The Horizontal Resolution Sensitivity of the Simple Convection-Permitting E3SM Atmosphere Model in a Doubly-Periodic Configuration

Peter Bogenschutz[1], Christopher Eldred[2], and Peter Caldwell[1]

[1]Lawrence Livermore National Laboratory
[2]Sanida National Laboratory

January 20, 2023

Abstract

We develop a doubly periodic version of the Simple Convection-Permitting E3SM Atmosphere Model (SCREAM) to provide an "efficient" configuration for this global storm resolving model (GSRM), akin to a single column model (SCM) often found in conventional general circulation models (GCMs). The design details are explained, in addition to the extensive case library associated with the doubly periodic SCREAM (DP-SCREAM) configuration. We demonstrate that doubly periodic cloud resolving models are useful tools to explore the scale awareness and scale sensitivity of GSRMs, in addition to replicating biases seen in the global models. Using DP-SCREAM, we show that SCREAM is a scale aware model as it is able to realistically partition between sub-grid scale (SGS) and resolved vertical transport across the gray zone of turbulence. We show that SCREAM is reasonably scale insensitive when run at resolutions from 1 to 5 km, but can exhibit sensitivity, particularly for the shallow convective regime, when run at resolutions approaching that of large eddy simulations. We conclude that SGS parameterization improvements are likely needed to reduce this scale sensitivity.

Horizontal Resolution Sensitivity of the Simple Convection-Permitting E3SM Atmosphere Model in a Doubly-Periodic Configuration

P. A. Bogenschutz[1], C. Eldred[2], and P. M. Caldwell[1]

[1]Lawrence Livermore National Laboratory, Livermore, CA, USA
[2]Sandia National Laboratory, Albuquerque NM, USA

Key Points:

- Doubly periodic configurations represent an efficient framework to supplement global CRMs
- SCREAM can credibly represent a variety of cloud regimes at a range of horizontal resolutions
- SCREAM is scale aware yet can exhibit scale sensitivity when run at very fine resolutions

Corresponding author: Peter A. Bogenschutz, bogenschutz1@llnl.gov

Abstract

We develop a doubly periodic version of the Simple Convection-Permitting E3SM Atmosphere Model (SCREAM) to provide an "efficient" configuration for this global storm resolving model (GSRM), akin to a single column model (SCM) often found in conventional general circulation models (GCMs). The design details are explained, in addition to the extensive case library associated with the doubly periodic SCREAM (DP-SCREAM) configuration. We demonstrate that doubly periodic cloud resolving models are useful tools to explore the scale awareness and scale sensitivity of GSRMs, in addition to replicating biases seen in the global models. Using DP-SCREAM, we show that SCREAM is a scale aware model as it is able to realistically partition between sub-grid scale (SGS) and resolved vertical transport across the gray zone of turbulence. We show that SCREAM is reasonably scale insensitive when run at resolutions from 1 to 5 km, but can exhibit sensitivity, particularly for the shallow convective regime, when run at resolutions approaching that of large eddy simulations. We conclude that SGS parameterization improvements are likely needed to reduce this scale sensitivity.

Plain Language Summary

Advances in computational resources have allowed climate simulations to be performed with very high resolution, which provides higher quality results. However, these simulations require a lot of time and computer resources to perform, which makes these models hard to use for the common scientist. In this paper we develop a high-resolution configuration which focuses on a specific point on the globe, enabling it to run fast and to use minimal computational resources. This allows users and developers to gauge how the model may perform before doing a computationally intensive global simulation. We show that this faster configuration is a useful tool to replicate problems that are found in the global model and is a valuable way to assess sensitivities of the model, particularly pertaining to choices made in its resolution.

1 Introduction

The next generation of general circulation models (GCMs) have arrived, taking the form of non-hydrostatic deep convection permitting global models. Pioneered more than fifteen years ago (Tomita et al., 2005; Satoh et al., 2008), the existence of global convection permitting models (GCPMs) has become increasingly commonplace within major modeling centers around the world and has recently culminated in the first intercomparison project of such models (Stevens et al., 2019). This progress is due to the rapid increase in computational power. These GCPMs are typically run with horizontal grid spacings of 1 to 5 km and do not have deep convective parameterizations, thus they rely on the dynamical core to represent motions associated with deep cumulus convection. There has even been recent activity of extending some of these GCPMs to use horizontal resolutions characteristic of Large Eddy Simulation (LES) either in a regionally refined context (Stevens et al., 2020) or multiscale modeling framework (Parishani et al., 2017); thus paving the way for potential global LES runs as we look towards the future.

One of the newest additions to the GCPM family is a 3 km model developed by the Department of Energy called the Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM; Caldwell et al. (2021)). A 40 day prescribed sea-surface temperature simulation (Jan 20-Feb 28, 2020) using an immature and untuned version of SCREAM demonstrates the benefits of moving to high horizontal resolution when compared to the conventionally parameterized Energy Exascale Earth System Model (E3SM) run with a comparatively coarse resolution of $\sim$100 km. Caldwell et al. (2021) reports that many long-standing biases typically associated with conventional GCMs are ameliorated simply by increasing the resolution to the kilometer scale; these include (but are not limited to) Amazon precipitation bias, frequency of light and heavy precipitation, the diurnal cycle of tropical precipitation, vertical structure of tropical convection, and coastal subtropical stratocumulus.

The results of Caldwell et al. (2021) and other recent high resolution modeling suggest many long-standing biases can simply be "resolved away". This is exciting and provides prospects for more accurate climate simulations to address pressing questions. However, as it currently stands, GCPMs are still expensive to run and typical simulations are both much shorter in duration and run less routinely than conventional GCMs. The relative increase in computational expense of GCPMs therefore introduces new challenges in terms of model debugging, parameterization implementation, evaluation, and tuning. Historically, many modeling centers have supported single column model (SCM) configurations for their conventionally parameterized GCMs (Bogenschutz et al., 2020; Gettelman et al., 2019) as a way to provide "rapid feedback" of model performance. These SCMs are often viewed as invaluable, if not essential, tools for model analysis and development (Bogenschutz et al., 2012; Park, 2014). However, a SCM is not appropriate for a GCPM since deep cumulus convection is expected to be resolved across multiple columns. While one could argue that a SCM could still be valid for boundary layer cloud regimes, the application of its use would be limited and situational.

Given the large computational expense of GCRMs, a SCM-like proxy is desired to facilitate fast feedback and encourage science that may be impractical using the global model. To minimize computational cost, options such as a regionally refined model (Tang et al., 2019) or limited area model (Giorgi, 2019) are often supported by modeling centers that provide GCPMs. However, regionally-refined and limited-area configurations are still relatively expensive to run and require expertise to set up for the desired region/regime of interest, so are not really suitable substitutions for SCM capability.

We argue that the so-called doubly-periodic CRM configuration is an ideal simple and efficient configuration for GCPM development. In this configuration, the model domain is configured on a cartesian planar grid with large-scale forcing provided from intensive observation period (IOP) field experiments and with lateral boundary conditions periodic in the x and y directions. Doubly-periodic cloud resolving configurations have already been used for many purposes including model validation against observations, gaining a better understanding of atmospheric processes, and to assess horizontal and vertical resolution sensitivity (Khairoutdinov & Randall, 2003; CHENG & XU, 2008; Krueger, 1988). Though doubly-periodic CRMs are certainly more expensive than a SCM would be in a conventional GCM, they are still significantly less expensive than running a GCRM and have been widely used by the community because of their usefulness and digestible computation cost. While there is a rich history of studies using doubly-periodic CRMs, such models have historically not been able to run in global configurations. However, global and doubly periodic configurations are mutually beneficial.

One of the most useful things a doubly-periodic CRM can be used for is to study the horizontal resolution sensitivity and the scale awareness of a GCPM in a computationally efficient manner (Bogenschutz & Krueger, 2013; Larson et al., 2012). Whereas traditional SCMs primarily exercise the GCM's physical components, a doubly-periodic CRM will exercise the model's full equation set (i.e. both physics and dynamics). In addition, it is trivial to configure a doubly-periodic CRM with a planar configuration to run with any desired domain size and resolution. This is counter to changing the resolution in a global model, which is often a time consuming process that requires expertise due to the generation of the necessary input files and configuration of the model to run at the new resolution.

Even if a particular GCPM is only configured to run at a particular resolution globally (i.e. 3 km), it is important to gain insights on the horizontal resolution sensitivity and scale awareness of the model to account for future changes in resolution and to ensure results are not dominated by discretization error. Should a particular model possess a large sensitivity in regards to the horizontal resolution, the doubly-periodic CRM can be an efficient vehicle to diagnose the cause while serving as a testbed to exploring modifications and potential parameterization deficiencies to reduce the sensitivity. In addition, a doubly-periodic CRM can be used to gauge if the GCPM is scale aware and diagnose any resolution limits. For instance, can a GCPM model that is configured to run at 3 km still be run at scales approaching that of $\sim$100 m without any necessary modifications to its equation set or changes in regards to the parameterizations used? As computational power increases and the resolutions of our GCPMs become progressively finer (as they already are in the aforementioned select works of Stevens et al. (2020) and Parishani et al. (2017)), these are critical questions we must ask of our GCPMs.

In this paper we introduce a doubly-periodic version of the SCREAM model (hereafter denoted as DP-SCREAM). While these doubly-periodic CRMs have existed for decades, this is the first time that the E3SM code base (in which SCREAMv0 was adapted from) has been modified to satisfy such a configuration. In addition, we use DP-SCREAM for five established and diverse cases to examine the horizontal resolution sensitivity as well as the scale awareness of the SCREAM model. In this paper we define "scale awareness" as the model's ability to adequately partition between the resolved and sub-grid scale transports as the resolution is modified. As an example, it is well established that processes associated with marine stratocumulus are largely sub-grid scale for a model with a horizontal resolution of 3 km (Cheng et al., 2010) and thus should be parameterized. As the horizontal resolution increases to that of LES ($\sim$100 m) the expectation is that the parameterized transport gradually shuts off and the resolved dynamics takes over. We define "scale insensitivity" to mean that as the model resolution changes, the representation of clouds and thermodynamics remains relatively robust. It is often considered a prerequisite that a model be scale aware in order to be scale insensitive (Bogenschutz & Krueger, 2013; Cheng et al., 2010; Larson et al., 2012). However, having a model that is scale aware does not guarantee that a model will be scale insensitive, as parameterization deficiencies, when run at relatively coarser resolutions, may degrade the simulation.

This paper is outlined as follows; section 2 will give a brief overview of the SCREAM model as well as introducing the DP-SCREAM configuration. In section 3 we discuss the cases run for our experiments as well as the range of horizontal resolutions we exploit DP-SCREAM to. Section 4 presents the results of these simulations to help us answer the questions of whether or not SCREAM is scale aware and scale insensitive. Finally, in section 5 we discuss the implication of our results for not only the SCREAM model, but for GCPMs at large.

2 Model Description

In this section we briefly discuss the SCREAM model (section 2.1) and give an overview of the doubly-periodic version of SCREAM (section 2.2).

2.1 SCREAM

The model version used in this study is very similar to SCREAMv0 as documented in Caldwell et al. (2021), so only a brief description given here. The development of SCREAM is designed to fulfill the US Department of Energy (DOE) mission of focusing on compute-intensive frontiers in climate science. The ultimate goal is to make SCREAM as computationally fast as possible on exascale machines by writing it in C++. However, the initial version of SCREAM, SCREAMv0, was written in Fortran using the existing E3SM atmosphere infrastructure. At the time of writing, the C++ (SCREAMv1) implementation is nearly ready for production runs but the infrastructure and abilities to run DP-SCREAM have not yet been converted to C++. Thus we use the Fortran version of SCREAM for this work.

The SCREAM model consists of nonhydrostatic fluid dynamics, a sub-grid scale (SGS) turbulence and cloud fraction scheme, a microphysics scheme, a radiation scheme, an energy fixer, and a prescribed-aerosol functionality. Specifically, the dynamical core uses the new nonhydrostatic version of the High Order Method Modeling Environment (HOMME-NH; Taylor et al. (2020)). The turbulence scheme is the Simplified Higher Order Closure (SHOC), which is a unified cloud macrophysics, turbulence, and shallow convective parameterization centered around a double-Gaussian assumed probability density function (PDF; Bogenschutz and Krueger (2013)). The microphysics scheme is based on the Predicted Particle Properties (P3) scheme of Morrison et al. (2015). The gas optical properties and radiative fluxes are computed using the RTE+RRTMGP radiative transfer package (Pincus et al., 2019). While Caldwell et al. (2019) used a prescribed-aerosol version of E3SM's modal aerosol model, the simulations used here employ an even simpler aerosol implementation that prescribes both cloud-condensation nuclei number and aerosol radiative properties from an E3SMv2 simulation. This new aerosol scheme is known as Simple Prescribed Aerosol (SPA).

2.2 Doubly Periodic SCREAM

The development of DP-SCREAM was broken into three pieces. First, the nonhydrostatic version of HOMME was extended to run on a planar domain. Second, infrastructure changes were needed to enable our code base to run on a domain of identically-forced columns and with the same location information. Third, the large library of cases and scripts developed for the E3SM SCM was extended to also work with DP-SCREAM.

2.2.1 Planar HOMME

The HOMME-NH dynamical core (Taylor et al., 2020) used in SCREAM solves the multicomponent compressible Euler equations in a rotating reference frame using Eulerian horizontal coordinates and a Lagrangian vertical coordinate, making the shallow atmosphere and traditional approximations. The lower boundary is a fixed material boundary, and the upper boundary is a (moving) constant pressure top material boundary. HOMME-NH uses mimetic finite differences (MFD) in the vertical with a Lorenz staggering and collocated compatible spectral elements (SEM) in the horizontal, and a vertical remapping for all variables to handle vanishing Lagrangian layer thickness.

Although in theory the SEM method works for arbitrary grids, the existing implementation was specialized to spherical grids. Therefore, in this work we extended the internal treatment of SEM to handle planar doubly periodic meshes. This involved the following changes to HOMME-NH:

1. Removed the dependence of the SEM derivative operators (divergence, gradient, curl, etc.) on spherical geometry and replaced with general versions valid for any geometry.
2. Added a planar doubly periodic mesh topology generation routine.
3. Added a uniform (constant $\Delta x/\Delta y$) planar doubly periodic mesh geometry generation routine.

By separating topology generation from geometry generation, it will be easy to add the ability to create non-uniform (but still topologically square) planar meshes in the future. Additionally, we implemented several commonly used planar test cases to validate the new model: the hydrostatic gravity wave (HGW) nonhydrostatic gravity wave (NHGW) and rising bubble (RB) tests from (Melvin et al., 2019). The main features still missing from the planar version of HOMME-NH are C++/Kokkos support, semi-Lagrangian advection of tracers and the ability to use separate grids for physics and dynamics (physgrid).

2.2.2 Infrastructure Design

Similar to the E3SM SCM, DP-SCREAM uses forcing files derived by IOPs to provide the necessary initial conditions, large-scale forcing, and surface fluxes (if available). DP-SCREAM makes extensive use of the existing E3SM SCM infrastructure, but with many modifications to suit the needs for this new configuration. Among these modifications is the need to make all SCM related routines work on multi-node parallelism. The E3SM SCM was coded with the intention that it would only be run on a single-processor, however it is essential that DP-SCREAM be run with multiple processors to ensure efficient run time. In addition, the interfaces of the atmosphere and land parallel input and output (PIO) routine also needed to be heavily modified to ensure that DP-SCREAM uses the same location and heterogeneous surface type throughout its domain for the particular case being run.

In DP-SCREAM, the domain size and horizontal resolution are determined by the user on the fly and the planar domain is set up to have the appropriate number of columns in the x and y direction to satisfy this. However, we still need to use E3SM domain files at initialization to determine on what point of the globe our domain will be set up at. By default, DP-SCREAM uses the files associated with ne30 resolution (corresponding to approximately 1° horizontal resolution) to determine the surface type of our domain, but not to initialize the atmospheric state. When the user submits a particular DP-SCREAM case the model uses the latitude and longitude specified in the IOP file to be the location for that particular run. This is to ensure consistent radiation computation across all columns, in addition to ensuring the correct surface type is used for that case.

The surface type is determined by searching the E3SM domain files. The grid cell in the ne30 file that is closest to the IOP latitude and longitude determines whether the model is operating over a land, ocean, or sea ice tile (or some combination/fraction of these). If operating over a land point, for example, then the land model is initialized identically for each column in the planar domain that matches the closest point to the IOP latitude and longitude. If operating over an ocean point then the data ocean model is initialized similarly. We note that at the time of this writing it is only possible to run DP-SCREAM with a data ocean model, as opposed to a fully interactive ocean. The land model can be run interactively or with surface fluxes specified (given they are provided in the IOP forcing file).

The atmosphere is initialized identically at all columns using the horizontal winds, temperature, and water vapor (u, v, T, and q, respectively) specified at the desired start time in the IOP forcing file. To spin up the turbulence, random perturbations are added to the initial profile of temperature in all cells below 900 hPa. The location and magnitude of these perturbations can be adjusted by the user in the namelist settings.

A new nudging routine has been added to provide the option of nudging DP-SCREAM to the IOP observations for u, v, T, and q. While the E3SM SCM has an existing routine to nudge to IOP observations, it is not suitable for use in DP-SCREAM where there are many active columns with large horizontal spatial variability. Considering only T (though treatment of q, u, and v are analogous), the horizontal domain average ($\overline{T}$) is computed at each level. The temperature relaxation is then computed at each model level as

$$\phi_T = -(\overline{T} - T_{obs})/\tau, \tag{1}$$

where τ is the relaxation time scale (set by default to 3 hr for DP-SCREAM, but easily modified by the user via namelist settings). The temperature at each grid point is then updated using this relaxation as

$$T_{forecast} = T_{before} + \phi_T * dt, \tag{2}$$

where dt is the model time step. We note that nudging is typically not turned on by default when using DP-SCREAM and its usage is case dependent (see section 2.2.3), with the ability to be switched on/off by the user via namelist option.

To account for the effects of subsidence or ascent from large-scale vertical velocity, which is often specified in the IOP forcing files, a simple routine was added to compute this effect on T, q, u, and v. Using T as an example (analogous for q, u, and v) this is computed as:

$$T_{forecast} = T_{before} - dt * \omega \left(\frac{dT}{dp} \right). \tag{3}$$

2.2.3 DP-SCREAM Case Library

The DP-SCREAM case library is shared with the E3SM SCM case library (Bogenschutz et al., 2020), which includes more than 25 cases ranging from continental and maritime deep convection to marine stratocumulus, mixed phase arctic clouds, and various flavors of shallow cumulus convection (see Tables 1 and 2 from Bogenschutz et al. (2020)). The IOP case library contains both well-established benchmark cases useful for gauging how SCREAM stacks up against other models as well as more modern cases for which novel observational constraints are available. For DP-SCREAM we have also added a radiative convective equilibrium (RCE, Wing et al. (2018)) case.

Our library is continuously growing and users can keep up-to-date on current case offerings and specifics by visiting https://github.com/E3SM-Project/scmlib/wiki/E3SM-Intensive-Observation-Period-(IOP)-Case-Library. At this location users can clone the Github repository to obtain scripts to run each case. These scripts are very similar to those developed for the E3SM SCM; we chose to provide and maintain separate scripts for each particular case rather than producing a universal script that can be used to run all cases, which would require hardcoding the specifics of each case into the SCREAM infrastructure as a particular run configuration (known as a “compset” in the CAM/E3SM parlance). We find that providing unique scripts for each case provides more transparency relative to compsets, which hide all settings from the average user.

Each script is set up to run with SCREAM's default 3.25 km horizontal grid spacing in the x and y direction and with the domain size that is most appropriate for that case (e.g. a larger domain for deep convection and smaller domain for boundary layer clouds). However, domain size and resolution can easily be modified by the user via the namelist.

3 Experiment Design

To help us determine whether SCREAM is scale aware and scale insensitive we run five cases spanning a range of cloud and convection regimes. In addition, we run these cases for horizontal grid spacings ranging from 100 m to 5 km. The exact choice of resolutions we select to run, including the domain size, is dependent on the actual case. For instance, cases which include deep cumulus convection will need to be run with a much larger domain than cases consisting primarily of boundary layer clouds. The specific domain size and resolution that is run for each case is mentioned in the case description and summarized in Table 1. For this study we include two cases of deep cumulus convection (one maritime and the other continental), one shallow cumulus case, one marine stratocumulus case, and one mixed-phase arctic cloud case.

We note that the primary motivating factor for this work is to assess the impacts of horizontal resolution sensitivity. Therefore, to eliminate any potential ambiguity in the results relating to time step, we choose to run each case with the same time step settings for all resolutions. All cases are run with SCREAM's standard 128 vertical levels as documented in Caldwell et al. (2021). We note that all cases at all resolutions are run with the exact same code base, tuning parameters, and parameterization suite.

3.1 ARM97 - Continental Deep Cumulus Convection

The Atmospheric Radiation Measurement (ARM) 1997 IOP occurred at the ARM southern great plain (SGP) site in June and July 1997. Similar to the predecessor IOP which took place in the summer of 1995, the ARM97 case features several distinct periods characterized by a wide range of summertime weather conditions. The data from this IOP formed the basis for the ARM/GCSS case SCM and CRM intercomparison (Xu et al., 2002; Xie et al., 2002). The forcing data was developed using the constrained variational analysis method described in Zhang and Lin (1997) and Zhang et al. (2001). This case features time varying forcing and prescribed surface sensible and latent heat fluxes. For this case, we nudge the u and v winds to observations using a three-hour time scale.

The IOP forcing file in the DP-SCREAM library is a 26 day case, however, here we focus on an 8-day active period featuring several strong deep convective events, starting on 23 June, 1997. This case is run in a horizontal domain of 200 km (Khairoutdinov & Randall, 2003) in the x and y directions and run with $\Delta x = \Delta y = 500$ m, 800 m, 1.5 km, 3 km, and 5 km. The physics and model time step is 50 s, while the dynamics time step is 2 s for all resolution settings. We note that due to the domain size and run duration of this case, we are unable to run with $\Delta x = \Delta y = 100$ m like we do for our boundary layer cloud cases. However, performing DP-SCREAM simulations at these fine resolutions for deep convection cases is something we plan to pursue in future work.

3.2 GATE - Maritime Deep Cumulus Convection

To simulate maritime deep cumulus convection we run the Global Atmospheric Research Program's Atlantic Tropical Experiment (GATE, Houze Jr. and Betts (1981)), phase III. GATE was an extensive field experiment that took place over the tropical Atlantic Ocean with the goal to improve the basic understanding of tropical convection and its role in the global atmospheric circulation. This case has been used extensively for CRM and LES related studies (Khairoutdinov et al., 2009; Fu et al., 1995; Xu et al., 1992). GATE features time varying forcing with surface sensible and latent heat fluxes computed interactively. For this case, we nudge the u and v winds to observations using a three-hour time scale.

The IOP forcing file in the DP-SCREAM library includes all 20 days of the GATE phase III, starting at 00Z 30 August 1974, and we run the case in its entirety. Similar to the ARM97 case, the horizontal domain is 200 km in the x and y directions and run with $\Delta x = \Delta y = 500$ m, 800 m, 1.5 km, 3 km, and 5 km. The physics and SCREAM time step is 50 s, while the dynamics time step is 2 s for all resolution settings.

3.3 RICO

The Rain in Cumulus over Ocean (RICO) field study (Rauber et al., 2007) is used to simulate precipitating maritime shallow convection. The RICO case is based on composite measurements over the trade-winds in the western Atlantic Ocean and the DP-SCREAM case setup follows that of the LES intercomparison study by vanZanten et al. (2011). The simulation starts with a 740 m deep sub-cloud mixed layer topped by a conditionally unstable layer. This case features steady state forcing with time constant sea surface temperature prescribed and the sensible and latent heat fluxes computed interactively.

As per vanZanten et al. (2011) the RICO case is run for a duration of 24 hours. This case is run in a horizontal domain of 50 km in the x and y directions and run with

$\Delta x = \Delta y = 100$ m, 500 m, 800 m, 1.5 km, 3 km, and 5 km. Unlike the ARM97 and GATE cases, RICO (as well as the remaining boundary layer cases to be described) uses a smaller domain and shorter run duration, which affords us the ability of running with horizontal grid spacings of 100 m. This allows us, as with the remainder of the boundary layer cloud cases to be described, to test how SCREAM handles simulations within the gray zone of turbulence. The physics and SCREAM time step is 10 s for this case, while the dynamics time step is 0.33 s. No nudging is applied for RICO simulations presented in this paper.

3.4 DYCOMS-RF01 Subtropical Marine Stratocumulus

The first research flight (RF01) of the second Dynamics and Chemistry of Marine Stratocumulus (DYCOMS; Stevens et al. (2003)) field study is used to evaluate SCREAM's ability to simulate subtropical marine stratocumulus. DYCOMS-RF01 was a nocturnal research flight that took place in marine stratocumulus west-southwest of San Diego, California in July 2001 and was the basis of an LES intercomparison study (Stevens et al., 2005). This case is often considered the gold-standard to evaluate a model's ability to adequately represent marine stratocumulus, as it is characterized by presence of mean conditions that some theories suggest should have dissipated the cloud deck. Thus, it is considered a difficult case to simulate with fidelity as the tendency for most models and parameterizations - and even some LES - is to dissipate the cloud (Stevens et al., 2005; Zhu et al., 2005).

As per Stevens et al. (2005) we run this case for a duration of four hours. This case features steady state forcing and prescribed surface sensible and latent heat fluxes. While the LES intercomparison of DYCOMS-RF01 and the official E3SM SCM case set up for RF01 (Bogenschutz et al., 2020) turns off the microphysics scheme, we choose to leave P3 active for this case so as to adequately test the scale sensitivity of the entire SCREAM cloud physics suite. This case is run in a horizontal domain of 50 km in the x and y directions and run with $\Delta x = \Delta y = 100$ m, 500 m, 800 m, 1.5 km, 3 km, and 5 km. The physics and SCREAM time step is 10 s for this case, while the dynamics time step is 0.33 s. No nudging is applied for DYCOMS-RF01 simulations presented in this paper.

3.5 MPACE-B

The Mixed-Phase Arctic Cloud Experiment (MPACE) was conducted from 27 September through 22 October 2004 over the Department of Energy's Atmospheric Radiation Measurement (ARM) Climate Research Facility on the North Slope of Alaska (Verlinde et al., 2007). The primary objectives of this field campaign were to collect a dataset suitable to study the interactions between microphysics, dynamics, and radiative transfer in mixed-phase Arctic clouds. We run the MPACE-B case, which represents a 12 hour subset of a cold-air outbreak single-layer mixed phase cloud case and was the basis for an intercomparison study featuring more than a dozen CRMs and SCMs (Klein et al., 2009). Klein et al. (2009) found that virtually all models underestimated the cloud liquid water of this case by nearly a factor of three when compared to observations.

MPACE-B features steady-state forcing and prescribed surface sensible and latent heat fluxes. This case is run in a horizontal domain of 50 km in the x and y directions and with $\Delta x = \Delta y = 100$ m, 500 m, 800 m, 1.5 km, 3 km, and 5 km. The physics and SCREAM time step is 10 s, while the dynamics time step is 0.33 s. No nudging is applied for MPACE-B simulations presented in this paper.

4 Results

We will present results starting with our deep convection cases, ARM97 and GATE, respectively. Following this we will present results for our boundary layer cloud cases for shallow cumulus convection (RICO), marine stratocumulus (DYCOMS-RF01), and then finally mixed-phase stratocumulus (MPACE-B).

While the focus of this paper is to assess the horizontal resolution sensitivity of SCREAM we will also compare the quality of the overall simulations to observations or large eddy simulations, where available.

4.1 ARM97 - Continental Deep Cumulus Convection

Our ARM97 experiment design is described in section 3.1. Figure 1 displays the evolution of the domain averaged precipitable water and surface precipitation rate for all resolution configurations compared to observations taken from the 1997 ARM summer IOP field campaign at the SGP site over the eight day simulated period. The observed temporal evolution of the precipitable water and surface precipitation are generally well captured by the model configurations, with a few exceptions. All resolution configurations are too moist during the third day and all configurations seem to generate too much precipitation during most of the convective events.

However, the main intent of this paper is not a rigorous comparison of DP-SCREAM simulations with observations but to examine the horizontal resolution sensitivity. In this regard, we see the various resolution configurations are fairly robust as the differences due to resolution are negligible compared to the differences relative to observations. The largest differences for surface precipitation seem to occur during the first large convective event. Though, we do note that the 500 m run tends to be a slight outlier in terms of the precipitable water and is slighter "wetter" than the rest of the simulations for the majority of the eight day run.

Figure 2 displays the horizontally and temporally averaged cloud profiles for the entirety of the eight day simulation. While all simulations show the same general characteristics in terms of the cloud fraction (Fig. 2a), the 500 m run shows sensitivity in regards to the low-level (below 3 km altitude) cloud amount, whereas the 1.5 km simulation is an outlier in regards to the upper-tropospheric clouds. The low-level cloud sensitivity is further demonstrated when looking at the liquid cloud mixing ratio profiles (Fig. 2b). Whereas the simulations with $\Delta x\ > 1$ km are robust within the boundary layer, the 800 and 500 m simulations produce significantly less cloud liquid, with a monotonic decrease as resolution increases. In addition, we see some sensitivity of cloud liquid between the model simulations at the mid-levels (3 km to 6 km in altitude), corresponding to cumulus congestus.

Figure 2c displays the averaged cloud ice profiles. It is important to note that the P3 microphysics scheme includes snow in the cloud ice mixing ratio. In general, we see a reasonable agreement between the various configurations in terms of the magnitude, however it is clear that the 500 m simulation has a peak of cloud ice that is a bit lower in altitude when compared to the simulations at the kilometer scale. These results suggest that SCREAM is reasonably scale insensitive in regards to the representation of cloud properties for this case when run with grid sizes greater than 1 km, though there is an apparent sensitivity in its representation of shallow convective clouds when run with finer resolutions. This sensitivity will be explored more when we analyze the results of GATE and RICO.

To gain an understanding of the scale awareness of SCREAM for continental deep convection, we examine the total moisture flux and how the partitioning is represented across scales. Figure 3a depicts the total moisture flux ($\overline{w'q_t'}$), which represents the sum of the resolved and sub-grid scale (SGS) contributions. Here we see robust agreement between the various resolution configurations, with some minor differences within the mid and lower layers of the troposphere. This is the type of behavior we would hope to see for a model that is scale insensitive. Examining how this total flux is partitioned between resolved and SGS contributions as the resolution changes, however, will give us insights on the scale awareness.

Figures 3b and 3c display the SGS and resolved contributions of $\overline{w'q'_t}$, respectively. Unlike the total $\overline{w'q'_t}$, we would expect there to be differences in the partitioning of the SGS and resolved fluxes as resolution changes (Cheng et al., 2010). Indeed, we see that for the simulations with $\Delta x > 1$ km SHOC is responsible for parameterizing the majority of the vertical transport of water within the boundary layer. However, for the 800 and 500 m resolutions, there is an inherent scale separation where the boundary layer turbulence and shallow convection is partially resolved and partially subgrid-scale. This is encouraging behavior as it demonstrates the hallmarks of a scale aware model in its ability to naturally partition between SGS and resolved processes as the resolution increases. This will be explored more in the following sections, especially when we examine boundary layer cloud cases where we are able to run at 100 m horizontal resolution to capture the full spectrum of the boundary layer gray zone.

The horizontally and temporally averaged differences in temperature and moisture, computed relative to observations at the ARM SGP site, are shown in Figure 4. The largest spread between the model configurations for temperature is in the lower troposphere. Generally, as model resolution increases the differences with observations becomes smaller. While it is intuitive that model performance is generally expected to become better as resolution increases, this also points to the need of possible improvements for the boundary layer parameterization (SHOC) to reduce the differences with observations at the coarser resolutions. The moisture differences between resolution configurations, however, generally tend to be greater in the mid-levels often where cumulus congestus is found. This is consistent with the relatively large spread between cloud liquid water at these levels.

Finally, figure 5 displays the temporal evolution of the horizontally-averaged top-of-atmosphere shortwave cloud forcing (SWCF) and longwave cloud forcing (LWCF) for our five DP-SCREAM resolution configurations. SWCF and LWCF are two important metrics in climate models that are often tuned as resolution changes to maintain radiation balance. Thus, the hope is that SWCF and LWCF are minimally scale insensitive so that time-consuming retuning of the model is not necessary should the resolution of the global model change. While the phase between each simulation is generally in agreement, there are definitely differences in magnitude for some individual convective events, particularly for SWCF during the first three days. The mean SWCF and LWCF values for the simulated period are depicted in Table 2, which shows that the 500 m run is generally the outlier for SWCF, with the remainder of simulations generally within a few W/m^2 of each other.

4.2 GATE - Maritime Deep Cumulus Convection

Switching to maritime deep cumulus convection we focus on the GATE case, for which our experiment design is described in section 3.2.

The evolution of the horizontally averaged total precipitable water and precipitation rate can be found in Fig. 6. Observations for precipitable water are not available for this case, therefore we focus on the sensitivity to horizontal grid spacing. Unlike the ARM97 case, where each resolution configuration was fairly robust for the representation of precipitable water, here we see an apparent sensitivity when moving from 3 km resolution to 500 m, where a monotonic increase occurs.

When examining the precipitation rate (Fig. 6b) we find that the 500 and 800 m simulations generally have good agreement with the estimated observed amount, albeit slightly overestimated. The remainder of the simulations, especially 3 km and 5 km, overestimate the precipitation rate and show an apparent oscillatory behavior that does not appear to be physical. This is interesting because Caldwell et al. (2021) found that SCREAM's representation of tropical convection at 3.25 km was not very realistic as it was generally unable to aggregate and could potentially be related to the oscillatory behavior we see in GATE. Caldwell et al. (2021) found that SCREAM tends to produce an abundance of precipitation clusters that are too small in size and with excessive rain rates when compared to observations. A more detailed process oriented study to investigate this problem could efficiently be carried out by DP-SCREAM in future work.

The temporally and horizontally averaged profiles of cloud fraction, liquid water mixing ratio, and ice mixing ratio are displayed in Fig. 7. In terms of upper tropospheric clouds, while we see some differences in terms of the magnitude of cloud fraction and ice mixing ratio, the simulations are generally characteristically similar. The largest sensitivity in terms of the horizontal resolution resides in the low-level clouds and we note that this sensitivity appears to be larger than that found in our continental convection case. Unlike ARM97, which showed a robust representation of low clouds for our 1.5, 3, and 5 km cases (Fig. 2), in GATE we see a near monotonic decrease in the cloud amount mixing ratio and depth of the shallow clouds as the resolution increases for all of our experiments. This could represent an apparent sensitivity of SCREAM in the representation of tropical shallow clouds, which will be explored in greater detail in section 4.3.

Figure 8 displays the temporally averaged profiles of the total moisture flux ($\overline{w'q_t'}$) as well as the SGS and resolved components. The expectation for a model that is scale insensitive is that the total flux (Fig. 8a) is robust when the resolution changes. Generally, we do see very good agreement between all of our simulations, however, the 500 m simulation does exhibit a lower magnitude in the lower-to-mid troposphere when compared to the coarse resolution simulations. This is generally in agreement with our findings of the 500 m simulation producing fewer clouds and precipitation.

The SGS and resolved profiles of $\overline{w'q_t'}$ are displayed in Figures 8b and c, respectively. In terms of the SGS component, we see similar behavior to that of the continental deep convective case, where the simulations with $\Delta x > 1$ km are very robust and with an apparent scale separation happening when the resolution is reduced to 500 m, indicating that SCREAM's SGS parameterization is doing less work as the resolution increases. In terms of the resolved transport, we see near equal contributions being produced by each configuration whereas the expectation is for the magnitude to increase as the resolution increases. This suggests that the lower resolution configurations could artificially be too strong in the resolved scales, possibility due to underactive SGS representation and could be a contributor towards the resolution sensitivity seen in the low clouds (Bogenschutz & Krueger, 2013).

Profiles of the differences in observed temperature and moisture over the twenty day run can be found in Fig. 9. In terms of temperature differences, there is remarkable agreement between each of the simulations which feature a warm bias in the lower troposphere, relatively unbiased mid-levels, and warm bias in the upper troposphere. Much larger sensitivity to resolution can be found when examining the differences in water vapor, especially in the boundary layer. While all simulations exhibit a fairly strong dry bias, the 500 m run is considerably more moist, which is in agreement with the analysis presented in Fig. 6. In general, Fig. 9b demonstrates, in addition to other analyses presented for this case, that the 500 m run has more skill in representing tropical convection when compared to the lower resolution counterparts run at 800 m to 5 km, which are resolutions that GCRMs are typically run at.

Finally, Figure 10 displays the temporal evolution of LWCF and SWCF. In agreement with our analysis of ice clouds in Fig. 7, we see very little scale sensitivity with respect to LWCF, suggesting that high clouds in the tropics may not need significant retuning as SCREAM's resolution is increased. In regards to the SWCF, we do see a bit more sensitivity to horizontal grid size for individual convective events, yet averaged values for SWCF and LWCF (Table 3) show fairly minimal scale sensitivity for simulations run with $\Delta x > 1$ km, with a bit larger sensitivity for the 800 and 500 m simulations.

4.3 RICO - Subtropical Precipitating Shallow Convection

Though SCREAM's default resolution of 3.25 km allows circulations associated with deep convection to be permitted, motions associated with shallow convection are still largely unresolved at this resolution (Cheng et al., 2010) and remains a challenge for GCPMs. The results from ARM97 and GATE suggest that SCREAM has a resolution sensitivity when representing shallow convection. Therefore, we focus on Rain in Cumulus Over Ocean (RICO), which represents a maritime precipitating shallow convective regime (experiment design described in section 3.3). Unlike our previously examined GATE and ARM97 cases, we can afford to run a 100 m resolution case, which puts us in the range of what is typically considered to be a large eddy simulation (LES). Where available, we compare DP-SCREAM simulations against the LES mean and spread from vanZanten et al. (2011). In that study, they show that the LES ensemble average could plausibly reproduce the characteristics of the observed clouds, and thus we treat LES as a numerical benchmark for this case.

The time evolution of the vertically integrated low cloud and cloud liquid water path is presented in Fig. 11. The LES mean and spread is characterized by a short spin-up period at the start of the simulation, which quickly transitions to a quasi-steady state of ∼20 percent cloud cover and 20 g/m^2 of vertically integrated cloud water. In terms of the DP-SCREAM simulations, while it is apparent that all resolutions can adequately produce vertically integrated cloud water and cover that is characteristic of a shallow convective regime, there are some key differences between the simulations. The first is that the coarser resolution simulations, chiefly 3 km and 5 km, tend to produce higher values of vertically integrated cloud fraction and liquid water when compared to the higher resolution simulations. Secondly, the coarser resolution simulations also seem to suffer from a longer spin-up time with more of a quasi-oscillatory behavior when compared to the more steady state solutions provided by the high resolution DP-SCREAM simulations and LES. Though the high resolution 500 and 100 m DP-SCREAM simulations tend to achieve a steady state solution, we note that these simulations underestimate the vertically integrated low cloud and liquid water.

More differences between the simulations emerge when we examine the vertical structure of the clouds in Fig. 12. While all DP-SCREAM simulations produce cloud fraction and cloud liquid water magnitudes that are characteristic to that of shallow cumulus, there is sensitivity in regards to the vertical structure of the clouds. It is clear that the coarse resolution simulations (3 and 5 km) tend to simulate clouds that are too shallow in vertical depth, with cloud tops that are nearly 1 km lower when compared to the LES ensemble. These results are consistent with the global simulation analysis presented in Caldwell et al. (2021). In fact, it is not until the resolution is increased to 100 m in DP-SCREAM when the representation of the vertical structure of the clouds is satisfactory. Furthermore, vanZanten et al. (2011) reports that most LES members have a double peak in cloud fraction and liquid water, one near cloud base and one near cloud top, which is not evident in any of the DP-SCREAM simulations.

The temporally averaged profiles relating to the thermodynamic structure are presented in Fig. 13, which shows clear differences among the DP-SCREAM simulations and compared to the LES ensemble. All SCREAM simulations are able to capture the well mixed sub-cloud layer and are in general agreement with LES, the exception being the 100 m run which is too dry. The larger differences occur within the cloud layer, which is not surprising given the differences in the vertical extent of clouds in Fig. 12. It is obvious that the 100 m simulation is the only DP-SCREAM experiment that is able to adequately capture the sharp increase in static stability, as compared to the initial profile

(Figure 4 in vanZanten et al. (2011)), while the remainder of the simulations struggle to break through the conditionally unstable layer. This is likely a result of the SGS parameterization to provide adequate countergradient gradient fluxes that cannot be compensated by the dynamics due to the coarse resolution of the 500 m to 5 km simulations, which cannot resolve the large eddies associated with shallow cumulus.

The total moisture flux profiles and their partitioning between SGS and resolved components are presented in Figure 14. In terms of the simulation of the total flux, results are consistent with the behaviors presented in Figures 12 and 13, where only the 100 m simulation has reasonable agreement with the LES ensemble. Though, it should be noted that the 500 m to 5 km simulations are reasonably robust in their representation of total $\overline{w'q_t'}$, even if the quality is poor. In addition, we find that even though DP-SCREAM struggles to capture some of the quantitative aspects of the trade cumulus regime with grid sizes of 500 m to 5 km, we do find that SCREAM is scale aware for this regime, depicted by the partitioning of SGS and resolved turbulent transports. This is encouraging, but also strongly suggests that the quality of coarse resolution simulations and scale sensitivity across the gray zone could be improved by addressing issues of SGS representation of the shallow cumulus regime.

4.4 DYCOMS-RF01 Subtropical Marine Stratocumulus

As described in section 3.4, we run the first research flight (RF01) of the second Dynamics and Chemistry of Marine Stratocumulus (DYCOMS; Stevens et al. (2005)) to examine SCREAM's ability to simulate marine Sc using a wide range of resolutions. For this case, where available, we compare our results to the LES mean and spread that is presented in Stevens et al. (2005).

The horizontally averaged time evolution of the vertically-integrated low cloud amount and liquid water path can be found in Fig. 15. For both of these variables, it is clear that SCREAM resolution sensitivity is small compared to inter-model spread in LES. All SCREAM simulations are able to maintain a near solid cloud deck, which was observed, throughout the four hour simulation (Fig. 15a). There is a bit more spread in the liquid water path (Fig. 15b) as the 800 and 500 m simulations have more cloud liquid than the remainder of the simulations, yet overall SCREAM simulations still have much less spread when compared to the LES ensemble.

Figure 16 displays the horizontally and temporally averaged profiles from the last simulated hour for the SCREAM simulations and LES (for cloud liquid). We generally see good agreement among the SCREAM simulations, though with some subtle characteristic differences. For instance, the simulations with $\Delta x > 1$ km tend to simulate a more solid cloud deck when compared to the simulations with $\Delta x < 1$ km. In terms of the cloud liquid water, all SCREAM simulations produce more cloud when compared to the LES mean (Stevens et al., 2005) and generally falls within uncertainty of observations. This is particularly impressive since SCREAM simulations have vertical resolution that is much coarser compared to that used in LES. The satisfactory simulation of marine Sc by DP-SCREAM is in agreement with the results of Caldwell et al. (2021).

The representation of the liquid water potential temperature ($\overline{\theta_l}$) and total water mixing ratio ($\overline{q_t}$) vertical structures are presented in figure 17. While all SCREAM simulations are able to reasonably produce the well mixed vertical structure when compared to the LES mean and spread, there are some differences. For example, all SCREAM simulations tend to be a bit warmer than LES within the boundary layer and the simulations with $\Delta x > 1$ km do not appear to be as well mixed in regards to $\overline{q_t}$ within the surface level. In addition, there are differences near the boundary layer top in terms of the thermodynamics structure for the coarser simulations relative to the simulations with $\Delta x < 1$ km.

While the simulation of cloud characteristics for this case is relatively scale insensitive, the resolution range between 100 m and 5 km represents a large theoretical gap for this regime (Cheng et al., 2010); thus we need to determine if SCREAM can gracefully handle the transition between parameterized and resolved turbulence. Figure 18 displays the total, SGS, and resolved moisture flux for all SCREAM resolutions and LES (for the total flux). While the agreement for the total $\overline{w'q_t'}$ is reasonable and mostly falls within the LES ensemble window, we do note that the simulations with $\Delta x < 1$ km tend to have a stronger flux throughout the depth of the boundary layer when compared to the coarse resolution simulations.

What we expect to be very different in the SCREAM simulations is how the SGS and resolved fluxes are partitioned as we move across scales. Figures 18b and c demonstrate that for the simulations with $\Delta x > 1$ km nearly all of the turbulent transport is provided by the sub-grid scale SHOC parameterization, with little resolved. This is to be expected for this case given the scale analysis of Cheng et al. (2010). As we move to the 800 and 500 m resolutions we see that we are clearly within the gray zone of turbulence, with vertical transport partially resolved and partially SGS. Furthermore, at our LES-like horizontal resolution of 100 m we note that nearly all turbulence is resolved, with the exception of near the surface. The fact that SCREAM is able to adequately partition between parameterized and resolved turbulence across scales without any adjustments to the code, tunable parameters, or changes to the parameterization suite is very encouraging, with the benefits further discussed in the summary and discussion (section 5).

4.5 MPACE-B - Mixed Phase Arctic Clouds

Many climate and weather models tend to have difficulty simulating the observed frequency and persistence of Arctic mixed-phase clouds (e.g. Morrison and Pinto (2006)), thus we simulate MPACE-B (as described in section 3.5) to determine the scale awareness and sensitivity of SCREAM for this challenging cold-air outbreak case. Klein et al. (2009) presented results from an MPACE-B intercomparison study with many participating cloud resolving models (CRMs). They found not only a large spread among the CRMs, but that a large majority of these models underpredicted the liquid water path by a factor-of-three, though models with sophisticated microphysics agreed better with the observed values of liquid and ice water path.

Table 4 represents the average values of cloud liquid and cloud ice water paths, over hours 4 through 12, for SCREAM simulations and the observed values reported in Klein et al. (2009), as well as the median value from the CRMs used in that study. In terms of liquid water path, we see that SCREAM simulations do not suffer from the large underestimate that plagued the CRM intercomparison study. Instead, SCREAM simulations tend to slightly overestimate this value. The 5 km and 100 m cases, which represent opposite ends of the resolution spectrum, are the outliers that slightly underestimate the liquid water path when compared to observations. In terms of the ice water path, SCREAM simulations have little resolution sensitivity and slightly overestimate ice mass. It is important to note, however, that P3 microphysics includes suspended ice and snow as one species, which is likely contributing to the higher values reported here.

The time evolution of the ice and cloud water paths are displayed in Fig. 19. In terms of the ice water path, all DP-SCREAM simulations have the same general characteristics, displayed by ice water that tends to increase over time. There is much more spread in terms of the liquid water path. As depicted by the averaged cloud liquid water values, it is clear that the 5 km and 100 m cases are relative outliers. The 100 m case simulates much less cloud liquid water, compared to the rest of the simulations, that appears to slowly deplete over time. While this spread is much greater than that seen with the marine stratocumulus case, recall that the spread of SCREAM simulations is still less than that reported by the CRM intercomparison of Klein et al. (2009) and that all simulations are in reasonable agreement with observations. However, it is a bit disconcerting that no apparent convergence with resolution is found.

The vertical structure of the observed and simulated cloud fraction is presented in Fig. 20. The observed cloud fraction (Fig. 20a) is provided by two aircraft flights and ground based radar/lidar averaged over hours 4 through 12 of the case. The simulated cloud fraction profiles (Fig. 20b) show the SCREAM results as well as the mean and spread of the CRM intercomparison study. While the simulated SCREAM cloud profiles are fairly robust, it is clear that SCREAM tends to simulate cloud base and height at a higher altitude than the CRM envelope. While it is not surprising that the characteristics of the simulated cloud profile are different between SCREAM and the CRM intercomparison, given the vastly different simulated liquid water paths, it is not clear if the SCREAM cloud profile agrees better with observations versus that of Klein et al. (2009). Though SCREAM simulates a cloud deck too high in altitude compared to the aircraft observations, most simulations agree well with the radar/lidar profiles. The exception is the outlier 100 m case, which simulates a cloud deck too high in altitude compared to all observational sources.

Figure 21 displays the vertical structure of the temporally averaged cloud liquid and cloud ice mixing ratios for the SCREAM simulations. In general, we see reasonable agreement for simulations in the 5 km to 500 m range, though we note a subtle shift of the cloud deck upward in altitude as the resolution increases. This is made most apparent when comparing the cloud liquid profiles of the 5 km simulation with the 100 m simulation. The 100 m simulation produces much less liquid water compared to most other simulations and with a much higher simulated cloud deck; both of which appear to be at odds with observations. Unfortunately, this is counterintuitive as we expect the simulation quality to improve as resolution increases. Thus, SCREAM's ability to simulate mixed phase clouds across scales clearly needs to be examined in greater detail. Nevertheless, the ability of SCREAM to simulate cloud liquid amounts at all resolutions that are in decent agreement with observations is encouraging. While profiles of cloud ice all show the same general characteristics, there are differences related to the progressively increasing altitudes as diagnosed. In addition, the $\Delta x < 1$ km simulations tend to produce more ice than the $\Delta x > 1$ km simulations.

The thermodynamic profiles for each resolution are displayed in figure 22. In general, the 100 m simulation is the clear outlier for both temperature and moisture when compared to the rest of the simulations. The 100 m simulation is characterized by a reduction of q_t in the upper half the boundary layer, which Klein et al. (2009) note in nearly all their CRM simulations and comment that this is unrealistic behavior, due to the underestimate of q_l. In this case, the coarser resolution simulations of 3 km and 5 km tend to produce the most well mixed boundary layer structures and the q_t profiles appear to be the most realistic given that we would expect a smaller jump between q_t in the cloud layer versus that in the sub-cloud layer (Klein et al., 2009). In terms of θ_l, the coarser resolution simulations, which tend to contain more cloud liquid water and less ice precipitation, act to keep the boundary layer more well mixed. This is counter to the 100 m simulation which has less cloud top cooling and more ice precipitation acting to show larger vertical gradients.

The temporally averaged total, SGS, and resolved moisture fluxes ($\overline{w'q_t'}$) are displayed in Fig. 23. While there is somewhat reasonable agreement in the total flux for most simulations (Fig. 23a), the 5 km simulations appears to be the outlier within the sub-cloud and cloud layer. When breaking down into components of SGS (Fig. 23b) and resolved (Fig. 23c) transports, we see the desired shift of energy from the SGS to resolved scales as the resolution increases. However, without energy spectra filter results from LES it is difficult to ascertain whether the magnitudes for each grid size is representative of what we expect or not and should be explored in future work.

The fact that a significant portion of transport is being carried out by resolved scales at the 5 km and 3 km simulations may be a bit unrealistic for a boundary layer cloud case such as MPACE-B and could potentially point to deficiencies in SCREAM's SGS turbulence scheme to handle this regime. Nonetheless, it is interesting that the simulations with greater SGS contributions (1.5 km, 3 km) tend to produce the most realistic cloud and thermodynamic structures. The 100 m simulation, on the other hand, which relies very little on SGS transport, tends to have the most unrealistic simulation which could point to deficiencies that need to be addressed with the microphysics scheme. There is potential that deficiencies exist in both the microphysics and turbulence schemes and compensating errors between the two are leading to more acceptable solutions for the 0.8 to 3 km range.

Nonetheless, while SCREAM certainly struggles in some aspects to simulate the MPACE-B case across scales, the satisfactory solution of clouds at many resolutions (including the default SCREAM resolution of 3.25 km) and less spread when compared to the CRM study of Klein et al. (2009) is encouraging.

5 Summary and Discussion

In this paper we develop a doubly periodic version of the SCREAM model that we call DP-SCREAM. Since SCREAM is a GCPM, it is therefore far more computationally expensive when compared to conventional GCMs, which often support very efficient configurations known as single column models to aid in model analysis at the process level and debugging of model development. Thus, DP-SCREAM fills the need and serves as SCREAM's proxy for a SCM-like utility, which offers rapid feedback of model performance at the process level.

One of the major benefits of SCREAM is that it allows the user to choose the model domain and grid size on the fly. This is unlike changing the resolution of a global model, which often requires time consuming generation and testing of the necessary input files. In DP-SCREAM, this is trivial and allows for one to easily explore the resolution sensitivity of the model. While SCREAM is currently run globally with a grid spacing of 3.25 km in the horizontal, higher resolutions are expected as computational capabilities advance. Even in the near term, scientists will likely run SCREAM with regional mesh refinement that pushes towards the sub-kilometer scale within their regions of interest. DP-SCREAM can give indications of how SCREAM will perform at these resolutions and what tuning (or code) requirements may be necessary as the resolution is pushed into uncharted territories.

In this paper we run DP-SCREAM for five cases, spanning a range of cloud regimes, and for grid sizes ranging from 100 m to 5 km. It is common for GCPMs to be run with horizontal grid sizes anywhere from 1 to 5 km. However, we choose to run many of our simulations at scales in which large eddy simulations are typically run to see how SCREAM handles moving across the "gray zone" of turbulence. In this work, we seek to answer two questions: 1) Is SCREAM scale aware and 2) is SCREAM scale insensitive?

Is SCREAM scale aware? We believe that we can conclusively say that SCREAM is a scale aware model based on the fact that it can reasonably partition between resolved and SGS transports as we move across scales. As an example, using the results from the marine stratocumulus case of DYCOMS-RF01, at 3 and 5 km resolution the vertical transport of moisture is almost completely parameterized by SHOC. As the resolution is increased towards 100 m, SHOC gradually shuts off while allowing resolved dynamics to take over. This general behavior is true for all cases examined in this paper. Whether or not the partitioning between SGS and resolved scales is of the correct magnitude for each grid size and case is an open question, however, and subject to further research and analysis using observations and LES.

This scale awareness is significant because it means that we can increase the resolution of SCREAM (either globally or in RRM mode) without the need of manually shutting off or swapping out parameterizations, avoiding any tricky ambiguities typically associated with gray zone modeling. It also means that it may be possible to use DP-SCREAM as an LES process model, though much validation would be needed to make this a reality. We believe that any GCPM using a PDF-based parameterization such as SHOC, Cloud Layers Unified by Bi-normals (CLUBB) (Golaz et al., 2002), or Intermediately Prognostics HOC (CHENG & XU, 2008) would likely be scale aware. However, many GCPMs use simple turbulence closures that were intended to be used at LES scales (Khairoutdinov & Randall, 2003) and are generally not scale aware at CRM resolutions (Bogenschutz & Krueger, 2013). We encourage other modeling centers to investigate the scale awareness of their GPCMs using a doubly periodic configuration.

Is SCREAM scale insensitive? Unfortunately, this question cannot be answered as cleanly as the former question as it appears to be regime dependent. While all cases experience at least some degree of scale sensitivity, stratiform clouds are the least sensitive to horizontal resolution. This is especially encouraging for the DYCOMS-RF01 case as the LES intercomparison study (Stevens et al., 2005) found a large sensitivity between the participating members, whereas the SCREAM model was generally robust while moving across scales.

The largest sensitivity in SCREAM is associated with the shallow convective regime. This is demonstrated by the results of RICO and the ARM97 and GATE cases when examining the lower troposphere; though it appears to be particularly exacerbated for tropical and subtropical oceanic cases. For simulations with $\Delta x > 1$ km SCREAM tends to produce shallow clouds that contain too much cloud water and are too shallow in depth. In other words, they appear to have characteristics of broken stratocumulus clouds. This is in agreement with the preliminary global assessment presented in Caldwell et al. (2021), so on one hand it is encouraging that DP-SCREAM can replicate biases seen in the global model. As the resolution decreases to $\Delta x < 1$ km for these cases, DP-SCREAM tends to simulate shallow convective clouds that are in better agreement with LES and observational reference, with clouds that are deeper in vertical extent. This suggests that SHOC, which serves as SCREAM's parameterization for shallow cumulus, should be improved or tuned to reduce the scale sensitivity seen in shallow convective regimes.

In general, the simulation of deep cumulus convection is relatively robust when SCREAM is run with $\Delta x > 1$ km, as represented by results of precipitation, top of atmosphere radiative fluxes, and upper-level clouds. However, some sensitivity can be seen when SCREAM is run with $\Delta x < 1$ km. This sensitivity appears to be introduced when SCREAM goes from a grid spacing of 1.5 to 0.8 km, representing an apparent scale separation. Indeed, this is often the gap where deep convection goes from being merely "permitted" to "resolved". The results related to our findings on scale sensitivity means that SCREAM will likely need some degree of retuning if resolution is changed globally, but it would likely be modest if that resolution is kept between 1 and 5 km and more substantial if resolution is reduced below 1 km.

While the priority of this paper was not to produce an in detail analysis of DP-SCREAM with observations or LES, we note that SCREAM does a credible job of simulating a wide range of cloud regimes as it seems to perform as well as previous CRM studies. The exception, however, is shallow cumulus convection. In addition, we note the general behavior that SCREAM simulations do appear to get better as resolution increases. While this result is not surprising, the exception to this rule appears to be for mixed phase Arctic clouds. For the MPACE-B case the majority of the simulations agree quite well with observations, yet the 100 m case produces cloud that is generally considered to be too high in altitude with smaller than observed cloud liquid water values. This apparent and unusual sensitivity could suggest a need to examine potential deficiencies with the microphysics treatment within SCREAM. In addition, while the 100 m RICO simulation produces much better cloud characteristics and thermodynamic structure than the lower resolution simulations, when compared to LES, it suffers from a dry bias in the representation of cloud liquid water.

Using a doubly periodic configuration of SCREAM we were able to address many questions pertaining to the SCREAM model. By running a particular model at a range of resolutions it not only provides information for how that model may perform globally at different resolutions, but also provides key information for its default resolution. Thus, we highly encourage other modeling centers to develop and support doubly periodic configurations for their GCPM.

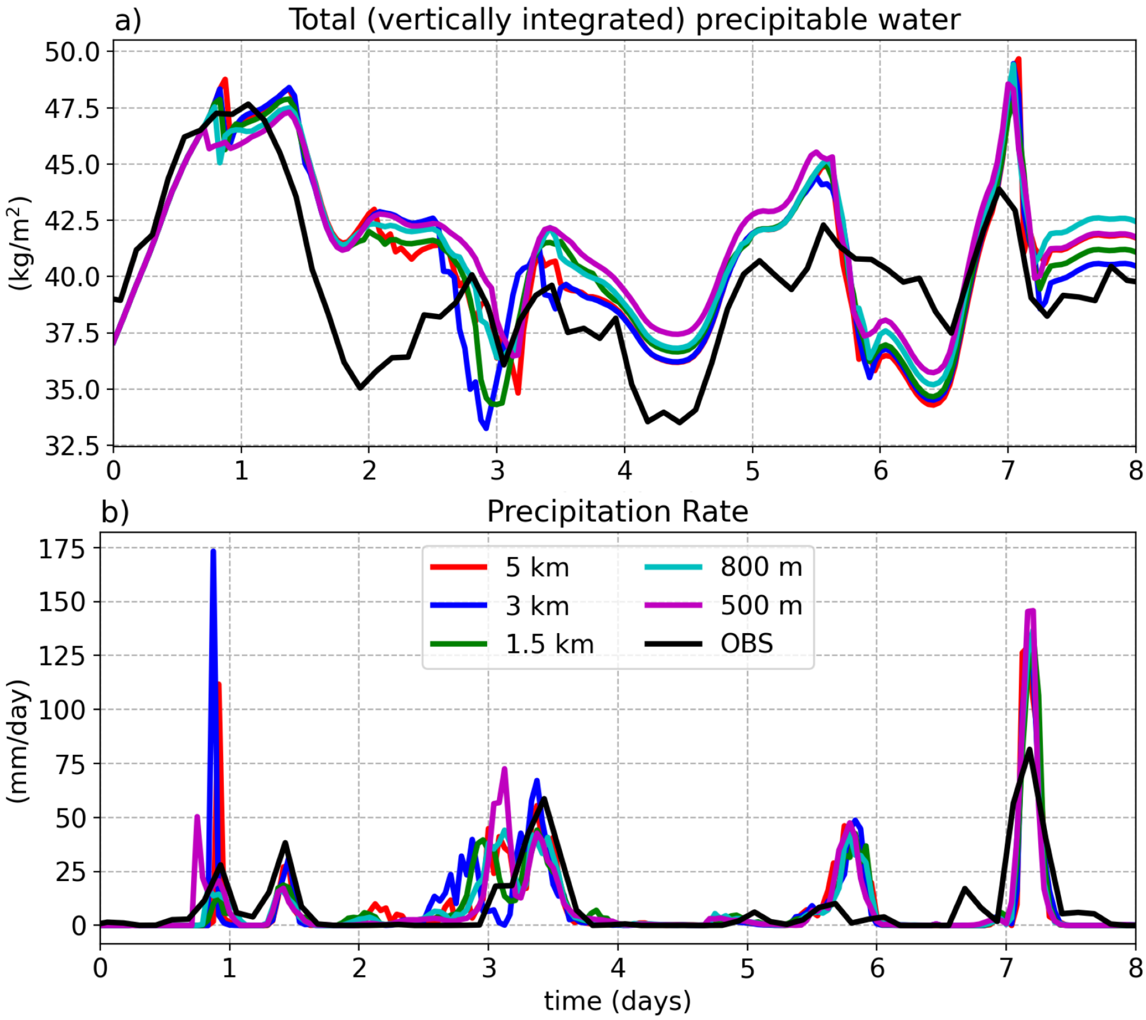

Figure 1. Temporal evolution of the horizontally averaged a) total vertically integrated precipitable water and b) precipitation rate over the 8 day simulation of the ARM97 case starting at 00Z 23 June 1997 for SCREAM simulations and observations taken at the Southern Great Plains (SGP) site.

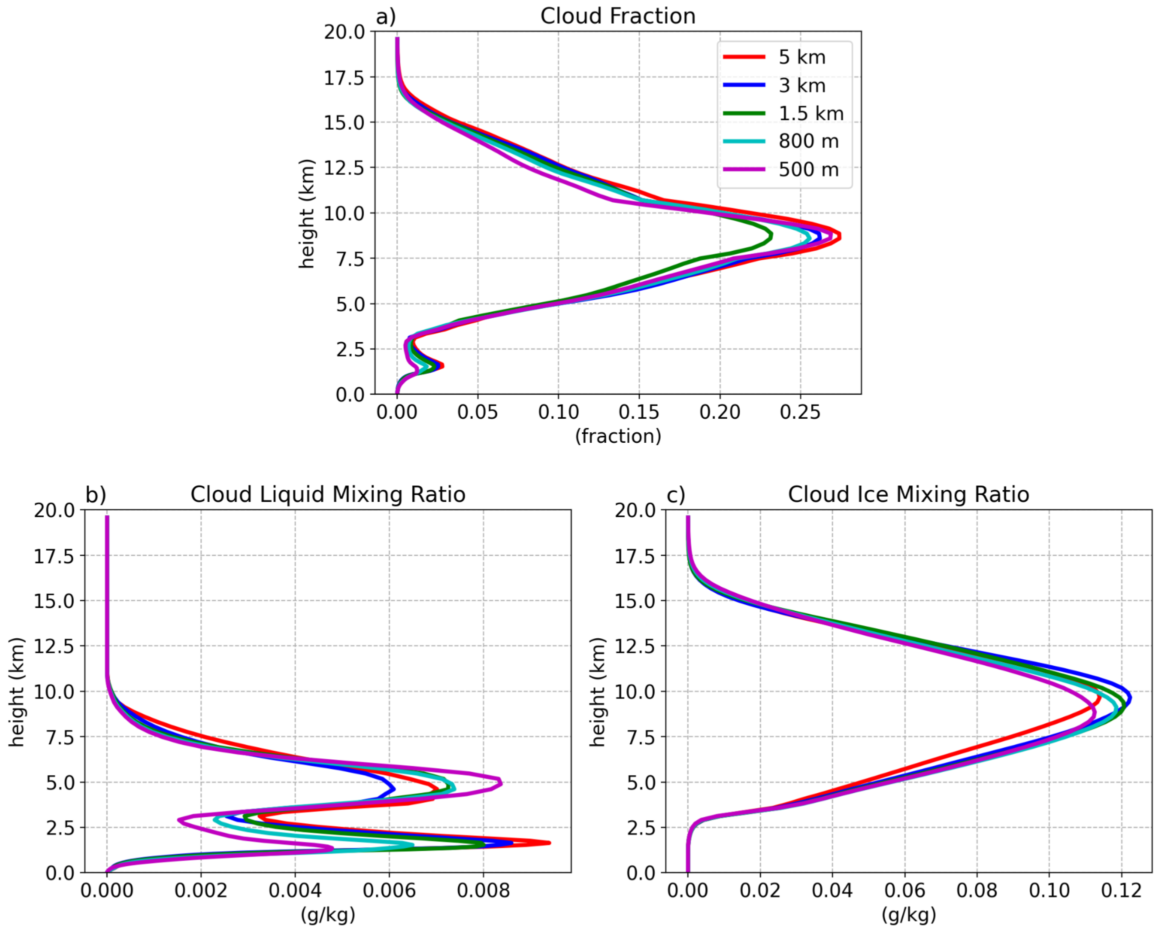

Figure 2. Temporally and horizontally averaged profiles of a) cloud fraction, b) cloud liquid mixing ratio, and c) cloud ice mixing ratio for the SCREAM simulations at various horizontal resolutions for the ARM97 case averaged over the entire eight day simulation.

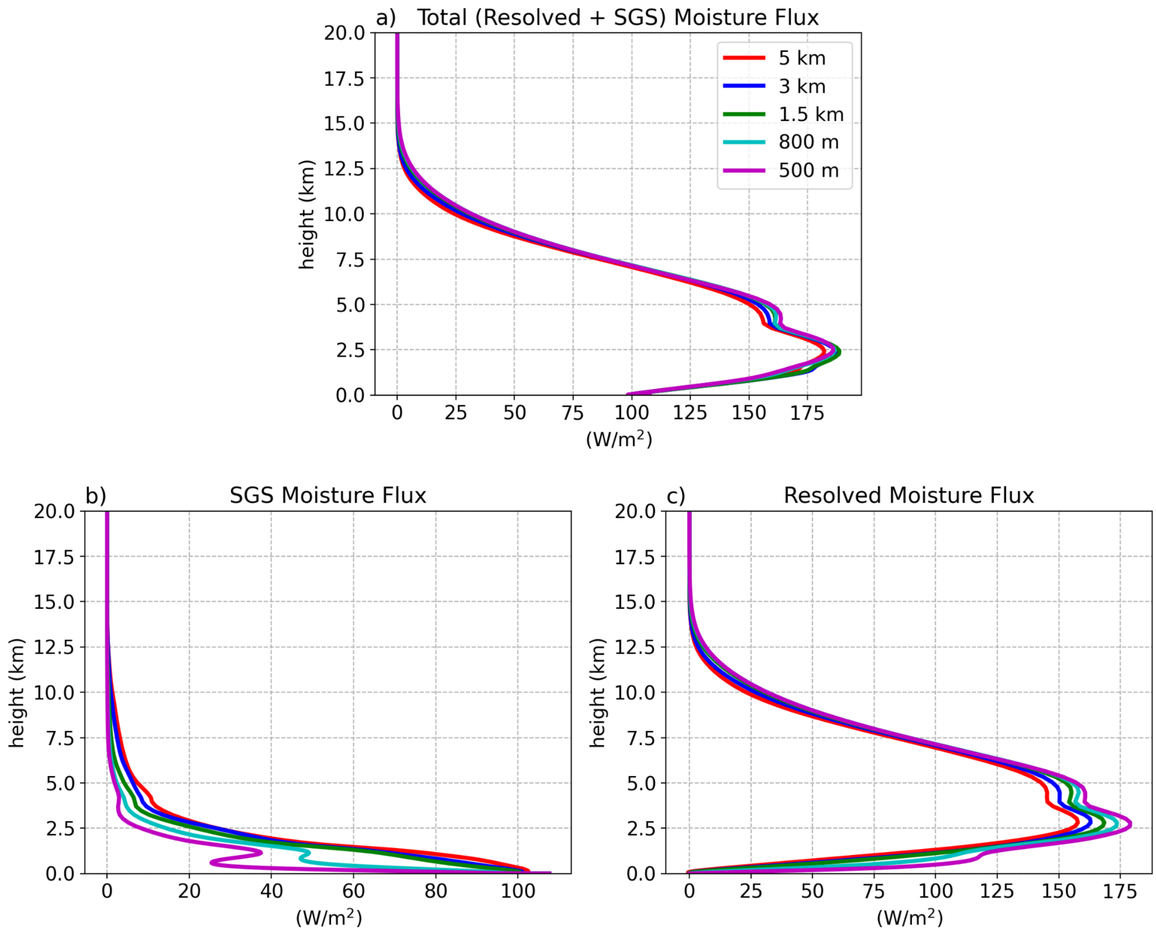

Figure 3. Temporally and horizontally averaged profiles of the a) total, b) subgrid-scale (SGS), and c) resolved moisture flux ($\overline{w'q_t'}$) for the ARM97 case averaged over the entire eight day simulation.

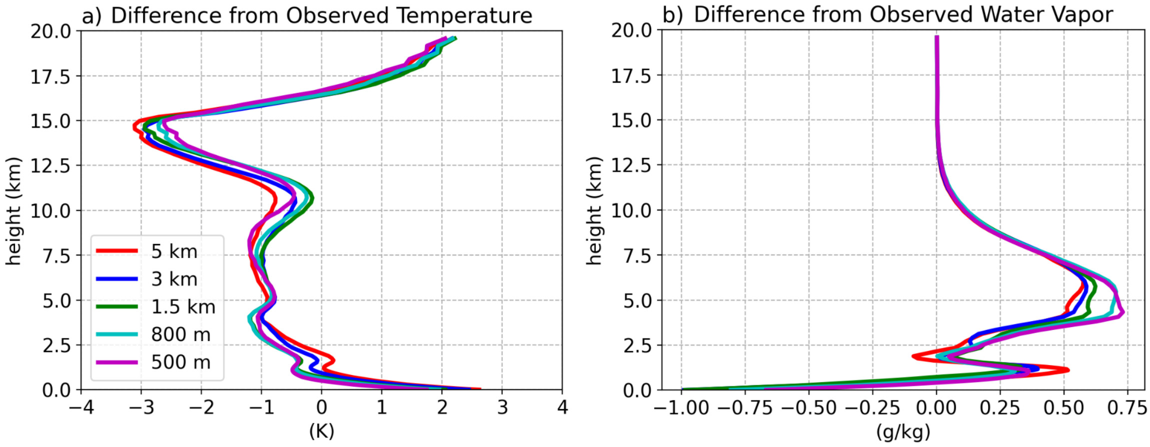

Figure 4. Temporally and horizontally averaged profiles of a) difference from observed temperature and b) difference from observed water vapor for the ARM97 case averaged over the entire eight day simulation.

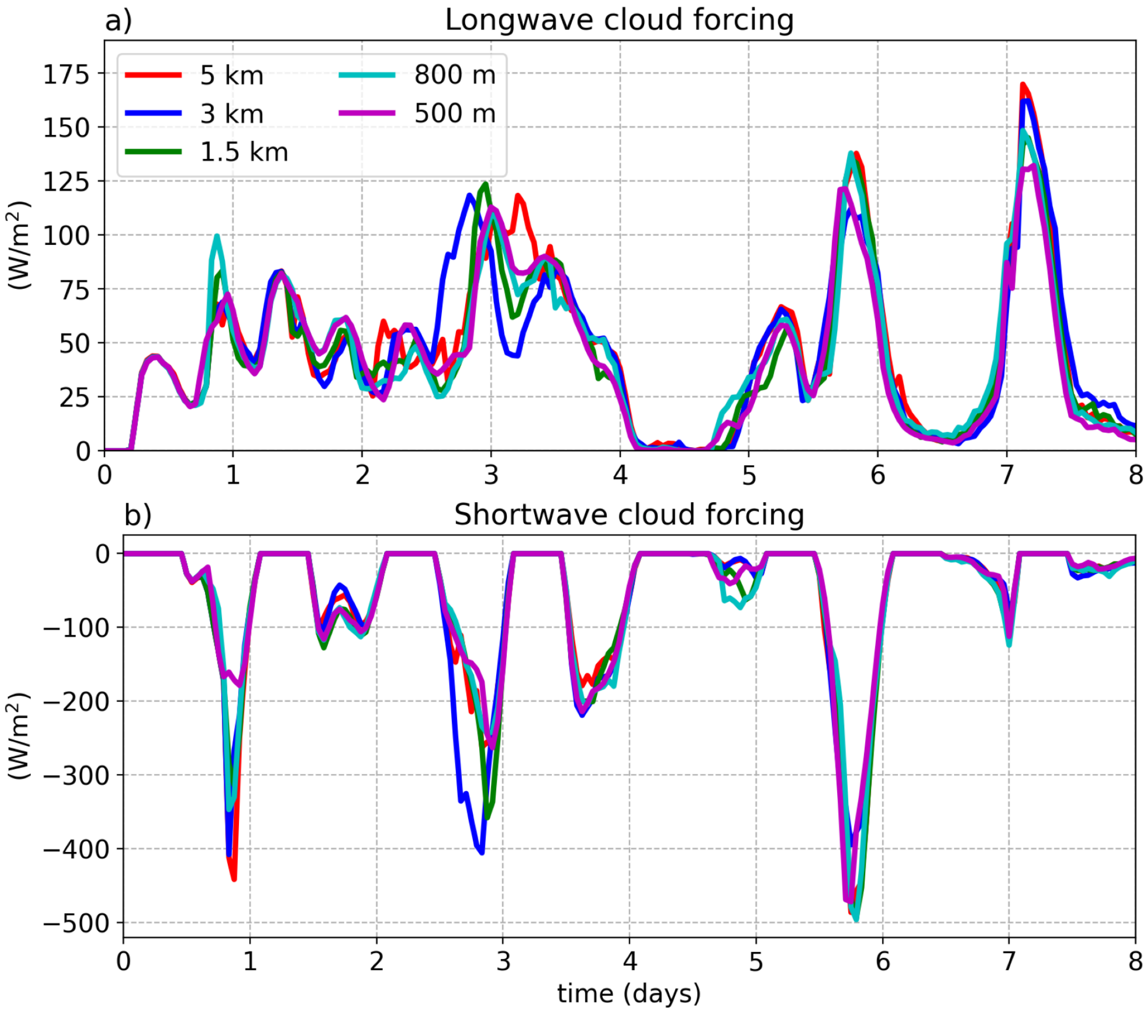

Figure 5. Temporal evolution of the horizontally averaged a) longwave cloud forcing and b) shortwave cloud forcing for the SCREAM simulations over the 8 day simulation of the ARM97 case.

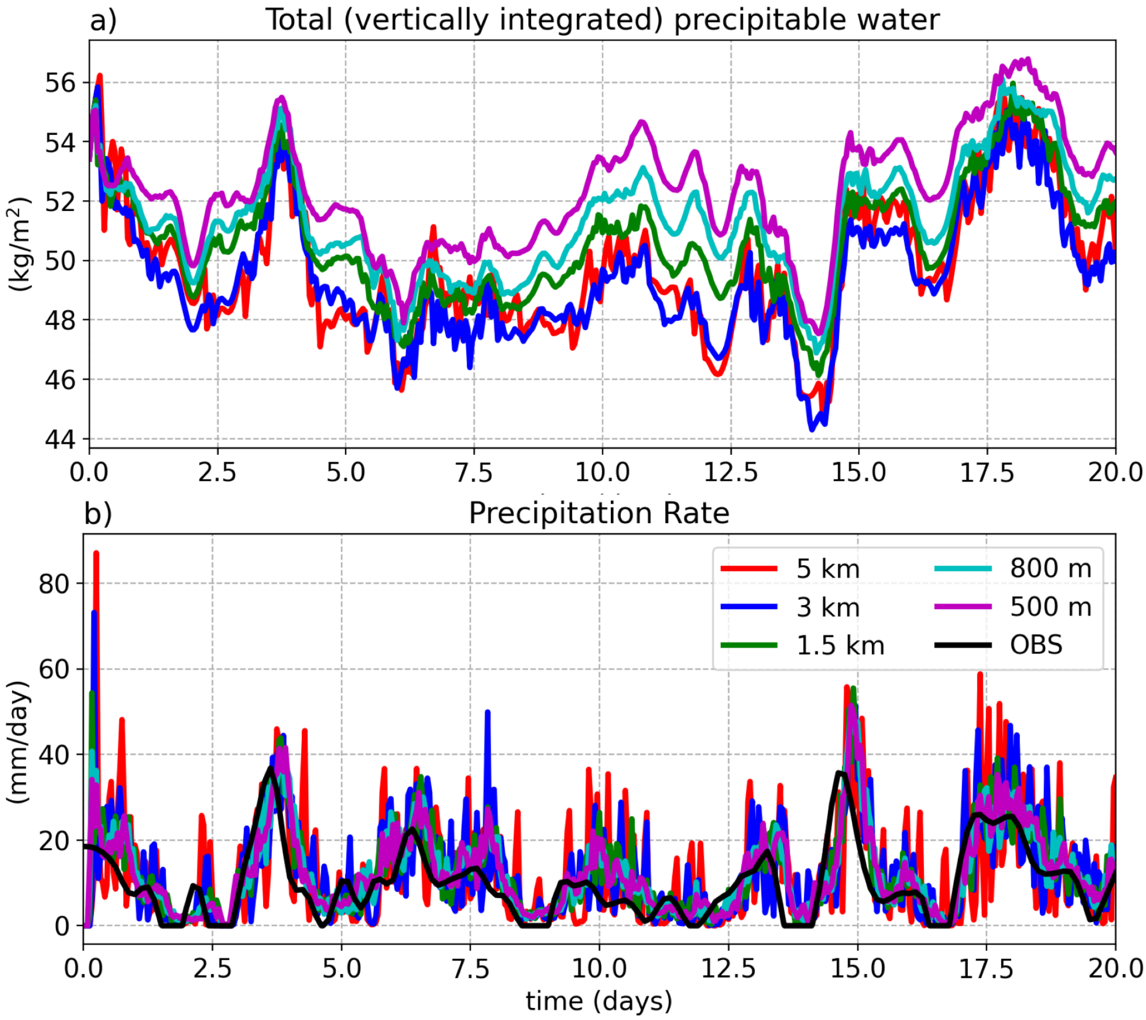

Figure 6. Temporal evolution of the horizontally averaged a) total vertically integrated precipitable water and b) precipitation rate over the 20 day simulation of the GATE case starting at 00Z 30 August 1974 for SCREAM simulations and observations. Precipitation observations are estimated by vertical integration of the observed moisture sink (Q_2) budget.

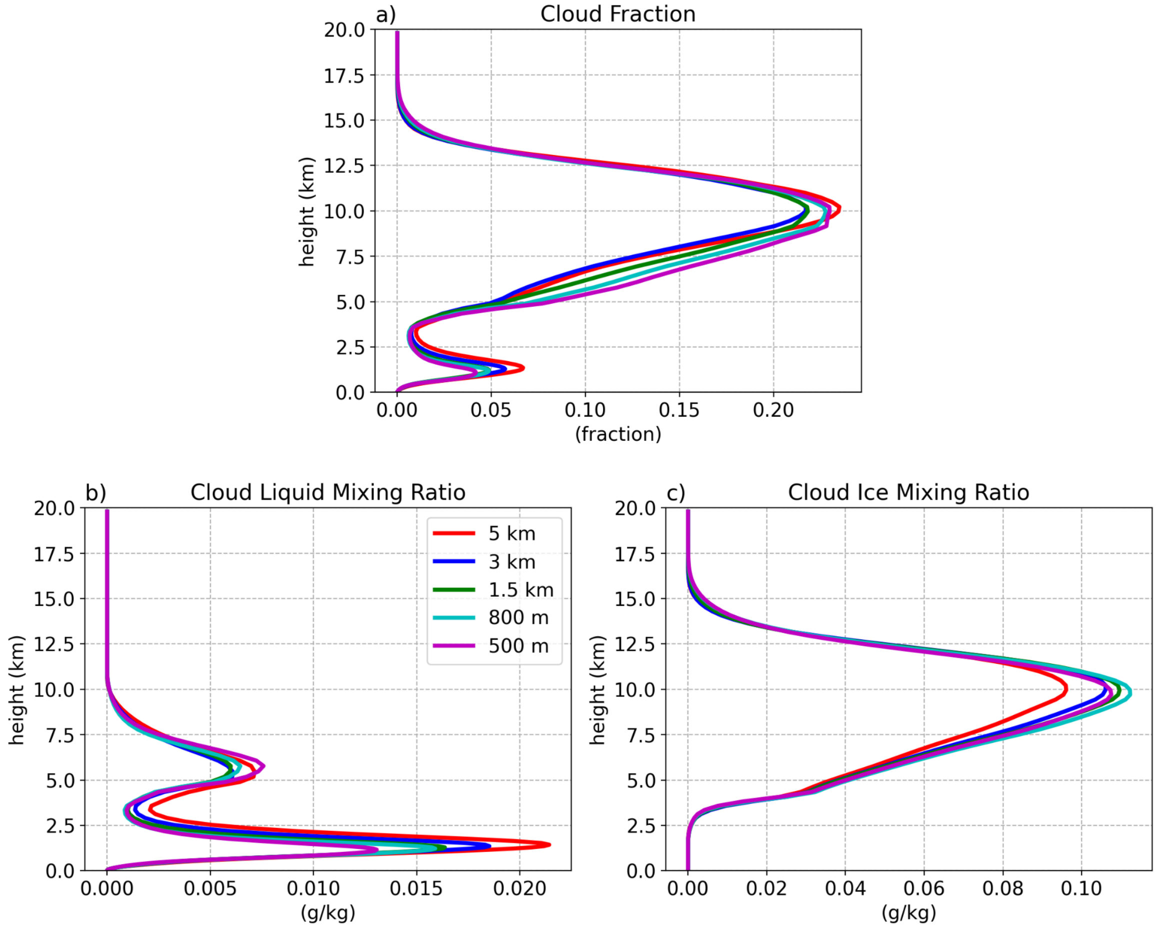

Figure 7. Temporally and horizontally averaged profiles of a) cloud fraction, b) cloud liquid mixing ratio, and c) cloud ice mixing ratio for the SCREAM simulations at various horizontal resolutions for the GATE case averaged over the entire 20 day simulation.

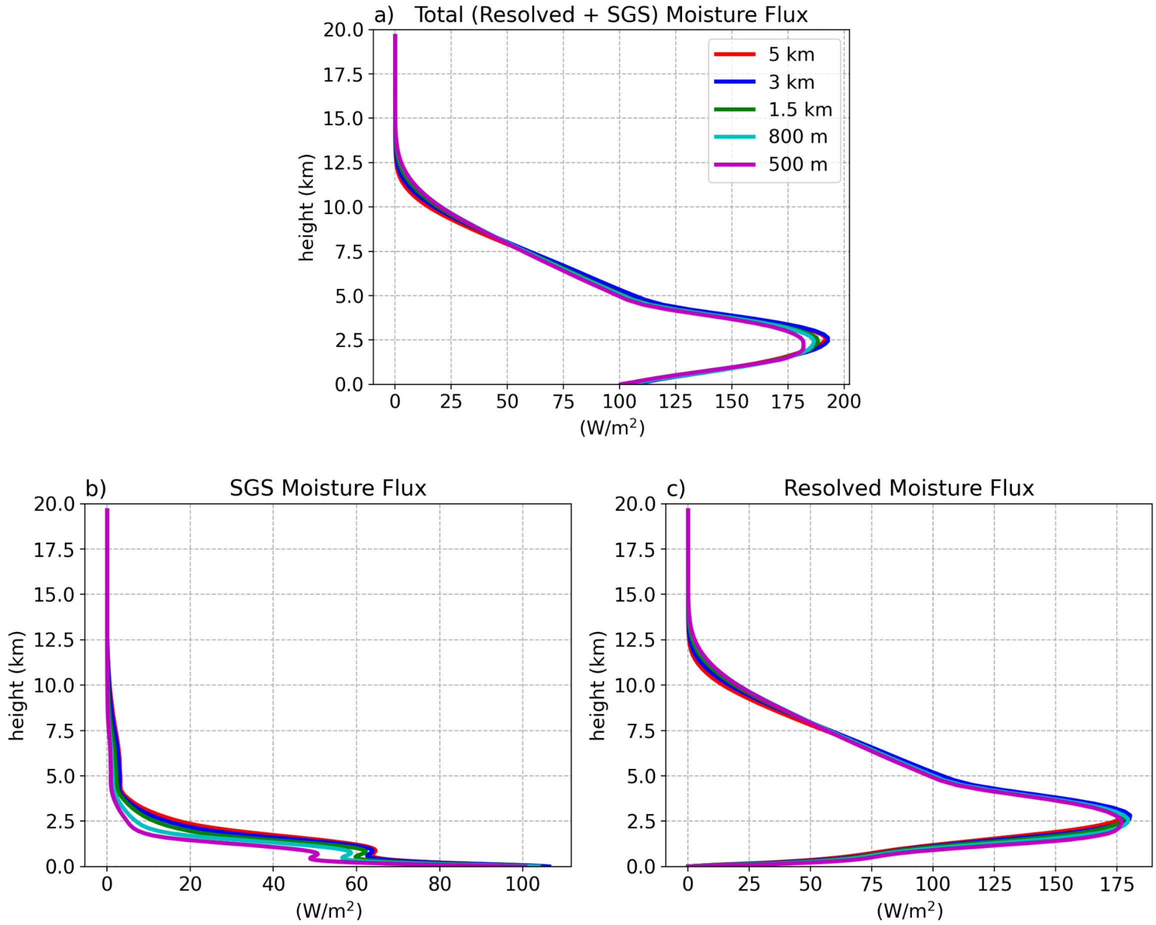

Figure 8. Temporally and horizontally averaged profiles of the a) total, b) subgrid-scale (SGS), and c) resolved moisture flux ($\overline{w'q_t'}$) for the GATE case averaged over the entire 20 day simulation.

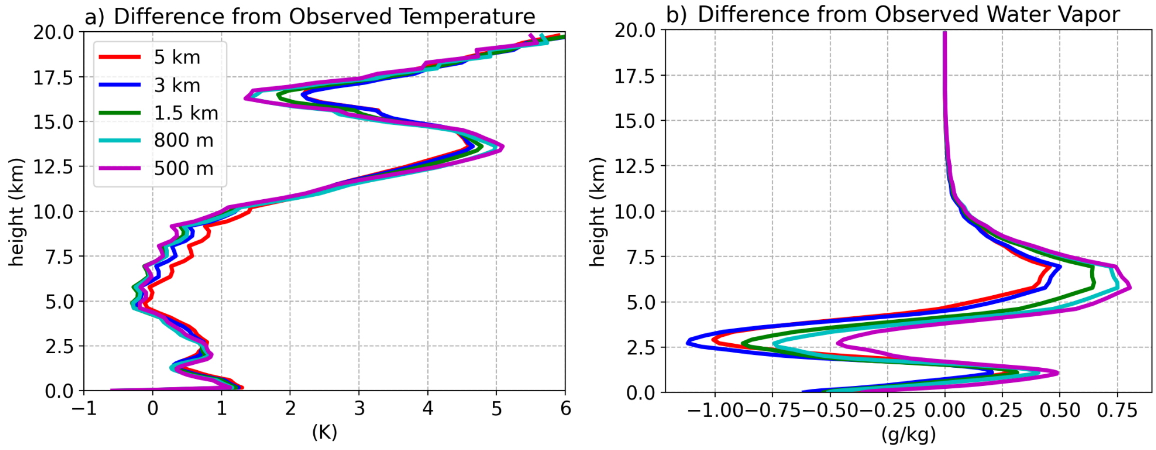

Figure 9. Temporally and horizontally averaged profiles of a) difference from observed temperature and b) difference from observed water vapor for the GATE case averaged over the entire 20 day simulation.

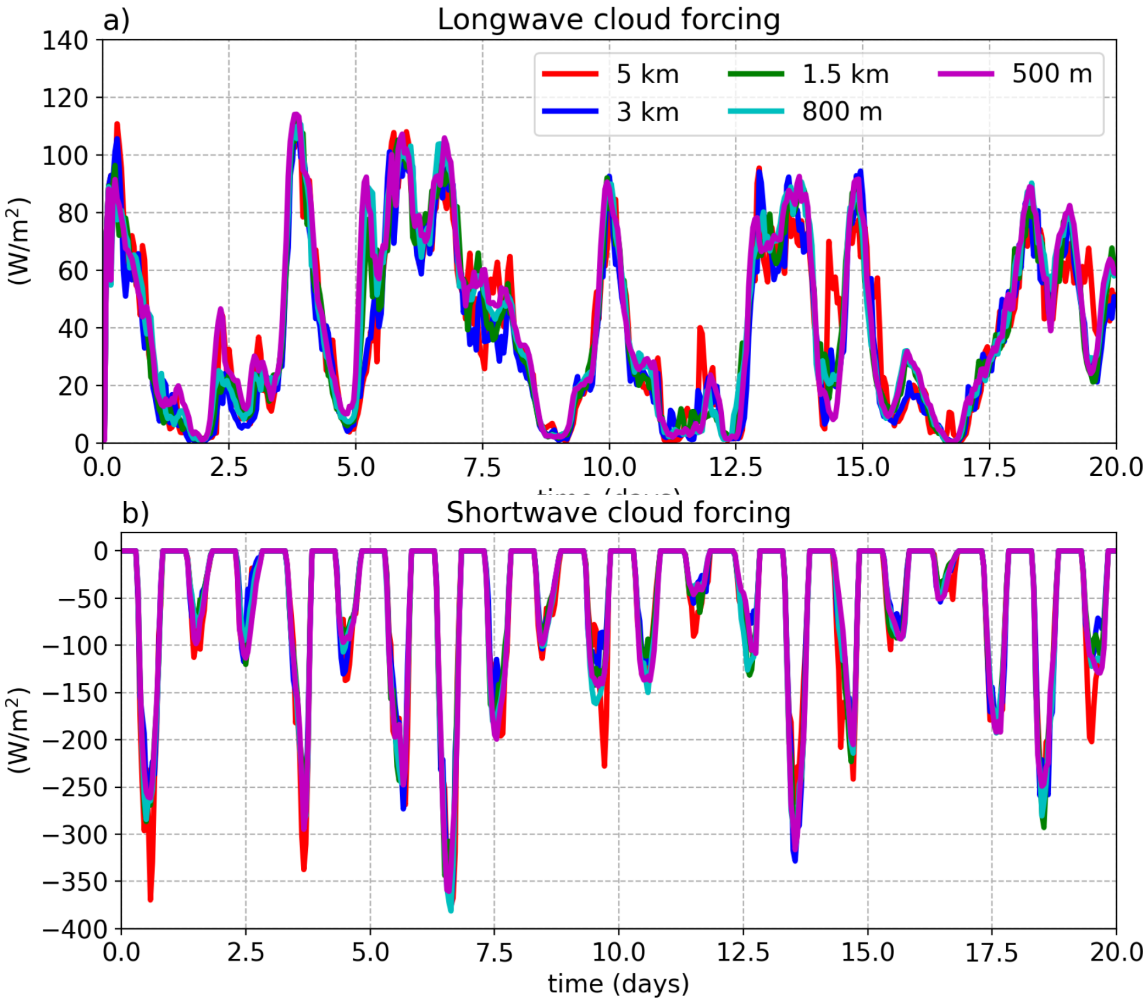

Figure 10. Temporal evolution of the horizontally averaged a) longwave cloud forcing and b) shortwave cloud forcing for the SCREAM simulations over the 20 day simulation of the GATE case.

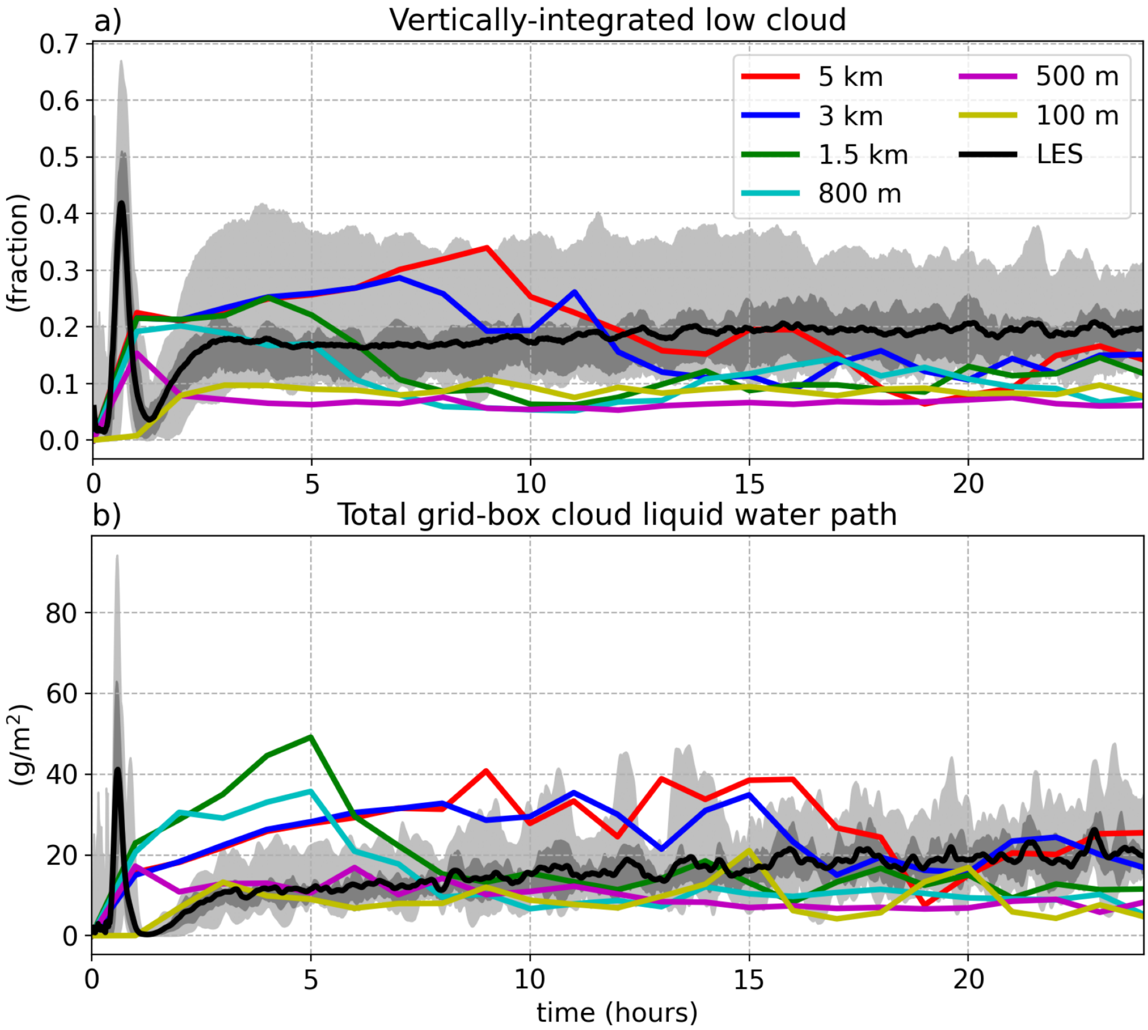

Figure 11. Temporally evolution of the horizontally averaged a) vertically-integrated low cloud and b) liquid water path for the SCREAM simulations (colored curves) over the duration of the 24 hour simulation of RICO. The black curve represents the LES mean from vanZanten et al. (2011), the dark shading represents the central half of the LES spread, while the light shading represents the full spread.

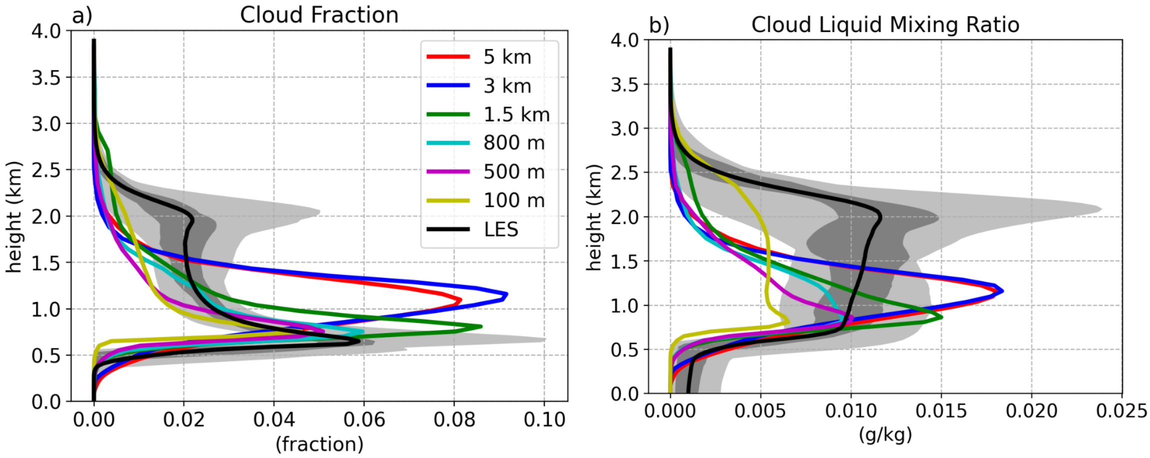

Figure 12. Temporally and horizontally averaged profiles of a) cloud fraction and b) cloud liquid mixing ratio for the RICO case, averaged over the last four hours of the case. LES results follow that as explained in Fig. 11.

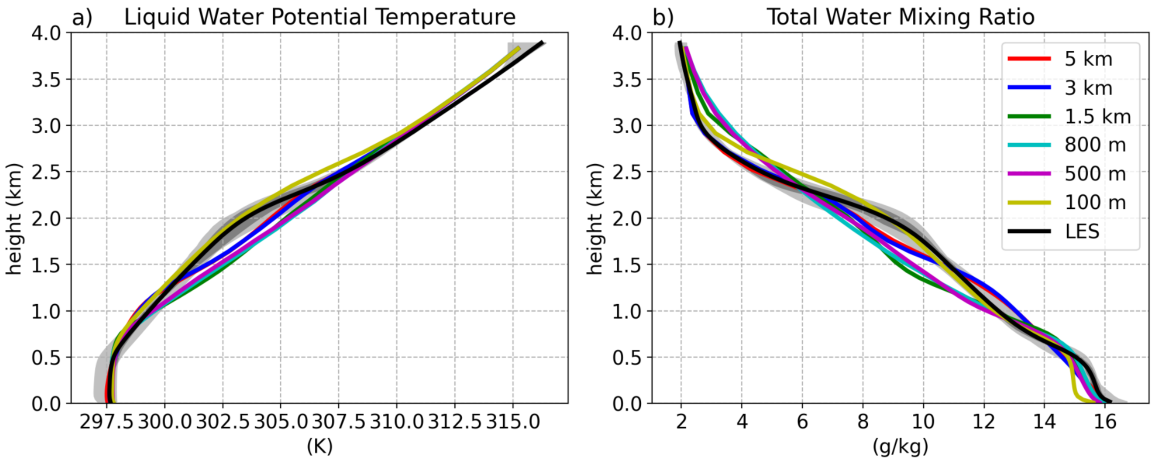

Figure 13. Temporally and horizontally averaged profiles of a) cloud fraction and b) cloud liquid mixing ratio for the RICO case, averaged over the last four hours of the case. LES results follow that as explained in Fig. 11.

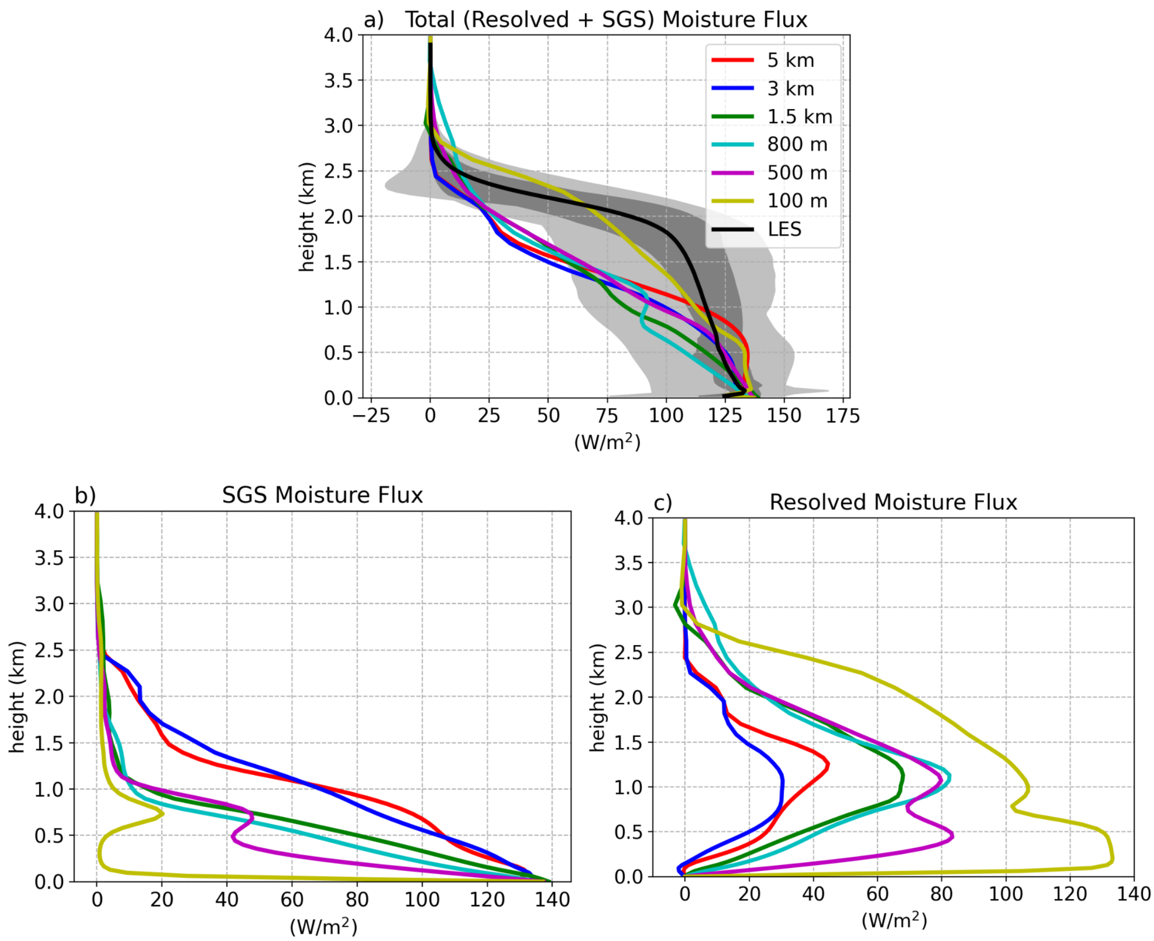

Figure 14. Temporally and horizontally averaged profiles of the a) total, b) subgrid-scale (SGS), and c) resolved moisture flux ($\overline{w'q_t'}$) for the RICO case averaged the last four hours of the case. LES results follow that as explained in Fig. 11.

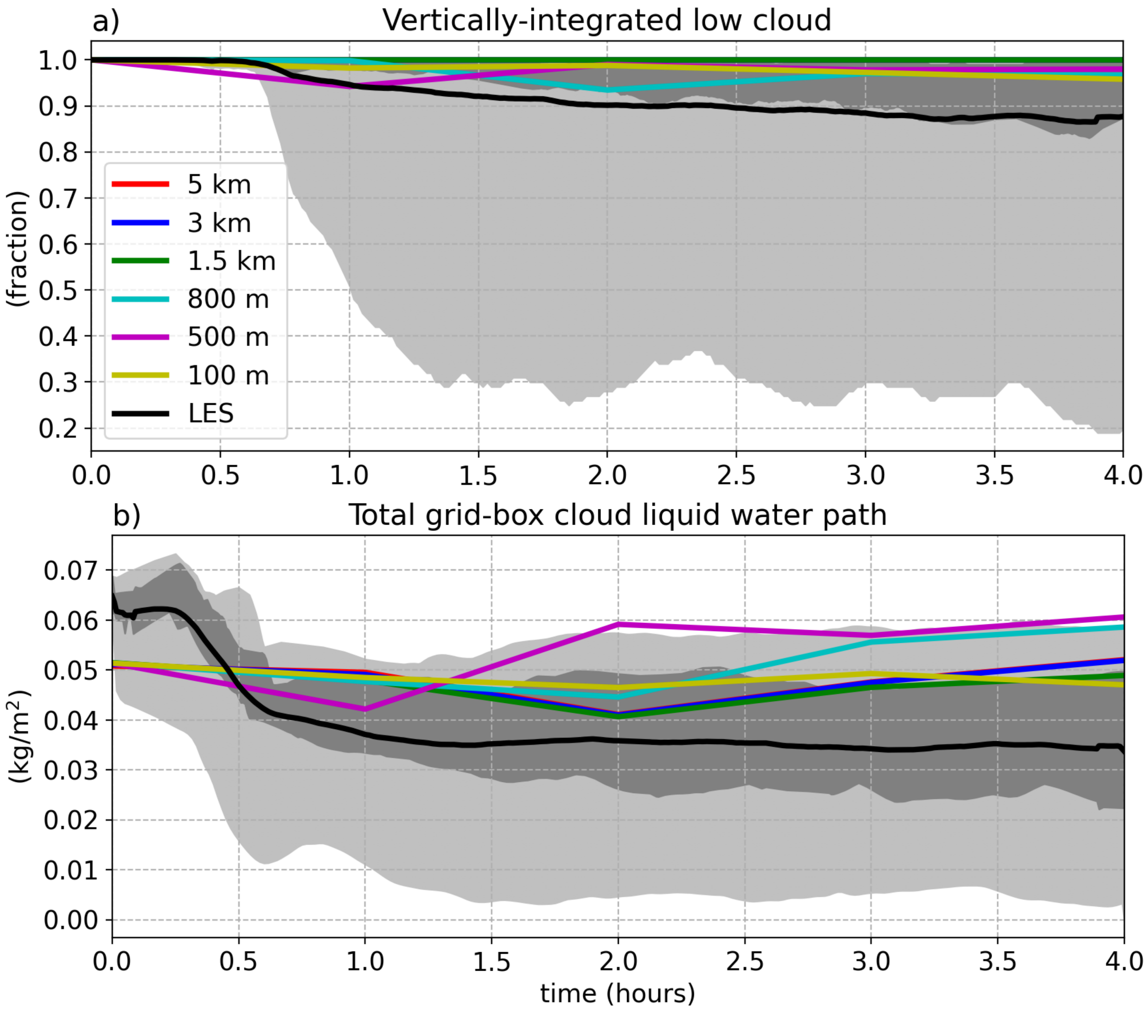

Figure 15. Temporally evolution of the horizontally averaged a) vertically-integrated low cloud and b) liquid water path for the SCREAM simulations (colored curves) over the duration of the 4 hour simulation of DYCOMS-RF01. The black curve represents the LES mean from Stevens et al. (2005), the dark shading represents the central half of the LES spread, while the light shading represents the full spread.

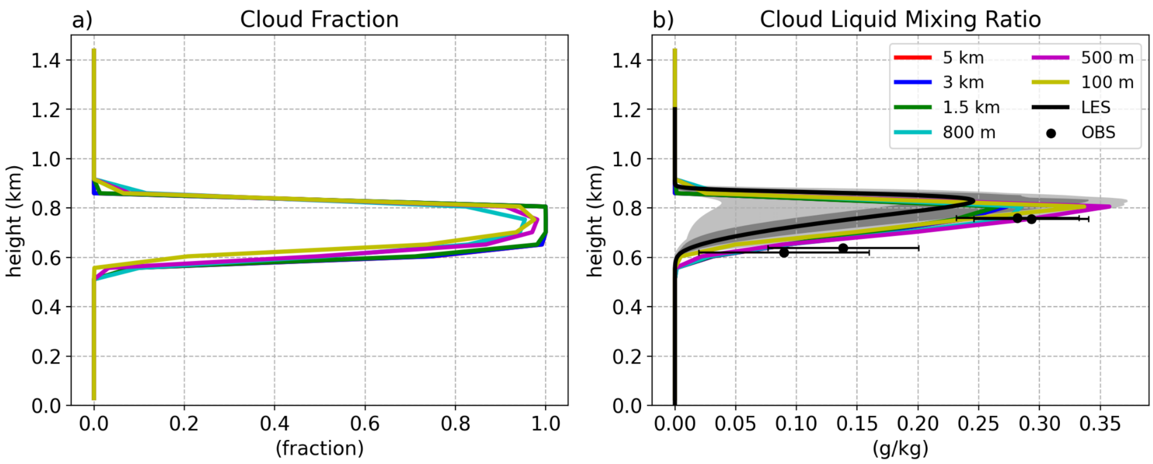

Figure 16. Temporally and horizontally averaged profiles of a) cloud fraction and b) cloud liquid mixing ratio for the DYCOMS-RF01 case, averaged over the last simulated hour. LES results follow that as explained in Fig. 15. Points denote observations, with bars representing the uncertainty.

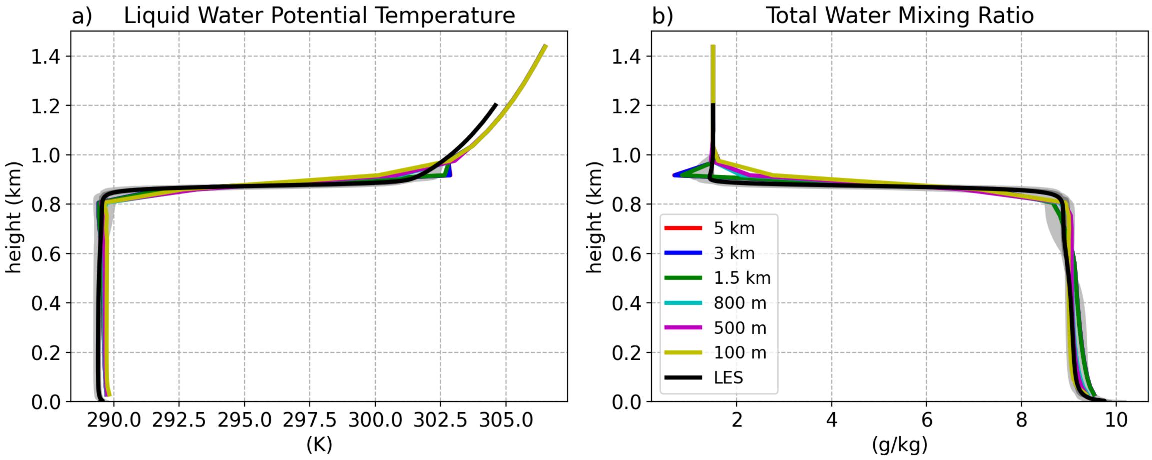

Figure 17. Temporally and horizontally averaged profiles of a) cloud fraction and b) cloud liquid mixing ratio for the DYCOMS-RF01 case, averaged over the last simulated hour. LES results follow that as explained in Fig. 15.

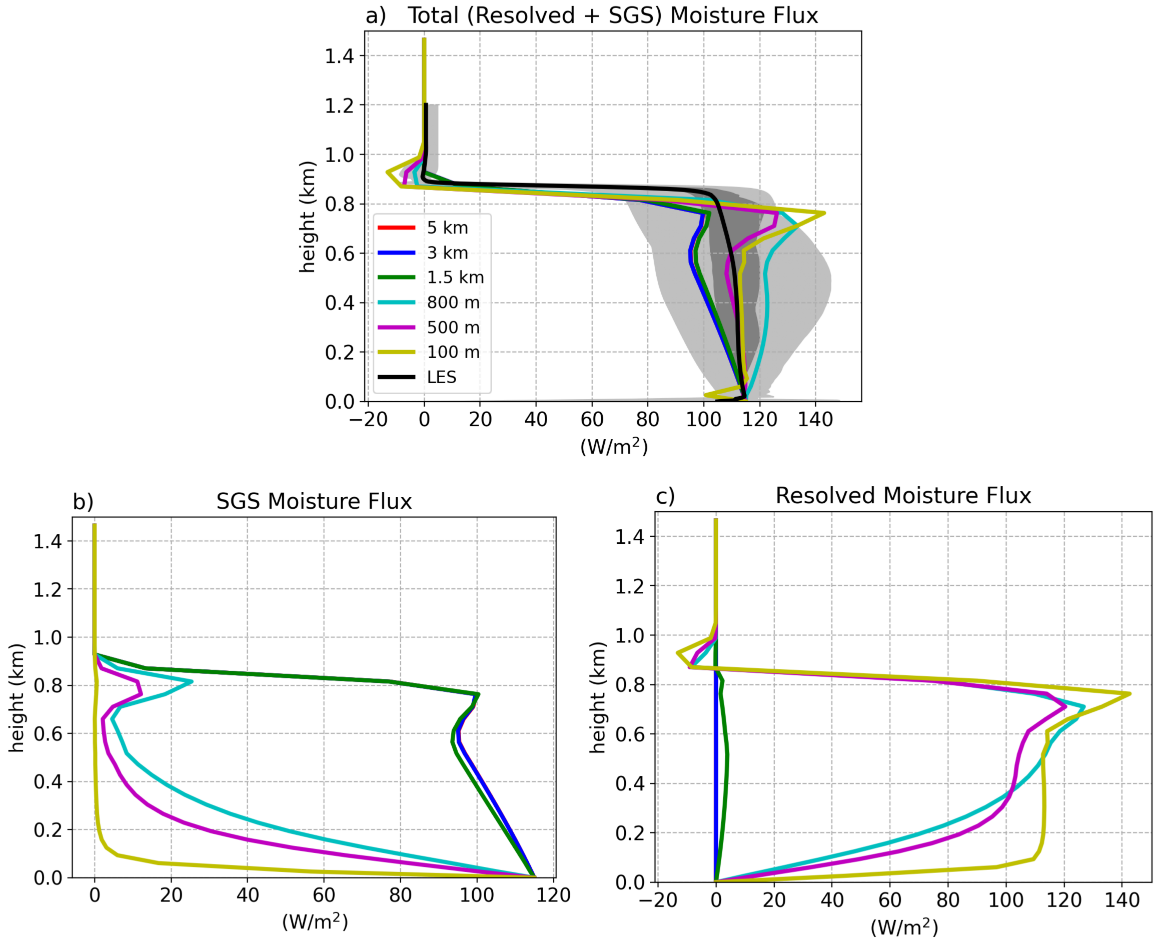

Figure 18. Temporally and horizontally averaged profiles of the a) total, b) subgrid-scale (SGS), and c) resolved moisture flux ($\overline{w'q_t'}$) for the DYCOMS-RF01 case averaged over the last simulated hour. LES results follow that as explained in Fig. 15.

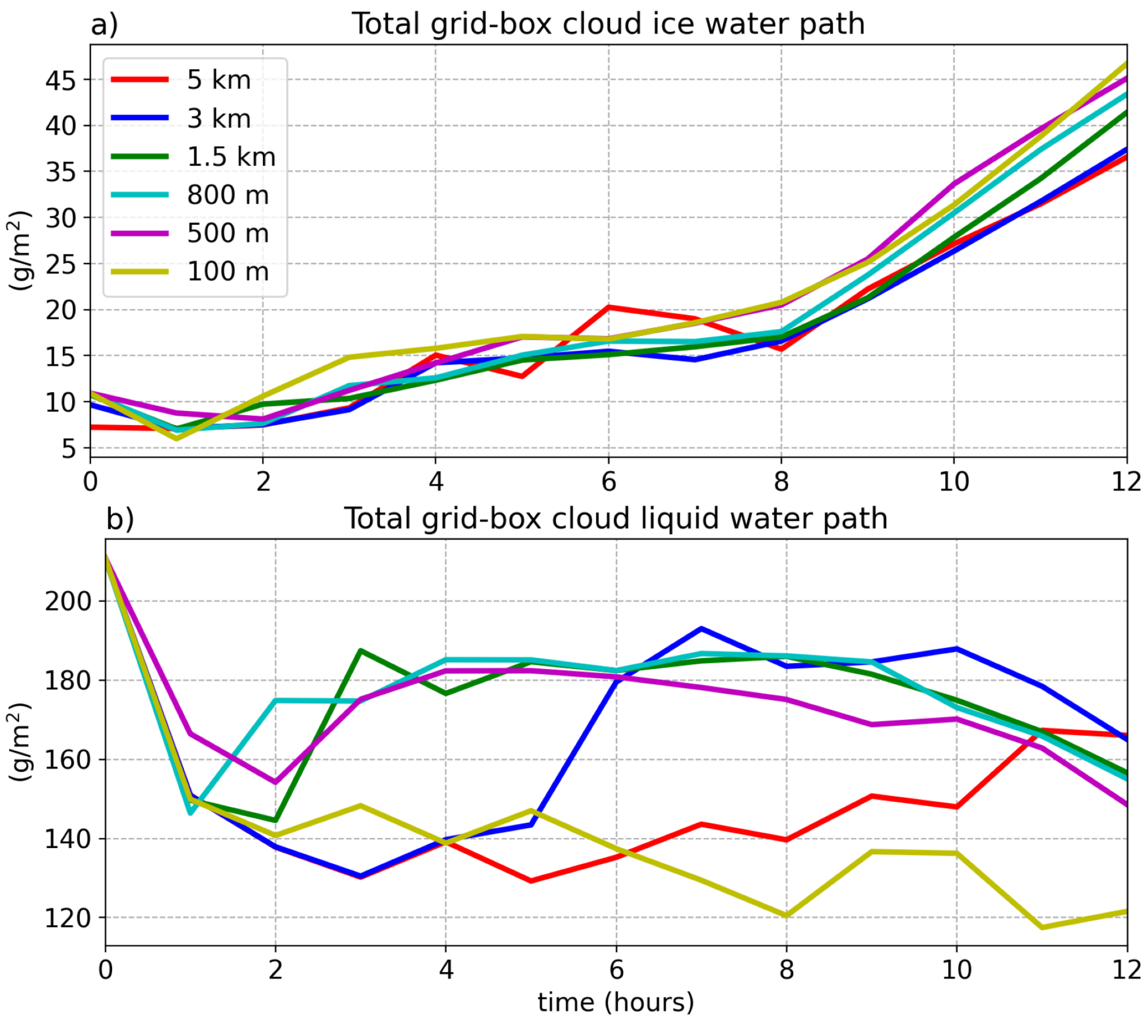

Figure 19. Temporal evolution of the horizontally averaged a) cloud ice mixing ratio and b) cloud liquid over the 12 hour simulation of the MPACE-B case for the SCREAM model simulations.

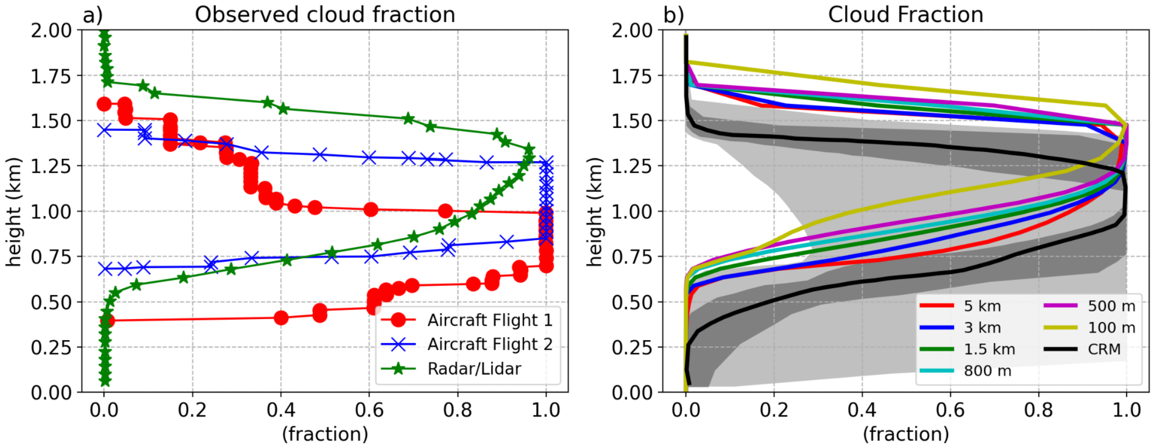

Figure 20. a) Observed cloud fraction profiles as presented in Klein et al. (2009) and b) temporally and horizontally averaged profiles of cloud fraction from the SCREAM simulations averaged over hours 4 to 12. The observation panel depicts the fraction of time at each height cloud was observed from remote sensors and two aircraft flights over the period of 1700 UTC 9 October to 0500 UTC 10 October 2004. In panel b) the solid black line indicates the mean from participating cloud resolving models (CRM) presented in Klein et al. (2009) with the dark shading representing the central half of the CRM spread and the dark shading representing the full spread of the CRMs.

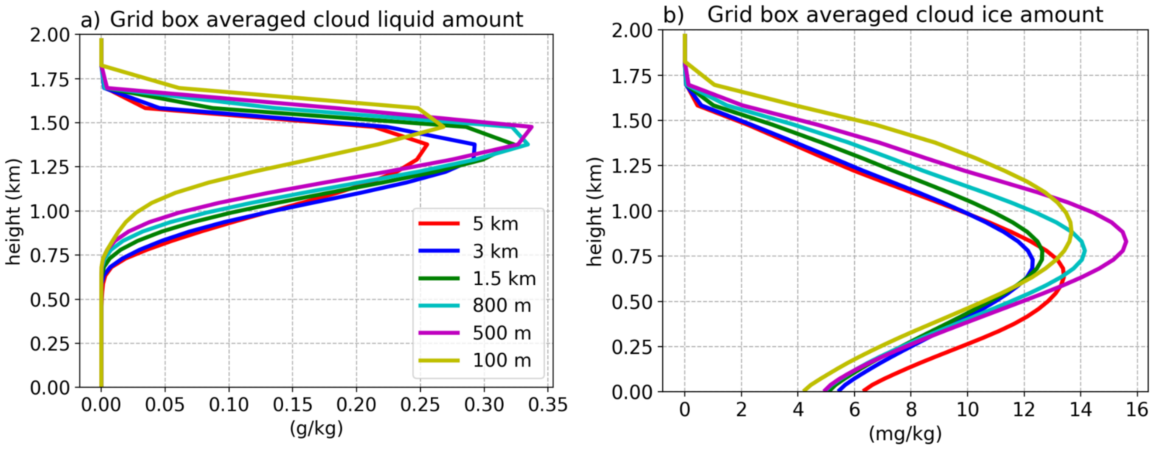

Figure 21. Temporally and horizontally averaged profiles of a) cloud liquid mixing ratio and b) cloud ice mixing ratio for the MPACE-B case, averaged over hours 4 to 12 of the model simulations.

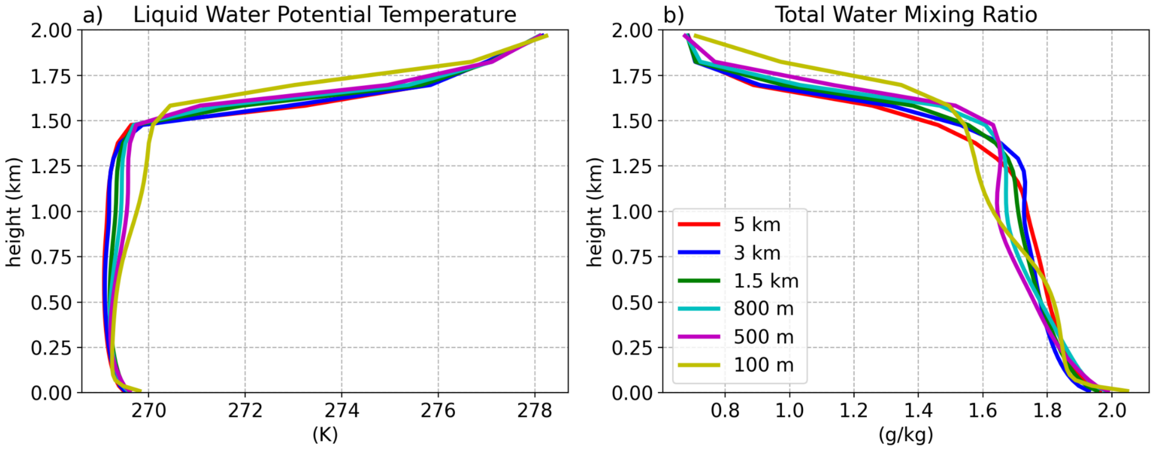

Figure 22. Temporally and horizontally averaged profiles of a) liquid water potential temperature and b) total water mixing ratio for the MPACE-B case, averaged over hours 4 to 12 of the model simulations.

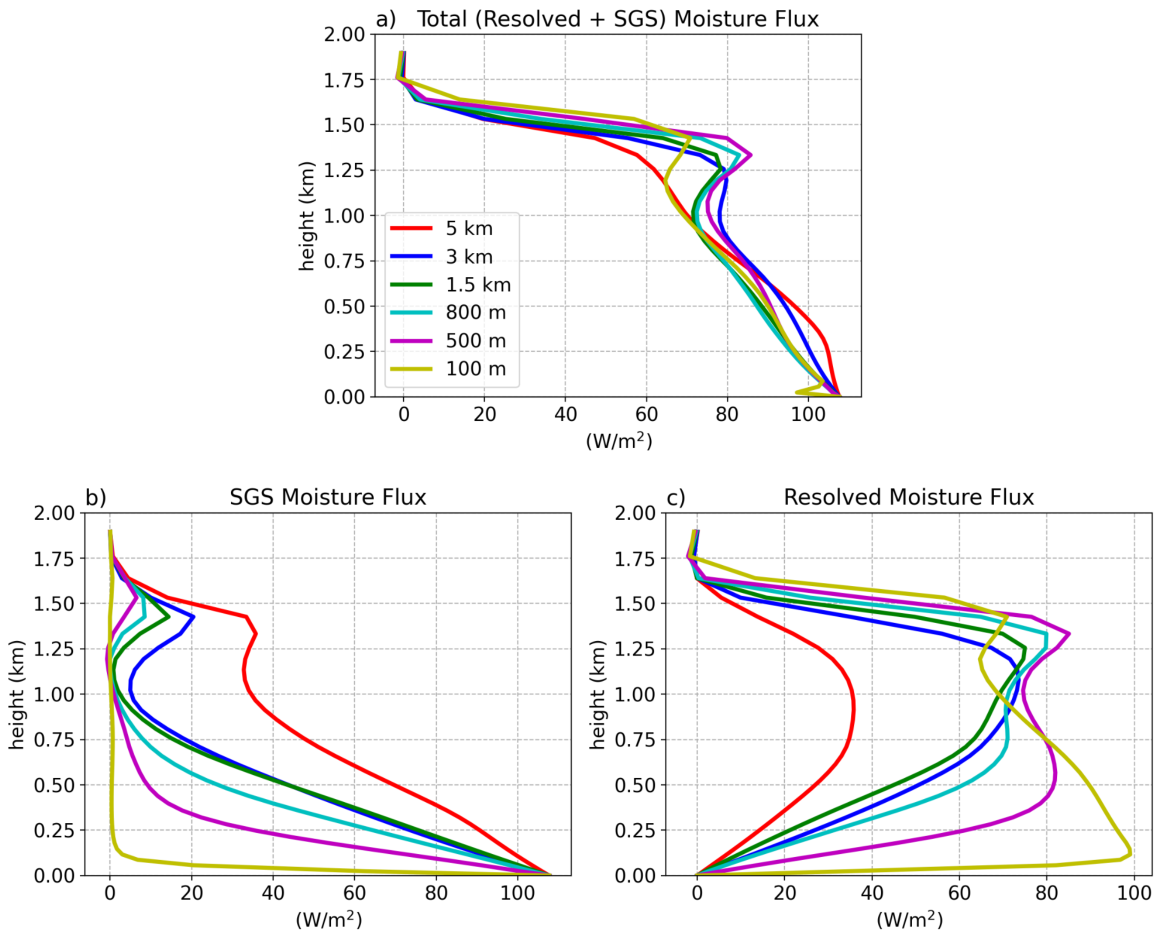

Figure 23. Temporally and horizontally averaged profiles of the a) total, b) subgrid-scale (SGS), and c) resolved moisture flux ($\overline{w'q_t'}$) for the MPACE-B case averaged over hours 4 to 12 of the model simulations.

Table 1. Specifics of the various cases used in this study. Note that all cases use the standard 128 vertical level configuration. The GATE and ARM97 cases are run with $\Delta x = \Delta y = 500$ m, 800 m, 1.5 km, 3 km, 5 km, 10 km, and 16 km while the DYCOMS-RF01, MPACE-B, and RICO cases are run with $\Delta x = \Delta y = 100$ m, 500 m, 800 m, 1.5 km, 3 km, 5 km.

Case Name	Regime	Run Duration	Horizontal domain size	Physics time step (s)	Dynamics time step (s)	Nudging
ARM97	continental deep cumulus	8 d	200 x 200 km	50	2	U,V
GATE	maritime deep cumulus	20 d	200 x 200 km	50	2	U,V
DYCOMS-RF01	marine stratocumulus	6 hr	50 x 50 km	10	0.33	none
MPACE-B	mixed-phase arctic stratocumulus	12 hr	50 x 50 km	10	0.33	none
RICO	maritime shallow cumulus	24 hr	50 x 50 km	10	0.33	none

Table 2. Temporally averaged values of shortwave and longwave cloud forcing over the simulated eight day period of the ARM97 case for each resolution.

DP-SCREAM resolution	Shortwave Cloud Forcing (W/m^2)	Longwave Cloud Forcing (W/m^2)
500 m	-55.8	44.0
800 m	-58.6	44.7
1.5 km	-59.1	44.4
3 km	-61.5	46.5
5 km	-57.5	47.9

Table 3. Temporally averaged values of shortwave and longwave cloud forcing over the simulated twenty day period of the GATE case for each resolution.

DP-SCREAM resolution	Shortwave Cloud Forcing (W/m^2)	Longwave Cloud Forcing (W/m^2)
500 m	-54.7	42.1
800 m	-54.9	41.0
1.5 km	-53.9	40.5
3 km	-53.0	40.2
5 km	-55.4	40.4

Table 4. Average liquid water and ice water paths from hours 4 through 12 of MPACE-B case from SCREAM simulations and observations. The observed values represents the averaged observed value from a blend of aircraft and ground-based observations as reported in Klein et al. (2009). We note that SCREAM ice water path values includes both non-precipitating ice and snow since P3 microphysics does not distinguish between the two. The Median Intercomparison CRM values reflect those reported by Klein et al. (2009).

Case	Liquid Water Path (g/m^2)	Ice Water Path (g/m^2)
Observations	160	15
Median Intercomparison CRM	57	29
DP-SCREAM 100 m	132	24
DP-SCREAM 500 m	174	24
DP-SCREAM 800 m	180	22
DP-SCREAM 1.5 km	180	21
DP-SCREAM 3 km	179	20
DP-SCREAM 5 km	150	21

Open Research

The software used to produce simulations presented in this paper can be found at https://doi.org/10.5281/zenodo.7218009, while the data used in the analysis of the paper can be found at https://doi.org/10.5281/zenodo.7218014.

Acknowledgments

This research was supported as part of the Energy Exascale Earth System Model (E3SM) project, funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. The authors thank Maria Chinta for comments on the manuscript. Work at LLNL was performed under the auspices of the U.S. DOE by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. LLNL IM: LLNL-JRNL-841261.

References

Bogenschutz, P. A., Gettelman, A., Morrison, H., Larson, V. E., Schanen, D. P., Meyer, N. R., & Craig, C. (2012). Unified parameterization of the planetary boundary layer and shallow convection with a higher-order turbulence closure in the community atmosphere model: single-column experiments. *Geoscientific Model Development*, *5*(6), 1407–1423. Retrieved from `https://www.geosci-model-dev.net/5/1407/2012/` doi: 10.5194/gmd-5-1407-2012

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

Bogenschutz, P. A., Tang, S., Caldwell, P. M., Xie, S., Lin, W., & Chen, Y.-S. (2020). The e3sm version 1 single-column model. *Geoscientific Model Development*, *13*(9), 4443–4458. Retrieved from `https://gmd.copernicus.org/articles/13/4443/2020/` doi: 10.5194/gmd-13-4443-2020

Caldwell, P. M., Mametjanov, A., Tang, Q., Van Roekel, L. P., Golaz, J.-C., Lin, W., ... Zhou, T. (2019). The doe e3sm coupled model version 1: Description and results at high resolution. *Journal of Advances in Modeling Earth Systems*, *11*(12), 4095-4146. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001870` doi: https://doi.org/10.1029/2019MS001870

Caldwell, P. M., Terai, C. R., Hillman, B., Keen, N. D., Bogenschutz, P., Lin, W., ... Zender, C. S. (2021). Convection-permitting simulations with the e3sm global atmosphere model. *Journal of Advances in Modeling Earth Systems*, *13*(11), e2021MS002544. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2021MS002544` (e2021MS002544 2021MS002544) doi: https://doi.org/10.1029/2021MS002544

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

Cheng, A., Xu, K.-M., & Stevens, B. (2010). Effects of resolution on the simulation of boundary-layer clouds and the partition of kinetic energy to subgrid scales. *Journal of Advances in Modeling Earth Systems*, *2*(1). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.3894/JAMES.2010.2.3` doi: https://doi.org/10.3894/JAMES.2010.2.3

Fu, Q., Krueger, S. K., & Liou, K. N. (1995). Interactions of radiation and convection in simulated tropical cloud clusters. *Journal of Atmospheric Sciences*, *52*(9), 1310 - 1328. Retrieved from `https://journals.ametsoc.org/view/`

journals/atsc/52/9/1520-0469_1995_052_1310_ioraci_2_0_co_2.xml doi: 10.1175/1520-0469(1995)052⟨1310:IORACI⟩2.0.CO;2

Gettelman, A., Truesdale, J. E., Bacmeister, J. T., Caldwell, P. M., Neale, R. B., Bogenschutz, P. A., & Simpson, I. R. (2019). The single column atmosphere model version 6 (scam6): Not a scam but a tool for model evaluation and development. *Journal of Advances in Modeling Earth Systems*, *11*(5), 1381-1401. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018MS001578 doi: https://doi.org/10.1029/2018MS001578

Giorgi, F. (2019). Thirty years of regional climate modeling: Where are we and where are we going next? *Journal of Geophysical Research: Atmospheres*, *124*(11), 5696-5723. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD030094 doi: https://doi.org/10.1029/2018JD030094

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

Houze Jr., R. A., & Betts, A. K. (1981). Convection in gate. *Reviews of Geophysics*, *19*(4), 541-576. Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/RG019i004p00541 doi: https://doi.org/10.1029/RG019i004p00541

Khairoutdinov, M. F., Krueger, S. K., Moeng, C.-H., Bogenschutz, P. A., & Randall, D. A. (2009). Large-eddy simulation of maritime deep tropical convection. *Journal of Advances in Modeling Earth Systems*, *1*(4). Retrieved from https://agupubs.onlinelibrary.wiley.com/doi/abs/10.3894/JAMES.2009.1.15 doi: https://doi.org/10.3894/JAMES.2009.1.15

Khairoutdinov, M. F., & Randall, D. A. (2003). Cloud resolving modeling of the arm summer 1997 iop: Model formulation, results, uncertainties, and sensitivities. *Journal of the Atmospheric Sciences*, *60*(4), 607 - 625. Retrieved from https://journals.ametsoc.org/view/journals/atsc/60/4/1520-0469_2003_060_0607_crmota_2.0.co_2.xml doi: 10.1175/1520-0469(2003)060⟨0607:CRMOTA⟩2.0.CO;2

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

Krueger, S. K. (1988). Numerical simulation of tropical cumulus clouds and their interaction with the subcloud layer. *Journal of Atmospheric Sciences*, *45*(16), 2221 - 2250. Retrieved from https://journals.ametsoc.org/view/journals/atsc/45/16/1520-0469_1988_045_2221_nsotcc_2_0_co_2.xml doi: 10.1175/1520-0469(1988)045⟨2221:NSOTCC⟩2.0.CO;2

Larson, V. E., Schanen, D. P., Wang, M., Ovchinnikov, M., & Ghan, S. (2012). Pdf parameterization of boundary layer clouds in models with horizontal grid spacings from 2 to 16 km. *Monthly Weather Review*, *140*(1), 285 - 306. Retrieved from https://journals.ametsoc.org/view/journals/mwre/140/1/mwr-d-10-05059.1.xml doi: 10.1175/MWR-D-10-05059.1

Melvin, T., Benacchio, T., Shipway, B., Wood, N., Thuburn, J., & Cotter, C. (2019). A mixed finite-element, finite-volume, semi-implicit discretization for atmospheric dynamics: Cartesian geometry. *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 2835-2853. Retrieved from https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3501 doi: https://doi.org/10.1002/qj.3501

Morrison, H., Milbrandt, J. A., Bryan, G. H., Ikeda, K., Tessendorf, S. A., & Thompson, G. (2015). Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. part ii: Case study comparisons with observations and other schemes. *Journal of the Atmospheric Sciences*, *72*(1), 312 - 339. Retrieved from `https://journals.ametsoc.org/view/journals/atsc/72/1/jas-d-14-0066.1.xml` doi: 10.1175/JAS-D-14-0066.1

Morrison, H., & Pinto, J. (2006, 07). Intercomparison of bulk cloud microphysics schemes in mesoscale simulations of springtime arctic mixed-phase stratiform clouds. *Monthly Weather Review - MON WEATHER REV*, *134*. doi: 10.1175/MWR3154.1

Parishani, H., Pritchard, M. S., Bretherton, C. S., Wyant, M. C., & Khairoutdinov, M. (2017). Toward low-cloud-permitting cloud superparameterization with explicit boundary layer turbulence. *Journal of Advances in Modeling Earth Systems*, *9*(3), 1542-1571. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017MS000968` doi: 10.1002/2017MS000968

Park, S. (2014). A unified convection scheme (unicon). part ii: Simulation. *Journal of the Atmospheric Sciences*, *71*(11), 3931 - 3973. Retrieved from `https://journals.ametsoc.org/view/journals/atsc/71/11/jas-d-13-0234.1.xml` doi: 10.1175/JAS-D-13-0234.1

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

Rauber, R. M., Stevens, B., Ochs, H. T., Knight, C., Albrecht, B. A., Blyth, A. M., ... Zuidema, P. (2007). Rain in shallow cumulus over the ocean: The rico campaign. *Bulletin of the American Meteorological Society*, *88*(12), 1912 - 1928. Retrieved from `https://journals.ametsoc.org/view/journals/bams/88/12/bams-88-12-1912.xml` doi: 10.1175/BAMS-88-12-1912

Satoh, M., Tomita, T., Miura, H., Nasuon, T., & co authors. (2008). Nonhydrostatic icosahedral atmospheric model (nicam) for global cloud resolving simulations. *J. Comput. Phys.*, *227*, 3486-3514. doi: 10.1016.j.jcp.2007.02.006

Stevens, B., ACQUISTAPACE, C., HANSEN, A., HEINZE, R., KLINGER, C., KLOCKE, D., ... ZÄNGL, G. (2020). The added value of large-eddy and storm-resolving models for simulating clouds and precipitation. *Journal of the Meteorological Society of Japan. Ser. II*, *98*(2), 395-435. doi: 10.2151/jmsj.2020-021

Stevens, B., Lenschow, D. H., Vali, G., Gerber, H., Bandy, A., Blomquist, B., ... van Zanten, M. C. (2003). Dynamics and chemistry of marine stratocumulus?dycoms-ii. *Bulletin of the American Meteorological Society*, *84*(5), 579 - 594. Retrieved from `https://journals.ametsoc.org/view/journals/bams/84/5/bams-84-5-579.xml` doi: 10.1175/BAMS-84-5-579

Stevens, B., Moeng, C.-H., & Co-authors. (2005). Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Mon. Wea. Rev.*, *133*, 1443-1462. doi: https://doi.org/10.1175/MWR2930.1

Stevens, B., Satoh, M., Auger, L., & Co-authors. (2019). Dyamond: the dynamics of the atmospheric general circulation modeled on non-hydrostatic domains. *Prog. Earth Planet Sci.*, *6*, 61. doi: https://doi.org/10.1186/s40645-019-0304-z

Tang, Q., Klein, S. A., Xie, S., Lin, W., Golaz, J.-C., Roesler, E. L., ... Zheng, X. (2019). Regionally refined test bed in e3sm atmosphere model version 1 (eamv1) and applications for high-resolution modeling. *Geoscientific Model Development*, *12*(7), 2679–2706. Retrieved from `https://gmd.copernicus.org/articles/12/2679/2019/` doi: 10.5194/gmd-12-2679-2019

Taylor, M. A., Guba, O., Steyer, A., Ullrich, P. A., Hall, D. M., & Eldred, C.

(2020). An energy consistent discretization of the nonhydrostatic equations in primitive variables. *Journal of Advances in Modeling Earth Systems*, *12*(1), e2019MS001783. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019MS001783` (e2019MS001783 10.1029/2019MS001783) doi: https://doi.org/10.1029/2019MS001783

Tomita, T., Miura, H., Iga, S., Nasuno, T., & Satoh, M. (2005). A global cloud-resolving simulaiton: Preliminary results from an aqua planet experiment. *Geophys. Res. Letters*, *32*, 3283.

vanZanten, M. C., Stevens, B., Nuijens, L., Siebesma, A. P., Ackerman, A. S., Burnet, F., . . . Wyszogrodzki, A. (2011). Controls on precipitation and cloudiness in simulations of trade-wind cumulus as observed during rico. *Journal of Advances in Modeling Earth Systems*, *3*(2). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011MS000056` doi: https://doi.org/10.1029/2011MS000056

Verlinde, J., Harrington, J. Y., McFarquhar, G. M., Yannuzzi, V. T., Avramov, A., Greenberg, S., . . . Schofield, R. (2007). The mixed-phase arctic cloud experiment. *Bulletin of the American Meteorological Society*, *88*(2), 205 - 222. Retrieved from `https://journals.ametsoc.org/view/journals/bams/88/2/bams-88-2-205.xml` doi: 10.1175/BAMS-88-2-205

Wing, A. A., Reed, K. A., Satoh, M., Stevens, B., Bony, S., & Ohno, T. (2018). Radiative–convective equilibrium model intercomparison project. *Geoscientific Model Development*, *11*(2), 793–813. Retrieved from `https://gmd.copernicus.org/articles/11/793/2018/` doi: 10.5194/gmd-11-793-2018

Xie, S., Xu, K.-M., Cederwall, R. T., Bechtold, P., Genio, A. D. D., Klein, S. A., . . . Zhang, M. (2002). Intercomparison and evaluation of cumulus parametrizations under summertime midlatitude continental conditions. *Quarterly Journal of the Royal Meteorological Society*, *128*(582), 1095-1135. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/003590002320373229` doi: https://doi.org/10.1256/003590002320373229

Xu, K.-M., Arakawa, A., & Krueger, S. K. (1992). The macroscopic behavior of cumulus ensembles simulated by a cumulus ensemble model. *Journal of Atmospheric Sciences*, *49*(24), 2402 - 2420. Retrieved from `https://journals.ametsoc.org/view/journals/atsc/49/24/1520-0469_1992_049_2402_tmboce_2_0_co_2.xml` doi: 10.1175/1520-0469(1992)049⟨2402:TMBOCE⟩2.0.CO;2

Xu, K.-M., Cederwall, R. T., Donner, L. J., Grabowski, W. W., Guichard, F., Johnson, D. E., . . . Zhang, M.-H. (2002). An intercomparison of cloud-resolving models with the atmospheric radiation measurement summer 1997 intensive observation period data. *Quarterly Journal of the Royal Meteorological Society*, *128*(580), 593-624. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1256/003590002321042117` doi: https://doi.org/10.1256/003590002321042117

Zhang, M. H., & Lin, J. L. (1997). Constrained variational analysis of sounding data based on column-integrated budgets of mass, heat, moisture, and momentum: Approach and application to arm measurements. *Journal of the Atmospheric Sciences*, *54*(11), 1503 - 1524. Retrieved from `https://journals.ametsoc.org/view/journals/atsc/54/11/1520-0469_1997_054_1503_cvaosd_2.0.co_2.xml` doi: 10.1175/1520-0469(1997)054⟨1503:CVAOSD⟩2.0.CO;2

Zhang, M. H., Lin, J. L., Cederwall, R. T., Yio, J. J., & Xie, S. C. (2001). Objective analysis of arm iop data: Method and sensitivity. *Monthly Weather Review*, *129*(2), 295 - 311. Retrieved from `https://journals.ametsoc.org/view/journals/mwre/129/2/1520-0493_2001_129_0295_oaoaid_2.0.co_2.xml` doi: 10.1175/1520-0493(2001)129⟨0295:OAOAID⟩2.0.CO;2

Zhu, P., Bretherton, C. S., K?hler, M., Cheng, A., Chlond, A., Geng, Q., . . .

Stevens, B. (2005). Intercomparison and interpretation of single-column model simulations of a nocturnal stratocumulus-topped marine boundary layer. *Monthly Weather Review*, *133*(9), 2741 - 2758. Retrieved from https://journals.ametsoc.org/view/journals/mwre/133/9/mwr2997.1.xml doi: 10.1175/MWR2997.1