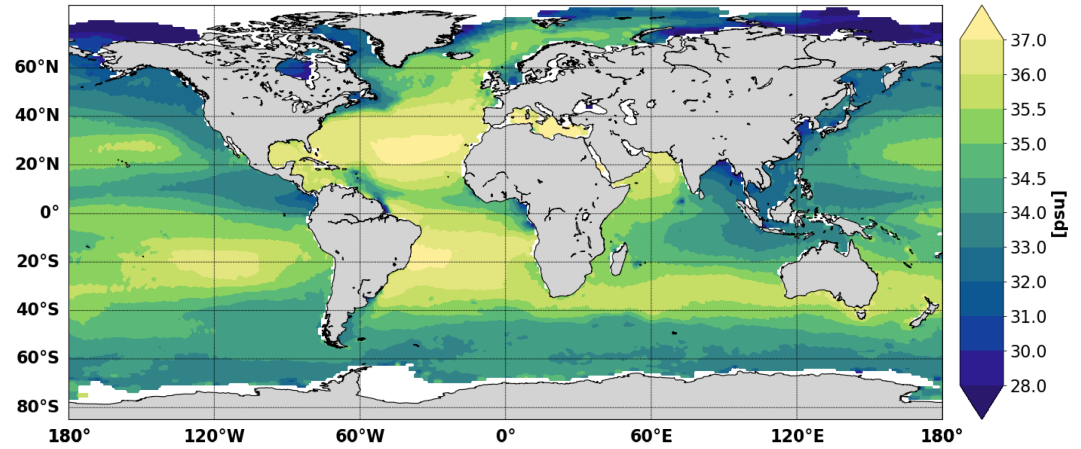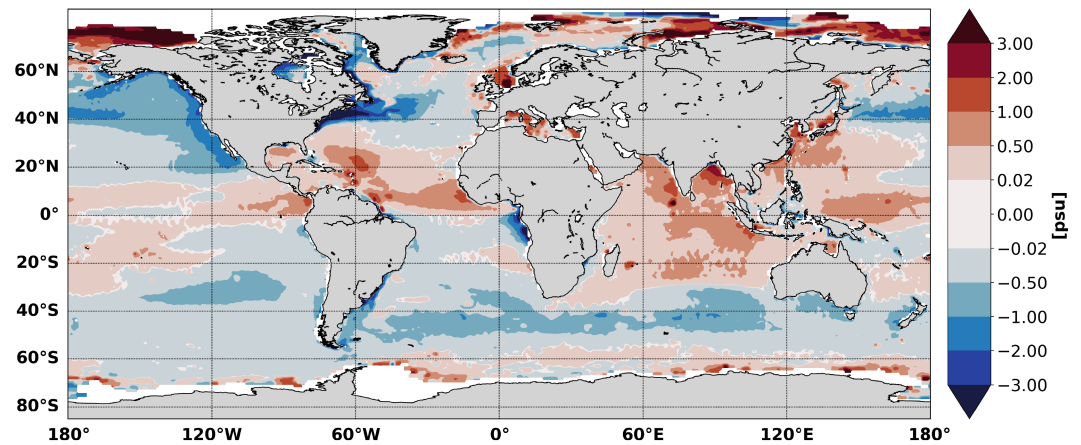(b)   Model-Obs bias (HR)

(c)   Model-Obs bias (LRtunedHR)

Figure 16. Annual-mean SST from (a) the Hadley/OI observational product (Hurrell et al., 2008); (b) HR bias with respect to observations (model-obs); and (c) LRtunedHR bias.
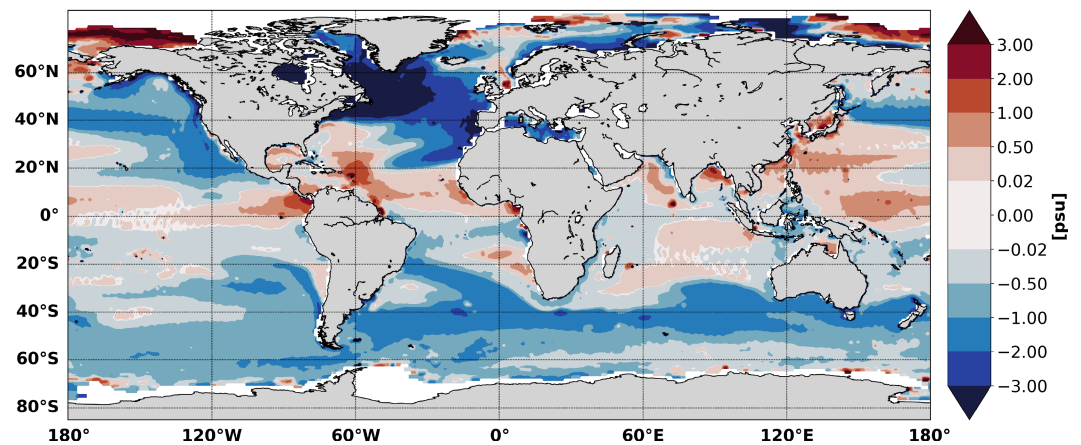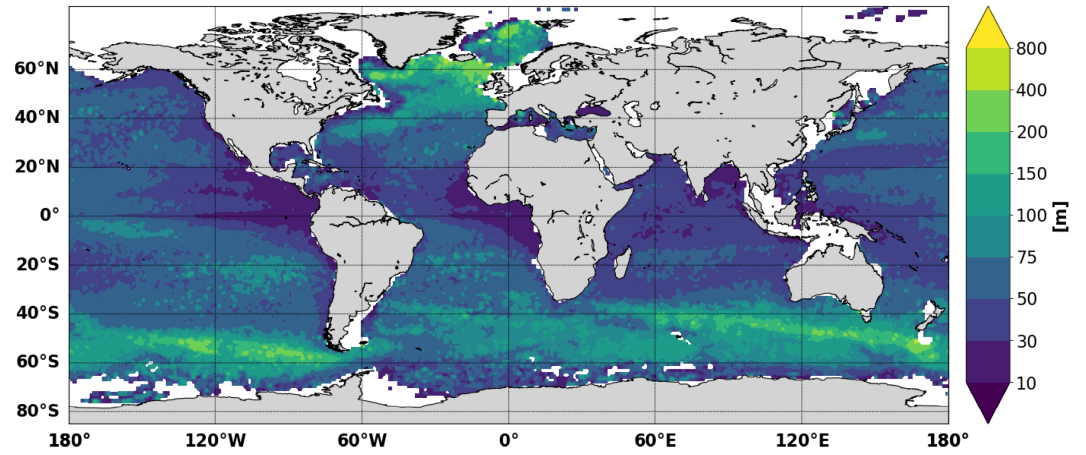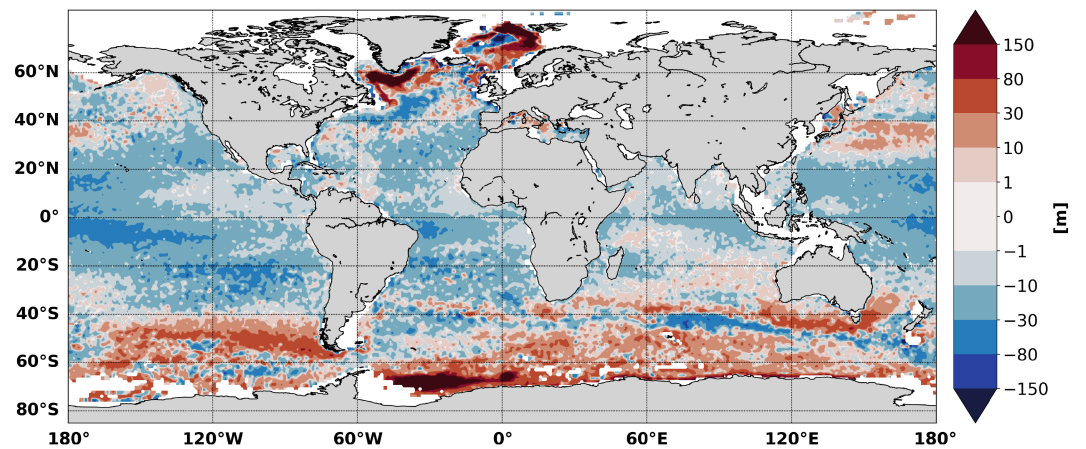
Figure 17. As for Figure 16 but for the annual-average SSS. Observations are from the Aquarius satellite product (Wentz et al., 2014, years 2012-2013).
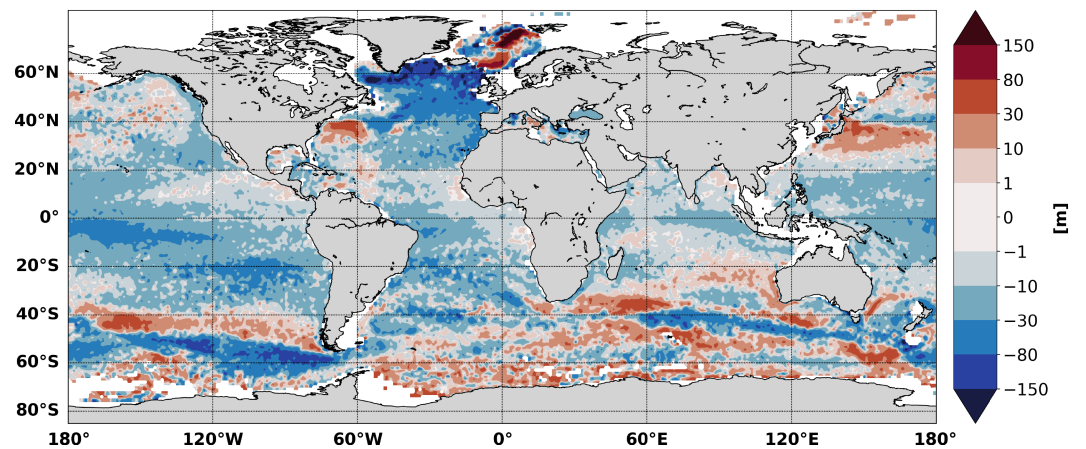
Figure 18. As for Figure 16 but for the annual-average MLD. Observations are from the Argo-based climatology of Holte et al. (2017).
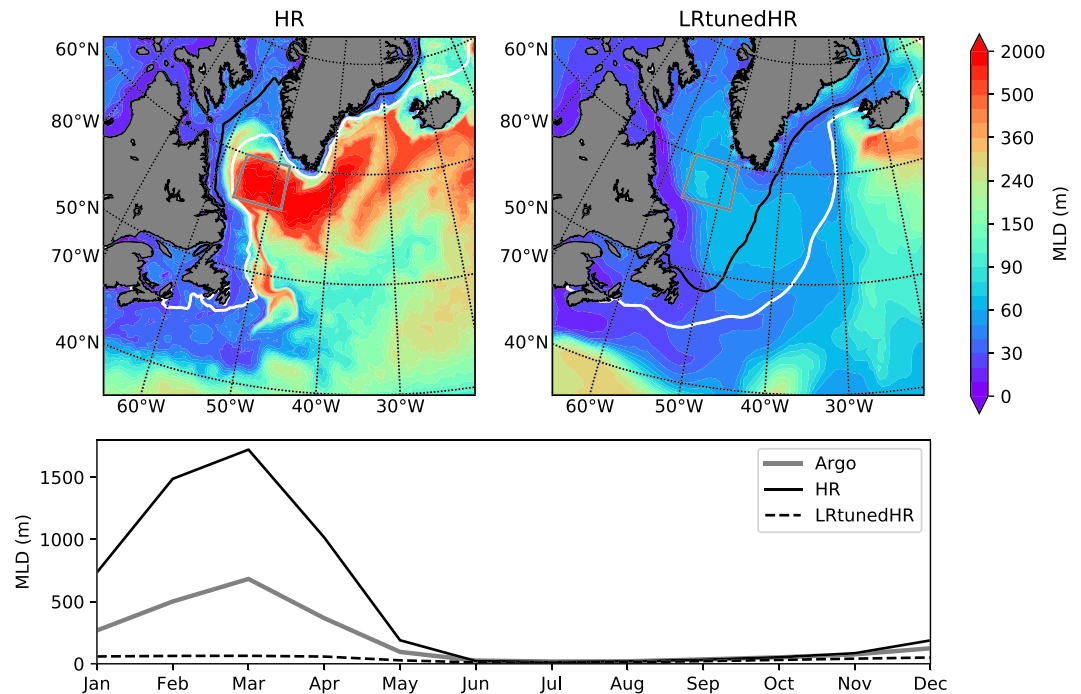
Figure 19. March mean MLD (0.03 kg/m^3 density threshold) in the Labrador Sea in HR (upper left panel) and LRtunedHR (upper right panel). White and black contours show the 15% and 85% sea ice fraction, respectively. Lower panel shows the seasonal cycle of MLD averaged over the gray box as shown in the upper panels for both simulations. Gray thick curve shows the same averaged MLD from Argo-based climatology of Holte et al. (2017).

with increasing resolution in E3SMv1 (not shown), but only along a very narrow strip just off the coast. Understanding why resolution sensitivity of eastern boundary upwelling has such small areal extent is ongoing research. We also note that while the SST bias patterns suggest a global average warm bias, the years 21–50 average has cooled relative to the initial condition (consistent with Figure 3a).

Mixed layer depth biases with respect to an Argo-based climatology (Holte et al., 2017; Figure 18) reflect a variety of upper ocean processes that are not explicitly resolved in models and whose representation depends on mixing parameterizations and shifts in buoyancy fronts (e.g., a shift in the Antarctic Circumpolar Current [ACC] relative to the Argo product). Similar to other Earth System Models, E3SM simulations tend to have larger MLD biases at higher latitudes where strong convection tends to occur. The large positive bias in HR in the Weddell Sea is due to an extensive Weddell Polynya that lasts a little over a decade (see also section 3.6) and is associated with strong oceanic convection, and a warm and salty surface anomaly. The MLD in the Labrador Sea is too shallow in LRtunedHR, but too deep in HR. This is also visible in the regional MLD seasonal cycle in Figure 19. While the lack of convection in the Labrador Sea in LRtunedHR is most likely due to the unrealistically high sea ice coverage in this region, associated with the accumulation of freshwater and thereby strong stratification probably resulting from too weak Irminger and West Greenland Currents, the overly convective Labrador Sea in HR cannot be easily explained by surface temperature and salinity biases and needs further investigation (a separate study focusing on the Labrador Sea is underway).

3.6. Sea Ice

The most important improvement to E3SM′s surface polar climate in HR relative to LRtunedHR manifests as a correction to the winter sea ice edge in both the Northern and Southern Hemispheres. Figure 20 summarizes this improvement, where we have used the 1979–1999 Meier et al. (2014) National Oceanic and Atmospheric Administration Climate Data Record (NOAA CDR) as an ice edge benchmark for which there are regular measurements, contrary to the 1950s. At the time of writing the Walsh et al. (2019) industrial-era ice extent data set was being revised and is not included in this study. Comparing the ice edge of HR with LRtunedHR and observations from the last 20 years of the twentieth century, sea ice extent is more realistic at high resolution owing to reduced bias in the Labrador Sea, the Sea of Okhotsk, and around most of the Southern Ocean cryospheric perimeter (Figure 21). Even though there is temporal mismatch between the
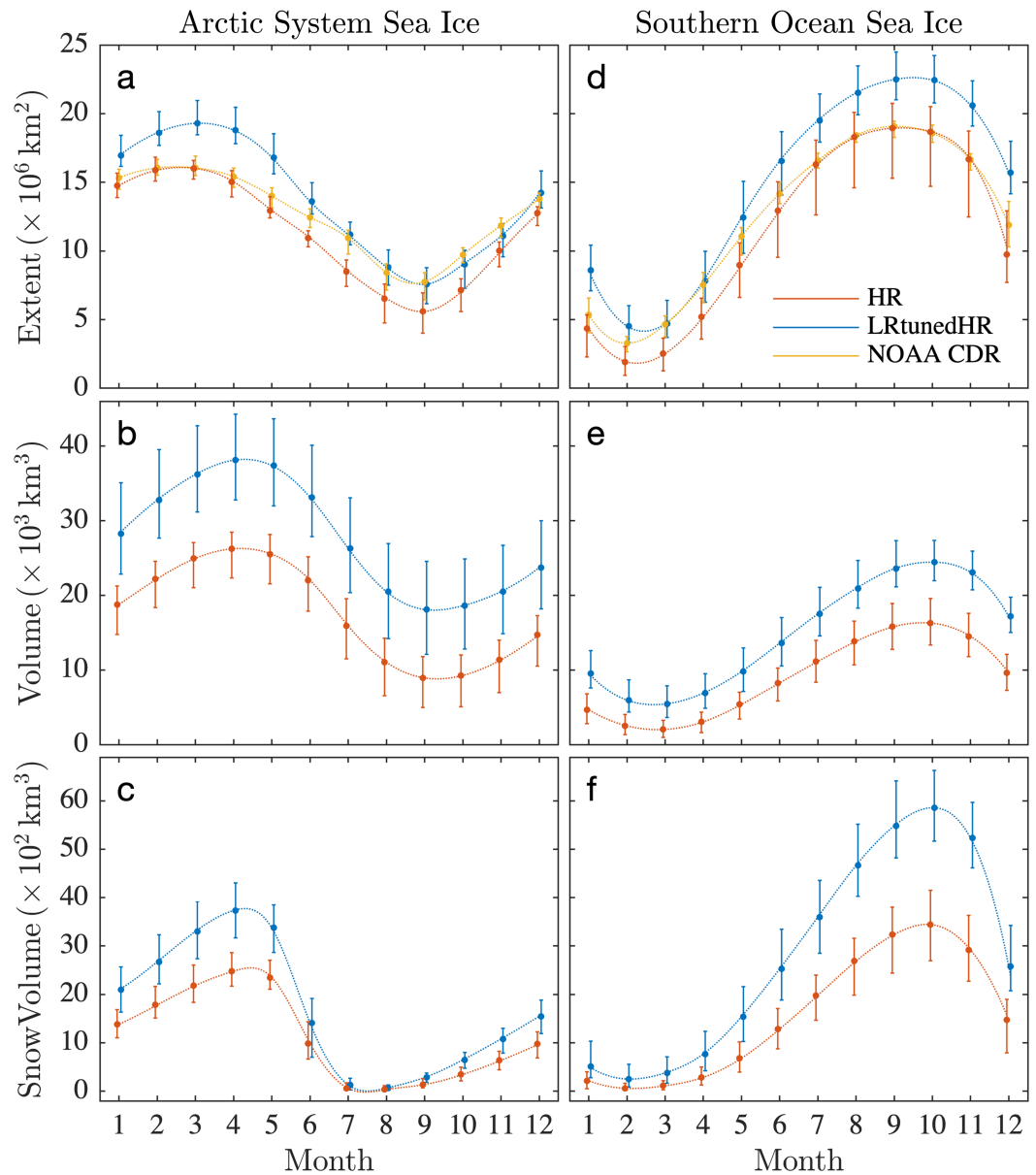
Figure 20. Seasonal evolution of sea ice for high- and low-resolution simulations for the Northern (a–c) and Southern (d–f) Hemispheres. Sea ice extent may be compared with the Meier et al. (2014) NOAA Climate Data Record (CDR) for the period of 1979–1999 as an approximate guide to real-world 20th century values. Points indicate monthly median for the designated 30-year analysis period, and lines indicate the maximum and minimum monthly means. The numeric values on the *x* axes correspond of consecutive months of the year, where 1 represents January and 12 represents December.

perpetual 1950 simulations and 1979–1999 measurements, sea ice extent in the 1950 LRtunedHR run is similar to predictions from the late 20th century in the LRv1 simulation, suggesting that the temporal mismatch in observed versus modeled eras is not a cause of LRtunedHR bias (Golaz et al., 2019). There is also a strong similarity between LRv1 historical ensemble members' ice edge in the Southern Ocean and LRtunedHR.

In the case of the Labrador Sea low-resolution sea ice bias, by far the largest marginal sea bias, the saline ice and its melt water separate early from the west Greenland current and spread toward Newfoundland along the streamlines in Figures 22e and 22f. As a consequence, the Labrador Sea becomes stably stratified, and oceanic mixing is shut down (Figure 19). By contrast, this does not occur at high resolution, where sea ice in eddies spinning off the West Greenland Current soon melt away, animated in Roberts et al. (2019a). The other significant Northern Hemisphere ice-edge bias occurs in the Sea of Okhotsk, where low-resolution
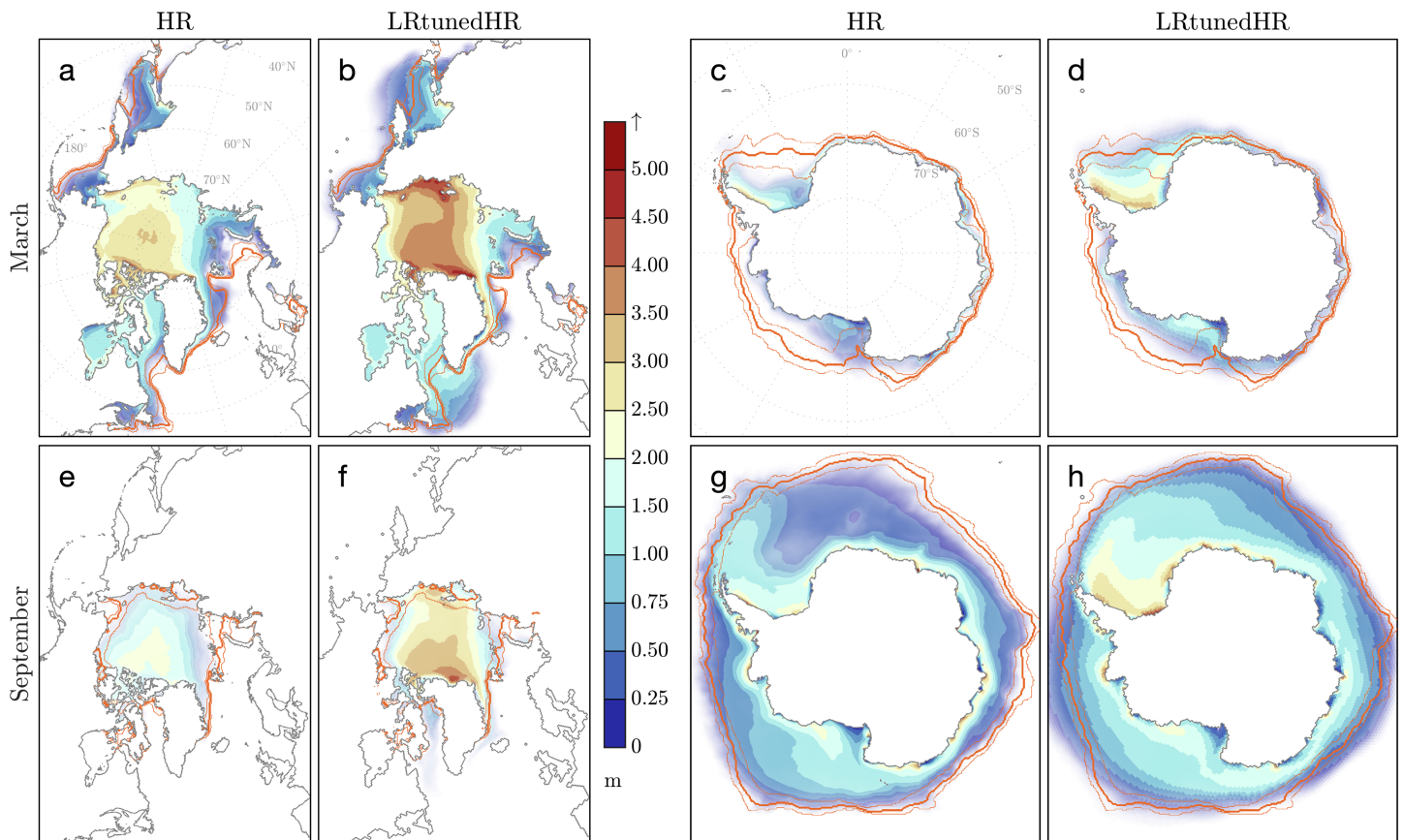
Figure 21. Mean March (a–d) and September (e–h) sea ice thickness for HR and LRtunedHR simulations over the 30-year climate averaging period. Data are rendered on native model grid cells with opacity (color density) determined by mean sea ice concentration, where fully saturated colors in the legend indicate 100% ice cover. The observed 1979–1999 Meier et al. (2014) NOAA sea ice extent is traced in bold orange, equivalent to the 15% concentration contour. Extent contours appear in thin orange either side of the main measured ice edge for ±1 standard deviation in the observed ice fraction, where a large deviation from the bold orange contour indicates high interannual variability. Native model coastlines indicate the closure of Nares Strait (b and f) and where the top of the Antarctic Peninsula is pinched off (d and h) in the low-resolution model.

bathymetry contribute to poor fidelity and where there is also high observed interannual ice-edge variability, as indicated by the thin orange ±1 standard deviation 15% concentration measurements in Figure 21.

High-resolution topobathymetry, greater extremes in wind and current forcing, and increased oceanic eddy kinetic energy combine to elevate the median and maximum sea ice drift speed in HR relative to LRtunedHR. Figures 22 and 23 demonstrate that sea ice is accelerated in HR relative to its low-resolution analog. In particular, sea ice reaches much greater drift speeds in coastal or slope currents surrounding Greenland, next to Baffin Island, along the East Antarctica coast and in the Weddell Sea, verified independently of Figures 22 and 23. These changes manifest from the sharper horizontal gradient $\nabla_h(f/H)$ in HR relative to LRtunedHR, for the Coriolis parameter f, and bathymetric depth H. There is also improved steering of the winter Southern Ocean ice edge, including from eddies downstream of Drake Passage, relative to LRtunedHR.

As seen in the E3SMv0 high-resolution coupled simulation (Kurtakoti et al., 2018), the HR simulation evolves intermittent Weddell Polynyas, imprinted in the model climate near Maud Rise (Figures 21j and 21k) but not present at low resolution (Figure 21l). This high-resolution feature results from Taylor columns and mesoscale eddies around the Maud Rise seamount, where polynyas are initiated, a further ramification of $\nabla_h(f/H)$ fidelity; the associated anticlockwise surface circulation around Maud Rise is evident in HR but not LRtunedHR sea ice drift (Figures 23d and 23h). Evolution of one such Weddell Polynya may be seen in an animation of the last 10 years of HR sea ice thickness in Roberts et al. (2019b). Improved coastal defini-tion also means that East Antarctic polynyas are better resolved in HR relative to LRtunedHR; an example of the high-resolution evolution of one such case, the Mertz Polynya, may be seen in Roberts et al. (2019c).
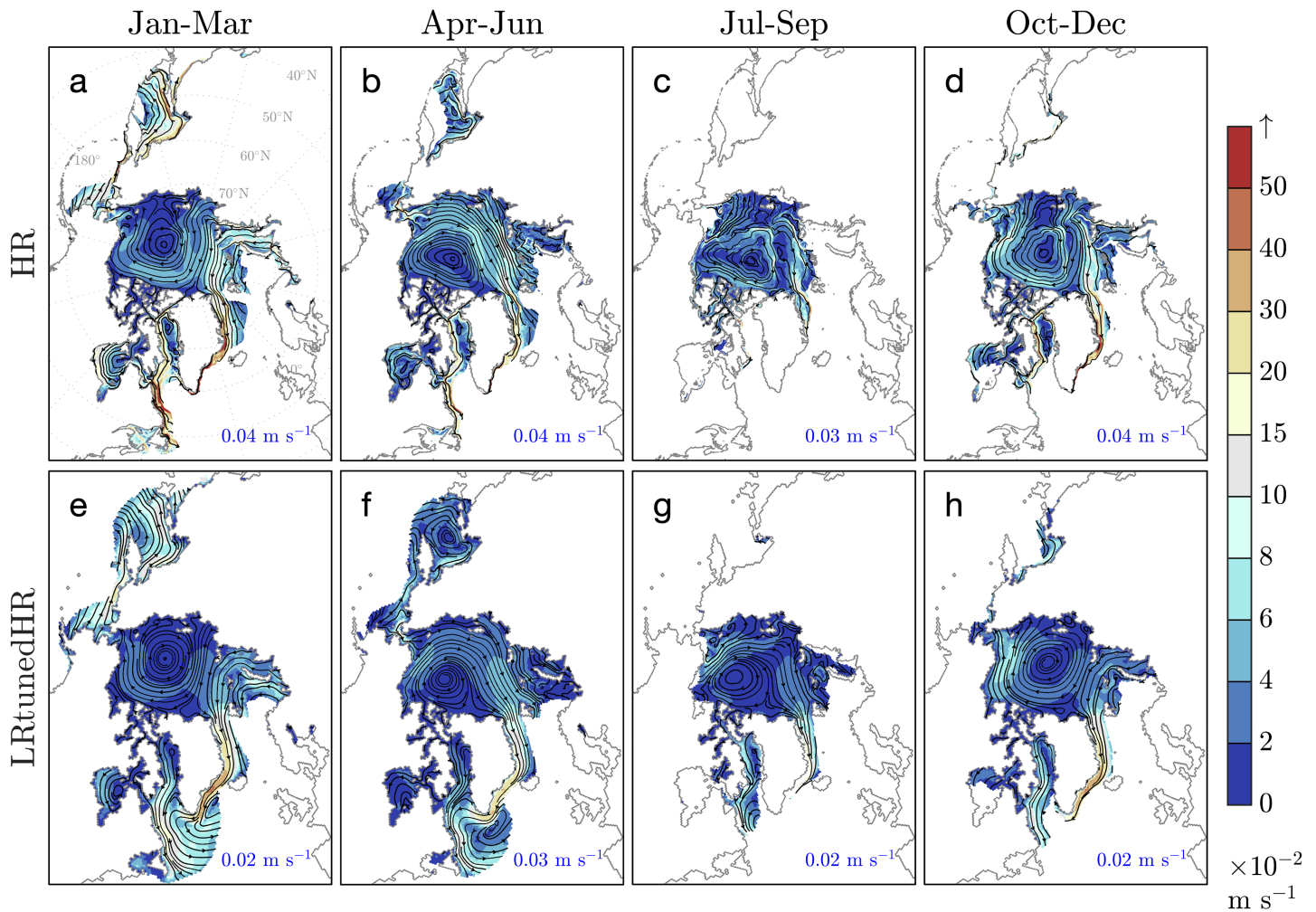
Figure 22. Seasonal mean Northern Hemisphere sea ice streamlines (black) and drift speed (shading) for years 21–50 of HR and LRtunedHR, truncated at 15% sea ice concentration. Median sea ice speeds from monthly model means within the respective sea ice zones appear in blue.

Increased ice production in polynyas at high resolution along the Antarctic coast is a function not just of coastal topology, but also of better resolved katabatics. In the Arctic, coastal definitions are important for refining ice and ocean fluxes through the Canadian Archipelago, including in Nares Strait, estimated to have a similar oceanic southward through-flow as Lancaster Sound (McGeehan and Maslowski, 2012). It may also have a significant southward freshwater flux in the form of sea ice (Kwok et al., 2010). However, Nares Strait is completely closed-off in low-resolution v1 E3SM simulations, as seen in the native model coastlines in Figure 21(c) and (i).

Even though key polar surface circulation features including the Beaufort Gyre, Transpolar Drift, Ross Sea Gyre and Weddell Sea circulation are better resolved at high resolution (Figures 22 and 23), HR exhibits an ice cover that is too thin. We are working to quantify this bias precisely in ongoing work using ICESat, CryoSat-2, and ICESat-2 emulators that compare satellite-derived and model sea ice freeboard. However, here, we surmise that thin ice is likely the cause of a negative ice extent bias everywhere in summer at high resolution (Figures 20 and 21). Mean ice thickness against the Canadian Archipelago never exceeds 3.5 m in HR except in slender coastal margins, even though submarine observations from 1958 to 1987 indicate vast year-round ice buildup exceeding 5-m draft, with mean draft of 7–8 m in some Canadian coastal margins (Bourke & McLaren, 1992). In the Weddell Sea, where Southern Ocean perennial sea ice is most prevalent, upward looking sonar deployed by Alfred Wegener Institute in the early 1990s revealed winter buildup exceeding 3-m draft against the Antarctic Peninsula (Harms et al., 2001; Strass and Fahrbach, 1998; Roberts, 2005) as compared with HR mean ~1,950 thickness of about 2 m or less. Ice advecting into the
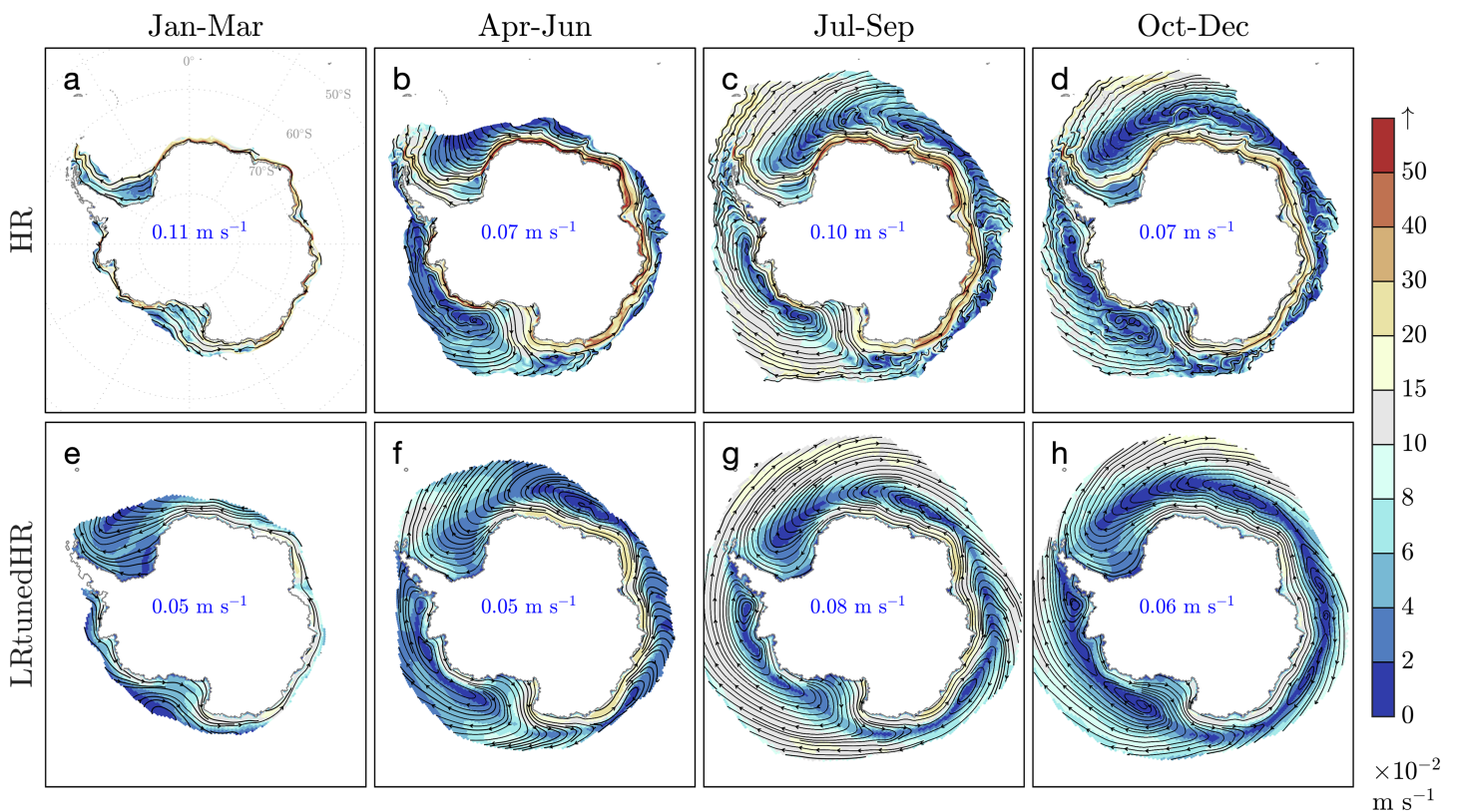
Figure 23. As for Figure 22 but for the Southern Hemisphere.

Weddell Gyre in 1992 along the West Antarctic coast near 15° W ranged in draft between about 0.9–1.7 m in March and 1.1–2.3 m in September (Roberts, 2005), whereas it remains thinner in E3SM in all mid-20th century high-resolution cases. By these measures, a much reduced thickness bias exists in LRtunedHR in both hemispheres, pointing to less bias in the polar surface heat budget, especially during summer, at low resolution.

4. Climate Variability

In addition to the mean state of the model, temporal variability is important. In this section we provide a short discussion of the main modes of climate variability.

4.1. SST/SSH Variability

As expected, the HR v1 ocean exhibits a robust mesoscale eddy field, with maximum SSH variability located primarily in western boundary current regions and the ACC (Figure 24). The globally averaged SSH variability of the model (4.0 cm) compares well with the satellite altimetry value of 4.2 cm, and the correlation between the model and data is 0.73. Both the Gulf Stream and Kuroshio maxima are shifted about 50
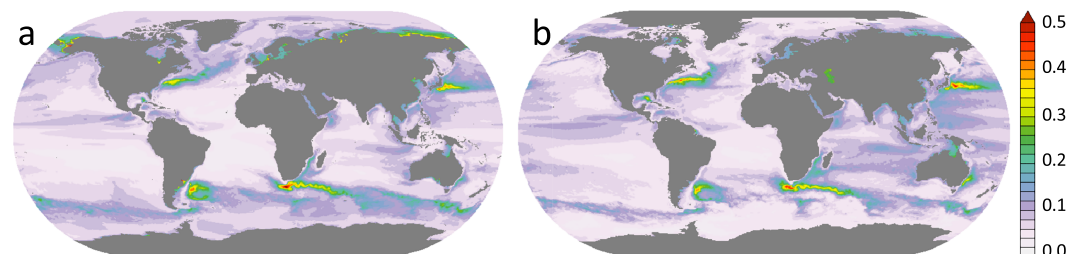

Figure 24. Sea surface height variability (m) for (a) E3SMv1 HR (years 21–50 calculated from 5-day snapshots) and (b) AVISO data (https://www.aviso.altimetry.fr) from satellite altimetry (years 1993–2013).
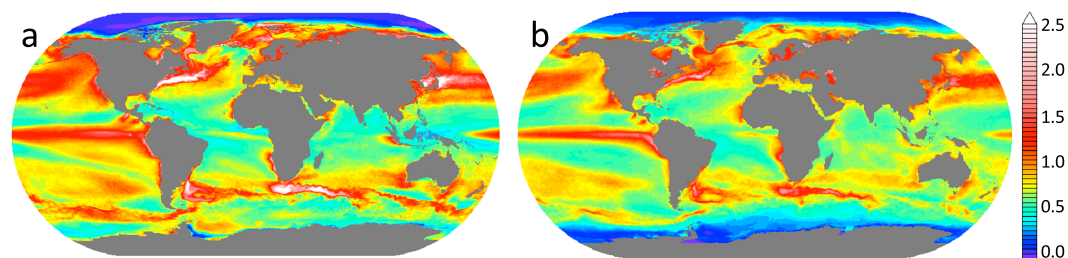
Figure 25. Sea surface temperature variability (°C) with the mean seasonal cycle removed for (a) E3SMv1 HR (years 21–50 calculated from 5-day snapshots) and (b) AVHRR (http://pathfinder.nodc.noaa.gov) satellite data (years 1982–2011 calculated from 5-day snapshots).

km equatorward in the model when compared to AVISO, whereas the variability pattern in East Australia Current (EAC) matches very well. However, the magnitude of the variability in these three Western Boundary Currents is low by about 25% for the Northern Hemisphere currents, and almost 50% for the EAC. As is typical of this class of coupled models (Delworth et al., 2012; McClean et al., 2011; Small, Bacmeister, et al., 2014), the Gulf Stream Extension reaches too far across the North Atlantic and has a weak representation of eddies hooking around the Northwest Corner. The model Kuroshio exhibits two of the observed modes of multiyear variability: the Large Meander to the south of Japan around 138E (Qiu & Miao, 2000) and periods of alternating high/low eddy energy in the Kuroshio Extension (Douglass et al., 2012; Qiu & Chen, 2005).

The low SSH variability in the Gulf of Mexico is due to lack of eddy shedding by the Loop Current for most of the simulation. Soon after the beginning of the run, the stratification in the core of the current drops, resulting in a baroclinically stable flow. It is not until year 38 that an intrusion of cooler water increases the stratification enough for the Loop Current to become unstable, with the first eddy that fully detaches and propagates westward occurring in year 42. The current then continues in a more realistic unstable state for the remainder of the simulation, detaching three more eddies during years 44–50.

Similarly to the Gulf Stream and Kuroshio, the SSH variability in the Agulhas Retroflection is shifted 50–100 km equatorward compared to AVISO. Somewhat remarkably, the meanders in the model match AVISO extremely well in both zonal position and wavelength. Agulhas eddies crossing the South Atlantic do not follow a single dominant path as appears to be the case in, for example, Delworth et al. (2012), but instead are more realistically distributed over a range of latitudes.

Figure 25 shows the analogous SST variability (with the mean seasonal cycle removed) compared with AVHRR satellite observations. Unlike the SSH, which has no direct influence on the atmosphere, mesoscale SST features do have discernible effects on fluxes of heat and moisture to the atmosphere. SST variability has been shown to modify cloudiness, precipitation, and surface winds (Bryan et al., 2010; Byrne et al., 2015), while the location and strength of oceanic fronts can influence the path of storm tracks (Small, Tomas, et al., 2014). The bias in globally averaged variability is about 40%, or 0.1 °C, though locally the model variations can be as much as 2 °C too large. This suggests the possibility of unrealistically high mesoscale feedbacks in the model, though it is beyond the scope of this paper to quantify this. Despite the large difference in normalized globally averaged variance, the correlation between the model and data for SST (0.72) is essentially the same as for SSH.

4.2. El Niño-Southern Oscillation

El Niño Southern Oscillation (ENSO) is a dominant mode of SST variability in the tropical Pacific (e.g., Figure 26a). Given the short simulation period, ENSO power spectra are not a reliable indicator of model performance (e.g., Wittenberg, 2009). Although the HR Nino 3.4 index looks reasonable in event magnitude and spacing, a single index only presents a small data point to the fidelity of ENSO in E3SM at HR. To better examine the influence of resolution on ENSO structure, we therefore examine the dominant patterns of SST variability in the HadISST data (Rayner et al., 2003), v1 HR, and v1 LRtunedHR. As expected, the ENSO SST pattern emerges as the first EOF of each SST data set (shown in Figure 26). Broadly, each model captures the ENSO pattern, with subtle but important differences. At low resolution, the warm SST anomalies are too strong and extend too far west relative to observations, which is a common model bias (e.g., Menary et al., 2018). Surprisingly, increased resolution has little impact on ENSO pattern in E3SMv1. This is in contrast to studies with other models, which found an improvement in westward extent (e.g., Kirtman et al., 2012;
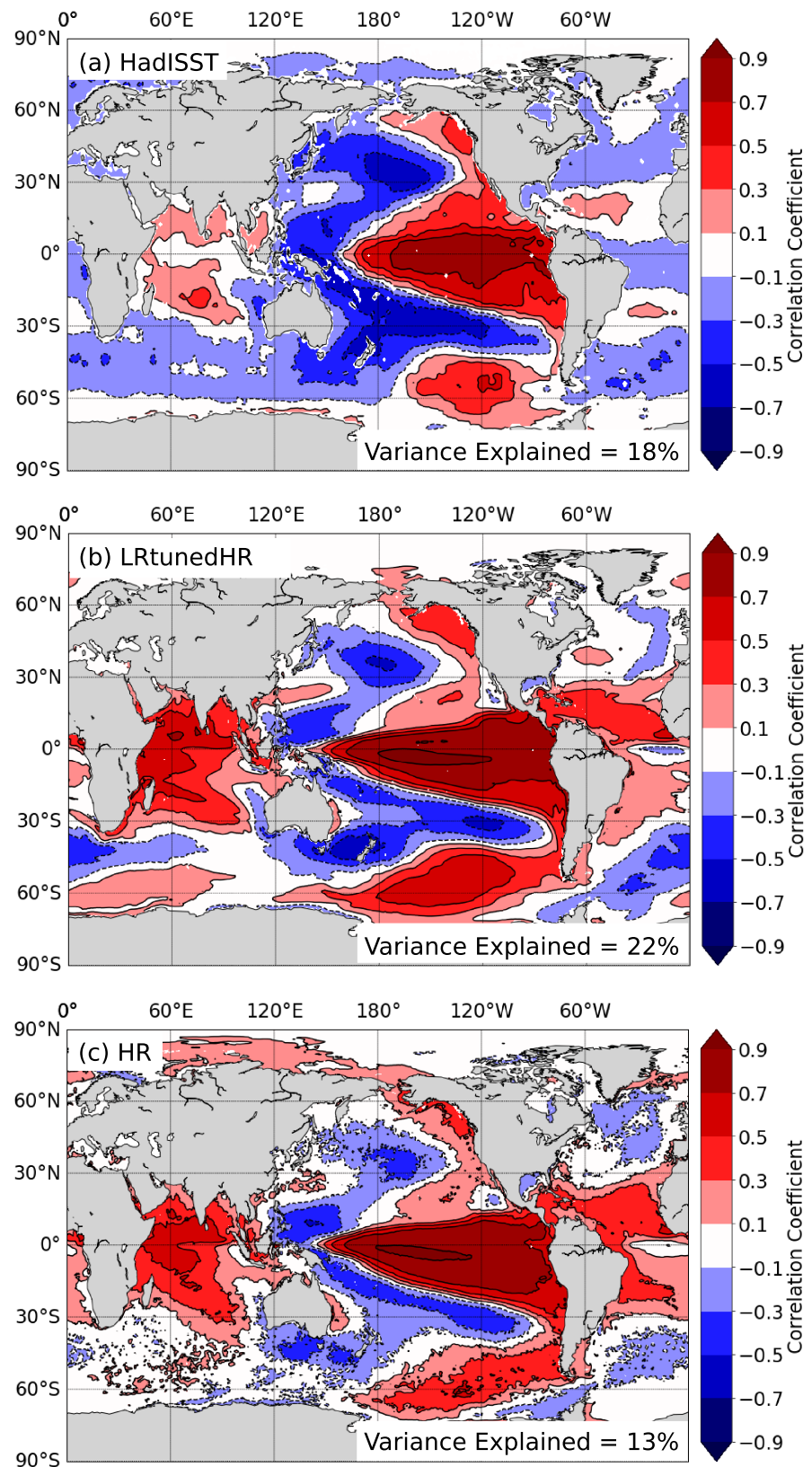
Figure 26. First Empirical Orthogonal Function (EOF) of sea surface temperature for (a) HadISST data (1870–2010), (b) LRtunedHR (years 1–50), and (c) HR (years 1–50). The variance explained by each EOF is shown in the respective panels.
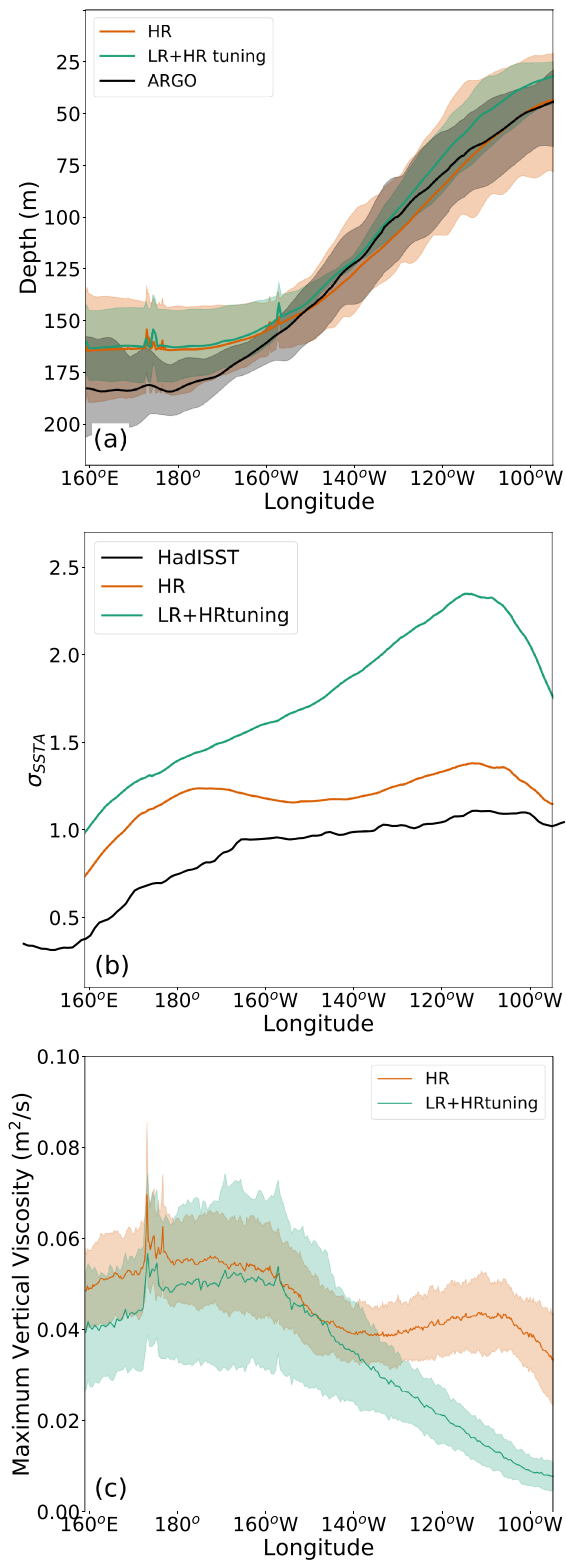
Figure 27. DJF averages of (a) thermocline depth (defined as the depth of the 20 °C isotherm; Z_{20}, where the shading represents the maximum and minimum depths observed during the period. (b) standard deviation of monthly SST anomalies and (c) average maximum vertical viscosity above the thermocline (solid lines) and standard deviation (shading). For all plots, fields have been averaged between 1° S and 1° N. The periods of analysis are years 1–50 for HR and for LRtunedHR. The ARGO data (Gaillard et al., 2016) in (a) run from 2013 to 2017, and the HadISST (c) is 1870 to 2015.
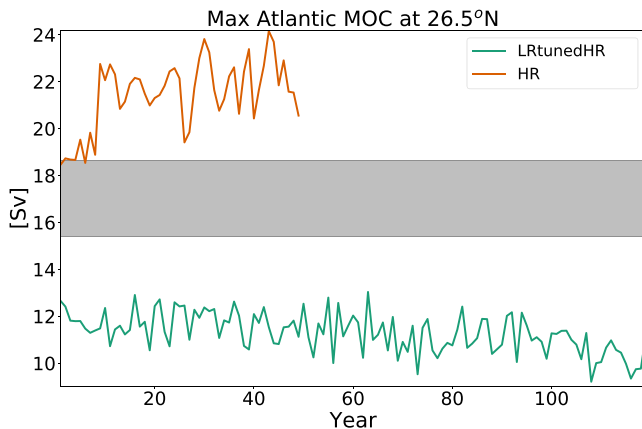
Figure 28. Maximum AMOC strength (Sv) at 26.5°N, The orange line is HR and teal is LRtunedHR. The gray shaded region represents the mean plus and minus the range of variability observed at the RAPID station (variability is based on monthly observations).

Menary et al., 2018; Sakamoto et al., 2012). Across resolutions, the off-equatorial ENSO response is muted, with the weakest response in HR. While the total tropical SST variance in HR is too large (Figure 25), the fraction of SST variance associated with ENSO in HR is too small (as noted in Figure 26) with the net result that the magnitude of ENSO-related SST variations in HR is about right. In other words, tropical SST variance in the HR model run is too large due to sources other than ENSO. This is confirmed in spectra of Nino3, 3.4, and 4 SST anomalies (not shown).

To explore possible mechanisms for the lack of improvement in ENSO in the HR simulation, we examine DJF thermocline depth and SST anomaly variability along the equator (Figures 27a and 27b), where we define the thermocline as the depth of the 20 °C isotherm (e.g., Kessler, 1990; Yang & Wang, 2009). The LRtunedHR simulation most faithfully represents the ARGO data, while the higher-resolution simulation has a mean thermocline depth gradient, which is too weak. Previous studies (e.g., Kirtman et al., 2012) show that SST anomalies are better represented when model resolution is increased. This is also evident in our simulations, although unlike previous work, our LR simulation has *more* variance along the equator than observed and increasing resolution reduces bias by *weakening* variability. We also note that the SST anomaly variance observed is consistent with the simulated thermocline depths in Figure 27a. The HR model has the deepest thermocline depths and a more muted SST anomaly variance than the lower resolution simulation (but higher than observed). Despite the deeper mean thermocline depths, the HR simulation has a larger range of thermocline depth variability than LRtunedHR or observations.

Deeper mean East Pacific thermocline depths and larger variability in the HR simulation may be responsible for the response seen in Figure 26 (e.g., Jin, 1996). Large fluctuations in the thermocline depth imply large variations in oceanic heat content (OHC). Excess variation in OHC could reduce the persistence and magnitude of individual ENSO events, which could explain the weakened SST-precipitation extremes correlation seen in Figure 31. Physically, we would expect the deeper depths in the East Pacific to result from either vertical advection or diffusion. The vertical velocity at the equator is roughly consistent between high and low resolution (not shown), given that the equatorial resolution in the LR configuration captures essential processes (e.g., Tropical Instability Waves). The vertical diffusivity above the thermocline in the tropical Pacific is shown in Figure 27c. In the Eastern Pacific, the maximum vertical diffusivity at and above the thermocline is higher throughout the tropical Pacific and increases toward the eastern boundary. It is possible that the increased diffusivity in the cold tongue results in the weaker east/west thermocline gradient, weakening the ENSO response at high resolution. However, it is still unclear what causes this increased diffusivity or if a reduction in diffusivity would improve the model representation of high resolution. This will be investigated in a future publication.
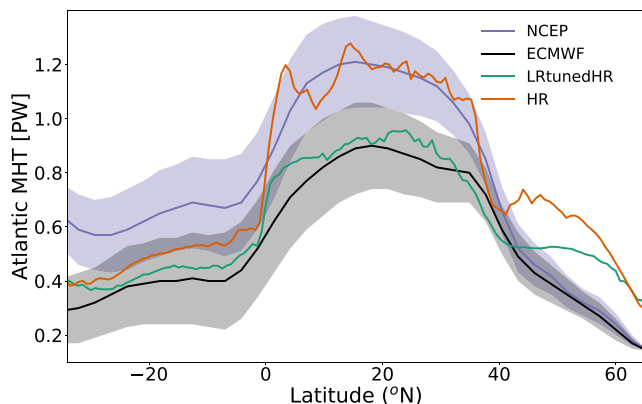
4.3. Atlantic Meridional Overturning Circulation

The Atlantic Meridional Overturning Circulation (AMOC) plays a dominant role in the Earth's climate (e.g., Cheng et al., 2013), where changes in AMOC can influence, among other things, North Atlantic storm tracks, Northern Hemisphere climate, and CO_2 sequestration. In LRv1 (Golaz et al., 2019), weak AMOC strength and variability was one of the most prominent model biases. Here we examine the influence of resolution on the modeled AMOC at the location of the RAPID array (Rayner et al., 2011). A few studies (Hirschi et al., 2013; Stepanov and Haines, 2014) have shown a mixed influence of resolution on the simulation of AMOC. These studies have found that there is a clear increase in AMOC variability across resolution, but the influence on the mean AMOC is less clear. Figure 28 shows the time series of the maximum AMOC at 26.5°N, the location of the RAPID array. When resolution is increased the mean AMOC strength increases, with v1 HR being slightly too high.

Figure 29. Atlantic Meridional Heat Transport computed from NCEP and ECMWF reanalysis (Trenberth and Caron, 2001), HR (orange), and v1 LRtunedHR (teal). An estimate of observational error is shown in the shaded envelope. For HR and LRtunedHR, years 21–50 are used.
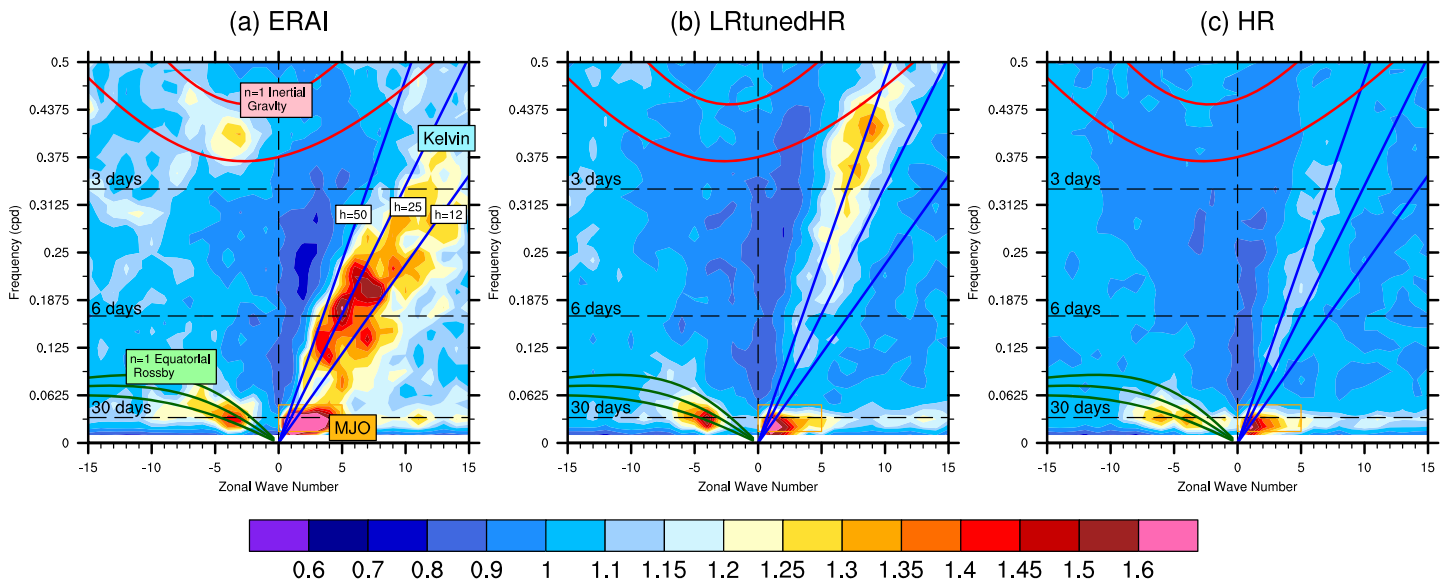
Figure 30. Filtered wave number frequency power spectra for total precipitation (mm^2/day^2), shown as a ratio of the smoothed pseudo-red background spectra of the same field following Wheeler and Kiladis (1999; 15° S to 15° N), for (a) TRMM, (b) LRtunedHR,and (c) HR with data interpolated to 1°.

Previous works (e.g., Cheng et al., 2013; Stepanov and Haines, 2014) suggest the AMOC strength could change in response to many variables, including surface forcing, ocean stratification, and northward heat transport. Given that the mean LRtunedHR AMOC is similar to LRv1, but with a very different atmosphere, it is unlikely that surface forcing is responsible for the change in magnitude. Figure 17 shows that the near surface stratification is changed dramatically with higher resolution, most notably, the Labrador Sea convection is vigorous in HR and absent at LR. However, it seems unlikely that increased convection in the Labrador Sea is responsible for 10 Sv of increased transport from LRtunedHR to v1 HR. We have also examined mass transport through the Florida Straits in each simulation (not shown), and the HR simulation has much stronger transport through this region, suggesting a possible relationship between increased northward transport (likely due to better simulation of the Gulf Stream as suggested by Joyce and Zhang, 2010 in the HR simulation) and AMOC strength. The increased northward heat transport could lead to the improvement in the Labrador Sea ice thickness and extent (consistent with Rugenstein et al. (2013)). This possible relationship is further explored by examining the Atlantic Meridional Heat Transport (MHT) for each simulation (Figure 29). All model results fall within the range of observational variability. Further, consistent with previous results (Volkov et al., 2008), MHT increases in magnitude due to the resolution of mesoscale eddies. It is possible that the increased heat transport encourages more evaporation in the North Atlantic, increasing deep convection, and hence the AMOC (e.g., Cheng et al., 2013; Stepanov and Haines, 2014).

4.4. Madden Julian Oscillation

The Madden Julian Oscillation (MJO) is the dominant signal of intraseasonal variability in the the tropical atmosphere (e.g., Zhang, 2005). The MJO has broad reaching impacts through various teleconnections, including North American temperature and precipitation (e.g., Donald et al., 2006; Lin et al., 2009). Accurate simulation of the MJO could also lead to improved predictability (e.g., Jones et al., 2004). Therefore, it is important for E3SM to exhibit skill in simulating the MJO in order to have improved projections of the North American hydrological cycle. In Figure 30 we see a robust simulation of the MJO at both resolutions. This is similar to the LRv1 simulations shown in Golaz et al. (2019). Increasing model resolution leads to no clear improvement in the modeled MJO, which is consistent with previous work (e.g., Kim et al., 2018).

Figure 30 also shows a clear underestimate of the tropical wave variability in the HR simulations. In particular the higher-requency Kelvin wave activity is near absent compared to observations. The tunings in the HR simulations do not themselves explain the lack of Kelvin Wave activity, since when these tunings are applied to a low-esolution configuration (LRtunedHR; Figure 30b), the wave activity remains very similar to the default 1° configurations seen in Golaz et al. (2019), which do retain some Kelvin wave activity, albeit at spuriously high frequencies.
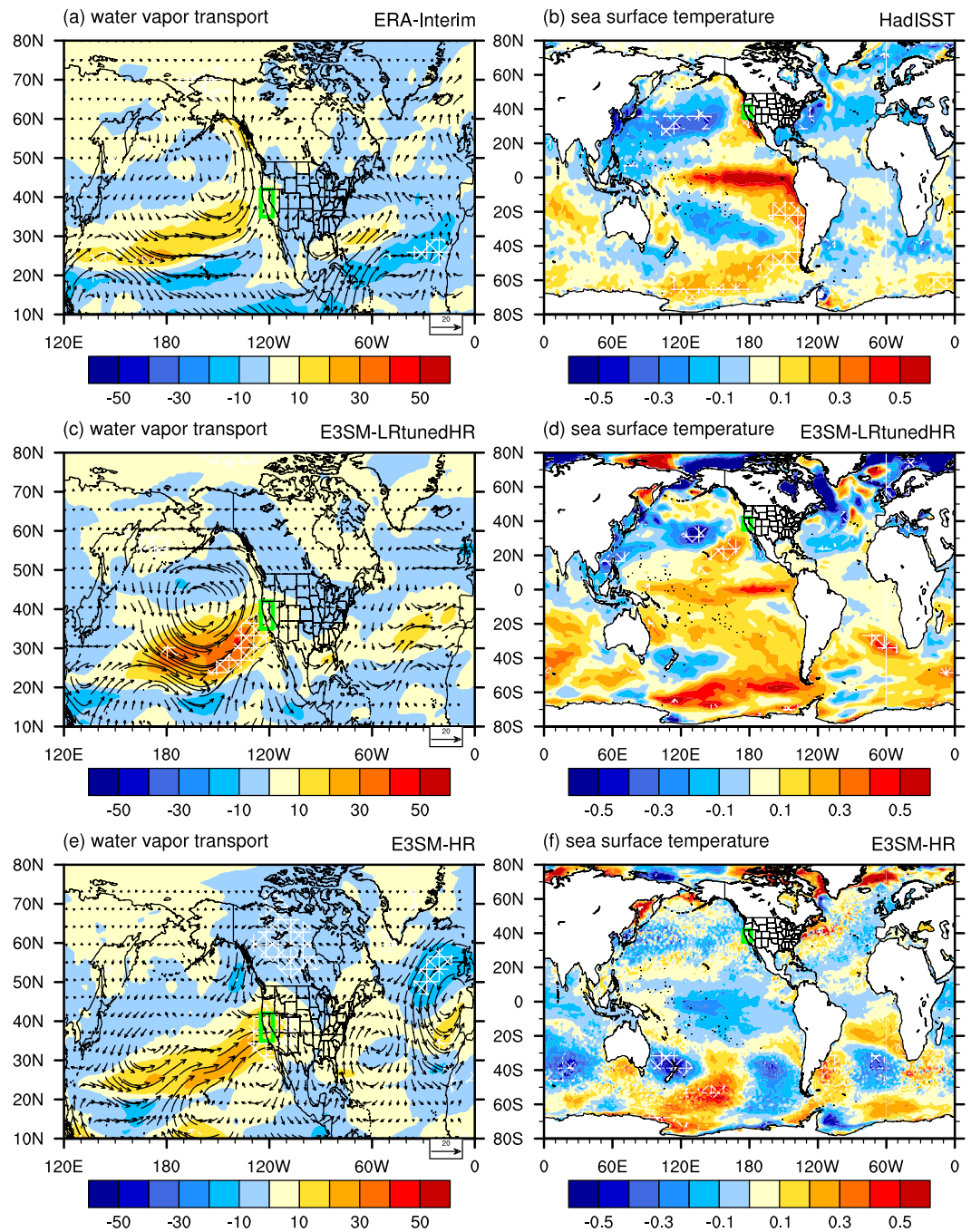
Figure 31. Regression onto the time series of winter extreme precipitation of (left) monthly vertically integrated water vapor transport (kg·m^{-1}·s^{-1}) and (right) SST (° C) averaged over DJF from (a) ERA-Interim, (b) HadISST, (c and d) LRtunedHR, and HR (e and f) runs. See text for details. Hatched areas indicate that the regression coefficients are statistically significant at the 95% level of confidence.

Recent theoretical work on the MJO (Chen & Wang, 2019; Wang et al., 2016) have proposed a dynamic moisture mode framework to fully encompass previous MJO theories. This work suggests that models that exhibit robust MJO power preferentially also have strong feedbacks between cloud radiative forcing, precipitation, and moisture convergence (Chen & Wang, 2019). Chen and Wang (2019) also find that for models to correctly simulate the Kelvin wave component of the MJO, the feedbacks between boundary layer convergence and low level moisture and free tropospheric waves and convective heating must be correctly captured. It is possible that E3SMv1 has biases in one or both of these feedbacks, leading to the anomalously weak Kelvin

wave signal seen in Figure 30. An exploration of these feedbacks is beyond the scope of this overview and is the subject of a future manuscript.

5. Extremes

Changes in the frequency and character of low-robability events have great impact on people and the environments in which they reside. Scaling and theoretical considerations would favor more accurate representations of these events by higher-esolution climate models. In this section, we test this assumption for a variety of extreme weather events.

5.1. Extreme Precipitation Over California

In this section, we evaluate the impact of horizontal resolution on extreme precipitation over California. California was chosen because of its large population, its importance to global food production, and its mountainous terrain, which one might hope is better captured at higher resolution.

Wintertime precipitation over California comes from moisture transported by the large-scale circulation. This leads us to evaluate the water vapor transport and sea surface temperature anomalies, which give rise to wintertime extreme precipitation in this region. To avoid the need for high-frequency output, this is done by first creating an extreme precipitation index by taking the sum of all daily precipitation values above the 95th percentile (computed separately for each model grid cell) within each month and calculating the DJF mean for each year, then computing the average index over California (defined here as 35–42° N, 125–118° W). The time series of DJF average of SST and vertically integrated water vapor transport (IVT) for every year are then regressed on this extreme precipitation index (separately for each model grid cell) and the resulting slope is plotted. This approach is presented in detail in Dong et al. (2018). Figure 31 shows the resulting regressions of SST and the vertically integrated water vapor transport between 1,000 and 500 hPa on the winter extreme precipitation averaged over California. Panels a and b show values expected from observations. Both precipitation and water vapor transport in this figure comes from ECMWF Interim Reanalysis Data (ERA-Interim) at a 1.5° × 1.5° resolution (Dee et al., 2011) and monthly SST is from the Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST, 1° × 1°, Rayner et al., 2003). Observations and reanalysis products from the 1981–2010 period are used. ERA-Interim data show that extreme precipitation is associated with enhanced cyclonic circulation over the North Pacific, which transports moisture-laden air from the tropics up the west coast of the United States. Both HR and LRtunedHR simulations are able to capture this basic pattern, though simulated moisture transport is a bit too strong in both models. In addition, moisture destined for California follows a slightly more zonal trajectory in HR due to the cyclonic trajectory over the North Pacific being shifted toward the coast and northward. This flow pattern is associated with an El Niño-like SST pattern with warm anomalies in the tropical central-to-eastern Pacific. This tropical SST pattern is reasonably reproduced at low resolution but is entirely absent at high resolution.

To further explore why the SST anomalies related to extreme precipitation in California are not reproduced in high-resolution simulations, we examine the atmospheric response to ENSO during boreal winter. Figure 32 shows the regression of precipitation and 500-hPa geopotential height onto the Nino3.4 index. Observations (Figures 32a and 32b) show that El Niño conditions cause a weakened Walker circulation with reduced precipitation in the tropical western Pacific and increased precipitation in the central-to-eastern tropical Pacific. The tropical ENSO signal can be propagated to the North Pacific via the Pacific-North American (PNA) pattern in geopotential height to influence precipitation in California (Figure 32b), consistent with Dong et al. (2018). The low-resolution (LRtunedHR) simulation reproduces the PNA pattern as well as the mean precipitation response to ENSO (Figures 32c and 32d). The HR simulation, on the other hand, has a weaker precipitation response over the tropical eastern Pacific and a weaker PNA teleconnection over the North Pacific (Figures 32e–32h), suggesting that the low resolution better reproduces the atmospheric response to ENSO events. These results are consistent with the weak relationship between winter extreme precipitation in California and SST in the high-resolution simulations (Figure 31) as well as the muted ENSO response seen in Figure 26.

5.2. Tropical Cyclones

Tropical Cyclones (TCs) are some of the most destructive weather systems in the global tropics and subtropics with the potential to impact climate (Emanuel, 2003). It is thus very important to improve our understanding of the response of these features to changes in air/sea temperatures and shifting meteorological patterns. A rigorous analysis of TCs based on observations is often constrained by the length
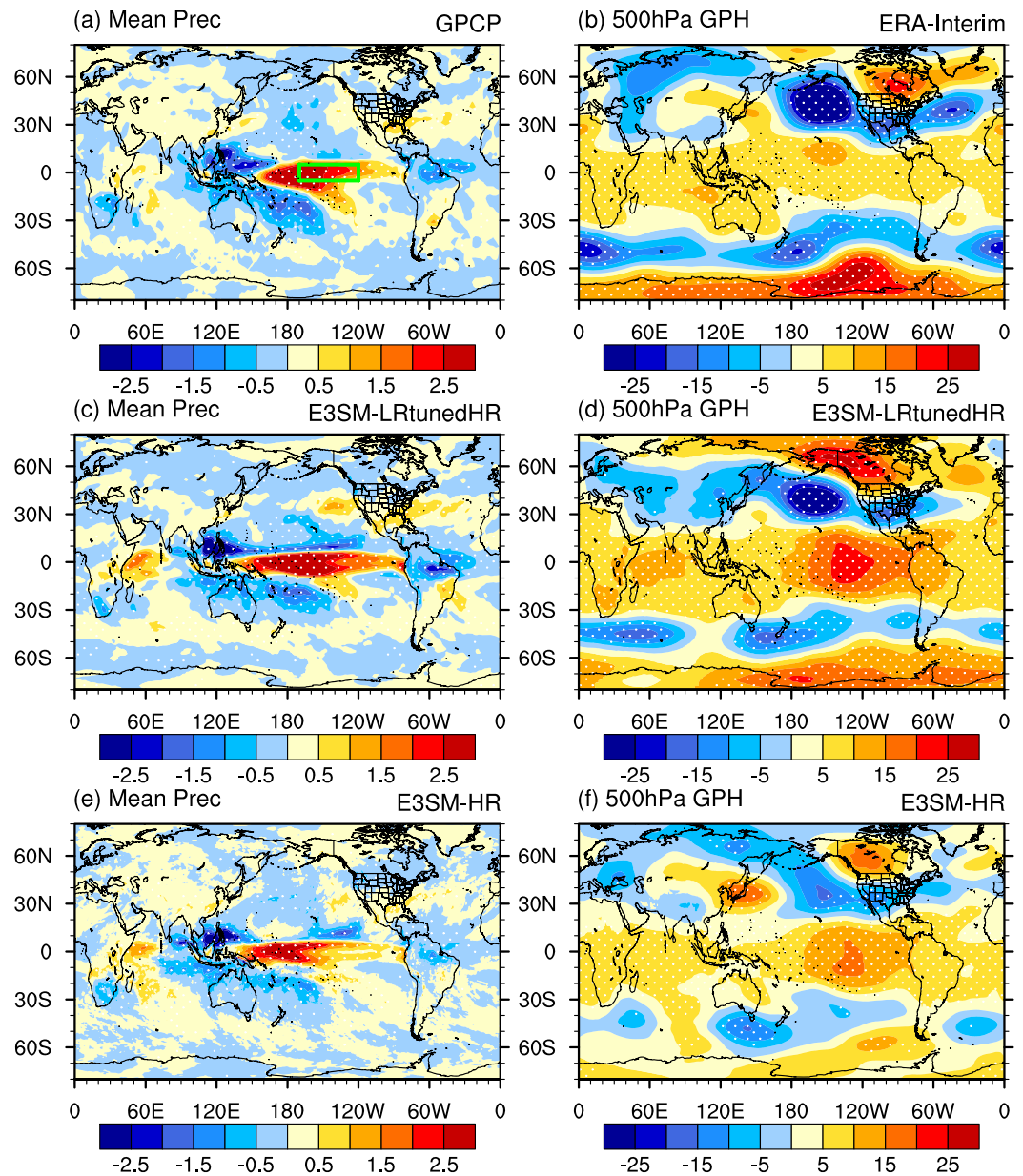
Figure 32. Regression onto time series of the Nino3.4 index during winter (DJF) of (left) monthly mean precipitation (mm/day) and (right) 500-hPa GPH (m; right) from (a) GPCP, (b) ERA-Interim, (c and d) low-resolution E3SM simulation, and (e and f) HR runs. The dotted areas indicate that the regression coefficients are statistically significant at the 95% level of confidence.

of the satellite record and by the lack of direct measurements of certain ocean-atmosphere parameters (Henderson-Sellers et al., 1998), motivating the use of high-resolution numerical models to fill this knowledge gap. Consequently, a key driver of high-resolution climate modeling efforts has been the realistic simulation of TCs and their climatology (Haarsma et al., 2016; Knutson et al., 2013; Walsh et al., 2010). With this consideration, we now examine TC simulation in E3SM.

In Figure 33, 50 years of TC tracks are shown from observations (Figure 33a) and E3SMv1 (Figures 33b and 33c). Observed TC tracks from the period 1966–2015 are obtained from the International Best Track Archive for Climate Stewardship (IBTrACS, Knapp et al., 2010). For the low-resolution TC tracks, we use a 50-year historical run of E3SM. In the model simulations, TCs are detected using TempestExtremes (Ullrich & Zarzycki, 2017; Zarzycki & Ullrich, 2017), a scale-aware feature tracking software. TC-like vortices are detected based on local minima in sea-level pressure, positive temperature anomaly at the 200- and 500-hPa
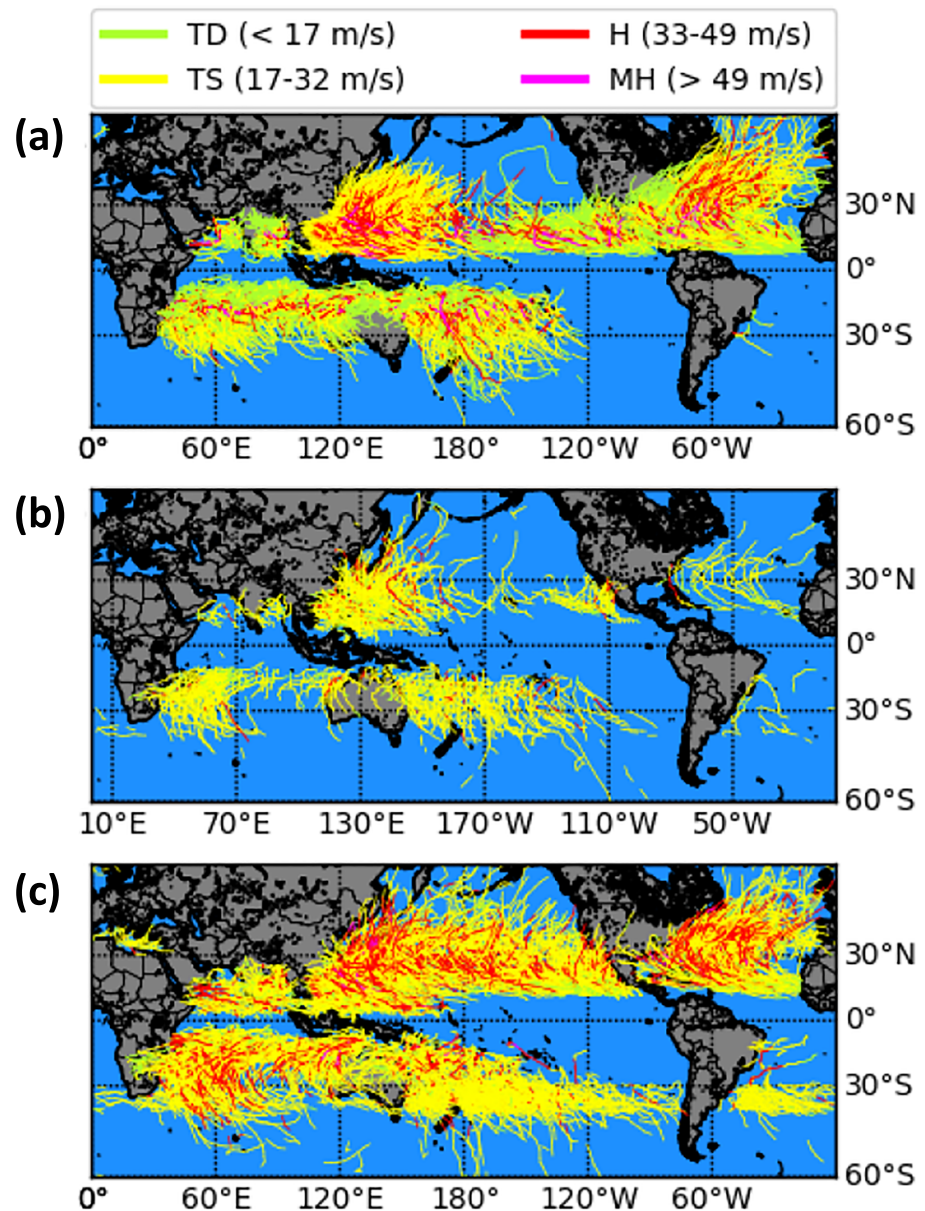
Figure 33. Global distribution of TC tracks from 50 years of (a) observations, (b) low-resolution E3SM simulations, and (c) high-resolution E3SM simulations. Tracks are color-coded based on the along-track intensity. The legend corresponding to the strength of the TCs is also shown.

levels to ensure a warm core, and the wind speed at the lowest model level. For further details regarding the TC-detection algorithm, see Ullrich and Zarzycki (2017). Not surprisingly, in the low-resolution version of the model, simulation of TCs is considerably underestimated (Figure 33b). On average, nearly 15 TC-like features are generated globally in the low-resolution version, which is a substantial underestimation when compared to observations (Figure 33a) where the annual global mean is about 92. However, this is in reasonable agreement with those produced by many other climate models at a similar resolution of about 100 km (Camargo, 2013). The annual mean global TC frequency is considerably improved in the high-resolution simulations of E3SM (Figure 33c). On average, about 64 TC tracks are produced in each model year. When it comes to the distribution of TC tracks in different basins, we find spurious TC activity being produced in the subtropical South Atlantic and Southeast Pacific. The model also produces TC tracks in the Mediterranean region, which are likely related to Medicanes or TC-like storms in that region (Cavicchia et al., 2014).
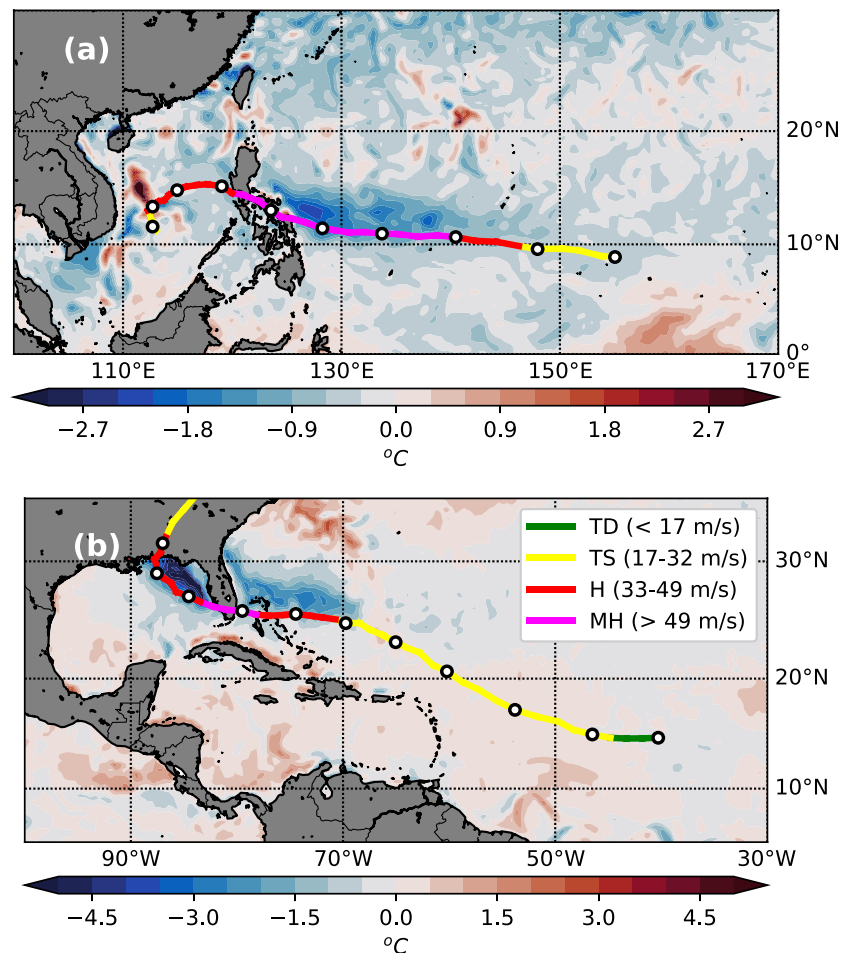
Figure 34. (a) Track of the most intense TC produced in the high-resolution E3SM simulations, color-coded by the along-track intensity. (b) An example of an intense TC produced in the Atlantic basin. In both (a) and (b), the color in the background represents the difference between SST on the day of the TC's formation and SST a week later. The legend shown in (b) corresponds to the along-track intensity of the TCs.

Next, when it comes to intensity, the strongest TC produced in the low-resolution E3SM is a Category 2 on the Saffir-Simpson scale with a maximum wind speed of about 46 m/s. On the other hand, the most intense TC from the high-resolution simulations nearly reaches Category 5 with a maximum lifetime wind speed of nearly 70 m/s (Figure 34a). The track of this TC is reminiscent of Supertyphoon Haiyan, which devastated the Philippines in 2013 (Lin et al., 2014). The cold SST wake produced by this TC is also shown in Figure 34a. Consistent with theory, the strongest cooling occurs on the right-hand side of the track (Price, 1981). The magnitude of the cold wake increases with the strength of the storm and a maximum cooling of about 3 ° C is produced just before landfall. Similarly, Figure 34b shows an example of an intense TC in the Atlantic. The track of this TC is almost identical to that of Hurricane Andrew (1992), one of the most destructive hurricanes in Atlantic history (Pielke & Landsea, 1998). After forming in the Atlantic, the TC goes around the Caribbean islands and makes two landfalls, first as a Category 3 TC over South Florida and subsequently as a high Category 1 TC over the Northern Gulf of Mexico. Again, a cold SST wake can be seen, stronger on the right-hand side of the track, with a maximum cooling of about 5 ° C in the Gulf of Mexico. While we have only presented an overview of TC simulation in this manuscript, we will follow this with another study focused on TCs with a more detailed description and analysis.

5.3. Atmospheric Blocking
Atmospheric blocking is a synoptic-scale weather phenomenon associated with relatively stationary regions of anomalously high pressure. Given that this feature is associated with other forms of extreme weather, including heat waves, cold air outbreaks, and extreme precipitation, it is desirable to capture the frequency
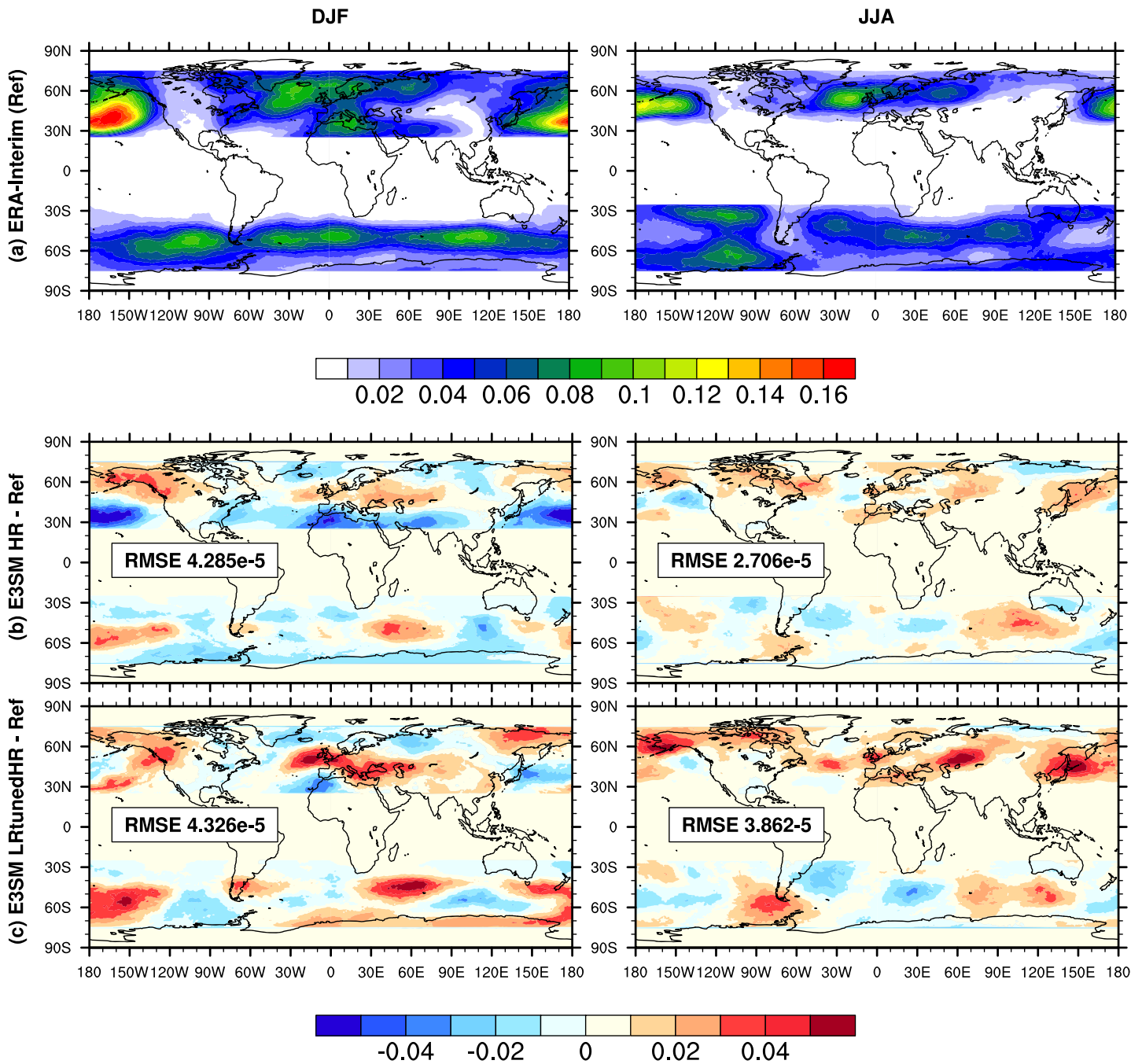
Figure 35. Pointwise frequency of 500-hPa blocking for (left) December-January-February (DJF) and (right) June-July-August (JJA). The rows indicate (a) ERA-Interim reference, and differences between (b) E3SM HR and reference, and (c) E3SM LR and reference.

and character of these features with high fidelity. Global climate models, while able to approximate broad spatial blocking patterns, have traditionally struggled with correctly replicating observed blocking frequencies (Dunn-Sigouin & Son, 2013). Because blocking typically emerges as the result of topographic drivers, improved model resolution (and subsequent improvements in the representation of topography) is expected to give rise to an improvement in model performance. However, Schiemann et al. (2017) studied the sensitivity of blocking quality to resolution and only documented significant improvements in Euro-Atlantic blocking in GCMs in the summer season, with practically no improvement in the Pacific basin.

The blocking detection algorithm used here is a modification of Dole and Gordon (1983) that is detailed in Pinheiro et al. (2019). The algorithm searches pointwise for latitude-adjusted 500-mb geopotential height

(Z500) anomalies that are greater than 1.5 times the standard deviation of the Z500 anomaly at that point, calculated for each day of the year. That is, we define

$$Z^* = Z - \overline{Z} \qquad Z^*_{adj} = Z^* \times \frac{\sin 45\%}{\sin \phi}, \qquad (1)$$

where Z^* is the Z500 anomaly, Z is the current Z500 value, $\overline{Z}$ is the long term daily mean (LTDM) of the Z500 field at each day of the year, Z^*_{adj} is the latitude-adjusted Z500 anomaly, and ϕ is the latitude. An additional minimum anomaly threshold of 100 m is imposed at coordinates where the calculated threshold value is low (due to anomaly time series with relatively low anomaly magnitudes and small standard deviation value).

Blocks from each data set were identified separately for the Northern and Southern Hemispheres, using the StitchBlobs program in TempestExtremes (Pinheiro et al., 2019; Ullrich & Zarzycki, 2017) with the normalized anomalies as the input. The detection algorithm was limited to a latitude range of 25–75° in each hemisphere, with blob persistence of 5 days, minimum size of 10^6 km^2, and a minimum 50% overlap of blobs between time steps. The results of this analysis are depicted in Figure 35. Performance in the Southern Hemisphere was largely unchanged regardless of model resolution, so we focus on analyzing the Northern Hemisphere. Analogous to the results of Schiemann et al. (2017), the NH summertime blocking density exhibits much smaller biases in HR compared to LRtunedHR across all regions. However, there is little improvement in the NH wintertime blocking density in HR compared with LRtunedHR. Further, a pattern of underestimated blocking density at lower latitudes and overestimated density at higher latitudes emerges in the high-resolution run. This pattern suggests that blocks are forming too far north, likely because of a bias toward weaker climatological winds that are more susceptible to curving at higher latitudes, and stronger zonally directed winds at lower latitudes. This pattern is particularly prominent over the Northern Pacific. Nonetheless, there is some improvement in the overestimation of wintertime blocking in the Euro-Atlantic region.

6. Model Sensitivity

In Golaz et al. (2019), LRv1 was shown to reproduce the observed total warming over the 20th century. This warming was, however, accomplished through a balance between extremely high equilibrium climate sensitivity (5.3 K) and very strong aerosol response as encapsulated in LRv1's effective radiative forcing (ERF) of -1.65 W/m^2. These sensitivities were argued to be too large because their combined effect results in underpredicted warming during years when aerosol emissions were particularly strong and overly strong warming in years when aerosol load was reduced. In this study, we will look at the impact of changing resolution and tuning on these sensitivities. Aerosol sensitivity is expected to decrease with resolution (Ma, Rasch, et al., 2015). Climate sensitivity is also likely to change because more scales are explicitly resolved at higher resolution.

Since we cannot afford to run the fully coupled HR model for the hundreds of years necessary to compute transient or equilibrium climate sensitivity, we instead diagnose the net climate feedback parameter (λ) from the difference between a 5-year atmosphere-land simulation with SST and sea ice extent prescribed at current-climate conditions and a simulation that is identical but has SST increased uniformly by 4 K. This approach, popularized by Cess et al. (1989), allows one to compute the effective net feedback from the change in TOA radiative imbalance ΔF_{TOA} caused by forcing global-average surface temperature $T_{glob\,ave}$ to change:

$$\lambda = -\Delta F_{TOA} / \Delta T_{glob\,ave}. \qquad (2)$$

Ringer et al. (2014) show that λ from Cess experiments provides a reasonable, though slightly high, approximation to values obtained by more expensive methods. Terms on the right-hand side of equation (2) are taken from 5-year simulations with annually repeating forcings representative of the decade centered on 2010. This choice differs from typical Cess experiments (which use preindustrial forcings) and was made because preindustrial fixed-SST model forcings were not yet available. Simulation year is unlikely to have a leading-order effect on the sensitivity of λ to resolution. SST and sea ice are actually prescribed from year 2013 because it was more representative of decadal-average behavior than 2010. For more details on configuration, see section 2.4.

Table 6
Net Climate Feedback Parameter (λ, unit: W/m²/K) and Aerosol-Related Effective Radiative Forcing (ERF, unit: W/m²) of Different E3SM Versions

	HR	LRtunedHR	LRv1
λ	−1.29	−1.20	−1.36
ERF	−1.37	−1.37	−1.64
ERF$_{aci}$	−1.49	−1.51	−1.76
ERF$_{ari}$	0.04	0.05	0.03
Surf. albedo ERF	0.09	0.08	0.10

The first row of Table 6 provides λ values for HR, LRtunedHR, and LRv1 simulations. The direct effect of increasing resolution causes a slight strengthening of λ, indicating stronger resistance to temperature change and therefore weaker climate sensitivity. Nevertheless, changing parameterization constants has a larger impact than resolution, as can be seen by the differences between the two low-resolution simulations.

Aerosol sensitivity was also computed from 15-month simulations with horizontal winds nudged toward ERA-Interim using a nudging timescale of 3 hr. Kooperman et al. (2012) demonstrate that nudged simulations of this type allow for accurate aerosol sensitivity predictions from 1-year runs by removing weather noise differences between the simulations. Direct aerosol effects are diagnosed by calling radiation twice in each run: once with modeled aerosols and again with clean conditions. Indirect aerosol effects are computed by repeating our present-day nudged simulation with anthropogenic aerosols and precursor gases replaced with preindustrial values. As explained and defended in Ghan (2013), this approach allows the total aerosol ERF to be decomposed into ERF due to aerosol-cloud interaction (ERF$_{aci}$), ERF due to aerosol-radiation interaction (ERF$_{ari}$), and ERF due to surface albedo change. Note that changes in land use/land cover may also induce resolution dependence in aerosol changes because higher-resolution results in stronger wind maxima that lofts more dust (section 3.3; this effect is not included in the present analysis).

ERF, ERF$_{ari}$, and ERF$_{aci}$ are provided for all three model configurations in Table 6. ERF$_{ari}$ is positive (warms the planet) due to the strong absorption of anthropogenic black carbon (BC) in the cloudy sky, especially when BC resides above clouds and bright snow/ice. The reduction in surface albedo straightforwardly leads to a positive ERF change. ERF$_{aci}$ is strongly negative in all simulations due to cloud brightening with increased cloud condensation nuclei. Surprisingly, increased resolution has no effect on ERF. ERF$_{aci}$ does weaken as expected, but its resolution sensitivity is very small and is counterbalanced by a corresponding weakening of ERF$_{ari}$. Tuning has a much larger effect on ERF than resolution change. ERF$_{aci}$ in particular is weakened by ∼15% due to tuning changes between LRtunedHR and LRv1. This emphasizes the need to understand aerosol sensitivity to parametric choices. More detailed analyses of E3SMv1 aerosol distribution and sensitivity will be presented in a future paper.

7. Conclusions

In this study, we describe the high-resolution (HR) configuration of E3SMv1 and document its simulation characteristics in a 50-year-long climatological control run with time-invariant 1950 forcings following the HighResMIP protocol. The atmosphere and land components have a horizontal resolution of approximately 0.25° longitude and latitude, which is adequate for realistically simulating most weather phenomena. The variable resolution ocean and sea ice grid employs an 18-km horizontal grid spacing near the equator, which smoothly reduces to approximately 6 km near the poles. This is sufficient to capture the most energetic motions in the ocean, which are poorly represented in standard resolution coupled climate models.

The HR version of E3SMv1 is superior to its low-resolution counterpart (and most CMIP5 models) in global RMSE for most important variables. Improvement results from a combination of resolution increases, related tuning changes, and perhaps from using a short initialization period. Improved resolution was particularly beneficial to ocean and ice components. Most strikingly, the excessive sea ice extent in the low-resolution (LR) simulation is eliminated in our HR run, with a resulting reduction of cold and fresh biases in that region. Increased resolution also strengthened (improved) the Atlantic Meridional Overturning Circulation and associated meridional heat transport, due to the improved Gulf Stream. Reduced ice coverage in the Labrador Sea also has important beneficial impacts on precipitation bias south of

Greenland. In this region and elsewhere, improvements come from better-resolved bathymetry, greater extremes in winds and currents, and strengthened ocean eddy kinetic energy. Unlike its low-resolution counterpart, the HR model is also able to capture large polynas typical of those observed in the Weddell Sea. Unfortunately, improvements in ice extent do not translate to improvements in ice thickness, which is generally too thin in our HR simulations. Another major improvement in HR relative to LR is in the spatial distribution of SST and salinity.

As expected, regional cloudiness and precipitation patterns are improved by introducing finer-scale topographic detail. Likewise, hurricanes are better captured, runoff and streamflow are generally improved, extreme precipitation events are better simulated, and the ratio of convective to stratiform precipitation decreases at higher resolution. For aerosol, increased extremes in winds at higher-resolution result in higher dust and sea-salt emissions, while enhanced spatial inhomogeneity in clouds and precipitation at finer resolution increases aerosol wet removal. Shift from parameterized convective transport to resolved-scale vertical advection reduces dry deposition at the surface as more aerosols are lofted into the upper troposphere. The net effect is an increase in global-average natural aerosol and a decrease in anthropogenic aerosol. The regional picture is more complicated. Clouds are relatively unaffected by resolution, though there is a slight tendency for cloudiness to decrease at all levels as resolution increases.

Interestingly, several features found to improve with resolution in other studies did not do so in our simulation. Most notably, stratocumulus and associated warm SST biases in the eastern subtropical oceans were not substantially improved at higher resolution. This lack of improvement occurs because unlike other studies, increased wind stress-induced oceanic upwelling with finer resolution is confined in E3SMv1 to an area that is negligibly small compared to the size of the region of deficient low clouds. Another surprise was degradation in the westward extent of ENSO with increasing resolution. Overly deeper East Pacific thermocline depths in the HR simulation offer one possible explanation for the lack of ENSO improvement, but more study is needed. Oddly, MJO-related Kelvin wave energy also disappears in E3SMv1 as resolution increases. Another surprise was that the diurnal cycle of precipitation was almost completely unaffected by resolution change, even in areas where sea breezes are expected. Additionally, while better topography did improve the frequency of Northern Hemisphere summertime blocking events, wintertime frequency actually got worse.

Another interesting finding is that large-scale patterns of climatological-average cloudiness and precipitation were relatively insensitive to resolution change. Climate and anthropogenic-aerosol sensitivity are also almost completely unaffected by changes in resolution when tuning parameters were held fixed. Lack of aerosol sensitivity change occurred even though ERF_{aci} is expected to weaken with resolution (Ma, Chuang, et al., 2015) because that weakening was slight and was completely compensated by increased ERF_{ari}. The tuning needed to maintain energy balance and model skill with changing resolution, in contrast, had a relatively large effect on climate and aerosol sensitivity. Improved prediction at global scales due to resolution of important fine-scale processes is a major motivator for the push to higher resolution. In this context, the finding that model sensitivity in E3SMv1 is relatively insensitive to resolution is important and perhaps reassuring. There are, however, several caveats to this result worth keeping in mind. First, this finding is from a single model. Other models may respond differently. In particular, models whose stratocumulus increases with resolution are likely to have stronger changes in climate and aerosol sensitivity with resolution because stratocumulus changes have such a powerful impact on TOA radiation. Second, this finding is specific to the transition from 100- to 25-km atmospheric grid spacing. It could be that even higher resolution is needed in order to capture the processes governing model sensitivity.

In summary, the high-resolution coupled version of E3SMv1 is a world-class climate model. Many aspects of model climate benefit from finer resolution. Ocean and sea ice behavior and fine-scale atmospheric features are particularly improved by increased grid spacing. Surprisingly, several features that improved with resolution in previous high-resolution studies did not improve in this study. In addition, anthropogenic-aerosol and climate sensitivity were almost completely unaffected by resolution change. These results suggest that while beneficial, increased resolution does not solve all model problems.

References

Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P.-P., Janowiak, J., et al. et al. (2003). The version-2 global precipitation climatology project (gpcp) monthly precipitation analysis (1979–present). *Journal of Hydrometeorology*, *4*(6), 1147–1167.

Atlas, R., Reale, O., Shen, B.-W., Lin, S.-J., Chern, J.-D., Putman, W., et al. (2005). Hurricane forecasting with the high-resolution nasa finite volume general circulation model. *Geophysical Research Letters*, *32*, L03807. https://doi.org/10.1029/2004GL021513

Bacmeister, J. T., Wehner, M. F., Neale, R. B., Gettelman, A., Hannay, C., Lauritzen, P. H., et al. (2014). Exploratory high-resolution climate simulations using the community atmosphere model (cam). *Journal of Climate*, *27*(9), 3073–3099.

Bader, D., W. Collins, R. Jacob, P. J. P. Rasch, M. Taylor, P. Thornton, and D. Williams (2014), Accelerated Climate Modeling for Energy, U.S. Department of Energy. Retrieved from https://climatemodeling.science.energy.gov/sites/default/files/publicat%ions/acme-project-strategy-plan.pdf

Berckmans, J., Woollings, T., Demory, M.-E., Vidale, P.-L., & Roberts, M. (2013). Atmospheric blocking in a high resolution climate model: influences of mean state, orography and eddy forcing. *Atmospheric Science Letters*, *14*(1), 34–40. https://doi.org/10.1002/asl.412

Bourke, R., & McLaren, A. (1992). Contour Mapping of Arctic Basin Ice Draft and Roughness Parameters. *Journal of Geophysical Research*, *97*(C11), 17,715–17,728.

Boyle, J., & Klein, S. A. (2010). Impact of horizontal resolution on climate model forecasts of tropical precipitation and diabatic heating for the twp-ice period. *Journal of Geophysical Research*, *115*, D23113. https://doi.org/10.1029/2010JD014262

Bryan, F. O., Tomas, R., Dennis, J. M., Chelton, D. B., Loeb, N. G., & McClean, J. L. (2010). Frontal scale air–sea interaction in high-resolution coupled climate models. *Journal of Climate*, *23*(23), 6277–6291. https://doi.org/10.1175/2010JCLI3665.1

Byrne, D., Papritz, L., Frenger, I., Münnich, M., & Gruber, N. (2015). Atmospheric response to mesoscale sea surface temperature anomalies: Assessment of mechanisms and coupling strength in a high-resolution coupled model over the south atlantic. *Journal of the Atmospheric Sciences*, *72*(5), 1872–1890. https://doi.org/10.1175/JAS-D-14-0195.1

Caldwell, P. (2010). California wintertime precipitation bias in regional and global climate models. *Journal of Applied Meteorology and Climatology*, *49*(10), 2147–2158. https://doi.org/10.1175/2010jamc2388.1

Caldwell, P., Chin, H.-N. S., Bader, D. C., & Bala, G. (2009). Evaluation of a WRF dynamical downscaling simulation over california. *Climatic Change*, *95*(3-4), 499–521. https://doi.org/10.1007/s10584-009-9583-5

Camargo, S. J. (2013). Global and regional aspects of tropical cyclone activity in the cmip5 models. *Journal of Climate*, *26*(24), 9880–9902.

Cavicchia, L., von Storch, H., & Gualdi, S. (2014). A long-term climatology of medicanes. *Climate Dynamics*, *43*(5-6), 1183–1195.

Cess, R. D., Potter, G. L., Blanchet, J. P., Boer, G. J., Ghan, S. J., Kiehl, J. T., et al. (1989). Interpretation of cloud-climate feedback as produced by 14 atmospheric general circulation models. *Science*, *245*, 513–516. https://doi.org/10.1126/science.245.4917.513

Chelton, D. B., Deszoeke, R. A., Schlax, M. G., El Naggar, K., & Siwertz, N. (1998). Geographical variability of the first baroclinic rossby radius of deformation. *Journal of Physical Oceanography*, *28*(3), 433–460.

Chen, G., & Wang, B. (2019). Dynamic moisture mode versus moisture mode in mjo dynamics: importance of the wave feedback and boundary layer convergence feedback. *Climate Dynamics*, *52*(9-10), 5127–5143.

Cheng, W., Chiang, J. C., & Zhang, D. (2013). Atlantic meridional overturning circulation (amoc) in cmip5 models: Rcp and historical simulations. *Journal of Climate*, *26*(18), 7187–7197.

Chepfer, H., Bony, S., Winker, D., Chiriaco, M., Dufresne, J.-L., & Sèze, G. (2008). Use of calipso lidar observations to evaluate the cloudiness simulated by a climate model. *Geophysical Research Letters*, *35*, L15704. https://doi.org/10.1029/2008GL034207

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*, 553–597.

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

Demory, M.-E., Vidale, P. L., Roberts, M. J., Berrisford, P., Strachan, J., Schiemann, R., & Mizielinski, M. S. (2014). The role of horizontal resolution in simulating drivers of the global hydrological cycle. *Climate Dynamics*, *42*(7), 2201–2225. https://doi.org/10.1007/s00382-013-1924-4

Dennis, J. M., Edwards, J., Evans, K. J., Guba, O., Lauritzen, P. H., Mirin, A. A., et al. (2012). Cam-se: A scalable spectral element dynamical core for the community atmosphere model. *The International Journal of High Performance Computing Applications*, *26*(1), 74–89. https://doi.org/10.1177/1094342011428142

Do, H. X., Gudmundsson, L., Leonard, M., & Westra, S. (2018). The global streamflow indices and metadata archive (gsim) – part 1: The production of a daily streamflow archive and metadata. *Earth System Science Data*, *10*(2), 765–785. https://doi.org/10.5194/essd-10-765-2018

Dole, R. M., & Gordon, N. D. (1983). Persistent anomalies of the extratropical northern hemisphere wintertime circulation: Geographical distribution and regional persistence characteristics. *Monthly Weather Review*, *111*(8), 1567–1586.

Donald, A., Meinke, H., Power, B., Maia, A. D. H., Wheeler, M. C., White, N., et al. (2006). Near-global impact of the madden-julian oscillation on rainfall. *Geophysical Research Letters*, *33*, L09704.https://doi.org/10.1029/2005GL025155

Dong, L., Leung, L. R., Song, F., & Lu, J. (2018). Roles of SST versus internal atmospheric variability in winter extreme precipitation variability along the u.s. west coast. *Journal of Climate*, *31*(19), 8039–8058. https://doi.org/10.1175/jcli-d-18-0062.1

Douglass, E. M., Jayne, S. R., Bryan, F. O., Peacock, S., & Maltrud, M. (2012). Kuroshio pathways in a climatologically forced model. *Journal of Oceanography*, *68*(5), 625–639. https://doi.org/10.1007/s10872-012-0123-y

Duffy, P. B., Govindasamy, B., Iorio, J. P., Milanovich, J., Sperber, K. R., Taylor, K. E., et al. (2003). High-resolution simulations of global climate, part 1: present climate. *Climate Dynamics*, *21*, 371–390.

Dunn-Sigouin, E., & Son, S.-W. (2013). Northern hemisphere blocking frequency and duration in the cmip5 models. *Journal of Geophysical Research: Atmospheres*, *118*, 1179–1188. https://doi.org/10.1002/jgrd.50143

DuVivier, A. K., & Cassano, J. J. (2013). Evaluation of WRF Model Resolution on Simulated Mesoscale Winds and Surface Fluxes near Greenland. *Monthly Weather Review*, *141*(3), 941–963. https://doi.org/10.1175/MWR-D-12-00091.1

Emanuel, K. (2003). Tropical cyclones. *Annual Review of Earth and Planetary Sciences*, *31*(1), 75–104.

Fekete, B. M., & Vorosmarty, C. J. (2011), Islscp ii unh/grdc composite monthly runoff, https://doi.org/10.3334/ornldaac/994.

Feng, Y., H. Wang, R. C. Easter, P. J. Rasch, K. Zhang, P.-L. Ma, et al. (2019). Global Dust Cycle and Direct Radiative Effect in the E3sm: Impact of Increasing Model Resolution, *Journal of Advances in Modeling Earth Systems*, joint submission to AGU Special Collection on E3SM,.

Gaillard, F., Reynaud, T., Thierry, V., Kolodziejczyk, N., & von Schuckmann, K. (2016). In situ–based reanalysis of the global ocean temperature and salinity with isas: Variability of the heat content and steric height. *Journal of Climate*, *29*(4), 1305–1323.

GarciaQuintana, Y., Courtois, P., Hu, X., Pennelly, C., Kieke, D., & Myers, P. G. (2019). Sensitivity of Labrador Sea Water formation to changes in model resolution, atmospheric forcing and freshwater input. *Journal of Geophysical Research: Oceans*, *124*, 2126–2152. https://doi.org/10.1029/2018JC014459

Gent, P. R., Yeager, S. G., Neale, R. B., Levis, S., & Bailey, D. A. (2010). Improvements in a half degree atmosphere/land version of the ccsm. *Climate Dynamics*, *34*(6), 819–833. https://doi.org/10.1007/s00382-009-0614-8

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

Gettelman, A., Morrison, H., Santos, S., Bogenschutz, P., & Caldwell, P. M. (2015). Advanced two-moment bulk microphysics for global models. Part II: Global model solutions and aerosol-cloud interactions. *Journal of Climate*, *28*(3), 1288–1307. https://doi.org/10.1175/jcli-d-14-00103.1

Ghan, S. J. (2013). Technical note: Estimating aerosol effects on cloud radiative forcing. *Atmospheric Chemistry and Physics*, *13*(19), 9971–9974. https://doi.org/10.5194/acp-13-9971-2013

Gleckler, P., Doutriaux, C., Durack, P., Taylor, K., Zhang, Y., Williams, D., et al. (2016). A more powerful reality test for climate models. *Eos*, *97*. https://doi.org/10.1029/2016eo051663

Gnanadesikan, A., & Hallberg, R. (2006). The Role of Eddies in Determining the Structure and Response of the Wind-Driven Southern Hemisphere Overturning: Results from the Modeling Eddies in the Southern Ocean (MESO) Project. *Journal of Physical Oceanography*, *36*(12), 2232–2252.

Golaz, J.-C., Caldwell, P. M., Van Roekel, L. P., Petersen, M. R., Tang, Q., Wolfe, J. D., et al. (2019). The DOE E3SM coupled model version 1: Overview and evaluation at standard resolution. *Journal of Advances in Modeling Earth Systems*, *11*, 2089–2129. https://doi.org/10.1029/2018MS001603

Golaz, J.-C., Larson, V. E., & Cotton, W. R. (2002). A PDF-based model for boundary layer clouds. Part I: Method and model description. *Journal of the Atmospheric Sciences*, *59*(24), 3540–3551. https://doi.org/10.1175/1520-0469(2002)059<3540:APBMFB>2.0.CO;2

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., et al. (2016). High resolution model intercomparison project (highresmip v1. 0) for cmip6. *Geoscientific Model Development*, *9*(11), 4185–4208. https://doi.org/10.5194/gmd-9-4185-2016

Hagemann, S., Arpe, K., & Roeckner, E. (2006). Evaluation of the hydrological cycle in the echam5 model. *Journal of Climate*, *19*(16), 3810–3827. https://doi.org/10.1175/JCLI3831.1

Harms, S., Fahrbach, E., & Strass, V. H. (2001). Sea ice transports in the Weddell Sea. *Journal of Geophysical Research*, *106*(C5), 9057–9073.

Henderson-Sellers, A., Zhang, H., Berz, G., Emanuel, K., Gray, W., Landsea, C., et al. (1998). Tropical cyclones and global climate change: A post-IPCC assessment. *Bulletin of the American Meteorological Society*, *79*(1), 19–38. https://doi.org/10.1175/1520-0477 (1998)079<0019:TCAGCC>2.0.CO;2

Hertwig, E., Storch, J.-S., Handorf, D., Dethloff, K., Fast, I., & Krismer, T. (2015). Effect of horizontal resolution on echam6-amip performance. *Climate Dynamics*, *45*. https://doi.org/10.1007/s00382-014-2396-x

Hertwig, E., von Storch, J.-S., Handorf, D., Dethloff, K., Fast, I., & Krismer, T. (2014). Effect of horizontal resolution on ECHAM6-AMIP performance. *Climate Dynamics*, *45*(1-2), 185–211. https://doi.org/10.1007/s00382-014-2396-x

Hirschi, J. J.-M., Blaker, A. T., Sinha, B., Coward, A. C., De Cuevas, B. A., Alderson, S. G., & Madec, G. (2013). Chaotic variability of the meridional overturning circulation on subannual to interannual timescales. *Ocean Science*, *9*, 805–823.

Holdsworth, A. M., & Myers, P. G. (2015). The Influence of High Frequency Atmospheric Forcing on the Circulation and Deep Convection of the Labrador Sea. *Journal of Climate*, *28*(12), 4980–4996. https://doi.org/10.1175/JCLI-D-14-00564.1

Holte, J., Talley, L. D., Gilson, J., & Roemmich, D. (2017). An argo mixed layer climatology and database. *Geophysical Research Letters*, *44*(11), 5618–5626.

Hourdin, F., Mauritsen, T., Gettelman, A., Golaz, J.-C., Balaji, V., Duan, Q., et al. (2017). The art and science of climate model tuning. *Bulletin of the American Meteorological Society*, *98*(3), 589–602. https://doi.org/10.1175/bams-d-15-00135.1

Huffman, G., Adler, R., Morrissey, M., Bolvin, D., Curtis, S., Joyce, R., et al. (2001). Global precipitation at one-degree daily resolution from multisatellite observations. *Journal of Hydrometeorology*, *2*, 36–50. https://doi.org/10.1175/1525-7541(2001)002<0036:GPAODD>2.0.CO;2

Huffman, G., & Bolvin, D. (2018). TRMM and other data precipitation dataset documentation, Documentation by NASA Goddard Space and Flight Center and Sciences Systems and Applications Inc.

Hurrell, J. W., Hack, J. J., Shea, D., Caron, J. M., & Rosinski, J. (2008). A new sea surface temperature and sea ice boundary dataset for the community atmosphere model. *Journal of Climate*, *21*(19), 5145–5153.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., et al. (2013). The Community Earth System Model: A Framework for Collaborative Research. *Bulletin of the American Meteorological Society*, *94*(9), 1339–1360. https://doi.org/10.1175/bams-d-12-00121.1

Iacono, M., Delamere, J., Mlawer, E., Shephard, M., Clough, S., & Collins, W. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *Journal of Geophysical Research*, *113*, D13103. https://doi.org/10.1029/2008JD009, 944

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

IPCC (2014). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In R. K. Pachauri, & L. A. Meyer (Eds.), *core Writing Team* (p. 151). Geneva, Switzerland: IPCC. https://www.ipcc.ch/site/assets/uploads/2018/02/SYR_AR5_FINAL_full.pdf

Jahn, A., Kay, J. E., Holland, M. M., & Hall, D. M. (2016). How predictable is the timing of a summer ice-free Arctic. *Geophysical Research Letters*, *43* 9113–9120. https://doi.org/10.1002/2016GL070067

Jin, F.-F. (1996). Tropical ocean-atmosphere interaction, the pacific cold tongue, and the el niño-southern oscillation. *Science*, *274*(5284), 76–78.

Jin, M., Deal, C., Maslowski, W., Matrai, P., Roberts, A., Osinski, R., et al. (2018). Effects of Model Resolution and Ocean Mixing on Forced Ice-Ocean Physical and Biogeochemical Simulations Using Global and Regional System Models. *Journal of Geophysical Research: Oceans*, *123*, 358–377. https://doi.org/10.1002/2017JC013365

Jones, C., Waliser, D. E., Lau, K., & Stern, W. (2004). The madden–julian oscillation and its impact on northern hemisphere weather predictability. *Monthly Weather Review*, *132*(6), 1462–1471.

Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). Cmorph: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, *5*, 487–503.

Joyce, T. M., & Zhang, R. (2010). On the path of the gulf stream and the atlantic meridional overturning circulation. *Journal of Climate*, *23*(11), 3146–3154.

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

Kessler, W. S. (1990). Observations of long rossby waves in the northern tropical pacific. *Journal of Geophysical Research, Oceans*, *95*(C4), 5183–5217.

Kiehl, J. T., & Williamson, D. L. (1991). Dependence of cloud amount on horizontal resolution in the national center for atmospheric research community climate model. *Journal of Geophysical Research*, *96*(D6), 10,955–10,980. https://doi.org/10.1029/91JD00164

Kim, H., Vitart, F., & Waliser, D. E. (2018). Prediction of the madden–julian oscillation: A review. *Journal of Climate*, *31*(23), 9425–9443.

Kirtman, B. P., Bitz, C., Bryan, F., Collins, W., Dennis, J., Hearn, N., et al. (2012). Impact of ocean model resolution on ccsm climate simulations. *Climate Dynamics*, *39*(6), 1303–1328.

Klein, S. A., & Jakob, C. (1999). Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Monthly Weather Review*, *127*, 2514–2531.

Klein, S. A., Zhang, Y., Zelinka, M. D., Pincus, R., Boyle, J., & Gleckler, P. J. (2013). Are climate model simulations of clouds improving? an evaluation using the ISCCP simulator. *Journal of Geophysical Research: Atmospheres*, *118*, 1329–1342. https://doi.org/10.1002/jgrd.50141

Knaff, J. A., Kossin, J. P., & DeMaria, M. (2003). Annular hurricanes. *Weather and Forecasting*, *18*(2), 204–223. https://doi.org/10.1175/1520-0434(2003)018<0204:AH>2.0.CO;2

Knapp, K. R., Kruk, M. C., Levinson, D. H., Diamond, H. J., & Neumann, C. J. (2010). The international best track archive for climate stewardship (ibtracs) unifying tropical cyclone data. *Bulletin of the American Meteorological Society*, *91*(3), 363–376.

Knutson, T. R., Sirutis, J., Vecchi, G. A., Garner, S., Zhao, M., Kim, H.-S., et al. (2013). Dynamical downscaling projections of twenty-first-century atlantic hurricane activity: Cmip3 and cmip5 model-based scenarios. *Journal of Climate*, *26*(17), 6591–6617.

Kooperman, G. J., Pritchard, M. S., Ghan, S. J., Wang, M., Somerville, R. C. J., & Russell, L. M. (2012). Constraining the influence of natural variability to improve estimates of global aerosol indirect effects in a nudged version of the community atmosphere model 5. *Journal of Geophysical Research*, *117*, D23204. https://doi.org/10.1029/2012jd018588

Kurtakoti, P., Veneziani, M., Stössel, A., & Weijer, W. (2018). Preconditioning and formation of maud rise polynyas in a high-resolution earth system model. *Journal of Climate*, *31*(23), 9659–9678. https://doi.org/10.1175/JCLI-D-18-0392.1

Kwok, R., Pedersen, L. T., Gudmandsen, P., & Pang, S. S. (2010). Large sea ice outflow into the Nares Strait in 2007. *Geophysical Research Letters*, *37*, L03502. https://doi.org/10.1029/2009gl041872

Large, W. G., & Yeager, S. (2009). The global climatology of an interannually varying air–sea flux data set. *Climate Dynamics*, *33*(2-3), 341–364.

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

Larson, V. E., & Golaz, J.-C. (2005). Using probability density functions to derive consistent closure relationships among higher-order moments. *Monthly Weather Review*, *133*(4), 1023–1042. https://doi.org/10.1175/MWR2902.1

Lenaerts, J. T., Van Den Broeke, M. R., Scarchilli, C., & Agosta, C. (2012). Impact of model resolution on simulated wind, drifting snow and surface mass balance in Terre Adélie, East Antarctica. *Journal of Glaciology*, *58*(211), 821–829. https://doi.org/10.3189/2012jog12j020

Li, H., Leung, L. R., Getirana, A., Huang, M., Wu, H., Xu, Y., et al. (2015). Evaluating global streamflow simulations by a physically based routing model coupled with the community land model. *Journal of Hydrometeorology*, *16*(2), 948–971. https://doi.org/10.1175/JHM-D-14-0079.1

Li, H., Wigmosta, M., Wu, H., Huang, M., Ke, Y., Coleman, A. M., & Leung, L. R. (2013). A physically based runoff routing model for land surface and earth system models. *Journal of Hydrometeorology*, *14*(3), 808–828.

Lin, H., Brunet, G., & Derome, J. (2009). An observed connection between the north atlantic oscillation and the madden–julian oscillation. *Journal of Climate*, *22*(2), 364–380.

Lin, I.-I., Pun, I.-F., & Lien, C.-C. (2014). Category-6 supertyphoon haiyan in global warming hiatus: Contribution from subsurface ocean warming. *Geophysical Research Letters*, *41*, 8547–8553. https://doi.org/10.1002/2014GL061281

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

Liu, X., Easter, R. C., Ghan, S. J., Zaveri, R., Rasch, P., Shi, X., et al. (2012). Toward a minimal representation of aerosols in climate models: description and evaluation in the Community Atmosphere Model CAM5. *Geoscientific Model Development*, *5*(3), 709–739. https://doi.org/10.5194/gmd-5-709-2012

Liu, X., Ma, P.-L., Wang, H., Tilmes, S., Singh, B., Easter, R. C., et al. (2016). Description and evaluation of a new four-mode version of the modal aerosol module (MAM4) within version 5.3 of the community atmosphere model. *Geoscientific Model Development*, *9*(2), 505–522. https://doi.org/10.5194/gmd-9-505-2016

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., et al. (2009). Towards optimal closure of the earth's top-of-atmosphere radiation budget. *Journal of Climate*, *22*, 748–766.

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

Ma, H.-Y., Chuang, C. C., Klein, S. A., Lo, M.-H., Zhang, Y., Xie, S., et al. (2015). An improved hindcast approach for evaluation and diagnosis of physical processes in global climate models. *Journal of Advances in Modeling Earth Systems*, *7*, 1810–1827. https://doi.org/10.1002/2015ms000490

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*, 5058–5065. https://doi.org/10.1002/2015GL064183

Marti, O., Braconnot, P., Dufresne, J.-L., Bellier, J., Benshila, R., Bony, S., et al. (2010). Key features of the ipsl ocean atmosphere model and its sensitivity to atmospheric resolution. *Climate Dynamics*, *34*(1), 1–26.

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

Mathiot, P., Goosse, H., Fichefet, T., Barnier, B., & Gallée, H. (2011). Modelling the seasonal variability of the Antarctic Slope Current. *Ocean Science*, *7*(4), 455–470. https://doi.org/10.5194/os-7-455-2011

Matsueda, M., Mizuta, R., & Kusunoki, S. (2009). Future change in wintertime atmospheric blocking simulated using a 20-km-mesh atmospheric global circulation model. *Journal of Geophysical Research*, *114*, D12114. https://doi.org/10.1029/2009JD011919

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

McGeehan, T., & Maslowski, W. (2011). Impact of Shelf–Basin Freshwater Transport on Deep Convection in the Western Labrador Sea. *Journal of Physical Oceanography*, *41*(11), 2187–2210. https://doi.org/10.1175/jpo-d-11-01.1

McGeehan, T., & Maslowski, W. (2012). Evaluation and control mechanisms of volume and freshwater export through the Canadian Arctic Archipelago in a high-resolution pan-Arctic ice-ocean model. *Journal of Geophysical Research*, *117*, C00D14. https://doi.org/10.1029/2011JC007261

McPhee, M. G. (1992). Turbulent Heat-Flux in the Upper Ocean under Sea Ice. *Journal of Geophysical Research*, *97*(C4), 5365–5379.

Meier, W. N., Peng, G., Scott, D. J., & Savoie, M. H. (2014). Verification of a new NOAA/NSIDC passive microwave sea-ice concentration climate record. *Polar Research*, *33*, 1–22. https://doi.org/10.3402/polar.v33.21004

Menary, M. B., Kuhlbrodt, T., Ridley, J., Andrews, M. B., Dimdore-Miles, O. B., Deshayes, J., et al. (2018). Preindustrial control simulations with hadgem3-gc3. 1 for cmip6. *Journal of Advances in Modeling Earth Systems*, *10*, 3049–3075.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., et al. (2006). North american regional reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), 343–360.

Meyers, M. P., DeMott, P. J., & Cotton, W. R. (1992). New primary ice-nucleation parameterizations in an explicit cloud model. *Journal of Applied Meteorology*, *31*(7), 708–721. https://doi.org/10.1175/1520-0450(1992)031<0708:npinpi>2.0.co;2

Mlawer, E., Taubman, S., Brown, P., Iacono, M., & Clough, S. (1997). Radiative transfer for inhomogenous atmospheres: RRTM, a validated correlated-k model for the longwave. *Journal of Geophysical Research*, *102*, 16,663–16,682.

Neale, R. B., Richter, J. H., & Jochum, M. (2008). The impact of convection on enso: From a delayed oscillator to a series of events. *Journal of Climate*, *21*(22), 5904–5924. https://doi.org/10.1175/2008JCLI2244.1

Petersen, M. R., Asay-Davis, X. S., Berres, A. S., Chen, Q., Feige, N., Hoffman, M. J., et al. (2019). An evaluation of the ocean and sea ice climate of e3sm using mpas and interannual core-ii forcing. *Journal of Advances in Modeling Earth Systems*, *11*, 1438–1458. https://doi.org/10.1029/2018MS001373

Pielke, R. A. Jr., & Landsea, C. W. (1998). Normalized hurricane damages in the united states: 1925–95. *Weather and Forecasting*, *13*(3), 621–631.

Pinheiro, M. C., Ullrich, P. A., & Grotjahn, R. (2019). Atmospheric blocking and intercomparison of objective detection methods: Flow field characteristics. *Climate Dynamics*, 1–28. https://doi.org/10.1007/s00382-019-04782-5

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

Price, J. F. (1981). Upper ocean response to a hurricane. *Journal of Physical Oceanography*, *11*(2), 153–175.

Qiu, B., & Chen, S. (2005). Variability of the kuroshio extension jet, recirculation gyre, and mesoscale eddies on decadal time scales. *Journal of Physical Oceanography*, *35*(11), 2090–2103. https://doi.org/10.1175/JPO2807.1

Qiu, B., & Miao, W. (2000). Kuroshio path variations south of japan: Bimodality as a self-sustained internal oscillation. *Journal of Physical Oceanography*, *30*(8), 2124–2137. https://doi.org/10.1175/1520-0485(2000)030<2124:KPVSOJ>2.0.CO;2

Rasch, P. J., Xie, S., Ma, P.-L., Lin, W., Wang, H., Tang, Q., et al. (2019). An Overview of the Atmospheric Component of the Energy Exascale Earth System Model. *Journal of Advances in Modeling Earth Systems*, *11*, 2377–2411. https://doi.org/10.1029/2019MS001629 joint submission to AGU Special Collection on E3SM

Rayner, D., Hirschi, J. J.-M., Kanzow, T., Johns, W. E., Wright, P. G., Frajka-Williams, E., et al. (2011). Monitoring the atlantic meridional overturning circulation. *Deep Sea Research Part II: Topical Studies in Oceanography*, *58*(17-18), 1744–1753.

Rayner, N., Parker, D. E., Horton, E., Folland, C. K., Alexander, L. V., Rowell, D., et al. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research*, *108*(D14) 4407. https://doi.org/10.1029/2002JD002670

Richter, J., & Rasch, P. (2008). Effects of convective momentum transport on the atmospheric circulation in the community atmosphere model, version 3. *Journal of Climate*, *21*. https://doi.org/10.1175/2007JCLI1789.1

Richter, J. H., Solomon, A., & Bacmeister, J. T. (2014). 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*, 357–383. https://doi.org/10.1002/2013ms000303

Ridley, D. A., Heald, C. L., Pierce, J. R., & Evans, M. J. (2013). Toward resolution-independent dust emissions in global models: Impacts on the seasonal and spatial distribution of dust. *Geophysical Research Letters*, *40*, 2873–2877. https://doi.org/10.1002/grl.50409

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*, 4035–4042. https://doi.org/10.1002/2014GL060347

Ringler, T., Petersen, M., Higdon, R. L., Jacobsen, D., Jones, P. W., & Maltrud, M. (2013). A multi-resolution approach to global ocean modeling. *Ocean Modelling*, *69*, 211–232.

Roberts, A. (2005). Medium-range numerical prediction of Antarctic sea ice, Ph.D. thesis, Antarctic Climate and Ecosystems Cooperative Research Centre and University of Tasmania.

Roberts, A., Bader, D., Caldwell, P., Feng, Y., Golaz, J., Jacob, R., et al. (2019a). Animated E3SM V1 High Resolution Labrador Sea Ice Thickness and Concentration with mid-20th Century Atmospheric Constituents. *Zenodo*. https://doi.org/10.5281/zenodo.2743414

Roberts, A., Bader, D., Caldwell, P., Feng, Y., Golaz, J., Jacob, R., et al. (2019b). Animated E3SM V1 High Resolution Full Coupled Sea Ice Thickness and Extent with mid-20th Century Atmospheric Constituents. *Zenodo*. https://doi.org/10.5281/zenodo.2741591

Roberts, A., Bader, D., Caldwell, P., Feng, Y., Golaz, J., Jacob, R., et al. (2019c). Animated E3SM V1 High Resolution Mertz Polynya Sea Ice Thickness and Concentration with mid-20th Century Atmospheric Constituents. *Zenodo*. https://doi.org/10.5281/zenodo.2747456

Roberts, A., Craig, A., Maslowski, W., Osinski, R., Duvivier, A., Hughes, M., et al. (2015). Simulating transient ice-ocean Ekman transport in the Regional Arctic System Model and Community Earth System Model. *Annals of Glaciology*, *56*(69), 211–228. https://doi.org/10.3189/2015AoG69A760

Roberts, M. J., Clayton, A., Demory, M.-E., Donners, J., Vidale, P. L., Norton, W., et al. (2009). Impact of resolution on the tropical pacific circulation in a matrix of coupled models. *Journal of Climate*, *22*(10), 2541–2556.

Rodell, M., Beaudoing, H. K., L'Ecuyer, T. S., Olson, W. S., Famiglietti, J. S., Houser, P. R., et al. (2015). The observed state of the water cycle in the early twenty-first century. *Journal of Climate*, *28*(21), 8289–8318. https://doi.org/10.1175/jcli-d-14-00555.1

Rossow, W. B., A. Walker, V. Golea, K. R. Knapp, A. Young, I. A. B. Hankins, and N. C. D. R. Program (2016), International satellite cloud climatology project (isccp) climate data record, h-series, https://doi.org/10.7289/v5qz281s.

Rugenstein, M. A., Winton, M., Stouffer, R. J., Griffies, S. M., & Hallberg, R. (2013). Northern high-latitude heat budget decomposition and transient warming. *Journal of Climate*, *26*(2), 609–621.

Sakamoto, T. T., Komuro, Y., Nishimura, T., Ishii, M., Tatebe, H., Shiogama, H., et al. (2012). Miroc4h: A new high-resolution atmosphere-ocean coupled general circulation model. *Journal of the Meteorological Society of Japan. Ser. II*, *90*(3), 325–359. https://doi.org/10.2151/jmsj.2012-301

Satoh, M., Tomita, H., Yashiro, H., Miura, H., Kodama, C., Seiki, T., et al. (2014). The non-hydrostatic icosahedral atmospheric model: description and development. *Progress in Earth and Planetary Science*, *1*(1), 18. https://doi.org/10.1186/s40645-014-0018-1

Schiemann, R., Demory, M.-E., Shaffrey, L. C., Strachan, J., Vidale, P. L., Mizielinski, M. S., et al. (2017). The resolution sensitivity of northern hemisphere blocking in four 25-km atmospheric global circulation models. *Journal of Climate*, *30*(1), 337–358. https://doi.org/10.1175/JCLI-D-16-0100.1

Schmidt, G. A., Bader, D., Donner, L. J., Elsaesser, G. S., Golaz, J.-C., Hannay, C., et al. (2017). Practice and philosophy of climate model tuning across six US modeling centers. *Geoscientific Model Development*, *10*(9), 3207–3223. https://doi.org/10.5194/gmd-10-3207-2017

Shaevitz, D. A., Camargo, S. J., Sobel, A. H., Jonas, J. A., Kim, D., Kumar, A., et al. (2014). Characteristics of tropical cyclones in high-resolution models in the present climate. *Journal of Advances in Modeling Earth Systems*, *6*, 1154–1172. https://doi.org/10.1002/2014MS000372

Small, R. J., Bacmeister, J., Bailey, D., Baker, A., Bishop, S., Bryan, F., et al. (2014). A new synoptic scale resolving global climate simulation using the community earth system model. *Journal of Advances in Modeling Earth Systems*, *6*, 1065–1094. https://doi.org/10.1002/2014MS000363

Small, R. J., Tomas, R. A., & Bryan, F. O. (2014). Storm track response to ocean fronts in a global high-resolution climate model. *Climate Dynamics*, *43*(3), 805–828. https://doi.org/10.1007/s00382-013-1980-9

Snape, T. J., & Forster, P. M. (2013). Decline of Arctic Sea Ice: Evaluation and weighting of CMIP5 projections. *Journal of Geophysical Research: Atmospheres*, *119*, 546–554. https://doi.org/10.1002/2013JD020593

Steele, M., Morley, R., & Ermold, W. (2001). Phc: A global ocean hydrography with a high-quality arctic ocean. *Journal of Climate*, *14*(9), 2079–2087.

Stepanov, V., & Haines, K. (2014). Mechanisms for amoc variability simulated by the nemo model. *Ocean Science*, *10*(4), 645–656.

Stephens, G. L., L'Ecuyer, T., Forbes, R., Gettelmen, A., Golaz, J.-C., Bodas-Salcedo, A., et al. (2010). Dreary state of precipitation in global models. *Journal of Geophysical Research*, *115*, D24211. https://doi.org/10.1029/2010JD014532

Stephens, G. L., Li, J., Wild, M., Clayson, C. A., Loeb, N., Kato, S., et al. (2012). An update on earth's energy balance in light of the latest global observations. *Nature Geoscience*, *5*(10), 691–696. https://doi.org/10.1038/ngeo1580

Stewart, K., Hogg, A. M., Griffies, S., Heerdegen, A., Ward, M., Spence, P., & England, M. (2017). Vertical resolution of baroclinic modes in global ocean models. *Ocean Modelling*, *113*, 50–65.

Strass, V., & Fahrbach, E. (1998). Temporal and regional variation in sea ice draft and coverage in the Weddell Sea obtained from upward looking sonars. In M. Jeffries (Ed.), *Antarctic sea ice: Physical processes, interactions and variability* (Vol. 74, pp. 123–139). Washington DC: American Geophysical Union.

Tang, Q., Klein, S. A., Xie, S., Lin, W., Golaz, J.-C., Roesler, E. L., et al. (2019). Regionally refined capability in e3sm atmosphere model version 1 (eamv1) and applications for high-resolution modelling. *Geoscientific Model Development Discussion*, *2019*, 1–36. https://doi.org/10.5194/gmd-2019-11

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

Tesfa, T. K., Ruby Leung, L., Huang, M., Li, H.-Y., Voisin, N., & Wigmosta, M. S. (2014). Scalability of grid-and subbasin-based land surface modeling approaches for hydrologic simulations. *Journal of Geophysical Research: Atmospheres*, *119*, 3166–3184. https://doi.org/10.1002/2013JD020493

Textor, C., Schulz, M., Guibert, S., Kinne, S., Balkanski, Y., Bauer, S., et al. (2006). Analysis and quantification of the diversities of aerosol life cycles within AeroCom. *Atmospheric Chemistry and Physics*, *6*, 1777–1813. https://doi.org/10.5194/acp-6-1777-2006

Trenberth, K. E., & Caron, J. M. (2001). Estimates of meridional atmosphere and ocean heat transports. *Journal of Climate*, *14*(16), 3433–3443.

Trenberth, K. E., Dai, A., Rasmussen, R., & Parsons, D. (2003). The changing character of precipitation. *Bulletin of the American Meteorological Society*, *84*, 1205–1218. https://doi.org/10.1175/BAMS-84-9-1205

Ullrich, P. A., & Zarzycki, C. M. (2017). Tempestextremes v1.0: A framework for scale-insensitive pointwise feature tracking on unstructured grids. *Geoscientific Model Development*, *10*, 1069–1090.

Volkov, D. L., Lee, T., & Fu, L.-L. (2008). Eddy-induced meridional heat transport in the ocean. *Geophysical Research Letters*, *35*, L20601. https://doi.org/10.1029/2008GL035490

Walsh, J. E., Chapman, W. L., Fetterer, F., & Stewart, J. S. (2019). Gridded monthly sea ice extent and concentration, 1850 onward, version 2. *Electron. Media*. https://doi.org/10.7265/jj4s-tq79

Walsh, K., Lavender, S., Murakami, H., Scoccimarro, E., Caron, L.-P., & Ghantous, M. (2010). The tropical cyclone climate model intercomparison project. In *Hurricanes and climate change* (pp. 1–24). NY city: Springer.

Wang, B., Liu, F., & Chen, G. (2016). A trio-interaction theory for madden–julian oscillation. *Geoscience Letters*, *3*(1), 34.

Wang, H., Burleyson, C. D., Ma, P.-L., Fast, J. D., & Rasch, P. J. (2018). Using the atmospheric radiation measurement (arm) datasets to evaluate climate models in simulating diurnal and seasonal variations of tropical clouds. *Journal of Climate*, *31*(8), 3301–3325.

Wang, Y., Liu, X., Hoose, C., & Wang, B. (2014). Different contact angle distributions for heterogeneous ice nucleation in the community atmospheric model version 5. *Atmospheric Chemistry and Physics*, *14*(19), 10,411–10,430. https://doi.org/10.5194/acp-14-10411-2014

Wehner, M. F., Reed, K. A., Li, F., Prabhat, J., Bacmeister, C.-T., Chen, C., 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*, 980–997. https://doi.org/10.1002/2013MS000276

Wehner, M. F., Smith, R. L., Bala, G., & Duffy, P. (2010). The effect of horizontal resolution on simulation of very extreme us precipitation events in a global atmosphere model. *Climate Dynamics*, *34*(2), 241–247. https://doi.org/10.1007/s00382-009-0656-y

Wentz, F., S. Yueh, and G. Lagerloef (2014), Aquarius level 3 sea surface salinity standard mapped image annual data v3.0. ver. 3.0. po.daac, ca, usa, dataset accessed [2015-11-05] at https://doi.org/10.5067/AQUAR-3SAPS.

Wild, M., Folini, D., Schär, C., Loeb, N., Dutton, E. G., & König-Langlo, G. (2012). The global energy balance from a surface perspective. *Climate Dynamics*, *40*(11-12), 3107–3134. https://doi.org/10.1007/s00382-012-1569-8

Wittenberg, A. T. (2009). Are historical records sufficient to constrain enso simulations. *Geophysical Research Letters*, *36*, L12702. https://doi.org/10.1029/2009GL038710

Wood, R., & Field, P. R. (2011). The distribution of cloud horizontal sizes. *Journal of Climate*, *24*(18), 4800–4816. https://doi.org/10.1175/2011JCLI4056.1

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*, 2618–2644. https://doi.org/10.1029/2018ms001350

Yang, H., & Wang, F. (2009). Revisiting the thermocline depth in the equatorial pacific. *Journal of Climate*, *22*(13), 3856–3863.

Zarzycki, C. M., & Ullrich, P. A. (2017). Assessing sensitivities in algorithmic detection of tropical cyclones in climate data. *Geophysical Research Letters*, *44*, 1141–1149. https://doi.org/10.1002/2016GL071606

Zhai, X., Greatbatch, R. J., & Eden, C. (2007). Spreading of near-inertial energy in a 1/12 degree model of the North Atlantic Ocean. *Geophysical Research Letters*, *34*, L10609. https://doi.org/10.1029/2007GL029895

Zhang, C. (2005). Madden-julian oscillation. *Reviews of Geophysics*, *43*, RG2003, https://doi.org/10.1029/2004RG000158.

Zhang, G., & McFarlane, N. A. (1995). Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian climate centre general circulation model. *Atmosphere-Ocean*, *33*(3), 407–446. https://doi.org/10.1080/07055900.1995.9649539

Zhang, K., Zhao, C., Wan, H., Qian, Y., Easter, R. C., Ghan, S. J., et al. (2016). Quantifying the impact of sub-grid surface wind variability on sea salt and dust emissions in cam5. *Geoscientific Model Development*, *9*(2), 607–632. https://doi.org/10.5194/gmd-9-607-2016

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*, 1253–1268. https://doi.org/10.1029/2018MS001562