AGU PUBLICATIONS

Journal of Advances in Modeling Earth Systems



RESEARCH ARTICLE

10.1002/2017MS000968

**Key Points:**
- Ultraparameterization (UP) is introduced as a global low-cloud-permitting GCM based on superparameterization (SP)
- Grid spacing of the embedded cloud-resolving models (CRM) is refined to explicitly capture the BL turbulence and the associated low clouds
- The world's first climate model with explicit embedded boundary layer turbulence is compared against satellite observations of stratocumulus and shallow cumulus clouds

**Correspondence to:**
H. Parishani,
h.parish@uci.edu

**Citation:**
Parishani, H., M. S. Pritchard, C. S. Bretherton, M. C. Wyant, and M. Khairoutdinov (2017), Toward low-cloud-permitting cloud superparameterization with explicit boundary layer turbulence, *J. Adv. Model. Earth Syst.*, *9*, 1542–1571, doi:10.1002/2017MS000968.

Received 20 MAR 2017
Accepted 19 MAY 2017
Accepted article online 19 JUN 2017
Published online 4 JUL 2017



# Toward low-cloud-permitting cloud superparameterization with explicit boundary layer turbulence

**Hossein Parishani[1], Michael S. Pritchard[1], Christopher S. Bretherton[2], Matthew C. Wyant[2], and Marat Khairoutdinov[3]**

[1]Department of Earth System Science, University of California, Irvine, California, USA, [2]Department of Atmospheric Sciences, University of Washington, Seattle, Washington, USA, [3]School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, New York, USA

**Abstract** Systematic biases in the representation of boundary layer (BL) clouds are a leading source of uncertainty in climate projections. A variation on superparameterization (SP) called "ultraparameterization" (UP) is developed, in which the grid spacing of the cloud-resolving models (CRMs) is fine enough (250 × 20 m) to explicitly capture the BL turbulence, associated clouds, and entrainment in a global climate model capable of multiyear simulations. UP is implemented within the Community Atmosphere Model using 2° resolution (∼14,000 embedded CRMs) with one-moment microphysics. By using a small domain and mean-state acceleration, UP is computationally feasible today and promising for exascale computers. Short-duration global UP hindcasts are compared with SP and satellite observations of top-of-atmosphere radiation and cloud vertical structure. The most encouraging improvement is a deeper BL and more realistic vertical structure of subtropical stratocumulus (Sc) clouds, due to stronger vertical eddy motions that promote entrainment. Results from 90 day integrations show climatological errors that are competitive with SP, with a significant improvement in the diurnal cycle of offshore Sc liquid water. Ongoing concerns with the current UP implementation include a dim bias for near-coastal Sc that also occurs less prominently in SP and a bright bias over tropical continental deep convection zones. Nevertheless, UP makes global eddy-permitting simulation a feasible and interesting alternative to conventionally parameterized GCMs or SP-GCMs with turbulence parameterizations for studying BL cloud-climate and cloud-aerosol feedback.

## 1. Introduction

Modern global climate models (GCMs) differ about how much the climate will warm in response to greenhouse gas increases. A leading source of uncertainty in climate sensitivity arises from the representation of low clouds [*Randall et al*., 2007; *Bony and Dufresne*, 2005] which are not resolved and depend on small-scale physical processes (e.g., entrainment, boundary layer (BL) turbulence, and sharp gradients at inversions) that are heavily parameterized [*Wood*, 2012]. The large spread of GCM-simulated low-cloud feedbacks on climate sensitivity has been a multidecade challenge [e.g., *Gettelman and Sherwood*, 2016]. A global simulation resolving boundary layer eddies and associated clouds, even though it would surely have its own uncertainties, biases, and limitations, would be a rich comparison for informing the development of conventionally parameterized climate models and possibly helping to constrain low-cloud feedbacks.

We explore a strategy to include explicit low-cloud physics using the multiscale modeling approach (also called superparameterization or SP for short) pioneered by *Grabowski* [2001] and *Khairoutdinov and Randall* [2001], which attempts to address shortcomings of conventional GCMs by embedding small cloud-resolving models (CRMs) in each GCM grid column. This method extends the range of scales that can be explicitly represented in climate models and consequently removes conventional parameterizations of deep moist convection, large-scale condensation, and boundary layer (BL) turbulence.

Implementations of SP have shown promising improvements in simulating several long-standing aspects of climate linked to deep convection such as (1) the diurnal cycle of precipitation [*Khairoutdinov et al*., 2005; *Pritchard and Somerville*, 2009a, 2009b; *Pritchard et al*., 2011], (2) the Madden-Julian Oscillation (MJO) and moist tropical Kelvin waves [*Khairoutdinov et al*., 2008; *Benedict and Randall*, 2009], and (3) more realistic precipitation intensity distribution and extremes [*Demott et al*., 2007; *Kooperman et al*., 2016]. Ongoing

challenges in SP models include (1) underprediction of low-cloud coverage and height and overprediction of cloud optical depth and drizzle in many regions [*Zhang et al*., 2008; *Marchand and Ackerman*, 2010], (2) overprediction of high clouds in some regions [*Marchand et al*., 2009; *Marchand and Ackerman*, 2010], and (3) seasonal mean precipitation biases comparable to those of the standard Community Atmosphere Model (CAM) [*Khairoutdinov et al*., 2005].

SP climate models have not been capable of properly resolving the turbulent eddies that form low clouds, due to grid resolution limitations. Typical SP implementations have used cloud-resolving models with 1–4 km horizontal resolution and a coarse vertical resolution encompassing 30–50 vertical layers. While this has produced promising effects for deep convection, it is known that accurate representation of cloud-top entrainment plays a crucial role in the realistic simulation of low clouds [*Bretherton et al*., 1999; *Stevens et al*., 2003; *Mellado*, 2017]. This requires extremely fine vertical grid spacing of $\mathcal{O}(5-25m)$ and horizontal grid spacing of $\mathcal{O}(5-100m)$ [*Stevens et al*., 2003]. SPGCMs are already $\sim$100X more computationally intensive than standard GCMs and adding this much additional resolution can seem prohibitive.

One pragmatic approach that is being pursued by several research groups is to develop advanced higher-order subgrid turbulence closure (HOC) parameterizations for the grid resolutions currently used in SPCAM [e.g., *Cheng and Xu*, 2011; *Xu and Cheng*, 2013a, 2013b; *Cheng and Xu*, 2013a, 2013b; *Bogenschutz and Krueger*, 2013; *Bogenschutz et al*., 2013; *Cheng and Xu*, 2015; *Painemal et al*., 2015; *Wang et al*., 2015] to handle the BL turbulence and clouds, while explicitly resolving the larger and more complex deep cumulus (Cu) clouds. This approach can improve the vertical structure of BL clouds [*Xu and Cheng*, 2013a] and may have potential for improving cloud radiative effect biases at top of atmosphere.

Drawbacks of the HOC approach are that it is quite complex and is sensitive to internal assumptions involving turbulence length scales, how turbulence is correlated with aerosol and precipitation, and vertical overlap of subgrid cloud in different model layers. Because the boundary layer turbulence is almost exclusively parameterized rather than being explicitly simulated, HOC is philosophically similar to the conventional parameterizations used in GCMs and has a similar set of uncertainties. Advances in both conventional BL and shallow convection parameterizations and HOC-enabled superparameterization have significantly improved low-cloud climatology in CAM/SPCAM [*Park and Bretherton*, 2009; *Kay et al*., 2012; *Bogenschutz et al*., 2013; *Cheng and Xu*, 2015] and other GCMs. However, consistently extending such approaches to precipitating clouds, aerosol-cloud interaction, and other realistic complexities of marine BL is mathematically cumbersome and computationally challenging.

These generalizations can be handled more naturally if BL eddies and associated cloud elements are explicitly simulated. Thus, our focus will be on achieving sufficient grid resolution within SPCAM to begin explicitly capturing the turbulent circulations that represent low clouds. While more computationally intense, this approach involves less assumptions and has potential to be more generally justifiable, especially for problems such as low-cloud feedbacks and cloud-aerosol interaction which require cloudy boundary layers to respond correctly to small perturbations in external forcing. It represents a previously unsampled complexity trade-off in global modeling, which is worthwhile to explore in its own right. *Grabowski* [2016] also speculated on the potential and value for a superparameterized GCM with sufficient embedded resolution to act as a global LES.

Refinements to the resolution of the embedded cloud-resolving models in SPCAM have been explored before, though not with a fully eddy-permitting grid resolution. Using SPCAM3.0 with 26 and 52 vertical levels and 64 column CRM arrays, *Marchand and Ackerman* [2010] investigated the cloud cover in a 1 km horizontal grid version of SPCAM; improved horizontal resolution decreased (degraded) low-cloud cover, while also increasing vertical resolution brought back similar low-cloud cover as in the default SPCAM and modestly increased (improved) cloud-top height.

Our approach, "ultraparameterization" (UP for short), is an ultra-high-resolution implementation of cloud superparameterization for global simulations of boundary layer and deep convective clouds at turbulent eddy-permitting grid resolutions. It uses 250 m horizontal resolution and 125 vertical levels with a grid spacing as fine as 20 m in each of the SPCAM's embedded CRMs. UP also uses the same fixed 125-level vertical grid in the host CAM so that there is no degradation of accuracy due to vertical interpolation between different CRM and host model grids. "UP" and "UPCAM," and similarly "SP" and "SPCAM" are used interchangeably throughout this paper.

In UP, the main purpose of ultrafine vertical resolution is to accurately simulate turbulence and entrainment processes near sharp temperature inversions that form atop stratocumulus (Sc) cloud layers. Large computational savings could be realized by using ultrafine resolution in each CRM only where it is needed. *Marchand and Ackerman* [2011] successfully implemented a vertically adaptive grid within a large-eddy simulation model for this purpose and showed it could reproduce most features of comparable fine-grid simulations for a variety of cloud-topped boundary layer benchmark cases. Similarly, *Grabowski* [2016] showed in idealized simulations of a 2-D mock-Hadley cell that a heterogeneous implementation of vertical resolution refinement can be coerced to simultaneously capture convincing deep convective systems over warm SSTs and Sc cloud layers over cold SSTs. While promising, these techniques have not been implemented or tested globally; a computational hurdle is developing a tailored load balancing that can dynamically allocate more parallel resources to locations experiencing high vertical refinement, a capability that the host GCM of SPCAM does not possess.

This paper outlines our UP implementation, its computational feasibility and scalability, and the results of hindcasts that reveal the benefits/consequences of increasing SPCAM's vertical versus horizontal resolution to this new limit. In section 2 the model detail and experiment design are presented. Global results of absorbed solar radiation and outgoing longwave radiation are discussed in section 3. Cloud development in subtropical stratocumulus and shallow cumulus regions are presented in sections 4 and 5. Results of a 3 month long UP simulation are discussed in section 6, followed by a discussion of some ongoing issues in section 7. The conclusions are summarized in section 8.

## 2. Methods

### 2.1. Model Description

We use the Super-Parameterized Community Earth System Model (SPCESM1.1) as a baseline model. The host GCM is CAM5.0 with a finite-volume dynamical core at $1.9^\circ \times 2.5^\circ$ horizontal resolution and a 5 min time step. The short GCM time step helps ensure a tight relationship between embedded cloud structures and their radiative impacts, which are calculated once per host GCM time step. The embedded CRMs are 2-D realizations of the System for Atmospheric Modeling (SAM) [*Khairoutdinov et al*., 2005], which replace standard CAM parameterizations for moist convection and large-scale condensation. The CRM is forced by the large-scale temperature and moisture tendencies arising from GCM-scale dynamical processes and runs continuously [*Khairoutdinov et al*., 2005; *Benedict and Randall*, 2009]. In our pilot implementation, the CRM uses a simplified one-moment microphysics described in *Khairoutdinov and Randall* [2003]. The model parameterizes subgrid-scale (SGS) transport by turbulence and subgrid total kinetic energy (TKE) using the Smagorinsky-type 1.5-order turbulence closure scheme. Radiative transfer is done on individual CRM columns using the CAM Radiative Transfer Model for GCMs (CAMRT).

This version of SPCAM5 is very similar to that described by *Wang et al*. [2011], except for the choice of simpler microphysics, made for reasons discussed in Appendix A. The model is available at https://svn-ccsm-models.cgd.ucar.edu/cam1/branches/UltraCAM-spcam2_0_cesm1_1_1 (revision 73250).

### 2.2. Target CRM Resolution: Considerations

The standard horizontal resolution (4 km) and number of vertical levels (30) in SPCESM are far too coarse to explicitly capture stratocumulus (Sc) eddy scales. For subtropical Sc, prior studies suggest a three-dimensional (3-D) large-eddy simulation (LES) with a 5 m vertical resolution, 25–50 m horizontal resolution, 64 or more grid columns in each horizontal direction, and an advanced scalar advection scheme is needed to simulate realistic turbulence statistics, entrainment rate, and cloud thickness [*Bretherton et al*., 1999; *Stevens et al*., 2005; *Van der Dussen et al*., 2013]. For trade cumulus, 40 m vertical and 100 m horizontal resolution appear adequate for results from differently formulated LES to be statistically similar [*Siebesma et al*., 2003]. However, use of a full LES in each SPCAM grid column is computationally infeasible for long simulations. Thus, we identified a CRM horizontal/vertical resolution representing a trade-off between acceptable resolution for cloud-topped BL simulation and acceptable computational expense.

Several past studies have experimented with coarser grids for eddy-permitting BL cloud simulations. One strategy exploits error trade-offs associated with horizontal versus vertical resolutions. Within some LES, including SAM, simulated Sc thickness increases with a coarser horizontal grid. This may be due to damping

of entrainment-stimulating turbulent velocity variations at the sharp inversion, which can compensate for numerical diffusion. *Berner et al*. [2011] obtained a larger and more realistic LWP for a Southeast Pacific Sc by using a horizontal grid spacing of 125 m compared to 50 m, for the same 5 m vertical resolution. *Wang and Feingold* [2009] used a similar strategy by employing a coarse vertical grid spacing of 30 m but compensating with an even coarser horizontal grid spacing of 350 m. *Cheng and Xu* [2008] found that simulations of GCSS shallow-Cu cases at 250 m horizontal resolution gave cloud statistics comparable to a high-resolution LES benchmark, while at coarser resolution a turbulence closure scheme was needed to obtain comparable results. *Cheng et al*. [2010] extended this to a large parameter space of horizontal and vertical resolutions over seven GCSS shallow cloud cases. For Sc, they confirm the counterbalancing effects on LWP of increasing horizontal and vertical grid spacing.

Together these results suggest that the horizontal grid spacing for UP should not exceed 250 m, which is roughly the width of an individual large BL or shallow cumulus updraft, that the vertical grid spacing should be as fine as can be afforded near Sc capping inversions, and that for any grid resolution, use of a three-dimensional domain and a state-of-the-art advection scheme helps prevent overentrainment through such inversions.

Another consideration is the minimum acceptable CRM domain size. There is a computational trade-off between the fine horizontal grid spacing required for a reasonable simulation of boundary layer clouds and a domain size large enough to reasonably simulate deep convective cloud statistics. While LES of stratocumulus have successfully employed domains as small as 2 km on a side [*Ackerman et al*., 2003], LES of shallow cumulus clouds typically use a domain size of 6–25 km so that at least one cumulus updraft is active throughout the simulation. Mesoscale organization of both stratocumulus and precipitating shallow cumulus is ubiquitous and affects the area-average cloud population [*Seifert et al*., 2015]. For deep convection, a small domain size is even more profoundly restrictive, since mesoscale organization and cold pools are central to its organization [*Tompkins*, 2001]. Nevertheless, *Pritchard et al*. [2014] showed that an eight-column 2-D CRM with 4 km grid spacing gave acceptable results in SPCAM for convectively-coupled waves and even global cloud climatology, except for excessive cloud liquid water and conditional instability in tropical deep convection regions. Ideally, one might choose a different CRM grid for different cloud regimes, but in this initial UP implementation we use a single CRM grid in all host-model grid columns.

### 2.3. Target CRM Resolution: Sensitivity Tests

With the above trade-offs and considerations in mind, we performed numerous sensitivity of benchmark cases to define a target grid for UP. We relied especially on two GEWEX Cloud System Study (GCSS)/Global Atmospheric System Study (GASS) LES intercomparison cases—the DYCOMS-RF01 nocturnal nonprecipitating marine Sc case [*Stevens et al*., 2005] and the BOMEX shallow cumulus case [*Siebesma et al*., 2003].

The DYCOMS simulations are particularly sensitive to CRM resolution, advection scheme and dimensionality, so they were a primary driver of our choices. We settled on a 125-level grid with horizontal grid spacing $\Delta x = 250$ m and a vertical grid spacing $\Delta z = 20$ m in the altitude range between 500 and 1800 m in which sharp inversions are most common, stretching the grid gradually below and above this level. We considered two horizontal geometries for UP, 32-column (C32) 2-D domains and an $8 \times 8$ (C8 $\times$ 8) 3-D domain. We also included simulations with the 30-level vertical grid of standard SP, which has $\Delta z \approx 200$ m below 2 km.

Figures 1a and 1b show the liquid water path (LWP) and cloud fraction for 6 h DYCOMS simulations with various grid configurations; aircraft observations of this case showed a quasi-steady Sc layer with 100% cloud cover and an LWP near 60 g/m$^2$ under a very sharp 7 K temperature inversion. The coarser-resolution simulations are compared to a 3-D "benchmark" with $192 \times 192$ horizontal columns, $\Delta x = \Delta y = 35$ m horizontal and inversion $\Delta z = 5$ m. For this configuration, we show simulations with both the default MPDATA (black solid line) and the more advanced (and somewhat more expensive) UM5 scalar advection scheme (black dashed) [*Yamaguchi et al*., 2011]. The UM5 simulation gives LWP comparable to the observations. The MPDATA simulation produces about 20% lower LWP than the UM5 case but maintains slightly higher cloud fraction.

As expected, increasing $\Delta z$ to 20 m (blue line) greatly decreases LWP and cloud fraction relative to the benchmark. Also, increasing $\Delta x$ to 250 m (magenta) increases the LWP and returns cloud fraction nearly to the benchmark values. Using a C8 $\times$ 8 grid (red) instead of C32 $\times$ 32 has little effect on the 250 m result.

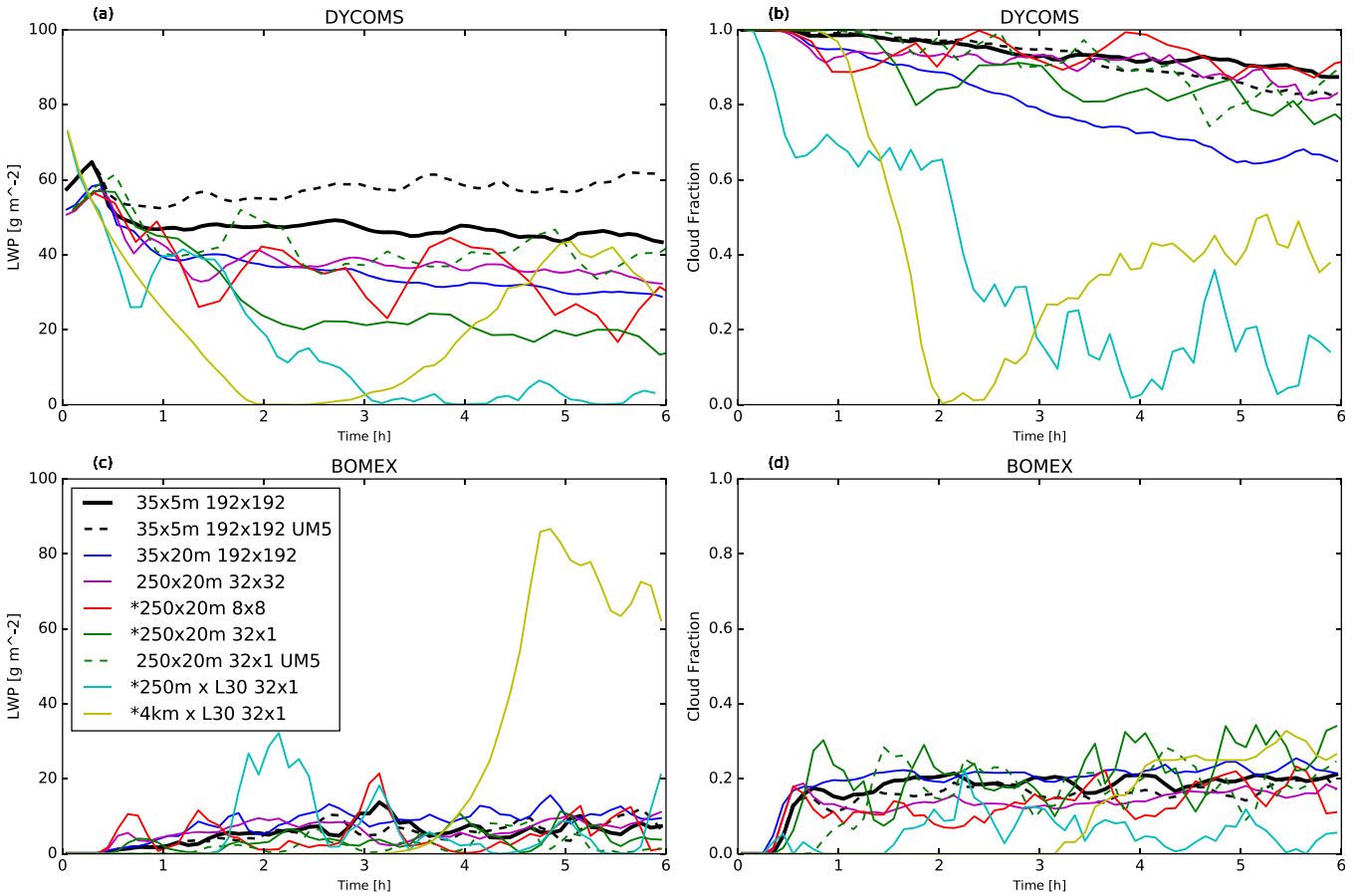


**Figure 1.** Sensitivity in offline SAM simulations of LWP and cloud fraction to grid resolution, configuration, and advection scheme for DYCOMS RF01 (a, b) and BOMEX trade cumulus (c, d). The legend gives horizontal × vertical grid resolution followed by the number of horizontal grid columns in the *x* and *y* directions. Simulations labelled with "UM5" used the UM5 advection scheme instead of the MPDATA scheme. Configurations with an asterisk are used in global UPCAM simulations presented in this study.

Use of 2-D C32 (green solid line) grids decrease the LWP and cloud fraction, especially with MPDATA advection.

Table 1 presents additional statistics for DYCOMS averaged over the final three hours of the simulations. The tabulated vertical velocity variance is the average of the peak of the 15 min mean vertical profiles. DYCOMS simulations with $\Delta x$=250 m resolution have less vertical velocity variance and generally less entrainment, reflecting their degraded resolution of boundary layer updrafts and downdrafts. Using SP-like 4 km horizontal resolution takes these trends to even more of an extreme. The relatively high tabulated entrainment in the DYCOMS 4 km case is due to an earlier collapse of the boundary layer then rapid entrainment to form a shallower boundary layer.

For the BOMEX case (Figures 1c and 1d and Table 1), the benchmark 35 × 5 m MPDATA case produces similar LWP and about 50% higher cloud fraction than the coarser 100 × 40 m simulations in *Siebesma et al.* [2003]. UM5 advection (black dashed line) produces no appreciable change in LWP but reduces the cloud fraction. A 3-D simulation with the UP candidate grid resolution of 250 ×20 m in 3-D has similar LWP and slightly smaller cloud fraction than the benchmark case. Of the two candidates, UP column configurations (C8 × 8 and C32 × 1), the C8 × 8 configuration provides the better match of BOMEX LWP and cloud fraction to the benchmark, with C32 × 1 producing lower cloud fraction. The peak vertical velocity variances in BOMEX (Table 1) for UP configurations are generally comparable to the benchmark case. The standard SP configuration has much higher LWP despite very small vertical velocity variance.

These tests confirm that a 250 × 20 m grid can perform acceptably across a range of subtropical cloud-topped boundary layer regimes. For UP, a CESM vertical grid was constructed with a 20 m vertical resolution within and below a height range typical of the marine inversion, transitioning to lower resolution in layers

**Table 1.** Hour 3–6 Statistics of Resolution and Configuration Sensitivity Tests Using Offline SAM for DYCOMS Stratocumulus and BOMEX Trade Cumulus Cases[a]

| | | | | DYCOMS | | | | BOMEX | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\Delta x$ [$m$] | $\Delta z$ [$m$] | Columns | Adv. | LWP [g m$^{-2}$] | CF | $\overline{w'w'}$ [m$^{-2}$ s$^{-2}$] | $w_e$ [mm s$^{-1}$] | LWP [g m$^{-2}$] | CF | $\overline{w'w'}$ [m$^{-2}$ s$^{-2}$] |
| +35 | 5 | 192 × 192 | MP | 45.6 | 0.91 | 0.36 | 4.3 | 7.1 | 0.19 | 0.16 |
| 35 | 5 | 192 × 192 | UM5 | 58.9 | 0.88 | 0.38 | 3.2 | 6.8 | 0.17 | 0.16 |
| 35 | 20 | 192 × 192 | MP | 31.3 | 0.70 | 0.29 | 3.8 | 10.2 | 0.22 | 0.18 |
| 250 | 20 | 32 × 32 | MP | 36.0 | 0.89 | 0.28 | 4.2 | 6.4 | 0.15 | 0.12 |
| *250 | 20 | 8 × 8 | MP | 31.4 | 0.92 | 0.24 | 4.6 | 6.1 | 0.16 | 0.13 |
| *250 | 20 | 32 × 1 | MP | 19.8 | 0.82 | 0.25 | 3.2 | 4.9 | 0.25 | 0.16 |
| 250 | 20 | 32 × 1 | UM5 | 39.6 | 0.86 | 0.18 | 2.1 | 3.0 | 0.19 | 0.18 |
| *250 | 200 | 32 × 1 | MP | 1.8 | 0.14 | 0.13 | 2.3 | 5.0 | 0.05 | 0.18 |
| *4000 | 200 | 32 × 1 | MP | 25.8 | 0.39 | 0.01 | 3.9 | 41.8 | 0.21 | 0.02 |

[a]"+" indicates the benchmark case and "*" indicates configurations of UPCAM global simulations in this study. "Adv." shows scalar advection scheme used, with "MP" indicating the MPDATA scheme.

above/below (as smoothly as possible in a log $(p)$ sense to avoid any numerical artifacts in GCM dynamics following *Williamson et al*. [1998]). Figure 2 compares 4 km wide segments of the 250 m 125-level UP grid (panel b) with the standard SPCAM 4 km 30-level grid (panel a).

For this UP grid, the C8 × 8 configuration best represents the benchmark LES of the three cases. However, it is more computationally expensive than a 2-D C32 × 1 configuration due to the larger column count and advection in the extra dimension. Thus, our preferred candidate UP configuration is 2-D with 32 columns, with the C8 × 8 3-D grid as an alternative if it significantly improves global performance. While UM5 advection would be a logical choice at the benchmark resolution, the MPDATA advection scheme was selected for the coarser C32 × 1 grid, for which it performed comparably well overall on the above test cases and in global hindcast tests (not shown here).

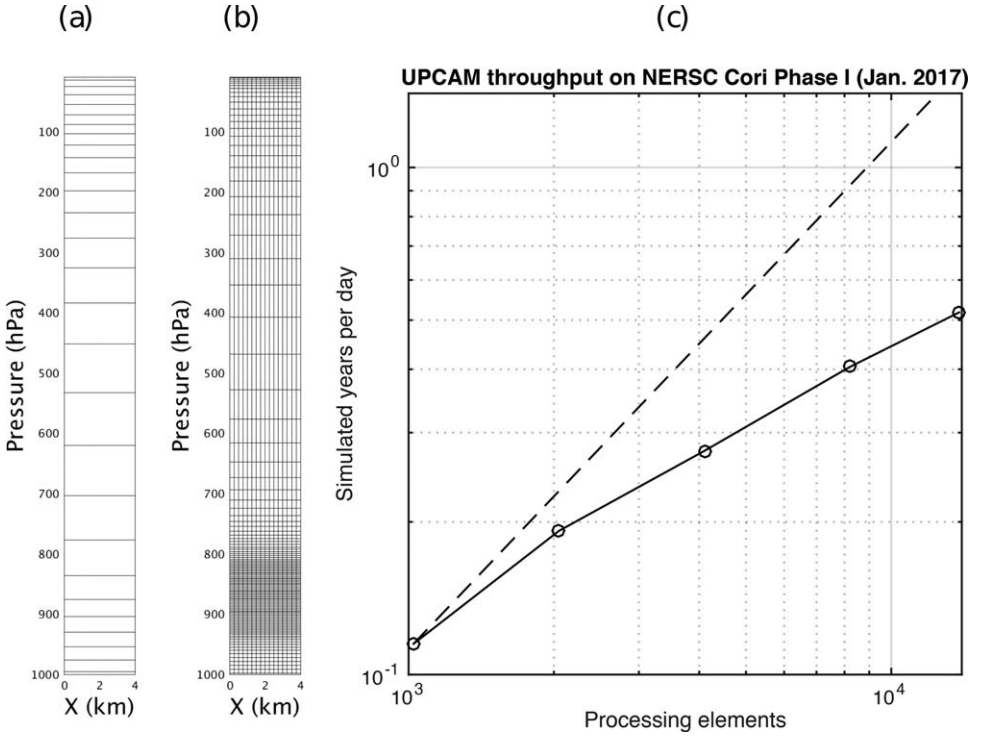


**Figure 2.** (a, b) A 4 km cut of the CRM grid structure for (a) standard superparameterization, i.e. L30-4km (with one column shown) and (b) ultraparameterization, i.e. L125-250m (with 16 columns shown). The horizontal bars are the model midlayers at which $u$, $v$, $T$, and $q$ are calculated. Both CRM and the host model share the vertical levels. (c) Scalability of the ultraparameterized CAM to its limit of 13,824 cores at 1.9°×2.5° horizontal resolution. Measured in January 2017 on the NERSC system Cori Phase I from a suite of 6 h pilot tests that included representative output. Dashed line is the ideal scaling (slope = 1).

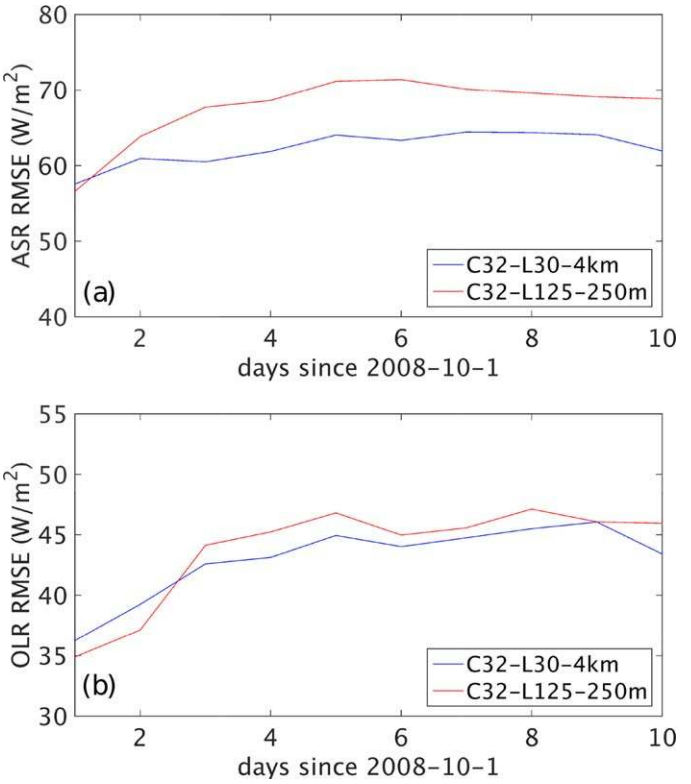


**Figure 3.** Unsteady evolution of root-mean-squared-error (RMSE) for daily maps of top-of-atmosphere global absorbed shortwave radiation (ASR) (a) and OLR (b) for the extended runs. RMSE is calculated relative to CERES-SYN 1-degree-daily satellite product. The first 12 h spin-up period is not shown.

### 2.4. Accommodating Computational Cost

Enhancing the resolution of SPCESM's CRMs is computationally costly. We performed a scalability analysis of the default SPCAM code [*Wang et al.*, 2011] and discovered that beyond 1000 cores there is a performance bottleneck, i.e., no speedup in model throughput with additional processors. This is not an expected feature of superparameterized simulations, which should be embarrassingly parallel [*Khairoutdinov et al.*, 2005]. We discovered that it is possible to bypass the bottleneck by deactivating the Explicit-Cloud-Parameterized-Pollutant (ECPP) scheme [*Gustafson et al.*, 2008; *Wang et al.*, 2011], which is still under development and has not yet been optimized for throughput; aerosols are prescribed instead. As shown in Figure 2c, this allows UPCAM to use over 13,000 cores. The model scales efficiently to its limit of 13,824 cores, which is set by the number of embedded CRMs at 1.9°×2.5° resolution, and can thus exploit ∼20% of a modern supercomputer at the National Energy Research Scientific Computing Center (NERSC).

Further 2–4X speed-up can be obtained through the CRM mean acceleration technique of *Jones et al.* [2015]. This algorithm is based on the fact that the horizontal mean state of limited-extent CRM domains evolves on a much longer time scale than the fast eddies that help regulate it. This scale separation is exploited to speed up SP calculations with insignificant loss of accuracy. The maximum applicable acceleration depends on the cloud regime being simulated, the rapidity of forcing changes, etc., so it must be chosen empirically via sensitivity tests. In this paper, we use 2X acceleration for the 36 h ensemble runs and 4X acceleration for the extended runs of Figure 3 and the 90 day runs of section 6.

### 2.5. Ensemble 36 h Hindcast Experiment Design

We will look at global maps of top-of-atmosphere radiation biases and other satellite-derived cloud products as measures of skill in simulating different cloud regimes. We are especially concerned with the subtropical shallow cumulus and stratocumulus regimes, which inspired our CRM grid design, but it is also important to simulate other regimes well, including deep cumulus convection and extratropical cloud systems.

We use a 10-member ensemble of 36 h hindcasts from different starting dates to study the skill of low-cloud simulations in different grid setups. This allows extensive use of daily and higher-frequency global satellite-derived cloud and radiation observations for validation (e.g., the CERES-Syn estimates of daily-mean top-of-atmosphere absorbed shortwave radiation (ASR) and outgoing longwave radiation (OLR) on a 1° lat-lon grid), in a context in which the simulated and actual large-scale atmospheric state are still very similar. This is particularly suitable for cloud-related biases, which develop very quickly into patterns similar to those in a free run of the model.

For the single-member extended hindcast simulations discussed in section 6, Figure 3a shows that shortwave biases with standard SP (blue line) and UP (red line) rapidly attain 80% of their asymptotic value of

**Table 2.** A Summary of the Simulations Performed in this Study[a]

| | Simulation ID | CRM Levels | CRM Columns | CRM *dx* [m] | CRM Extent [km] | CRM *dt* [s] |
|---|---|---|---|---|---|---|
| 2-D CRM | C32-L30–4 km (SP) | 30 | 32 × 1 | 4000 | 128 | 20 |
| | C32-L30–250 m | 30 | 32 × 1 | 250 | 8 | 3 |
| | C32-L125–4 km | 125 | 32 × 1 | 4000 | 128 | 3 |
| | C32-L125–250 m (UP) | 125 | 32 × 1 | 250 | 8 | 1 |
| 3-D CRM | C8 × 8-L125–250 m (3-D UP) | 125 | 8 × 8 | 250 | 2 × 2 | 1 |

[a]Each case consists of 10 ensemble members starting at different calendar dates of October 2008. *dx* and *dt* are the CRM horizontal grid spacing and time step size, respectively. For the 3-D CRM, *dx* = *dy*. The exterior model shares the vertical grid with the CRM.

~60–70 W/m$^2$ within the first day. In contrast, the slow portion of error growth associated with decorrelation of the initialized flow only contributes 10 W/m$^2$ of this total TOA shortwave error. For each hindcast, the CERES observations are interpolated to the model grid and differenced with the simulated fields to obtain an area-weighted global root-mean square error (RMSE) and global-mean bias. In the longwave, asymptotic errors are lower and the fraction contributed by the 1 day timescale is smaller. That is, longwave radiative biases are smaller than shortwave biases in SPCAM, and take longer to spin up because they are due to deep convective clouds that interact strongly with developing circulation biases. During the first hindcast day, the large-scale atmospheric state in which the clouds develop is most accurate. However, for the first several hours, the CRM clouds and turbulence must spin up, and the GCM undergoes an initialization shock while it adapts the ECMWF global initial state to be compatible with its internal dynamics and physics. We find that the SPCAM cloud field has largely equilibrated from this spin-up phase within 12 h. Thus, we analyze 12–36 h hindcasts for most of our model development and analysis.

The next sections of the paper test and compare candidate choices of CRM grid in the superparameterized CAM5. We use 10-member ensemble hindcast simulations spanning 1–29 October 2008 (with 3 day separation between ensemble members). Namely, the initialization dates are 1st, 4th, 7th, 10th, 13th, 16th, 19th, 22nd, 25th, and 28th of October 2008. Simulations are initialized at 12 UTC from high-resolution ECMWF observational/reanalysis boundary conditions available through the Year of Tropical Convection (YOTC) [*Waliser et al*., 2012]. Sea surface temperatures are prescribed based on daily ECMWF data. Climatological aerosol concentrations varying in space and time are prescribed as per *Lamarque et al*. [2010]. However, in our simulations aerosols are not permitted to induce indirect cloud radiative effects; rather cloud droplet sizes are constrained to ramp between polluted land values of 8 μm and pristine ocean values of 14 μm from the perspective of the radiation scheme.

Table 2 summarizes the CRM configurations. We separately vary the CRM horizontal resolution (the SP default of 4 km versus our UP default of 250 m) and the vertical resolution (the SP default of L30 versus our UP default of L125), for a set of four permutations, using 32-column (C32) CRMs. For the UP resolution, we also show results from a small 3-D 8 × 8 CRM domain (C8 × 8), following *Khairoutdinov et al*. [2005]. We also repeat the UP configuration without the use of the mean state CRM acceleration to demonstrate that the algorithm does not impact essential model behavior.

## 3. Global Hindcast Results

Figures 4 and 5 show maps of the ensemble mean biases of absorbed shortwave radiation (ASR) and outgoing longwave radiation (OLR) for forecast hours 12–36 (i.e., 0–24 UTC of the target days in October 2008) for five selected model configurations. In this figure, and many that follow, panel a shows standard SP, panel d shows standard UP, and panel e shows UP with 8 × 8 CRM columns. Panels b and c show the intermediate configurations between SP and UP. Table 3 lists the global mean and RMSE summary statistics of the OLR and ASR biases for these configurations.

There is considerable similarity between the spatial patterns of ASR and OLR error between the configurations. This is especially true of small-scale ASR and OLR features in midlatitudes and of tropical OLR features between the L30 configurations (a and b), or between the L125 configurations (c–e), suggesting these features are connected to vertical velocity errors common to all simulations.

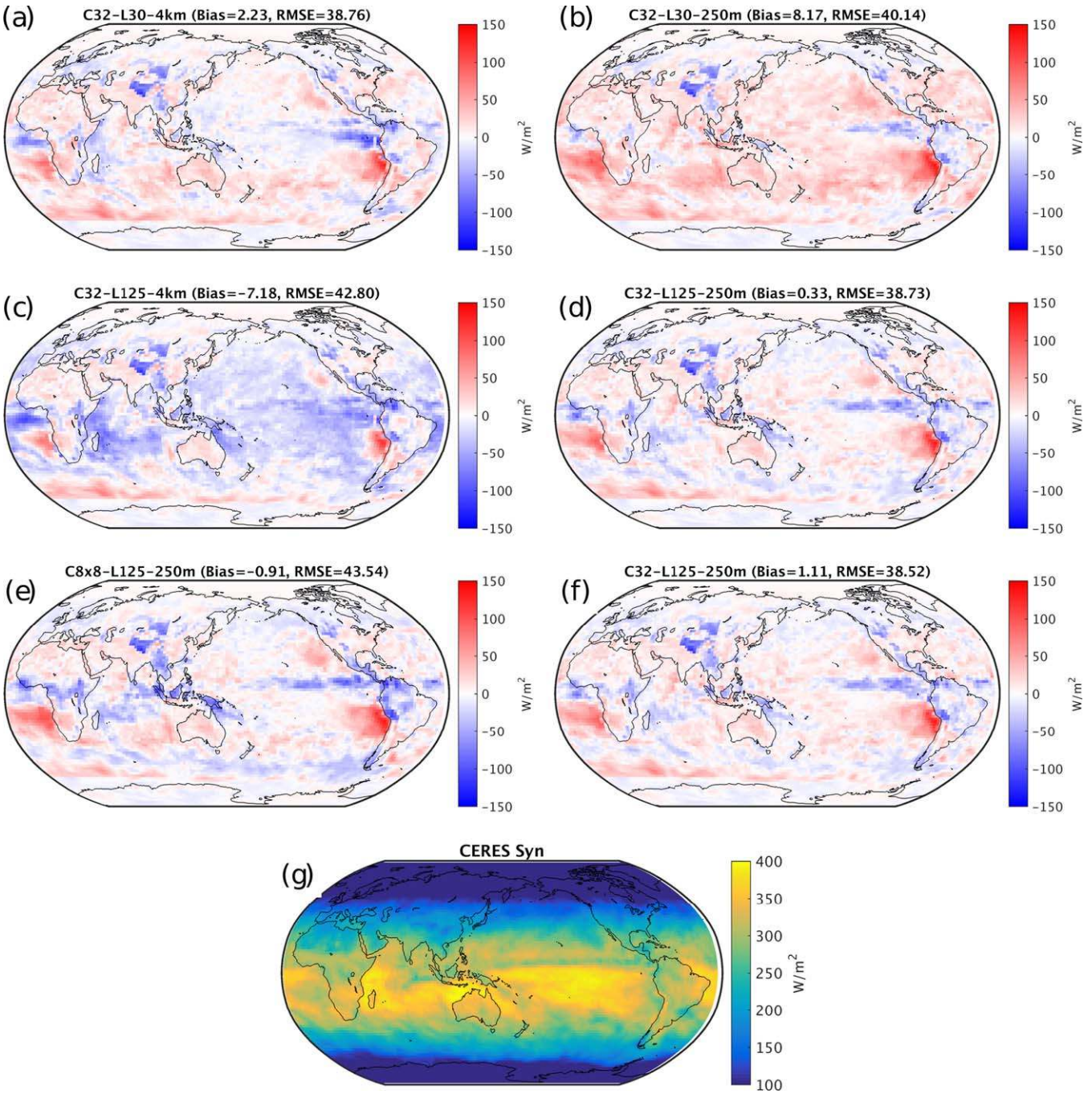


**Figure 4.** Day 1 ensemble mean ASR bias of the model: (a) C32-L30–4 km (SP), (b) C32-L30–250 m, (c) C32-L125–4 km, (d) C32-L125–250 m (UP), (e) C8 × 8-L125–250 m, and (f) C32-L125–250 m (UP) but without 2X CRM acceleration, relative to (g) the 10 day mean CERES-Syn observation of ASR.

All configurations have much smaller longwave than shortwave errors. The UP C32 configuration has marginally larger OLR RMSE (22.1 W/m$^2$) than default SP (20.9 W/m$^2$) but a lower global mean bias (2.1 W/m$^2$, compared to standard SP's 4.7 W/m$^2$). UP tends to systematically reduce OLR biases in subtropical low-cloud regions while amplifying positive OLR biases in regions of deep tropical convection. Since a C32 UP domain with $\Delta x = 250$ m is only 8 km wide, it is unsurprising that UP distorts cloud formation by deep convection, compromising OLR from the simulated upper level cloud.

The ASR errors and global biases discriminate more between the model configurations. Table 3 shows that standard SP and UP have the smallest ASR RMSE and that global mean ASR bias is smaller for UP (0.3 W/m$^2$)

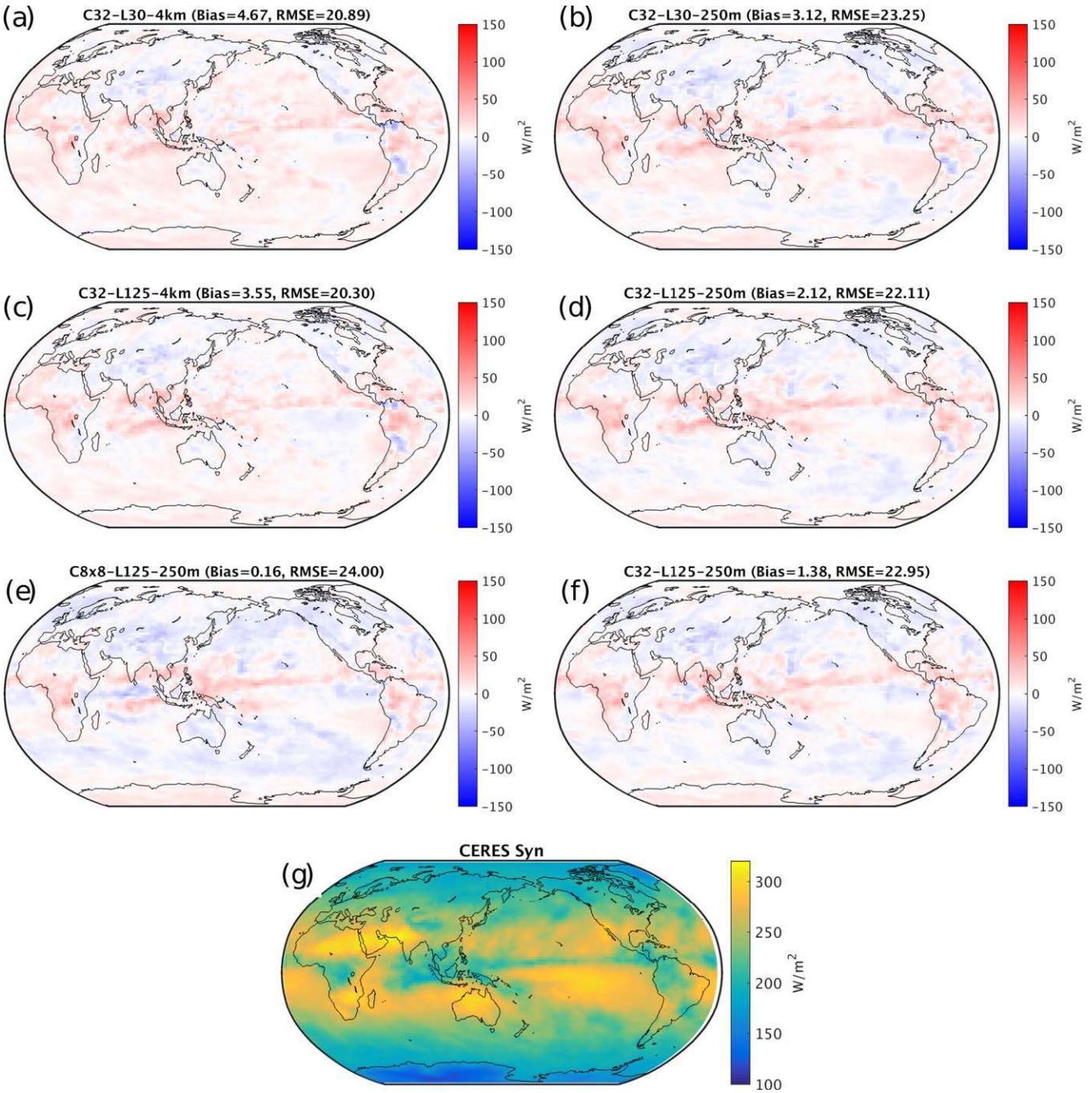


**Figure 5.** Day 1 ensemble mean OLR bias of the model: (a) C32-L30–4 km (SP), (b) C32-L30–250 m, (c) C32-L125–4 km, (d) C32-L125–250 m (UP), (e) C8 × 8-L125–250 m, and (f) C32-L125–250 m (UP) but without 2X CRM acceleration, relative to (g) the 10 day mean CERES-Syn observation of OLR.

than for SP (2.2 W/m$^2$), primarily due to reduced dim biases over the Southern Indian, South Pacific, and Southern Oceans. Increasing vertical resolution from L30 to L125 makes global-mean ASR substantially more negative. The effect of increasing horizontal resolution from 4 km to 250 m is roughly the reverse. Both of these trends were also noted by *Marchand and Ackerman* [2010]. That is, carefully balancing horizontal and vertical grid refinement is important to retaining global radiation balance in an UP-like model. The C32 configuration of UP has generally less RMSE than the C8 × 8 configuration.

We stress that the UP configurations have not been retuned in any way from standard SP by changing adjustable constants within the CRM microphysics such as autoconversion efficiency from cloud to rain, ice

**Table 3.** Global Mean, RMSE, and Standard Deviation (r) Biases of the Simulations for the 10 Ensemble Members Spanning October 2008 in Units of $W/m^2$[a]

| Simulation ID | OLR | | | ASR | | |
|---|---|---|---|---|---|---|
| | Mean | RMSE | $\sigma$ | Mean | RMSE | $\sigma$ |
| C32-L30–4 km (SP) | 4.67 | 20.89 | 16.44 | 2.23 | 38.76 | 32.38 |
| C32-L30–250 m | 3.12 | 23.25 | 18.63 | 8.17 | 40.14 | 32.17 |
| C32-L125–4 km | 3.55 | 20.30 | 16.18 | −7.18 | 42.80 | 34.11 |
| C32-L125–250 m (UP) | 2.12 | 22.11 | 17.97 | 0.33 | 38.73 | 31.66 |
| C8 × 8-L125–250 m (3-D UP) | 0.16 | 24.00 | 19.67 | −0.91 | 43.54 | 35.08 |

[a]OLR and ASR biases are relative to CERES-SYN.

fall speed, etc., or by fixing known issues with UP discussed in Appendix B, which might reduce their errors. There are also robust biases across SP-like and UP-like model configurations, including too negative (strong) ASR in the Atlantic and East Pacific ITCZ, and anomalously positive (weak) ASR over the Peruvian and Namibian subtropical Sc regions. Compared to SP, C32 UP gives a particularly unbiased simulation of ASR in the trade cumulus regimes over the south-central and south-eastern Pacific and the southeastern Indian Ocean. UP also reduces the positive ASR biases over the high latitude Southern Ocean. In the L30–250 m configuration (Figure 4b) the dim-trade bias in SP is exaggerated while at L125–4 km (Figure 4c) a bright-trade bias dominates instead. The C8 × 8 UP configuration produces similar errors to its C32 analogue over the subtropical Sc regions but worsens the bright (negative) ASR bias in the Atlantic and East Pacific ITCZs, as noted by *Pritchard et al.* [2014].

## 4. Analysis of Subtropical Stratocumulus Regions

This section focuses on evolution of the boundary layer vertical structure and the associated cloud fields in the stratocumulus regions off the coasts of Peru and Namibia. One might hope that ultraparameterization would have a large benefit in these regions, because its fine vertical and horizontal grid can better resolve the interaction of BL eddies with the sharp capping inversion that typically overlays the boundary layer clouds. This interaction is crucial for entrainment, which is a key process regulating Sc clouds. The previous section showed that UP does not significantly improve on the SP TOA radiation fields in the subtropical Sc regions across the Sc-Cu transition, but we will see in this section that the vertical cloud and BL structure is considerably improved with UP. Both conclusions are consistent with *Marchand and Ackerman* [2010]'s study of the sensitivity of SP to use of 1 km resolution and 52 vertical levels.

### 4.1. Improved Vertical Cloud Profile in Subtropical Sc Regions

Figure 6 compares the vertical structure of subtropical Sc in UPCAM (red) and SPCAM (blue) against daily mean swath-level data from NASAs A-Train satellites (C3M product, produced by a radar and lidar looking

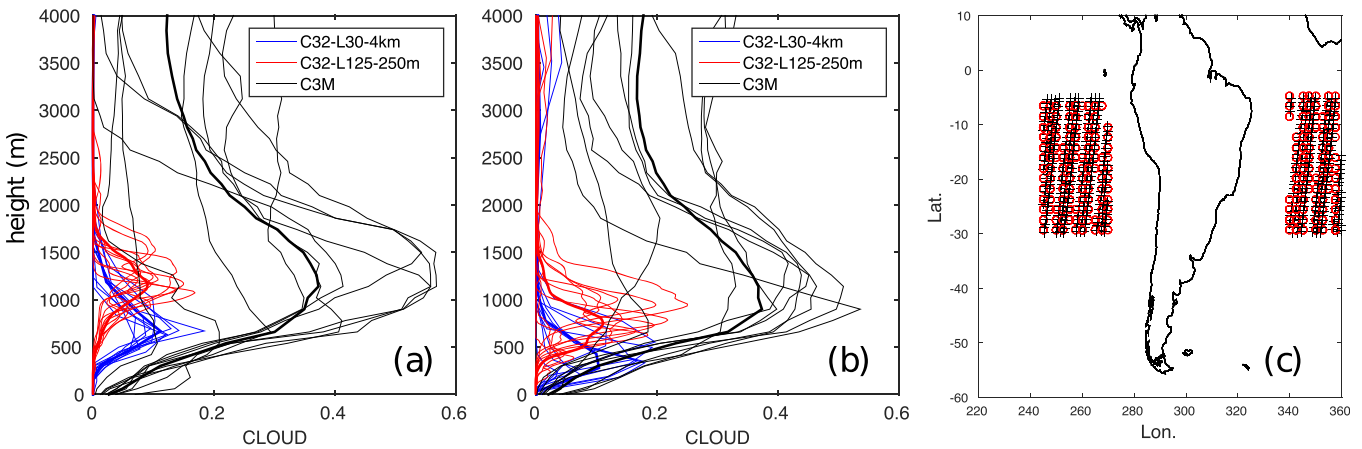


**Figure 6.** Daily mean UP/SP cloud height profile comparison against C3M observations (black) averaged over the closest grid points to satellite swaths (panel c) for 10 ensemble members spanning October 2008. Thin lines are the individual model results and the thick lines are the ensemble mean. (a) Peruvian points, (b) Namibia points and (c) is the map of the C3M satellite swaths (black cross shows every four retrieval) and associated simulation grid points (red circle) on which averaging is performed. Figures 6a and 6b correspond to a swath set on Figure 6c.

down from space [*Kato et al*., 2010, 2011] and https://ceres.larc.nasa.gov/products.php?product=CCCM). Because of the sensor characteristics, the shape of the vertical cloud profile in C3M is more reliable than the absolute cloud fraction for comparison to the model output. Model hindcast output is collocated to the satellite overpass location/time and cloud profiles composited independently across two separate regional populations of A-Train swaths (panel c) that sampled the Peruvian and Namibian low-lying cloud decks on the 10 target dates spanning October 2008. Each calendar date is shown superimposed, with the ensemble mean in bold.

Whereas SPCAM stratocumulus clouds tend to be biased 500 m too low, the vertical location of UPCAM peak cloud fraction is in good agreement with the satellite estimates in both regions. As will be discussed in section 4.3.2, UPCAM improves the vertical position of low-level cloud, due to its better representation of cloud-forming turbulent eddies and their effects.

### 4.2. Area-Mean and Cross-Section Biases in Subtropical Sc Regions

All model configurations capture broadly realistic ensemble mean vertical profiles of equivalent potential temperature and water vapor during the initial simulation day in the Peruvian region (Figure D1), subject to the caveat that the YOTC reanalysis may not be a reliable proxy for observations of the detailed vertical boundary layer thermodynamic structure. Overall, in all configurations there is a systematic drift to a slightly drier and warmer boundary layer, and too moist a free troposphere.

We earlier noted that enhancing BL vertical and horizontal resolution via UP has only marginal impacts on daily mean ASR biases in the subtropics (Figures 4a and 4d). Differences in ASR can be interpreted as the combined effects of refined resolution on time mean cloud fraction (CF) and area-mean LWP, given that cloud droplet number concentration (CDNC) does not enter into the one-moment microphysics used in these simulations. Figure 7 shows that model differences in LWP bias (calculated with respect to observation estimates of MAC-LWP, available at https://disc.gsfc.nasa.gov/uui/datasets?keywords=lwp&start=1920-01-01&end=2017-02-07) largely mirror model differences in ASR biases. MAC-LWP combines data from 13 passive microwave retrievals with corrections applied to biases in regions where optical imager LWPs are more trustworthy; the first two harmonics of the diurnal cycle are sampled by MAC-LWP. With both the L30 and L125 grids, increasing horizontal resolution from 4 km to 250 m (Figures 7a and 7c versus Figures 7b and 7d) reduces LWP throughout the subtropical trade-Cu and Sc latitudes, consistent with an increase in ASR (Figures 4b and 4d). Enhancing vertical resolution (Figures 7a and 7c) increases LWP to unrealistically large values in the trade-Cu regions in the 4 km simulations, consistent with a decrease in ASR. Overall, the LWP bias pattern in SP and UP look remarkably similar, with too little LWP over the subtropical latitudes, and too much over the ITCZ and Southern Ocean. When UP is used (Figure 7d) LWP RMSE is similar to that of standard SP (Figure 7a), but global mean LWP biases decrease from 9 to 3 g/m$^2$, partly due to reduced overestimation of LWP over the Southern Ocean.

### 4.3. Twelve to Thirty-Six Hour Cloud and BL Evolution at a SE Pacific Grid Column

To better understand differences in the UP versus SP cloud vertical structure in subtropical Sc regions, we analyze the transient evolution 12–36 h into the hindcast of selected horizontal-mean quantities in the CRM corresponding to a single CAM5 grid column in the Peruvian region at 19.9°S, 275.0°E; local noon is 17 UTC at this location, averaged over the hindcast ensemble. The essential grid resolution sensitivities we will demonstrate at this grid point are broadly representative of the whole Peruvian Sc region and can be discerned in arbitrary individual hindcasts in addition to the ensemble mean.

#### 4.3.1. Cloud Fraction and Liquid Water Path

Figure 8 shows a full day's evolution of horizontal-mean cloud fraction in this CRM, beginning from 0 UTC (around 7 P.M. local time), 12 h after initialization (i.e., forecast hours 12–36). Standard SP (C32-L30–4 km) (panel a) simulates resilient but very low lying cloud in the 400–800 m layer. It exhibits too small a cloud fraction (generally less than 15% compared to 100% typically observed) and is better thought of as "pancake" cumulus with weak, 4 km wide updrafts rather than stratocumulus. However, it is able to persist past sunrise (12 UTC) throughout most of the first simulated day and thus it produces a nontrivial shortwave cloud radiative effect. When only horizontal resolution is increased to the UP value of 250 m (panel b) a similarly low Sc layer occurs but begins to exhibit high-frequency periods of intermittently low-cloud fraction close to zero. If instead vertical resolution is increased to UP's L125 (panel c), the Sc layer occurs higher in the troposphere (600–1000 m) and maintains larger ensemble mean cloud fraction (20–25% during the

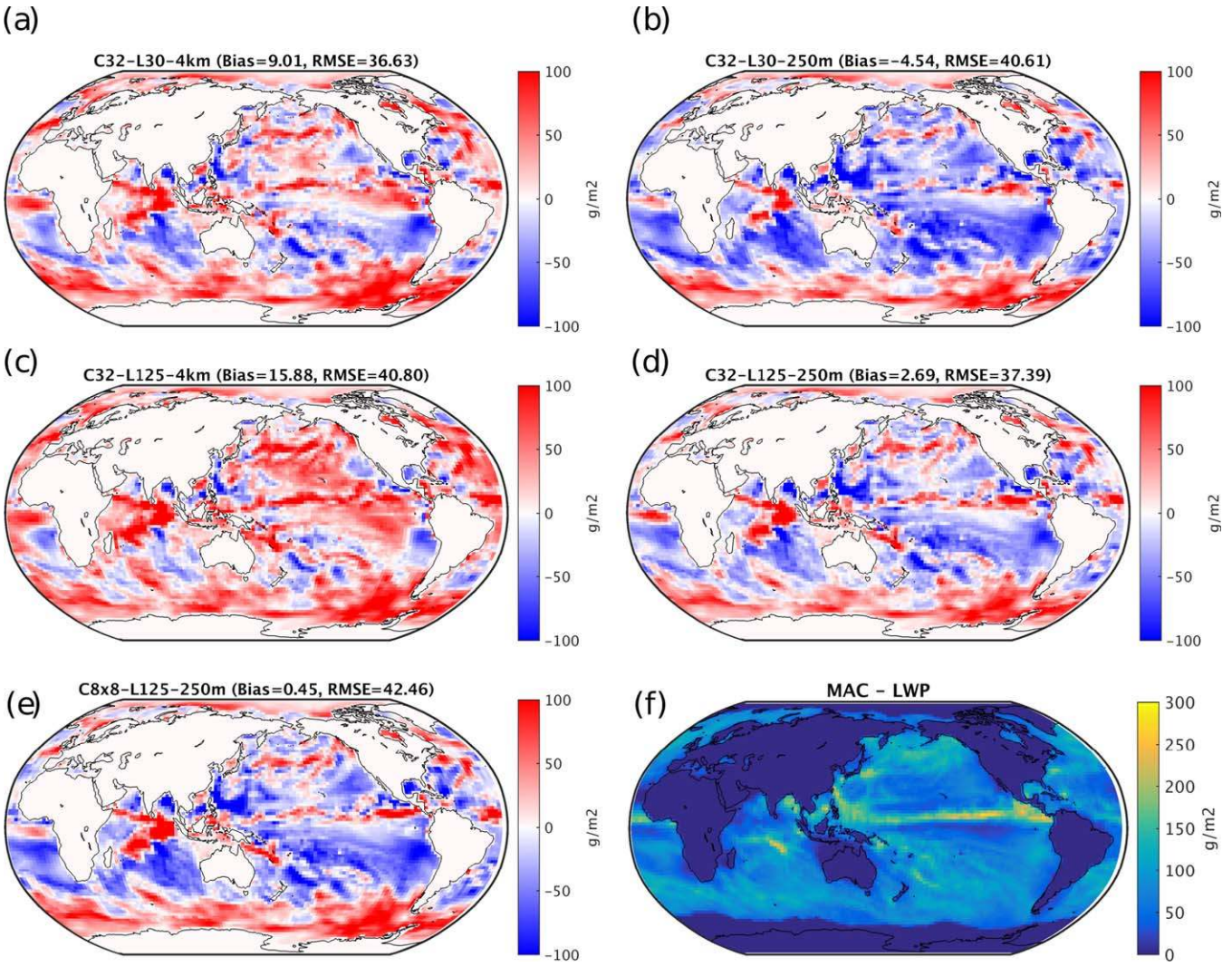


**Figure 7.** Ensemble mean oceanic LWP bias of the model: (a) C32-L30–4 km (SP), (b) C32-L30–250 m, (c) C32-L125–4 km, and (d) C32-L125–250 m (UP), and (e) C8 × 8-L125–250 m, relative to (f) the mean October 2008 LWP observations (based on MAC-LWP).

night, decreasing to 5–10% after sunrise) and a realistically sharp inversion forms. With both UP vertical and horizontal resolution (panel d) cloud fraction is slightly lower (10–15%) but with increased intermittency and within thin layers that are internally variable within the ensemble (such that multiple distinct traces are visible in the ensemble mean), and which are also located at a higher altitude than all other configurations (800–1300 m). Thus, both vertical and horizontal resolution play important roles in the amount and vertical structure of the boundary layer cloud.

Offline analysis of individual members of the UP hindcast ensemble (not shown) indicate symptoms of potential overentrainment, such as quasiperiodic cycles of developing Sc and rapid inversion rise followed by clearing and collapse. Under UP the system struggles to maintain liquid after sunrise (∼12 UTC). The C8 × 8 UP version (panel e) behaves similarly to UP but exaggerates symptoms of overentrainment and cloud intermittency, leading to extremely low time-mean cloud fraction.

Figure 9 shows the corresponding time evolution of LWP and low-cloud fraction (CF) for all model configurations at the same Peruvian grid point. As expected from Figure 8, there are striking model differences in the daily evolution of LWP and CF. Standard SP (solid blue) LWP (CF) rarely exceeds 30 g/m$^2$ (50%) even at night (before 12 UTC) when observations here frequently exceed 100 g/m$^2$ (90%) [*de Szoeke et al*., 2012; *Burleyson et al*., 2013]. The L30–250 m configuration has an even lower CF and even less LWP. The L125 configurations (red) maintain encouragingly higher nocturnal CF exceeding 60%. For standard UP (red dashed),

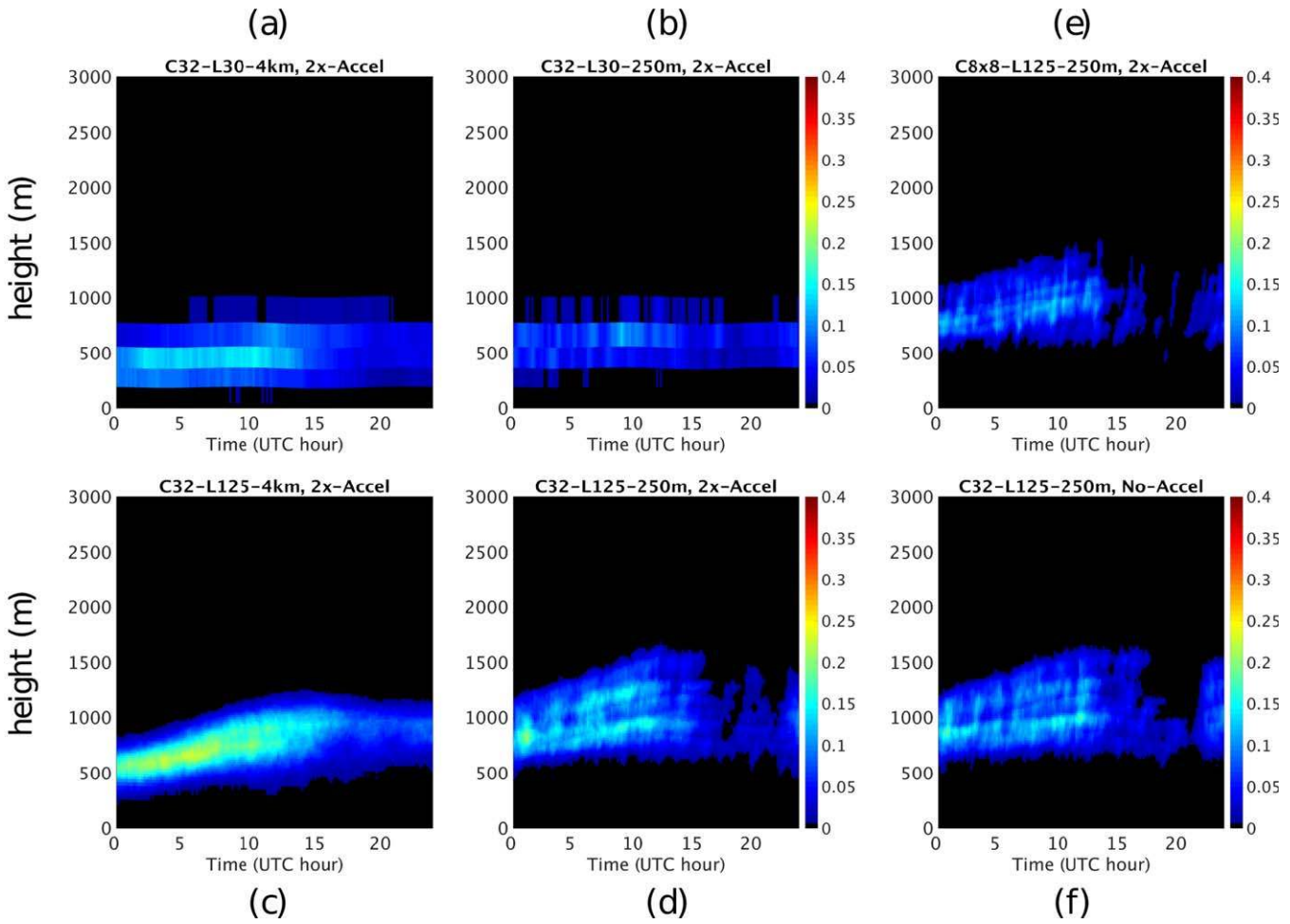


**Figure 8.** Ensemble-mean height time evolution of the cloud fraction at −19.9°N, 275.0°E. (a) C32-L30–4 km (SP), (b) C32-L30–250 m, (c) C32-L125–4 km, (d) C32-L125–250 m (UP), (e) C8 × 8-L125–250 m, and (f) C32-L125–250 m (UP) but without CRM acceleration.

area-mean LWP is variable and desirably 50% higher than SP throughout the night. The L125–4 km configuration (red solid) behaves similarly at this location, but it greatly overestimates LWP and SWCRE in trade cumulus locations. This indicates that despite the fact that none of the simulations yet reproduce satellite observations of 90–100% nocturnal cloud fraction with LWP frequently exceeding 100 $g/m^2$, added UP resolution is beginning to produce fledgling Sc-like cloud layers, especially at night; something that SP is incapable of doing during hours 12–36 of a realistically initialized hindcast.

#### 4.3.2. Turbulence

These resolution sensitivities are connected to how characteristics of turbulence and its generation differ among these simulations. Turbulent eddies transport water vapor upward from the ocean surface to form cloud and are the physical driver for entrainment. We analyze the generation of turbulent kinetic energy (TKE), its partitioning between eddy horizontal and vertical velocity variance, and its relationship to entrainment.

Figure 10 shows the evolution (12–36 h) of the CRM-simulated buoyancy flux profile for these configurations. Buoyancy flux is the principal TKE source in most subtropical cloud-topped BLs [*Moeng et al*., 1992; *Bretherton and Wyant*, 1997]. Figure 11 shows the corresponding profiles of the vertical velocity variance.

Aircraft observations during the VOCALS-REX field experiment, which took place over the SE Pacific Ocean during October 2008, generally showed decoupled BLs in this region with cumulus rising into stratocumulus [*Bretherton et al*., 2010]. These are characterized by a surface-driven BL with maximum buoyancy flux near the ocean surface, slightly negative buoyancy flux at the Cu cloud base, above which buoyancy flux increases with height up to a maximum in the Sc layer [*Nicholls*, 1984; *Van der Dussen et al*., 2013]. LES and

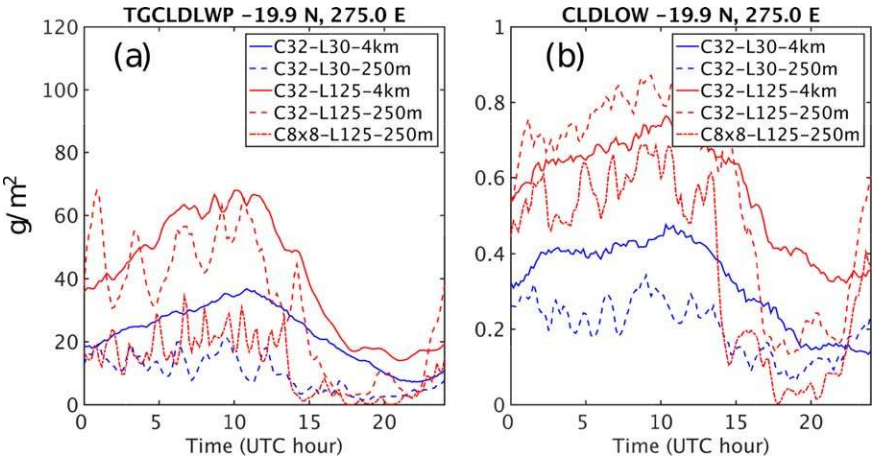


**Figure 9.** Time series of the (a) ensemble mean LWP and (b) low-cloud fraction.

observations also show separate peaks in the vertical velocity variance and TKE in the subcloud and Sc layers [*Nicholls*, 1984; *Van der Dussen et al*., 2013]. Decoupling is driven in large part by entrainment of warm dry air through the capping inversion [*Bretherton and Wyant*, 1997]. Boundary layers with a deep inversion 500 m or more above the surface lifted condensation level, commonly seen at this location, are more susceptible to decoupling [*Bretherton and Wyant*, 1997; *Jones et al*., 2011].

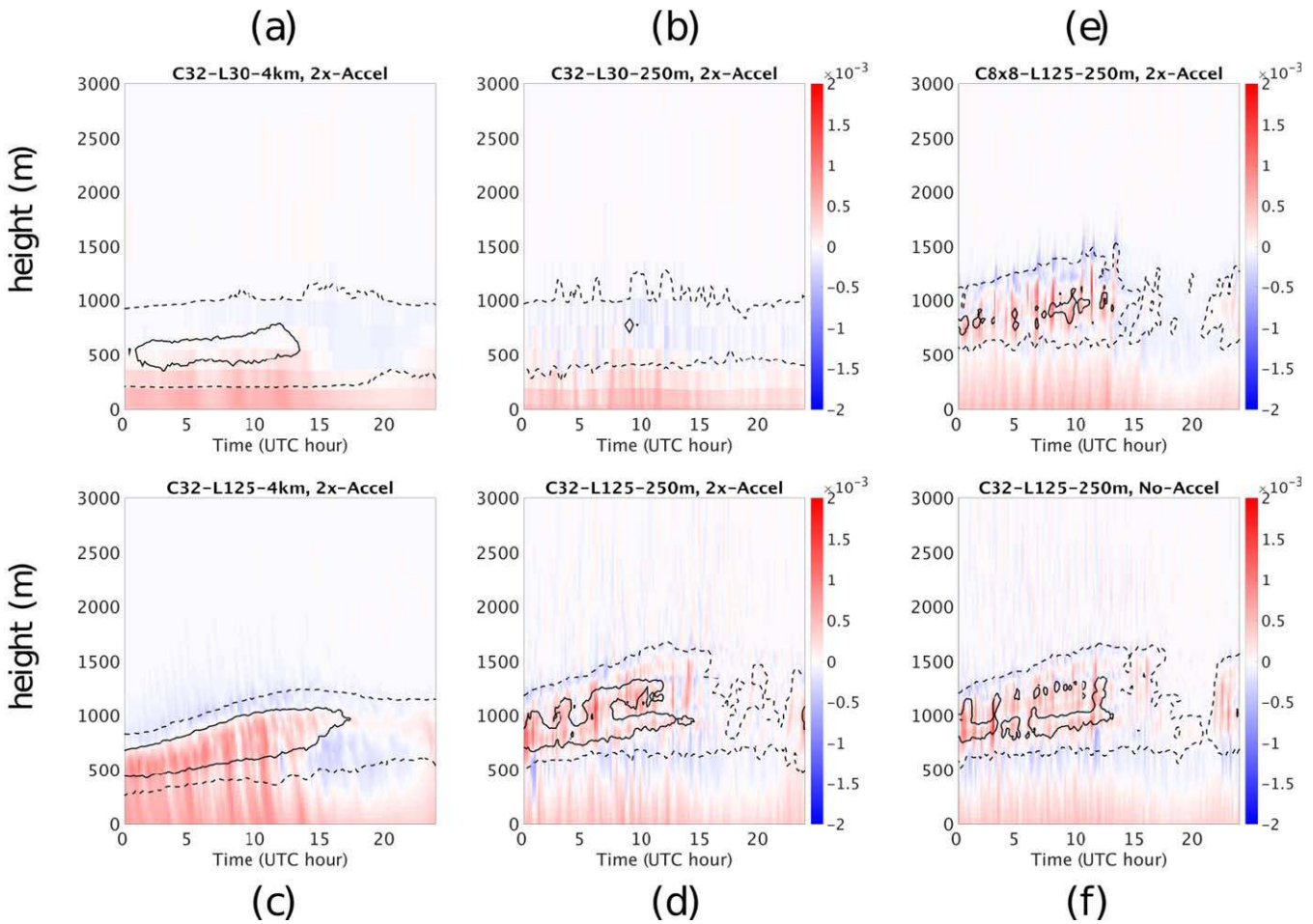


**Figure 10.** As in Figure 8, but for buoyancy flux (in units of $m^2/s^3$). The black contours mark the 0.01 (dashed) and 0.1 (solid) cloud fraction.

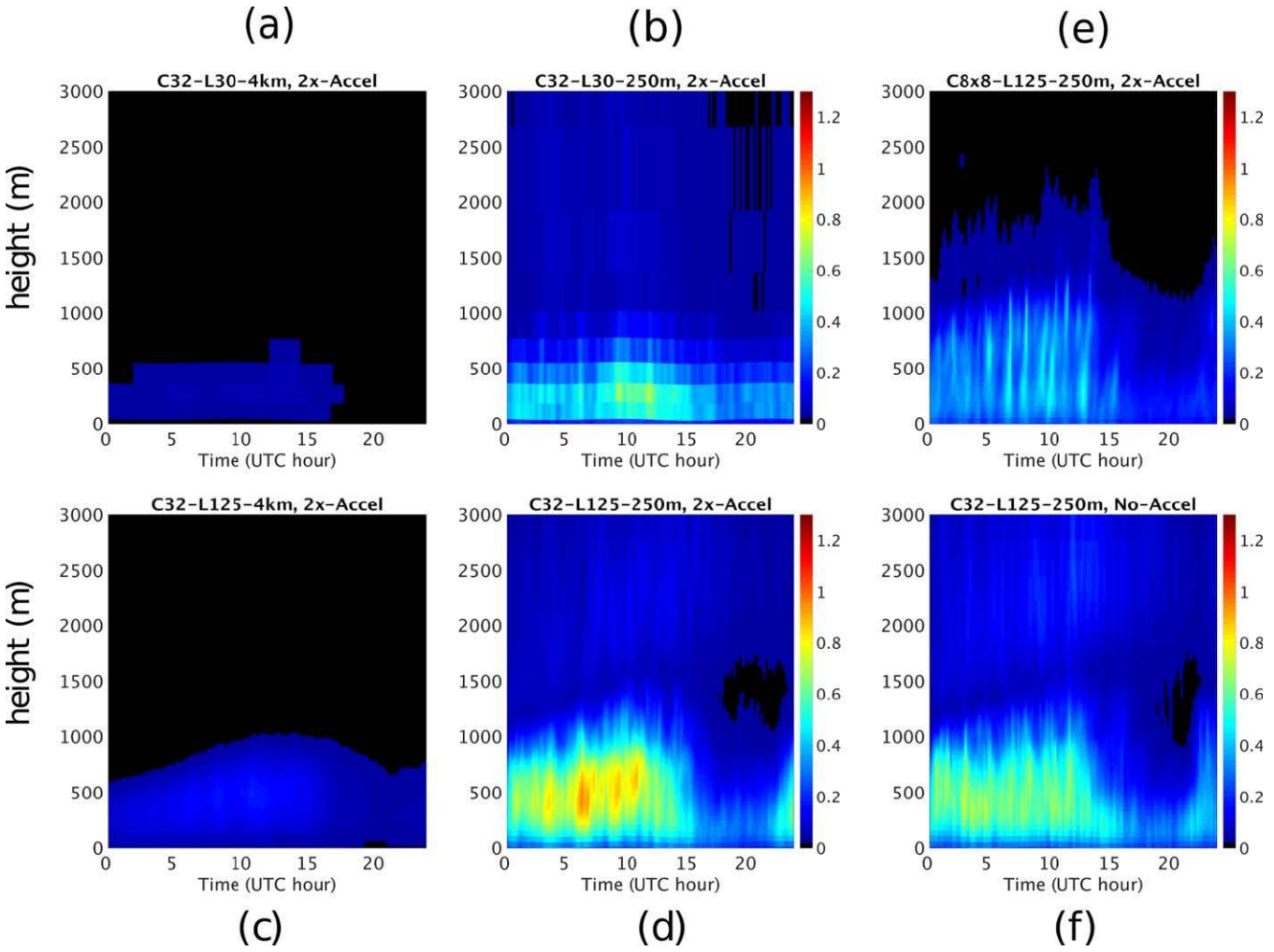


**Figure 11.** As in Figure 8, but for vertical velocity variance $\langle w'w' \rangle$ (in units of $m^2/s^2$).

Figure 10a shows the buoyancy flux for standard SP (L30–4 km) at the offshore Peruvian grid point. The expected decoupled vertical structure of the BL is not captured. During the night when weak Sc is present, there is only a surface buoyancy flux maximum. This behavior is due to the relatively weak boundary layer updrafts supported by the broad 4 km horizontal grid spacing (Figure 11a); most of the TKE is instead partitioned into overly strong (compared to LES) horizontal velocity perturbations (not shown) maximizing near the top and bottom of the BL, again suggestive of a single layer of convective cells. The weak updrafts do not disturb the inversion enough to drive realistic entrainment. Instead, numerical diffusion is the main source of entrainment, and at this location it does not produce enough entrainment to induce decoupling. Thus, the inversion height is biased low and the BL has a vertically well-mixed structure and relatively weak entrainment of dry, warm air; together these help keep the top of the BL moist and filled with a diffuse patchy cloud layer, but via physical processes that are distorted from reality and unable to sustain observed liquid water contents or realistically strong diurnal variability.

Figure 10c shows that the buoyancy flux profile for the L125–4 km case, which supported the highest LWP and CF, is somewhat more realistic but still not indicative of a satisfyingly decoupled BL. Positive buoyancy fluxes adjacent to the surface extend to higher altitudes. While an in-cloud source of buoyancy fluxes is visible, it is intermittent and does not dominate the buoyancy flux profile. Furthermore, vertical velocity variance profiles (Figure 11c) are only marginally enhanced relative to those from L30 SP and do not contain the expected double-peaked structure.

Figure 10d shows that the buoyancy flux profile for standard UP (C32-L125–250 m) maintains a more realistic double-maximum indicative of a decoupled BL. The profiles of buoyancy flux and vertical velocity variance (Figure 11d) have the expected double-peaked vertical structure. The vertical velocity variance is episodically much larger, associated with updrafts and downdrafts that now seem too strong and that induce the bursts of overentrainment. With the C8 × 8 grid (small 3-D, Figure 11f), the maxima in Sc-layer vertical velocity variance are desirably weaker, but the cloud liquid content remains low, suggesting that this has not fixed the overentrainment issue. The height-time evolution of TKE (not shown) and buoyancy flux in UP is resilient to the use of 2X CRM acceleration (Figure 10d versus Figure 10f), other than a slight increase of the most extreme vertical velocity variance values (Figure 11), but this effect is generally only weakly detectable at the majority of grid columns.

## 5. Analysis of a Trade Cumulus Region

The transition from stratocumulus to cumulus clouds is a fundamental characteristic of the Hadley circulation that is challenging to capture in current GCMs, many of which suffer from a "bright-trades" bias [*Flato et al*., 2013]. In this section, we analyze the details of cloud vertical structure and cloud development for a trade cumulus regime in the grid column closest to the Barbados Cloud Observatory (BCO at 13°N, 300°E). We have verified that the model differences here are representative of other grid points in the surrounding area and in other trade-Cu regimes including the southeastern Pacific.

Shallow-Cu in the Barbados region form predominantly in a weak subsidence regime (500 hPa $\omega$ roughly 10 hPa/d), in which satellite lidar indicates cloud top is usually observed between 1 and 2.8 km altitude and cloud base can be as low as 500 m [*Medeiros and Nuijens*, 2016]. Observed low-cloud fraction in October is 20–40% [*Nuijens et al*., 2015, 2014].

The ensemble mean height-time evolution of the simulated Barbados cloud fraction is plotted in Figures 12a–12e. Standard SP (panel a) yields unrealistically thick Sc-like clouds at the Barbados grid point (and in the surrounding region); a propensity for this configuration to simulate overly low trade cumulus cloud was previously noted by *Marchand and Ackerman* [2010]. Vertical refinement alone (panel c) leads to clouds that are slightly deeper but still generally lower than observed. The standard UP cloud field (panel d) is more reminiscent of shallow Cu with deeper updrafts attaining 2–3 km. In the UP C8 × 8 setup, the shallow cumulus layer is similar to the standard UP setup.

Figure 12f shows that the resolution effects can be explained by changes in turbulence characteristics. Standard SP does not simulate deeper trade-Cu because its vertical velocity variance is weak (Figure 12f). The C32 UP configurations produce much more vertical velocity variance (maximizing $\sim$0.25 $m^2/s^2$ in the subcloud layer). This subcloud vertical velocity variance is at least 25% larger than the C8 × 8 UP configuration. A 2-D grid exaggerates vertical velocity variance since there is only one horizontal dimension, so a larger fraction of TKE is partitioned into vertical motions. However, the second maximum of the UP vertical velocity variance in the Cu layer, seen in both C32 and C8 × 8 versions, is consistent with BOMEX LES of *Siebesma et al*. [2003] in a similar regime, suggesting that UP is capturing plausible turbulence dynamics of a shallow cumulus BL.

Figure 13 uses satellite data from A-Train swaths to validate the realism of the SP/UP cloud profiles in the trade cumulus regime (panels a and b for Barbados and the Pacific region, respectively). Comparing Figures 6 and 13, UP is realistically promoting more differentiation in the vertical structure of shallow Cu versus Sc clouds, whereas SP is biased toward producing overly bottom heavy cloud structures in both regimes, with insufficient diversity between the shallow Cu and Sc regions.

## 6. Analysis of 90 Day Simulations

### 6.1. Mean State Climatology

Table 4 summarizes statistics of the UP versus SP global mean climate from 90 day integrations. These runs are initialized on 1 October 2008 with climatological SSTs using high-resolution ECMWF initial conditions. The most striking sensitivities are an increase of low-cloud fraction from 36 to 50%, a 7 $W/m^2$ (14%) increase in shortwave cloud radiative effect (SWCRE) and reduced ASR, and a 4 $W/m^2$ (20%) increase in longwave cloud radiative effect. Global mean biases of LWCRE, SWCRE (relative to CERES for October–December 2008) and precipitation (relative to GPCP climatology of October–December 2008) are tabulated in Table 5.

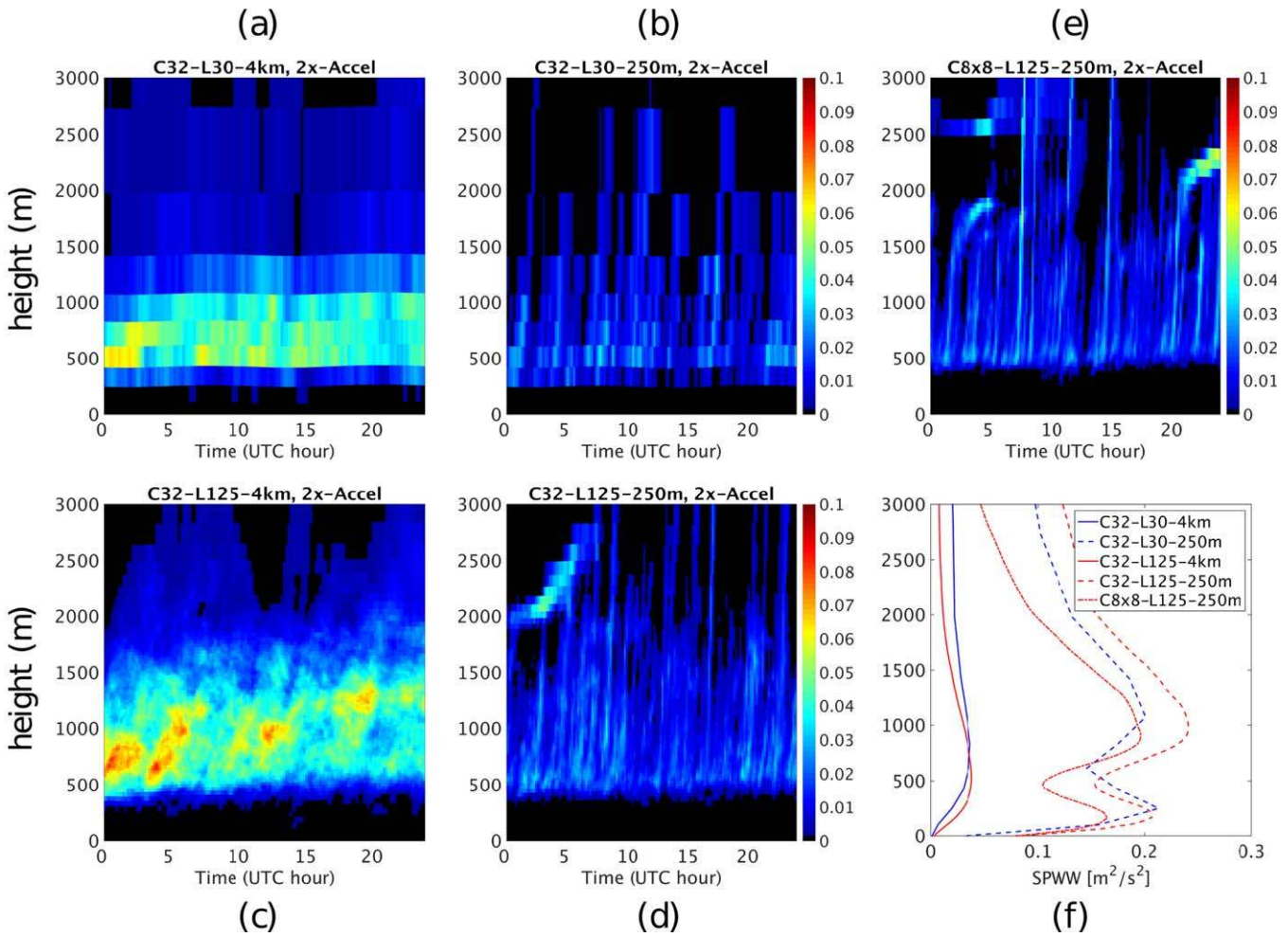


**Figure 12.** (a–e) Height time evolution of cloud fraction in trade cumulus region (at 12.3°N, 300°E) averaged over ensemble members spanning the October 2008. (a) C32-L30–4 km (SP), (b) C32-L30–250 m, (c) C32-L125–4 km, (d) C32-L125–250 m (UP), and (e) C8 × 8-L125–250 m. (f) Profiles of ensemble-time-mean vertical velocity variance $\langle w'w' \rangle$ at the same grid point.

The impacts of UP versus SP on time-mean precipitation rate (Figure 14) are reassuringly small. The global mean precipitation rate of UP (2.72 mm/d) is virtually indistinguishable from SP (2.75 mm/d). Other than a northward shift of the ITCZ in the Atlantic and Pacific, and a general tendency for rainfall to be more patchy

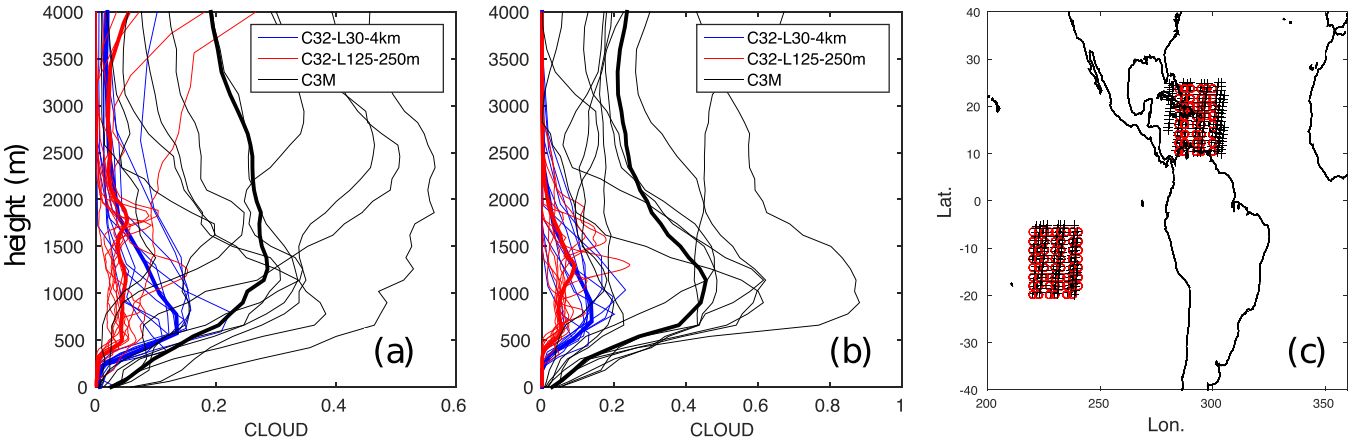


**Figure 13.** As in Figure 6, for the trade-Cu regions. The model data are scaled by factor of 3 for visibility. (a) Barbados region, (b) Pacific region, and (c) map of the C3M satellite swaths (black) and simulation points (red). Figures 13a and 13b correspond to a set on Figure 13c.

**Table 4.** Global Mean Statistics of the 90 day SP and UP Simulations

| Quantity | C32-L30–4 km (SP) | C32-L125–250 m (UP) |
|---|---|---|
| Precipitable water [mm] | 24.48 | 25.65 |
| Precipitation [mm/d] | 2.72 | 2.75 |
| Liquid water path [$g/m^2$] | 89.15 | 88.76 |
| Ice water path [$g/m^2$] | 46.05 | 50.73 |
| Total cloud fraction | 0.53 | 0.63 |
| High cloud fraction | 0.24 | 0.23 |
| Low-cloud fraction | 0.36 | 0.50 |
| LWCRE [$W/m^2$] | 24.92 | 28.99 |
| SWCRE [$W/m^2$] | −52.22 | −59.30 |
| OLR [$W/m^2$] | 220.57 | 217.80 |
| ASR [$W/m^2$] | 194.97 | 189.85 |

in UP, geographic changes in rainfall clearly distinct from internal variability on a 3 month timescale are not apparent.

For OLR, when examining the 90 day results (Figures 15b, 15d, and 15f) effects of UP are more striking than originally noted in Figure 5 as teleconnected effects of UP on deep convection spin up. UP tends to reduce outgoing longwave radiation over the majority of the planet leading to a global mean bias that is −6 W/m$^2$, higher in magnitude than that of SP. This is likely due to a bias in underlying thermodynamic profiles that impacts clear-sky OLR given that UP actually improves the climatology of longwave cloud forcing (Table 4) improving SP's LWCRE bias of −3.2 W/m$^2$ by raising the global mean LWCRE from 25 to 29 W/m$^2$. This results in a global mean bias of only 0.9 W/m$^2$ relative to CERES-EBAF data for October–December 2008.

Effects of UP on time-mean low cloud and SWCRE are more pronounced. UP's 3 month climatology exhibits a dramatic increase in cloud fraction over the majority of the ocean (see Figures 16b, 16d, and 16f). The global mean total cloud fraction increases from 53% in SP to 63% in UP, especially due to increases in the offshore sectors of subtropical Sc regions (Table 4) and extratropical storm tracks, where regional cloud fraction enhancements are frequently as high as 30–40%. These global effects can also be detected in the 2 day hindcasts (not shown). UP is more consistent with satellite estimates using both active and passive sensors, which suggest that mean global cloud fraction of optical depth greater than 0.1 is 0.68 ± 0.02 [*Stubenrauch et al*., 2013]. Given that UP underestimates ASR (Figures 15a, 15c, and 15e) and significantly overestimates the magnitude of global SWCRE compared to CERES (Table 5), this suggests that the UP clouds must on average be too optically thick. In future, activation of SPCAM's ISCCP and MISR simulators [*Marchand and Ackerman*, 2010] within UP could be a fruitful way to compare its vertical cloud distribution with satellite observations in more detail.

Two unintended regional consequences of UP are excessive cloud brightening over tropical rainforests and dimming of Sc immediately adjacent to the coasts of Peru, Namibia, and W. Australia (Figures 15a, 15c, and 15e). Fast emergence of similar regional biases can also be seen to dominate the 12–36 h hindcast results (Figure 4) but are increased in magnitude in the 90 day results. The tropical rainforest bias and oceanic ASR biases also take more than 36 h to spin up. Offline analysis of tropical rainforest regions indicates brightening is associated with strong local enhancement of low and mid-level cloudiness. While important, we do not expect these distortions of continental deep convection teleconnect to the mean offshore subsidence environment given that UP and SP simulate similar mean precipitation over tropical rainforests (Figure 14e). A decrease in the realism of time-mean Sc brightness is discouraging and will be discussed in detail in the following sections. However, first we demonstrate below an encouraging improvement in UP's stratocumulus diurnal variability and mean vertical positioning.

**Table 5.** Global Mean Biases of the 90 day C32-L30–4 km (SP) and C32-L125–250 m (UP) Simulations With Respect to Observations of the Same Period

| | | SP | | UP | |
|---|---|---|---|---|---|
| Quantity | Observation | Mean | RMSE | Mean | RMSE |
| Precipitation [mm/d] | GPCP | 0.057 | 2.25 | 0.014 | 2.88 |
| Longwave cloud forcing [$W/m^2$] | CERES-EBAF | −3.2 | 10.3 | 0.9 | 10.2 |
| Shortwave cloud forcing [$W/m^2$] | CERES-EBAF | −3.8 | 21.0 | −10.8 | 25.9 |
| Outgoing longwave radiation [$W/m^2$] | CERES-EBAF | −1.4 | 11.7 | −6.0 | 14.6 |
| Absorbed solar radiation [$W/m^2$] | CERES-EBAF | −0.8 | 18.7 | −7.4 | 22.5 |

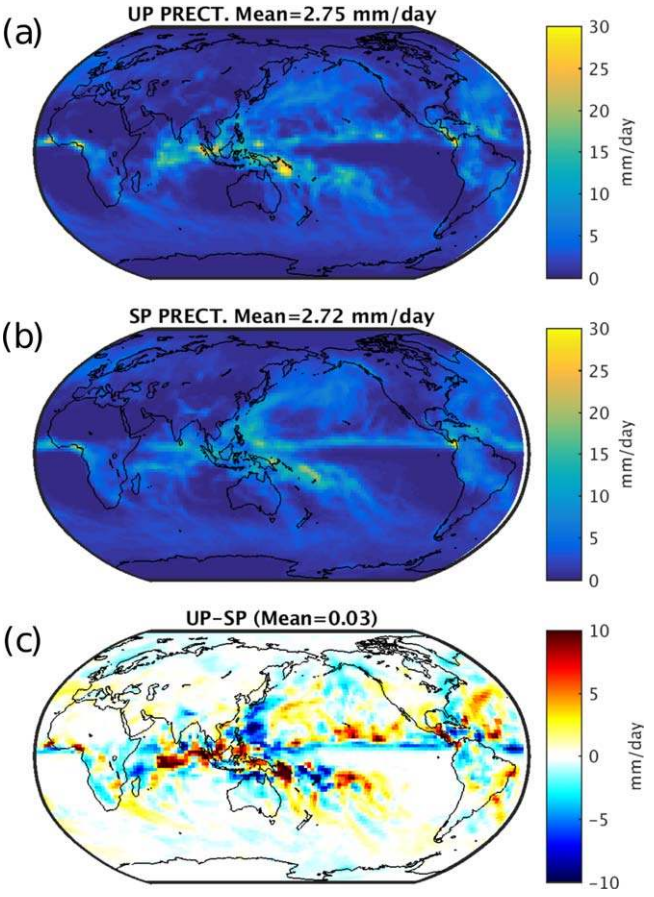


**Figure 14.** 90 day mean precipitation for (a) C32-L125–250 m (UP), (b) C32-L30–4 km (SP), and (c) UP-SP difference.

### 6.2. Improved Stratocumulus Diurnal Cycle

A desired property of UP is improved sensitivity of Sc to changes in radiative forcing. In this section, we demonstrate that the UP low-cloud response to the most simply forced mode of radiative variability—the diurnal cycle—is significantly improved relative to SP.

Figure 17 shows the global structure (amplitude and phase) of the liquid water path diurnal cycle in the models and in merged satellite observations (MAC-LWP product available at https://disc.gsfc.nasa.gov/uui/datasets?keywords=lwp&start=1920-01-01&end=2017-02-07) for the same time period as the simulations, October–December 2008. For consistency with the MAC-LWP data set, the 90 day composite diurnal cycle in the models is fit with a superposition of 12 and 24 h harmonics using a five-parameter least squares linear regression. The resulting diurnal range and local hour of maximum LWP are visualized where the amplitude is statistically distinguishable from zero (i.e., where the statistical error in the fit is less than half the diurnal amplitude). Detectability is weaker in the satellite data than model output due to more limited sampling. Whereas both UP and SP succeed in capturing the timing of peak LWP near the end of the night (Figures 17b and 17d) as observed (Figure 17f) differences in skill are evident in their ability to capture observed amplitude (Figures 17a and 17c).

In the observations, the most detectable geographic action center of highest amplitude diurnal LWP variability is located in the offshore sector of the Peruvian Sc deck (Figure 17e). Encouragingly, this is also the highest amplitude region in UP, and both its magnitude and geographic position are successfully simulated (Figure 17a). The next two most significant regions of observed Sc LWP diurnal variation are located off the coasts of Namibia and W. Australia. Again, UP successfully simulates their geographic positions and relative amplitude rank compared to the dominant Peruvian Sc deck, though it overdoes their absolute amplitude. These are definite improvements relative to SP (Figure 17c), which mispositions the regions of highest Sc LWP diurnal variability, confining high-amplitude Sc diurnal variation unrealistically to zones immediately adjacent to the Peruvian and Namibian coastlines, with excessive amplitude. Under SP there is too little diurnal LWP variability across the large offshore sectors of each major Sc deck where it is observed and a spurious diurnal variation in LWP in the equatorial Eastern Pacific. Analogous improvements can be seen in the diurnal cycle of the low-cloud fraction (see Figure D2). This confirms the initial understanding from 12 to 36 h hindcast diurnal composites that despite its low bias in daytime cloud leading to exaggerated time-mean ASR biases, UP improves the physical fidelity of low-cloud variability.

### 6.3. Cloud Vertical Structure Along the VOCA Transect

A classic test of low-cloud fidelity is the vertical structure of cloud fraction along the 20S VOCA transect off the coast of Peru. Figure 18 demonstrates the extent of improvement due to UP in this region for the 90 day October–December climatology. As seen in the 12–36 h hindcasts, UP elevates Sc altitude and the slope of cloud altitude rises with longitude, both of which are in agreement with C3M data. It is also reassuring

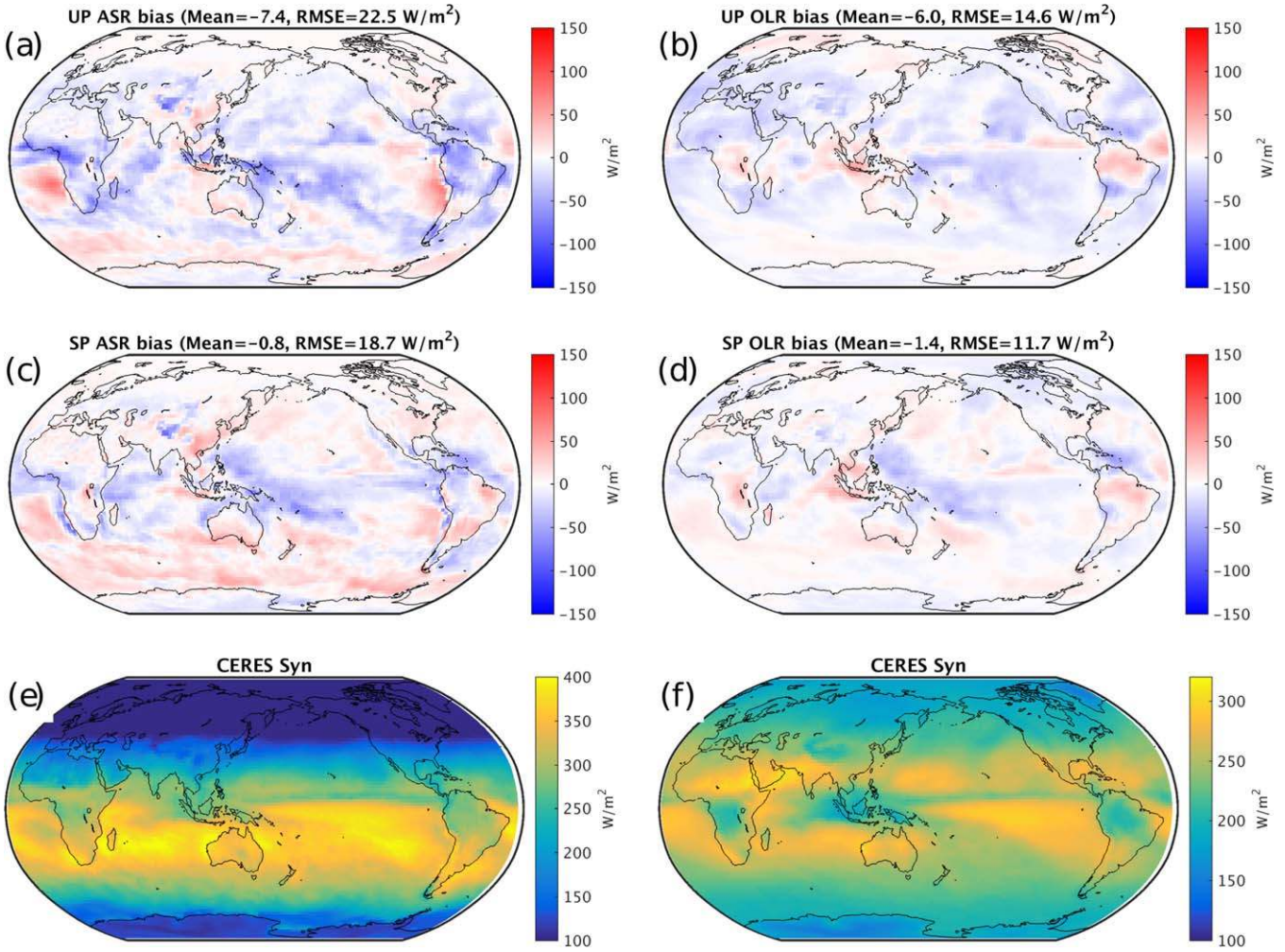


**Figure 15.** 90 day mean biases of ASR (a, c) and OLR (b, d) with respect to CERES Syn observations for October–December 2008 (e, f). Top row: C32-L125–250 m (UP) and middle row: C32-L30–4 km (SP).

that UP captures the zonal location of peak LWP in the 10–20S band. We emphasize that in C3M, the shape of the cloud profile is more reliable than the absolute magnitude for comparison to the model output. Similar improvements in low-cloud vertical structure along this transect have been noted by others using higher order closure parameterizations of the sub-km regime [e.g., *Cheng and Xu*, 2011; *Wang et al*., 2015].

## 7. Discussion

We have found both more and less beneficial effects of using ultraparameterization in place of standard superparameterization, based on our primary analysis metrics of vertical cloud fraction profiles and global top-of-atmosphere shortwave radiation biases from an ensemble of ten 36 h global hindcasts, and additional metrics of 90 day climatology and composite diurnal variability during October–December 2008.

As might be anticipated from offline LES intercomparison studies of trade cumulus and stratocumulus BLs, not only can explicitly capturing BL dynamics improve local turbulence statistics, but also the associated vertical cloud distribution. Ultraparameterization removes a key bias of SP, which mispositions both stratocumulus and shallow cumulus clouds too low in the atmosphere. UP enhances vertical eddy motions and cloud-top entrainment, supporting a more realistically deep, decoupled cumulus-under-stratocumulus state. UP also realistically increases low-cloud fraction across the subtropics compared to SP. Over the subtropical oceans, UP gives similar errors as SP compared to observations of TOA shortwave radiation and grid-mean LWP.

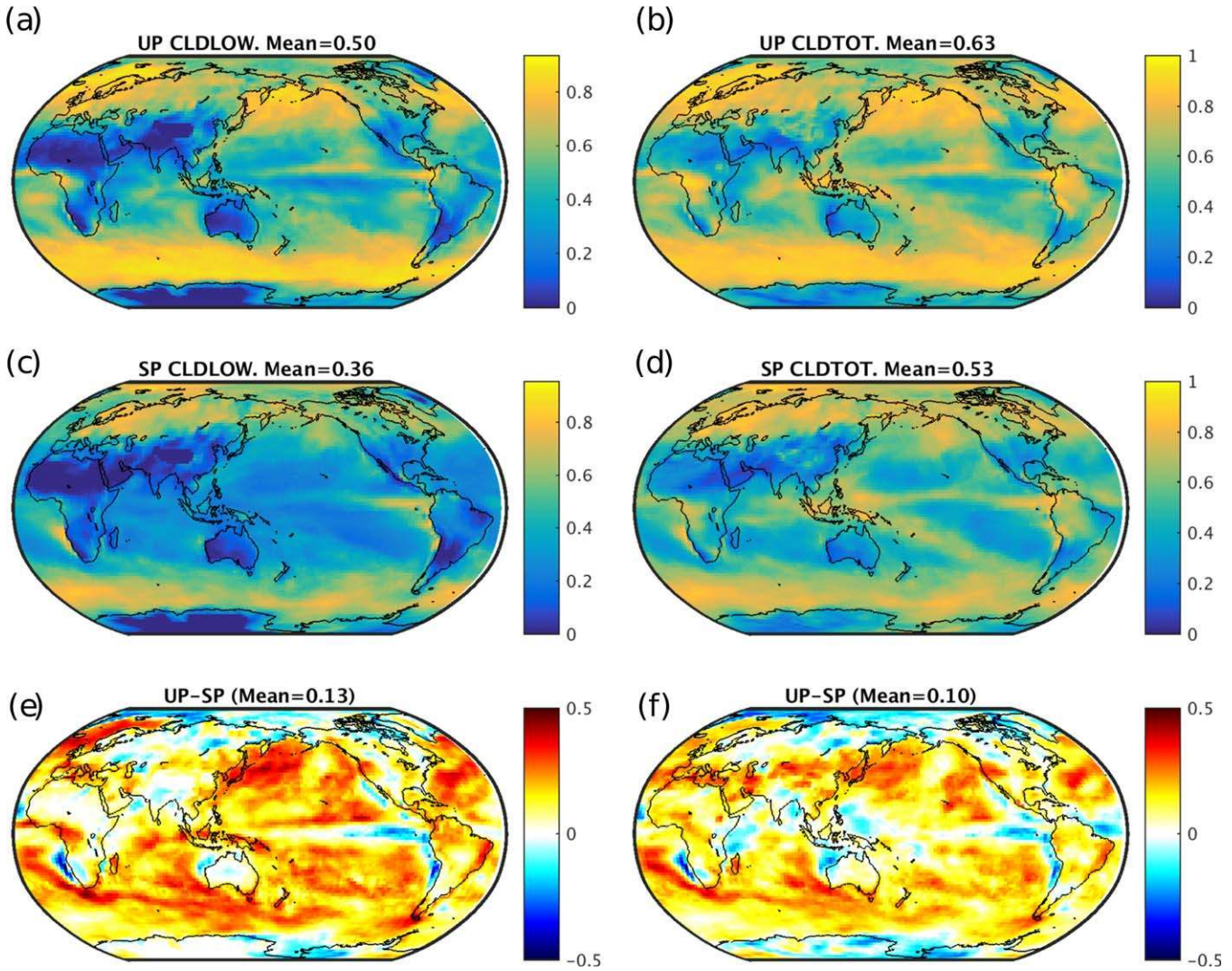


**Figure 16.** 90 day mean of low-cloud fraction (a, c, e) and total cloud fraction (b, d, f). Top row: C32-L125–250 m (UP), middle row: C32-L30–4 km (SP) and bottom row: UP-SP difference.

The sensitivity of low clouds to diurnal variations in radiative forcing is also improved when using UP. UP captures the observed geographic structure and relative amplitude of action centers of strong liquid water diurnal variability within each of the world's main Sc decks. An improved sensitivity of low clouds to diurnal radiative forcing builds confidence in the potential skill of UP in simulating low-cloud feedback on climate change.

There are ongoing problems with using our current version of UP. The boost in low-cloud fraction over most ocean areas, while desirable, tends to produce a global mean bright bias absent any retuning. Over tropical rainforests, UP accumulates more liquid water in shallow clouds, producing regional bright biases that contribute to elevated SWCRE RMSE. On the other hand, we expected a substantial degradation of upper level cloud and outgoing longwave radiation distribution over deep convective regions due to the small domain size of UP compared to SP, but UP simulations exceed the skill of SP for OLR. This is encouraging, but it would be philosophically more defensible to enlarge the UP domain or incorporate a dual-CRM approach involving a "switch function" that blends SP and UP depending on the vertical structure of the simulated clouds and turbulence.

Other obstacles are less straightforward and worthy of more discussion. In the near-coastal Sc environment, UP does not resolve SP's undersimulation of time-mean LWP and exaggerates regional dim biases. It is puzzling that unlike offline LES tests that inspired its grid, and despite its promising effects on Sc vertical

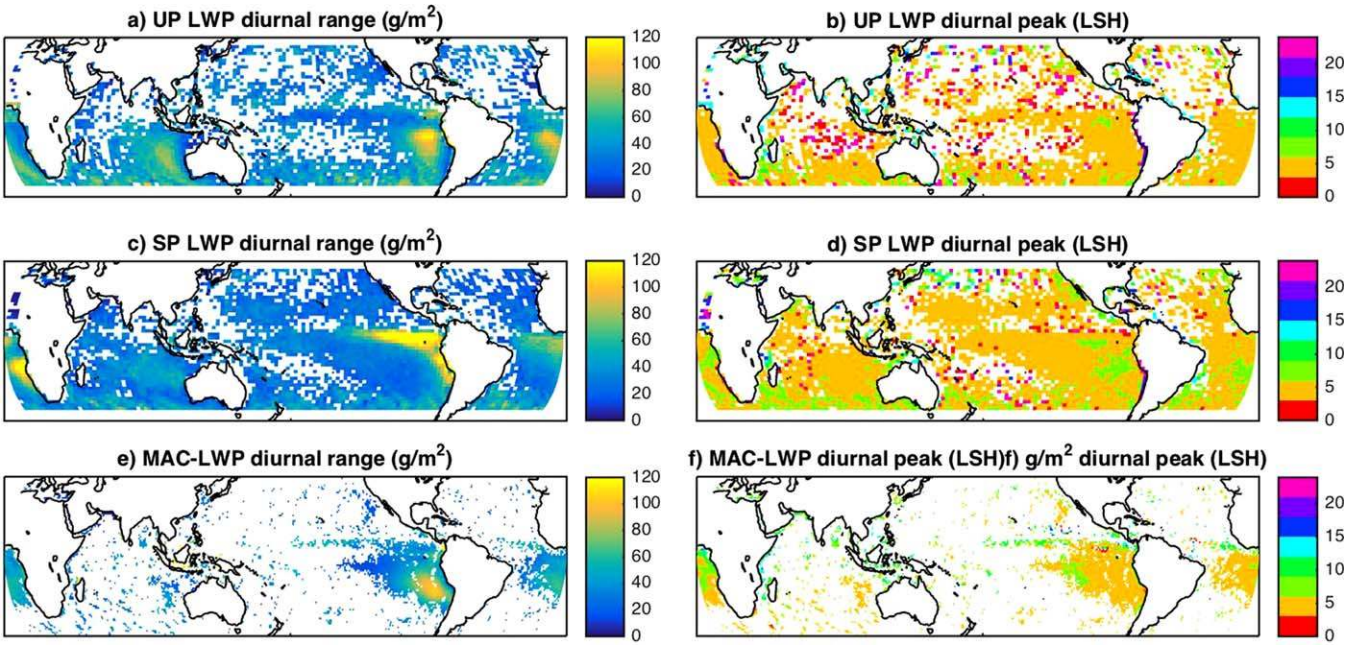


**Figure 17.** Composite diurnal (left) range and (right) local hour of maximum liquid water path calculated from 90 day simulations using (top) UP (C32-L125–250 m), compared to (middle) SP (C32-L30–4 km), validated against (bottom) MAC-LWP merged satellite observations for the simulated time period of October–December 2008.

structure and diurnal variation, UP did not lead to larger LWP gains and associated shortwave bias reductions in all Sc regions. In the following, we speculate on causes.

### 7.1. Possible Explanations of the Marine Stratocumulus Low-LWP Bias

Why does adding so much horizontal and vertical resolution have little impact on time-mean stratocumulus LWP and daytime ASR biases, in contrast to our findings from offline LES simulations?

#### 7.1.1. Precipitation

Feedbacks associated with rainfall production can affect simulated stratocumulus clouds [*Wood*, 2012]. However, this is unlikely to have played a major role in hours 12–36 of our hindcast simulations. The percentage of nonzero surface precipitation values across the ensemble of all 15 min surface samples within the entire oceanic Peruvian region is less than 1.5% and for precipitating grid columns the median rain rate is only 0.1 mm/d. Above the surface, the mixing ratio of precipitating condensate is almost always less than 0.01 g/kg and typically less than 0.0001 g/kg.

#### 7.1.2. Surface Fluxes

While offline LES simulations tend to constrain either surface winds or fluxes, in SPCAM interactive winds and near-surface humidity provide another possible source of sensitivities. Such surface feedbacks do not appear to play a leading role in orchestrating the LWP sensitivities seen in our simulations. We have verified

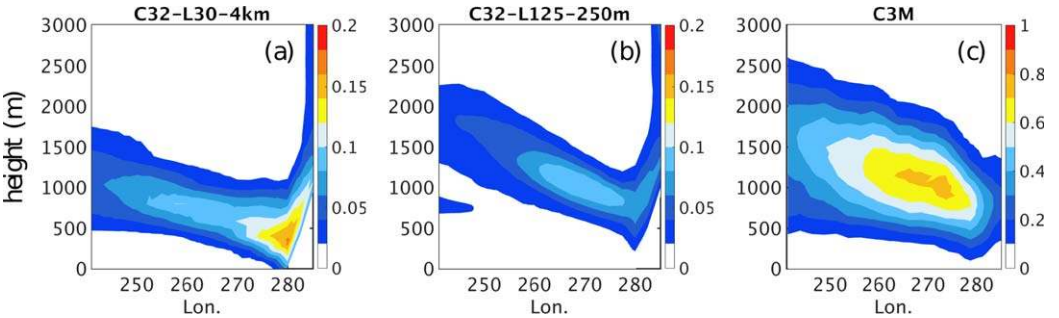


**Figure 18.** Longitudinal cross sections of cloud fraction along 10°S–20°S off the Peruvian coast. (a) C32-L30–4 km (SP), and (b) C32-L125–250 m (UP) are the simulated mean CF of October–December 2008. (c) The 2006–2010 mean C3M cloud fraction of October–December.

that all the configurations give fluxes and temperatures reassuringly close to observations across the target regions.

### 7.1.3. Overentrainment

Detailed analysis reveals that despite its enhancement of nocturnal LWP, UP struggles to maintain its LWP after sunrise. Daytime stratocumulus are especially challenging to simulate explicitly because cloud-top longwave radiative cooling (which helps drive water transport to the cloud layer by energizing boundary layer eddies) is countered after sunrise by shortwave heating. This more quiescent, delicate turbulence environment can become dominated by secondary entrainment from artificial sources like numerical diffusion across the sharp capping inversion. By mixing dry air down into the BL when the turbulence is too weak to mix moist air up from below, this leads to the LES underpredicting daytime cloud liquid water.

Figure 19 shows evidence of such overentrainment acting to inhibit UP daytime LWP by comparing height-time sections of total water (top row), which clearly shows the inversion height, against corresponding vertical velocity variance (bottom row), for three members in the UPCAM ensemble at the Peruvian Sc grid point. Clouds are overlaid (black) and the variations of the inversion height in the model (white dotted) are contrasted with 6 hourly YOTC ECMWF analyses (white solid), with the caveat that ECMWF's inversion is biased 100–200 m low with respect to sounding and satellite data. The UP simulations capture the expected daytime reduction of cloud-layer $\langle w'w' \rangle$ (turbulence) after 12 UTC, which should correspondingly weaken entrainment. However, after 12 UTC, the UP-simulated daytime inversion height (white dotted) remains high despite the striking shutdown of turbulence, implying that entrainment in the model remains strong enough to balance the mean subsidence.

The implication is that an unintended source of diffusion in UP is leading to daytime overentrainment. Offline analysis of diurnally segregated C3M data confirms that UP overestimates (underestimates) daytime (nighttime) inversion height (not shown). Further research investigating details of the time dependence of the inversion height is needed to better clarify the realism of UP's entrainment rate. Meanwhile, Appendix B elaborates on possible pathways that could conspire to produce such a symptom in the multiscale framework, and relevant sensitivity tests.

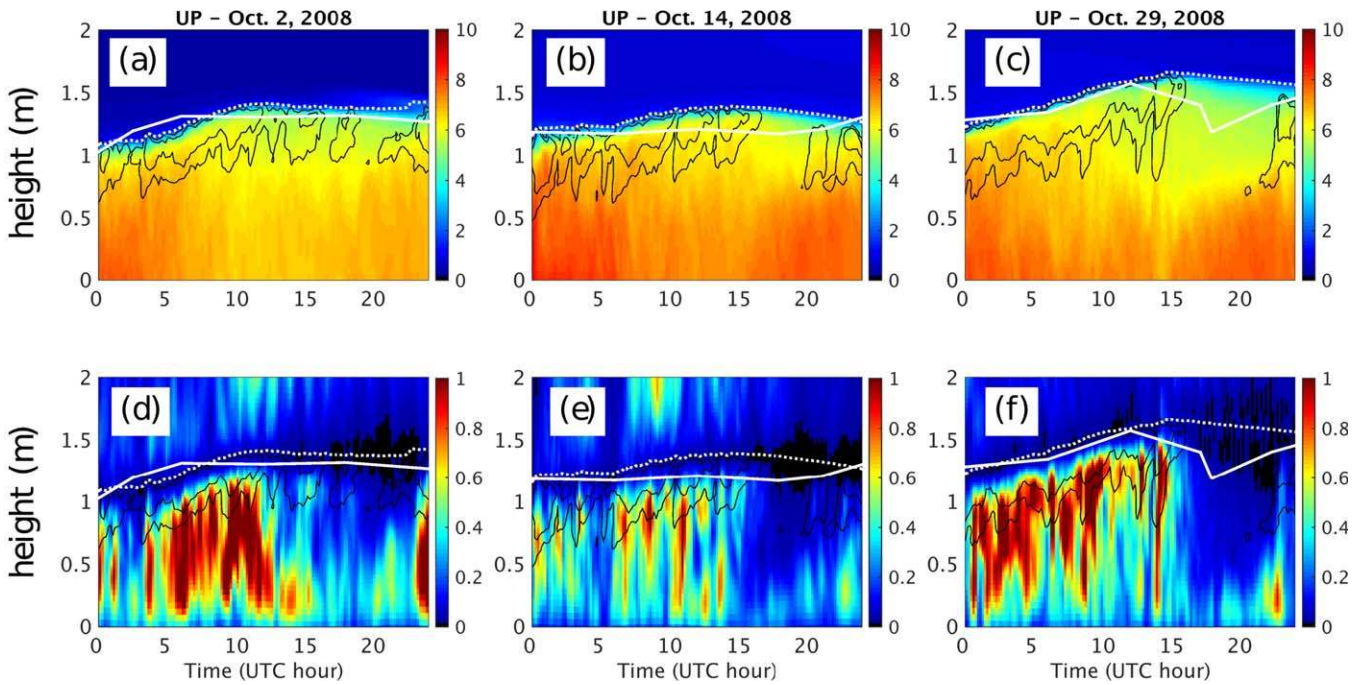


**Figure 19.** Top row: Height time evolution of the UP total water mixing ratio (in units of g/kg) at −19.9°N, 275.0°E for the days of (a) 2 October 2008, (b) 14 October 2008, and (c) 29 October 2008. Cloud liquid is contoured (solid black) at 0.01, 0.05, 0.1, and 0.2 g/kg. White lines track the inversion height via the 2 g/kg water vapor mixing ratio for the ECMWF analyses (solid) versus model (dotted). Bottom row: Same as top but for vertical velocity variance $\langle w'w' \rangle$ (in units of $m^2/s^2$).

### 7.2. Validation of CRM Mean Acceleration

Verification of the CRM-acceleration algorithm (see section 2.4) throughout this paper has been demonstrated in the form of similar predictions between two sets of UP ensemble hindcast simulations: one that used 2X CRM mean acceleration and one that did not. This has been shown for ensemble mean cloud fraction, buoyancy flux, and vertical velocity variance (e.g., Figure 8d versus Figure 8f, Figure 10d versus Figure 10f, and Figure 11d versus Figure 11f). For additional validation that the CRM mean acceleration does not substantially alter model predictions, we contrast ASR and OLR biases in Figures 4d, 4f, 5d, and 5f, respectively. Reassuringly, the global patterns of ASR and OLR biases and associated RMSE statistics of 2X accelerated runs (panels f) are also in good agreement with those of nonaccelerated simulations (panels d). Overall, these results affirm that the CRM mean acceleration can be successfully used to accommodate the high computational cost of UP.

## 8. Conclusion

We demonstrated the computational and technical feasibility of ultraparameterization, a superparameterized framework for global modeling in which the embedded CRMs have a fine enough grid (250 m $\Delta x$, 20 m $\Delta z$, and 1 s $\Delta t$) to explicitly simulate the outer scales of BL turbulence and associated clouds. To test the performance of various possible UP configurations relative to standard superparameterization (32 columns, 4 km $\Delta x$, 200 m $\Delta z$, and 20 s $\Delta t$), we analyzed top-of-atmosphere radiation, LWP and vertical cloud fraction profiles relative to satellite data in 12–36 h global hindcast pilot ensembles. We began with a global analysis, and then honed our attention on clouds and turbulence within the subtropical stratocumulus and trade cumulus regimes that UP is especially designed to better simulate. The 32-column UP version with the default MPDATA advection scheme somewhat outperformed a 3D 8 $\times$ 8-column configuration in these tests, despite stand-alone LES benchmark tests in which the 8 $\times$ 8-column configuration and a more advanced advection scheme produced thicker and more realistic Sc clouds. Having validated that the behavior of UP is not altered by the use of CRM acceleration, we performed accelerated 90 day free-running simulations contrasting 32-column UP with SP.

In both stratocumulus and trade cumulus regimes, UP simulates better vertical profiles of cloud cover, eliminating SP's tendency to misposition boundary layer clouds $\sim$500 m too low relative to colocated A-Train swath data. UP's better-resolved BL turbulence helps increase BL eddy vertical velocities to more realistic values. In combination with a better resolved inversion this helped induce realistically larger cloud fractions and higher entrainment rates in Sc regions. Unlike SP, UP is able to simulate the Cu-under-Sc decoupled boundary layers with a sharp inversion typical of observed conditions in the regions of maximum subtropical Sc cloud cover. Importantly for low-cloud feedback applications, the low clouds in UP respond much more realistically to variations in TOA radiation than in SP, as revealed by a significant improvement in the realism of 90 day composite diurnal LWP variability relative to satellite constraints.

The hindcast global fields and 3 month climatology of TOA reflected longwave radiation and precipitation are comparably skillful using UP as with SP, despite the 8 km wide UP CRM domain, which is too small for simulating deep convection and its associated high clouds. For shortwave radiation, UP enhances a global mean bright bias associated with increased low-cloud coverage in most oceanic regions, and exaggerates existing dim stratocumulus biases, despite the vertical structure and low-cloud fraction improvements. UP also develops a pronounced bright bias over the deep convecting tropical rainforest regions, which probably reflects excess shallow cumulus buildup due to the small UP CRM domain size, as previously noted in eight-column SP configurations [*Pritchard et al*., 2014]; such regional symptoms could be contained with a larger CRM domain extent.

In the near future, we plan to use a variant of the current UP implementation to simulate BL cloud response to climate perturbations such as an increase in SST and $CO_2$. Our motivation is to enrich diversity in model intercomparisons of low-cloud feedback. Since low-cloud-climate interaction is a complex problem of primary importance to climate sensitivity, it is prudent to employ a wide range of simulation techniques that use as diverse a range of trade-offs as possible for handling the issue of unresolved BL turbulence.

We acknowledge the value and proven merit of using a higher-order closure (HOC) of the sub-km scale under standard SP [*Cheng and Xu*, 2011; *Xu and Cheng*, 2013a, 2013b; *Cheng and Xu*, 2013a, 2013b; *Bogenschutz and Krueger*, 2013; *Bogenschutz et al*., 2013; *Cheng and Xu*, 2015; *Painemal et al*., 2015; *Wang et al*.,

2015] as another interesting complexity trade-off, with clear computational advantages. This improves important aspects of the physics and climatology of low clouds relative to SP, and provides a rich comparison of low-cloud feedback. We emphasize that our main interest in UP (and the justification for its computational expense) is that in relying entirely on explicit BL eddies to mediate its cloud feedback, it represents a new complexity trade-off, which may thus produce cloud responses to climate forcing that are different than other approaches, including HOC-augmented SP; if so, this would be worth knowing. Unsurprisingly, in its pioneering form, the approach also comes with some unintended consequences, such as a disappointing exaggeration of a time-mean dim bias in the near-coastal stratocumulus environment. However, UP already performs sufficiently in tests of current climate to merit preliminary application toward its intended aim.

In the longer term, further research is needed to improve the cloud-topped BL simulations in the embedded CRMs. This includes testing whether a coarser-resolution host GCM with larger, possibly 3-D, CRMs might be an attractive trade-off, further investigating other CRM numerical formulations and grids (especially that increasingly resolve the marine inversion) to see if they can more consistently maintain high Sc LWP, and finding stand-alone CRM tests that better predict the performance of a CRM configuration in an UP framework. To avoid unintended effects on deep convection, it could prove helpful to (1) expand the CRM domain extent or (2) introduce a "dual CRM" approach to help marry UP with a companion SP-like CRM responsible for handling convection outside of subsiding regimes in the subtropics. Further research is also needed on realizing the full potential of CRM acceleration, and testing software engineering strategies to enhance the parallel scalability of the multiscale modeling approach even farther than its current limit, for instance through MPI-decomposition or GPU-acceleration of embedded CRMs.

## Appendix A: Unintended Droplet Nucleation Feedbacks When Using Two-Moment Microphysics

We have used one-moment microphysics in this study. This is despite the fact that one of our eventual aims for UP is to use it to study aerosol-cloud feedback in a setting that includes explicit boundary layer eddies, which requires the use of two-moment microphysics. The reason is that our pilot tests in a two-moment configuration with prescribed aerosol produced an unrealistic increase in cloud droplet number concentrations leading to shortwave cloud brightening globally, especially over midlatitude oceans. One possibility is that the droplet nucleation schemes used in SPCAM5 may need updating to behave realistically when interfaced with updrafts that are more realistically intense, as in UP.

## Appendix B: Potential Causes of Overentrainment

In section 7.1.3 a symptom of over-entrainment was identified in UP simulations, with unknown cause. Here we speculate on possible causes and summarize results of relevant offline sensitivity tests.

*Advection*. One possible contributor to overentrainment is GCM-scale advection. Typical LES Sc studies are forced by a subsidence profile with large-scale vertical advection calculated locally. But in UP and SP, a total vertical advective heating and moistening forcing computed from the GCM is used, which limits the control of internal turbulence in selecting the inversion height. GCM-scale numerical diffusivity could be transmitted through vertical advection and might contribute to overentrainment.

Related issues are horizontal advection and inter-CRM interactions. Many offline LES tests assume a quasi-Lagrangian setup in which the grid is assumed to be horizontally translating with a representative parcel of air. This includes the DYCOMS case that inspired our LES grid design. But in SP/UP, the climatological southeasterlies near Peru can act to advectively warm/dry the subcloud through the GCM-to-CRM coupling terms. It is natural to wonder whether this has a first order control on UP's unsatisfyingly low Sc LWP, or if nonlocal feedbacks between incipient cloud layers at adjacent grid columns have the capacity for unintended advective effects not encountered in offline LES.

Consistent with any of these views, offline SAM—even when fully resolved using full 3-D LES resolution—reproduces the undersimulation of Peruvian Sc LWP when driven by total advective forcings derived from the full UPCAM simulation (not shown). Causative tests that isolate the role of horizontal advection by modifying the GCM-to-CRM coupling terms in UP would thus be useful to further test the role of advection in our results.

*Horizontal wind sensitivity*. We have also interrogated less obvious sources of discrepancy in investigating the differences in Sc LWP between UP and offline LES. Unexpectedly, we found that the simulated Sc are fairly sensitive to changing the horizontal wind speed profile used within the CRMs, even with specified surface fluxes. This form of sensitivity is not expected from offline LES cases—in which most experience from using fully resolved simulations indicate results are not strongly sensitive to the choice of U,V profiles. For instance, for DYCOMS cases, idealized linear or even zero wind profiles are routinely used interchangeably [*Wood*, 2012]. Insensitivity to horizontal wind profiles may be a familiar property of well-resolved LES of Sc that breaks down for our grey-zone resolution of 250 m × 20 m. More research on these LES-scale dynamics is underway and will be reported separately. This is needed to better inform effective use of grey-zone LES in SP simulations, in which the background wind field can depend sensitively on the assumed orientation of embedded CRMs and can change dramatically in space/time.

## Appendix C: Additional Insensitivities Related to Issues of SP Implementation

As a result of this work a variety of other aspects of SPCAM's formulation were interrogated. While none except for the above wind sensitivity has revealed a striking change in LWP suggesting the possible root cause of UPCAM's undersimulation, they are relevant to constraining the possible origins of the issue and deal with unreported aspects of SPCAM's formulation that may be relevant to others. Table C1 summarizes statistics for many of the following sensitivity tests. A caveat is that the following sensitivity tests focus on hours 0–24 of hindcasts initialized at 0 UTC (instead of hours 12–36 of hindcasts initialized 12 h earlier) so they do include a component of initialization shock.

*Formulation of CRM-scale advection*. Offline LES tests suggest that SAM is better able to capture Sc at marginal vertical grid resolutions when a higher order advection scheme (UM5) is used. However, porting a fifth-order advection scheme into UPCAM's CRMs did not immediately lead to any improvements in LWP. This again points to a broader issue as limiting at first order.

*Formulation of GCM-scale advection*. We tested UP in a legacy version of SPCAM that used a semi-Lagrangian dynamical core, and varied various GCM-scale advection parameters in the FV dycore (thanks to Peter Lauritzen for suggesting these tests). However, effects for Sc were found to be minimal. For instance, the global cloud radiative effect of a simulation with a positive definite advection constraint is tabulated in Table C1 showing a difference of ∼1 W/m$^2$ with baseline experiment in the mean OLR/ASR bias.

*Frequency of radiative transfer*. Since radiative cooling is one of the main players in driving Sc turbulence, we reduced the GCM time step (and associated radiative transfer frequency) from 5 min to an unusually small 60 s. Offline analysis of the full CRM state showed this had a desirable effect of improving the horizontal colocation of radiative cooling anomalies with incipient cloud patches by limiting the time available for horizontal translation to separate them through the multiscale coupling. However, overall LWP and associated cloud radiative properties were almost unchanged (see Table C1), perhaps since the overall driver of Cu-under-Sc is the horizontal mean CRM radiative cooling rate.

*Implementation of surface fluxes*. We discovered an issue in the order of operations through which surface fluxes are transmitted to the CRMs in SPCESM and its predecessors. Unintentionally, the hydrostatic

**Table C1.** Global Mean and RMSE Bias of OLR and ASR Relative to CERES-SYN on Day 1 for Various Methods That Are Tried in This Study But Proved to be Ineffective in Mitigating the Low LWP Bias in Sc Regions[a]

| | | OLR | | ASR | |
|---|---|---|---|---|---|
| Parameter | Case | Mean | RMSE | Mean | RMSE |
| | Baseline[b] | 2.65 | 18.22 | −7.59 | 40.90 |
| CRM dimension | C8 × 8 | 1.46 | 18.21 | −12.40 | 46.53 |
| CRM column width | 25 m | 3.33 | 21.05 | −10.30 | 54.90 |
| Advection scheme | UM5 | 3.11 | 18.93 | −8.46 | 41.05 |
| Radiation time step | 60 s | 3.11 | 18.78 | −6.77 | 40.01 |
| GCM advection constraint | Positive definite | 2.57 | 17.90 | −8.37 | 46.14 |

[a]All values are in $W/m^2$. For the data in this table, the simulations are initialized at 00 UTC 15 October 2008.
[b]Baseline is the C32-L125–250 m (UP) case (with MPDATA advection scheme and 300 s radiation time step using a piecewise parabolic GCM advection).

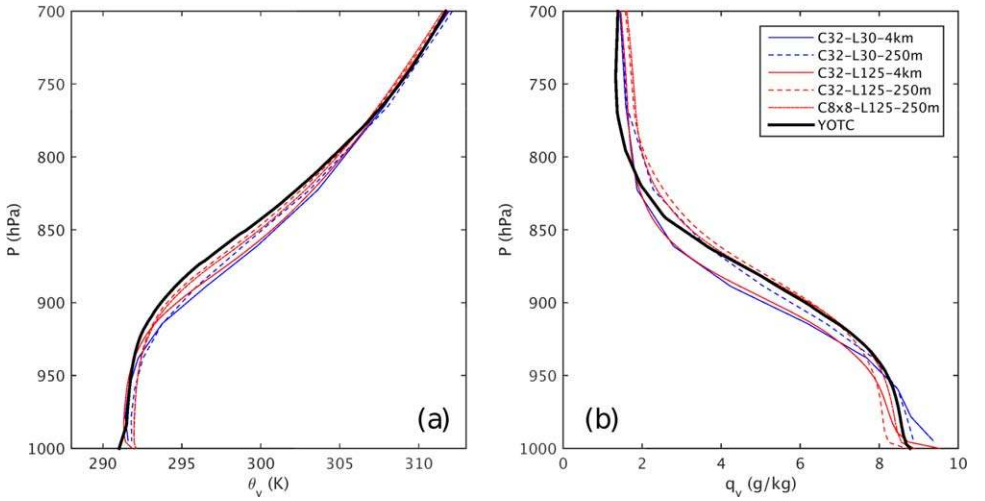


**Figure D1.** The ensemble mean thermodynamic profiles, averaged over the Peruvian region ($-28 < lat < -10$ and $270 < lon < 287$), is compared against that of YOTC reanalysis for the target days spanning October 2008. (a) Equivalent potential temperature and (b) specific humidity.

dynamical core was being exposed to incompletely adjusted surface fluxes concentrated as a tendency on the lowest model layer, prior to their transmission to the CRM which is meant to vertically diffuse them in lieu of a conventional boundary layer scheme. However, correcting this order of operations by directly feeding the surface fluxes to the CRMs had minimal effects on results.

*GCM-CRM inconsistency in wet/dry saturation mixing ratios*. An inconsistency between the use of GCM/CRM dry/wet saturation mixing ratios was discovered (thanks to Peter Blossey). Internal to SAM, dry mixing ratios (mass of water/mass of dry air) are used for saturation calculations while inside CAM moist mixing ratios (mass of water/mass of moist air) are used. The difference in the saturation mixing ratio resulting from these two methods may be as large as 1–2% for the model parameter range, with conceivably larger impacts on surface flux feedbacks. However, correcting this issue did not affect our model statistics of interest, consistent with evidence shown above against a prominent role for surface flux feedbacks in regulating our results.

## Appendix D

Figure D1 shows the ensemble mean vertical profiles of equivalent potential temperature and water vapor during the initial simulation day in the Peruvian region. Further details are discussed in section 4.2. Figure D2 shows the diurnal cycle of the low-cloud fraction. See details in section 6.2.

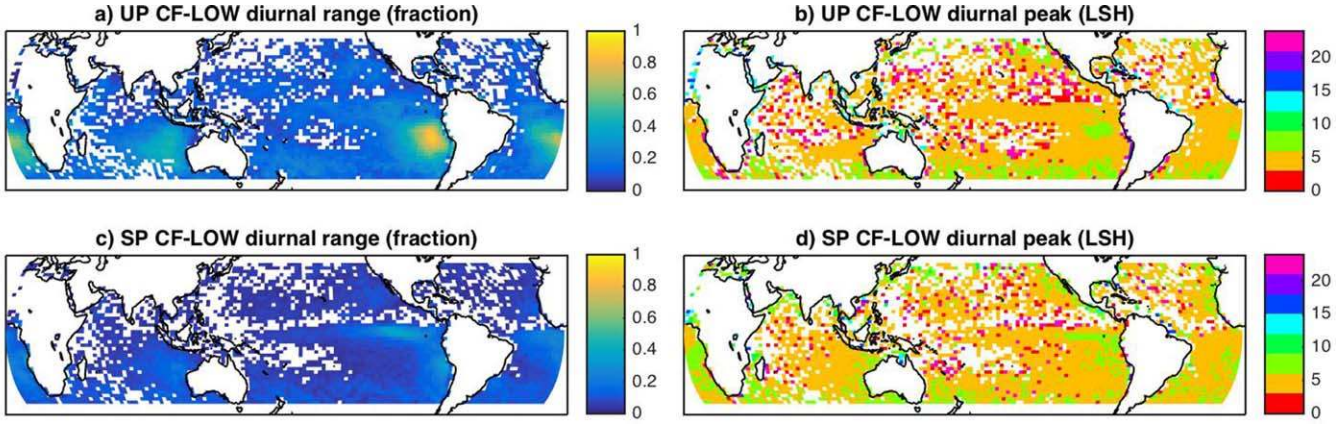


**Figure D2.** As in Figure 17 but for the model-diagnosed diurnal cycle of low-cloud fraction.

**Acknowledgments**
Funding for this study is provided by the Department of Energy under DE-SC0012548, DE-SC0012451, and DE-SC0012488. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under DE-AC02–05CH11231. All simulations were performed on NERSC Edison and Cori machines under repository No. m2222. The simulation data in this study is stored and can be made available to scientific researchers upon request. We acknowledge YOTC data provided on NCAR Yellowstone. The CCCM (C3M) and other CERES data are obtained from the NASA CERES using the NASA Langley Atmospheric Science Data Center (ASDC) ordering tool (http://ceres.larc.nasa.gov/). We gratefully acknowledge many useful conversations with Peter Blossey, Mark Branson, Peter Caldwell, Bill Collins, Cheryl Craig, Tony Craig, Brian Eaton, Jim Edwards, Andrew Gettelman, Steve Ghan, Wojciech Grabowski, David Hall, Walter Hannah, Cecile Hannay, Helen He, Isaac Held, Robert Jacob, Chris Jones, James Kent, Jean-Francois Lamarque, Peter Lauritzen, Roj Marchand, Georgios Matheou, Hugh Morrison, Jerry Olson, Robert Pincus, David Randall, Phil Rasch, Richard Rood, Daniel Rosenfeld, Masaki Satoh, Wayne Schubert, Balwinder Singh, Bill Skamarock, Joao Teixeira, Stefan Tulich, Paul Ullrich, Steve Vavrus, Hilary Weller, David Williamson, Patrick Worley, Shang-Ping Xie, and Takanobu Yamaguchi. We thank Minghuai Wang for developing and sharing the baseline version of SPCAM5, and the four anonymous reviewers for their constructive comments and suggestions.

## References

Ackerman, A., O. Toon, D. Stevens, and J. Coakley Jr. (2003), Enhancement of cloud cover and suppression of nocturnal drizzle in stratocumulus polluted by haze, *Geophys. Res. Lett.*, *30*(7), 1381, doi:10.1029/2002GL016634.

Benedict, J. J., and D. A. Randall (2009), Structure of the Madden-Julian Oscillation in the superparameterized CAM, *J. Atmos. Sci.*, *66*(11), 3277–3296, doi:10.1175/2009JAS3030.1.

Berner, A., C. Bretherton, and R. Wood (2011), Large-eddy simulation of mesoscale dynamics and entrainment around a pocket of open cells observed in VOCALS-REx RF06, *Atmos. Chem. Phys.*, *11*, 10,525–10,540.

Bogenschutz, P., and S. Krueger (2013), A simplified PDF parameterization of subgrid-scale clouds and turbulence for cloud-resolving models, *J. Adv. Model. Earth Syst.*, *5*, 195–211, doi:10.1002/jame.20018.

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

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

Bretherton, C., et al. (1999), An intercomparison of radiatively-driven entrainment and turbulence in a smoke cloud, as simulated by different numerical models, *Q. J. R. Meteorol. Soc.*, *125*, 391–423.

Bretherton, C., R. Wood, R. George, D. Leon, G. Allen, and X. Zheng (2010), Southeast Pacific stratocumulus clouds, precipitation and boundary layer structure sampled along 20 S during VOCALS-REx, *Atmos. Chem. Phys.*, *10*, 10,639–10,654.

Bretherton, C. S., and M. C. Wyant (1997), Moisture transport, lower tropospheric stability and decoupling of cloud-topped boundary layers, *J. Atmos. Sci.*, *54*, 148–167.

Burleyson, C., S. de Szoeke, S. Yuter, M. Wilbanks, and W. A. Brewer (2013), Observations of the diurnal cycle of Southeast Pacific marine stratocumulus clouds and precipitation, *J. Atmos. Sci.*, *70*, 3876–3894.

Cheng, A., and K. Xu (2008), Simulation of boundary-layer cumulus and stratocumulus clouds using a cloud-resolving model with low- and third-order turbulence closures, *J. Meteorol. Soc. Jpn.*, *86A*, 67–86.

Cheng, A., and K. Xu (2013a), Diurnal variability of low clouds in the Southeast Pacific simulated by a multiscale modeling framework model, *J. Geophys. Res. Atmos.*, *118*, 9191–9208, doi:10.1002/jgrd.50683.

Cheng, A., and K. Xu (2013b), Evaluating low cloud simulation from an upgraded multiscale modeling framework. Part III: Tropical and subtropical cloud transitions over the northern Pacific, *J. Clim.*, *26*, 5761–5781.

Cheng, A., and K.-M. Xu (2011), Improved low-cloud simulation from a multiscale modeling framework with a third-order turbulence closure in its cloud-resolving model component, *J. Geophys. Res.*, *116*, D14101, doi:10.1029/2010JD015362.

Cheng, A., and K.-M. Xu (2015), Improved low-cloud simulation from the Community Atmosphere Model with an advanced third-order turbulence closure, *J. Clim.*, *28*, 5737–5762.

Cheng, A., X. K-M., and B. Stevens (2010), Effects of resolution on the simulation of boundary-layer clouds and the partition of kinetic energy to subgrid scales, *J. Adv. Model. Earth Syst.*, *2*, 3, doi:10.3894/JAMES.2010.2.3.

Demott, C. A., D. A. Randall, and M. Khairoutdinov (2007), Convective precipitation variability as a tool for general circulation model analysis, *J. Clim.*, *20*(1), 91–112.

de Szoeke, S., et al. (2012), Observations of stratocumulus and their effect on the eastern Pacific surface heat budget along 20S, *J. Clim.*, *25*, 8542–8567.

Flato, G., et al. (2013), Evaluation of climate models, Chapter 9, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker, eds., Cambridge University Press, Cambridge, U. K. and New York.

Gettelman, A., and S. C. Sherwood (2016), Processes responsible for cloud feedback, *Curr. Clim. Change Rep.*, *2*, 179–189, doi:10.1007/s40641-016-0052-8.

Grabowski, W. (2001), Coupling cloud processes with the large-scale dynamics using the cloud-resolving convection parameterization (CRCP), *J. Atmos. Sci.*, *58*(9), 978–997.

Grabowski, W. (2016), Towards global large eddy simulation: Super-parameterization revisited, *J. Meteorol. Soc. Jpn.*, *94*, 327–344.

Gustafson, W., L. Berg, R. Easter, and S. Ghan (2008), The explicit-cloud parameterized-pollutant hybrid approach for aerosol-cloud interactions in multiscale modeling framework models: Tracer transport results, *Environ. Res. Lett.*, *3*, 025005.

Jones, C., C. Bretherton, and D. Leon (2011), Coupled vs. decoupled boundary layers in VOCALS-REx, *Atmos. Chem. Phys.*, *11*, 7143–7153.

Jones, C., C. Bretherton, and M. Pritchard (2015), Mean-state acceleration of cloud-resolving models and large eddy simulations, *J. Adv. Model. Earth Syst.*, *7*, 1643–1660, doi:10.1002/2015MS000488.

Kato, S., S. Sun-Mack, W. F. Miller, F. G. Rose, Y. Chen, P. Minnis, and B. A. Wielicki (2010), Relationships among cloud occurrence frequency, overlap, and effective thickness derived from CALIPSO and CloudSat merged cloud vertical profiles, *J. Geophys. Res.*, *115*, D00H28, doi:10.1029/2009JD012277.

Kato, S., et al. (2011), Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties, *J. Geophys. Res.*, *116*, D19209, doi:10.1029/2011JD016050.

Kay, J., et al. (2012), Exposing global cloud biases in the Community Atmosphere Model (CAM) using satellite observations and their corresponding instrument simulators, *J. Clim.*, *25*, 5190–5207, doi:10.1175/JCLI-D-11-00469.1.

Khairoutdinov, M., D. Randall, and C. DeMott (2005), Simulations of the atmospheric general circulation using a cloud-resolving model as a superparameterization of physical processes, *J. Atmos. Sci.*, *62*(7), 2136–2154.

Khairoutdinov, M., C. DeMott, and D. Randall (2008), Evaluation of the simulated interannual and subseasonal variability in an AMIP-style simulation using the CSU multi-scale modeling framework, *J. Clim.*, *21*(3), 413–431.

Khairoutdinov, M. F., and D. A. Randall (2001), A cloud resolving model as a cloud parameterization in the NCAR community climate system model: Preliminary results, *Geophys. Res. Lett.*, *28*, 3617–3620.

Khairoutdinov, M. F., and D. A. Randall (2003), Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities, *J. Atmos. Sci.*, *60*(4), 607–625.

Kooperman, G., M. Pritchard, M. Burt, M. Branson, and D. Randall (2016), Robust effects of cloud superparameterization on simulated daily rainfall intensity statistics across multiple versions of the Community Earth System Model, *J. Adv. Model. Earth Syst.*, *8*, 140–165, doi:10.1002/2015MS000574.

Lamarque, J.-F., et al. (2010), Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: Methodology and application, *Atmos. Chem. Phys.*, *10*(15), 7017–7039.

Marchand, R., and T. Ackerman (2010), An analysis of cloud cover in multiscale modeling framework global climate model simulations using 4 and 1 km horizontal grids, *J. Geophys. Res.*, *115*, D16207, doi:10.1029/2009JD013423.
Marchand, R., and T. Ackerman (2011), A cloud-resolving model with an adaptive vertical grid for boundary layer clouds, *J. Atmos. Sci.*, *68*, 1058–1074.
Marchand, R., J. Haynes, G. Mace, and T. Ackerman (2009), A comparison of simulated cloud radar output from the multiscale modeling framework global climate model with cloudsat cloud radar observations, *J. Geophys. Res.*, *114*, D00A20, doi:10.1029/2008JD009790.
Medeiros, B., and L. Nuijens (2016), Clouds at Barbados are representative of clouds across the trade wind regions in observations and climate models, *Proc. Natl. Acad. Sci. U. S. A.*, *113*, E3062–3070, doi:10.1073/pnas.1521494113.
Mellado, J. (2017), Cloud-top entrainment in stratocumulus clouds, *Annu. Rev. Fluid Mech.*, *49*, 145–169.
Moeng, C., S. Shen, and D. Randall (1992), Physical processes within the nocturnal stratus-topped boundary layer, *J. Atmos. Sci.*, *49*, 2384–2401.
Nicholls, S. (1984), The dynamics of stratocumulus: Aircraft observations and comparisons with a mixed layer model, *Q. J. R. Meteorol. Soc.*, *110*, 783–820.
Nuijens, L., I. Serikov, L. Hirsch, K. Lonitz, and B. Stevens (2014), The distribution and variability of low-level cloud in the North Atlantic trades, *Q. J. R. Meteorol. Soc.*, *140*, 2364–2374.
Nuijens, L., B. Medeiros, I. Sandu, and M. Ahlgrimm (2015), Observed and modeled patterns of covariability between low-level cloudiness and the structure of the trade-wind layer, *J. Adv. Model. Earth Syst.*, *7*, 1741–1764, doi:10.1002/2015MS000483.
Painemal, D., K. Xu, A. Cheng, P. Minnis, and R. Palikonda (2015), Mean structure and diurnal cycle of southeast Atlantic boundary layer clouds: Insights from satellite observations and Multiscale Modeling Framework simulations, *J. Clim.*, *28*, 324–341.
Park, S., and C. S. Bretherton (2009), The University of Washington shallow convection and moist turbulence schemes and their impact on climate simulations with the Community Atmosphere Model, *J. Clim.*, *22*, 3449–3469.
Pritchard, M. S., and R. C. J. Somerville (2009a), Empirical orthogonal function analysis of the diurnal cycle of precipitation in a multi-scale climate model, *Geophys. Res. Lett.*, *36*, L05812, doi:10.1029/2008GL036964.
Pritchard, M. S., and R. C. J. Somerville (2009b), Assessing the diurnal cycle of precipitation in a multi-scale climate model, *J. Adv. Model. Earth Syst.*, *2*, 12, doi:10.3894/JAMES.2009.1.12.
Pritchard, M. S., M. W. Moncrieff, and R. C. J. Somerville (2011), Orogenic propagating precipitation systems over the US in a global climate model with embedded explicit convection, *J. Atmos. Sci.*, *68*, 1821–1840.
Pritchard, M. S., C. S. Bretherton, and C. A. DeMott (2014), Restricting 32–128 km horizontal scales hardly affects the MJO in the Superparameterized Community Atmosphere Model v. 3.0 but the number of cloud resolving grid columns constrains vertical mixing, *J. Adv. Model. Earth Syst.*, *6*, 723–739, doi:10.1002/2014MS000340.
Randall, D., et al. (2007), Climate models and their evaluation, in Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, chap. 8, pp. 589–662, Cambridge Univ. Press, Cambridge, U. K.
Seifert, A., T. Heus, R. Pincus, and B. Stevens (2015), Large-eddy simulation of the transient and near-equilibrium behavior of precipitating shallow convection, *J. Adv. Model. Earth Syst.*, *7*, 1918–1937, doi:10.1002/2015MS000489.
Siebesma, A., et al. (2003), A large eddy simulation intercomparison study of shallow cumulus convection, *J. Atmos. Sci.*, *60*(10), 1201–1219.
Stevens, B., et al. (2003), On entrainment rates in nocturnal marine stratocumulus, *Q. J. R. Meteorol. Soc.*, *129*, 3469–3493.
Stevens, B., et al. (2005), Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus, *Mon. Weather Rev.*, *133*, 1443–1462.
Stubenrauch, C., et al. (2013), Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX radiation panel, *Bull. Am. Meteorol. Soc.*, *94*, 1031–1049.
Tompkins, A. M. (2001), Organization of tropical convection in low vertical wind shears: The role of cold pools, *J. Atmos. Sci.*, *58*, 1650–1672.
Van der Dussen, J. J., S. R. de Roode, A. S. Ackerman, P. N. Blossey, C. S. Bretherton, M. J. Kurowski, A. P. Lock, R. A. J. Neggers, I. Sandu, and A. P. Siebesma (2013), The GASS/EUCLIPSE model intercomparison of the stratocumulus transition as observed during ASTEX: LES results, *J. Adv. Model. Earth Syst.*, *5*, 483–499, doi:10.1002/jame.20033.
Waliser, D., et al. (2012), The year of tropical convection (May 2008–April 2010) climate variability and weather highlights, *Bull. Am. Meteorol. Soc.*, *93*, 1189–1218.
Wang, H., and G. Feingold (2009), Modeling mesoscale cellular structures and drizzle in marine stratocumulus. Part I: Impact of drizzle on the formation and evolution of open cells, *J. Atmos. Sci.*, *66*, 3257–3275, doi:10.1175/2009JAS3120.1.
Wang, M., et al. (2011), The multi-scale aerosol-climate model PNNL-MMF: Model description and evaluation, *Geosci. Model Dev.*, *4*, 137–168.
Wang, M., V. E. Larson,S. Ghan, M. Ovchinnikov, D. P. Schanen, H. Xiao, X. Liu, P. Rasch, and Z. Guo (2015), A multiscale modeling framework model (superparameterized CAM5) with a higher-order turbulence closure: Model description and low-cloud simulations, *J. Adv. Model. Earth Syst.*, *7*, 484–509, doi:10.1002/2014MS000375.
Williamson, D. L., J. G. Olson, and B. A. Boville (1998), A comparison of semi-Lagrangian and Eulerian tropical climate simulations, *Mon. Weather Rev.*, *126*, 1001–1012.
Wood, R. (2012), Stratocumulus clouds, *Mon. Weather Rev.*, *140*, 2373–2423.
Xu, K., and A. Cheng (2013a), Evaluating low-cloud simulation from an upgraded Multiscale Modeling Framework model. Part I: Sensitivity to spatial resolution and climatology, *J. Clim.*, *26*, 5717–5740.
Xu, K., and A. Cheng (2013b), Evaluating low-cloud simulation from an upgraded multiscale modeling framework model. Part II: Seasonal variations over the Eastern Pacific, *J. Clim.*, *26*, 5741–5760.
Yamaguchi, T., D. Randall, and K. M. F. (2011), Cloud modeling tests of the ULTIMATEMACHO scalar advection scheme, *Mon. Weather Rev.*, *139*, 3248–3264.
Zhang, Y., S. A. Klein, C. Liu, B. Tian, R. T. Marchand, J. M. Haynes, R. B. McCoy, Y. Zhang, and T. P. Ackerman (2008), On the diurnal cycle of deep convection, high-level cloud, and upper troposphere water vapor in the multiscale modeling framework, *J. Geophys. Res.*, *113*, D16105, doi:10.1029/2008JD009905.