# Statistically Steady State Large-Eddy Simulations Forced by an Idealized GCM: 1. Forcing Framework and Simulation Characteristics

**Zhaoyi Shen**[1], **Kyle G. Pressel**[2], **Zhihong Tan**[3], **Tapio Schneider**[1,4]

[1]Department of Environmental Science and Engineering, California Institute of Technology, Pasadena, CA, USA.
[2]Pacific Northwest National Laboratory, Richland, WA, USA.
[3]Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA.
[4]Jet Propulsion Laboratory, Pasadena, CA, USA.

**Key Points:**

- A framework in which LES is driven by large-scale forcing from a GCM is developed
- LES with large-scale forcing reaches steady states without nudging to thermodynamic GCM profiles
- LES driven by the GCM is used to simulate different tropical cloud regimes across the Walker circulation

---

Corresponding author: Zhaoyi Shen, `zhaoyi@caltech.edu`

## Abstract

Using large-eddy simulations (LES) systematically has the potential to inform parameterizations of subgrid-scale (SGS) processes in general circulation models (GCMs), such as turbulence, convection, and clouds. Here we show how LES can be run to simulate grid columns of GCMs to generate LES across a cross-section of dynamical regimes. The LES setup approximately replicates the thermodynamic and water budgets in GCM grid columns. Resolved horizontal and vertical transports of heat and water and large-scale pressure gradients from the GCM are prescribed as forcing in the LES. The LES are forced with prescribed surface temperatures, but atmospheric temperature and moisture are free to adjust, reducing the imprinting of GCM fields on the LES. In both the GCM and LES, radiative transfer is treated in a unified but idealized manner (semi-gray atmosphere without water vapor feedback or cloud radiative effects). We show that the LES in this setup reaches statistically steady states without nudging to thermodynamic GCM profiles. The steady states provide training data for developing GCM parameterizations. The same LES setup also provides a good basis for studying the cloud response to global warming.

## Plain Language Summary

Clouds and their feedbacks remain one of the largest uncertainties in predictions of future climate changes. High-resolution models can provide faithful simulations of clouds and their underlying turbulence in limited areas, but they have primarily been used in select locations, with limited success in reducing uncertainties in climate predictions. This study presents a framework for driving high-resolution simulations by a global climate model, which allows us to generate a library of high-resolution simulations across different cloud regimes. The framework leverages the potential of high-resolution models to improve parameterizations of clouds and turbulence in climate models and to better understand the cloud feedback mechanisms.

## 1 Introduction

General circulation model (GCM) predictions of the equilibrium climate sensitivity (ECS)—the equilibrium global mean surface temperature change after doubling of $CO_2$ concentrations—range between 2 K and 5 K across climate models (?, ?). This 3 K uncertainty has remained virtually unchanged for forty years (?, ?), despite increases in the complexity of climate models and exponential increases in the computational power of high-performance computing (HPC) systems (?, ?). Similar uncertainties exist in the transient climate response (?, ?). The largest uncertainties in model projections of climate change are traceable to the way in which these models represent boundary layer clouds and their climate feedbacks (?, ?, ?, ?, ?, ?, ?, ?, ?). The challenge posed by clouds is that, on the one hand, their dynamics depend on small-scale processes, such as three-dimensional turbulence and microphysics, but on the other hand, they respond to and shape larger-scale atmospheric circulations. Thus, in coarse resolution GCMs, boundary-layer turbulence, convection, and clouds must be parameterized in terms of a GCM's resolved fields. Inadequacies of these parameterizations are at the core of climate model uncertainties.

At the same time as the physical and computational complexity of GCMs has increased, expanding HPC capabilities have also enabled high-resolution, limited-area models, such as large-eddy simulations (LES), which directly resolve essential dynamics of turbulence, convection, and clouds. LES are being run with ever increasing resolution, on ever larger domains, and for ever longer durations. The ability of LES to provide high-fidelity simulations of boundary-layer turbulence, convection, and clouds is well established (?, ?, ?, ?, ?), making it a primary numerical laboratory for informing and validating climate model parameterizations. However, LES have so far mostly been used in selected locations, usually associated with field campaigns, such as the often studied BOMEX (?, ?), DYCOMS-II (?, ?), and RICO (?, ?) field campaigns. Some previous studies have used LES with idealized boundary conditions and large-scale forcings to investigate boundary layer dynamics in different dynamical regimes (e.g., ?, ?). Despite these attempts, the potential of LES to provide more systematic information for parameterizations across a broad cross-section of realistic dynamical regimes in a GCM has not been fully realized.

Here we describe an LES experimental design that allows us to explore a large range of atmospheric conditions encountered in a GCM. While driving limited-area models in idealized conditions in the tropics is well established (?, ?, ?), driving limited-area models with large-scale forcing from a GCM or a numerical weather prediction model in principle can be done anywhere in the atmosphere. However, it is less well established how to do so; see the discussion of potential pitfalls in ? (?). A framework to drive LES with host models (GCMs or weather models) or reanalyses data is described by ? (?). In this framework, single-column models (SCMs) and/or LES are driven by prescribed large-scale forcing from host models or reanalyses, while being nudged toward the state of the host model or reanalyses. The nudging prevents the LES from drifting away from the host model or reanalysis state. But it also constrains the turbulence developed by the LES, which limits the independent insights that can be gained from the LES in some dynamical regimes. Some early studies have investigated subtropical low cloud response to climate change with two-dimensional LES driven by prescribed large-scale forcings from a superparameterized GCM, while moisture nudging and adjustment of vertical motions are used to prevent model drift (?, ?, ?). The large-scale forcing approach has been used in recent years to study boundary layer dynamics and cloud feedbacks in selected locations and over long times (e.g., ?, ?, ?, ?). Another approach to force LES with GCMs is one-way nesting (e.g., ?, ?, ?), where states at lateral boundaries of the LES domain are relaxed toward those in a host model. The nesting approach allows for heterogeneous boundaries. But it requires LES on large domains, which is computationally expensive. In this study, we drive LES by dynamically consistent large-scale forcings given by a GCM and allow LES to evolve freely, without nudging of thermodynamic variables in the atmosphere. We use the LES systematically to simulate clouds in different dynamical regimes.

To prove the concepts and for simplicity in illustrating them, the GCM here is the idealized moist GCM used in a number of studies of large-scale dynamics going back to ? (?) and ? (?). The GCM has simplified radiation and convection schemes, and there are no clouds in the model. We treat each LES as representing an idealized single grid column of the host GCM by prescribing terms in the water and thermodynamic budgets resolved by the GCM as forcing within the LES. The treatment of radiative transfer is identically simplified between the LES and GCM, allowing the LES to achieve a physically realizable atmospheric energy balance that is consistent with the GCM (albeit not necessarily realistic). Because the GCM's resolved-scale fluxes that are applied to the LES do not directly include contributions from parameterized turbulent and convective processes in the GCM, the LES responds by generating a turbulent flow that closes the budgets. To the extent that the parameterizations in the GCM are an accurate representation of the dynamics resolved by the LES, the LES and GCM should reach the same statistically steady states. The extent to which they do not do so reflects inadequacies in the parameterizations in the GCM, and the mismatches can be used to improve the parameterizations.

One unique aspect of this experimental design is that it allows an iterative workflow to refine parameterizations that are implemented in a GCM. The iterative workflow consists of the following steps:

1. Integrate the GCM to generate forcing data from selected GCM grid columns.
2. Integrate LES driven by forcing from the GCM columns.

3. Improve the GCM parameterizations by minimizing mismatches between the LES and the GCM simulations for the selected columns.
4. Go to step 1.

GCM parameterizations can learn from offline LES simulations using data assimilation and machine learning approaches, and eventually, this workflow may be automated by spinning LES off the GCM simulation. Furthermore, experimental design approaches can be used to optimize the selection of GCM grid columns so as to be maximally informative about the parameterizations (?, ?). We do not demonstrate this workflow in this paper; rather, this paper is a contribution within a larger research program with the goal to realize such an automated workflow.

The rest of the paper is organized as follows. Section 2 discusses the GCM and LES used to simulate the transition from shallow to deep convection across a Walker circulation. Section 3 describes how the GCM grid column budgets of temperature and water vapor are decomposed and then used to force the LES. Section 4 describes the simulation characteristics in this forcing framework. Section 5 summarizes the conclusions and discusses the potential application of this framework to a range of cloud-climate problems.

## 2 Model Descriptions

### 2.1 Moist Idealized General Circulation Model

The GCM simulations used in this work are performed using the moist idealized GCM described in ? (?) and ? (?). Moist convective processes are represented using a simplified quasi-equilibrium (Betts–Miller) scheme (?, ?), which assumes convection relaxes convectively unstable profiles towards moist-adiabatic reference profiles with a relative humidity of 70%. Precipitation is formed both by the parameterized convection and by large-scale condensation when a GCM grid box reaches saturation. Only liquid water (rain) is considered, and precipitation is assumed to fall instantaneously to the surface. There are no clouds in the GCM. A $k$-profile eddy diffusivity scheme similar to that developed by ? (?), with a dynamically determined boundary layer height, is used to model turbulent transport within the atmospheric boundary layer. The atmosphere is assumed to overlay a 40-m deep slab ocean, with which it interacts through radiative transfer and through latent and sensible heat fluxes. Surface fluxes of heat, moisture, and momentum are modeled using bulk aerodynamic formulas with drag coefficients determined from Monin-Obukhov similarity theory, with a roughness lengths of $1\times10^{-5}$ m for all quantities. Atmospheric radiative transfer is represented by a two-stream semi-gray radiation scheme, where longwave and shortwave absorption is calculated using prescribed latitudinally dependent absorber profiles, as described in ? (?). The model does not have seasonal cycles or diurnal cycles.

The GCM simulations we perform are of a zonally asymmetric aquaplanet, in which zonal asymmetry is imparted by prescribing an ocean heat flux divergence (see **??**). The prescribed ocean heat flux divergence produces an off-equatorial Walker circulation, with suppressed convection in the region of maximum ocean heat flux divergence and enhanced convection in the region of maximum ocean heat flux convergence. Having a Walker circulation allows the LES driven by the aquaplanet GCM to generate somewhat more realistic tropical cloud regimes than would be possible in a zonally symmetric setting. The GCM is run at T42 horizontal spectral resolution with 72 vertical levels (30 levels in the lowest 1.5 km) for a total of 3000 days. We accumulate forcing data to drive the LES from the last 500 days of the simulation. The dipole pattern of the prescribed ocean heat flux divergence (Figure **??**a) induces zonal variations in surface temperatures (Figure **??**b) in the GCM. The zonally varying surface temperatures in turn drive an atmospheric Walker circulation, with strong ascent over warm temperatures and large-scale subsidence over cold temperatures (Figure **??**). To drive the LES, we extract forcing data from selected GCM grid-columns along a transect connecting the points of maximum and minimum ocean heat flux divergence (dots in Figure **??**a).

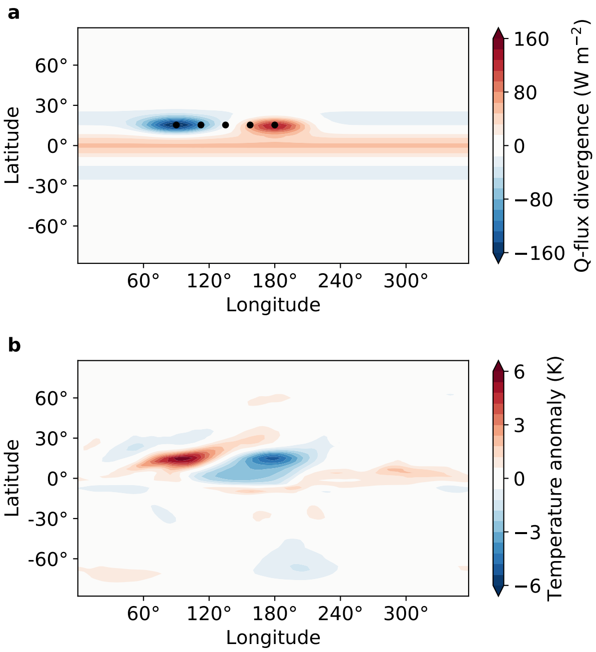

**Figure 1.** Surface fields in the GCM simulation. (a) Ocean heat flux divergence. (b) Surface temperature anomaly about zonal mean. Black circles indicate GCM grid columns along a transect connecting the regions of maximum and minimum ocean heat flux divergence.

## 2.2 Large-Eddy Simulation

The forcing framework is implemented in the Python Cloud Large-Eddy Simulation (PyCLES) code (?, ?). PyCLES is a parallel, three-dimensional LES code that solves the moist anelastic equations of motion (?, ?), with prognostic equations for moist specific entropy $s$, total non-precipitating water specific humidity $q_t$, precipitating liquid water $q_r$ (rain), and precipitating frozen water $q_s$ (snow). The total water specific humidity includes contributions from water vapor $q_v$, non-precipitating liquid water $q_l$, and non-precipitating ice $q_i$, which are assumed to be in thermodynamic equilibrium. Along with dry air, these four components make up the thermodynamic system. The precipitating species are not assumed to be in thermodynamic equilibrium with $q_t$.

In PyCLES, the equations of motion are solved on a staggered, Arakawa-C grid (?, ?). The advection of both scalar and momentum fields are discretized using nominally fifth-order weighted essentially non-oscillatory (WENO) schemes implemented as described in ? (?) and based on the schemes developed by ? (?). Because of grid staggering, the actual accuracy of the WENO schemes is limited to second order. The pressure is diagnosed following ? (?), and time-stepping is performed using a second-order strong stability preserving Runge-Kutta scheme (?, ?). The numerical formulation does not ensure monotonicity or positivity of solutions, so clipping is employed when needed to ensure that fields such as specific humidities that need to remain positive do indeed remain positive. Subgrid-scale (SGS) fluxes of momentum and scalars are modeled though the Smagorinsky-Lilly closure (?, ?, ?), with the Smagorinsky coefficient $c_s = 0.17$ and the

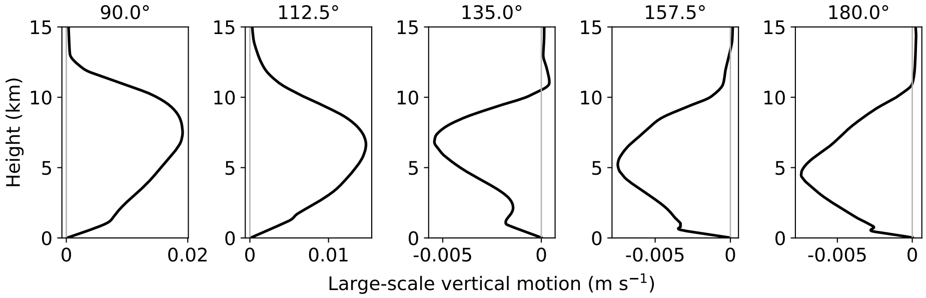

**Figure 2.** Vertical profiles of large-scale vertical motion in the GCM along the Walker circulation transect. The longitudes of the simulations are indicated above each panel.

turbulent Prandtl number $\mathrm{Pr}_t = 1/3$. A Rayleigh friction sponge layer near the LES top is used to absorb waves, with horizontal velocities damped toward the domain mean with the height-dependent damping coefficient described in **??**.

The formation and fate of precipitating species is represented using a single-moment, bulk microphysical model based on the scheme developed by ? (?); differences between their scheme and that used here are enumerated in **??**. Precipitation that reaches the surface is assumed to fall out of the atmosphere, with no transfer of energy or mass to the ocean. The LES are forced with prescribed sea surface temperatures from the GCM. It is tempting to allow the surface temperatures in LES to vary by coupling the simulations with a slab ocean; however, we found that doing so leads to significant drift of the LES away from the GCM state in deep convection regions, presumably because of inadequacies of the simple turbulence and convection parameterizations employed in the GCM. Surface fluxes are modeled using bulk aerodynamic formulas based on Monin-Obukhov similarity theory. Radiative transfer is represented identically to the scheme used in GCM, with no effects of clouds on the radiative transfer.

We run LES driven by the GCM fields into a statistically steady state. This is computationally expensive for several reasons. First, we simulate a large, three-dimensional, doubly periodic domain that is 64 km wide and 25.6 km high, to be able to simulate deep convection. Second, the timescale for the LES to reach a statistically steady state is set by the timescale over which the atmosphere equilibrates, which can take tens of days. To reduce the computational cost we run the LES at a relatively coarse resolution of 250 m. The vertical grid is stretched and consists of 256 grid points. The vertical resolution decreases from about 50 m near the surface to around 200 m at the domain top. The simulations are run for 60 days (90.0° and 112.5°), 90 days (135.0°), or 180 days (157.5° and 180.0°). The results in this study are averaged over the last 15 simulated days, when the simulations have reached statistically steady states.

## 3 Forcing Framework

### 3.1 GCM Equations

Our goal is to develop an LES forcing framework in which the LES emulates a single grid column of a GCM, directly resolving processes that are parameterized in the GCM while satisfying the same large-scale water and thermodynamic balances as the GCM. To achieve this, we begin by considering the equations for the grid-scale temperature $\tilde{T}$ and specific humidity $\tilde{q}_t$ that are resolved by the GCM. Tildes $\widetilde{(\cdot)}$ denote variables resolved on the GCM grid. The thermodynamic equation in the GCM's $\sigma$ coordinates is given by (?, ?)

$$\underbrace{\partial_t \tilde{T} + \tilde{u}\partial_x \tilde{T} + \tilde{v}\partial_y \tilde{T} + \tilde{\varpi}\partial_\sigma \tilde{T} - \frac{\tilde{\alpha}\tilde{\omega}}{c_p}}_{\text{GCM−Resolved Dynamics}} = \underbrace{\tilde{J}_{\text{conv}} + \tilde{J}_{\text{cond}} + \tilde{J}_{\text{diff}} + \tilde{J}_{\text{numerical}} + \tilde{J}_{\text{rad}}}_{\text{GCM Parameterized}}, \tag{1}$$

which we have written such that the left hand side of the equation is the sum of the time-tendency of temperature and GCM-resolved advection terms, while the right hand side consists of parameterized source terms arising from the deep convection scheme ($\tilde{J}_{\text{conv}}$), large-scale condensation ($\tilde{J}_{\text{cond}}$), diffusive parameterizations such as the boundary layer turbulence scheme ($\tilde{J}_{\text{diff}}$), the radiation scheme ($\tilde{J}_{\text{rad}}$), and numerical damping and spectral filtering ($\tilde{J}_{\text{numerical}}$). Here, $\sigma = \tilde{p}/\tilde{p}_s$, with pressure $\tilde{p}$ and surface pressure $\tilde{p}_s$, $\tilde{\varpi} = d\sigma/dt$, $\tilde{\omega} = d\tilde{p}/dt$, $\tilde{\alpha}$ is the specific volume, and $c_p$ is the specific heat capacity of dry air; the notation is otherwise standard. (We use Cartesian coordinates to simplify the notation; however, the GCM uses spherical coordinates, whereas the LES is Cartesian.) Similarly, the GCM's specific humidity budget is given by

$$\partial_t \tilde{q}_t + \underbrace{\tilde{u}\partial_x \tilde{q}_t + \tilde{v}\partial_y \tilde{q}_t + \tilde{\varpi}\partial_\sigma \tilde{q}_t}_{\text{GCM−Resolved Dynamics}} = \underbrace{\tilde{S}_{\text{conv}} + \tilde{S}_{\text{cond}} + \tilde{S}_{\text{diff}}}_{\text{GCM Parameterized}}. \tag{2}$$

In a statistically steady state, the explicit time derivatives vanish upon long-term time averaging, and the parameterized processes on the right-hand sides balances with resolved-scale flows of $\tilde{T}$ and $\tilde{q}_t$ on the left-hand sides. Our goal is to impose the resolved-scale budget terms on the left-hand side of equations (**??**) and (**??**) on the LES, while allowing the LES to generate a three-dimensional turbulent flow, microphysical processes, and radiative energy fluxes that determine the terms on the right-hand sides of the same equations. Here we implement long-time mean forcing from statistically steady states of the GCM, without the explicitly time dependent terms; however, it is straightforward to drive the LES using time-varying forcing from the GCM consistent with equations (**??**) and (**??**). We limit the present study to simulations using long-time mean forcing because time varying forcing includes synoptic scale variability which would require much longer LES simulations to achieve statistically robust results.

### 3.2 Specific Humidity Forcing

In the LES, we modify the prognostic equation for total water specific humidity to include additional source terms $S_{\text{hadv}}$ and $S_{\text{vadv}}$ arising from GCM-resolved horizontal and vertical advection, giving

$$\begin{aligned}\partial_t q_t + \partial_x (u q_t) + \partial_y (v q_t) + \rho_0^{-1}\partial_z (\rho_0 w q_t) = -\partial_x (\gamma_{q,x}) - \partial_y (\gamma_{q,y}) - \rho_0^{-1}\partial_z (\rho_0 \gamma_{q,z}) \\ + E - P + S_{\text{hadv}} + S_{\text{vadv}},\end{aligned} \tag{3}$$

where $\rho_0$ is the reference density, $\gamma_{q,x}$, $\gamma_{q,y}$, and $\gamma_{q,z}$ are the SGS fluxes of $q_t$, and $E$ and $P$ are source terms due to evaporation and production of hydrometeors. Consistent with ? (?), we prescribe the horizontal advective source term directly from the GCM such that

$$S_{\text{hadv}} = -\langle \tilde{u}\partial_x \tilde{q}_t \rangle - \langle \tilde{v}\partial_y \tilde{q}_t \rangle, \tag{4}$$

where $\langle \cdot \rangle$ indicates a long-time mean on $\sigma$ surfaces in the statistically steady state of the GCM.

Specification of the vertical advection source term is more complicated because the LES produces its own vertical advection, acting on its own domain-mean vertical gradients. We compute the vertical advection source term by rewriting the vertical advection term in equation (**??**) using the hydrostatic relation and decomposing it into time mean and fluctuating components as

$$\langle \tilde{\varpi}\partial_\sigma \tilde{q}_t \rangle \approx \langle \tilde{w} \rangle \partial_z \langle \tilde{q}_t \rangle + \langle \tilde{w}' \partial_z \tilde{q}_t' \rangle, \tag{5}$$

where $w = dz/dt$ is the vertical velocity and primes $(\cdot)' = (\cdot) - \langle\cdot\rangle$ denote fluctuations about the time mean $\langle\cdot\rangle$. We approximate the time-averaged vertical velocity by $\langle\tilde{w}\rangle \approx -\langle\tilde{\omega}\rangle\langle\tilde{\alpha}\rangle/g$ where we have assumed that $\tilde{\omega}$ and $\tilde{\alpha}$ are uncorrelated. Replacing the GCM specific humidity $\tilde{q}_t$ with the LES specific humidity $q_t$ then gives the vertical advection source term we apply to the LES:

$$S_{\mathrm{vadv}} = -\langle\tilde{w}\rangle\partial_z q_t - \langle\tilde{w}'\partial_z\tilde{q}_t'\rangle. \tag{6}$$

The first term on the right-hand side accounts for vertical advection of $q_t$ by GCM-resolved time-mean vertical motion. The second term accounts for vertical fluxes associated with fluctuations about the long-time mean and is prescribed directly from the GCM; it is typically much smaller than the mean component.

### 3.3 Specific Entropy Forcing

We modify the LES prognostic equation for moist specific entropy to include source terms owing to GCM-resolved transports of total water specific humidity and temperature, such that

$$\begin{aligned}\partial_t s + \partial_x(us) + \partial_y(vs) + \rho_0^{-1}\partial_z(\rho_0 ws) = &-\partial_x(\gamma_{s,x}) - \partial_y(\gamma_{s,y}) - \rho_0^{-1}\partial_z(\rho_0\gamma_{s,z}) \\ &+ \frac{c_p}{T}J_{\mathrm{rad}} + \dot{S} + \frac{c_p}{T}(J_{\mathrm{hadv}} + J_{\mathrm{vadv}}) + (s_v - s_d)(S_{\mathrm{hadv}} + S_{\mathrm{vadv}}),\end{aligned} \tag{7}$$

where $\gamma_{s,x}$, $\gamma_{s,y}$, and $\gamma_{s,z}$ are the SGS fluxes of $s$ and $J_{\mathrm{rad}}$ is the source term due to radiation. The term $\dot{S}$ represents irreversible entropy sources associated with the SGS dynamics and precipitation processes, as described in ? (?). The last two terms on the right-hand side arise from the GCM-resolved advective tendencies of temperature ($J_{\mathrm{hadv}}$ and $J_{\mathrm{vadv}}$) and total water specific humidity ($S_{\mathrm{hadv}}$ and $S_{\mathrm{vadv}}$), and $s_v$ and $s_d$ are specific entropies of dry air and water vapor, respectively. Consistent with ? (?) and the treatment of total water specific humidity in section **??**, we take the horizontal advective temperature tendency $J_{\mathrm{hadv}}$ directly from the GCM, such that

$$J_{\mathrm{hadv}} = -\langle\tilde{u}\partial_x\tilde{T}\rangle - \langle\tilde{v}\partial_y\tilde{T}\rangle + \langle\tilde{J}_{\mathrm{numerical}}\rangle. \tag{8}$$

The vertical advective temperature tendency $J_{\mathrm{vadv}}$, including the pressure-volume work term $\tilde{\alpha}\tilde{\omega}/c_p$, is derived by using the hydrostatic approximation to approximate the last term on the left-hand side of (**??**),

$$\left\langle \tilde{\varpi}\partial_\sigma\tilde{T} - \frac{\tilde{\alpha}\tilde{\omega}}{c_p} \right\rangle \approx \langle\tilde{w}\rangle\partial_z\langle\tilde{T}\rangle + \langle\tilde{w}'\partial_z\tilde{T}'\rangle + \langle\tilde{w}\rangle\frac{g}{c_p}. \tag{9}$$

Here, we have used again the approximation $\langle\tilde{w}\rangle \approx -\langle\tilde{\omega}\rangle\langle\tilde{\alpha}\rangle/g$. Replacing the GCM temperature $\tilde{T}$ with the LES temperature $T$ in the first term on the right-hand side then gives the vertical transport applied to the LES as

$$J_{\mathrm{vadv}} = -\langle\tilde{w}\rangle\partial_z T - \langle\tilde{w}'\partial_z\tilde{T}'\rangle - \langle\tilde{w}\rangle\frac{g}{c_p}, \tag{10}$$

with the second and third terms on the right-hand side taken directly from the GCM.

While we have discussed the forcing framework for an LES using specific entropy as a prognostic variable, it is straightforward, and in some ways easier, to implement it in models using other formulations of moist thermodynamics. For example, in an LES using some variety of liquid-ice potential temperature as a prognostic variable, the formulation of the specific humidity forcing is identical to that described here, and the GCM-resolved transport terms in the potential temperature equations are analogous to the terms involving $J_{\mathrm{hadv/vadv}}$ on the right-hand side of equation (**??**) but with the inverse Exner function replacing $1/T$.

### 3.4 Momentum Forcing

The LES horizontal momentum field is driven by the large-scale pressure gradient from the GCM and Coriolis acceleration. The LES horizontal momentum equations are modified to be

$$\partial_t u + \partial_x (uu) + \partial_y (uv) + \rho_0^{-1}\partial_z (\rho_0 uw) = -\partial_x \left(\rho_0^{-1} p'\right) \\ - \partial_x \tau_{xx} - \partial_y \tau_{xy} - \rho_0^{-1}\partial_z (\rho_0 \tau_{xz}) - f \left(\langle \tilde{v}_g \rangle - v\right), \quad (11)$$

and

$$\partial_t v + \partial_x (uv) + \partial_y (vv) + \rho_0^{-1}\partial_z (\rho_0 vw) = -\partial_y \left(\rho_0^{-1} p'\right) \\ - \partial_x \tau_{yx} - \partial_y \tau_{yy} - \rho_0^{-1}\partial_z (\rho_0 \tau_{yz}) + f \left(\langle \tilde{u}_g \rangle - u\right) \quad (12)$$

where $\tau$ represents the SGS stress, and $\langle \tilde{u}_g \rangle$ and $\langle \tilde{v}_g \rangle$ are the mean zonal and meridional geostrophic winds from the GCM.

The prescribed momentum forcing leads to strong inertial oscillations in upper troposphere in the LES. In order to reduce the magnitude of the oscillation, the LES mean horizontal winds are relaxed to GCM mean in the free troposphere, with a timescale of 2 days.

### 3.5 Radiation

One essential aspect of this framework is that the representation of radiative transfer implemented in the LES should be identical to that implemented in the GCM. This enables controlled numerical experiments that isolate differences between resolved turbulence, convection, and clouds in the LES from their parameterized counterparts in the GCM. Furthermore, using the same representation of radiative transfer guarantees that, at least initially, free-troposphere large-scale transport imposed on the LES from the GCM is balanced by radiative heating/cooling, thus preventing the free-troposphere temperature from drifting. This is why we employ the same two-stream gray radiation scheme, without cloud radiative effects, in both the LES and the GCM.

### 3.6 Surface fluxes

Surface fluxes are modeled using a bulk scheme with drag coefficients obtained from Monin-Obukhov similarity theory (?, ?). Since we are using the time-mean forcing from the GCM, the variation in surface wind speed in the LES is much smaller than in the GCM. This leads to significant differences in surface fluxes between the GCM and LES, which affects the temperature profile. To account for synoptic variability, we modify the surface wind speed to add a gustiness when calculating surface fluxes in the LES. The gustiness is estimated by the standard deviation of 6-hourly wind speed at the lowest model level in the GCM, which decreases from 3.20 m s$^{-1}$ over the warmest surface temperature to 1.49 m s$^{-1}$ over the coldest surface temperature.

### 3.7 Interpolation

The GCM and LES use different vertical grids ($\sigma$ vs. $z$ coordinates). Therefore, vertical profiles extracted from the GCM must be interpolated onto the LES vertical grid. We do so by interpolating from the GCM's $\sigma$ levels onto the LES's height levels using the time-mean height of each $\sigma$ level at the horizontal location in question. This represents an additional approximation because averages at constant $\sigma$ and constant $z$ differ. One of two interpolation methods is used depending on whether the gradient of the interpolated field appears explicitly in the forcing. For variables that do not appear inside a gradient (e.g., $u$, $v$, or $J_{\mathrm{hadv}}$), the GCM profiles are interpolated using a monotonic piecewise cubic Hermite interpolating polynomial (PCHIP). For interpolated variables that do appear inside a gradient (e.g., $q_t$ or $T$), the vertical derivative of the field on the GCM grid is interpolated to the LES vertical grid using a PCHIP interpolation, and the value of the field is recovered by numerically integrating the interpolated derivative using the trapezoidal rule. This approach ensures a nearly monotonic interpolation of discrete derivatives of the field.

### 3.8 LES Initial Condition

The LES initial conditions are specified from the GCM statistically steady-state mean vertical profiles of temperature, specific humidity, and horizontal geostrophic wind speed. Starting with this initial condition helps because it ensures that where parameterized processes in the GCM vanish, the GCM-resolved forcing terms balance (up to small differences arising mostly from numerical interpolation error), provided the treatment of radiative transfer is identical in both models.

## 4 Simulation Characteristics

Figure **??** shows the timeseries of outgoing longwave radiation in the LES and illustrates the time it takes to equilibrate to a statistically steady state. The simulations reach steady state more rapidly in deep convection regions than in shallow convection regions, and all cases are in statistically steady state by the end of the simulations. As the LES are free to evolve from their initial condition, which is specified to be the long-time mean of the GCM, their equilibrated solution can diverge substantially from their initial state. The differences between the equilibrated LES solution and GCM long-time means can be used to identify deficiencies in GCM parameterizations.

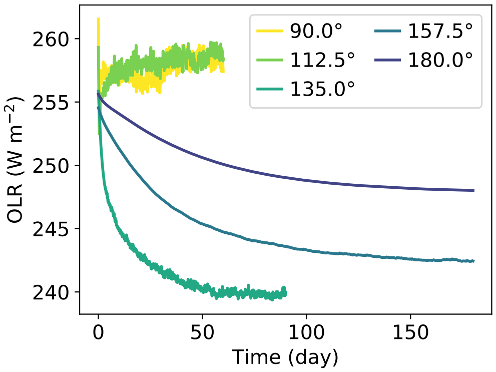

**Figure 3.** Timeseries of LES outgoing longwave radiation (OLR) at various longitudes along the Walker circulation transect.

Figure **??** shows the equilibrium temperature and relative humidity profiles in the GCM and in the LES along the Walker circulation transect. The temperature profiles are in general very similar between the GCM and LES, except that in the free troposphere the LES is slightly cooler than the GCM at 135.0° longitude, which is consistent with the lower OLR in the LES (Figure **??**). There are larger differences in the relative humidity profiles. The GCM-simulated relative humidity in deep convection regions is close to 70% in the free troposphere. The relative humidity has a local maximum of about 90% near 1 km and decreases to about 60% near the surface. The relative humidity in the LES is higher than that in the GCM both in the free troposphere and near the surface. The difference in relative humidity between GCM and LES arises because in deep convection regions, convection is occurring most of the time, and the difference shows the deficiency of the simple GCM convection scheme. In the LES, convection keeps the relative humidity at a high value, while in the GCM, the simple convection scheme keeps the relative humidity close to the reference value (0.7). The jump in relative humidity at about 5 km in the LES results from the cooling effect of snow melting. In shallow convection regions, the relative humidity in the GCM has local maxima near the surface and the tropopause and a local minimum in the middle of the troposphere. The LES produces a much moister free troposphere at 135.0° longitude, which may be related to the lower tropospheric temperature in the LES. The differences between the GCM and LES at 135.0° longitude may partially result from the lack of coupling to large-scale motions in the LES. The simulations with lower surface temperatures (157.5° and 180.0°) in general resemble the GCM, with a slightly larger relative humidity in the LES.

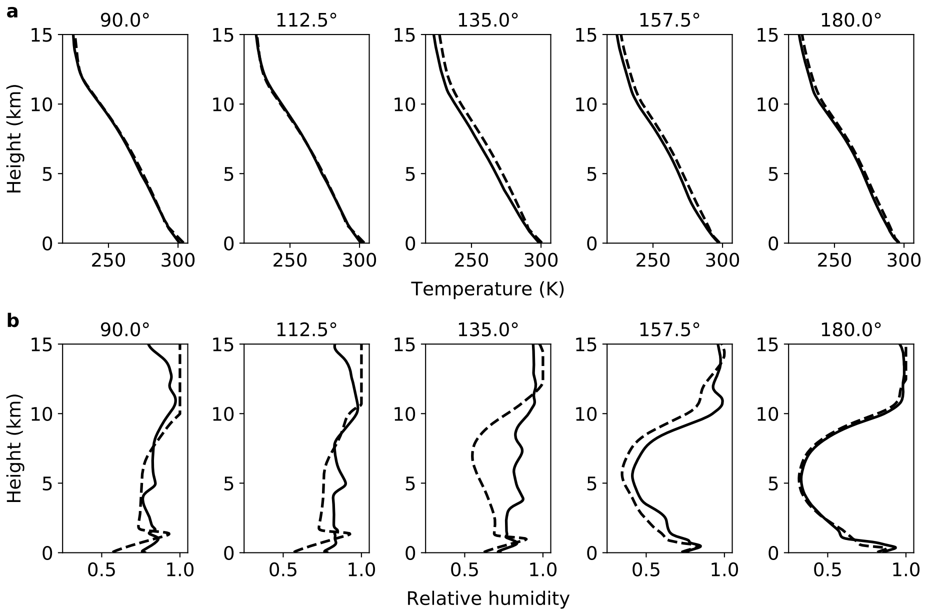

**Figure 4.** Vertical profiles of (a) temperature and (b) relative humidity along the Walker circulation transect in the GCM (dashed) and in the LES (solid). The relative humidity is the weighted average with respect to liquid and ice. The longitudes of the simulations are indicated above each panel.

Figure **??** shows the equilibrium surface fluxes and surface precipitation in the GCM and LES. The latent heat flux decreases from about 300 W m$^{-2}$ in deep convection regions to less than 100 W m$^{-2}$ in shallow convection regions in the GCM. The sensible heat flux and net longwave radiative flux show similar differences between deep and shallow convection regions, although the magnitudes are smaller. In the LES, the latent heat flux is weaker by 20–25% in deep convection regions, consistent with the higher relative humidity in the boundary layer (Figure **??**b). The decrease in the latent heat flux is partially balanced by increases in the net upward longwave radiative flux and in the sensible heat flux, which result from an increase in the temperature difference between the surface and near-surface air. The differences in the sum of surface radiative and turbulent heat fluxes between the GCM and LES are 46 W m$^{-2}$ and 38 W m$^{-2}$ at 90.0° and 112.5° longitudes, respectively. The surface energy budget is not closed in LES simulations with fixed surface temperatures; the results imply that significant changes in ocean

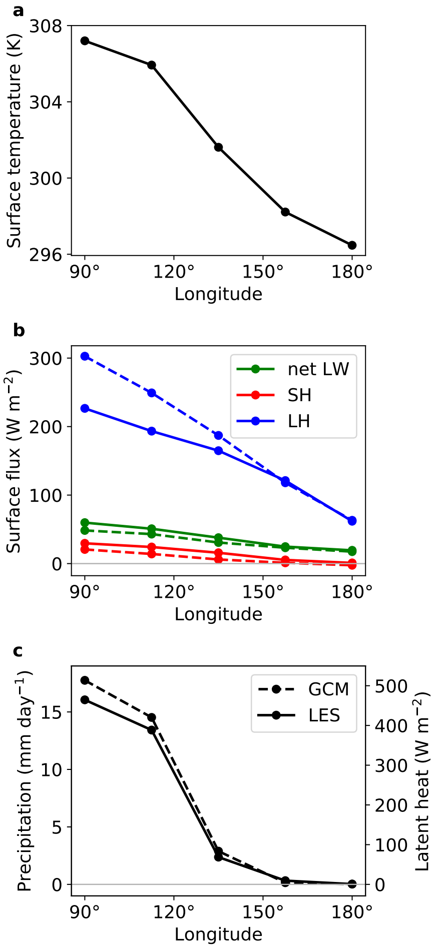

**Figure 5.** (a) Surface temperatures along the Walker circulation transect. (b) Equilibrium surface fluxes in the GCM (dashed) and LES (solid). Green lines denote net longwave radiative flux, red lines sensible heat flux, blue lines latent heat flux. (c) Surface precipitation in the GCM (dashed) and LES (solid). The latent heat associated with precipitation is shown on the right vertical axis.

heat uptake are required for LES simulations with a mixed-layer ocean to reach similar steady states. The surface fluxes in shallow convection regions are similar between the GCM and LES. The LES generally reproduces the spatial pattern of surface precipitation in the GCM, with strong precipitation in deep convection regions and weak or no precipitation in shallow convection regions along the transect. In deep convection regions, the surface precipitation is weaker in the LES than that in the GCM.

To further understand the difference between the GCM and LES, we compare vertically integrated moisture and dry enthalpy budgets (Figure **??**). Note that the tendencies from horizontal advection in the LES are directly prescribed and are the same as those in the GCM. Since we focus on a statistically steady state, the sum of different terms affecting the budget is approximately zero. In deep convection regions, the moisture budget is kept in balance mainly by vertical advection, precipitation, and evaporation. The decreased evaporation in the LES is balanced by the decreased precipitation and the strengthened vertical advection of water vapor out of the moister boundary layer (Figure **??**b).

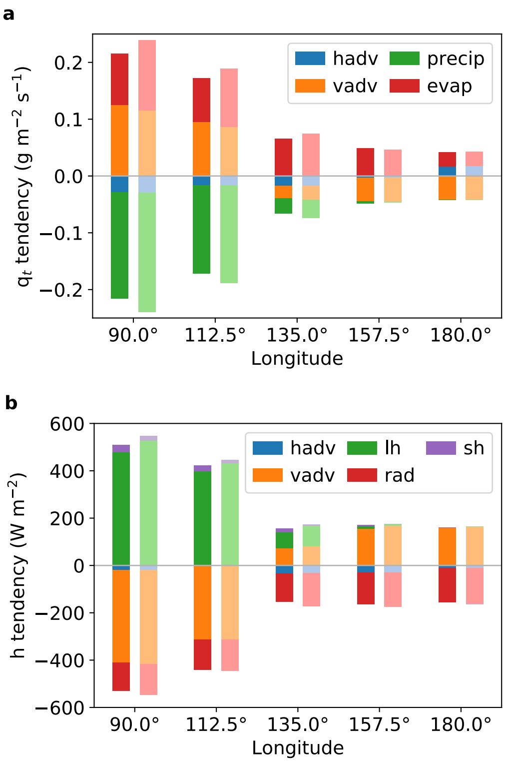

**Figure 6.** (a) Vertically integrated moisture tendencies from horizontal advection (hadv), vertical advection (vadv), precipitation (precip), and evaporation (evap). (b) Vertically integrated dry enthalpy tendencies from horizontal advection, vertical advection, latent heat of precipitation (lh), surface sensible heat flux (sh), and radiation (rad). Upward bars show sources and downward bars show sinks. LES and GCM budgets are shown in darker and lighter colors, respectively.

The dry enthalpy budget is kept in balance mainly by vertical advection, latent heat of precipitation, and radiation. The decrease in latent heat from precipitation in the LES is mainly balanced by the weaker radiative cooling and weaker adiabatic cooling from vertical motion. In the shallow convection regions, the main terms controlling the moisture budget are subsidence drying and evaporation, and the main terms controlling the dry enthalpy budget are subsidence warming and radiative cooling. The budgets are consistent between the GCM and LES in shallow convection regions.

Figure **??** shows the vertical profiles of different terms balancing moisture and temperature budgets at three longitudes along the Walker cell transect. At 90.0°, the moisture budget is mainly balanced by vertical advection and turbulent and convective drying, and the temperature budget is mainly balanced by vertical advection, turbulent and convective heating, and radiative cooling. The vertical structure of the tendencies is broadly consistent between the GCM and LES, with some differences in the boundary layer and near the melting level. The mismatch between the GCM and LES near the boundary layer top is related to the local minimum in moisture in the LES (Figure **??**b). At 135.0°, all three physical terms (horizontal advection, vertical advection, and turbulence and convection) contribute to balancing the moisture budget. The dominant balance in the temperature budget is between subsidence warming, turbulent and convective heating, and radiative cooling. There are stronger mismatches between the GCM and LES in the free

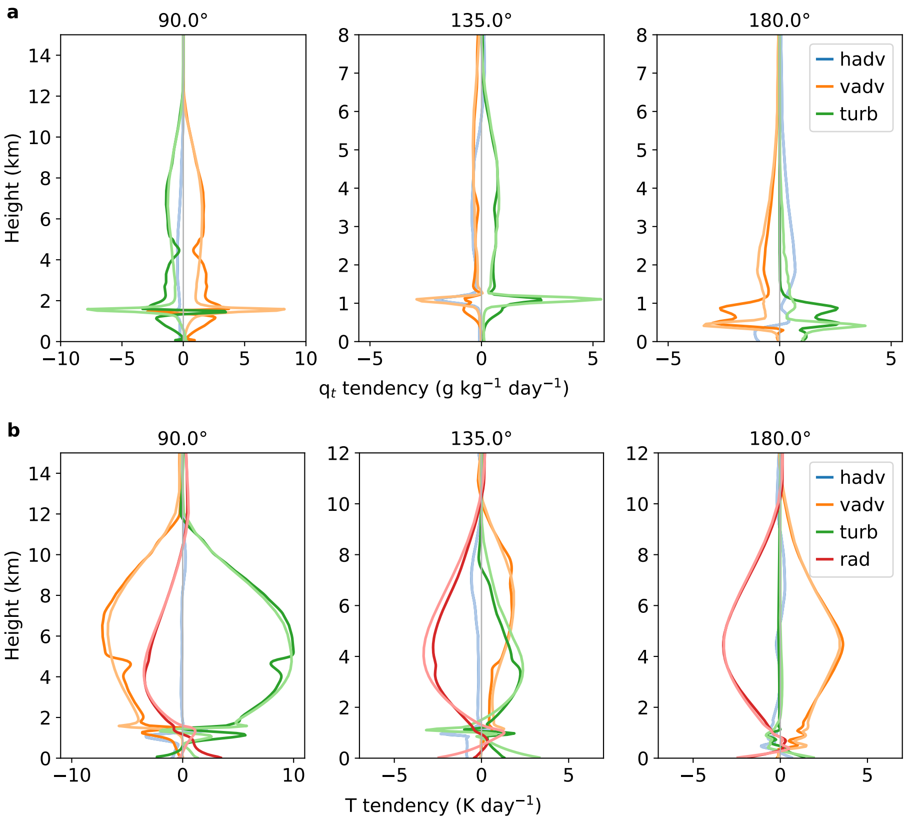

**Figure 7.** (a) Vertical profiles of moisture tendencies from horizontal advection (hadv), vertical advection (vadv), and turbulence and convection (turb). (b) Vertical profiles of temperature tendencies from horizontal advection, vertical advection, turbulence and convection, and radiation (rad). LES and GCM budgets are shown in darker and lighter colors, respectively. The longitudes of the simulations are indicated above each panel.

troposphere, which is related to the warm and dry biases in the GCM at 135.0° (Figure **??**). At 180.0°, the dominant balance in the moisture budget is between subsidence drying and turbulent and convective moistening in the cloud layer, and subsidence drying and horizontal advection in the free troposphere. The GCM–LES difference in the moisture budget below 1 km suggests that the GCM has a shallower boundary layer and weaker convection than the LES, which is associated with the smaller surface buoyancy flux dominated by the weaker sensible heat flux in the GCM. The temperature budget is mainly balanced by subsidence warming and radiative cooling. The GCM and LES agree well on the temperature budget at 180.0°, as expected from the similar temperature profiles (Figure **??**a).

Figure **??** shows the vertical profiles of cloud fraction and cloud water specific humidity along the Walker cell transect in the LES. The simulation with the coldest surface temperature (180.0°) produces a very shallow and thin cloud layer with cloud base around 300 m and cloud top around 1000 m. Below the cloud base, there is a well-mixed boundary layer, as seen in the vertical profiles of liquid-ice potential temperature and total water specific humidity (Figure **??**). The well-mixed boundary layer and the shallow-cumulus layer deepen as the surface temperature increases. The simulation at 135.0° longitude develops a secondary maximum in cloud fraction at about 4 km, which originates

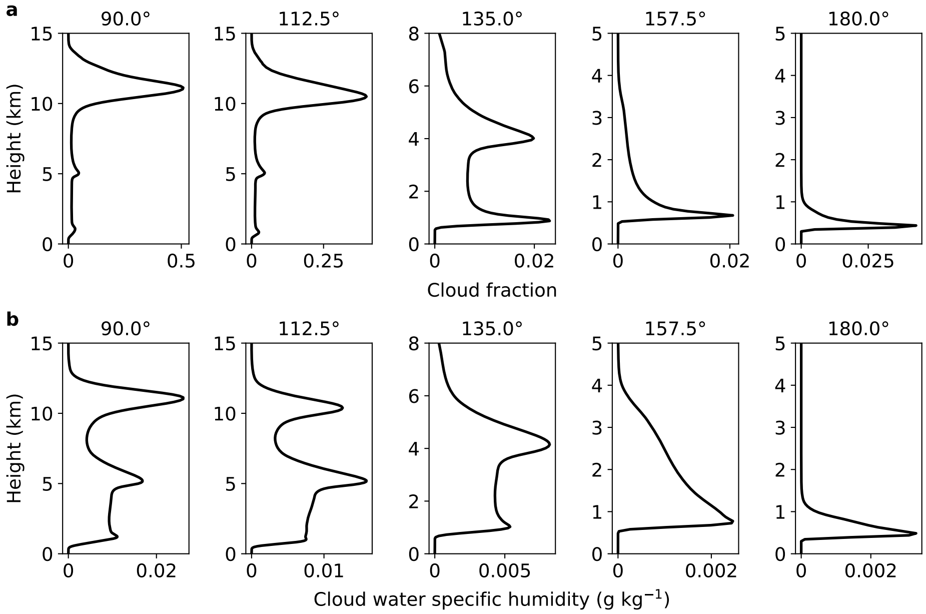

**Figure 8.** Vertical profiles of (a) cloud fraction and (b) cloud water specific humidity along the Walker circulation transect in the LES. The longitudes of the simulations are indicated above each panel.

from the cooling effects of snow melting. The cloud water specific humidity (the sum of the cloud liquid and ice specific humidities) has a similar vertical structure as the cloud fraction in shallow convection regions. As the surface temperature increases along the transect, the LES produce a transition to deep convection at 90.0° and 112.5° longitudes. The simulations produce anvil clouds with a peak in cloud fraction at about 11 km. There is a secondary maximum in cloud fraction near the melting level at about 5 km, which has also been reported in previous observational and high-resolution modeling studies (?, ?, ?). The vertical profile of cloud water specific humidity is more bottom-heavy than that of cloud fraction, with a more obvious peak near the melting level. In general, the LES produce little cloud ice in the upper troposphere in deep convection regions, which may be related to the relatively simple microphysics scheme and the lack of cloud-radiation interactions (?, ?) in this study.

## 5 Discussion and Conclusions

We have described a framework in which LES are driven by large-scale forcing taken from an idealized aquaplanet GCM, in a one-way coupling setup. The LES are forced with prescribed surface temperatures from the GCM, but they are otherwise freely evolving, without the direct constraints on temperature or humidity profiles that are usually imposed in LES studies. Instead, we imposed GCM-resolved, large-scale energy and moisture budget terms as forcing terms in the LES. We also imposed the large-scale pressure gradient from the GCM to drive the LES horizontal wind, and we relax the mean LES horizontal winds to those from the GCM to damp inertial oscillations that otherwise arise. Both the LES and GCM include the same radiative transfer schemes. This allows us to generate LES in grid columns of GCMs and to provide suitable benchmarks for testing and calibration of parameterization schemes in GCMs.

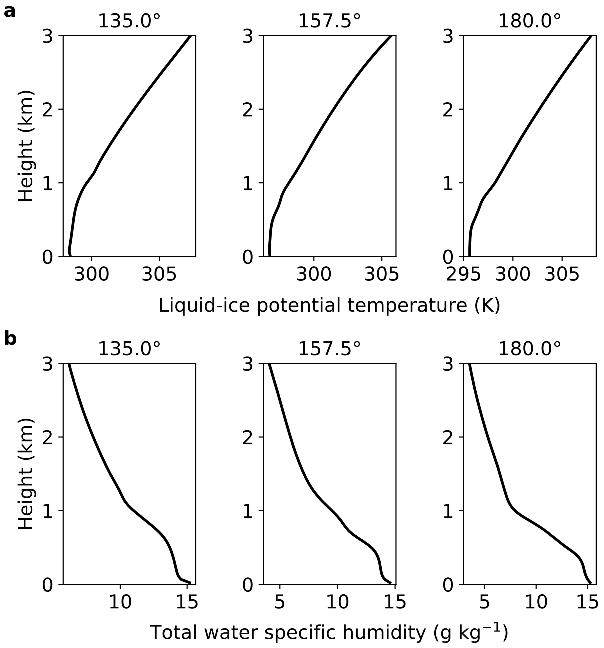

**Figure 9.** Vertical profiles of (a) liquid-ice potential temperature and (b) total water specific humidity in shallow convection regions. The longitudes of the simulations are indicated above each panel.

To illustrate the concepts, we used this forcing framework to simulate the transition from shallow to deep convection along an idealized Walker circulation transect. The Walker circulation is induced by prescribing a dipole structure of ocean heat flux convergence/divergence in the slab ocean of the GCM. LES driven by large-scale forcing from the GCM reach statistically steady states without nudging of thermodynamic variables toward reference profiles. The LES are able to reproduce the longitudinal variation of relative humidity, surface fluxes, and precipitation, with some mismatches to the GCM climatology, especially in deep convection regions. The mismatches are indicative of inadequacies in the GCM's convection parameterization, which is unsurprising given the simplicity of the parameterization. The mismatches can be used to systematically improve parameterization schemes, for example, with Bayesian calibration methods (?, ?, ?).

Our idealized setup has limitations. The gray radiation scheme we used to prove the concept does not depend on water vapor concentration or cloud properties. This distorts the radiative driving of the flows relative to what would happen in reality. The coupling of clouds, water vapor, and radiation has been shown to affect the Intertropical Convergence Zone and anvil clouds in deep convection regions (?, ?, ?, ?), as well as boundary layer clouds, the diurnal cycle of shallow cumulus, and the stratocumulus-to-cumulus transition (?, ?, ?). The resolution of LES used in this study (250 m) is not fine enough to fully resolve low clouds. Deep convective clouds are sensitive to microphysical processes, which are parameterized in a simple way in our model. These factors may contribute to the relatively small cloud fraction and cloud water path in our simulations of shallow cumulus, and to the relatively small amount of cloud ice in deep convection regions. These shortcomings limit the applicability of the results to some extent. But the simplifications facilitated the development and testing of the framework and suffice in a proof-of-concept that illustrates its usability. They also simplify interpretation, for example, of the cloud response to climate change.

We are currently using this framework to study cloud–climate feedbacks in an idealized setting and to generate a library of LES simulations for developing and training more comprehensive GCM parameterizations than the ones we used here. A library containing a wide range of LES across different cloud regimes, including LES of changed climates, provides a powerful opportunity to systematically train complex dynamical parameterizations, with minimal imprinting of the GCM's parameterizations onto the LES forcing (the primary imprinting occurs through the dependence of the large-scale advection terms on the dynamical parameterizations). It thereby enables direct and meaningful tests of the parameterizations against LES. An LES library including simulations of changed climates not only enables tests of parameterizations but also provides an opportunity to investigate cloud feedbacks in LES that are not subject to nudging of thermodynamic variables to reference profiles or ad hoc assumptions about how large-scale dynamics change with climate. The results of climate change simulations will be described in a companion paper.

## Appendix A Ocean Heat Flux

The time-independent ocean heat flux divergence ($O$) prescribed in the GCM is of the following form:

$$\begin{aligned} O(\phi,\eta) = {} & Q_0\left(1-\frac{2\phi^2}{\phi_0^2}\right)\exp\left(-\frac{\phi^2}{\phi_0^2}\right) \\ & - Q_1\exp\left(-\frac{D\left(\eta,\eta_e\right)^2}{\eta_1^2}-\frac{\left(\phi-\phi_n\right)^2}{\phi_1^2}\right) + Q_1\exp\left(-\frac{D\left(\eta,\eta_w\right)^2}{\eta_1^2}-\frac{\left(\phi-\phi_n\right)^2}{\phi_1^2}\right). \end{aligned} \quad \text{(A1)}$$

Here, $\phi$ is latitude and $\eta$ is longitude. The first term on the right-hand side represents the hemispherically and zonally symmetric component of ocean heat flux divergence away from the equator, as in ? (?) and ? (?). The second and third terms represent the zonally asymmetric components, similar to ? (?). The asymmetric component of the heat fluxes takes the form of a dipole in the northern hemisphere centered at $\phi_n = 15.0°$ latitude, with cooling and warming lobes centered at $\eta_w = 90.0°$ and $\eta_e = 180.0°$ longitude. $D$ is the distance between two longitudes $\eta_a$ and $\eta_b$, defined as $D\left(\eta_a,\eta_b\right) = \min\left(\left|\eta_a-\eta_b\right|, 360° - \left|\eta_a-\eta_b\right|\right)$. The parameters we chose are $\phi_0 = 16.0°$, $\eta_1 = 30.0°$, $\phi_1 = 6.0°$, $Q_0 = 50 \text{ W m}^{-2}$, and $Q_1 = 150 \text{ W m}^{-2}$.

## Appendix B Rayleigh damping

The Rayleigh damping coefficient ($\Gamma_z$) for horizontal velocities in the sponge layer is

$$\Gamma_z = \begin{cases} \Gamma_{max}\sin^2\left[\frac{\pi}{2}\left(1-\frac{z_{\text{top}}-z}{z_d}\right)\right] & z \geq z_{\text{top}} - z_d, \\ 0 & z < z_{\text{top}} - z_d, \end{cases} \quad \text{(B1)}$$

where $z_{\text{top}}$ is the depth of the LES domain (25.6 km) and $z_d$ is the depth of the sponge layer (8 km). $\Gamma_{max}$ is set to 0.01 $\text{s}^{-1}$. The depth of the sponge layer does not significantly affect the LES results.

## Appendix C Microphysics

Our microphysics scheme is similar to Arctic mixed-phase scheme described in ? (?) in that it is a one-moment microphysical parameterization with prognostic equations for precipitating water and ice. Our microphysics differs from that in ? (?) primarily in the liquid fraction function $\lambda(T)$, which determines the phase partitioning between cloud liquid and ice:

$$\lambda(T) = \begin{cases} 0 & T < T_i, \\ \frac{T-T_i}{T_f-T_i} & T_i < T \leq T_f, \\ 1 & T > T_f. \end{cases} \tag{C1}$$

Here, we use $T_i = 263.15$ K and $T_f = 273.15$ K. Further, our scheme replaces the cloud ice and snow auto-conversion rate closures used in ? (?) with those used in ? (?), which are more appropriate for use in simulations of subtropical and tropical convection.

## Acknowledgments

We gratefully acknowledge the generous support of Eric and Wendy Schmidt (by recommendation of Schmidt Futures), Mountain Philanthropies, EarthRise Alliance, Charles Trimble, the Paul G. Allen Family Foundation, and the National Science Foundation (grant 1835860). The simulations were performed on Caltech's High Performance Cluster, which is partially supported by a grant from the Gordon and Betty Moore Foundation. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. The GCM codes are available on the GitHub repository `https://github.com/szy21/fms_GCMForcing`. The LES codes are available on the GitHub repository `https://github.com/szy21/pycles_GCM/tree/shen2020`. Primary GCM and LES data that may be used to produce the plots are available at `https://data.caltech.edu/records/1337`.