Can We Use 1D Models to Predict 3D Physics?

Yi-Ling Hwong[1], Steven Sherwood[2], and David Fuchs[2]

[1]Climate Change Research Centre, University of New South Wales
[2]University of New South Wales

November 22, 2022

Abstract

Single-column models (SCMs) are often used to evaluate model physics and aid parameterization development. However, few studies have systematically compared the results obtained using 1D setups with those of their corresponding 3D models, and examined what factors potentially impact their comparability. This paper addresses these questions. We focus on the application of SCMs under idealized RCE conditions and use a multi-column model (MCM) setup as stepping stone for a 3D model. We find that convective organization in the MCM depends at least as much on the convection scheme used as on other mechanisms known to organize convection (e.g., radiative feedback). Moreover, convective organization emerges as a robust factor affecting SCM-MCM comparability, with more aggregated states in 3D associated with larger behavior deviations from the 1D counterpart. This is found across five convection schemes and applies to simulated mean states, linear responses to small tendency perturbations, and adjustments to doubled-CO2 forcing. Applying a "model-as-truth" approach, we find that even when convection is organized, behavior differences between pairs of schemes in the SCM are largely preserved in the MCM. This indicates that when model physics produces accurate behavior in a 1D setup, it will be more likely to do so in a 3D setup. We also demonstrate the practical value of linear responses by showing that they can accurately predict an SCM's tropospheric adjustment to doubled-CO2 forcing.

Can We Use 1D Models to Predict 3D Physics?

Y.-L. Hwong[1], S.C. Sherwood[1], D. Fuchs[1]

[1]Climate Change Research Centre, University of New South Wales, Sydney, Australia

Key Points:

- Convective organization in large-scale simulations depends on convection scheme, sometimes more so than on radiative feedback
- 1D and 3D behavior is very similar if convection is not organized in 3D
- A convection scheme's linear responses can be used to predict its tropospheric adjustments to doubled-CO_2 forcing

Corresponding author: Yi-Ling Hwong, `yiling.hwong@gmail.com`

Abstract

Single-column models (SCMs) are often used to evaluate model physics and aid parameterization development. However, few studies have systematically compared the results obtained using 1D setups with those of their corresponding 3D models, and examined what factors potentially impact their comparability. This paper addresses these questions. We focus on the application of SCMs under idealized RCE conditions and use a multi-column model (MCM) setup as stepping stone for a 3D model. We find that convective organization in the MCM depends at least as much on the convection scheme used as on other mechanisms known to organize convection (e.g., radiative feedback). Moreover, convective organization emerges as a robust factor affecting SCM-MCM comparability, with more aggregated states in 3D associated with larger behavior deviations from the 1D counterpart. This is found across five convection schemes and applies to simulated mean states, linear responses to small tendency perturbations, and adjustments to doubled-CO_2 forcing. Applying a "model-as-truth" approach, we find that even when convection is organized, behavior differences between pairs of schemes in the SCM are largely preserved in the MCM. This indicates that when model physics produces accurate behavior in a 1D setup, it will be more likely to do so in a 3D setup. We also demonstrate the practical value of linear responses by showing that they can accurately predict an SCM's tropospheric adjustment to doubled-CO_2 forcing.

Plain Language Summary

To study various climate processes, scientists often use 3D climate models, but these simulations use huge amounts of computing time and resources. One way to alleviate this problem is to use a single-column model, which is a 1D vertical column extracted from a 3D model. Although these 1D simulations are very efficient, we cannot always be sure that their results are comparable to those of their parent 3D model. In this study, we find that when clouds are randomly spread across the sky (when convection is disorganized), results of 1D and 3D simulations are very similar. However, when clouds are clustered into clumps (when convection is organized), we cannot always trust the results of 1D simulations as they tend to be different from those of 3D simulations. Nevertheless, we find that when two models show very different behavior in their 1D setup, they will tend to also behave differently in their 3D setup. This tells us that 1D simulations can still be useful. We also discover that the way a model responds when it is lightly tickled (perturbed) can be used to predict its responses to a situation where the amount of carbon dioxide in the atmosphere is suddenly doubled.

1 Introduction

One-dimensional, single-column models (SCMs) of the atmosphere are often used as a research tool to understand climate and climate change, as well as the behavior of model physics to inform parameterization development. Experiments conducted in SCMs are commonly used as proxies to evaluate how parameterizations would perform at individual grid points in a 3D model (Randall et al., 1996). Other studies take an SCM to be a representative 1D model of the whole atmosphere, starting from the classic study of Manabe and Wetherald (1967). However, few studies have attempted to directly evaluate to what extent understanding obtained with an SCM carries over to a 3D setup and identify what factors potentially enter that may alter the conclusions. This paper addresses these questions. Specifically, we focus on the applicability of SCM tests under idealized radiative-convective equilibrium (RCE) conditions to the behavior of a 3D model also run in regional RCE.

Commonly, an SCM is a single vertical column from a General Circulation Model (GCM), using the same physical parameterizations of the parent GCM to represent unresolved subgrid-scale processes. Model dynamics are prescribed as boundary forcings

(Randall & Cripe, 1999). Since there is only one vertical column it is computationally very cheap to run an SCM. This offers great benefits to modelers, for example when exploration of a large parameter space is unfeasible in a full GCM (M. Zhang et al., 2016). Moreoever, the SCM allows model physics to be isolated and tested in controlled conditions. Hence, SCMs are often used as a tool to facilitate parameterization development. Outputs from an SCM can be compared with observational data or results from cloud-resolving models (CRMs) to evaluate the performance of a physics scheme. Early studies using SCMs in this way include Iacobellis and Somerville (1991) and Lee et al. (1997). SCMs have been used to study cloud and convective processes—a major source of uncertainty in climate predictions (Boucher et al., 2013)—including the representation of boundary layer clouds (Blossey et al., 2013; Dal Gesso & Neggers, 2018; M. Zhang et al., 2013; P. Zhu et al., 2005), the diurnal cycle of shallow (Lenderink et al., 2004) and deep precipitating clouds (Guichard et al., 2004), various convective regimes (Petch et al., 2007), the evolution of tropical convection (Petch et al., 2014), tropical squall lines (Bechtold et al., 2000), the representation of shallow convection (Bogenschutz et al., 2012) and cloud microphysics (Gettelman et al., 2008). As research tools, SCMs have, for example, been employed for parameter sensitivity analysis of cloud properties (Guo et al., 2014) and to test the sensitivity of subgrid-scale physical processes to model vertical resolution (Lane et al., 2000).

The lack of feedback between model physics and large-scale circulation is a key limitation of SCMs. For example, if forced only by estimated large-scale tendencies they can drift away from a realistic state in their temperature and humidity fields (M. Zhang et al., 2016). This can be overcome by, for example, driving the SCM with large-scale forcings derived from coarse-graining high-resolution model outputs (Christensen et al., 2018), or nudging to a target state (Dal Gesso & Neggers, 2018; Neggers et al., 2017; Randall & Cripe, 1999). An alternative approach utilizes SCMs in more idealized setups, where knowledge obtained in simpler contexts can hopefully inform more complex systems (Maher et al., 2019). Examples include parameterizing large-scale circulations using the weak temperature gradient (WTG; Sobel & Bretherton, 2000; H. Zhu & Sobel, 2012) or quasi-geostrophic (QG) approximations (Nie & Sobel, 2016).

Another idealization extensively applied in both SCMs and atmospheric general circulation models (AGCMs) is to run model simulations in an RCE configuration (see Wing et al., 2018, for a comprehensive review). RCE is a statistical equilibrium state where net radiative cooling is balanced by net convective heating. There is no horizontal energy transport, and background vertical motion is assumed to be zero ($w = 0$), effectively eliminating large-scale dynamics. In the hierarchy of modeling approaches, a single-column RCE is the simplest representation of the climate system (Jeevanjee et al., 2017; Maher et al., 2019) and is often used to represent tropical regions as a whole (Ghan et al., 2000; Pakula & Stephens, 2009; Wing et al., 2018). RCE has been extensively applied to study various aspects of the atmosphere, including to understand climate change, where it was, for example, used for early estimates of equilibrium climate sensitivity (ECS; Manabe & Wetherald, 1967; Ramanathan & Coakley, 1978).

It has been argued that the linearized behavior in a single column near RCE could characterize physical processes sufficiently to replicate (i.e., parameterize) their impact on slowly varying tropical circulations at larger scales, based on the linear response function (LRF) framework (Kuang, 2010, 2018). Subsequent studies of SCMs taken from climate models have reported substantial discrepancies among their linearized behavior and departures of their linear responses from explicit numerical simulations (Herman & Kuang, 2013; Hwong et al., 2021, henceforth H21), suggesting that such near-RCE behavior could be a valuable test of model physics. In general, however, RCE almost never occurs locally because of the presence of large-scale circulations (w is not zero, or at least does not remain zero for a time sufficient for the local column to attain equilibrium). Therefore, one reason to question the relevance and validity of SCMs in RCE is that there is little to no ascent and descent or large-scale condensation in this 1D configuration but a fair amount of them in reality. Even if a larger-scale domain is in regional RCE, the individual columns within this domain will usually experience large departures from local RCE. This begs the question: if we know an SCM behaves correctly near RCE according to some reference standard (based on e.g., a large-eddy simulation [LES] calculation, or observations collected at a field site), how helpful is this? If the most important errors in a 3D simulation only occur far away from local RCE, then SCM RCE tests could—although useful in principle—be unhelpful in practice.

One interesting phenomenon emerging from numerical simulations of RCE in 3D models, and relevant for our objectives, is the organization of convection. Under certain conditions, numerically simulated convection sometimes spontaneously organizes into distinct wet and dry regions. This can happen even under homogeneous boundary conditions and forcing, and has been observed in both cloud-system-resolving simulations (e.g., Holloway & Woolnough, 2016; Muller & Bony, 2015; Wing & Emanuel, 2014) and those in which convection is parameterized (e.g., Arnold & Randall, 2015; Becker & Stevens, 2014; Coppin & Bony, 2015; Hohenegger & Stevens, 2016; Reed et al., 2015). Coppin and Bony (2015) describe the emergence of convective organization in equilibrium conditions as an instability of the RCE state, and its potential causes have been investigated in various studies (Beucler & Cronin, 2016; Bretherton et al., 2005; Craig & Mack, 2013; Emanuel et al., 2014). In an RCE model intercomparison project (RCEMIP), Wing et al. (2020) showed that all 3D models in their study exhibit organization behavior, albeit to different degrees. Convective organization can also be observed in the real world, for example as squall lines (Bryan, 2005), the Madden-Julian Oscillation (MJO; Madden & Julian, 1994), and all the way up to the planetary scale where rain is organized into structures such as the Inter-Tropical Convergence Zone (ITCZ) and midlatitude rain bands. These planetary-scale organizations make the global humidity significantly less sensitive to model physics than it would be in a 1D or horizontally homogeneous situation (Sherwood & Meyer, 2006), again suggesting caution in the application of SCMs to represent heterogeneous domains. Given its association with interactions between clouds, moisture and large-scale circulation, there is reason to believe that convective organization might be an important factor to consider when comparing the results of 1D and 3D models.

The overarching goal of this paper is to assess the utility of 1D simulations for a 3D world. We used a multi-column model (MCM) setup as a stepping stone for a 3D setup and evaluate the similarity between the SCM vs. MCM setups by comparing their mean states, linear responses to small tendency perturbations (the LRF method explored recently by H21), and adjustment responses to doubled-CO_2 forcing. All experiments are conducted on five widely-used convection schemes. The specific objectives are:

- To assess how informative SCM model-physics tests are about more realistic 3D scenarios under RCE conditions, and the potential role of convective organization in modulating the comparability of 1D vs. 3D setups;
- To determine if the LRF method can be used to predict doubled-CO_2 responses, and if so, what are the implications for using it in an SCM vs. an MCM.

This paper is organized as follows: the models and simulation details are described in Section 2; the RCE mean state and organization patterns of the five convection schemes under different experimental configurations are presented in Section 3; the SCM and MCM responses to imposed tendency perturbations are compared in Section 4, and their adjustment responses to doubled-CO_2 forcing are compared in Section 5; the conclusions from these experiments are presented in Section 6.

2 Methods

2.1 Models and RCE Simulation

We largely followed the procedures of H21 for our RCE simulations. All simulations were performed with the Weather Research and Forecasting (WRF) model (version 4.0.2). The single-column model version of WRF was used for the 1D simulations (Hacker & Angevine, 2013). For the 3D simulations, the model was run in a multi-column model (MCM) setup with a square domain, with 20 grid points in both x and y directions and a horizontal resolution of 100 km. There are 74 vertical levels, with model top at around 6 hPa. Configurations were kept consistent between corresponding SCM and MCM runs, except for elements only applying to a 3D setup (sea surface temperature [SST] hot spot and water vapor homogenization, described below).

The simulations were initialized with the same sounding at every grid point, using the initial profiles of Wing et al. (2018), which are based on a moist tropical sounding of Dunion (2011). There is no diurnal cycle and we simulate a non-rotating RCE with the coriolis parameter set to zero. We used an SST of 28°C for the experiments with uniform SST. Zonal and meridional winds were initialized with zero values and relaxed to zero throughout the simulations with a relaxation time constant of 3 h (except for experiments with imposed vertical wind shear, described below). Convection was kick-started by applying random low-level perturbations to the temperature field, although this does not appear to affect the RCE state. Surface fluxes were computed using a bulk aerodynamic formula with a fixed value of 0.001 for the heat and moisture exchange coefficients, and a constant near surface wind speed of 4.8 m s^{-1} to remove any wind-induced surface heat exchange effects (WISHE).

In terms of physical parameterization, we tested five convection schemes: Kain-Fritsch (KFETA; Kain, 2004), New-Tiedtke (NTIEDTKE; C. Zhang & Wang, 2017), New-Simplied-Arakawa-Schubert (NSAS; Han & Pan, 2011), Betts-Miller-Janjic (BMJ; Betts, 1986; Betts & Miller, 1986; Janjić, 1994), and Zhang-McFarlane (CAMZM; G. Zhang & McFarlane, 1995). We refer to H21 for a description of their main features. The Zhang-McFarlane scheme is additionally paired with the University of Washington shallow convection scheme (Park & Bretherton, 2009). All convection schemes were paired with the same planetary-boundary-layer (PBL) and microphysics schemes: the YSU PBL scheme (Hong et al., 2006) and the WSM6 microphysics scheme (Hong & Lim, 2006). We did not examine sensitivity to these schemes since H21 found them to play only a minor role in the RCE responses examined. For the simulations with interactive radiation, the RRTMG longwave and shortwave schemes were used (Iacono et al., 2008), with a solar constant of 544 $\mathrm{W\,m^{-2}}$ and a fixed zenith angle of 37°, yielding a solar insolation of around 436 $\mathrm{W\,m^{-2}}$ to match equatorial conditions. For the simulations with fixed radiation, we prescribed a radiative cooling profile of -1.5 K day^{-1} from the surface to around 200 hPa and then linearly decreasing to zero at around 100 hPa and kept at zero above that (following Herman & Kuang, 2013). The relaxation inverse time constant is zero from surface to around 160 hPa and then increases linearly to 0.5 day^{-1} at and above 100 hPa.

2.2 Convective Organization Configurations

Our initial MCM experiments revealed that convective organization depends on the choice of convection scheme. For example, we found that certain schemes produced organized convection (and others did not) even when radiative feedback was disabled. To remove potential confounding factors arising from varying degrees of organization across the schemes, we wanted the schemes to display relatively similar organization behavior (i.e., all organized or all disorganized) for our analyses. To achieve this, we experimented with three mechanisms that have been shown in previous studies to have an impact on convective organization: an imposed SST hot spot, vertical wind shear, and water vapor homogenization. Each mechanism was paired with either idealized (described in the previous section) or interactive radiation. In total, eight experimental configurations were tested (Table 1). Note that in doing so our motivation was not to provide physical explanations for the manifestation of convective organization of the schemes under these configurations, but to establish a common baseline so that comparisons between the schemes can be made on the basis of relatively similar degrees of organization.

Table 1. Experimental configurations for convective organization.

Configuration	**Simulation details**
`idealrad_ctrl`	idealized radiation (same radiative cooling profile at every grid point), control simulation
`idealrad_hotspot`	`idealrad_ctrl` + SST hot spot at center of domain
`idealrad_windshear`	`idealrad_ctrl` + vertical wind shear
`idealrad_qvhomo`	`idealrad_ctrl` + water vapor homogenization above 2 km
`intrad_ctrl`	interactive radiation (RRTMG LW and SW schemes), control simulation
`intrad_hotspot`	`intrad_ctrl` + SST hot spot at center of domain
`intrad_windshear`	`intrad_ctrl` + vertical wind shear
`intrad_qvhomo`	`intrad_ctrl` + water vapor homogenization above 2 km

The first mechanism employed, which was to promote aggregation, was to conduct MCM experiments with an SST hot spot in the form of a circular warm pool at the center of the domain. Studies have shown that SST gradients have the effect of organizing convection (e.g., Liu & Moncrieff, 2008; Shamekh et al., 2020; Tompkins, 2001). The hot spot has a gaussian surface with a half-width of around 400 km and covers around 10% of the domain area. The peak SST anomaly is 2 K, with the edge of the hot spot having approximately the same SST as in the control simulations (28°C). To homogenize convection across all schemes, we tested two mechanisms. First we followed the "strong shears" procedure of Tompkins (2001), which the authors found to disrupt the water vapor feedback responsible for convective organization. The zonal winds were relaxed to a target profile with a relaxation time constant of 1 h. The target winds increase from 8 m s^{-1} at the lowest level to 12 m s^{-1} at around 1 km, and then reduce linearly to -10 m s^{-1} at around 12 km, increasing linearly thereafter to 0 at around 14.5 km. Second, in the water vapor homogenization experiments, we followed the procedures of Grabowski and Moncrieff (2004), where we removed moisture variations in the free troposphere (above 2 km) by applying a relaxation term to the water vapor equation as follows

$$\left(\frac{\partial q_v}{\partial t}\right)_{relax} = -\frac{q_v - \overline{q_v}}{\tau} \tag{1}$$

where an overbar denotes the domain mean value at a given level and τ the relaxation time constant. We used a value of 1 day for τ. All simulations were run for 1,000 days for the SCM and 100 days for the MCM simulations. These simulations are henceforth referred to as the `PreRCE` runs.

2.3 Forcing Experiments

We ran two sets of forcing experiments to test the comparability between the SCM and MCM setups. In the first set of experiments, we probed the linear responses of the schemes in both setups using the LRF framework. According to this framework, the responses of a cumulus ensemble to small perturbations to its large-scale environment can be considered to be approximately linear even though convection involves many nonlinear processes (Kuang, 2010). This linear assumption drastically simplifies the representation of convection, as the behavior of a cumulus ensemble around a reference state can be approximated as

$$\frac{d\mathbf{x}}{dt} = \mathbf{M}\mathbf{x} \tag{2}$$

where $\mathbf{x}$ is the anomalous state vector of temperature (**T'**) and moisture (**q'**) around a reference state (e.g., RCE), $d\mathbf{x}/dt$ is the corresponding anomalous tendency perturbations, and M is the LRF matrix. We followed the procedures of H21, which we briefly summarize here. We initialized the control runs of the eight configurations described in the previous section (Table 1) from their respective `PreRCE` states, maintaining the same experimental configurations to get the respective organization patterns. These simulations are referred to henceforth as the `CTRL` runs. The perturbation runs were initialized and run the same way, but additionally with small perturbations applied to the temperature ($d\mathbf{T}/dt$) and moisture ($d\mathbf{q}/dt$) tendencies in separate runs and at every timestep, until the models reached a new steady state. These simulations are referred to henceforth as the `PerturbLRF` runs. The perturbation profile takes the form of the sum of a delta and gaussian functions, following Equation 4 in Herman and Kuang (2013). For brevity a profile that peaks at pressure level p is referred to simply as "perturbation at p". For this study, we selected a perturbation level of 850 hPa. The perturbation amplitudes are 0.5 K day^{-1} and 0.2 g kg^{-1} day^{-1} for temperature and moisture tendency perturbations, respectively. Positive and negative perturbations were applied in separate runs. The steady state responses (**T'** and **q'**) are the differences of the time-averaged temperature and moisture profiles between the `PerturbLRF` and `CTRL` runs. We averaged the responses of the positive and negative perturbation runs to obtain the final **T'** and **q'** profiles.

In the second set of forcing experiments, we doubled the atmospheric CO_2 concentration. This set of experiments was conducted with interactive radiation and fixed SST (28°C). We initialized the models from the `PreRCE` states of `intrad_ctrl` but doubled the CO_2 concentration from the default 379 ppm to 758 ppm and ran the experiments until a new equilibrium was reached. These simulations are referred to as the `PerturbCO2` runs. The adjustment response of a variable resulting from this doubling of CO_2 is the difference between its time-averaged profiles between the `PerturbCO2` and `CTRL` runs of `intrad_ctrl`.

All experiments described above were run in both SCM and MCM setups and for the five convection schemes. The simulation periods of the `CTRL`, `PerturbLRF` and `PerturbCO2` experiments were 1,000 days for the SCM and 100 days for the MCM runs. Unless specified otherwise, the equilibrium quantities presented in this paper are the averages over the final 700 days of the `CTRL` runs (and, where applicable, their corresponding `PerturbLRF` or `PerturbCO2` runs) for the SCM and final 20 days (and averaged over the domain) for the MCM setup.

Additionally, to verify the usefulness of the LRF method in predicting model responses to other types of forcings, we tested if the M^{-1} matrices can be used to predict the temperature and humidity responses to a doubling of CO_2 in the atmosphere. We constructed the M^{-1} matrices of the five convection schemes, which show their steady state responses per unit perturbation (H21). To construct M^{-1}, we applied small tendency perturbations at all model levels, successively and in separate runs. Rearranging Equation 2, we get $\mathbf{x} = \mathbf{M}^{-1}\frac{d\mathbf{x}}{dt}$. In principle, we should be able to predict the temperature and moisture responses ($\mathbf{x}$) to doubled CO_2 by multiplying the M^{-1} matrices by the radiative forcing ($d\mathbf{x}/dt$) resulting from the doubling of CO_2. Given the resource intensity of these matrix runs, we conducted this part of the study only in the SCM setup.

3 RCE Mean State and Convective Organization

3.1 Convective Organization Across Schemes and Configurations

We first investigate the organization patterns of the five convection schemes at RCE for the eight configurations listed in Table 1. The schemes reach RCE in the `PreRCE` runs latest by around day 300 for the SCM and around day 60 for the MCM runs. In RCE, the radiative cooling rates largely balance the total surface heat fluxes and precipitation rates roughly equal the latent heat fluxes. In the MCM simulations, large-scale circulations develop and there are positive and negative vertical winds in the individual grids—indicating local RCE instability—but there are no net vertical motions over the whole domain ($\overline{w} = 0$). Snapshots of the daily accumulated convective rain on the final day of the `CTRL` runs for simulations with idealized radiation are shown in Figure 1 and interactive radiation in Figure 2.

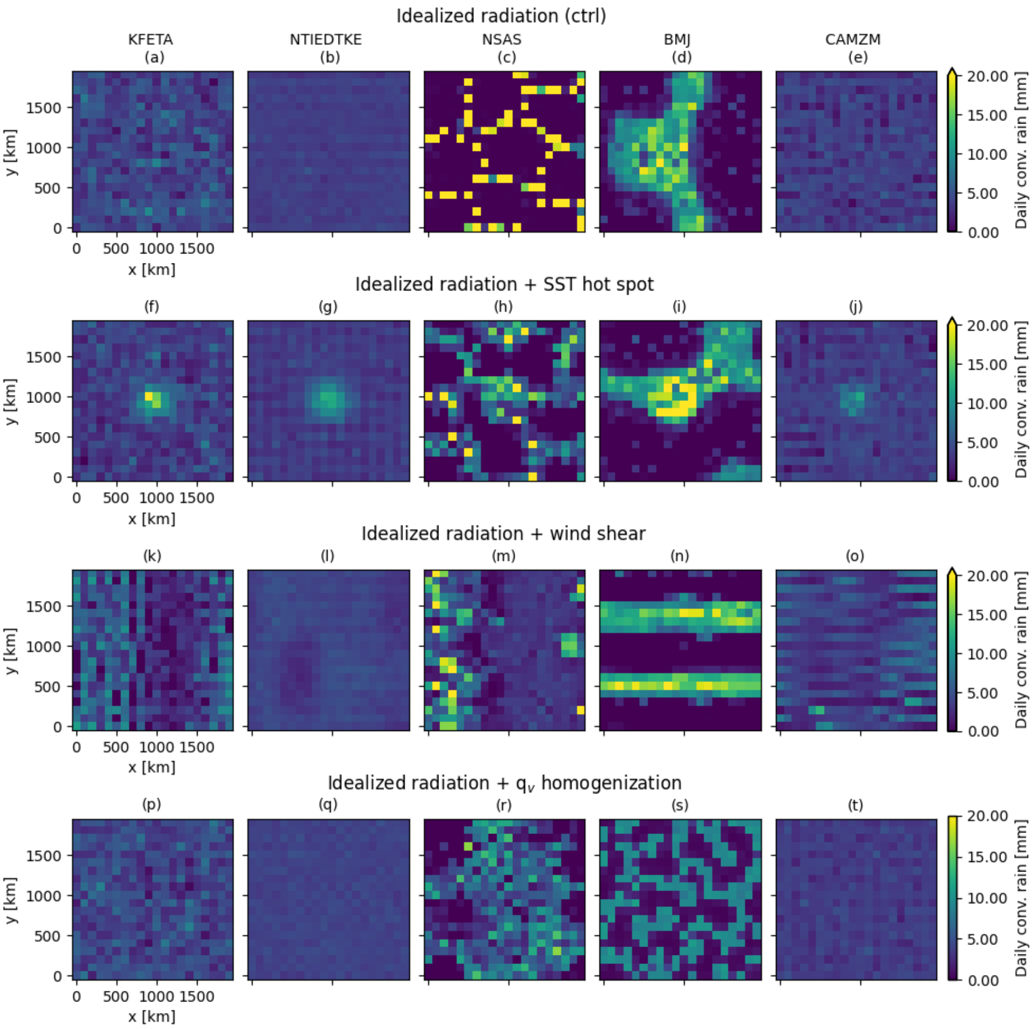

Figure 1. Daily accumulated convective rain on day 100 of the MCM `CTRL` runs with idealized radiation of (left–right) KFETA, NTIEDTKE, NSAS, BMJ and CAMZM for the (a–e) control simulations, and simulations with imposed (f–j) SST hot spot, (k–o) vertical wind shear, (p–t) water vapor homogenization.

Generally, simulations that organize tend to have clearer separations between the dry (subsidence) and moist (ascending motion) cells. The clusters of rainfall regions usually correspond to slightly higher latent heat fluxes. In simulations with more homogeneous convection, precipitation is more evenly spread out with no clear distinction between dry and wet cells, indicating that convection does not organize despite the presence of resolved circulations. As expected, simulations with idealized radiation tend to be more disorganized

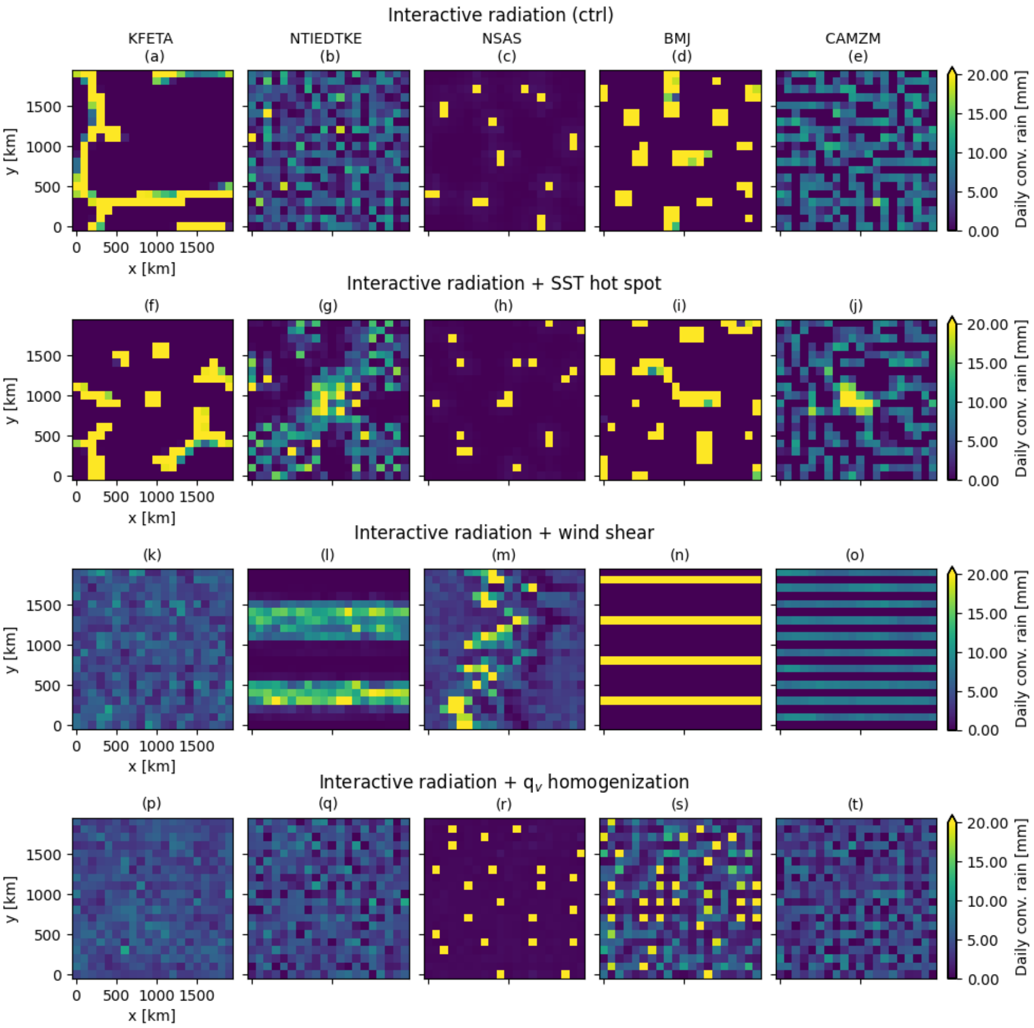

Figure 2. As in Figure 1, but for simulations with interactive radiation.

compared to the ones with interactive radiation. However, there are a few notable exceptions. NSAS and BMJ produce strong organization even in the case of prescribed uniform radiation (c and d in Figure 1). This is an interesting observation, as numerous studies have shown that radiative (especially longwave) feedback is a key factor in convective organization and homogenizing radiation typically leads to disaggregation (e.g., Arnold & Randall, 2015; Coppin & Bony, 2015; Holloway & Woolnough, 2016; Muller & Held, 2012; Wing & Emanuel, 2014). A few studies have shown that convection can self-aggregate in the absence of interactive radiation if evaporation of precipitation is artificially suppressed (hence weakening cold pool feedback) (Holloway & Woolnough, 2016; Muller & Bony, 2015). This is a type of "moisture memory aggregation" as opposed to the more commonly known "radiative aggregation" (Muller & Bony, 2015). So it is possible that some form of moisture feedback is responsible for the organization of NSAS and BMJ under uniform radiative forcing. On the other end of the spectrum, NTIEDTKE and CAMZM appear to be fairly disorganized even in the presence of interactive radiation (b and e in Figure 2). We suspect this could be due to the fairly small domain size that we used (2000 × 2000 km). Larger domain sizes have been found to be more conducive to aggregation (Jeevanjee & Romps, 2013; Muller &

Held, 2012). In any case, these overall results suggest that the convection scheme can exert a stronger influence than interactive radiation in determining convective organization.

Imposing an SST gradient tends to organize convection toward the location of the hot spot, although the effect again varies across the convection schemes, as well as between idealized and interactive radiation. The organization effect of SST hot spot is more prominent when interactive radiation is used. In the case of idealized radiation, convection is still largely disorganized despite the stronger convection observed at the location of the hot spot for KFETA, NTIEDTKE and CAMZM (f, g and j in Figure 1). The schemes also respond very differently to imposed wind shear forcing: this homogenizes convection for NSAS (m in Figures 1 and 2) and KFETA (Figure 2k), but increases organization in the case of BMJ (n in Figures 1 and 2) and NTIEDTKE (Figure 2l). This ambiguity has also been shown in other studies, which have found vertical wind shear to make organization either more (LeMone et al., 1998; Muller, 2013; Rotunno et al., 1988) or less likely (Abbot, 2014; Bretherton et al., 2005; Held et al., 1993; Tompkins, 2000). When organized by wind shear, the schemes tend to aggregate into band-like structures, reminiscent of squall lines observed in nature as well as CRM simulations (Muller, 2013). Lastly, homogenizing free tropospheric moisture is found to have a dramatic effect on convective organization, especially in the case of idealized radiation (p–t in Figure 1). By effectively disabling the feedback between convection and free-tropospheric moisture—which is critical for convective organization—this procedure causes convection to become disorganized across all schemes in the case of idealized radiation. In the case of interactive radiation, however, NSAS and, to a lesser degree, BMJ, still display organization behavior (r and s in Figure 2). This could be because when these two schemes were used the homogenization of water vapor was insufficient to overcome the fluctuations created by interaction between convection and the large-scale circulations.

3.2 Quantifying the Degree of Organization

To quantify the degree of organization in the MCM simulations, we use the spatial variance of precipitable water scaled by its average value, averaged in time over the last 20 days of the `CTRL` runs, following one of the metrics used by Wing et al. (2020). We refer to this metric as org_{pw}. We also tested another metric, the subsidence fraction (Becker et al., 2017; Coppin & Bony, 2015; Wing et al., 2020), which is calculated as the fraction of domain area where the vertically integrated mass-weighted vertical wind is directed downward. There is a high correlation between the two metrics ($r = 0.78$). For simplicity we only report the results using org_{pw}.

The time series (Figure 3) and mean values (Table 2) of org_{pw} largely agree with visual impressions of organization in Figures 1 and 2. For example, the NSAS and BMJ runs, which appeared more organized in the `idealrad_ctrl` simulations, have comparatively higher org_{pw} values (Figure 3a). All simulations were initialized with the same disorganized state; some remain disaggregated, while others develop aggregation after 10–20 days (or later) and usually stabilize thereafter. NSAS and BMJ runs tend to organize sooner than those with the other schemes. NSAS runs also display oscillations in org_{pw}. All model runs stabilize by around day 60, except for BMJ in the `intrad_windshear` configuration, which appears to stabilize only after day 90 (Figure 3g). Overall, convective organization does not appear to be binary but varies in a continuum depending on convection scheme and experimental configuration. This has also been found in previous studies (Holloway & Woolnough, 2016; Wing et al., 2020) and is probably due to the myriad physical mechanisms and feedbacks at play between the schemes and configurations.

To recall, our aim was to find a configuration where all five schemes display relatively similar degrees of organization or disorganization in order to remove potential confounding factors in our analyses. Judging by their equilibrium org_{pw} values, `intrad_hotspot` appears to lead to organization across all five schemes, while `idealrad_qvhomo` consistently leads to disorganization (d and f in Figure 3). We henceforth refer to these two configurations, respectively, as the `all_org` and `all_disorg` simulations.

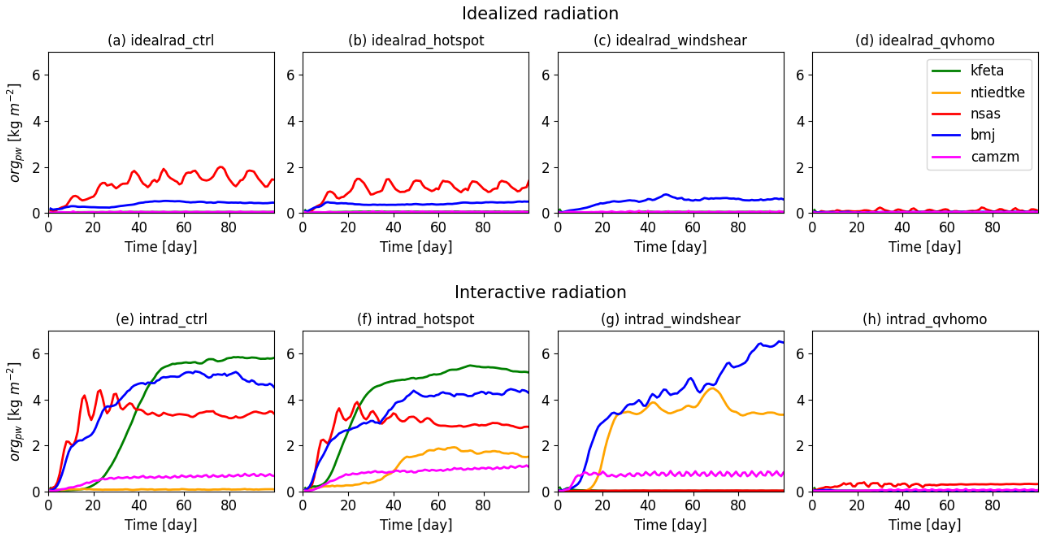

Figure 3. Time series of the degree of organization (org_{pw}) of the five convection schemes in the eight experimental configurations for the `PreRCE` runs (100 days). Top row (a–d) shows results with fixed radiation and four different configurations; bottom row (e–h) shows the same but with interactive radiation.

Table 2. Mean equilibrium org_{pw} values for the convection schemes and experimental configurations. The bolded rows are the `all_disorg` and `all_org` simulations. Units are $kg\ m^{-2}$.

	KFETA	**NTIEDTKE**	**NSAS**	**BMJ**	**CAMZM**
`idealrad_ctrl`	0.042	0.01	1.688	0.445	0.044
`idealrad_hotspot`	0.059	0.017	1.353	0.588	0.05
`idealrad_windshear`	0.047	0.011	0.04	0.564	0.049
`idealrad_qvhomo`	**0.028**	**0.011**	**0.108**	**0.054**	**0.036**
`intrad_ctrl`	5.052	0.083	3.627	4.45	0.819
`intrad_hotspot`	**4.144**	**1.176**	**3.23**	**4.262**	**1.254**
`intrad_windshear`	0.021	3.758	0.04	6.398	0.741
`intrad_qvhomo`	0.03	0.015	0.327	0.029	0.064

3.3 Mean State Temperature and Humidity Profiles

Convective organization is known to significantly affect the atmospheric mean state (Wing & Cronin, 2016; Wing et al., 2017). Figure 4 shows the mean RCE temperature and relative humidity (RH) profiles of the MCM simulations averaged over the final 20 days of the `CTRL` runs. For temperature, we show the saturation equivalent potential temperature (θ_{es}) as this is more informative and shows the differences between the profiles better. For a given pressure there exists a unique and monotonic relationship between absolute temperature T and θ_{es}.

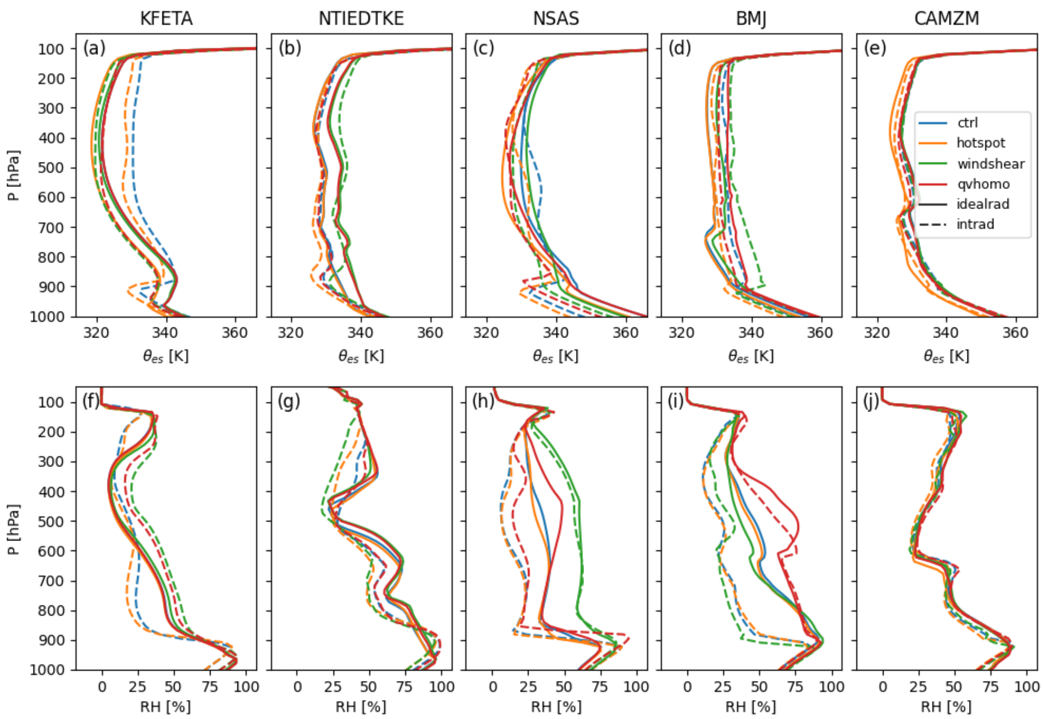

Figure 4. (a–e) Saturation equivalent potential temperature (θ_{es}) and (f–j) relative humidity (RH) profiles at RCE of the five convection schemes for the eight experimental configurations in the MCM setup. Profiles of simulations with idealized radiation are shown in solid curves and interactive radiation in dashed curves.

More organized mean states are generally associated with a warmer and drier free troposphere, as expected (Bretherton et al., 2005; Muller & Held, 2012; Wing & Cronin, 2016). For example, simulations with interactive radiation, which are generally more organized than their corresponding simulations with the same scheme but prescribed radiation for the control and SST hot spot simulations (`idealrad_ctrl` vs. `intrad_ctrl`, `idealrad_hotspot` vs. `intrad_hotspot`), correspondingly display higher θ_{es} and lower RH values in the free troposphere (the blue and orange dashed curves are warmer and drier than the corresponding solid curves in the individual panels of Figure 4). The warmer mean state when convection is organized is consistent with the fact that the increase in boundary-layer moisture in the convecting regions shifts the moist adiabat warmer there, and these warmer temperatures are then propagated to the entire domain through gravity waves (Bretherton et al., 2005; Muller & Held, 2012). The drier state caused by aggregation is brought on by the fact that the subsidence regions are much drier than anywhere in the runs with disorganized convection, reducing the mean (Wing & Cronin, 2016; Wing et al., 2017). However, we note that a drier mean free troposphere is not always warmer in our simulations. The `intrad_qvhomo` simulation for NSAS, for example, displays lower RH but a slightly cooler free troposphere compared to its corresponding `idealrad_qvhomo` mean state (red solid and dashed curves of c and h in Figure 4). We further note that simulations with more organization also often exhibit strong temperature inversions near the cloud base level, which often coincide with kinks in the RH profiles and most frequently in simulations with interactive radiation (e.g., the orange and blue dashed curves for KFETA and NSAS; a, c, f and h in Figure 4). These inversions are probably caused by the strong subsidence present in these highly organized simulations.

For the SCM simulations, their mean-state temperature and humidity profiles are very similar to those of their MCM counterparts in the cases where the MCM simulations are disorganized. The SCM-MCM mean-state discrepancy appears to widen with increasing organization in the MCM simulations. To illustrate this, we show the SCM-MCM profile differences for the `all_org` and `all_disorg` runs in Figure 5. The deviations between the SCM and MCM mean state profiles are significantly more pronounced for the `all_org` runs (all schemes are organized in MCM). Nevertheless, we note that the mean-state differences *between* the schemes are still on average larger than the SCM-MCM deviations for the `all_org` case. The MCM simulations are consistently drier in the free troposphere than their corresponding SCMs, which is expected due to the general drying caused by aggregation in the MCM. The temperature differences are also larger for the `all_org` case, albeit with inconsistent signs in the free troposphere. The MCM simulations are cooler at the near surface levels, but higher up they can be warmer (KFETA and NSAS) or cooler (NTIEDTKE, BMJ and CAMZM) than the SCMs. For the `all_disorg` simulations, the SCM and MCM simulations display almost identical profiles, except for small differences in the upper tropospheric humidity for NSAS and BMJ. Overall, we can conclude that the organization propensity of a scheme under specific experimental configuration can significantly impact the comparability of its SCM and MCM mean state, with increasing organization associated with growing discrepancy of equilibrium temperature and humidity profiles between its 1D and 3D setups.

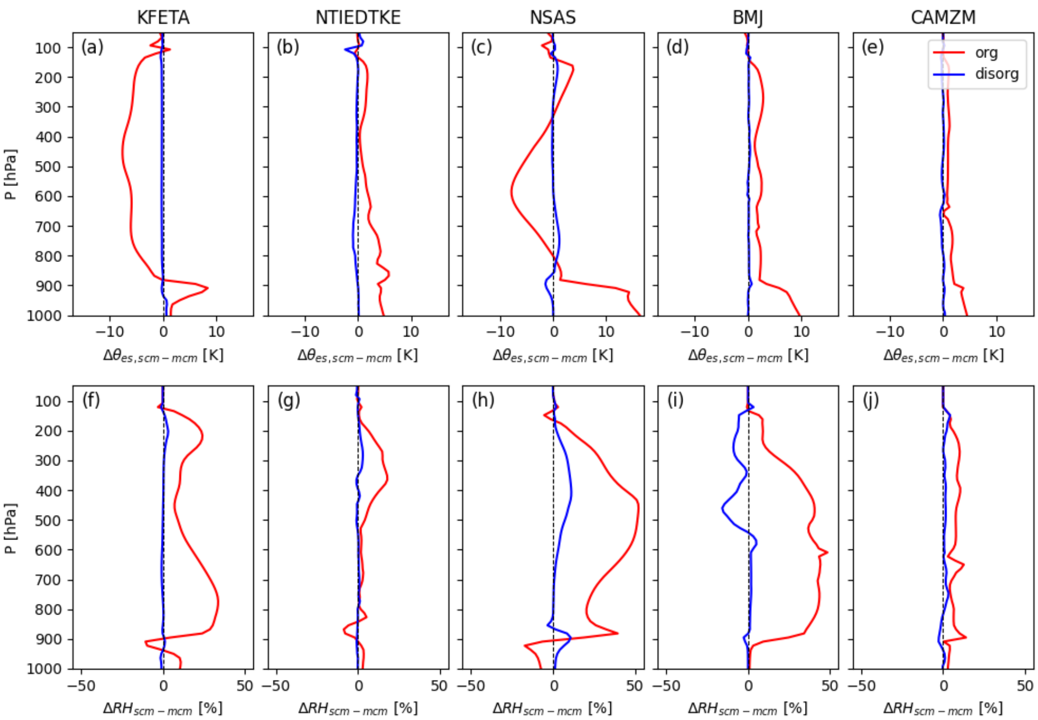

Figure 5. Difference of the (a–e) θ_{es} and (f–j) RH profiles between the corresponding SCM and MCM simulations for the five convection schemes. Red curves show the SCM-MCM differences of the `all_org` and blue curves the `all_disorg` simulations.

4 Response to Tendency Perturbations

4.1 Comparison of SCM-MCM Perturbation Responses

In this section we present the results from the tendency perturbation experiments. Recall that these are the profile responses to small, steady temperature or moisture tendency perturbations as per the linear response function framework of Kuang (2010). For a detailed description of the responses of the CRM used in Kuang (2010) and the five convection schemes in this study in an SCM setup we refer readers to H21. Here, we only note that many of the general trends of the CRM responses are observable in the SCM responses, though there is disagreement on some of them among the SCMs, and the SCM responses are consistently less smooth than the CRM's, often exhibiting sharp twists and kinks, especially around the cloud base and freezing levels (H21).

We again find that the responses are very similar between the SCM and MCM when convection is organized in the MCM, but diverge with increasing degree of MCM organization. This is true for all four perturbation quadrants: T response to dT/dt (T_DT), q response to dT/dt (Q_DT), T response to dq/dt (T_DQ), and q response to dq/dt perturbations (Q_DQ). To illustrate this we show the results for the `all_org` and `all_disorg` simulations (Figures 6 and 7 for dT/dt and dq/dt perturbation, respectively). For the `all_disorg` simulations (blue curves), the profiles are quite similar between the SCM and MCM setups, at times almost identical. By contrast, the SCM and MCM response profiles are often vastly different in both shape and magnitude for the `all_org` simulations (red curves). This is most prominent for the schemes with the highest org_{pw} values (KFETA and BMJ) while the schemes with relatively low org_{pw} values (NTIEDTKE and CAMZM) differ by less. Additionally, the MCM responses often display massive kinks around the cloud base level that are significantly sharper than those of their SCM counterparts (dashed red curves in Figures 6 and 7). These MCM kinks almost always coincide with strong inversions in the mean state temperature (described in previous section), which are also less pronounced in the SCM. This shows that the RCE mean state can affect the linear responses, as postulated by H21.

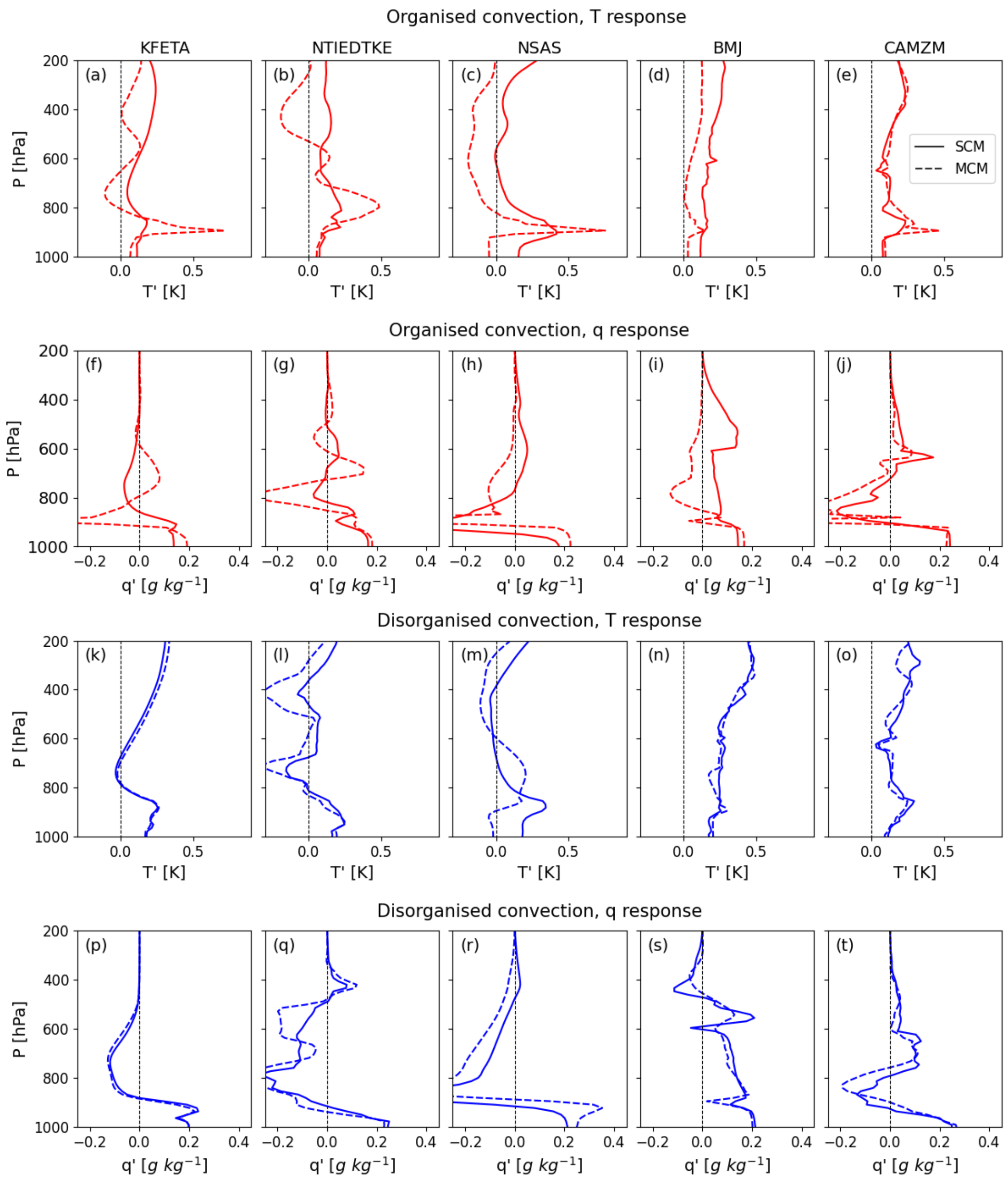

Figure 6. Comparison of the profiles of (a–e, k–o) temperature and (f–j, p–t) moisture responses to temperature tendency perturbations for the (red) `all_org` and (blue) `all_disorg` simulations. Solid curves are SCM and dashed curves MCM responses.

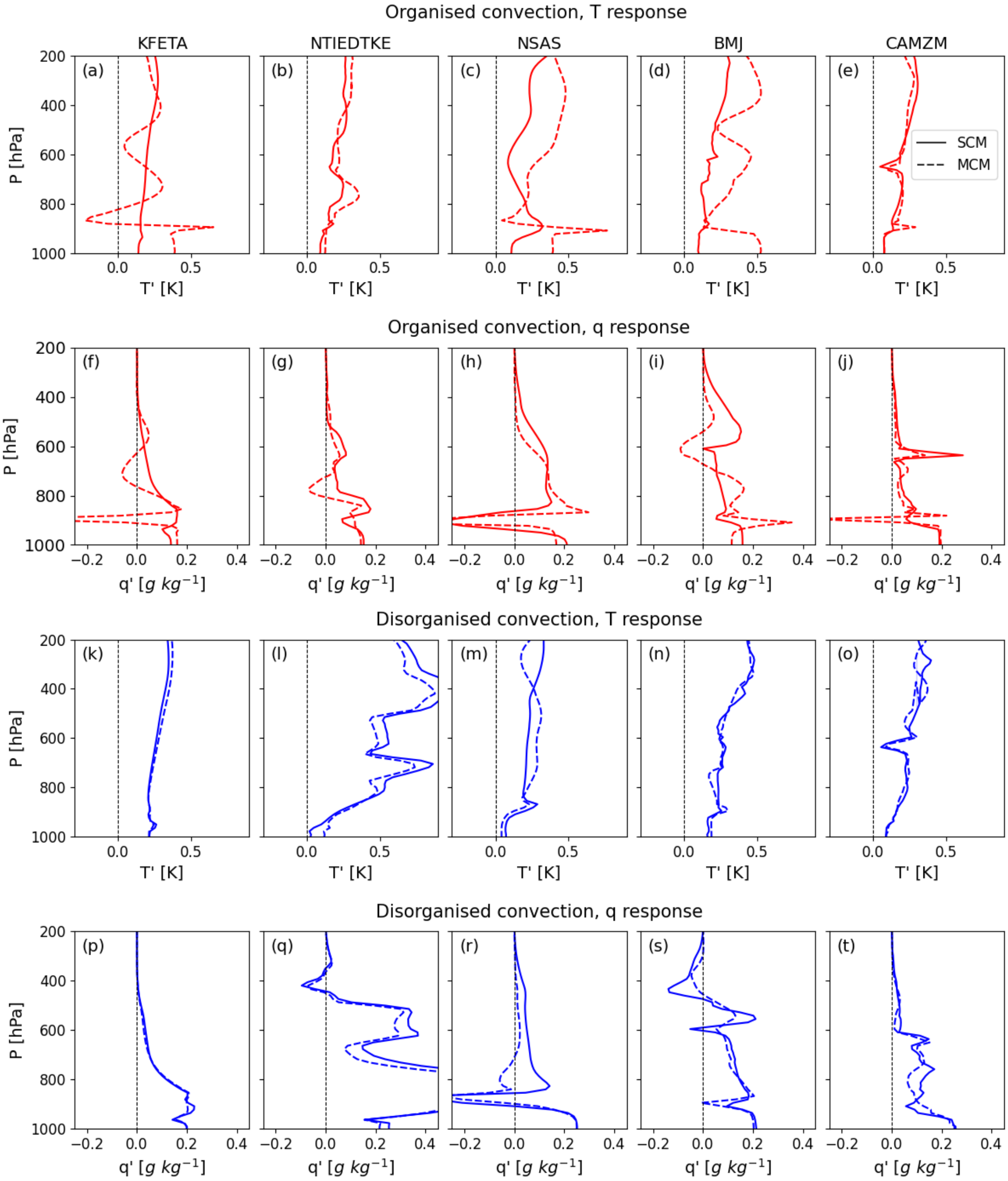

Figure 7. As in Figure 6, but for responses to moisture tendency perturbations.

We quantify the effect of convective organization for all experimental configurations by fitting a linear mixed-effects (*lme*) statistical model. The fixed effects (independent predictors) are the org_{pw} values and the mean state temperature and RH difference between the SCM and MCM setups ($\Delta T_{\mathrm{scm-mcm}}$ and $\Delta RH_{\mathrm{scm-mcm}}$). The dependent variable is the deviation of the SCM-MCM linear responses (SCM-MCM-deviation). The random effects are the five convection schemes and eight experimental configurations, which the *lme* model controls for by taking into account the random variability due to individual differences between them. We measure profile differences (SCM-MCM-deviation, $\Delta T_{\mathrm{scm-mcm}}$ and $\Delta RH_{\mathrm{scm-mcm}}$) by using a simple root-mean-square deviation (RMSD) between the profiles, normalized by their mean (NRMSD). We log-transformed org_{pw} and $\Delta RH_{\mathrm{scm-mcm}}$ as initial scatter plots revealed a nonlinear relationship between them and SCM-MCM-deviation. Results are presented in Figure 8. We found a high correlation between org_{pw} and $\Delta RH_{\mathrm{scm-mcm}}$ ($r = 0.89$; Figure 8a), confirming an effect of convective organization on SCM-MCM mean state difference as described in the previous section. Note that due to this high correlation it is inappropriate to use both org_{pw} and $\Delta RH_{\mathrm{scm-mcm}}$ as independent predictors in the same *lme* model. We deem org_{pw} a more suitable predictor as a high RH difference between the SCM and MCM is likely a result of a highly aggregated MCM mean state. Our *lme* model confirms a strong effect of org_{pw} on SCM-MCM-deviation ($p < 0.001$; Figure 8b). Interestingly, we did not find an effect of $\Delta T_{\mathrm{scm-mcm}}$ on SCM-MCM-deviation ($p > 0.05$). $\Delta T_{\mathrm{scm-mcm}}$ is also only moderately correlated with org_{pw} ($r = 0.5$), indicating a stronger influence of convective organization on mean state humidity than on temperature.

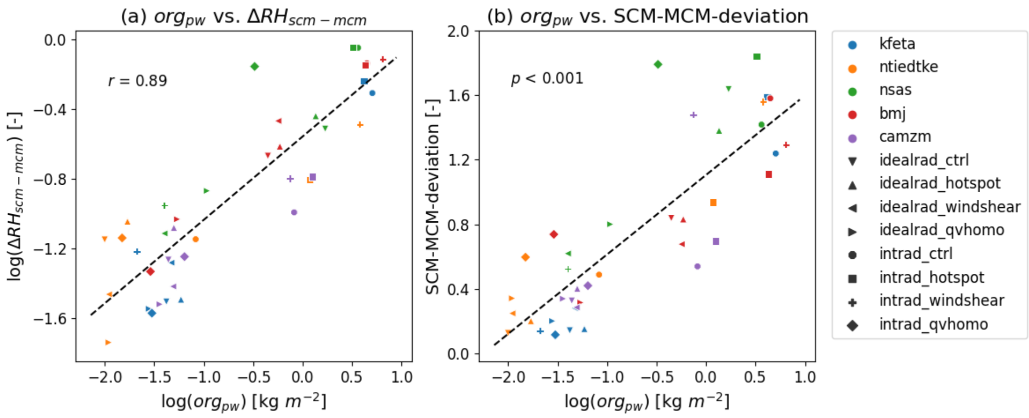

Figure 8. Scatter plots of (a) $log(org_{\mathrm{pw}})$ vs. $log(\Delta RH_{\mathrm{scm-mcm}})$, (b) $log(org_{\mathrm{pw}})$ vs. SCM-MCM-deviation. The black dashed line in (b) is the regression line from the fitted *lme* model after all fixed and random effects are controlled for.

The results in this section suggest that in idealized and homogeneous conditions where convection is disorganized, the behavior of a 3D model in RCE is likely predictable from its 1D counterpart. However, when convection is organized, multiple factors that complicate matters can come into play, and an SCM cannot automatically be expected to be able to accurately capture the behavior of a more realistic 3D setup. The strong correlation between the degree of organization in MCM and the SCM-MCM mean state RH difference shows that the drier mean state associated with highly aggregated conditions leads to a larger disparity between the humidity profiles of the two setups. These factors tend to then disrupt the comparability of behavior between 1D and 3D models.

4.2 SCM-MCM Relative Difference

So far we have shown that when convection is organized, the SCM and MCM responses to small tendency perturbations tend to become dissimilar. However, one important question remains about their relative comparability – can the relative difference between a pair of schemes in SCM predict anything meaningful about the difference between them in their corresponding MCM setup? Concretely, for the five convection schemes in our study—10 unique pairs of scheme combination—if the difference between a given pair of schemes (e.g., KFETA-NTIEDTKE) is larger in an SCM setup compared to another pair (e.g., KFETA-NSAS), will it also be larger in an MCM setup? In addressing this question, we invoke an approach akin to the "model-as-truth" concept (Herger et al., 2018), where we test whether an SCM-based model evaluation (selecting the model physics that gives an SCM closest to "truth") would also provide better MCM results, even though the SCM and MCM results may not align perfectly, for example due to convective organization.

We use the response profiles of the `all_org` and `all_disorg` configurations for our analysis. In addition to the responses to perturbations at 850 hPa, we conducted an additional set of experiments where we applied perturbations at 650 hPa for the `all_org` and `all_disorg` configurations to obtain more data points and improve statistical confidence. We measure the difference between the response profiles of a pair of schemes using the NRMSD metric mentioned in the previous section, referred to as SCM-pair-difference and MCM-pair-difference for the SCM and MCM setups respectively.

For statistical analysis we again fit a *lme* model. The independent predictor is the SCM-pair-difference and the dependent variable the MCM-pair-difference. The random effects that we control for are the scheme pairs (10 combinations), perturbation levels (850 and 650 hPa), and the quadrants (T_DT, Q_DT, T_DQ and Q_DQ). We fit two separate *lme* models for the `all_org` and `all_disorg` configurations (`lme_org` and `lme_disorg`). Each model includes 80 data points (10 scheme combinations $\times$ 2 perturbation levels $\times$ 4 quadrants).

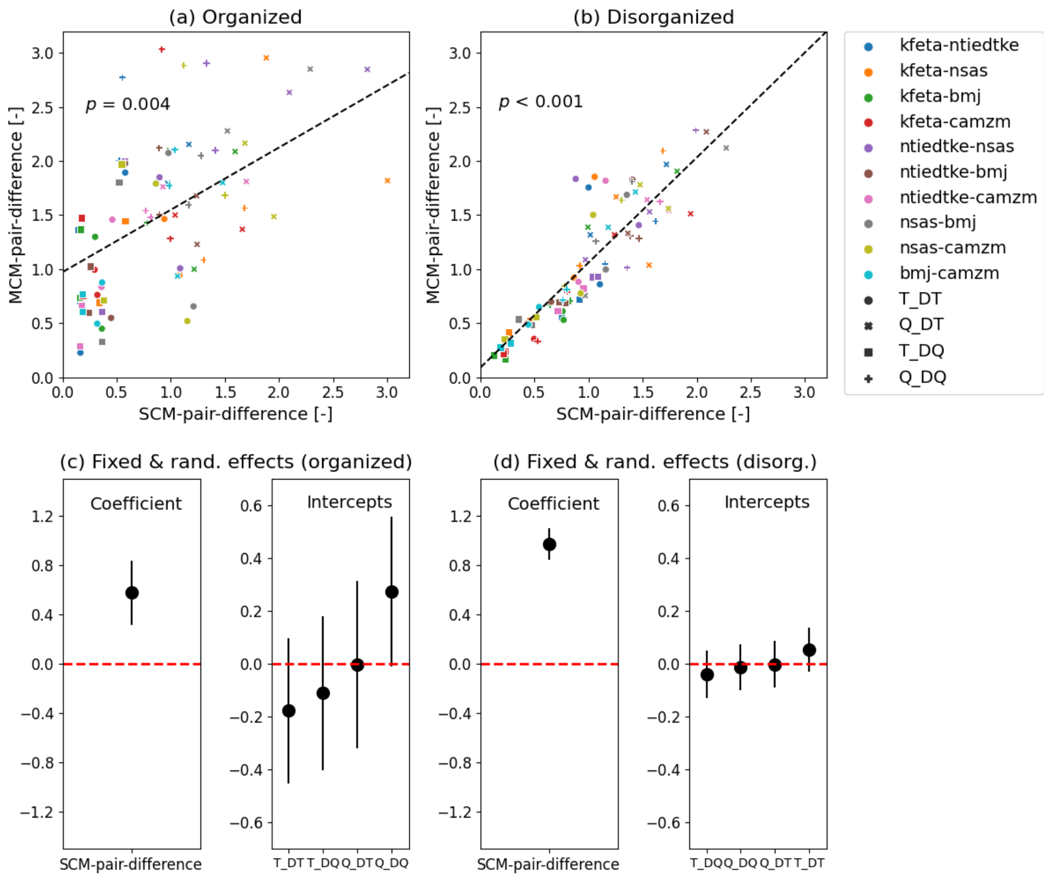

Figure 9. Scatter plots of SCM vs. MCM pair differences for the (a) `all_org` and (b) `all_disorg` simulations. The black dashed lines are the regression lines from the fitted *lme* models and the p-values for SCM-pair-difference are annotated in the top left corners. The 95% confidence intervals of the coefficient for SCM-pair-difference and the intercepts of the "quadrant" random effect are shown for (c) `lme_org` and (d) `lme_disorg`. Intervals that do not overlap zero are significant at a 95% confidence level.

Figure 9 shows a summary of the analysis. As expected, for `lme_disorg` there is a very high correlation between SCM-pair-difference and MCM-pair-difference (Figure 9b). There is strong evidence of SCM-pair-difference as an effective predictor of MCM-pair-difference, after taking into account the variations of the random effects ($p < 0.001$; Figure 9d, 95% confidence interval of the coefficient for SCM-pair-difference does not overlap zero). The random effects do not have a significant impact on the outcome (Figure 9d; only the effects of "quadrant" are shown and not the other random effects as they all have zero intercepts). This is consistent with our results so far that show high SCM-MCM similarity when convection is disorganized. For `lme_org`, a faint linear relationship between SCM-pair-difference and MCM-pair-difference can still be observed, albeit substantially weaker (Figure 9a). The statistical analysis reveals there is evidence for using SCM-pair-difference to predict MCM-pair-difference ($p = 0.004$; Figure 9c, 95% confidence interval of the coefficient does not overlap zero). This implies that even when the responses do not match up perfectly between the SCM and MCM setups (e.g., as a result of organization in MCM), two schemes that are very different in a 1D setup can also be expected to be very different in a 3D setup. However, for `lme_org` there is a non-negligible effect of the "quadrant" random effect (Figure 9c, 95% confidence interval of the intercepts of the four quadrants do not always overlap zero). Specifically, the Q_DQ quadrant has a strong impact on the outcome (i.e., considerable variance of the outcome comes from this quadrant), which skews the relationship between the SCM and MCM pairs. The unpredictability of moisture responses again points to the important role of moisture in convective organization: the interplay or feedback between convection and moisture becomes more apparent and important in a 3D setup, thereby weakening the predictive power of the corresponding SCM.

The results presented in this section extend the conclusions drawn in the previous section and suggest that, even in a more realistic scenario (i.e., when convection is organized), the SCM can still serve as a useful tool to investigate model physics. For example, we have a reference profile (e.g., responses of a "good" convection scheme), and the relative distance from this reference is a measure of the performance of a new convection scheme X. Our results here imply that if the difference between this reference profile and the profile of scheme X is large in an SCM setup (compared to an old scheme Y, for example), we can expect a correspondingly large difference between them in an MCM setup (which might be unavailable or impractical to calculate). However, this interpretation should be applied with caution when it concerns moisture responses.

5 Response to Doubled-CO_2 Forcing

5.1 Comparison of MCM-SCM Response to Doubled CO_2

In this section we compare the SCM and MCM responses to doubled-CO_2 forcing ($2\times CO_2$), with interactive radiation. The response profiles are computed as the time-averaged difference between the `PerturbCO2` and `CTRL` experiments (described in Section 2.3). Surface temperature is held constant, so the response computed here is the CO_2 adjustment responses (Sherwood et al., 2015). Results are shown in Figure 10. The response profiles of temperature (**T'**), relative humidity (**RH'**), cloud fraction (**CLDFRA'**) and radiative heating ($\mathbf{Q'}_{rad}$) are shown.

Overall, the schemes respond quite differently in their SCM vs. MCM setups. The tropospheric temperature increases in all SCMs, but not all MCMs (the KFETA run cools in the free troposphere, Figure 10a). The net radiative heating rate responses (p–t in Figure 10) are the atmospheric effective radiative forcing (AERF), which is analogous to the top-of-atmosphere (TOA) effective radiative forcing (ERF) following Boucher et al. (2013). The AERF shown here includes both the instantaneous (initial) atmospheric radiative forcing (ARF) to $2\times CO_2$ forcing—generally a warming peaking in the lower troposphere and cooling in the upper troposphere (dashed curves in p–t in Figure 10; see also Iacono et al., 2008)—and adjustments due to cloud and other responses (a–o in Figure 10). In most cases, AERF and ARF are quite similar, indicating that the CO_2 adjustments alter the instantaneous forcings by only a small, albeit non-negligible, amount. The same has been found in

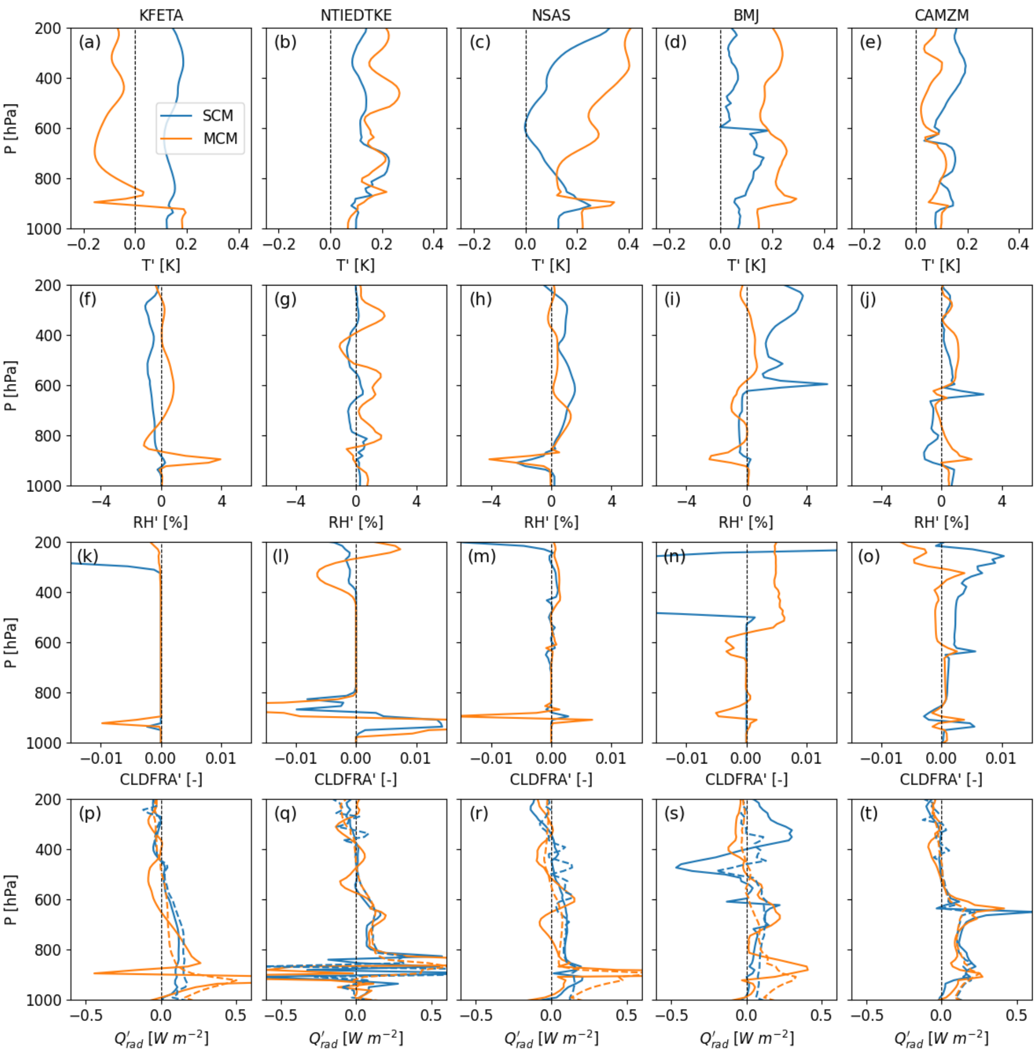

Figure 10. Responses of the five convection schemes to doubled-CO_2 forcing for (a–e) temperature, (f–j) relative humidity, (k–o) cloud fraction, and (p–t) radiative heating. Light blue curves show the SCM responses and orange curves the MCM responses. Dashed curves in (p–t) show the instantaneous (initial) radiative forcing.

previous studies for TOA effective radiative forcing, or ERF (Vial et al., 2013; Zelinka et al., 2013), but as far as we know, this is the first study that shows the vertical distribution of this forcing in terms of AERF. The general trend of radiative forcing (heating increases in lower- to mid-troposphere and decreases in the upper troposphere) can be observed in both SCM and MCM setups, but frequently with different forcing shapes. Sharp responses in clouds (**CLDFRA'**) and **RH'** are often observed around the cloud base and freezing levels, albeit sometimes with inconsistent signs between the SCM and MCM setups (f–o in Figure 10). Notably, the MCMs tend to exhibit larger spikes compared to SCMs with the same scheme, in particular around the cloud base level in their **T'** and **RH'** (orange curves, a–j in Figure 10). These spikes are more noticeable for schemes that produce higher org_{pw} values in the `intrad_ctrl` configuration (KFETA, NSAS and BMJ in Figure 3e) and are probably related to temperature inversions.

Nevertheless, similarities can sometimes be observed between the SCMs and MCMs, particularly for models that are more disorganized in their MCM setup. Specifically, NTIEDTKE and CAMZM display relatively homogeneous convective organization patterns in the `intrad_ctrl` configuration (Figure 3e), and their 2×CO_2 response profiles are also comparatively more similar between SCM and MCM. For NTIEDTKE—which produces the most disorganized convection amongst the schemes in `intrad_ctrl`—the general shape and magnitude of **T'** and **RH'** are comparable between SCM and MCM (b and g in Figure 10). As for **CLDFRA'**, substantial low cloud changes tend to occur in the same direction and altitude in SCM and MCM, albeit with different magnitudes (Figure 10l). Some discrepancies are observed in the boundary layer heating ($\mathbf{Q'}_{rad}$), but the profiles are remarkably similar in the mid-troposphere (Figure 10q). For CAMZM—which is slightly more organized than NTIEDTKE—similar shapes and magnitudes are observed in both setups for **T'** and to a lesser extent **RH'**, but with the kinks around the cloud base level trending in opposite directions (e and j in Figure 10), again probably due to temperature inversion in the MCM. Changes in cloud fraction are quite similar, especially in the mid-troposphere, with the spike around the freezing level (600 hPa) captured in both SCM and MCM (Figure 10o). The radiative heating responses are remarkably similar, with the big spike around the freezing level again represented in both setups, albeit with slightly different magnitudes (Figure 10t). By contrast, big disparities between the SCM and MCM setups are observed in the adjustment responses of the three other schemes, which are considerably more organized in their MCM setups. Slight resemblance can sometimes be observed in the shapes of their **T'**, but not in magnitudes (NSAS and BMJ, c and d in Figure 10).

Using a different forcing scenario, this section again highlights the role of convective organization in influencing SCM-MCM comparability. We find that the higher the aggregation proclivity of a scheme, the less reliable its corresponding SCM results, although certain important features are sometimes preserved, e.g., general trend in radiative forcing and spikes in low cloud changes, which are noticeable in the SCM and accentuated in the MCM setup. Overall, we show the potential usefulness of SCMs in climate change research by demonstrating that in an idealized, relatively homogeneous scenario, an SCM can predict how a 3D model would respond to a doubling of CO_2 in the atmosphere.

5.2 Predicting Response to Doubled $\mathbf{CO_2}$ with $\mathbf{M^{-1}}$

In this section we test the feasibility of using the LRF method to predict model adjustment responses to doubled-CO_2 forcing, using the experiments described in Section 2.3. The aim is to verify the usefulness of the LRF method in a climate change scenario and explore the potential implications for SCM-MCM comparability.

Recall that in the LRF framework (Equation 2), the matrices M^{-1} show the model responses per unit perturbation, e.g., [K/(W m^{-2})] for temperature and [(g kg^{-1})/(W m^{-2})] for moisture responses. Figure 11 shows the M^{-1} matrices of the five convection schemes in the SCM setup for the `intrad_ctrl` configuration. The x-axis is the perturbation level and y-axis the response level, so each column is the response profile per unit perturbation at a given level, and each row is the response at a given level as a function of where the perturbation is applied. Generally, the SCMs exhibit stronger responses near the perturbations (diagonal of the matrices), except for the BMJ scheme, which displays uniform responses across a broad range of perturbation levels (d and i in Figure 11). The non-local responses (off-diagonal) vary greatly across the five schemes, with the biggest disparities observed in the boundary layer. Negative responses are sometimes observed (e.g., for NSAS, c and h in Figure 11). Discontinuities in responses (horizontal stripes) are found consistently around the cloud base level and sometimes also the freezing level (BMJ and CAMZM).

Note that the M^{-1} shown here are equivalent to those shown in Figures 4 and 5 of H21, except that here the SCMs have interactive radiation whereas in H21 radiation was fixed. Most of the above features are similar to those noted by H21, indicating that many peculiarities of convection schemes are preserved when radiation is interactive, but not all. We note that a detailed investigation into the matrices of the individual schemes is beyond the scope of this study. Our intention is to see whether these LRF matrices can help predict CO_2 adjustment responses.

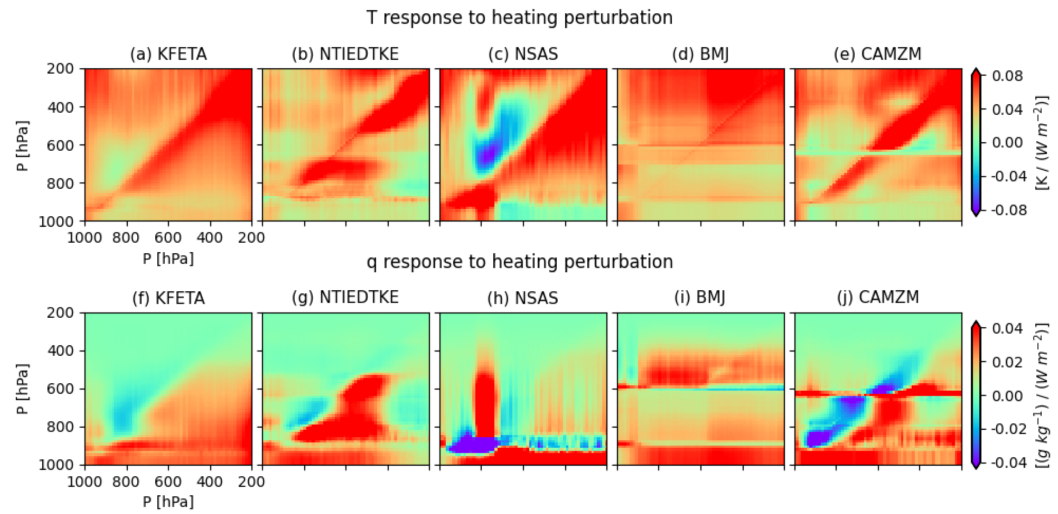

Figure 11. M^{-1} matrices of the (a–e) temperature and (f–j) moisture responses to heating perturbations for the five convection schemes in an SCM setup.

As described in Section 2.3, we compute the LRF-predicted responses of the five schemes to doubled-CO_2 forcing by multiplying their M^{-1} matrices by their respective $2\times CO_2$ instantaneous radiative forcing vector (dashed light blue curves in p–t in Figure 10). This is a vector of heating and moistening, with the heating set equal to the radiative forcing and the moistening set to zero. Figure 12 shows a comparison of the model-simulated and LRF-predicted temperature and RH responses with the five schemes. Overall, the simulated and predicted profiles are strikingly similar, except for the BMJ scheme. Notably, the kinks around the cloud base level—for CAMZM also the freezing level—are reproduced in the predicted profiles. For BMJ, the predicted responses depart considerably from the simulation (d and i in Figure 12). The kinks around the cloud base and freezing levels are predicted, but often in opposite directions as in the simulated profiles. The BMJ simulated responses show a marked discontinuity around the freezing level ($\sim$ 600 hPa), which might reflect threshold-related mechanisms embedded in the scheme to demarcate between shallow and deep convection. This discontinuity is somewhat reproduced by the LRF-predicted responses, albeit with opposite trend above the freezing level for the temperature response. It is unclear what is the source of this discrepancy for BMJ. One potential explanation could be that a switch-like mechanism around the freezing level causes nonlinearity that disrupts its linearized behavior, leading to diverging responses between the simulation and prediction.

To further test if the M^{-1} matrices constructed in an SCM setup can be used to predict the MCM responses to doubled CO_2, we compared the simulated temperature and RH responses in MCM to the ones predicted by multiplying the SCM matrices with the MCM radiative forcing vector (dashed orange curves in p–t in Figure 10). Results (not shown) show that simulated and predicted profiles differ considerably for the schemes that are highly organized in MCM, as expected, whereas for the relatively disorganized schemes they are comparable.

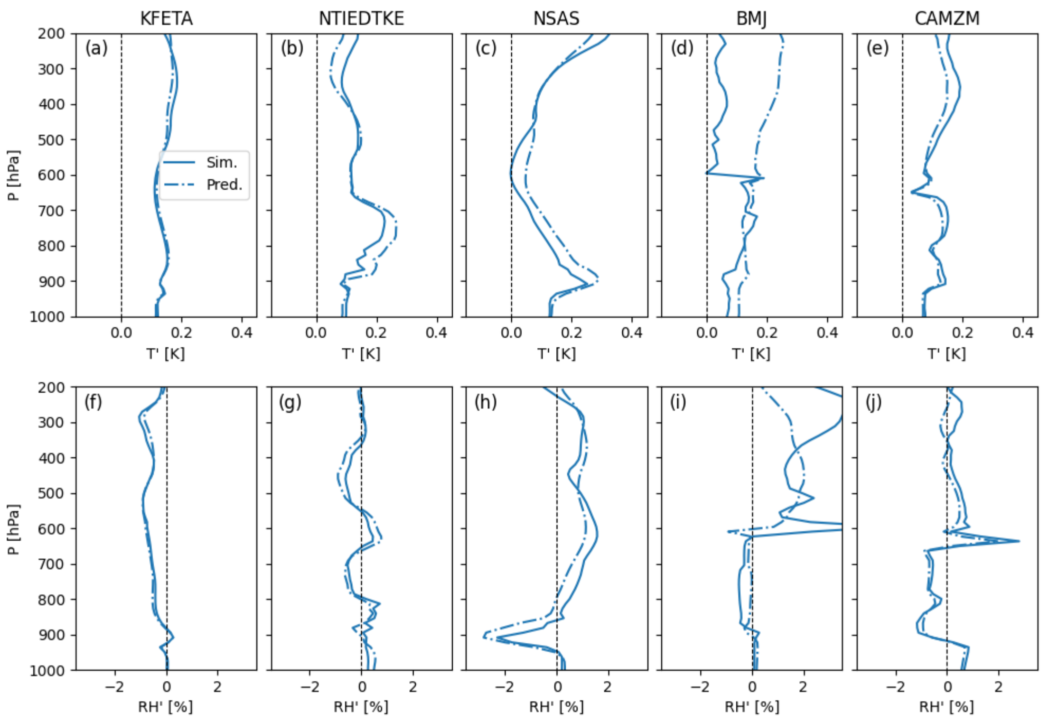

Figure 12. Comparison of (solid) model-simulated and (dot-dashed) LRF-predicted (a–e) temperature and (f–j) relative humidity responses to doubled-CO_2 forcing for the five convection schemes.

Our results confirm the relevance of the LRF framework in predicting model adjustment responses in a climate change scenario, thus substantiating our decision to use it as one of the tests in this study. We did not run the matrix simulations in the MCM setup due to resource constraints. However, our findings suggest that, under relatively homogeneous conditions, these expensive matrix simulations need only be run in an SCM setup, and the results can be used to predict the responses to various forcing scenarios in a more realistic 3D setup with relative accuracy. For organized convection, on the other hand, this might not be the case. We have shown that the linear responses diverge between the SCM and MCM when convection is organized. In such cases, matrices constructed using an SCM cannot be reliably used to predict a model's responses in a 3D scenario. It may be possible with further effort to find ways of better predicting 3D results using SCM linear responses, for example by adding a parameterization of a larger-scale environment (see Brenowitz et al., 2020) or by applying some forcing in the SCM, but this is deferred to future work.

6 Conclusions

The main objective of this paper is to investigate the relevance of single-column models (SCMs) in a radiative-convective equilibrium (RCE) setting for testing model physics and/or predicting model responses in a more realistic 3D setup, termed here "comparability". We also explore the influence of convection schemes on convective organization and the role of convective organization in the aforementioned comparability. We use a 20 × 20 multi-column model (MCM) configuration with periodic boundary conditions and model columns matching the SCM as a stepping stone to a 3D setup, testing five widely-used convection schemes. We examine the behavior of both model setups by probing their responses to small tendency perturbations following the linear response function (LRF) framework of Kuang

(2010), as well as their adjustment responses to doubled-CO_2 forcing. Four main conclusions can be drawn from our results:

1. Convection schemes strongly influence convective organization, comparably to other factors known to organize convection such as interactive radiation.
2. Convective organization in turn has a strong impact on SCM-MCM comparability, with more organization associated with less comparability.
3. When convection is organized, differences in linear responses between schemes are nonetheless largely preserved between the SCM and MCM, albeit less so for moisture than for temperature responses.
4. The LRF matrix of the SCMs can be used to predict their adjustment responses to doubled CO_2, suggesting practical applications of the LRF method.

Regarding conclusion (1), we find that even when a fixed horizontally homogeneous radiative cooling profile is imposed—thus denying the longwave radiative feedback which has been found to be key to aggregation in numerous studies—two schemes (NSAS and BMJ) still produce organization. In the same vein, two schemes (NTIEDTKE and CAMZM) produce disorganized convection even *with* interactive radiation, perhaps because the limited domain size inhibits large-scale circulations. Vertical wind shear is found to have opposite effects depending on scheme: it homogenizes convection with KFETA and NSAS, but causes convection with NTIEDTKE and BMJ to become more organized, indicating that much is left to be understood about the effect of wind shear on aggregation, echoing the review by Wing et al. (2017). To obtain disorganized convection across all schemes required deploying horizontal moisture homogenization and fixing radiation, while to consistently obtain organized convection required interactive radiation and an imposed SST hot spot. Our results thus show that for models that do not resolve convection explicitly, convection schemes appear to have a bigger impact on aggregation than factors commonly accepted to organize convection (e.g., radiative feedback). It is difficult to untangle the root causes of these differences, as a range of processes can affect aggregation; explaining the variations in behavior identified here requires extensive sensitivity tests and probably intimate understanding of the schemes, which is beyond the scope of this study.

Convective organization has been shown to impact extreme precipitation (Bao et al., 2017; Pendergrass et al., 2016) and tropical cyclones (Muller & Romps, 2018). Hence its proper representation in models is important, and our results show that there are still considerable uncertainties associated with it arising from parameterizations (see also Bador et al., 2018). Past studies have found that organization in global atmospheric models can be model-dependent (Wing et al., 2017; Wing, 2019; Wing et al., 2020), but we believe this is the first study to show this can be specifically attributed to convection schemes.

Regarding conclusion (2), SCM-MCM comparability is strongly influenced by the organization seen in the MCM: the more organized it is, the larger the divergence between SCM and MCM behavior as measured by their mean states, linear responses to small tendency perturbations and adjustments to doubled-CO_2 forcing. On the other hand, when convection is disorganized in the MCM, these measures are very similar (at times indistinguishable). Consistent with previous studies, we find that a more aggregated state is drier and frequently also warmer. This leads to a larger difference in the mean state profiles between the SCM and MCM setups and hence also their linear responses, as these responses are correlated with model mean state, as shown in Hwong et al. (2021). Schemes that are relatively more disorganized in their MCM setup also display more similar adjustment responses to doubled-CO_2 forcing. This has important implications for the understanding of climate change, as SCMs are sometimes used to examine climate sensitivity (e.g., Kluft et al., 2019; Wing et al., 2020). Our findings on the influence of convective organization on SCM-MCM comparability have important implications for the use of SCM in an RCE setting. The lack of dynamical feedback in SCMs may cause them to behave differently from their corresponding full GCMs if interaction between model physics and large-scale circulation is important, which is the case when convection is organized (Bretherton et al., 2005; Muller & Bony, 2015). Our findings suggest that when convection is organized, studies drawing conclusions from SCM experimentes need to be interpreted with prudence.

Regarding (3), it could be argued that for an SCM test to be relevant to realistic models, if it yields similar results in an SCM with two different physics packages, then a similar test in the two corresponding MCMs should also produce relatively similar results (whether we can predict them or not). Our results (measuring similarity by a simple RMS distance metric) confirm this to a certain extent: in an organized state, a pair of schemes (e.g., KFETA-NTIEDTKE) that produces more similar (but not necessarily identical) behavior in the SCM also does so in the MCM, compared to other pairs of schemes that are less similar in the SCM. As such, SCMs can be a useful tool in the development of parameterizations when we have an SCM verification standard: if a scheme produces behavior closer to the standard in an SCM setup, it will probably also perform better in a 3D setup. However, a few caveats must be noted. In contrast to disorganized cases, notable features of the responses are usually not preserved between the SCM and MCM when convection is organized, even though their averaged distances are similar. Hence, if these features (e.g., kinks around cloud base level) are important, the SCM is less useful. Further, our conclusion is less reliable when it concerns moisture responses. These responses tend to trend in more unpredictable directions between the SCM and MCM setups, probably because they are more affected by convective organization.

Regarding the final conclusion, we tested the practical value of the LRF approach by examining whether it could predict the SCM adjustments responses to a doubling of CO_2 in the atmosphere. The answer is yes, except for the BMJ scheme whose matrix predicts adequate humidity but inaccurate temperature response, particularly above the freezing level. For the other schemes, prominent features are reproduced, such as spikes around the cloud base and freezing levels. Since the radiative forcing caused by different climate change agents can be estimated from radiative transfer calculations (Clough & Iacono, 1995; Collins et al., 2006), our results imply that the adjustment responses of different schemes to these agents can be compared by simple linear algebra calculations using their LRF matrices, without having to import different schemes into the same host model. This could potentially be helpful in climate change research, where parameterizations are a major contributor to intermodel spread in climate sensitivity predictions (e.g., Geoffroy et al., 2017; Ringer et al., 2014; Sherwood et al., 2014; Webb et al., 2013). Moreover, combined with conclusion (2), our results suggest that when convection is disorganized in MCM (hence high SCM-MCM comparability), the LRF matrices can be constructed using SCMs, thereby drastically reducing the computing overhead, yet still ensuring adequate representation of the parameterization behavior in a 3D setting.

This study represents only a first step in exploring the extent to which atmospheric behavior in a small isolated column can be used to learn anything about the broader atmosphere. There is a long tradition of attempting this to better understand climate and to test process models, but little systematic exploration of behavior comparability. We acknowledge that our MCM setup is still relatively idealized, and may not be sufficiently representative of a realistic 3D RCE scenario. Additionally, the domain size (2000 $\times$ 2000 km) might not be large enough to allow more realistic circulations to form, potentially impeding certain feedback mechanisms. Also, we have only explored a limited set of tests and measures of success. The conclusions drawn here thus deserve further investigation using more realistic setups and other tests. Apart from convective organization, there may be other factors that contribute to the comparability of results in 1D vs. 3D setups. For results obtained in SCM in RCE to be more convincing, more research is needed to explore these factors so that they can be properly controlled for.

Acknowledgments
The authors thank Isaac Held, Tim Raupach, Maxime Colin and Abhnil Prasad for use-

ful discussions. The authors were supported by the Australian Research Council (grant FL150100035). The authors would also like to thank the Australian Research Council Centre of Excellence for Climate Extremes (CLEX) and the Australian National Computational Infrastructure (NCI) for providing computing facility and technical support to this project.

The data and scripts required to reproduce the results described in this paper are available at: `https://zenodo.org/record/5225430` (DOI: 10.5281/zenodo.5225430)

References

Abbot, D. S. (2014). Resolved snowball earth clouds. *Journal of Climate*, *27*(12), 4391–4402.

Arnold, N. P., & Randall, D. A. (2015). Global-scale convective aggregation: Implications for the madden-julian oscillation. *Journal of Advances in Modeling Earth Systems*, *7*(4), 1499–1518.

Bador, M., Donat, M. G., Geoffroy, O., & Alexander, L. V. (2018). Assessing the robustness of future extreme precipitation intensification in the cmip5 ensemble. *Journal of Climate*, *31*(16), 6505–6525.

Bao, J., Sherwood, S. C., Colin, M., & Dixit, V. (2017). The robust relationship between extreme precipitation and convective organization in idealized numerical modeling simulations. *Journal of Advances in Modeling Earth Systems*, *9*(6), 2291–2303.

Bechtold, P., Redelsperger, J.-L., Beau, I., Blackburn, M., Brinkop, S., Grandper, J.-Y., ... others (2000). A gcss model intercomparison for a tropical squall line observed during toga-coare. ii: Intercomparison of single-column models and a cloud-resolving model. *Quarterly Journal of the Royal Meteorological Society*, *126*(564), 865–888.

Becker, T., & Stevens, B. (2014). Climate and climate sensitivity to changing co 2 on an idealized land planet. *Journal of Advances in Modeling Earth Systems*, *6*(4), 1205–1223.

Becker, T., Stevens, B., & Hohenegger, C. (2017). Imprint of the convective parameterization and sea-surface temperature on large-scale convective self-aggregation. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1488–1505.

Betts, A. (1986). A new convective adjustment scheme. part i: Observational and theoretical basis. *Quarterly Journal of the Royal Meteorological Society*, *112*(473), 677–691.

Betts, A., & Miller, M. (1986). A new convective adjustment scheme. part ii: Single column tests using gate wave, bomex, atex and arctic air-mass data sets. *Quarterly Journal of the Royal Meteorological Society*, *112*(473), 693–709.

Beucler, T., & Cronin, T. W. (2016). Moisture-radiative cooling instability. *Journal of Advances in Modeling Earth Systems*, *8*(4), 1620–1640.

Blossey, P. N., Bretherton, C. S., Zhang, M., Cheng, A., Endo, S., Heus, T., ... Xu, K.-M. (2013). Marine low cloud sensitivity to an idealized climate change: The cgils les intercomparison. *Journal of Advances in Modeling Earth Systems*, *5*(2), 234–258.

Bogenschutz, P., Gettelman, A., Morrison, H., Larson, V., Schanen, D., Meyer, N., & 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.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., ... others (2013). Clouds and aerosols. In *Climate change 2013: the physical science basis. contribution of working group i to the fifth assessment report of the intergovernmental panel on climate change* (pp. 571–657). Cambridge University Press.

Brenowitz, N. D., Beucler, T., Pritchard, M., & Bretherton, C. S. (2020). Interpreting and stabilizing machine-learning parametrizations of convection. *Journal of the Atmospheric Sciences*, *77*(12), 4357–4375.

Bretherton, C. S., Blossey, P. N., & Khairoutdinov, M. (2005). An energy-balance analysis of deep convective self-aggregation above uniform sst. *Journal of the atmospheric sciences*, *62*(12), 4273–4292.

Bryan, G. H. (2005). Spurious convective organization in simulated squall lines owing to moist absolutely unstable layers. *Monthly weather review*, *133*(7), 1978–1997.

Christensen, H. M., Dawson, A., & Holloway, C. E. (2018). Forcing single-column models using high-resolution model simulations. *Journal of advances in modeling earth systems*, *10*(8), 1833–1857.

Clough, S. A., & Iacono, M. J. (1995). Line-by-line calculation of atmospheric fluxes and cooling rates: 2. application to carbon dioxide, ozone, methane, nitrous oxide and the halocarbons. *Journal of Geophysical Research: Atmospheres*, *100*(D8), 16519–16535.

Collins, W., Ramaswamy, V., Schwarzkopf, M. D., Sun, Y., Portmann, R. W., Fu, Q., ... others (2006). Radiative forcing by well-mixed greenhouse gases: Estimates from climate models in the intergovernmental panel on climate change (ipcc) fourth assessment report (ar4). *Journal of Geophysical Research: Atmospheres*, *111*(D14).

Coppin, D., & Bony, S. (2015). Physical mechanisms controlling the initiation of convective self-aggregation in a general circulation model. *Journal of Advances in Modeling Earth Systems*, *7*(4), 2060–2078.

Craig, G., & Mack, J. (2013). A coarsening model for self-organization of tropical convection. *Journal of Geophysical Research: Atmospheres*, *118*(16), 8761–8769.

Dal Gesso, S., & Neggers, R. (2018). Can we use single-column models for understanding the boundary layer cloud-climate feedback? *Journal of Advances in Modeling Earth Systems*, *10*(2), 245–261.

Dunion, J. P. (2011). Rewriting the climatology of the tropical north atlantic and caribbean sea atmosphere. *Journal of Climate*, *24*(3), 893–908.

Emanuel, K., Wing, A. A., & Vincent, E. M. (2014). Radiative-convective instability. *Journal of Advances in Modeling Earth Systems*, *6*(1), 75–90.

Geoffroy, O., Sherwood, S. C., & Fuchs, D. (2017). On the role of the stratiform cloud scheme in the inter-model spread of cloud feedback. *Journal of Advances in Modeling Earth Systems*, *9*(1), 423–437.

Gettelman, A., Morrison, H., & Ghan, S. J. (2008). A new two-moment bulk stratiform cloud microphysics scheme in the community atmosphere model, version 3 (cam3). part ii: Single-column and global results. *Journal of Climate*, *21*(15), 3660–3679.

Ghan, S., Randall, D., Xu, K.-M., Cederwall, R., Cripe, D., Hack, J., ... others (2000). A comparison of single column model simulations of summertime midlatitude continental convection. *Journal of Geophysical Research: Atmospheres*, *105*(D2), 2091–2124.

Grabowski, W. W., & Moncrieff, M. (2004). Moisture–convection feedback in the tropics. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *130*(604), 3081–3104.

Guichard, F., Petch, J., Redelsperger, J.-L., Bechtold, P., Chaboureau, J.-P., Cheinet, S., ... others (2004). Modelling the diurnal cycle of deep precipitating convection over land with cloud-resolving models and single-column models. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *130*(604), 3139–3172.

Guo, Z., Wang, M., Qian, Y., Larson, V. E., Ghan, S., Ovchinnikov, M., ... Zhou, T. (2014). A sensitivity analysis of cloud properties to clubb parameters in the single-column community atmosphere model (scam5). *Journal of Advances in Modeling Earth Systems*, *6*(3), 829–858.

Hacker, J., & Angevine, W. (2013). Ensemble data assimilation to characterize surface-layer errors in numerical weather prediction models. *Monthly weather review*, *141*(6), 1804–1821.

Han, J., & Pan, H.-L. (2011). Revision of convection and vertical diffusion schemes in the ncep global forecast system. *Weather and Forecasting*, *26*(4), 520–533.

Held, I. M., Hemler, R. S., & Ramaswamy, V. (1993). Radiative-convective equilibrium with explicit two-dimensional moist convection. *Journal of Atmospheric Sciences*, *50*(23), 3909–3927.

Herger, N., Abramowitz, G., Knutti, R., Angélil, O., Lehmann, K., & Sanderson, B. M. (2018). Selecting a climate model subset to optimise key ensemble properties. *Earth*

System Dynamics, *9*(1), 135–151.
Herman, M. J., & Kuang, Z. (2013). Linear response functions of two convective parameterization schemes. *Journal of Advances in Modeling Earth Systems*, *5*(3), 510–541.
Hohenegger, C., & Stevens, B. (2016). Coupled radiative convective equilibrium simulations with explicit and parameterized convection. *Journal of Advances in Modeling Earth Systems*, *8*(3), 1468–1482.
Holloway, C. E., & Woolnough, S. J. (2016). The sensitivity of convective aggregation to diabatic processes in idealized radiative-convective equilibrium simulations. *Journal of Advances in Modeling Earth Systems*, *8*(1), 166–195.
Hong, S.-Y., & Lim, J.-O. J. (2006). The wrf single-moment 6-class microphysics scheme (wsm6). *Asia-Pacific Journal of Atmospheric Sciences*, *42*(2), 129–151.
Hong, S.-Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly weather review*, *134*(9), 2318–2341.
Hwong, Y.-L., Song, S., Sherwood, S. C., Stirling, A., Rio, C., Roehrig, R., ... others (2021). Characterizing convection schemes using their responses to imposed tendency perturbations. *Journal of Advances in Modeling Earth Systems*, *13*(5), e2021MS002461.
Iacobellis, S. F., & Somerville, R. C. (1991). Diagnostic modeling of the indian monsoon onset. part i: Model description and validation. *Journal of Atmospheric Sciences*, *48*(17), 1948–1959.
Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., & Collins, W. D. (2008). Radiative forcing by long-lived greenhouse gases: Calculations with the aer radiative transfer models. *Journal of Geophysical Research: Atmospheres*, *113*(D13).
Janjić, Z. I. (1994). The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Monthly weather review*, *122*(5), 927–945.
Jeevanjee, N., Hassanzadeh, P., Hill, S., & Sheshadri, A. (2017). A perspective on climate model hierarchies. *Journal of Advances in Modeling Earth Systems*, *9*(4), 1760–1771.
Jeevanjee, N., & Romps, D. M. (2013). Convective self-aggregation, cold pools, and domain size. *Geophysical Research Letters*, *40*(5), 994–998.
Kain, J. S. (2004). The kain–fritsch convective parameterization: an update. *Journal of applied meteorology*, *43*(1), 170–181.
Kluft, L., Dacie, S., Buehler, S. A., Schmidt, H., & Stevens, B. (2019). Re-examining the first climate models: Climate sensitivity of a modern radiative–convective equilibrium model. *Journal of Climate*, *32*(23), 8111–8125.
Kuang, Z. (2010). Linear response functions of a cumulus ensemble to temperature and moisture perturbations and implications for the dynamics of convectively coupled waves. *Journal of the atmospheric sciences*, *67*(4), 941–962.
Kuang, Z. (2018). Linear stability of moist convecting atmospheres. part i: From linear response functions to a simple model and applications to convectively coupled waves. *Journal of the Atmospheric Sciences*, *75*(9), 2889–2907.
Lane, D. E., Somerville, R. C., & Iacobellis, S. F. (2000). Sensitivity of cloud and radiation parameterizations to changes in vertical resolution. *Journal of climate*, *13*(5), 915–922.
Lee, W.-H., Iacobellis, S. F., & Somerville, R. C. (1997). Cloud radiation forcings and feedbacks: General circulation model tests and observational validation. *Journal of climate*, *10*(10), 2479–2496.
LeMone, M. A., Zipser, E. J., & Trier, S. B. (1998). The role of environmental shear and thermodynamic conditions in determining the structure and evolution of mesoscale convective systems during toga coare. *Journal of the Atmospheric Sciences*, *55*(23), 3493–3518.
Lenderink, G., Siebesma, A. P., Cheinet, S., Irons, S., Jones, C. G., Marquet, P., ... others (2004). The diurnal cycle of shallow cumulus clouds over land: A single-column model intercomparison study. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*,

130(604), 3339–3364.
Liu, C., & Moncrieff, M. W. (2008). Explicitly simulated tropical convection over idealized warm pools. *Journal of Geophysical Research: Atmospheres*, *113*(D21).
Madden, R. A., & Julian, P. R. (1994). Observations of the 40–50-day tropical oscillation—a review. *Monthly weather review*, *122*(5), 814–837.
Maher, P., Gerber, E. P., Medeiros, B., Merlis, T. M., Sherwood, S. C., Sheshadri, A., ... Zurita-Gotor, P. (2019). Model hierarchies for understanding atmospheric circulation. *Reviews of Geophysics*, *57*(2), 250–280.
Manabe, S., & Wetherald, R. T. (1967). Thermal equilibrium of the atmosphere with a given distribution of relative humidity. *Journal of the Atmospheric Sciences*.
Muller, C. (2013). Impact of convective organization on the response of tropical precipitation extremes to warming. *Journal of climate*, *26*(14), 5028–5043.
Muller, C., & Bony, S. (2015). What favors convective aggregation and why? *Geophysical Research Letters*, *42*(13), 5626–5634.
Muller, C., & Held, I. M. (2012). Detailed investigation of the self-aggregation of convection in cloud-resolving simulations. *Journal of the Atmospheric Sciences*, *69*(8), 2551–2565.
Muller, C., & Romps, D. M. (2018). Acceleration of tropical cyclogenesis by self-aggregation feedbacks. *Proceedings of the National Academy of Sciences*, *115*(12), 2930–2935.
Neggers, R., Ackerman, A. S., Angevine, W., Bazile, E., Beau, I., Blossey, P., ... others (2017). Single-column model simulations of subtropical marine boundary-layer cloud transitions under weakening inversions. *Journal of Advances in Modeling Earth Systems*, *9*(6), 2385–2412.
Nie, J., & Sobel, A. H. (2016). Modeling the interaction between quasigeostrophic vertical motion and convection in a single column. *Journal of the Atmospheric Sciences*, *73*(3), 1101–1117.
Pakula, L., & Stephens, G. L. (2009). The role of radiation in influencing tropical cloud distributions in a radiative–convective equilibrium cloud-resolving model. *Journal of the atmospheric sciences*, *66*(1), 62–76.
Park, S., & Bretherton, C. S. (2009). The university of washington shallow convection and moist turbulence schemes and their impact on climate simulations with the community atmosphere model. *Journal of Climate*, *22*(12), 3449–3469.
Pendergrass, A. G., Reed, K. A., & Medeiros, B. (2016). The link between extreme precipitation and convective organization in a warming climate: Global radiative-convective equilibrium simulations. *Geophysical Research Letters*, *43*(21), 11–445.
Petch, J., Hill, A., Davies, L., Fridlind, A., Jakob, C., Lin, Y., ... Zhu, P. (2014). Evaluation of intercomparisons of four different types of model simulating twp-ice. *Quarterly Journal of the Royal Meteorological Society*, *140*(680), 826–837.
Petch, J., Willett, M., Wong, R., & Woolnough, S. (2007). Modelling suppressed and active convection. comparing a numerical weather prediction, cloud-resolving and single-column model. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *133*(626), 1087–1100.
Ramanathan, V., & Coakley, J. A. (1978). Climate modeling through radiative-convective models. *Reviews of geophysics*, *16*(4), 465–489.
Randall, D., & Cripe, D. G. (1999). Alternative methods for specification of observed forcing in single-column models and cloud system models. *Journal of Geophysical Research: Atmospheres*, *104*(D20), 24527–24545.
Randall, D., Xu, K.-M., Somerville, R. J., & Iacobellis, S. (1996). Single-column models and cloud ensemble models as links between observations and climate models. *Journal of Climate*, *9*(8), 1683–1697.
Reed, K. A., Medeiros, B., Bacmeister, J. T., & Lauritzen, P. H. (2015). Global radiative–convective equilibrium in the community atmosphere model, version 5. *Journal of the Atmospheric Sciences*, *72*(5), 2183–2197.
Ringer, M. A., Andrews, T., & Webb, M. J. (2014). Global-mean radiative feedbacks and

forcing in atmosphere-only and coupled atmosphere-ocean climate change experiments. *Geophysical Research Letters*, *41*(11), 4035–4042.

Rotunno, R., Klemp, J. B., & Weisman, M. L. (1988). A theory for strong, long-lived squall lines. *Journal of Atmospheric Sciences*, *45*(3), 463–485.

Shamekh, S., Muller, C., Duvel, J.-P., & d'Andrea, F. (2020). How do ocean warm anomalies favor the aggregation of deep convective clouds? *Journal of the Atmospheric Sciences*, *77*(11), 3733–3745.

Sherwood, S. C., Bony, S., Boucher, O., Bretherton, C., Forster, P. M., Gregory, J. M., & Stevens, B. (2015). Adjustments in the forcing-feedback framework for understanding climate change. *Bulletin of the American Meteorological Society*, *96*(2), 217–228.

Sherwood, S. C., Bony, S., & Dufresne, J.-L. (2014). Spread in model climate sensitivity traced to atmospheric convective mixing. *Nature*, *505*(7481), 37–42.

Sherwood, S. C., & Meyer, C. (2006). The general circulation and robust relative humidity. *Journal of climate*, *19*(24), 6278–6290.

Sobel, A. H., & Bretherton, C. S. (2000). Modeling tropical precipitation in a single column. *Journal of climate*, *13*(24), 4378–4392.

Tompkins, A. (2000). The impact of dimensionality on long-term cloud-resolving model simulations. *Monthly Weather Review*, *128*(5), 1521–1535.

Tompkins, A. (2001). Organization of tropical convection in low vertical wind shears: The role of water vapor. *Journal of the atmospheric sciences*, *58*(6), 529–545.

Vial, J., Dufresne, J.-L., & Bony, S. (2013). On the interpretation of inter-model spread in cmip5 climate sensitivity estimates. *Climate Dynamics*, *41*(11-12), 3339–3362.

Webb, M. J., Lambert, F. H., & Gregory, J. M. (2013). Origins of differences in climate sensitivity, forcing and feedback in climate models. *Climate Dynamics*, *40*(3), 677–707.

Wing, A. A. (2019). Self-aggregation of deep convection and its implications for climate. *Current climate change reports*, *5*(1), 1–11.

Wing, A. A., & Cronin, T. W. (2016). Self-aggregation of convection in long channel geometry. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 1–15.

Wing, A. A., Emanuel, K., Holloway, C. E., & Muller, C. (2017). Convective self-aggregation in numerical simulations: A review. *Shallow clouds, water vapor, circulation, and climate sensitivity*, 1–25.

Wing, A. A., & Emanuel, K. A. (2014). Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations. *Journal of Advances in Modeling Earth Systems*, *6*(1), 59–74.

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.

Wing, A. A., Stauffer, C. L., Becker, T., Reed, K. A., Ahn, M.-S., Arnold, N. P., ... others (2020). Clouds and convective self-aggregation in a multimodel ensemble of radiative-convective equilibrium simulations. *Journal of advances in modeling earth systems*, *12*(9), e2020MS002138.

Zelinka, M. D., Klein, S. A., Taylor, K. E., Andrews, T., Webb, M. J., Gregory, J. M., & Forster, P. M. (2013). Contributions of different cloud types to feedbacks and rapid adjustments in cmip5. *Journal of Climate*, *26*(14), 5007–5027.

Zhang, C., & Wang, Y. (2017). Projected future changes of tropical cyclone activity over the western north and south pacific in a 20-km-mesh regional climate model. *Journal of Climate*, *30*(15), 5923–5941.

Zhang, G., & McFarlane, N. A. (1995). Sensitivity of climate simulations to the parameterization of cumulus convection in the canadian climate centre general circulation model. *Atmosphere-ocean*, *33*(3), 407–446.

Zhang, M., Bretherton, C. S., Blossey, P. N., Austin, P. H., Bacmeister, J. T., Bony, S., ... others (2013). Cgils: Results from the first phase of an international project to understand the physical mechanisms of low cloud feedbacks in single column models. *Journal of Advances in Modeling Earth Systems*, *5*(4), 826–842.

Zhang, M., Somerville, R. C., & Xie, S. (2016). The scm concept and creation of arm forcing datasets. *Meteorological Monographs*, *57*, 24–1.

Zhu, H., & Sobel, A. H. (2012). Comparison of a single-column model in weak temperature gradient mode to its parent agcm. *Quarterly Journal of the Royal Meteorological Society*, *138*(665), 1025–1034.

Zhu, P., Bretherton, C. S., Köhler, M., Cheng, A., Chlond, A., Geng, Q., ... others (2005). Intercomparison and interpretation of single-column model simulations of a nocturnal stratocumulus-topped marine boundary layer. *Monthly weather review*, *133*(9), 2741–2758.