# Combining Regional Mesh Refinement With Vertically Enhanced Physics to Target Marine Stratocumulus Biases as Demonstrated in the Energy Exascale Earth System Model Version 1

Peter A. Bogenschutz[1], Hsiang-He Lee[1], Qi Tang[1], and Takanobu Yamaguchi[2,3]

[1]Lawrence Livermore National Laboratory, Livermore, CA, USA
[2]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA
[3]NOAA Earth System Research Laboratories, Chemical Science Laboratory, Boulder, CO, USA

**Correspondence:** Peter Bogenschutz (bogenschutz1@llnl.gov)

**Abstract.** In this paper we develop a novel framework aimed to significantly reduce biases related to marine stratocumulus clouds in general circulation models (GCMs) while circumventing excessive computational cost requirements. Our strategy is to increase the horizontal resolution using a regionally refined mesh (RRM) over our region of interest in addition to using the Framework for Improvement of Vertical Enhancement (FIVE) to increase the vertical resolution only for specific physical processes that are important for stratocumulus. We apply the RRM off the coast of Peru in the Southeast Pacific, a region that climatologically contains the most marine stratocumulus in the subtropics. We find that our new modeling framework is able to replicate the results of our high resolution benchmark simulation with much fidelity, while reducing the computational cost by several orders of magnitude. In addition, this framework is able to greatly reduce the longstanding biases associated with marine stratocumulus in GCMs when compared to the standard resolution control simulation.

## 1 Introduction

Computational advances continue to allow for general circulation models (GCMs) to be run with finer horizontal and vertical resolution. This is beneficial because it is found that running at deep convective permitting resolutions, for example, seems to solve many problems associated with conventional GCMs that are run with horizontal grid sizes on the order of 100 km (Caldwell et al., 2021). In addition, it is found that running GCMs with higher vertical resolution, characteristic of those found in large eddy simulations, can substantially help to ameliorate certain long standing cloud biases that are endemic to conventional GCMs (Bogenschutz et al., 2021; Lee et al., 2021). While this progress is encouraging and exciting, the goal of attaining GCMs that can be run globally for long duration simulations with kilometer scale resolution in the horizontal and adequately high vertical resolution at a reasonable computational cost is something that is seemingly still in the distant future.

Despite the fact that global high resolution models are currently impractical for long term climate projections, tools exist that allow for a fraction of the globe, or a region of interest, to be run at higher horizontal resolution. This is known as a regionally refined model (RRM; Ringler et al., 2008; Tang et al. 2019). The RRM simulation cost is usually dominated by the high resolution region, so the model cost is relative to the size of the region with the finer mesh. For example, a high resolution mesh that covers 10% of the globe would roughly be equal to 10% of the cost of running the entire globe at the higher resolution.

The RRM approach has been used in the past decade for several applications. For example, Zarzycki and Jablonowski (2015) used a refined mesh over the Atlantic Ocean in the Community Atmosphere Model (CAM) to improve the skill of simulating tropical cyclones relative to a CAM simulation that uses the default horizontal resolution. Tang et al. (2019) used the Department of Energy's (DOE's) Energy Exascale Earth System Model (E3SM) with a refined mesh over the continental United States (CONUS) and showed that this configuration can reproduce the benefits, including improved precipitation skill scores, over CONUS that a high resolution model simulates but at a fraction of the computational cost.

The RRM approach is an efficient framework and test bed for high horizontal resolution model development and analysis. However, the RRM approach only refines the resolution in the horizontal direction and not the vertical direction. While this is reasonable for many applications, as most atmospheric processes seem to benefit more from changes in the horizontal rather vertical resolution, this is not the case for all cloud regimes on the globe. An obvious example is the simulation of subtropical marine stratocumulus clouds.

Subtropical marine stratocumulus typically form over oceans with relatively cold sea surface temperatures (SSTs) and are characterized by sharp gradients of temperature and moisture at the top of the boundary layer (Klein and Hartmann, 1993). Most GCMs tend to poorly simulate these clouds, manifested by a severe underrepresentation (e.g. Zhang et al., 2019; Medeiros et al., 2012). This is particularly problematic due to the fact that subtropical stratocumulus have large radiative effects and a strong relationship to climate sensitivity (Bony and Dufresne, 2005).

Despite many improvements found in higher resolution simulations (i.e. 0.25° horizontal resolution) using E3SM, versus the low resolution configuration (i.e. 1.0° horizontal resolution), Caldwell et al. (2019) found only marginal improvements in regards to the simulation of marine stratocumulus clouds. However, Bogenschutz et al. (2021) found that a key ingredient towards improving the marine stratocmulus bias is increased vertical resolution in the lower troposphere. The study of Bogen-schutz et al. (2021) found that running E3SMv1, at the standard 1.0° horizontal resolution, with vertical resolution on the order of 10 m from the surface to 700 hPa led to a substantial reduction of biases associated with marine stratocumulus, due to better representation of the cloud top cooling and turbulence feedback that drives this particular cloud regime.

While the results of Bogenschutz et al. (2021) provide key insights on how to improve marine stratocumulus, they note that the vertical resolution necessary for significant improvement comes with an excessively high computational expense. A companion paper (Lee et al., 2021) presents results from a configuration of E3SM interfaced with the Framework for Improvement by Vertical Enahancement (FIVE). FIVE allows for select processes to be run at a higher vertical resolution that is independent of the vertical grid used by the host model. For example, Lee et al. (2021) ran E3SM-FIVE with the turbulence, microphysics, radiation, and large-scale vertical advection on the same high resolution vertical grid advocated by Bogenschutz

et al. (2021), with the remaining processes (i.e. dynamics, deep convection, etc.) run on the standard E3SM vertical grid. Lee et al. (2021) found that E3SM-FIVE was able to replicate the results of the benchmark runs of Bogenschutz et al. (2021) and at a much reduced computational cost.

While Bogenschutz et al. (2021) and Lee et al. (2021) found that increasing the vertical resolution in the lower troposphere does lead to significant improvements in the representation of marine stratocumulus, it did not completely ameliorate biases. While more off-shore stratocumulus were improved, stubborn biases remained directly along the coastlines. However, a recent study by Lee et al. (2022) found that the combination of both high horizontal resolution and vertical resolution resulted in a dramatic reduction in stratocumulus biases, that cannot be achieved by refining just one dimension of the model.

Therefore, it is clear that substantial increases to the horizontal and vertical resolutions are required to ameliorate biases related to marine stratocumulus clouds. However, the horizontal and vertical resolution requirements established by Bogenschutz et al. (2021) and Lee et al. (2022) are computationally very expensive and make climate projection simulations all but infeasible. The solution we propose is a novel framework that combines RRM in the horizontal with FIVE in the vertical. The idea is to provide high horizontal resolution only where we want it (i.e., we will apply RRM for one subtropical marine stratocumulus region) and provide high vertical resolution only for the particular physical process schemes that are important for representing marine stratocumulus. We seek to determine if the same improvements can be achieved in such a framework, for our region of interest, and at a reduced computational cost compared to running the model with high horizontal resolution everywhere and high vertical resolution for all processes.

This paper is organized as follows. Section 2 describes the methodology used, including the RRM we generated and a description of FIVE. Section 3 describes the details on the experiments we performed in this paper, while section 4 presents the results of our new framework. Finally, section provides a discussion and summary, including the potential implications of this work for GCMs and global convection permitting models.

## 2  Methods

In this section we provide description of the model used in this study as well as the tools used to refine the resolution in the horizontal and vertical directions.

### 2.1  The Energy Exascale Earth System Model

In this study we use the Energy Exascale Earth System Model version 1 (E3SMv1; Golaz et al., 2019), which was developed with funding by the United States Department of Energy (DOE). The E3SM atmosphere model (EAM; Xie et al., 2018; Rasch et al., 2019) was branched off from the National Center for Atmospheric Research's (NCAR's) CAM, but with significant changes made thereafter to both the dynamical and physical components. In terms of its physics packages, E3SM uses the Cloud Layers Unified by Bi-normals (CLUBB) parameterization (Golaz et al., 2002; Bogenschutz et al., 2013) to serve as a unified treatment of planetary boundary layer turbulence, shallow convection, and cloud macrophysics. For large-scale microphysics, CLUBB is coupled to version two of the Morrison and Gettlemen microphysics scheme (MG2; Morrison and Gettelman 2008; Gettelman

and Morrison, 2015). Deep convection is treated using the parameterization as described by Zhang and McFarlane (1995) (hereafter ZM).

## 2.2 RRM in the Southeast Pacific

Motivated by the results of Lee et al. (2022), we have implemented a new regionally refined mesh (RRM) over a stratocumulus (Sc) region. Their results show that marine stratocumulus situated off the coast of South America appear to be the most responsive to changes in the horizontal and vertical resolution. This region also represents the most persistent area of marine boundary layer clouds in the subtropics and thus makes it the ideal location to implement our experimental marine Sc domain. We refer to this as the Southeast Pacific RRM (SEP-RRM).

The SEP-RRM consists of low resolution (LR) and high resolution (HR) regions and a transition area between them (Fig. 1). The HR grid is located over the heart of the Southeast Pacific stratocumulus area (Klein and Hartman 1993; Bogenschutz et al. 2021; Lee et al. 2021). In addition, the HR grid includes much of the Andes mountains over South America. We created the regionally refined grid with the offline software tool Spherical Quadrilateral grid Generator (SQuadGen, https://github.com/ClimateGlobalChange/sqadgen). The effective resolutions between the LR and HR regions are 1 and 0.25°, respectively. SEP-RRM has a total of 8,351 quadrilateral elements on the globe, with these elements depicted in Fig. 1. As described in Dennis et al. (2012), the Spectral Element (SE) dynamical core operates on quadrilateral elements whose Gauss-Lobatto-Legrendre (GLL) quadrature points form the physics columns in E3SM. The SEP-RRM covers approximately 4.5% of the Earth's surface. In section 3 we compare this number to the standard resolution configurations of E3SM, however, we note that the SEP-RRM is smaller in size to the RRM generated for the Continental United States (CONUS; Tang et al. 2019), which contains 9,905 elements.

## 2.3 Framework for Improvement by Vertical Enhancement

We utilize the novel Framework for Improvement by Vertical Enhancement (FIVE; Yamaguchi et al., 2017), which has been implemented into E3SM to improve marine stratocumulus clouds while substantially saving on computational cost (Lee et al., 2021). FIVE is unique because it allows for select processes or parameterizations to be run at a different vertical resolution than the host model. In the configuration of E3SM-FIVE used in this paper, three physics schemes are interfaced with vertically enhanced physics (VEP) to better represent low clouds: the CLUBB turbulence parameterization, the MG2 microphysics scheme, and the Rapid Radiative Transfer Module for GCMs (RRTMG) longwave and shortwave radiation schemes (Iacono et al., 2018; Mlawer et al., 1997). In addition to the aforementioned physics schemes, large-scale vertical advection in the dynamical core is computed on the high-resolution grid, which is necessary to accurately balance entrainment via the turbulence scheme.

The tendencies calculated in VEP with higher vertical resolution are averaged and applied to the host model (in this case E3SM) for prediction. The VEP calculation does not interfere with the order of the computation of processes in the host model, so the calculation of processes is not repeated between the host model and VEP. Lee et al. (2021) provides a detailed description of how FIVE is interfaced and implemented into E3SM.

## 3 Experiment Design

We present results from a total of six experiments. Our low resolution control model is E3SM run with the ne30 configuration (approximately 1-degree horizontal resolution at the equator) and with the standard 72 vertical layers (hereafter referred to ne30-CNTL). The high resolution control model is E3SM run with the ne120 configuration (approximately 0.25-degree horizontal resolution) and with the standard 72 vertical layers (hereafter referred to as ne120-CNTL).

In the first set of experiment simulations we run with FIVE turned on for each of the horizontal resolution configurations (ne30 and ne120). We refer to these experiments as ne30-FIVE and ne120-FIVE. These simulations are the same as their CNTL experiment counterparts except that FIVE is turned on and thus the turbulence, microphysics, radiation, and large-scale vertical advection calculations are done on the VEP grid. In all FIVE simulations presented in this paper the VEP grid represents an octupling of the vertical resolution compared to the standard vertical resolution in the lower troposphere (between 995 hPa and 700 hPa; figure 2). Thus, the VEP grid contains 212 vertical levels, compared to the E3SM vertical grid of 72 levels. This is the exact vertical grid advocated by Bogenschutz et al. (2021) to achieve the best possible simulation of marine Sc.

For the second set of experiment simulations we use the newly generated Southeast Pacific regionally refined model (SEP-RRM) over the Peruvian region both with the standard 72 vertical layers (hereafter referred to as SEP-RRM-CNTL) and with FIVE turned on with the same vertical grid as the ne30-FIVE and ne120-FIVE experiments (hereafter referred to as SEP-RRM-FIVE). Table 1 provides details on the horizontal and vertical resolution configurations for each experiment.

For this study we treat ne30-CNTL as the default control simulation and the ne120-FIVE simulation as the high resolution benchmark we strive to replicate within the Southeast Pacific with the SEP-RRM-FIVE simulation. We recognize that a true high resolution benchmark simulation would involve running ne120 with 212 vertical levels. However, this simulation would have been computationally burdensome to perform (see section 3.3). Furthermore, the work of Lee et al. (2021) has already established that simulations with FIVE can replicate the results of the high resolution benchmark simulations in regards to the representation of marine stratocumulus.

In this study we perform atmosphere-only simulations forced by observed present day climatology of aerosol emissions, greenhouse gases, sea surface temperatures (SSTs), and sea ice concentrations. The simulations use the interactive E3SM land model on the same grid as the atmosphere in each configuration. All simulations presented in table 1 are run for five years.

### 3.1 Time Steps

Table 2 presents the physics and dynamical time steps used for each simulation. We note that ne30-FIVE and ne120-FIVE simulations use the exact same time step settings as their control simulation counterparts (ne30-CNTL and ne120-CNTL, respectively). Despite the fact that simulations with FIVE use eight times the vertical resolution, compared to the standard E3SM 72 layer grid in the lower troposphere for several processes, no time step reduction is needed as the schemes interfaced with VEP grid use implicit schemes and thus not constrained by the Courant-Friedrichs-Lewy (CFL) condition. Lee et al. (2021) report this as being a significant advantage of using FIVE versus refining the grid of the entire E3SM model (Bogenschutz et al., 2021), for the application of stratocumulus clouds, and results in a major savings in terms of the computational cost.

We also note that simulations using SEP-RRM use the same time step settings as the ne120 simulations, to retain the numerical stability in the high resolution region. Though, we do note that this could introduce some degree of time step sensitivity in the LR region of our SEP-RRM experiments as E3SM clouds and convection are known to exhibit some sensitivity to time step (Wan et al., 2020).

In our study the ne30 (ne120) simulations are run with an 1800 s (900 s) physics time step. The dynamics use three layers of substepping. For the ne30 (ne120) simulations, the Lagrangian vertical discretization time step is 15 min (2.5 min), the horizontal discretization time step is 5 min (75 s), and the explicit numerical diffusion time step is 100 s (18.75 s). As already stated, for the SEP-RRM simulations, we use the same physics and dynamics time step settings as the ne120 simulations.

## 3.2 Tunable Parameters

We note that all simulations presented in this paper are run with the exact same tunable parameters. While default ne30 and ne120 configurations of E3SMv1 are run with a slightly different set of tunable parameters in the CLUBB and Zhang-McFarlane (ZM) deep convection scheme (Zhang and McFarlane, 1995) to achieve radiation balance and optimal skill scores for coupled simulations, we chose to run all simulations with the default parameters for the ne30 configuration. This was a conscious decision we made because otherwise it could be very difficult to disentangle whether differences in our regime of interest (marine stratocumulus) would be due to tuning or resolution differences. This is a similar approach Caldwell et al. (2019) took when comparing low and high resolution E3SMv1 simulations, when they wanted to avoid potential ambiguities that could arise from different tunable parameter sets.

In addition, it is common practice for RRM simulations to run with tunable parameters that are default for the HR mesh (Tang et al., 2019), with the argument being that they desire to mimic the HR behavior as closely as possible in their RRM runs. However, for our work, we feel it is most appropriate to maintain the ne30 tunable parameters for SEP-RRM, regardless of whether or not we chose to maintain these tunable parameters for our ne120 simulations. The reason being that SEP-RRM is focused on a marine stratocmulus regime where the deep convection scheme is not active and most of the work is being done by the turbulence scheme, CLUBB. CLUBB is a scale aware parameterization that has been shown to exhibit little sensitivity to the horizontal grid size (Larson et al., 2012), whereas the deep convection scheme in E3SMv1 is not scale aware. Therefore, since our area of interest does not contain deep convection, choosing to retain the ne30 tunable parameters so as not to modify the deep convection occurring in our LR region makes the most sense. Running the deep convective regime at ne30 and with ne120 tunable parameters could potentially modify the Hadley circulation and thereby influence the results in our SEP-RRM domain.

## 3.3 Computational Cost

Table 3 displays the computational cost of each simulation. All of the runs presented in this paper were performed on Livermore Computing's (LC) Quartz machine. Table 3 provides the node hours required for each simulation per simulated year and is computed as the total number of nodes used multiplied by the hours to get a simulated year. As expected, the ne30-CNTL

185 simulation is the least computationally expensive, while the ne120-FIVE simulation is the most computationally expensive (i.e. two orders of magnitude more expensive than ne30-CNTL).

In general, runs using FIVE tend to be 2-3 times as expensive as their CNTL run counterparts using the same horizontal resolution. As reported by Lee et al. (2021), the added cost associated with FIVE is small when compared to the cost of running the entire model at high vertical resolution, which would have yielded costs an entire order of magnitude higher. We estimate

190 that a simulation with ne120 resolution and 212 vertical levels would have costed nearly 300,000 node hours per simulated year, with the majority of this expense being attributed by the need to run with a very small time step. Thus, without FIVE this would make running ne120 simulations with high vertical resolution infeasible for this study.

Table 3 demonstrates that while both SEP-RRM and SEP-RRM-FIVE are roughly three times more expensive than ne30-CNTL and ne30-FIVE, respectively, they are both an order magnitude less expensive than their ne120 counterpart simulations.

195 This represents the reasonable trade off in cost that is expected with these RRM simulations (Tang et al. 2019).

## 4  Results

### 4.1  Low Cloud Climatology

Since our target regime is subtropical marine stratocumulus, we will first focus on results relating to the simulated low cloud amount and the shortwave cloud radiative effect (SWCRE). Figure 3 displays the climatological observed values for these

200 two quantities. Observations for low cloud amounts are provided from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) from January 2007 to January 2010 while those for SWCRE are provided by the Clouds and Earth's Radiant Energy System (CERES) Energy Balance and Filled (CERES-EBAF) product from January 2001 to January 2018. Focusing on the subtropics, we see three prominent local maxima of low cloud amounts off the western coasts of South America, California, and Namibia with two lesser maxima near the western coasts of Australia and North Africa. These cloud

205 maxima are due to the subtropical marine stratocmulus (Klein and Hartman 1993) and area also indicated in the observed figure of SWCRE by areas of strong reflectivity.

Figure 4 displays the geographical bias of simulated low level cloud amount for our six configurations computed relative to CALIPSO observations. We note that our E3SM simulations use the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP; Bodas-Salcedo et al., 2011), which contains several independent satellite simulators

210 to diagnose model clouds in a similar way that satellites would view an atmosphere. This allows for an apples-to-apples comparison of our simulated low cloud amount with CALIPSO observations. Low level cloud amount represents the vertically integrated cloud fraction from the surface to 680 hPa as defined in Chepfer et al. (2008). As already indicated by Bogenschutz et al. (2021) and Lee et al. (2021), the ne30-CNTL (fig. 4a) simulation is characterized by a dearth of low level cloud in the three main sub-tropical stratocumulus regions. This is a bias endemic to most GCMs and is not specific to E3SM. While increasing

215 vertical resolution through FIVE (ne30-FIVE) helps to ameliorate the bias associated with low cloud for stratocumulus that resides offshore (Fig. 4b), biases near the coast remain. In addition, while increasing horizontal resolution (ne120-CNTL) alone

appears to marginally improve biases of stratocmulus along the coast, it does little towards reducing the overall bias associated with this cloud regime.

The concurrent increase of horizontal and vertical resolution (ne120-FIVE) has the most substantial impact on the local bias reduction of subtropical marine stratocumulus (Fig. 4d; Lee et al. submitted). Here we seek to answer the question if RRM can help to replicate these results for one of the subtropical stratocumulus regimes, specifically that residing in the SE Pacific.

Figures 4e-f display the low level cloud amount biases for the SEP-RRM-CNTL and SEP-RRM-FIVE simulations, respectively. Our focus region is the stratocumulus in the Southeast Pacific, as denoted by the bounding box in these figures. The hope is that the bias displayed within the bounding box of Fig. 4e (SEP-RRM-CNTL) resembles that of the corresponding region of Fig. 4c (ne120-CNTL) and the same for the sets of simulations using FIVE (Fig. 4f versus Fig. 4d). While the degree that the results can be replicated in the RRM simulations will be quantified in further figures, it appears that this can be successfully done. This is especially obvious when comparing the SEP-RRM-FIVE results in the SEP to that of the ne120-FIVE simulation where the local bias reduction is substantial and very similar when comparing these two simulations.

Figure 5 directly compares the SEP-RRM-FIVE simulation with the ne120-FIVE and ne30-FIVE simulations for the refined and outer regions. Focusing on the differences with respect to the high resolution benchmark ne120-FIVE simulation (Fig. 5a) we see virtually no difference in the change of low level cloud amount within the refined region, while there are relatively large differences around the rest of the globe, including the subtropical stratocumulus regions near California and Namibia. On the other hand, the differences with respect to ne30-FIVE (Fig. 5b) show nearly the opposite behavior, with relatively large differences in the refined region and the smaller difference in the rest of the globe. Overall, this is the expected result and provides encouraging evidence that the new SEP RRM is performing as hypothesized. In Fig. 5b, since the horizontal and vertical resolutions in the outer domain are identical, this figure could presumably show the global impact of reduced stratocumulus biases in the SEP. To explore this, the stippled areas indicate where differences are statistically significant at the 95% confidence level. This suggests improving stratocumulus biases in the SEP could have implications on the simulation within the ITCZ (and associated double ITCZ problem) and southern hemisphere storm tracks. Though, we caution that these two simulations do have different time step settings (Table 2), which could be contributing to these differences. In addition, these simulations use prescribed SSTs which could potentially be muting the global impact. Exploring the global effects from improved marine stratocumulus biases is something that could be pursued in future work. The SEP-RRM-FIVE configuration would provide the ideal framework to do such a study.

The bias and root mean squared error (RMSE) scores for low level cloud amount can be found in Fig. 6. Here we have segregated the scores into two categories, one which includes the bias and RMSE for the entire globe for each configuration and the other which represents the skill scores exclusively for oceanic columns in SE Pacific region. Focusing first on the global skill scores, we see that all simulations using FIVE are characterized by an increase in the global low level cloud. This is mostly due to an increase in stratocumulus and trade wind cumulus clouds across the entire subtropical oceans when FIVE is activated. However, in terms of RMSE, there is generally a less dramatic difference between the simulations, though the simulations using FIVE score slightly better than their CNTL counterparts.

Focusing on the skill scores for the SE Pacific region, however, we see a much more distinct behavior as the horizontal and vertical resolution changes. For the SE Pacific the expectation is that the skill metrics would be very similar for the SEP-RRM-CNTL and ne120-CNTL set of simulations and the SEP-RRM-FIVE and ne120-FIVE set of simulations. For both the bias and RMSE skill scores we see that both sets of simulations agree very well with each other, suggesting that the SEP-RRM is able to well replicate the behavior of the ne120 configurations for marine stratocumulus.

Beyond demonstrating the effectiveness of the SEP-RRM to replicate the behavior of high resolution runs for marine stratocumulus, Fig. 6 also illustrates the importance of concurrent increases of horizontal and vertical resolution when simulating marine Sc. All three simulations that use the default 72 vertical levels for all processes greatly underestimates cloud cover with associated large errors for this region. Simply increasing the horizontal resolution, as evident in the ne120-CNTL simulation, does little to ameliorate error. The introduction of FIVE in the low horizontal resolution configuration, ne30-FIVE, demonstrates a noticeable reduction in regards to the bias and error. However, it is not until the horizontal and vertical resolutions are concurrently increased (ne120-FIVE and SEP-RRM-FIVE) where substantial reductions in the bias and RMSE are seen.

We are most concerned with the simulation of low clouds because of their impact on shortwave radiation. Therefore, we also examine the simulation of SWCRE for each simulation. The annually averaged observed SWCRE can be found in Fig. 3b with the differences for each simulation compared to observations found in Fig. 7. In general, the same behaviors found in the analysis with low level cloud amount can be found in SWCRE. The ne30-CNTL simulation is characterized by too little reflectivity in the subtropics, with bull's eyes on the subtropical marine Sc. These biases are only marginally reduced in the ne120-CNTL simulation, with slight improvements seen in the coastal regions of the marine Sc. This follows the general conclusions of Caldwell et al. (2019). While the introduction of FIVE to the ne30 simulation goes a long way in helping to ameliorate the SWCRE biases in the subtropics, it is not until the horizontal resolution is concurrently increased in the ne120-FIVE simulation when large reductions in the biases are found. In addition, similar to the result of low level cloud amount, the results for the two SEP-RRM runs (CNTL and FIVE) appear to behave as expected; with bias patterns and magnitude that look very similar to each ne120/ne30 counterpart within the refined/outer region.

To exploit the behavior of the of the SEP-RRM for SWCF we directly compare the SEP-RRM-FIVE simulation with the counterpart ne120-FIVE (Fig. 8a) and ne30-FIVE (Fig. 8b) simulations. Similar to low level cloud amount, we see excellent agreement of SWCRE with the ne120-FIVE experiment within the refined region and relatively large differences for the rest of the globe. Conversely, there are large differences of SWCRE with the ne30-FIVE experiment within the refined region. Again, while this is the result we hoped for, a very interesting feature is the robustness of the simulated SWCRE within the refined region for SEP-RRM-FIVE with respect to ne120-FIVE, despite there being large differences in the simulated SWCRE in the rest of the globe. This is especially true in the tropics where there are relatively large differences in SWCRE. While this is not especially surprising given that the resolutions of these two simulations are different in these regions that will cause a change in the simulated convection due to the convective parameterization of E3SM not being scale aware, it is often thought that modified deep convection simulation will alter the large scale circulation thereby inducing changes to the simulated marine Sc. However, here we see that refining the local resolution is the first order sensitivity for marine Sc.

285     Focusing on the quantitative skill score metrics for SWCRE we see that most simulations have minimal bias with respect to global domain (Fig 9a). This is not surprising since ne30-CNTL was tuned to achieve radiation balance, though the remainder of simulations which were not tuned tend to exhibit either a slightly positive or negative bias. Conversely, when just examining the SE Pacific region, we see that all three CNTL simulations exhibit a large positive bias. This bias is substantially reduced with the introduction of FIVE but is especially better and comparable for the ne120-FIVE and SEP-RRM FIVE simulations,

290 as expected. The global RMSE skill scores for all six simulations are very similar for the global domain (Fig. 9b) but again, demonstrates considerable differences for the SE Pacific region. Here we see that ne30-CNTL has the largest error, with the ne120-FIVE and SEP-RRM-FIVE simulations having the best scores and being very comparable to one another.

    Figure 10 presents five year annual averaged profiles for oceanic columns from the SEP RRM domain. The profiles of cloud fraction and cloud liquid amount (figs. 10a-b) demonstrate that the simulations including FIVE tend to simulate thicker

295 cloud decks that contain more coverage and liquid water relative to their CNTL counterpart configurations. This is in general agreement with the results presented in Lee et al. (2021). Here we demonstrate that the SEP-RRM-FIVE simulation bears remarkable agreement to the simulation characteristics of ne120-FIVE. This agreement is also present in the averaged profiles for the longwave cooling rate and the vertical velocity variance ($\overline{w'^2}$) predicted by CLUBB. The agreement provided by SEP-RRM-FIVE and ne120-FIVE suggests that the simulated turbulence and cloud top feedback at the process level are roughly

300 the same between the two simulations.

    Thus far we have demonstrated that the SEP-RRM is effective in replicating the results that global high resolution simulations provide, but with a refined patch over the region of interest. We have also demonstrated that the combination of increased horizontal resolution and vertical resolution (through the use of FIVE) is required to produce the best possible results with regards to the simulation of marine stratocumulus. Bogenschutz et al. (2021) found that LES-like vertical resolution was needed

305 to achieve satisfactory simulation of marine stratocumulus and the primarily mechanism was a positive feedback between the cloud top cooling, turbulence, and cloud macrophysics processes. They found that the high vertical resolution was better able to resolve the sharp cloud top inversion, which leads to a more accurate representation of the height of maximum cloud and liquid when compared to satellite retrievals. This helps to generate copious amounts of cloud top cooling, which encourages turbulence of the boundary layer to be primarily driven from cloud top. The lower vertical resolution simulations lacked

310 sufficient cloud cover and allowed an abundance of solar radiation to reach the surface during the daytime and caused turbulence generation to be primarily surface driven. In turn, this led to decoupled profiles of $\overline{w'^2}$ and positive values of third moment of vertical velocity ($\overline{w'^3}$) which were more indicative of a shallow cumulus boundary layer and helped to further thin the cloud deck as vertical velocity skewness (defined as $(\overline{w'^3}/\overline{w'^2})^{3/2}$) is used in CLUBB to parameterize cloud fraction and liquid water mixing ratio (Golaz et al. 2002; Larson et al. 2002). On the other hand, the higher vertical resolution case, where turbulence is

315 primarily generated from cloud top, produces neutral or negative values of vertical velocity skewness, which encourages more abundant cloud formation and a positive feedback.

    Lee et al. (2021) found that this positive feedback loop to improve stratocumulus could be replicated in FIVE as long as CLUBB, microphysics, radiation, and large scale vertical advection are all interfaced with the vertically enhanced physics. Finally, Lee et al. (2022) found a substantial improvement to the simulated coastal stratocumulus with concurrent increases in

320   horizontal and vertical resolution; to better capture the coastlines and sharp terrain gradients, which is particularly important for the Southeast Pacific straocumulus region given its proximity to the Andes Mountains.

## 4.2   Seasonal Cycle of Low Clouds in the Southeast Pacific

While we have thus far demonstrated the ability of SEP-RRM to represent stratocumulus through the climatological mean, in this section we seek to address how our highest resolution configuration (ne120-FIVE) can represent the seasonal cycle of
325   low clouds in the Southeast Pacific and how well the RRM-FIVE configuration can replicate this. Because we found that the simulations with FIVE are required to obtain superior results for the climatological mean, we will only focus our analysis on comparing the ne120-FIVE and SEP-RRM-FIVE simulations against observations and the ne30-CNTL model.

    Figure 11 displays the observed seasonal cycle of low-level cloud amount and SWCRE for the SE Pacific region that is subject to regional refinement in our simulations. The figures are segregated by season; namely boreal winter (December,
330   January, February; DJF), spring (March, April, May; MAM), summer (June, July, August; JJA), and autumn (September, October, November; SON). While marine stratocumulus persists over the ocean in all seasons for this region, the climatological maximum occurs during SON, with a minimum during MAM. The JJA and SON seasons see a relative northward shift and maximum of coastal Sc along the coast of Peru, while DJF and MAM sees a southward shift of marine Sc and a maximum of coastal Sc along the Chilean coast.

335     Differences of low cloud amount for the ne30-CNTL and ne120-FIVE simulations, computed relative to CALIPSO observations, are displayed in the first two columns of Fig. 12. It is clear that the ne30-CNTL simulation systematically underestimates low level cloud for all seasons. This bias is most severe along the coast but also persists off-shore, particularly during DJF and MAM. It is also noteworthy that widespread regional deficits along the coast upwards of 50% exist during the peak seasons of JJA and SON, highlighting the severity of the Sc bias in E3SM that is not as obvious when only considering the global annual
340   averages.

    Examining the biases of the ne120-FIVE simulation we see noteworthy improvement for all seasons, though some biases still remain to a degree. The most stubborn bias in the ne120-FIVE simulation is slight underestimation of low cloud along the Chilean coast for all seasons, though the magnitude of this bias is significantly smaller when compared to ne30-CNTL. In addition, whereas ne30-CNTL has a persistent underrepresentation of low clouds along the Peruvian coast, ne120-FIVE
345   now has a slight overestimation of low level clouds in this area during the MAM and SON seasons. Can SEP-RRM-FIVE replicate the improvements that ne120-FIVE demonstrates at the seasonal timescale? An examination of the biases relative to observations of SEP-RRM-FIVE (third column of fig. 12) shows very similar bias patterns and magnitude to that of ne120-FIVE. The similarity of these two simulations is further confirmed when looking at the differences between SEP-RRM-FIVE and ne120-FIVE (fourth column of fig. 12) which shows near perfect agreement for all seasons. This further demonstrates the
350   ability of RRM to replicate the high resolution behaviors of a Sc regime.

    Figure 13 displays the seasonal biases for the SWCRE. Overall, we see the expected result given the biases presented for the low level cloud amounts in Fig. 12. The ne30-CNTL simulation suffers from systematic bias of too little reflectivity that is most severe along the coast. Whereas the ne120-FIVE simulation improves upon this bias, it is notable that relatively large biases

exist for this configuration both along the coast and offshore during the DJF and SON seasons. These biases are somewhat larger

355    than the biases in the low level cloud amount may suggest and points to the need for further parameterization improvements or tuning to the parameterization suite in E3SM to enhance the agreement with observations. In addition, further refinement of the vertical grid in VEP could lead to additional improvements, as suggested in Lee et al. (2021). However, similar to the analysis of low level cloud amount, we see that the SEP-RRM-FIVE has excellent agreement with the much more expensive ne120-FIVE simulation (third and fourth columns of Fig. 7).

360    **4.3   General Climatology**

This paper has been focused on the simulation of low-level marine clouds and their sensitivity to horizontal and vertical resolution. However, it is important to assess the credibility of the overall simulated global climate to ensure that potential significant degradations in our experimental simulations are not providing unrealistic boundary conditions to our region of interest. We have assessed the credibility of the global simulated climate and have verified that no configuration contains any

365    major pathological error or bias relative to ne30-CNTL. Here, we succinctly present our findings to support this conclusion.

Table 4 displays the top of atmosphere radiation budgets for each experiment. Since these simulations use present day forcing, we cannot expect a near zero top of atmosphere energy balance. The ne30-CNTL simulations shows an energy balance (RESTOM) of about 2.77 $\mathrm{W}m^{-2}$. As already noted, we chose to run our ne120 simulations with the same tunable parameters as the ne30 simulations, therefore it is not surprising that these simulations differ in terms of their RESTOM values. Therefore,

370    it is clear that any experiment simulation presented in this paper would be subjected to some tuning to achieve radiation balance should they be run in coupled mode, which could have some minor impact on the overall results presented in this paper.

Figure 14 displays diagrams (Taylor 2001) to succinctly compare model skill between the various experiment simulations for climatologically important variables such as SWCRE, LWCRE, precipitation, and reference temperature height. It is expected that ne30-CNTL would tend to produce the best results since that simulation was tuned for E3SM production runs and overall

375    this does appear to be the case here. However, all simulations tend to produce similar skill score metrics, further highlighting the fact that our experimental simulations are credible at the global scale. The biggest spread between the simulations appears to be for LWCRE and precipitation, where the ne120 simulations tend to show slightly less skill than the remainder of the simulations.

## 5   Conclusions

380    In this paper we develop a novel framework that refines the horizontal resolution of E3SM in the southeast Pacific (SEP) while simultaneously refining the vertical resolution for select physics processes to reduce biases related to subtropical marine stratocumulus clouds. Our regional refined mesh (RRM) is centered off the coast of Peru and Chile with an effective resolution of $0.25°$, with the standard $1.0°$ resolution covering the remainder of the globe. Our vertical resolution refinement takes advantage of the Framework for Improvement by Vertical Enhancement (FIVE; Yamachuci et al., 2017) to run E3SM's turbulence,

385    microphysics, and radiation schemes on a much higher vertical resolution than the standard 72 layer E3SM grid. In this study

we refine the vertical resolution to near 10 m in the lower troposphere for FIVE, which is representative of the resolution often used in large eddy simulation for stratocumulus.

We find that the superior results produced by the high resolution benchmark simulation (ne120-FIVE) are reproduced with great fidelity in the SEP region with our new SEP-RRM-FIVE configuration. In addition, the computational cost of the SEP-RRM-FIVE simulation is much reduced compared to ne120-FIVE, as expected, as it uses only 5% of the computational resources as the high resolution benchmark simulation. We recognize that in this paper we did not perform a true high resolution benchmark, which would have involved running ne120 with 212 vertical levels, due to the excessive computational cost that would demand (Bogenschutz et al., 2021) and due to the fact that it has already been established that FIVE can reproduce the results for marine stratocumulus clouds compared to running high vertical resolution for all processes (Lee et al., 2021). However, we project that compared to a true high resolution benchmark (ne120 horizontal resolution with 212 vertical levels), the cost of the SEP-RRM-FIVE configuration would represent only 0.05% of the total cost of said simulation. This highlights the major cost benefits of using horizontal mesh refinement in conjunction with FIVE to target marine stratocumulus biases.

We demonstrate that concurrent increases in the horizontal and vertical resolution are needed for substantial reduction of the stratocumulus biases and realistic seasonal representation in the Peruvian region. We also show that the SEP-RRM-FIVE configuration can replicate the results of the ne120-FIVE configuration in regards to this. This provides an affordable framework to perform experiments related to climate sensitivity and aerosol affects in a simulation with improved and realistic stratocumulus representation. Albeit, we recognize that this work refines only one stratocumulus region. However, it is completely feasible for one to horizontally refine multiple regions of the globe. For instance, one could conceivably refine the Californian, Peruvian, and Namibian low cloud regions simultaneously. Given the relatively small horizontal extent that marine stratocumulus cover, the size requirement of the horizontal refined patches would likely still allow for big computational cost savings.

Another opportunity to increase throughput of these simulations is to only activate FIVE in the geographical regions of interest and run with the standard 72 layer grid for all columns and processes outside of these regions. Currently, the FIVE implementation into E3SM can only be activated in all (or no) columns globally and with the same number of VEP levels. An extension to make FIVE activate only in select columns is theoretically possible, but made challenging by the fact that all parameterizations in E3SM (i.e. CLUBB, MG2, RRTMG), and likely in most GCMs, are initialized with the same number of vertical levels in all columns and assumes that all columns will use the same number of vertical levels throughout the duration of the simulation. This provides a software engineering challenge and one that would likely be worthwhile to overcome, given that stratocumulus are particularly sensitive to vertical resolution and cover a small portion of the globe.

Finally, while our prototype framework has been implemented in the context of the conventionally parameterized GCM, we expect this framework to be useful for the next generation of GCMs as we gradually move to cloud resolving scales. In terms of the vertical resolution, most global cloud resolving models (GCRMs; i.e. Caldwell et al., 2021) run with relatively coarse vertical resolution in the lower troposphere. Thus, this highlights the importance of tools such as FIVE to ensure that the minimum vertical resolution requirement to simulate stratocumulus is met (Lee et al., submitted) in these models. Though GCRMs often have horizontal resolutions on the order of 3 km, which is enough to adequately resolve the coastlines, boundary layer turbulence is still not resolved at these resolutions. While some GCRMs contain sufficient boundary layer

parameterizations to represent stratocumulus reasonably enough, refining a horizontal mesh over a particular stratocumulus region could allow for simulations where the large eddies are resolved, providing a computationally affordable framework for assessing the performance and potential benefits of LES in the context of a global simulation.

*Code and data availability.* The code used to generate the model output, in addition to the model output itself, can be found at
425    https://doi.org/10.5281/zenodo.6986304.

*Author contributions.* Bogenschutz created the Southeast Pacific regionally refined mesh and led the analysis, Lee performed some of the simulations presented in this paper and contributed to analysis, Tang assisted in the generation of the new regionally refined mesh, and Yamaguchi helped to conceive the idea of this work and provided support for FIVE.

*Competing interests.* The authors declare that they have no competing interests.

430    *Acknowledgements.* This work was funded by Scientific Discovery through Advanced Computing (SciDAC), award number DE-SC0018650, by the U.S. Department of Energy office of Biological and Environment Research. The authors thank two anonymous reviewers for providing valuable comments to improve this manuscript. Work at LLNL was performed under the auspices of the U.S. DOE by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. IM release: LLNL-JRNL-836836-DRAFT.

# References

435 Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J.-L., and Klein, S. A.: COSP: Satellite simulation software for model assessment. Bull. Amer. Met. Soc., 92(8), 1023-1043, https://doi.org/10.1175/2011BAMS2856.1.

Bogenschutz, P.A., Gettelman A., Morrison H., Larson V. E., Craig C., and Schanen D. P.: Higher-Order Turbulence Closure and Its Impact on Climate Simulations in the Community Atmosphere Model, J. Climate, 26, 9655-9676, https://doi.org/10.1175/JCLI-D-13-00075.1, 2013.

440 Bogenschutz, P. A., Yamaguchi, T., and Lee, H.-H.: The Energy Exascale Earth System Model Simulations With High Vertical Resolution in the Lower Troposphere, 13(6), e2020MS002239, https://doi.org/10.1029/2020MS002239, 2021.

Bony, S., and Dufresne, J.-L., Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models, Geophys. Res. Lett., 32, L20806, doi:10.1029/2005GL023851, 2005.

Caldwell, P. M., Mametjanov, A., Tang, Q., Van Roekel, L. P., Golaz, J.-C., Lin, W. et al., The DOE E3SM coupled model
445 version 1: Description and results at high resolution. Journal of Advances in Modeling Earth Systems, 11, 4095– 4146. https://doi.org/10.1029/2019MS001870, 2019.

Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschutz, P., Lin, W., et al.: Convection-permitting simulations with the E3SM global atmosphere model. Journal of Advances in Modeling Earth Systems, 13, e2021MS002544. https://doi.org/10.1029/2021MS002544, 2021.

450 Chepfer, H., Bony, S., Winker, D., Chiriaco, M., Dufresne, J.-L., and Seze, G.: Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model. Geophys. Res. Letters, 35, L15704, https://doi.10.1029/2008GL034207, 2008.

Dennis, J. M., Edwards, J., Evans, K. J., Guba, O., Lauritzen, P. H., Mirin, A. A., and Worley, P. H.: CAM-SE A scalable spectral element dynamical core for the Community Atmosphere Model, Int. J. High Perform, 26, 74-89, https://doi.org/10.1177/1094342011428142, 2012.

Gettelman, A., Morrison, H., Santos, S., Bogenschutz, P., and Caldwell, P. M.: Advanced Two-Moment Bulk Microphysics for Global
455 Models. Part II: Global Model Solutions and Aerosol–Cloud Interactions, Journal of Climate, 28(3), 1288-1307. doi:10.1175/jcli-d-14-00103.1, 2015.

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

Golaz, J., Caldwell, P. M., Roekel, L. P. V., Petersen, M. R., Tang, Q., and Wolfe, J. D.: The DOE E3SM coupled model version 1: Overview
460 and evaluation at standard resolution, J. Adv. Model. Earth Syst., 11, 2089-2129, https://doi.org/10.1029/2018MS001603, 2019.

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113(D13). doi:10.1029/2008jd009944, 2008.

Klein, S. A., Hartmann D. L.: The seasonal cycle of low stratiform clouds, J. Climate, 6, 1587-1606. https://doi.org/10.1175/1520-0442.
465 2002 Larson, V. E., Golaz, J.-C., and Cotton, W. R.: Small-scale and mesoscale variability in cloudy boundary layers: Joint probability density functions, J. Atmos. Sci., 59, 3519-3539. doi: https://doi.org./10.1175/1520-0469(2002)059<3519:SSAMVI>2.0.CO;2, 2002.

Lee, H.-H., Bogenschutz, P., and Yamaguchi, T.: The implementation of Framework for Improvement by Vertical Enhancement into Energy Exascale Earth System Model, J. Adv. Model. Earth Syst., 13(6), e2020MS002240. doi:https://doi.org/10.1029/2020MS002240, 2021.

Lee, H.-H., Bogenschutz, P., and Yamaguchi, T.: Resolving away stratocumulus biases in modern GCMS, Geophys. Res. Letters., 49,
470 e2002GL099422, https://doi.org/10.1029/2022GL099422, 2022.

Medeiros, B., Williamson, D. L., Hannay, C., and Olson, J. G.: Southeast pacific stratocumulus in the Community Atmosphere Model. J. Climate, 25(18), 6175-6192, https://doi.org/10.1175/JCLI-D-11-00502.1.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, Journal of Geophysical Research: Atmospheres, 102(D14), 16663-16682. https://doi:10.1029/97jd00237, 1997.

Morrison, H., and Gettelman, A.: A New Two-Moment Bulk Stratiform Cloud Microphysics Scheme in the Community Atmosphere Model, Version 3 (CAM3). Part I: Description and Numerical Tests, J. Climate, 21(15),3642-3659, https://doi:10.1175/2008jcli2105.1, 2008.

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

Ringler, T., Ju, L. and Gunzburger, M.: A multiresolution method for climate system modeling: application of spherical centroidal Voronoi tessellations. Ocean Dynamics 58, 475–498, https://doi.org/10.1007/s10236-008-0157-2, 2008.

Tang, Q., Klein, S. A., Xie, S., Lin, W., Golaz, J. C., Roesler, E. L., and Zheng, X.: Regionally refined test bed in E3SM atmosphere model version 1 (EAMv1) and applications for high-resolution modeling, Geosci. Model Dev., 12(7), 2679-2706, https://doi:10.5194/gmd-12-2679-2019, 2019.

Taylor, K. E.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183-7192, https://doi.org/10.1029/2000JD900719, 2001.

Wan, H., Woodward, C. S., Zhang, S., Vogl, C. J., Stinis, P., Gardner, D. J., et al.: Improving time step convergence in an atmosphere model with simplified physics: The impacts of closure assumption and process coupling. Journal of Advances in Modeling Earth Systems, 12, e2019MS001982, https://doi.org/10.1029/2019MS001982, 2020.

Xie, S., Lin, W., Rasch, P. J., Ma, P.-L., and Neale, R.: Understanding cloud and convective characteristics in version1 of the E3SM atmosphere model. J Adv. Model. Earth Syst., 10, 2618-2644, https://doi.org/10.1029/2018MS001350, 2018.

Yamaguchi, T., Feingold, G., and Larson V. E.: Framework for improvement by vertical enhancement: A simple approach to improve representation of low and high-level clouds in large-scale models, J. Adv. Model. Earth Syst., 9(10), 627-646, https://doi:10.1002/2016MS000815, 2017.

Zarzycki, C. M., and Jablonowski, C.: Experimental Tropical Cyclone Forecasts Using a Variable-Resolution Global Model, Monthly Weather Review, 143(10), 4012-4037. Retrieved Feb 16, from https://journals.ametsoc.org/view/journals/mwre/143/10/mwr-d-15-0159.1.xml, 2015.

Zhang, G. and McFarlane, N. A.: Sensitivity of climate simulations to the parameterization of cumulus convection in the canadian climate center general circulation model, Atmosphere-OCean, 33(3), 407-446, https://doi.org/10.1080/07055900.1995.9649539, 1995.

Zhang, Y., Xie, S., Lin, W., Klein, S. A., Zelinka, M., and Ma, P.-L.: Evaluation of clouds in version 1 of the E3SM atmosphere model with satellite simulators. J. Adv. Model. Earth Syst., 11(5), 1253-1268, https://doi.org/10.1029/2018ms001562.
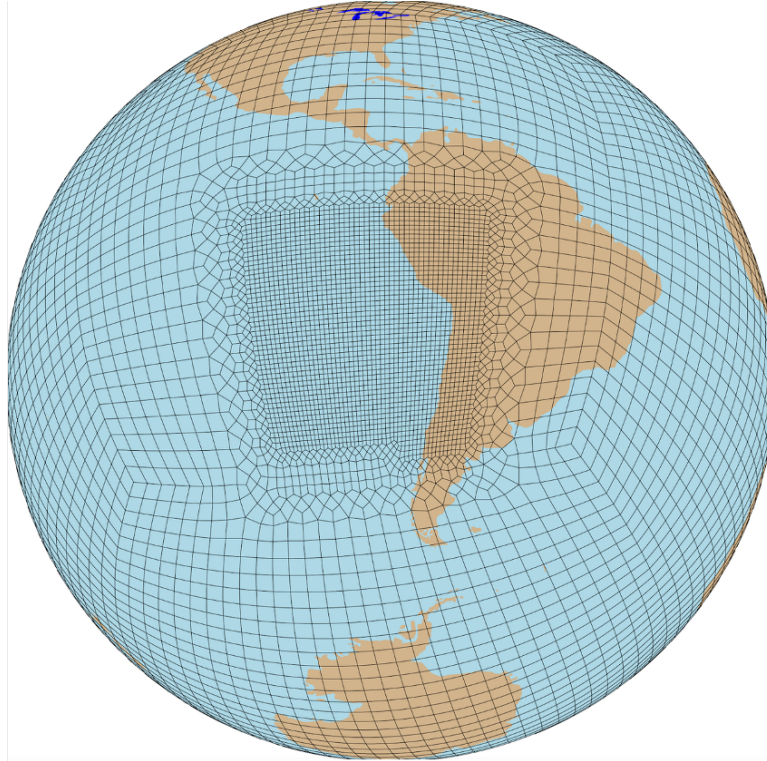
**Figure 1.** Depiction of the grid mesh used for the SEP-RRM-CNTL and SEP-RRM-FIVE simulations in this paper with ne120 resolution used within the refined region and ne30 resolution outside of this region. Note that each element shown contains additional 3 x 3 collocation cells.
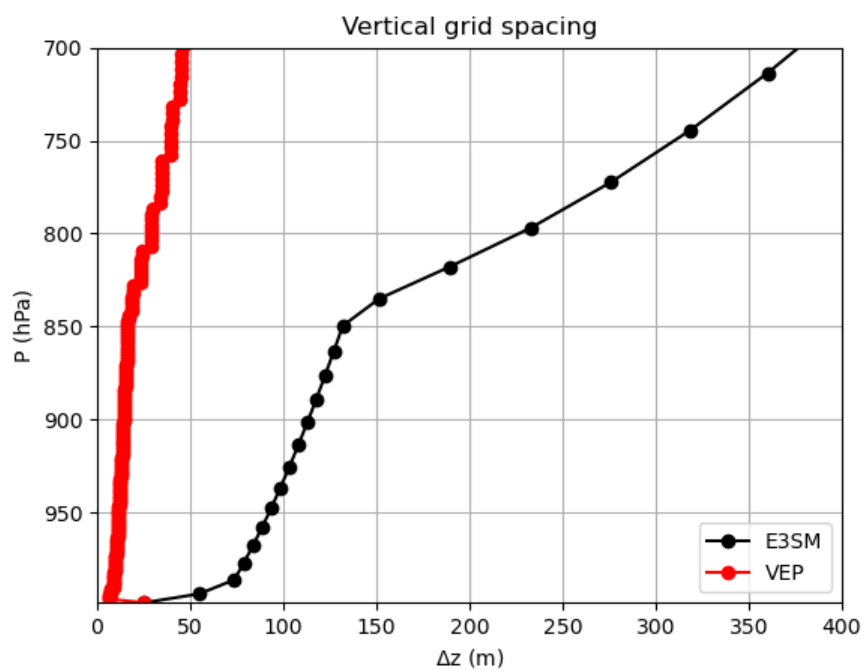
**Figure 2.** Estimated vertical grid spacing for the vertical grid used in E3SM (black curve) and for the vertically enhanced physics (VEP) used in the FIVE simulations in this paper.
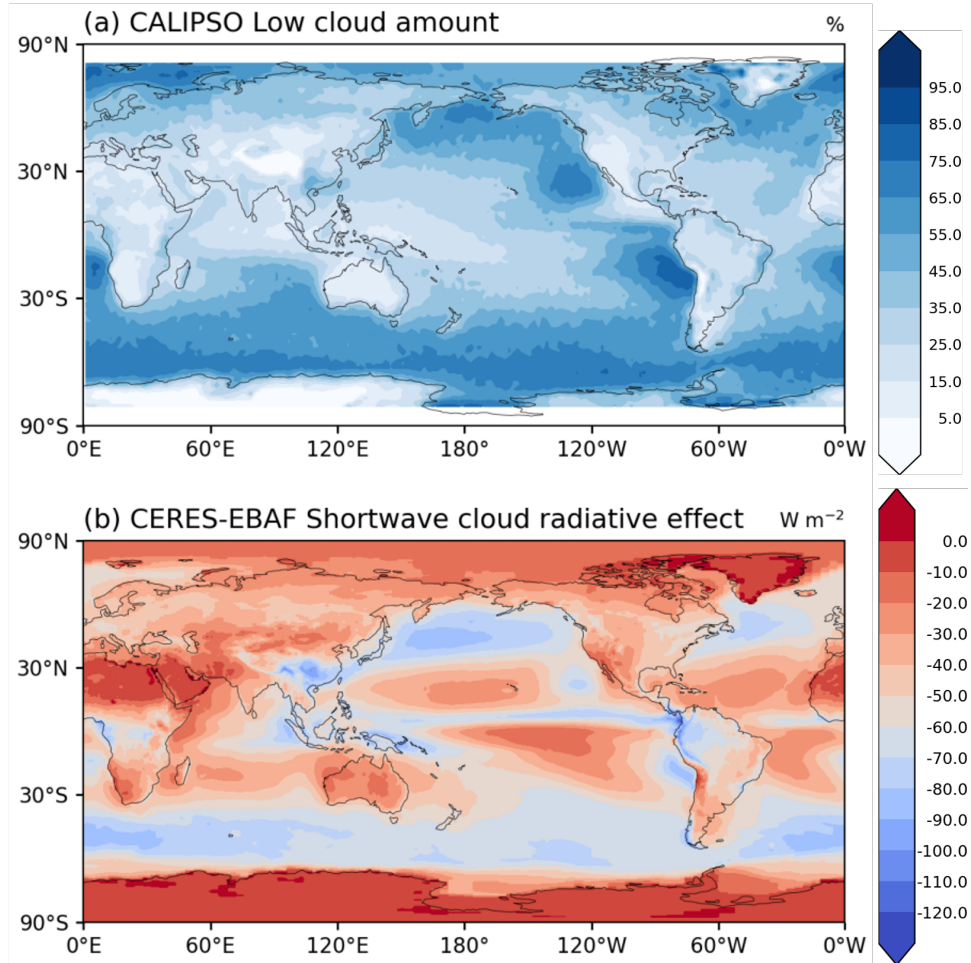
**Figure 3.** a) Low cloud amounts from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) from January 2007 to January 2010 and b) shortwave cloud radiative effect (SWCRE) from Clouds and Earth's Radiant Energy System (CERES) Energy Balance and Filled (CERES-EBAF) product from 2001 to 2018.

# Low level cloud amount
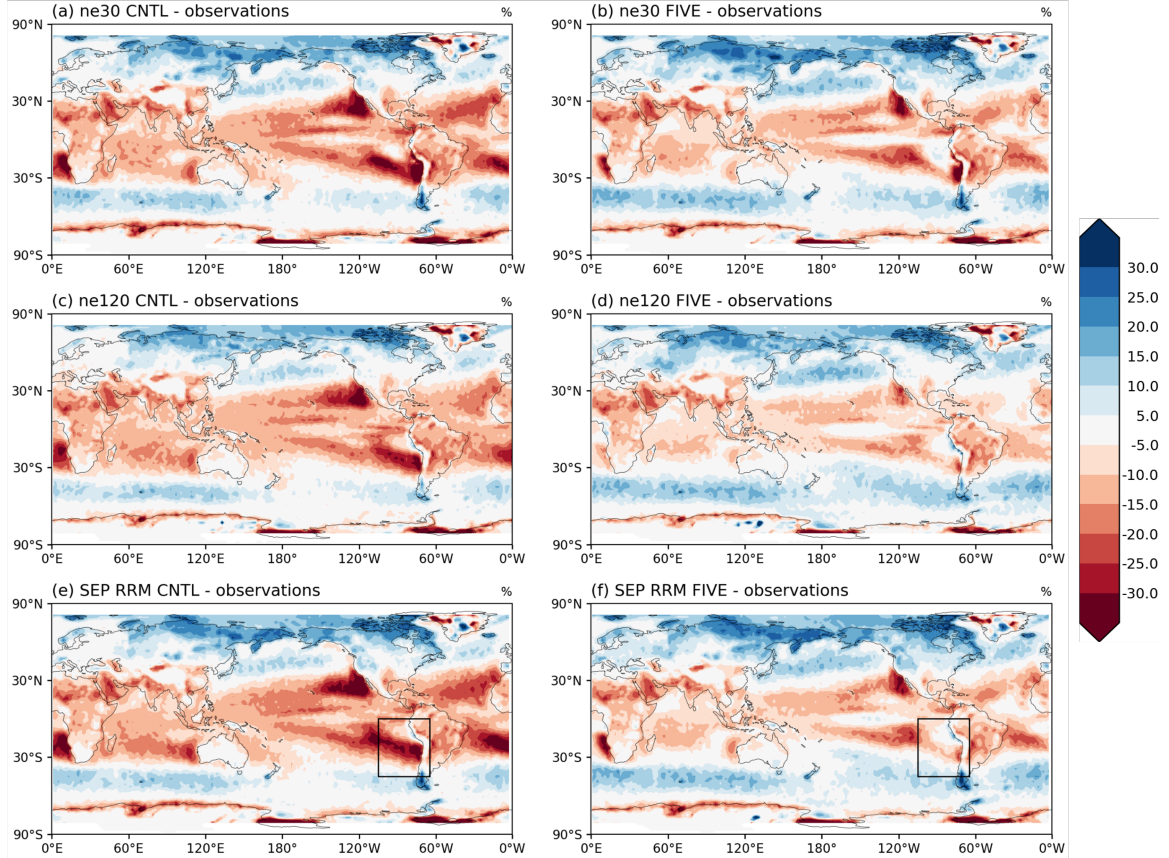
**Figure 4.** Annual low cloud amount biases computed relative to observations for a) ne30-CNTL, b) ne30-FIVE, c) ne120-CNTL, d) ne120-FIVE, e) SEP-RRM-CNTL, and f) SEP-RRM-FIVE. The box in figures e) and f) represents the area of the refined mesh in these simulations.
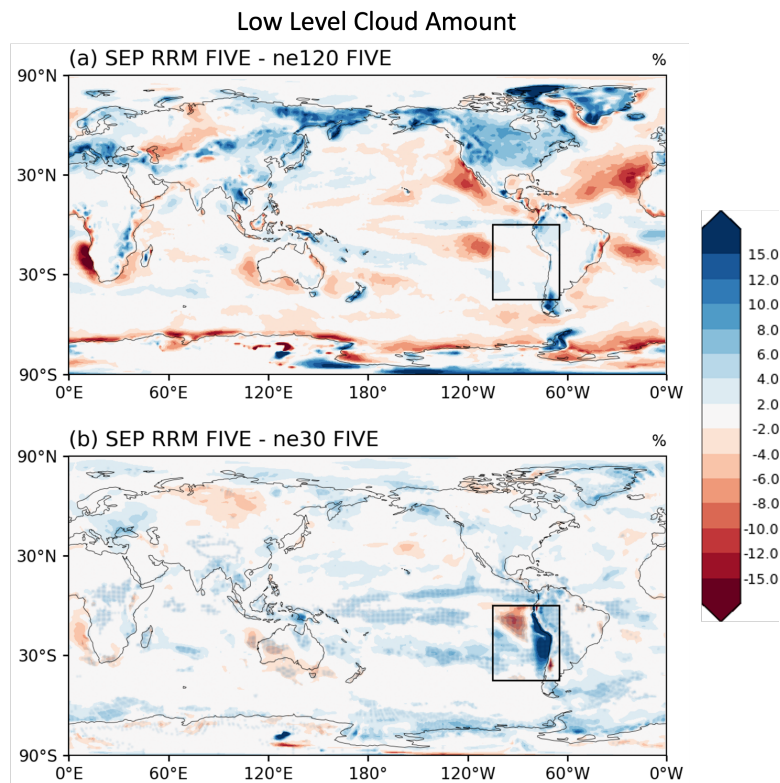
**Figure 5.** Differences of low cloud amount for the SEP-RRM-FIVE simulation calculated with respect to a) ne120-FIVE and b) ne30-FIVE simulations. The bounded box represents the area of the refined mesh in these simulations. In b) the the light blue stipple indicate locations where differences are statistically significant at the 95% confidence level.
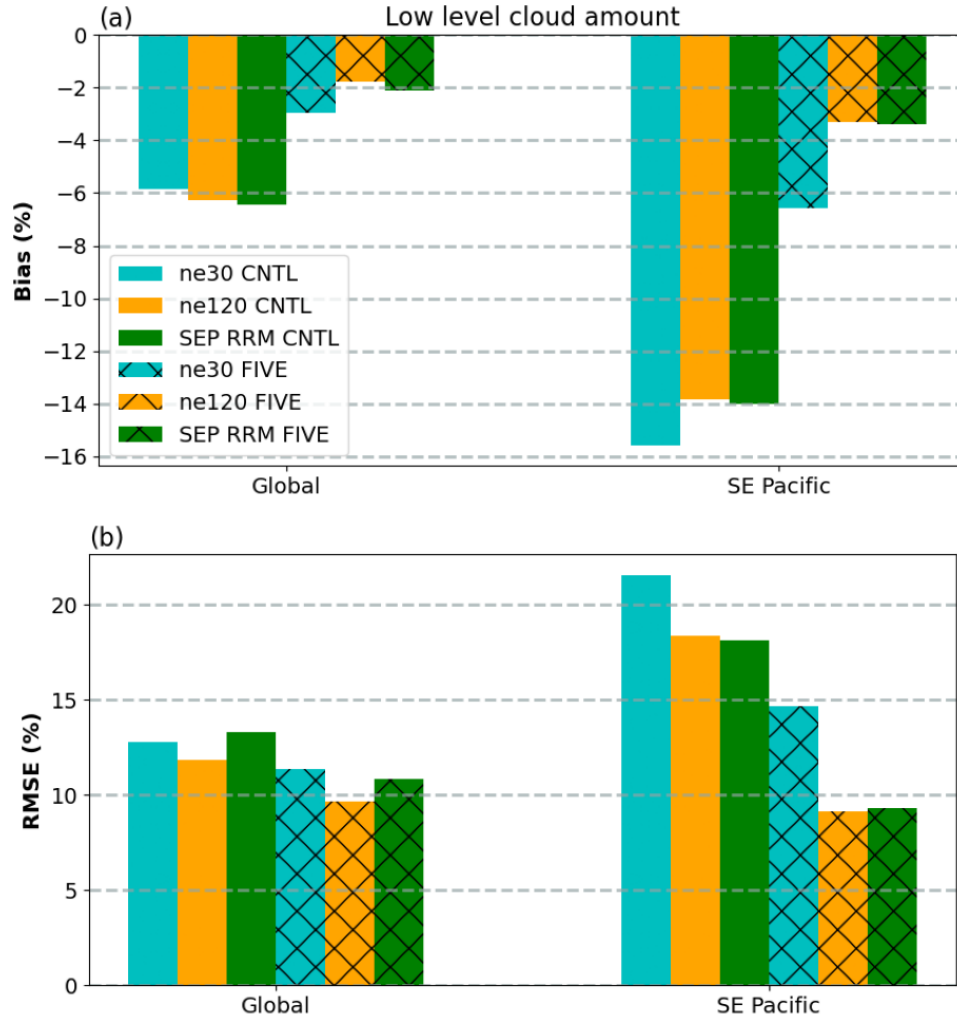
**Figure 6.** a) Bias and b) root mean squared error (RMSE) of low level cloud amounts for each simulation, computed relative to CALIPSO observations. Each skill metric is computed for the global domain (left) and the SE Pacific region (right) as bounded by the box in Figure 4e. We note that for skill scores computed for the SE Pacific region only ocean columns (which we define as columns having less than 0.5 land fraction) are considered to focus on scores for marine boundary layer clouds. Hatched bars represent simulations run with FIVE.

## Shortwave cloud radiative effect
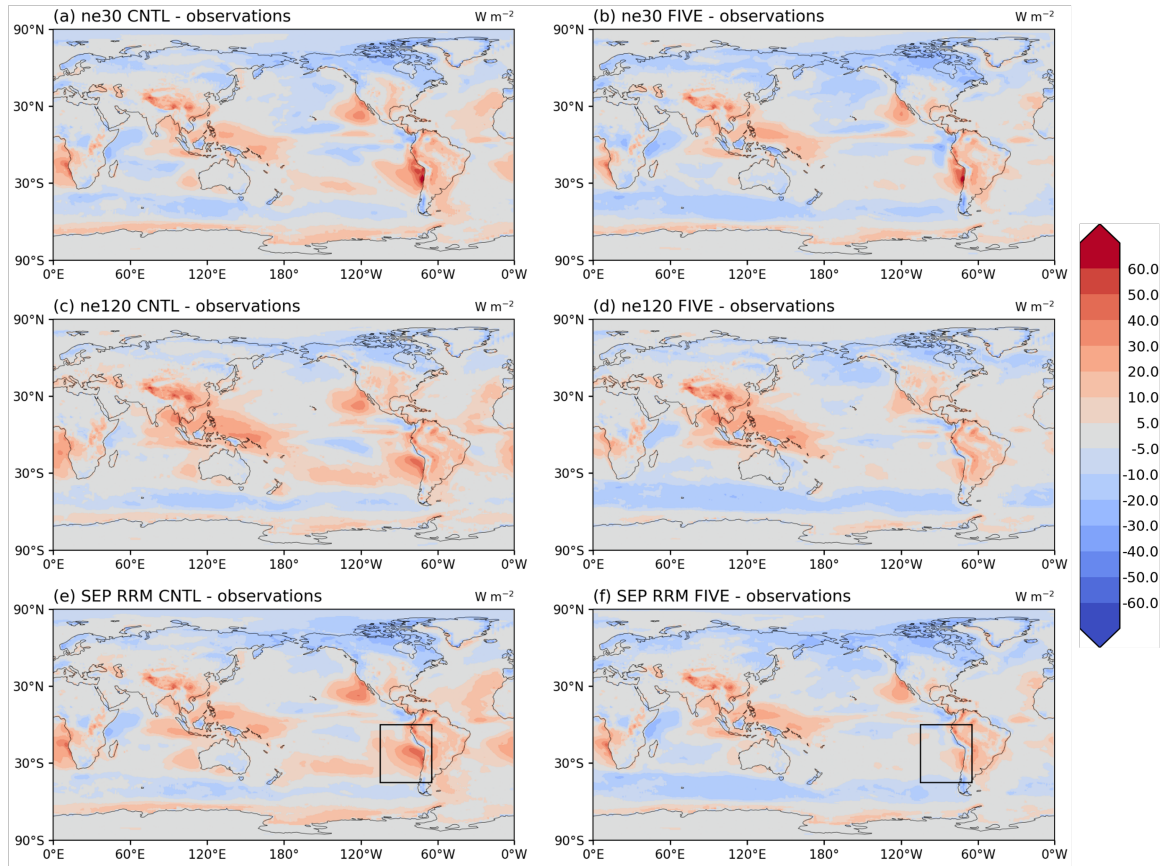
**Figure 7.** Same as Fig. 4 but for shortwave cloud radiative effect.
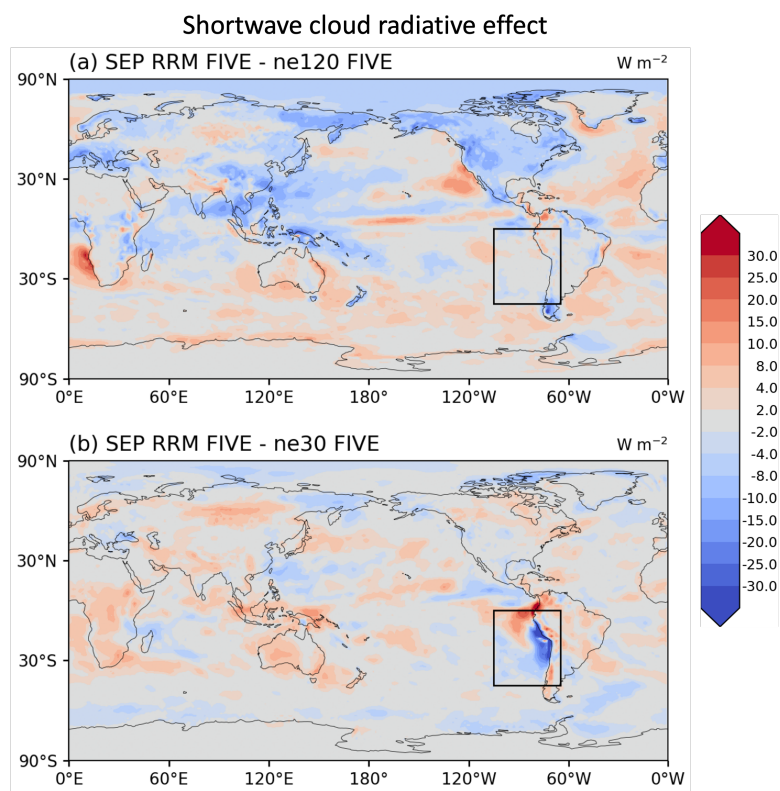
**Figure 8.** Same as Fig. 5 but for shortwave cloud radiative effect.
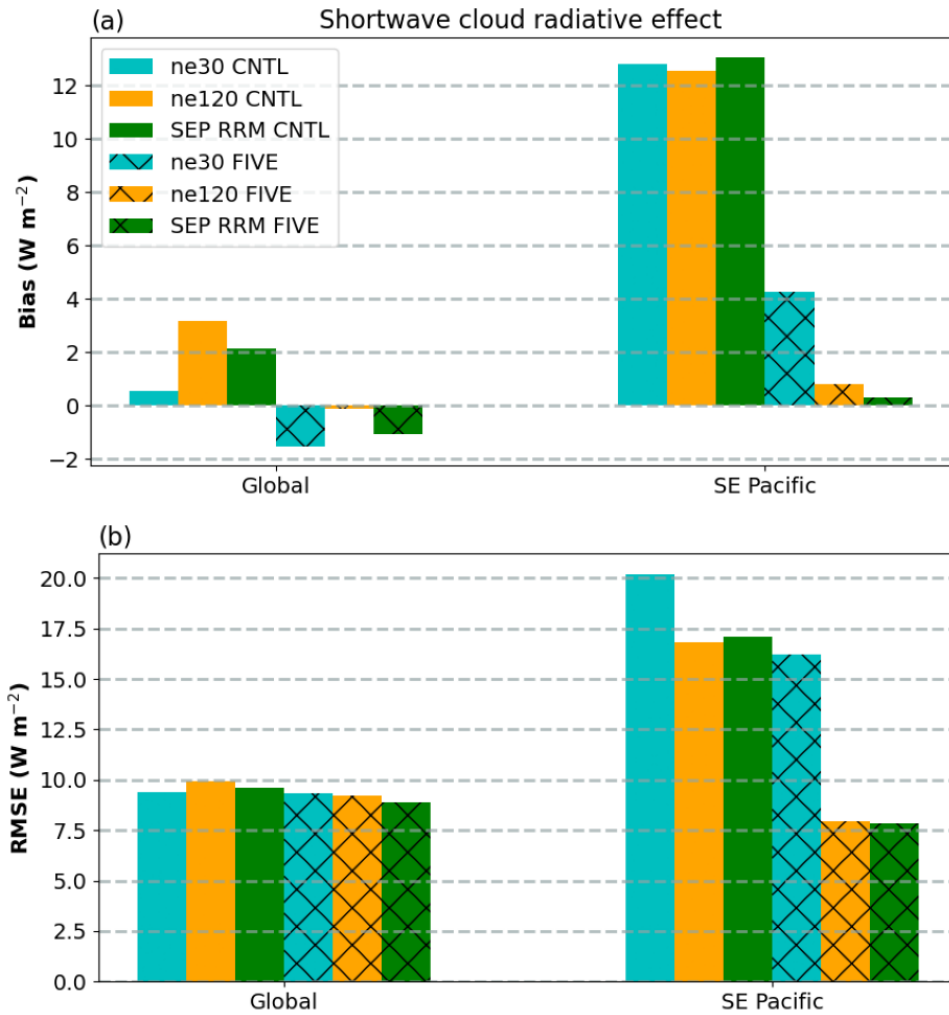
**Figure 9.** Same as Fig. 6 but for shortwave cloud radiative effect, computed relative to CERES-EBAF observations.
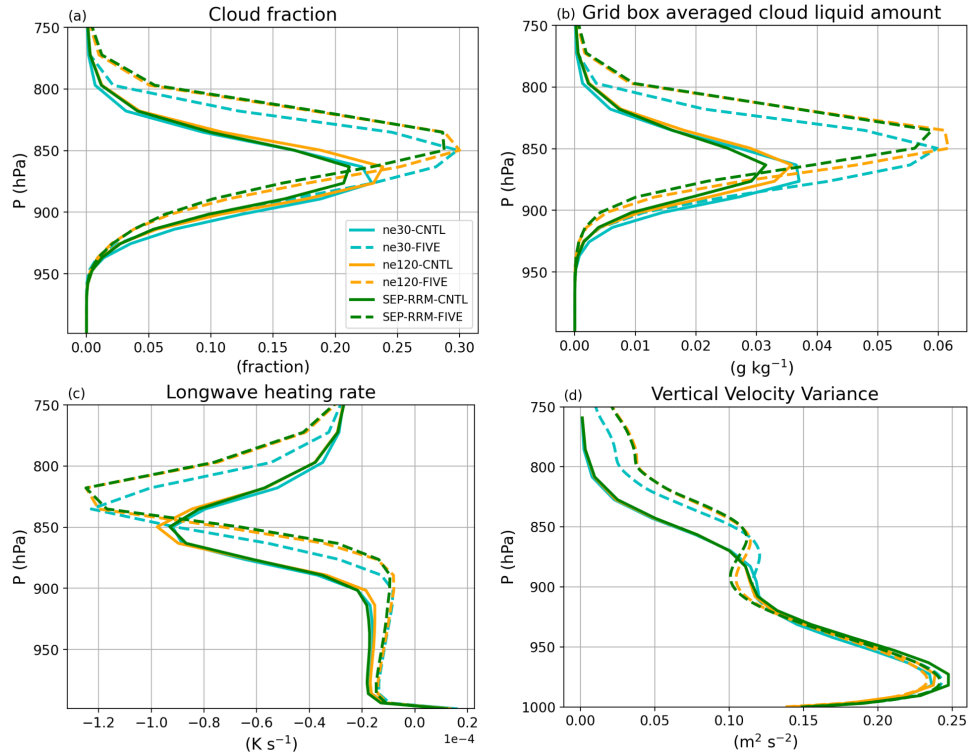
**Figure 10.** Profiles of a) cloud fraction, b) cloud liquid water, c) longwave heating rate, and d) vertical velocity variance predicted by CLUBB for each simulation averaged over five years. The profiles are spatially averaged over the domain used for regional refinement, illustrated in 1. Only oceanic columns are included in the averages, defined as those with land fraction less than 0.5.
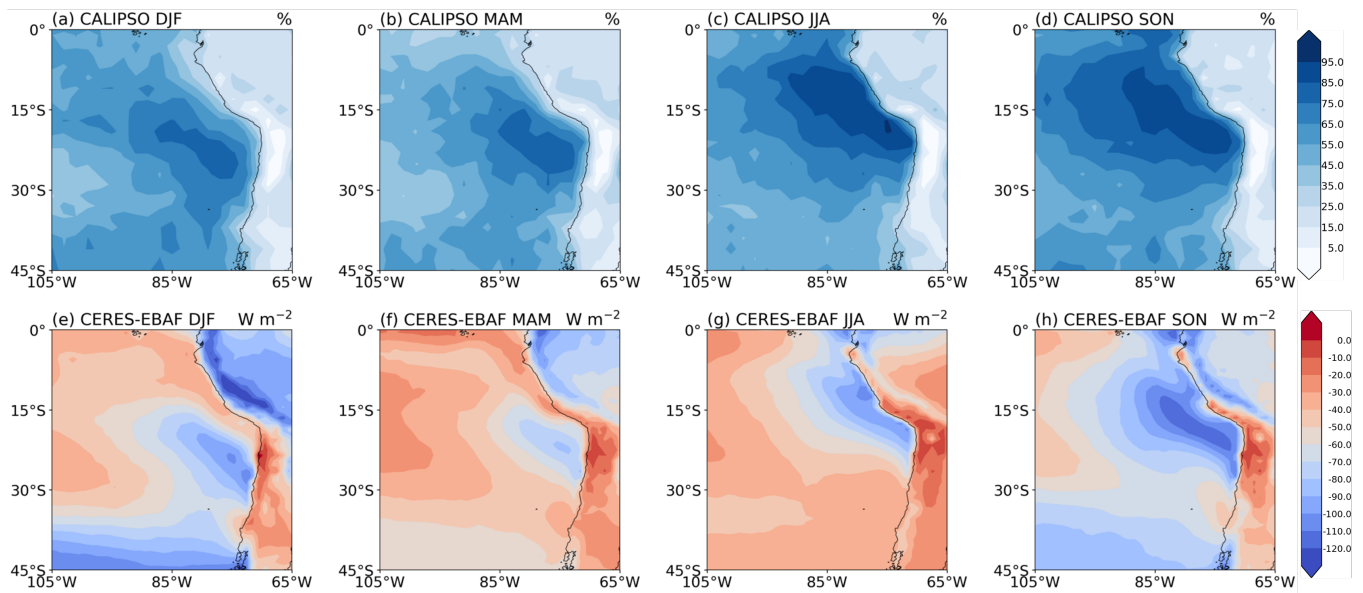
**Figure 11.** Observed low level cloud amount from CALIPSO (top row) and shortwave cloud radiative effect from CERES-EBAF (bottom row) averaged over December, January, February (DJF; first column); March, April, May (MAM; second column); June, July, August (JJA; third column); September, October, November (SON; fourth column).

# Low level cloud amount (%)
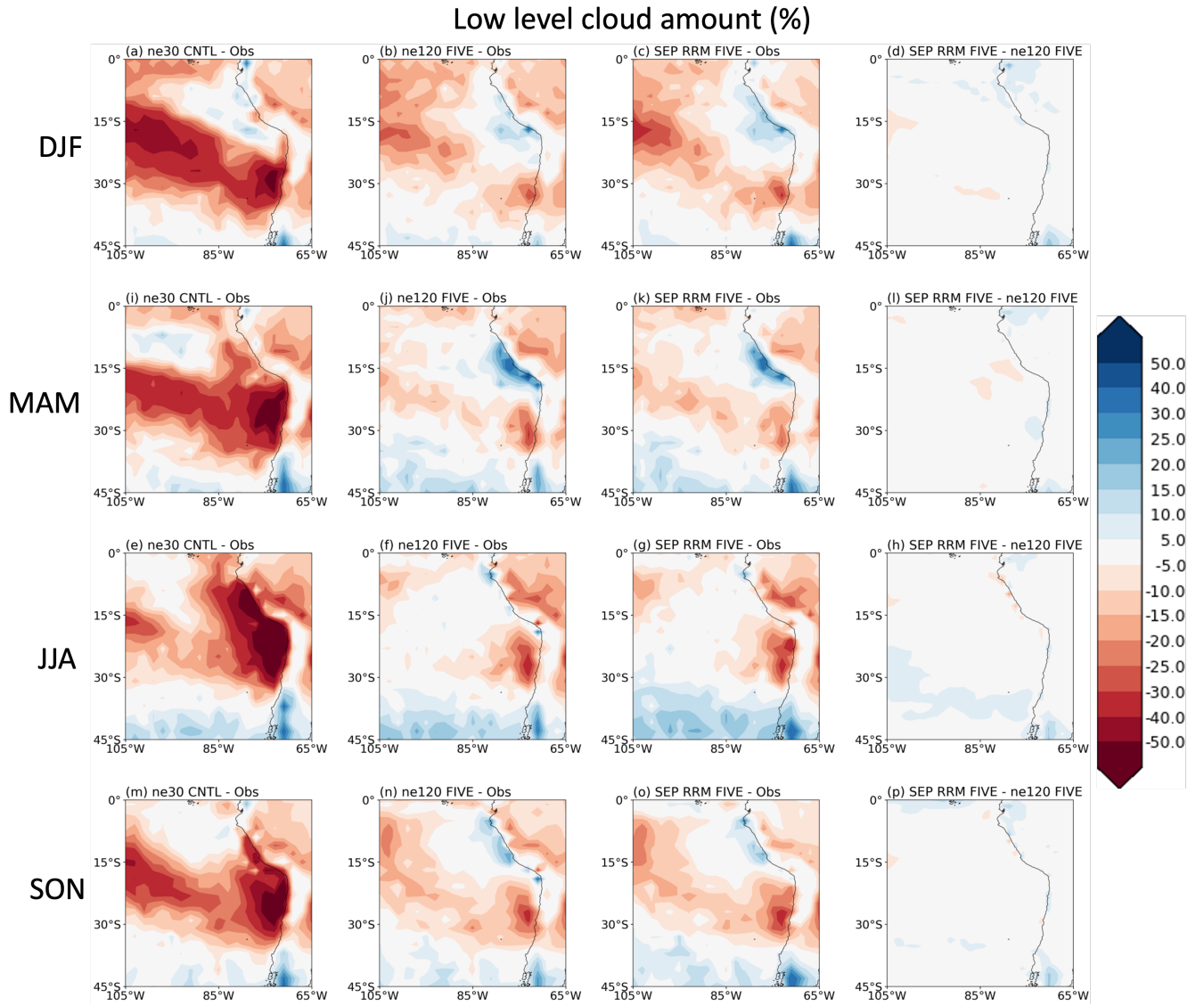
**Figure 12.** Differences of low level cloud amount computed relative to CALIPSO observations for ne30-CNTL (first column), ne120-FIVE (second column), and SEP-RRM-FIVE (third column). Also displayed are the differences of SEP-RRM-FIVE computed relative to ne120-FIVE (fourth column). Results are displayed for DJF (first row), MAM (second row), JJA (third row), and SON (fourth row) seasons.

# Shortwave cloud radiative effect (W m$^{-2}$)
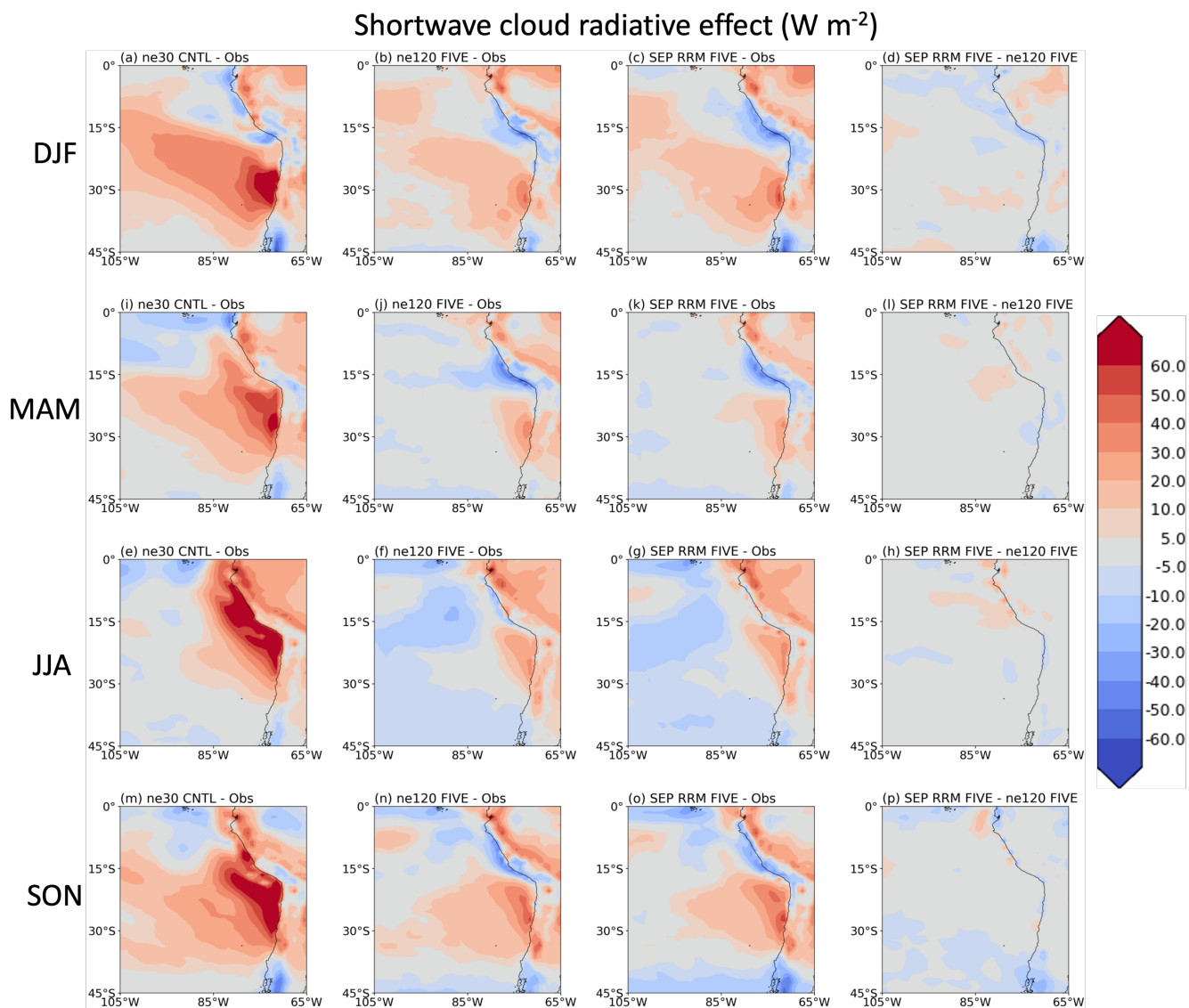
**Figure 13.** Same as Fig. 12 but for shortwave cloud radiative effect.
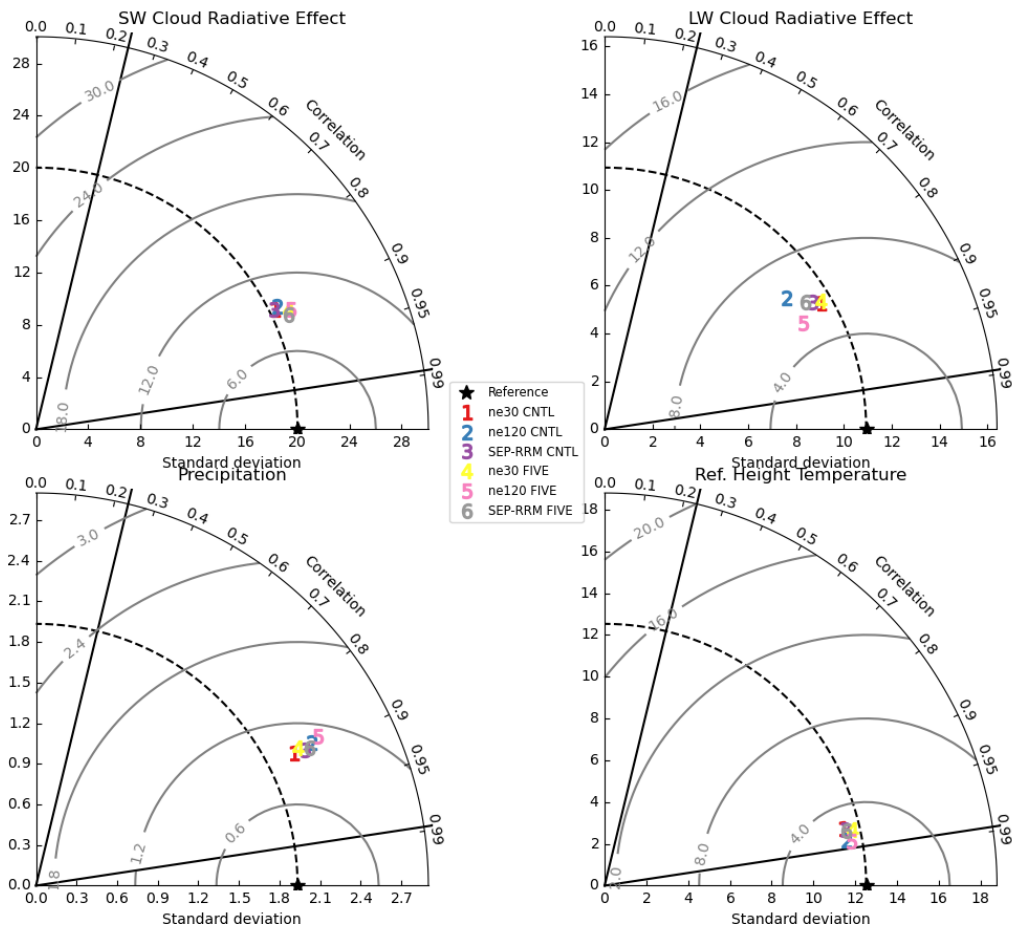
**Figure 14.** Taylor diagrams depicting the global skill for each simulation for five years for shortwave cloud radiative effect (SWCRE), longwave cloud radiative effect (LWCRE), precipitation, and reference height temperature.

**Table 1.** Description of the simulations performed.

| Simulation | Effective resolution | Number of elements | Number of vertical layers | Number of VEP layers |
|---|---|---|---|---|
| ne30-CNTL | 1° | 5400 | 72 | N/A |
| ne120-CNTL | 0.25° | 86400 | 72 | N/A |
| ne30-FIVE | 1° | 5400 | 72 | 212 |
| ne120-FIVE | 0.25° | 86400 | 72 | 212 |
| SEP-RRM | 1 to 0.25° | 8351 | 72 | N/A |
| SEP-RRM-FIVE | 1 to 0.25° | 8351 | 72 | 212 |

**Table 2.** Description of the time steps used in each simulation.

| Simulation | Physics Time step | Dynamics Lagrangian vertical time step | Dynamics horizontal time step | Explicit numerical diffusion time step |
|---|---|---|---|---|
| ne30-CNTL | 1800 | 900 | 300 | 100 |
| ne120-CNTL | 900 | 150 | 75 | 18.75 |
| ne30-FIVE | 1800 | 900 | 300 | 100 |
| ne120-FIVE | 900 | 150 | 75 | 18.75 |
| SEP-RRM | 900 | 150 | 75 | 18.75 |
| SEP-RRM-FIVE | 900 | 150 | 75 | 18.75 |

**Table 3.** Computational expense of each simulation as run on Livermore Computing (LC). SYPD stands for simulated years per day. Node hours is computed as nodes used multiplied by the number of hours required to simulate one year.

| Simulation | Nodes | SYPD | Node hours per year |
|---|---|---|---|
| ne30-CNTL | 29 | 4.3 | 162 |
| ne120-CNTL | 56 | 0.12 | 11200 |
| ne30-FIVE | 29 | 1.21 | 621 |
| ne120-FIVE | 56 | 0.06 | 22400 |
| SEP-RRM | 29 | 0.96 | 725 |
| SEP-RRM-FIVE | 29 | 0.55 | 1265 |

**Table 4.** Top of atmosphere radiation budgets for the simulations performed. All units in W/m$^2$.

| Simulation | RESTOM | SWCRE | LWCRE |
|---|---|---|---|
| ne30-CNTL | 2.7 | -44.7 | 22.8 |
| ne120-CNTL | 1.2 | -42.1 | 20.3 |
| ne30-FIVE | 1.1 | -45.8 | 23.0 |
| ne120-FIVE | 0.5 | -45.4 | 21.9 |
| SEP-RRM | 3.0 | -43.2 | 22.0 |
| SEP-RRM-FIVE | 0.9 | -45.5 | 22.8 |