Characterizing Convection Schemes Using Their Responses to Imposed Tendency Perturbations

Yi-Ling Hwong[1], Siwon Song[2], Steven Sherwood[3], Alison Stirling[4], Catherine Rio[5], Romain Roehrig[6], Chimene Laure Daleu[7], Robert Stephen Plant[8], David Fuchs[3], Penelope Maher[9], and Ludovic Touzé-Peiffer[10]

[1]Climate Change Research Centre, University of New South Wales
[2]Center for Climate/Environment Change Prediction Research, Ewha Womans University
[3]University of New South Wales
[4]UK Met Office
[5]Centre national des recherches météorologiques (CNRM), Université de Toulouse, Météo-France, CNRS
[6]CNRM, Université de Toulouse, Météo-France, CNRS
[7]Reading University, UK
[8]University of Reading
[9]University of Exeter
[10]Laboratoire de Météorologie Dynamique, Sorbonne Université, CNRS

November 24, 2022

Abstract

Convection is usually parameterized in global climate models, and there are often large discrepancies between results obtained with different convection schemes. Conventional methods of comparing convection schemes using observational cases or directly in 3D models do not always clearly identify parameterization strengths and weaknesses. In this paper we evaluate the response of parameterizations to various perturbations rather than their behavior under particular strong forcing. We use the linear response function method proposed by Kuang (2010) to compare twelve physical packages in five atmospheric models using single-column model (SCM) simulations under idealized radiative-convective equilibrium conditions. The models are forced with anomalous temperature and moisture tendencies. The temperature and moisture departures from equilibrium are compared with published results from a cloud-resolving model (CRM). Results show that the procedure is capable of isolating the behavior of a convection scheme from other physics schemes. We identify areas of agreement but also substantial differences between convection schemes, some of which can be related to scheme design. Some aspects of the model linear responses are related to their RCE profiles (the relative humidity profile in particular), while others constitute independent diagnostics. All the SCMs show irregularities or discontinuities in behavior that are likely related to switches or thresholds built into the convection schemes, and which do not appear in the CRM. Our results highlight potential flaws in convection schemes and suggest possible new directions to explore for parameterization evaluation.

Characterizing Convection Schemes Using Their Responses to Imposed Tendency Perturbations

Y. L. Hwong[1], S. Song[1,2], S. C. Sherwood[1], A. J. Stirling[3], C. Rio[4], R. Roehrig[4], C. L. Daleu[5], R. S. Plant[5], D. Fuchs[1], P. Maher[6], L. Touzé-Peiffer[7]

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

[2] Center for Climate/Environment Change Prediction Research, Ewha Womans University, Seoul, South Korea

[3] Met Office, Exeter, UK

[4] CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France

[5] Department of Meteorology, University of Reading, Reading, UK

[6] Department of Mathematics, University of Exeter, Exeter, UK

[7] Laboratoire de Météorologie Dynamique, Sorbonne Université, CNRS, Paris, France

Corresponding author: Yi-Ling Hwong (yiling.hwong@gmail.com)

Key Points:

- The linear response function method is applied in SCM simulations and is able to isolate the behavior of convection schemes.
- Linear responses of the models are related to the mean state relative humidity in both shape and magnitude.
- All SCMs show discontinuities in their responses which are likely related to switches or threshold built into convective parameterization.

Abstract

Convection is usually parameterized in global climate models, and there are often large discrepancies between results obtained with different convection schemes. Conventional methods of comparing convection schemes using observational cases or directly in 3D models do not always clearly identify parameterization strengths and weaknesses. In this paper we evaluate the response of parameterizations to various perturbations rather than their behavior under particular strong forcing. We use the linear response function method proposed by Kuang (2010) to compare twelve physical packages in five atmospheric models using single-column model (SCM) simulations under idealized radiative-convective equilibrium conditions. The models are forced with anomalous temperature and moisture tendencies. The temperature and moisture departures from equilibrium are compared with published results from a cloud-resolving model (CRM). Results show that the procedure is capable of isolating the behavior of a convection scheme from other physics schemes. We identify areas of agreement but also substantial differences between convection schemes, some of which can be related to scheme design. Some aspects of the model linear responses are related to their RCE profiles (the relative humidity profile in particular), while others constitute independent diagnostics. All the SCMs show irregularities or discontinuities in behavior that are likely related to switches or thresholds built into the convection schemes, and which do not appear in the CRM. Our results highlight potential flaws in convection schemes and suggest possible new directions to explore for parameterization evaluation.

Plain Language Summary

The transport of heat up and down the atmosphere, called atmospheric convection, is a complex process. To simplify the representation of convection in global climate models (GCMs) scientists use “parameterization”, which is essentially mathematical equations of physical processes. However, there are many different ways to formulate these equations, and no agreement on which is better. In this work we aim to understand a few popular ways to parameterize convection. We extract one vertical column from five different GCMs and lightly tickle (perturb) it and then observe its responses. We found that different models respond very differently to the same tickling, and this tells us a lot about the model. Importantly, the specific perturbation that we used can single out the responses of convection-related equations from equations of other processes. All the models in our study have one thing in common: they are quite jumpy when tickled, especially at the top of the boundary layer where clouds start to form. We suspect the culprits are switches placed in the models that sometimes lead to sudden changes in their response. Our work highlights potentially problematic behavior that can give clues on how to make climate models better.

1 Introduction

Atmospheric deep convection is an important process that is still imperfectly understood. It generates most of the observed precipitation and is the main source of heating to balance radiative cooling. Global climate models (GCMs) usually have a horizontal resolution that is much bigger than individual convective clouds. This makes the representation of convection in

GCMs particularly challenging as it cannot be explicitly resolved. The collective effect of subgrid-scale convection on the resolved flow is expressed through parameterizations, which are approximate equations to capture the essence of unresolved processes in a realistic way. Arakawa (2004) defines convective parameterization as "an attempt to formulate the statistical effects of cumulus convection without predicting each individual cloud". Convection parameterizations typically simulate subgrid-scale precipitation and adjust the vertical distribution of heat, moisture, and momentum (Kain & Fritsch, 1990). Most convection schemes used in GCMs today are mass-flux based and updated from schemes developed in the 1980s and 1990s (Rio et al., 2019). More recently, new approaches to parameterize convection have been proposed, for example with the introduction of stochastic elements (e.g., Berner et al., 2017; Grell & Freitas, 2014) and new processes such as cold pools (e.g., Del Genio et al., 2015; Grandpeix & Lafore, 2010; Rio et al., 2013). There are also now attempts based on machine learning (e.g., Gentine et al., 2018; O'Gorman & Dwyer, 2018).

The wide array of convection schemes employing different underlying assumptions is one of the major sources of uncertainties in GCMs. For instance, schemes often use different trigger functions and closure assumptions. As Arakawa (2004) points out, there are at least six types of convection schemes based on their closure assumptions alone. Trigger functions can be constructed using various variables such as convective available potential energy (CAPE), vertical velocity at the lifting condensation level (Bechtold et al., 2001; Kain & Fritsch, 1990), cloud work function (Arakawa & Schubert, 1974), and surface temperature and moisture (Tawfik & Dirmeyer, 2014). Certain assumptions that are widely used in convective parameterization have been found to be flawed. The quasi-equilibrium assumption (Arakawa & Schubert 1974; Emanuel et al., 1994), for example, has been recognized to be incomplete in some cases (Bechtold et al., 2014; Davies et al., 2013; Mapes, 1997; Raymond, 1995; Yano & Plant, 2012). Further, convection schemes inherently have adjustable parameters that can be "tuned", in particular to allow simulation results to better match certain observed features of the Earth system such as clouds, temperature, and winds (e.g., Kain & Fritsch, 1990; Mauritsen et al., 2012). All these factors have led to considerable differences in model outputs when different convection schemes were employed (e.g., Emanuel & Živković-Rothman, 1999). Convective parameterization has also been identified as one of the major contributors to the discrepancies in climate sensitivity predictions between GCMs (e.g., Bony & Dufresne, 2005; Boucher et al., 2013; Vial et al., 2013). Studies have attributed the biases in various simulated variables, such as precipitation variability (DeMott et al., 2007; Wang & Zhang, 2013; Zhang & Mu, 2005), clouds (Chepfer et al., 2008; Zhang et al., 2010), convective organization (Bony et al., 2015), and the diurnal cycle of convection (Bechtold et al., 2014; Langhans et al., 2013; Rio et al., 2009), to the parameterization of convection.

Conventional methods of comparing convection schemes typically use observational case studies, where model outputs are compared with a selection of relevant observed properties in the atmosphere (e.g., Grell & Freitas, 2014; Han & Pan, 2011; Kwon & Hong, 2017; Zhang & Wang, 2017; Zhang et al., 2011). However, this method relies on a sometimes difficult derivation of large-scale forcing and is based on a limited selection of observed situations. An alternative approach was suggested by Arakawa (2004), wherein he notes that differences between convection schemes could perhaps be better understood if they were expressed in a common mathematical framework instead of the physical theories they were based on. Along these lines, Kuang (2010, hereafter K10) proposed the linear response function as an assessment method for convective parameterizations based on their behavior, i.e., how they actually react to atmospheric variations. There have been many studies that examined the convective responses of cloud-resolving models (CRMs) as well as convection schemes to perturbation of its large-scale environment (e.g., Derbyshire et al. 2004; Lambert et al., 2020; Redelsperger et al. 2002; Takemi et al. 2004; Tulich & Mapes, 2010).

In this study, we base our approach on K10's method and assess how it can be applied to explore the behavior of convection schemes in a systematic way. K10 points out that the responses of a cumulus ensemble to weak perturbations of its large-scale environment can be quite linear even though cumulus convection involves many non-linear processes. The behavior of a cumulus ensemble (i.e., its variation around a reference state) can therefore be approximated with a linear response function (or linear response matrix), M, which can be used to probe the mean response of a non-linear system to small imposed perturbations. The anomalous convective tendencies are given as

$$\frac{d\mathbf{x}}{dt} = \mathsf{M}\mathbf{x} \tag{1}$$

where $\mathbf{x}$ is the anomalous state vector, i.e., vertical profiles of anomalous temperature **T'** and moisture **q'** corresponding to the vector of the anomalous temperature or moisture tendency ($d\mathbf{T'}/dt$ or $d\mathbf{q'}/dt$). Prime indicates departure from the equilibrium state of the control (unperturbed) run and bold characters denote column vectors, e.g., $\mathbf{q'} = q'(k)$, where k is the vertical levels. In K10's experiments, small perturbations are applied to the tendencies of the thermodynamic variables, and maintained until the system reaches a new equilibrium. The anomalous convective tendencies ($d\mathbf{x}/dt$) in this new equilibrium state then balance the additional perturbed forcing applied. The deviation of the temperature and moisture profiles from their profile in the control unperturbed run is $\mathbf{x}$. To construct the matrix M, perturbations are applied to the temperature and moisture tendencies separately, using similarly shaped profiles that peak at successive models levels. The resulting vectors of $d\mathbf{x}/dt$ and $\mathbf{x}$ are stacked together so that

$$\mathsf{Y} = \mathsf{M}\mathsf{X} \tag{2}$$

In this matrix formulation, each column of Y represents a profile of the prescribed tendency perturbation that peaks at a given model level $(dT'/dt_{sfc}, \ldots\ dT'/dt_{top}, dq'/dt_{sfc}, \ldots, dq'/dt_{top})^{\mathrm{T}}$, where the subscripts *sfc* and *top* denote the lowest and highest model levels, and the corresponding column of X is the corresponding state responses $(T'_{sfc}, \ldots, T'_{top}, q'_{sfc}, \ldots, q'_{top})^{\mathrm{T}}$.

Our study focuses on the temperature and moisture responses to small perturbations of convective tendencies using single-column model (SCM) simulations, following Herman and Kuang (2013, hereafter HK13). To be precise, we present the "response per unit perturbation" of the models, i.e., the M^{-1} matrix (see Appendix A of HK13). The overarching goal is to characterize and compare some widely used convection schemes using K10's linear response function method. Further efforts to investigate the underlying mechanisms and assumptions of the individual schemes that may explain their behavior presented here form part of our ongoing work and will be presented in future publications. Twelve physical packages in five SCMs are tested. We also compare our results with the corresponding CRM (SAM6.8.2, 2 km resolution) results of K10. The focus on the steady state responses (M^{-1} matrix) of the SCMs in this paper allows us to easily recognize salient features of the schemes and locate discrepancies between them to gain insights into their behavior.

The mean state used in this study is that of a radiative-convective equilibrium (RCE), in which the climate system is represented by a balance between radiative cooling and convective heating. RCE resembles the tropical atmosphere on a large scale, where there is no vertical motion on average (Manabe & Strickler, 1964). It is the simplest framework to describe the atmosphere and has been applied to study a myriad of climate phenomena such as convective self-aggregation (Wing et al., 2020), precipitation extremes (Pendergrass et al., 2016), and convective updraught velocities (Singh & O'Gorman, 2015). Besides comparing between convection schemes, we also compare simulations with different planetary boundary layer (PBL) and microphysics (MP) schemes.

The specific objectives of this paper are: (1) to compare the RCE mean states of the different SCMs, (2) to examine and compare the steady state responses (**T'** and **q'**) of the different schemes to small convective tendency perturbations, and (3) to test the sensitivity of the RCE mean state and the responses to the types of parameterization typically used in global models (convection, PBL, and MP).

2 Methods

2.1 Participating models and simulation setup

The participating SCMs and their model physics are listed in Table 1. Further details on the convection schemes of the SCMs are presented in Table 2. For the Weather Research and Forecasting (WRF) model, five convection schemes are tested; for the Unified Model (UM), two convection schemes are tested; for the LMDZ model, three physical packages for convection and clouds are tested. This brings the total number of SCM cases to 12 (for brevity hereafter we will refer to these cases simply as "SCMs"). The Zhang-McFarlane deep convection in combination with the University of Washington (UW) shallow convection schemes are used in two SCMs – WRF and SCAM (the SCM version of the Community Atmosphere Model, CAM). In both cases the same PBL and MP schemes are also used so their model physics are matched as closely as possible. Two variations of the Betts-Miller convection scheme are tested: the Simplified Betts-

Miller (SBM) scheme in UM and the Betts-Miller-Janjic (BMJ) scheme in WRF. These cases make for interesting comparisons of the same (or similar) scheme in two different models.

Table 1. SCMs and their model physics

SCM cases and versions		**Convection scheme**	**PBL scheme**	**Microphysics / large-scale scheme**	**Other schemes**
LMDZ	5A	Emanuel scheme (Emanuel, 1993)	Eddy diffusion (Laval et al., 1981) with counter-gradient term (Deardorff, 1972)	Sundqvist (1978) for liquid water, Zender and Kiehl (1997), Heymsfield and Donner (1990) for ice	Log-normal cloud scheme of Bony and Emanuel (2001)
	6A	Modified Emanuel scheme (Grandpeix et al., 2004) + cold pool parameterization (Grandpeix & Lafore, 2010; Rio et al., 2013)	Pronostic eddy diffusion (Yamada, 1983) + mass-flux representation of thermals (Rio et al., 2010)	Same as above	Bi-gaussian cloud scheme of Jam et al. (2013) for cumulus clouds , log-normal cloud scheme of Bony and Emanuel (2001) for deep and LS clouds
	6Ab	Same as above	Same as above	Same as above + Jakob and Klein (2000) for the evaporation of precipitation	Same as above
SCAM (CAM, v.5.3)		Zhang-McFarlane deep convection (Zhang & McFarlane, 1995) + UW shallow convection scheme (Park & Bretherton, 2009)	UW Moist Turbulence scheme (Park & Bretherton, 2009)	Stratiform microphysical processes (Morrison & Gettelman, 2008)	Cloud macrophysics scheme (Park et al., 2014)
WRF (v. 4.0.2)	ZM	Zhang-McFarlane (Zhang & McFarlane, 1995) + UW shallow convection scheme (Park & Bretherton, 2009)	UW Moist Turbulence scheme (Park & Bretherton, 2009)	Stratiform microphysical processes (Morrison & Gettelman, 2008)	
	KF	Kain-Fritsch (Kain, 2004)			

	NT	New-Tiedtke (Zhang & Wang, 2017)	Yonsei University (Hong et al., 2006)	WRF Single-Moment 6-class (Hong & Lim, 2006)	
	NSAS	New Simplified Arakawa-Schubert (Han & Pan, 2011)			
	BMJ	Betts-Miller-Janjic (Betts, 1986; Betts & Miller, 1986; Janjic, 1994, 2000)			
UM (v.11.6)	SBM	Simplified Betts-Miller (Frierson, 2007)	Lock et al. (2000)	Single-moment scheme based on Wilson and Ballard (1999)	PC2 cloud scheme (Wilson et al., 2008)
	MF	UM 6A Mass-Flux scheme (Walters et al., 2019)			
CNRM (ARPEGE-Climat v.6.4.1)		Prognostic Condensates and Microphysics Transport (PCMT; Guérémy, 2011; Piriou et al., 2007; Roehrig et al., 2020)	Prognostic eddy-diffusion (Cuxart et al., 2000) and dry and shallow convection with PCMT	Single-moment, 5-class (Lopez, 2002)	Cloud macrophysics (Bougeault, 1981; Ricard & Royer, 1993)

Table 2. Convection schemes and their main features

SCM	**Convection scheme**	**Brief description**	**Closure assumption**	**Triggering**	**Entrainment / detrainment**	**Shallow convection (Yes / No)**	**Interaction with large-scale (LS) cloud scheme (Yes / No)**
LMDZ5A	Emanuel scheme	Episodic mixing and buoyancy sorting mass-flux scheme. Representation of an unsaturated downdraft.	CAPE-based	CIN-based	Episodic mixing and buoyancy sorting.	No	No

LMDZ6A	Modified Emanuel scheme + cold pool parameterization for deep convection	Episodic mixing of Grandpeix et al. (2004) and coupling with cold pools: saturated updrafts and downdrafts happen in the environment of cold pools while the unsaturated downdraft falls into the cold pool region. EDMF type scheme for shallow convection.	Available Lifting Power (ALP) at cloud base provided by boundary layer thermals and cold pools.	When Available Lifting Energy (ALE) provided either by thermals or cold pools exceeds convective inhibition (ALE > CIN).	Episodic mixing as described by Grandpeix et al. (2004) and detrainment of mixtures at level of neutral buoyancy.	Yes. Thermal plume model that represents dry and shallow convection in a unified way.	Yes. Bi-gaussian cloud scheme based on thermal properties to compute cloud fraction and precipitation in the LS scheme.
LMDZ6Ab	Same as 6A but with the parameterization of Jakob and Klein (2000) to account for cloud overlap in evaporation of precipitation.	Same as above	Same as above	Same as above	Same as above	Same as above	Same as above with the parameterization of Jakob and Klein (2000) to account for cloud overlap in evaporation of precipitation.
SCAM	Zhang-McFarlane (ZM) deep convection + UW shallow convection	ZM is a mass-flux scheme and only considers deep convection. Specifies distribution of updrafts assuming all categories in the plume spectrum have same cloud base mass-flux. Assumes convection removes CAPE at exponential rate with specified	ZM: CAPE-based UW: CIN-based mass-flux closure	ZM: Threshold value for CAPE exceeded for air parcel lifted from level of highest moist static energy UW: CIN-based trigger	ZM: Each updraft has characteristic entrainment rate (ER). Detrains cloud liquid and ice.	No, when ZM scheme is used alone. Yes, when combined with the UW shallow scheme.	Yes. Convection scheme detrains cloud and ice at cloud top, which are then used by the MP scheme.

		adjustment time scale. UW is a mass-flux shallow convection scheme. Includes momentum mixing. Entrainment depends on vertical velocity of updraft.					
WRF	ZM + UW	Same as above	Same as above	Same as above	Same as above	Same as above	Yes, same as above
	Kain-Fritsch (KF)	A simple cloud model with moist updrafts and downdrafts. Includes simple microphysics. Perturbation temperature based on horizontal and vertical moisture convergence. Minimum cloud depth varies according to cloud base temperature. Updated downdraft formulation from original KF scheme.	CAPE-based	Parcel vertical velocity (which has dependence on LS *w*) is positive over a specified cloud depth (typically 3 km).	Minimum ER imposed and variable ER based on sub-cloud layer convergence.	Yes	Yes, same as above
	New-Tiedtke (NT)	Updated version of the Tiedtke mass-flux based scheme. Updates include trigger functions and closure for deep and shallow convection.	CAPE-based	Net moisture convergence is positive and unstable parcels present in air in lower layers.	Entrainment and detrainment depends on environmental relative humidity (RH).	Yes	Yes, same as above

		Convective adjustment time depends on vertical velocity averaged in updraft and cloud depth.					
	New Simplified Arakawa-Schubert (NSAS)	Updated from Simplified AS scheme that uses only one type of cloud (deepest) instead of an ensemble. Shallow and deep convection uses a mass-flux scheme. Increased threshold for mass-flux at cloud base and remove random cloud-top selection to enhance deep convection.	Based on cloud work function quasi-equilibrium.	Based on threshold for cloud work function, also has some dependency on LS *w* and low-level moisture.	Entrainment and detrainment depend on environmental RH.	Yes	Yes, same as above
	Betts-Miller-Janjic (BMJ)	Based on Betts-Miller convective adjustment scheme. Parameters for target moisture profile and relaxation time are variable and depend on cloud efficiency. Moisture profile for shallow convection requires entropy change to be small and non-negative.	CAPE-based	Three conditions: CAPE available, threshold for cloud depth exceeded, moist soundings.	N/A	Yes	No
UM	Simplified Betts-Miller (SBM)	A simplified version of the Betts-Miller	CAPE-based	Positive CAPE trigger	N/A	Yes	No

		scheme. Profiles of temperature and moisture are relaxed to a fixed RH (typically 70%) over a given relaxation time.					
	6A Mass-Flux (MF)	Based on Gregory and Rowntree (1990). Single bulk plume mass-flux scheme with diagnosis for shallow, deep states depending on the depth reached by an undilute parcel ascent. Also a mid-level scheme for initiation above the boundary layer top.	CAPE-based, with a variable timescale for CAPE removal dependent on resolved ascent (although no ascent assumed in these SCM experiments).	Convection triggered from top of boundary layer when a dilute parcel ascent exceeds a threshold buoyancy at the next level (currently set to 0.2 K).	Fixed entrainment profile, a mixing detrainment profile that depends on the entrainment, and an adaptive detrainment term that acts to increase parcel buoyancy once it starts to decline to reflect detrainment of less buoyant plumes from the 'bulk' population. (Derbyshire et al., 2011).	Yes	No initiation from LS cloud, but detrained liquid and frozen condensate from convective plume provides increments to LS cloud condensates and fractions (Wilson & Ballard, 1999).
CNRM	PCMT	Main concepts based on buoyancy (triggering, mass-flux, entrainment-detrainment). Condensates and convective vertical velocity are prognostic, providing memory effect.	CAPE-based	Buoyancy-based. Triggered when convective updraft vertical velocity is positive.	Buoyancy sorting (Bretherton et al., 2004)	Yes (continuous approach)	Yes. Convective MP consistently mirrors the LS MP. LS condensates are entrained in the convective updraft, while convective condensates are detrained in the environment.

Following HK13, we first perform an RCE simulation (referred to below as PreRCE) for each model to find its steady state (radiative cooling equals convective heating), then use this state to initiate the control and perturbation experiments. For the control and perturbation runs, we replace the interactive radiative scheme in all models with an idealized constant radiative cooling profile of Q_{rad} = -1.5 K day^{-1} from the surface to 200 hPa; from there, decreasing linearly to zero at and above 100 hPa. A temperature and moisture relaxation to the models' respective PreRCE profiles is imposed near and above the tropopause. The inverse relaxation time constant is zero from surface to approximately 160 hPa and then increases linearly to 0.5 day^{-1} at and above the 100 hPa level (see Figure 1 in HK13). This adjustment serves to prevent unrealistic temperature and moisture values due to weak convective activity in these regions (HK13). Note that for the PreRCE run we leave the handling of stratospheric temperature and moisture profiles to the judgement of each modeller. Tests using the five WRF cases reveal that M^{-1} is not sensitive to this part of the profile (not shown).

The sea surface temperature (SST) used in all models is 28°C. Surface sensible and latent heat fluxes (*SH* and *LH*, respectively) are computed using a bulk aerodynamic formula:

$$SH = \rho_1 c_p C_h U \left[T_s - \frac{\pi_s}{\pi_1} T_1\right] \tag{3}$$

$$LH = \rho_1 L_v C_e U[q_{sat}(T_s, p_s) - q_1] \tag{4}$$

where ρ, p, T, q and $\boldsymbol{\pi}$ are, respectively, the air density, pressure, temperature, specific humidity and the Exner function, with their subscripts s and 1 referring to surface and lowest model level, respectively; C_h and C_e are the surface exchange coefficients for heat and moisture, respectively; U is the near surface wind speed; $q_{sat}(T_s, p_s)$ is the saturation specific humidity at surface temperature and pressure; c_p is the heat capacity of dry air and L_v is the latent heat of water vaporization. We used a fixed value of 0.001 for the exchange coefficients C_h and C_e and constant of 4.8 m s^{-1} for the near surface wind U. This removes any surface exchange feedback caused by winds. The horizontal mean wind speeds are relaxed to a vertically uniform value of 4.8 m s^{-1} for zonal and 0 m s^{-1} for meridional wind, with a relaxation time constant of 3 h.

Our approach assumes that closely examining model behavior under RCE conditions (w = 0) will be helpful for characterizing model physics behavior. However, a few convection schemes in WRF—specifically, the Kain-Fritsch and New Simplified Arakawa-Schubert schemes—use mechanisms that involve the large-scale vertical velocity in their convection triggering functions (see Table 2), even though this is arguably unphysical (Emanuel et al., 1994). Our experimental setup is possibly not well suited to such schemes, since they require non-zero vertical velocity (i.e., a departure from local RCE) to behave properly. The WRF SCM, however, does have small fluctuating w values in its individual grids due to the 3 x 3 horizontal grid stencil that it uses (described in Hacker & Angevine, 2013), which are sufficient to trigger convection in those schemes. Although the w values remain small (~ 0.1 cm s^{-1} in individual grids, almost zero averaged over all grids) compared to those in nature, we believe that this is a reasonable test of any scheme since the average condition of the atmosphere on a large scale is close to RCE (i.e., no large scale w).

2.2 Perturbation experiment

We apply the method described in HK13 to get the T and q responses to small perturbation of convective tendencies ("inverse technique"). The procedure is briefly described here. We first use the PreRCE state to initiate a control run with no perturbations. For the perturbation runs, we initiate the same way but force the models with small, steady perturbations, separately, in temperature (*dT/dt*) and moisture (*dq/dt*) tendencies at every time step until a new RCE is reached. The applied perturbation follows that of HK13 and is the sum of a delta and Gaussian function. The form of the perturbation applied at the j-th model level is as follows:

$$f_j(p_i) = \frac{1}{2}\left\{\delta_{ij} \,+\, exp\left[-\left(\frac{p_j - p_i}{75\ hPa}\right)^2\right]\right\} \tag{5}$$

where p_i is the local pressure, p_j is the pressure at model level j, and δ_{ij} is a delta function at the j-th model level. The amplitudes of the perturbations are 0.5 K day^{-1} for temperature tendency perturbations and 0.2 g kg^{-1} day^{-1} for moisture tendency perturbations. The profile of a perturbation that peaks at a given model level is hence the respective amplitude multiplied by the function in Equation 5 (see Figures 2a, 3a). For brevity, in this paper we refer to a perturbation profile that peaks at pressure level p as "perturbation at pressure level p". For instance, "perturbation at 850 hPa" denotes a perturbation profile where the magnitude of the perturbation peaks at 850 hPa.

Positive and negative perturbations are applied at every model level in separate runs. The anomalous state response vectors **T'** and **q'** are then the differences of the time-averaged T and q profiles between the perturbation and control runs. We ensure that the simulation lengths and averaging windows used in the models are long enough to attain sufficient signal-to-noise ratio (see Table 3). The anomalies of the positive and negative perturbation runs are averaged to obtain the best-estimate T and q responses presented in this paper; they can also be compared to assess linearity. Note that linearity is assessed following formula B1 in HK13:

$$D_j'(z) = x'_{j+}(z) + x'_{j-}(z) \tag{6}$$

where $D'_j(z)$is the discrepancy for perturbation applied at the j-th model level, $x'_j(z)$is the T or q anomaly for the perturbation at that level, with the +/– subscript denoting positive or negative perturbation, respectively. $D = 0$ indicates perfect linearity. Detailed investigation into linearity is beyond the scope of this study. We merely ensure that the linearity of our models is satisfactory and comparable to that of the SCMs in HK13 (Figure B7 in HK13). For a few models (UM-MF, SCAM, and LMDZ) we reduced the perturbation amplitudes to 0.2 K day^{-1} and 0.1 g kg^{-1} day^{-1} to improve linearity. Additionally for SCAM, results are the average of an ensemble of five members after a series of random noise is added to specific humidity over the whole perturbation period, based on the procedure described in Appendix B4 in HK13 but with a longer period for each random perturbation. This additional step improved the linearity of the system, bringing the linearity of SCAM closer to that of the other SCMs.

Table 3 summarizes the simulation details of the SCMs.

Table 3. Simulation details of the SCMs

SCM	Time step (sec)	Vertical resolution	Perturbation amplitudes (K d^{-1}, g kg^{-1} d^{-1})	Perturbation application period (day)[a]	Time for control and perturbation runs to reach new RCE (day)	Averaging window for mean state and anomaly calculations (day)
LMDZ (x3)	600	79 levels, up to 1.5 hPa	0.2, 0.1	600	100	500
SCAM	300	60 levels, up to 3 hPa	0.2, 0.1	6,500[b]	300	3,000
WRF (x5)	300	74 levels, up to 6 hPa	0.5, 0.2	1,000	300[c]	700
UM-SBM	600	55 levels, up to 48 hPa	0.5, 0.2	500	250	250
UM-MF	600	55 levels, up to 48 hPa	0.2, 0.1	500	250	250
CNRM	900	91 levels, up to 1 hPa	0.5, 0.2	1,000	200	800

[a] After reinitialization from PreRCE state. Models require different simulation lengths to reach new equilibrium, which we leave to the judgement of individual modellers

[b] Longer runs needed for equilibrium to be reached due to random noise application

[c] Varies between convection schemes, but all WRF schemes attain new RCE by around day 300

As mentioned in Section 1, we present the responses in the form of the matrix M^{-1}, which shows the steady state responses per unit perturbation. To construct M^{-1}, we multiply both sides of Equation 2 by Y^{-1} and then again by M^{-1} to get $\mathsf{M}^{-1} = \mathsf{XY}^{-1}$. Y^{-1} is a diagonal matrix where the diagonal elements are the inverses of the total power input for perturbation of a given model level in the units of W m^{-2}. Additionally, we multiply M^{-1} by the standard power inputs of the SAM CRM (noting that the total power input to each model is different owing to the different vertical resolutions) so that the matrices of the SCMs are expressed in the more intuitive units of K or g kg^{-1} (instead of K / [W m^{-2}] or [g kg^{-1}] / [W m^{-2}]) and are directly comparable to that of the CRM.

2.3 Individual scheme sensitivity tests

We anticipate that the SCM behaviors examined here will largely be determined by their convective schemes but this is not guaranteed *a priori*. To test this, in addition to comparing the behavior of the SCMs as configured in Table 1, we also run two separate sets of simulations with alternate PBL and MP schemes. We run this part of the study only in WRF, as it is the only model system that provides multiple options for each parameterization and allows switching between schemes. We run these tests for only two perturbation levels: 850 and 650 hPa. As the radiative profile is prescribed, radiation schemes are not considered here. Four PBL schemes are tested: Yonsei University (YSU), Mellor-Yamada-Nakanishi-Niino level 2.5 (MYNN2) with the eddy-diffusivity mass-flux (EDMF) option enabled, Asymmetrical Convective Model version 2 (ACM2), and Grenier-Bretherton-McCaa (GBM). Four MP schemes are also tested: the WRF Single-Moment 6-class (WSM6), Kessler, Thompson, and Morrison 2-moment schemes. Each of the four convection schemes in WRF (Kain-Fritsch, Betts-Miller-Janjic, New-Tiedtke, New Simplified Arakawa-Schubert) is paired with the four PBL (with default MP) and then four MP (with default PBL) schemes, yielding a total of 32 combinations. The WRF Zhang-McFarlane scheme is excluded from this part of the study as it can only be paired with one PBL scheme. The results of these sensitivity tests are presented in Section 6.

3 RCE mean states

We begin by examining the RCE mean state of the SCMs for temperature and relative humidity (RH), as shown in Figure 1. These are calculated from the temporal averages of the state variables after the models have reached RCE in the control run (see Table 3). For the temperature profiles, saturation equivalent potential temperatures (θ_{es}) are shown instead of temperatures as they are more informative and show the spread better (in a temperature plot the curves are indistinguishable from each other). Note that for a given pressure there is a unique, monotonic relationship between θ_{es} and absolute temperature T. The mean states of the CRM are shown for comparison (Figure 1a, d). The SCMs are generally colder than the CRM, probably due to the warmer SST used in K10's experiment (K10 used an SST of 29.5°C as opposed to 28°C in his SCM experiments in HK13. Sensitivity tests show that using SST = 29.5°C does not change the pattern of the perturbation results by much. For consistency with HK13's SCM experiments we used SST = 28°C). The profiles are all near moist-adiabatic but there are significant departures (Figure 1b, c). In the region of scientific interest to this study (below 160 hPa), a maximum θ_{es} difference of around 25 K (~ 5 K in T) is detected around the surface regions (below 900 hPa) and around 20 K (~ 8 K in T) in the free troposphere (except for UM-SBM). UM-SBM has an outlying RCE temperature profile that is consistently warmer than the other SCMs between the lifting condensation level (LCL) and tropopause. As UM-MF and UM-SBM simulations are identical except for the convection scheme, it is realistic to assume this is not an implementation error. Despite the warm bias in UM-SBM, this SCM is included in this study as the pattern of the perturbation results is the primary interest (we further show in Section 5 that no correlation was found between the mean state temperature and the perturbation results). Nevertheless, this warm bias should be borne in mind in interpreting UM-SBM's results. Apart from UM-SBM, the spread in RCE temperature profiles among the SCMs is consistent with other similar studies (Daleu et al., 2015; Wing et al., 2020). Even among the WRF cases, which use the same experimental setups except for the convection scheme, there is a similar spread throughout the column.

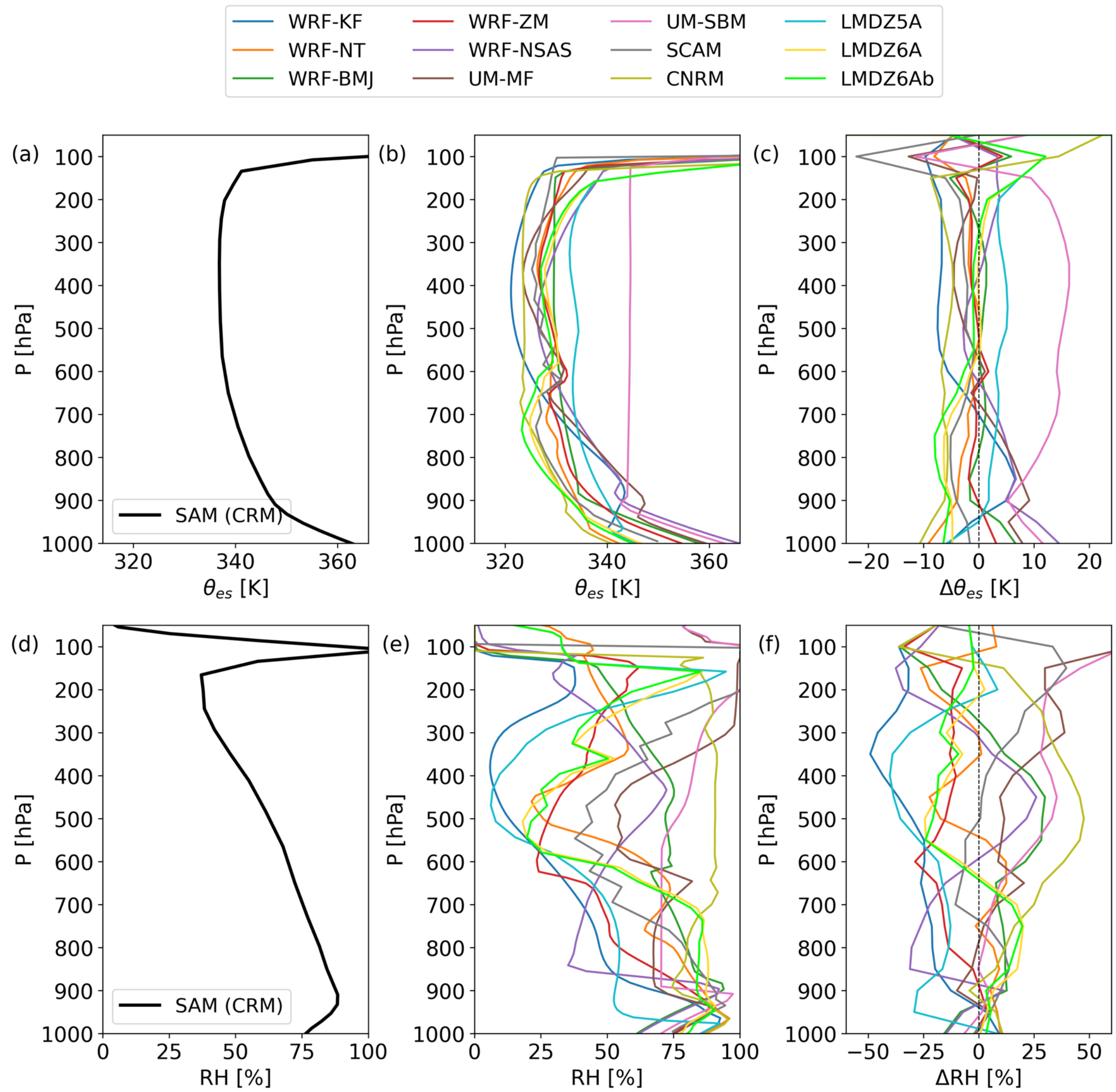

Figure 1. RCE profiles for saturation equivalent potential temperature (a – c) and relative humidity (d – f) of the SAM CRM (a, d) and the SCMs (b, e). The anomalies of the SCMs from their ensemble mean (mean of all SCMs) are shown in c, f.

A large spread is also found in the RCE RH profiles (Figure 1e, f), similar to what HK13 found, and consistent with results of comparable studies (Emanuel & Živković-Rothman, 1999; Rennó et al., 1994; Sobel & Bretherton, 2000; Wing et al., 2020). The RH values of the SCMs range between 56% and 88% at the surface levels and between 6% and 85% in the mid-troposphere. CNRM, UM-MF, UM-SBM, and WRF-BMJ are generally moister than the other SCMs in the free troposphere, while WRF-KF, WRF-ZM, and LMDZ5A are generally drier. Again, the WRF cases diverge considerably in their RH profiles despite identical simulation setups. A kink in the RH profile around the cloud base level (~ 850 – 950 hPa) is detected in the CRM and the SCMs, albeit generally steeper in the SCMs. In a few SCMs these coincide with a slight inversion in their temperature profiles, although this is not always evident. The SCMs in our experiment are generally drier than the CRM, except for CNRM. The RH profiles of the SCMs also frequently display kinks in the free troposphere, which are not found in the CRM, e.g., ~ 600 hPa for UM-MF and WRF-ZM, ~ 700 hPa for WRF-NT. The RCE mean precipitation rates of the SCMs lie between 3.92 – 5.14 mm day^{-1} ($\bar{x}$ = 4.78, σ = 0.38), similar to the SCM values of the RCE Model Intercomparison Project of Wing et al. (2020) and consistent with the expected precipitation rates diagnosed from the prescribed radiative profile.

The two cases involving the Zhang-McFarlane convection scheme (WRF-ZM and SCAM) display similar temperature profiles and comparable shape in their RH profiles, although WRF-ZM is consistently drier than SCAM by around 10 – 20% in the free troposphere. Given that these two SCMs use largely the same model physics (Convection, PBL, and MP schemes), the differences in their mean state could be due to numerics or the way the schemes are implemented. The same applies for the two Betts-Miller cases (WRF-BMJ and UM-SBM), which also display quite different temperature and RH profiles, although in this case the models use different PBL and MP schemes. Additionally, the BMJ and SBM convection schemes—although based on the same concept of relaxation toward a reference profile—differ considerably in their implementation. The two LMDZ6A versions (6A and 6Ab) display almost identical temperature and RH profiles, while the profiles of LMDZ5A differ considerably from those of the LMDZ6A versions.

It is difficult to diagnose the cause of the diverse RCE mean states among the SCMs using only their profiles in Figure 1. In order to investigate this further, we next present in Section 4 the linear responses outlined in Section 2, which convey richer information about the models behavior. We will explore whether the RCE mean states and linear responses are related in Section 5, and investigate the impact of PBL and MP schemes on the RCE mean states in Section 6.

4 Temperature and moisture responses to perturbations

4.1 Key aspects of the SCM responses

In this section we present vertical profiles of the T and q responses (i.e., departure from RCE profiles presented in Section 3) resulting from temperature and moisture tendency perturbations at two particular levels (850 and 650 hPa), for the SAM CRM and four selected SCMs (Figures 2 and 3). The goal is to illustrate a few high-level observations in a more intuitive format before delving into the full results. The complete M^{-1} matrices of all models and a more detailed analysis of their behavior are presented in Section 4.2. Overall, the responses vary greatly among the models. Here, for each variable (T or q response) we show the responses of the SAM CRM from K10, one SCM that closely resembles the CRM (CNRM for T response, WRF-BMJ for q response), and one that differs greatly from it (UM-MF for T response, WRF-NSAS for q response).

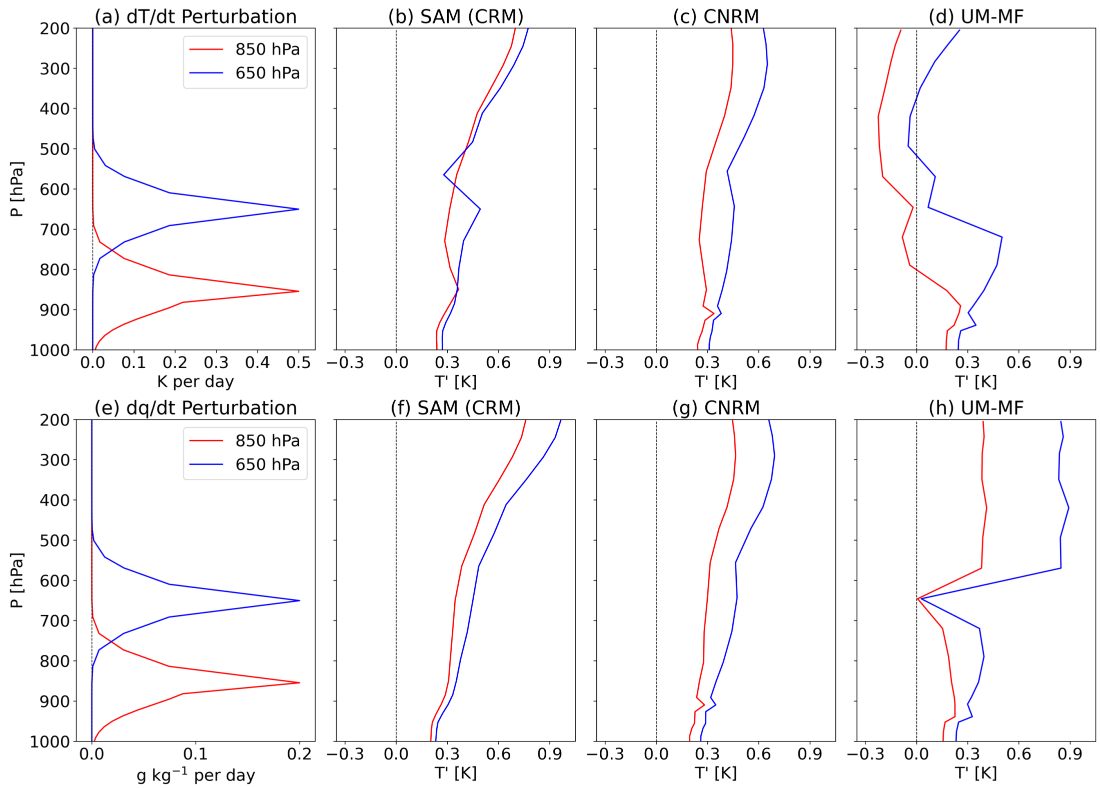

Figure 2. Profiles of the *T* responses to temperature (top) and moisture (bottom) tendency perturbations at 850 (red) and 650 (blue) hPa. The shapes of the perturbations are shown in (a) and (e). Responses of the SAM CRM (b, f), CNRM (c, g) and UM-MF (d, h) are shown here.

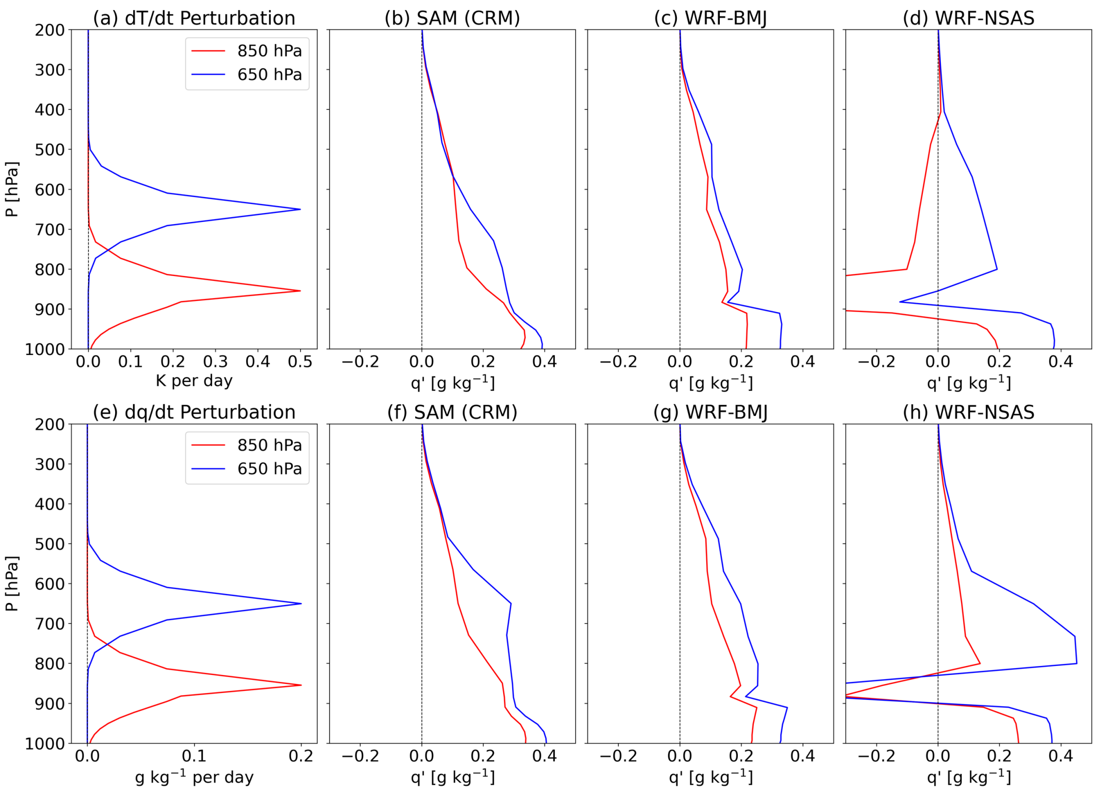

Figure 3. As in Figure 2 but for q responses of the SAM CRM (b, f), WRF-BMJ (c, g) and WRF-NSAS (d, h).

As K10 pointed out, the CRM responds to both heating and moistening perturbations by warming throughout the column, approximating the difference between two moist adiabats (Figures 2b, f). The attendant q responses roughly resemble the expected change in specific humidity computed using the corresponding change in T, if RH remains the same as in the reference state (Figures 3b, f). CNRM and WRF-BMJ largely echo this CRM behavior in their T and q responses, respectively (Figures 2c, g; Figures 3c, g). The observation that WRF-BMJ responds in a similar way to the CRM is perhaps unsurprising, given that the shift in the CRM's response profiles largely conforms to the difference between two moist adiabats. This is the way Betts-Miller type schemes are constructed, where convective activity acts to relax the atmospheric state back to a reference profile, often a moist adiabat (Betts 1986; Betts and Miller, 1986). We elaborate further on the behavior of WRF-BMJ and CNRM in Section 4.2.2.

By contrast, UM-MF and WRF-NSAS exhibit significantly different behavior compared to the CRM. UM-MF shows cool anomalies above the heating levels (Figure 2d). When moistening is applied, its T response drops abruptly to zero around 650 hPa, above which the change in T appears to intensify (Figure 2h). This happens for both perturbation levels. WRF-NSAS shows sharp negative anomalies in its q responses around 850 hPa when heating or moistening is applied (Figure 3d, h), again for both perturbation levels.

Nevertheless, there are a few similarities between the four SCMs and the CRM. Perturbations applied at the higher level (650 hPa) induce stronger responses, likely because convective damping is weaker at higher altitudes, making the convection less able to counter the applied forcing at those levels. A greater change in the equilibrium state is then required to sufficiently alter the convection. All four SCMs display the greatest q responses at the surface levels where the specific humidity itself is largest.

One notable difference between the four SCMs and the CRM is the sharp kinks in SCM responses, commonly around the model-predicted cloud base level (850 – 950 hPa), but also in the mid-troposphere in UM-MF. These kinks often appear to divide the responses into distinctive regions, signalling a level shift in sensitivity. In UM-MF, for example, the T responses either decrease (for heating perturbation, Figure 2d) or increase (for moistening perturbation, Figure 2h) dramatically above the kink around 600 hPa. This characteristic is not observed in the CRM, whose responses are generally smoother and do not appear to have discontinuities, except for a slight kink in its T response when perturbing 650 hPa (Figure 2b), which could be because applied heating produces a small inversion that reduces the T response just above it. The presence of sharp kinks in the SCMs and not the CRM suggests that the kinks probably reflect “switches” or other threshold behavior common in convective parameterizations.

4.2 Matrices of T and q responses

In this section, we present the M^{-1} matrix, which gives a more complete overview of the SCMs’ behavior. For plotting, we divide M^{-1} into four quadrants: T response to heating perturbation (Figure 4), q response to heating (Figure 5), T response to moistening (Figure 6), and q response to moistening (Figure 7). Basically, the quadrants show the T or q response profiles for successive perturbation levels stacked next to each other, with the main diagonal representing the local responses (i.e., responses at pressure level p to perturbation applied at p). The profiles in Figures 2 and 3 comprise two columns of these matrices: the x-axis in these figures is the perturbation level and the y-axis the response level. First, we present the broad features that are largely similar between the models (Section 4.2.1); then, notable differences between the models are presented (Section 4.2.2); finally, we compare the matrices of SCMs with similar or comparable model physics (Section 4.2.3).

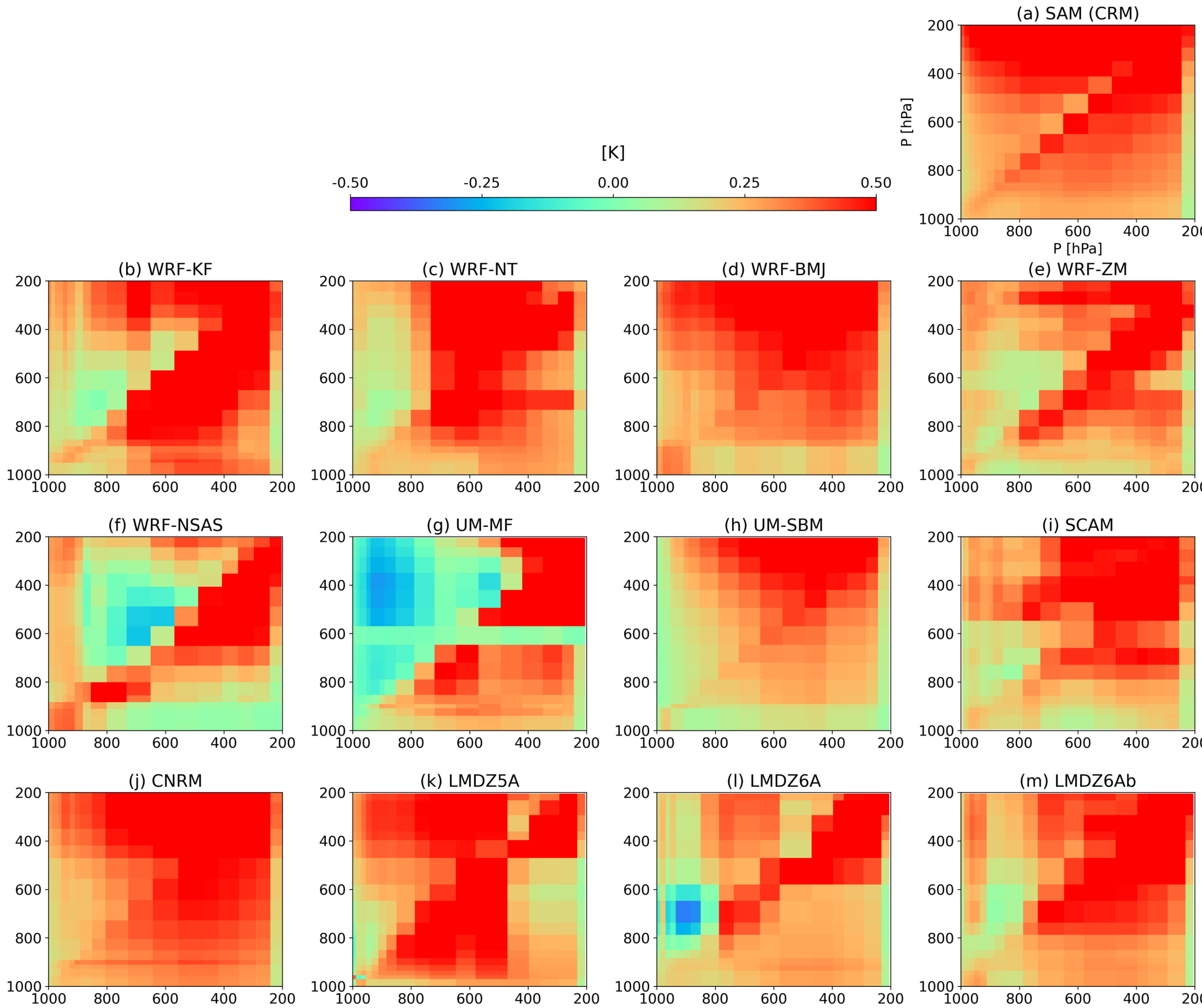

Figure 4. M^{-1} quadrants of *T* responses to temperature tendency perturbation, in the units of K. *x*-axis is perturbation level, *y*-axis is response level.

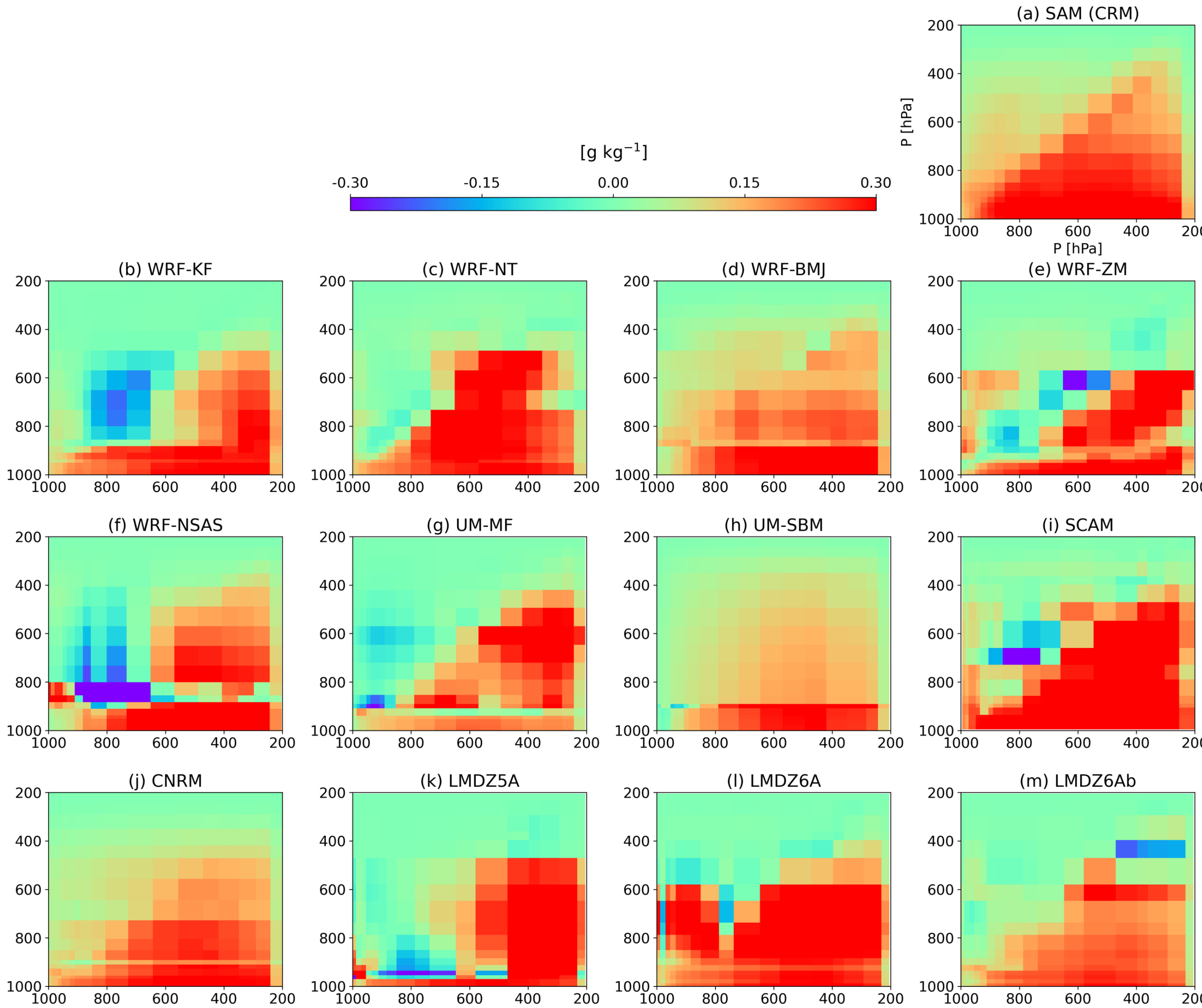

Figure 5. As in Figure 4, but for q responses to temperature tendency perturbation, in the units of g kg^{-1}.

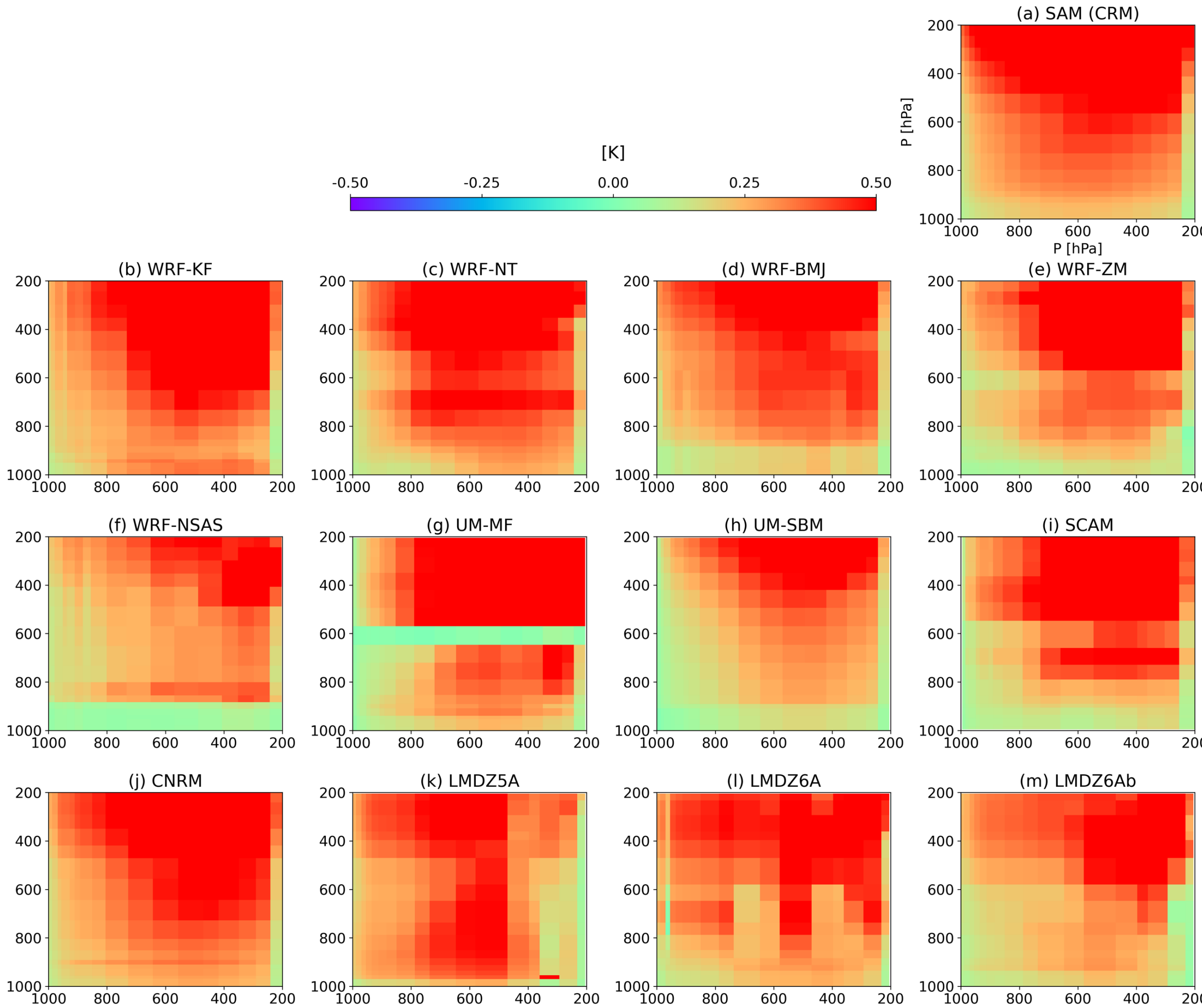

Figure 6. As in Figure 4, but for T responses to moisture tendency perturbation, in the units of K.

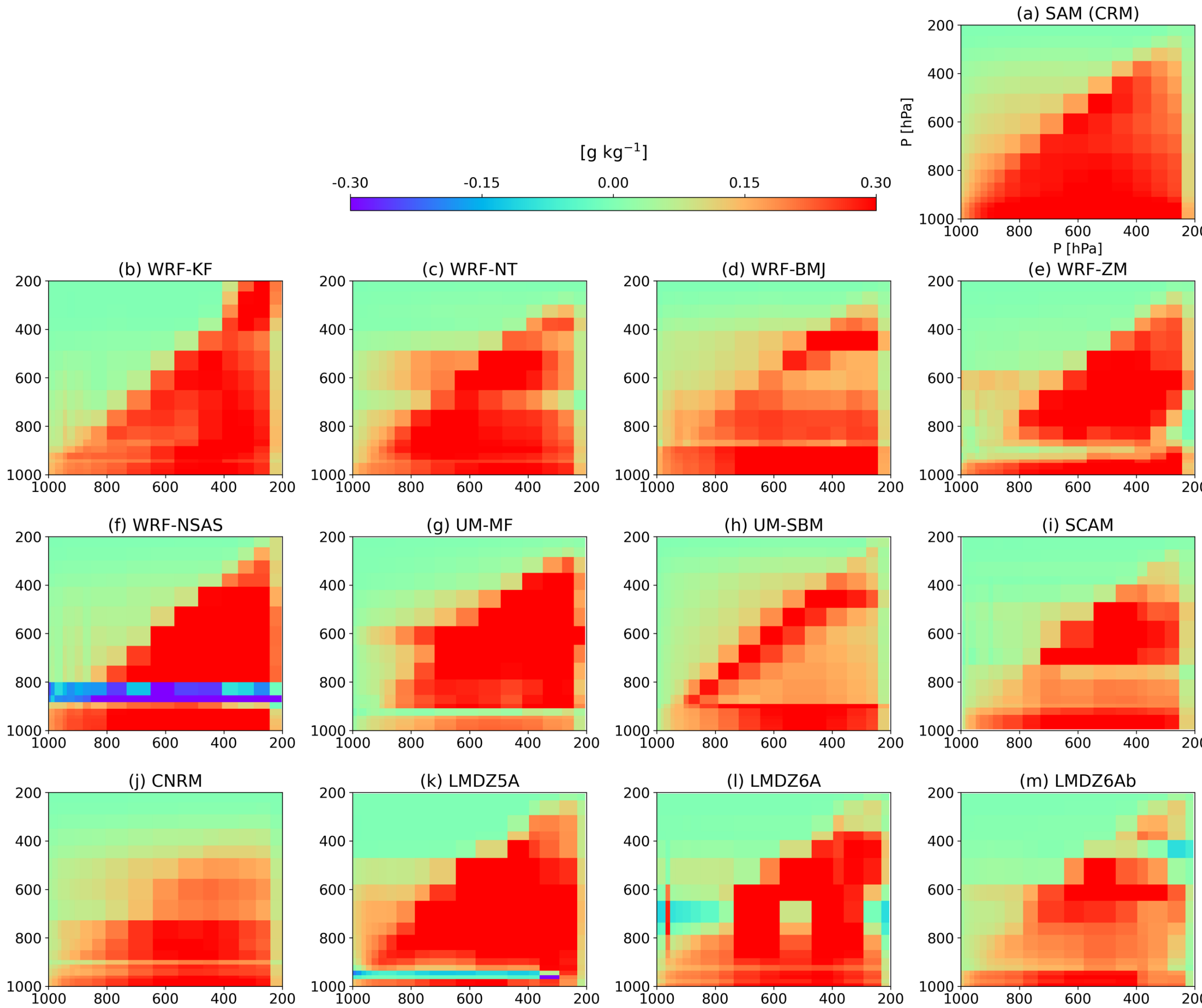

Figure 7. As in Figure 4, but for q responses to moisture tendency perturbation, in the units of g kg^{-1}.

4.2.1 Similarities between models

We first examine if the features presented in Section 4.1 are valid across all perturbation levels and models. Overall, as noted before, the CRM and SCMs all show a general tendency toward stronger T and q responses when perturbations are applied higher in the troposphere (Figures 4 – 7, increasing warmer colors towards the right columns of the matrices), and changes in q responses are generally biggest at the surface levels where moisture content is the biggest (Figures 5 and 7, warmer-colored horizontal layers close to surface), although sudden surges in q response are sometimes observed higher up. CNRM and Betts-Miller type schemes (WRF-BMJ and UM-SBM) behave most similarly to the CRM (d, h, and j in Figures 4 – 7), especially in their T responses. We offer a potential explanation for this in Section 4.2.2.

Additional similarities between the models can now also be observed when scrutinizing their complete M^{-1} matrices. In general, upper-tropospheric heating produces strong upper-tropospheric warming responses in all models (Figure 4, warmer colors in upper right corners, indicating stronger positive *T* responses), but inconsistent lower-tropospheric warming. Lower-tropospheric heating, on the other hand, leads to weak lower-tropospheric warming, but usually bigger upper-tropospheric warming. In other words, in the upper troposphere larger *T* responses are required to balance the imposed heating there, while heating applied in the lower troposphere requires much smaller *T* responses to stabilize. Also, heating applied at any level tends to increase the moisture below the perturbation level (Figure 5, red lower right triangles indicating positive *q* responses) and reduce it above, but to varying degrees among the models.

Next we examine the responses to moistening perturbations (Figures 6 and 7). Overall, the *T* responses to moistening are the most consistent across models of the four response types, and moreover are relatively uniform across a wide range of perturbation levels (Figure 6). This indicates that moistening tends to produce warming responses that are independent of where forcing is applied, while (as with the response to heating perturbations) increasing with height. Moistening also tends to provoke a stronger *q* response at and/or below the moistening level, sometimes with a weaker response above (Figure 7).

The above observations can be explained with the following physical interpretation. The difference in local *T* responses to heating perturbations in the upper and lower troposphere indicates strong lower tropospheric damping and weak upper tropospheric damping as noted earlier. Note that weaker damping is indicated by warmer colors in the figures (i.e., bigger responses required to compensate for the imposed perturbation). The increase in moisture below a heating level is also expected since heating stabilizes the atmosphere locally, inhibiting convection and trapping moisture below the heating level, leading to drying of the air above. The near-invariance of the response of *T* to the moistening level is interesting and the reason not obvious, but suggests that moisture added at any level ends up benefitting deep convection throughout the column.

4.2.2 Differences between models

Next we analyze the notable differences between the models. First, we note that the outlying behavior of UM-MF in its *T* response (horizontal discontinuity around 600 hPa and cool anomalies above heating levels; Figures 4g and 6g) and WRF-NSAS in its *q* responses (exceedingly strong, mostly negative, *q* responses around 850 hPa, Figures 5f and 7f) described in Section 4.1 is now observable across all perturbation levels. In general, the matrices of the SCMs are not as smooth as the CRM, containing more splotchy patterns that indicate jumpy responses, with discontinuities sometimes evident with respect to forcing level (vertical stripes) and sometimes with respect to response level (horizontal stripes). This is most apparent in the lower troposphere, possibly because responses in these layers are more dependent on contributions from different physics schemes (e.g., PBL and convection schemes). The inconsistent responses in the lower levels could be reflective of the different ways schemes represent shallow convection, downdrafts, and the evaporation of precipitation.

The kink around cloud base (~ 900 hPa) noted in Section 4.1 is clearly visible as a horizontal stripe across all perturbation levels and in all SCMs, most prominently in their *q* responses (Figures 5 and 7) but in a few cases also their *T* responses (Figures 4 and 6). Responses below this divide are often near constant and weaker than the rest of the column for *T* change (cooler-colored horizontal layers near surface in Figures 4 and 6) and stronger for *q* change (warmer-colored horizontal layers near surface in Figures 5 and 7). As mentioned, these discontinuities are not observed in the CRM, indicating that they probably reflect switches or threshold-like behavior common in convective parameterization, or perhaps deficiencies in the coupling to the PBL schemes. Our speculation of switches as the cause for these discontinuities is also supported by analyzing the linearity of the SCMs' responses (not shown). As mentioned in Section 2.2, we ensure that the responses are linear to a large extent (calculated with Equation 6). Nevertheless, non-linearities are sometimes detected and we found that they often coincide with the heights of the discontinuities. This suggests that switches—which are inherently non-linear—could be the common cause for both the discontinuities and response non-linearity.

It is noteworthy that the discontinuity around the LCL is more pronounced in the *q* responses than the *T* responses. This echoes findings of GCM studies, where moisture errors in convective regions are usually larger than temperature errors, possibly a consequence of deficiencies in the formulation of the entrainment and detrainment processes of moisture in some convection schemes (Gregory, 1997). For example, in a mass-flux based approach, errors in estimating the apparent moisture sink (*Q2* in the notation of Yanai et al., 1973) can arise when the effect of entrainment into the areas near cloud base is not properly represented, leading to an underestimation of drying in regions below 800 hPa and overestimation above this level (Gregory & Miller, 1989). This has consequences in the way a convection scheme behaves when additional heating or moistening is imposed in our experiment.

A few SCMs also display kinks or discontinuities in their *T* responses around the freezing level, which are not present in the CRM: around 650 hPa for WRF-NT, and around 600 hPa for WRF-ZM, UM-MF, and SCAM (c, e, g, and i in Figures 4 and 6). For the latter models *T* responses near the freezing level are generally weak (cooler color stripe), while for WRF-NT they are strong (warmer color stripe). All four SCMs use plume-based mass-flux schemes with CAPE closure, although the location of these anomalies near the freezing level suggests a possible role for microphysics and phase transitions around the freezing level.

Overall, we note two main groups of SCMs: the first displays smoother responses (especially in their *T* responses) that are more similar to the CRM, and the second exhibits more jumpy and disjointed behavior. As mentioned, the former consists of SCMs employing Betts-Miller adjustment type schemes (WRF-BMJ and UM-SBM) and CNRM. The remaining models belong to the latter group, and all employ mass-flux based convection schemes. A steep decrease in *T* response (at times negative) immediately above the imposed heating is often detected in the second group, most evident in WRF-KF, WRF-NSAS, UM-MF, and LMDZ6A (blue hues in b, f, g, and i in Figure 4). The discontinuity in responses (horizontal stripe) mentioned before is also more prominent in the second group. These behaviors may be a reflection of the way convection balances the imposed forcing. In mass-flux schemes, where it is mainly the subsidence term that balances the forcing, this can be achieved either through modification of the mass-flux shape or the environmental profile (or a mixture of the two). Where the mass-flux shape is less flexible, the environment has to be modified substantially to accommodate the forcing; where the mass-flux shape is more adaptive, less modification to the environment is required. It will be a subject of future research to identify the correct balance between these two.

The simpler assumptions of Betts-Miller adjustment-type schemes might result in more efficient balancing of the applied perturbations. We speculate that this could be due to how the closures are applied in the BM schemes for deep convection. In UM-SBM, the CAPE closure is applied by ensuring enthalpy conservation, which is achieved by either shifting the temperature reference profile (a cooling effect), or by reducing the precipitation rate computed from the moisture relaxation (a drying effect). Both methods are applied with a constant change to the convective tendencies at each vertical level between the ascent level and the level of neutral buoyancy. In WRF-BMJ, the enthalpy conservation is broadly similar to UM-SBM, as the applied enthalpy correction is smooth between the vertical levels. The closure of the BM schemes might explain why they are more effective in balancing the imposed forcing. The smooth CRM-like response of CNRM is interesting, as it is the only mass-flux scheme that exhibits smooth responses. What sets it apart from the schemes in the second group is its consistent use of buoyancy as the forcing term in the scheme design, including triggering condition, mass-flux calculation, entrainment and detrainment rates (Guérémy, 2011). It is possible that this smoother and continuous treatment of convection enhances the scheme's ability to respond locally to perturbations and could have contributed to its CRM-like responses. However, further tests are required to confirm this.

4.2.3 Comparison of SCMs with similar physics

We now analyse the M^{-1} matrices of similar or comparable SCMs: the three LMDZ cases (LMDZ5A, 6A, and 6Ab), the two Betts-Miller cases (WRF-BMJ and UM-SBM) and two Zhang-McFarlane cases (SCAM and WRF-ZM). Since these groups of SCMs share related convection schemes, they might be expected to produce similar results.

The three LMDZ versions share the same deep convection scheme but with different ways of handling shallow convection and associated clouds, and cold pools (Tables 1 and 2). They display significantly different responses (k, l, m in Figures 4 – 7). Two additional parameterizations are introduced in LMDZ6A that were not available in LMDZ5A: the representation of dry and shallow convection by a thermal plume model, and near-surface cold pools created by the evaporation of precipitation. Indeed, differences in response between LMDZ5A and LMDZ6A are the largest at low levels (below 800 hPa), with LMDZ6A displaying weaker T and q responses. The big discontinuity in the q responses of LMDZ5A around cloud base (950 hPa, purple line in Figures 5k, 7k) appears to be attenuated in LMDZ6A, perhaps an effect of the new parameterizations which are active at this level in LMDZ6A. The T responses to perturbations above 500 hPa are also stronger at high levels in LMDZ6A than

LMDZ5A, which could be related to the different representation of entrainment between the two versions (Grandpeix et al., 2004).

LMDZ6A displays unusually strong *q* responses within its shallow convective cloud layer (between 800 and 600 hPa) when perturbations are applied at certain levels (dark red blocks in Figures 5l, 7l). Its *T* responses also show unusual behavior in this layer, with a clear horizontal discontinuity at around 600 hPa and irregular responses below (800 – 600 hPa, Figures 4l, 6l), for example negative anomalies are observed when heating perturbations are applied below 800 hPa (blue hues in Figure 4l). In fact, our perturbation experiments have shed light on a problematic behavior of LMDZ6A that was not identified earlier with traditional 1D case-studies nor 3D experiments. Following these results, further investigation pointed to potential flaws in the representation of the evaporation of precipitation in the large-scale cloud scheme of LMDZ6A, which also handles shallow clouds. In LMDZ6A, evaporation of precipitation has two limitations: (1) it assumes that precipitation falls into clear sky, which could potentially overestimate evaporation in the shallow cloud layer, (2) at a given level it is not possible to saturate a fractional area greater than the maximum cloud fraction above, which can lead to underestimation of evaporation. LMDZ6Ab is a slightly modified version of LMDZ6A where a new scheme, inspired by Jakob and Klein (2000), has been introduced to take into account the overlap between clouds in the formation and evaporation of precipitation, thus addressing the two limitations outlined above. Our results here show that this new development has a significant effect on the model behaviour, improving its *q* responses between 800 and 600 hPa (l and m in Figures 4 – 7), as well as the linearity of the responses (not shown). This implies that the representation of the evaporation of precipitation may be an important factor in the response of a model to a modification of its environment. Note also that the RCE mean states of LMDZ6A and LMDZ6Ab are almost identical, showing that their M^{-1} matrices have captured important features of the models which are not obvious by only scrutinizing their mean states.

The two Betts-Miller SCMs employ related convection schemes but in two different SCM architectures and with otherwise different model physics. Even though they both exhibit behavior that is close to the CRM, there are telling differences between them. While their *T* responses are largely similar, they display quite different *q* responses. The *q* responses of UM-SBM to moistening applied at mid-levels (800 – 400 hPa) are more localized (dark red diagonal in Figure 7h), i.e., a peak in forcing is attenuated by a peak in response in the same region, whereas WRF-BMJ displays more uniform *q* responses to moistening (Figure 7d), more similar to the CRM. While these two models have similar convection schemes, the implementation of the two schemes is different enough that we would not expect *a priori* for the perturbation responses to be the same. Compared to the original Betts-Miller scheme, UM-SBM is a simplified version while WRF-BMJ is more complex. One possible explanation for the different *q* responses in these two cases could be that our experiments have picked up on the changes implemented by Janjic (2000) in the BMJ scheme that include a more sophisticated formulation of the moisture profile and variable relaxation time. It is also possible that other model differences play a role, although we think this is less likely (see Section 6).

The two Zhang-McFarlane SCMs employ nominally identical model physics in two different SCM model systems (WRF vs. CAM). As we would hope, they exhibit largely similar behavior (e and i in Figures 4 – 7). Although not identical, they are still significantly more similar to each other than to the other SCMs. In any case, their M^{-1} matrices are more similar than their RCE profiles in Figure 1, again suggesting that linear responses may provide a clearer window into model physics than mean profiles. A slight horizontal discontinuity around the freezing level (~ 600 hPa) in the *T* responses is visible in both SCMs (e and i in Figures 4 and 5). Given the location of this discontinuity, one possible explanation could be the interaction between the Zhang-McFarlane deep and the UW shallow convection schemes. To test this, using WRF we reran the experiments with only the Zhang-McFarlane deep convection scheme switched on and the UW shallow convection scheme switched off. Results show that the horizontal discontinuity around the freezing level remains present (not shown). We further tested altering a constant that defines the freezing level in the ZM scheme, which shifted the horizontal stripe to the new specified freezing level, confirming that it is caused by the ZM scheme. As mentioned before, such discontinuities could indicate threshold setting in the scheme; in the ZM scheme, for example, a threshold is implemented to restrict precipitation production only to clouds that extend beyond the freezing level (Zhang & McFarlane, 1995), which could explain our results.

We note that the physical explanations presented in Section 4 are preliminary and speculative at this point. Nonetheless, they serve as useful hypotheses to guide ongoing research. The main takeaway from this section is that the idealized framework based on the linear response function that we have applied is able to illuminate and locate areas of agreement and differences between model physics, which can provide insights into physical processes or ways to simplify or improve current convective parameterizations.

5 Relationship between RCE mean states and responses

We noted in the previous section a couple of examples where aspects of model behavior changes were more evident in the linear responses than in RCE mean states (temperature and RH) described in Section 3. In this section we examine more generally if the linear responses can be linked to the RCE mean states in any way. One aspect that is well documented is the interaction between environmental humidity and convection. Convective activity has been shown to be sensitive to environmental humidity in observational studies (Brown & Zhang, 1997; Parsons et al., 2000; Sherwood et al., 2004) and experimental analyses using CRMs (Tompkins, 2001; Grabowski 2003). Derbyshire et al. (2004), for example, found a significant impact of mid-tropospheric humidity on convective activity, where a dry RH inhibits deep convection and encourages shallow convection instead. A recent study by Wolding et al. (2020) found a cyclical behavior of moisture and convection which points to a joint evolution of the two variables. In our experiment, we found a large spread in the SCMs' RH profiles as well as their responses. Convection plays a role in influencing both. However, we do not know if they (RH and *T*, *q* responses) respond in similar fashion. This section addresses this question. Specifically, *do a model's temperature and moisture responses to heating and moistening perturbations correlate with its RCE mean state?*

The first aspect we examine is whether the shape of a model's mean RH profile is linked to the shape of its responses. As pointed out in Section 3, the mean RH profiles often contain kinks. We found that these kinks almost always coincide with discontinuities in the linear responses (horizontal stripes in the M^{-1} matrices). These kinks are ubiquitous at cloud base but can also be observed at ~ 700 hPa for WRF-NT, ~ 800 hPa for WRF-NSAS, ~ 600 hPa for WRF-ZM, SCAM, and UM-MF, and ~ 500 hPa for LMDZ5A. This collocation of RH kinks and discontinuities in linear responses are found in both T and q responses to heating and moistening perturbations. The smooth M^{-1} quadrants of the CRM are likely related to its smooth RH profile. Additionally, the size of the kinks in the RH profile appears to have an impact on the size of the kinks in the responses: for example, the big RH kinks around 800 hPa of WRF-NSAS and around 600 hPa of UM-MF (Figure 1e) coincide with strong discontinuities in their responses at the same heights (f and g in Figures 4 – 7), while the smaller RH kinks around 600 hPa of WRF-NT, WRF-ZM, and SCAM coincide with smaller or less obvious discontinuities in their responses. The even smaller RH kink of WRF-BMJ around 600 hPa hardly registers in its responses. The correspondence appears to fade away in higher altitudes (above 500 hPa): for example, the RH kinks around 450 and 350 hPa of WRF-NT, around 450 hPa of WRF-NSAS, and around 400 hPa of WRF-KF are not noticeably associated with discontinuities in their respective model responses. This is probably because the amount of moisture available at these higher altitudes is too small for any sharp changes to be registered in the responses.

Now that we have seen that the *shape* of the RCE mean RH profile is linked to the linear responses, next we examine if there is a correlation between the *magnitude* of these two components in either temperature or moisture. This will tell us whether, if a model's environment is warmer or moister, it will also respond more strongly to heating or moistening perturbations. To this end, we correlate the RCE θ_{es} and RH values of all the models at specific pressure levels with their responses at various levels and averaged over all perturbation levels, i.e., the average of a horizontal stripe in a M^{-1} quadrant (negative anomalies are set to zero to avoid ambiguity in interpreting the correlations). We compute the Spearman correlation coefficient of these correlations as it is more suitable for non-parametric data and less sensitive to outliers than Pearson correlation coefficient (Kokoska & Zwillinger, 2000), although both methods for computing the coefficients return similar results for our experiment. Eight common pressure levels were selected between 1000 and 200 hPa, in intervals of 100 hPa. This yields eight 8 x 8 correlation matrices, one for each combination of RCE variable (θ_{es} or RH), forcing variable (dT/dt or dq/dt) and response variable ($\overline{T'}$ or $\overline{q'}$), with the matrix entry in the i-th column and j-th row representing the correlation coefficient between the RCE variable at pressure level p_i and response at pressure level p_j. In other words, an entry in our correlation matrix denotes the Spearman correlation coefficient r_{ij} between two data series A_i and B_j:

$$r_{ij} = corr(A_i, B_j) \quad (7)$$

where $A_i = [a^i_{m1}, a^i_{m2}, \ldots, a^i_{m13}]$, with a^i representing the RCE value (θ_{es} or RH) at pressure level p_i and $m1, m2, \ldots, m13$ denoting the 13 models in our study (CRM and 12 SCMs);

and $B_j = [b^j_{m1}, b^j_{m2}, \ldots, b^j_{m13}]$, with b^j representing the mean response ($\overline{T'}$ or $\overline{q'}$ to *dT/dt* or *dq/dt* perturbation) at pressure level p_j for the models.

We found significant correlations (p-value < .05) only between the RCE RH and q responses to applied heating, shown in Figure 8. The other correlation matrices contain mostly weak correlations ($|r_{ij}| < .5$) and are not explored here. Apart from in the boundary layer, RCE RH is positively correlated with q responses locally and at levels higher up, evident by the red tiles in and above the main-diagonal. That is, a high RCE RH at level p tends to correspond to strong q responses at p and above, or a strong q response at p tends to correlate with high RH values at p and below. The local correlations suggest that high RCE RH values at certain levels indicate that convection is acting strongly and introducing moisture near those levels, and thus when convection is slightly enhanced via a temperature or moisture tendency perturbation, the q responses at those levels are also bigger due to the bigger effect of convection there. The strong positive correlations above the main diagonal are interesting. These results suggest high RH at level p permits convection to penetrate that level more easily, which leads to stronger q responses above p. Another interpretation, albeit more ambiguous, is that a strong influence of convection at level p causes a big q response at p (i.e., local correlations), as well as higher RH at p and below due to convectively induced subsidence. Interestingly, the same correlations are not observed for T responses. In other words, while the shape of the RCE RH profile reflects that of both T and q responses, the magnitude of mean RH reflects only the magnitude of q responses.

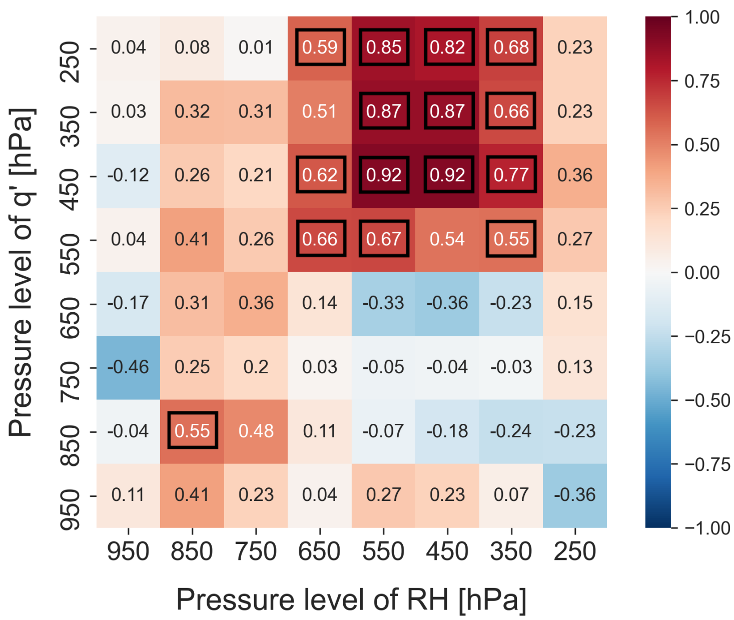

Figure 8. Correlation matrix of RCE RH and q responses to temperature tendency perturbation. An entry in the i-th column and j-th row represents the correlation between the RCE RH values of the SCMs at pressure level i and their mean q responses at pressure level j. Significant correlations (p-value < .05) are shown in black boxes.

6 Sensitivity to PBL and MP schemes

Here we present results from the tests to determine the role of schemes other than convection schemes in a model's linear response, as described in Section 2.3. Specifically, this section addresses the question: *do a model's RCE mean state and responses to heating and moistening perturbations change significantly when different PBL or microphysics (MP) schemes are used?*

We first present the sensitivity of the RCE mean states to the choice of PBL and MP schemes (Figures 9 and 10). Figure 9 shows clearly that the impact of the other schemes on the mean state temperature, especially the microphysics scheme, is small compared to that of the convection scheme. The RCE profiles of RH do show some sensitivity to choice of PBL and MP schemes, but at different heights of the troposphere (Figure 10). For PBL sensitivity, differences are more prominent in the lower- to mid-troposphere (below 500 hPa). For MP sensitivity, divergence between the MP schemes appears mostly in the upper troposphere (above 500 hPa). This is consistent with expectations that the treatment of convective outflows and cloud hydrometeors will be most important to the water vapor budget in the upper troposphere where vapor amounts are smallest. Overall, the RCE temperature (θ_{es}) profiles are predominantly decided by the convection scheme while the RH profiles can be influenced by the PBL and MP schemes at different elevations.

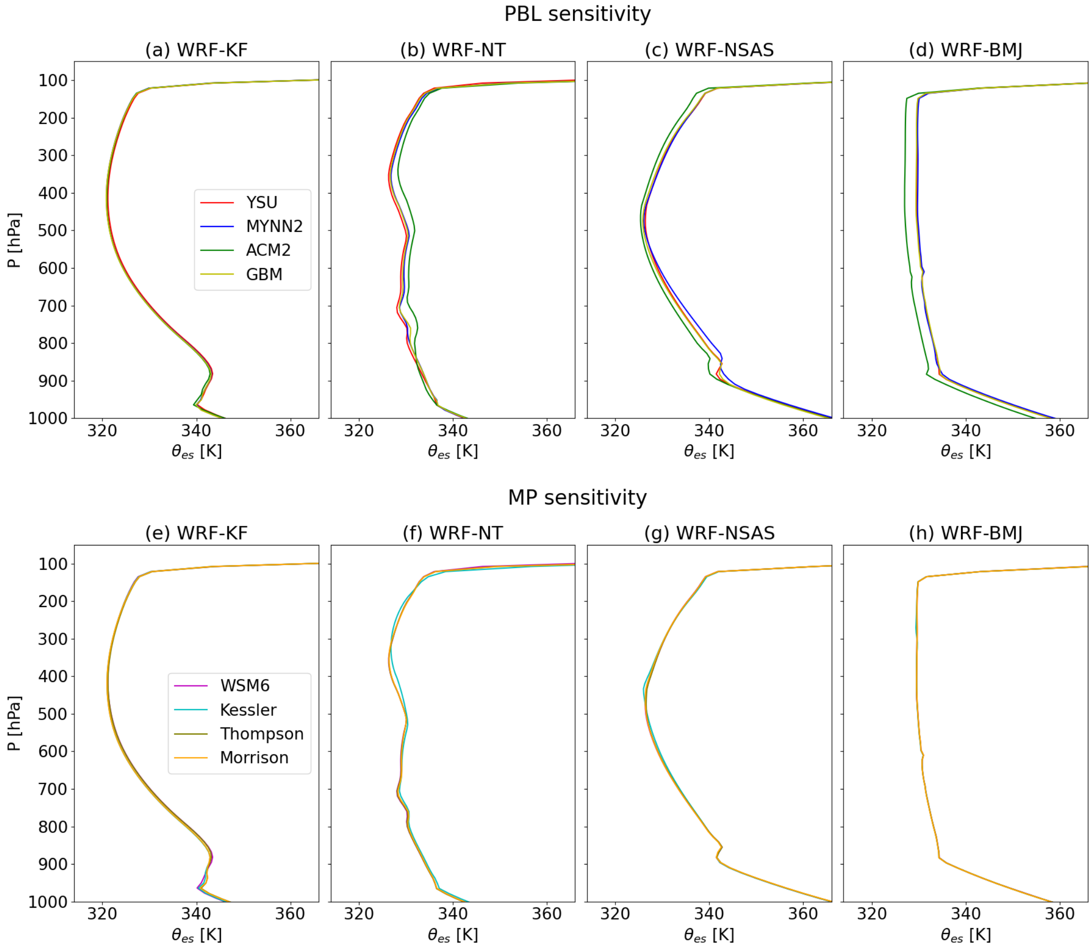

Figure 9. RCE saturation equivalent potential temperatures sensitivity of WRF-KF (a, e), WRF-NT (b, f), WRF-NSAS (c, g), and WRF-BMJ (d ,h) to PBL (top) and MP (bottom) schemes.

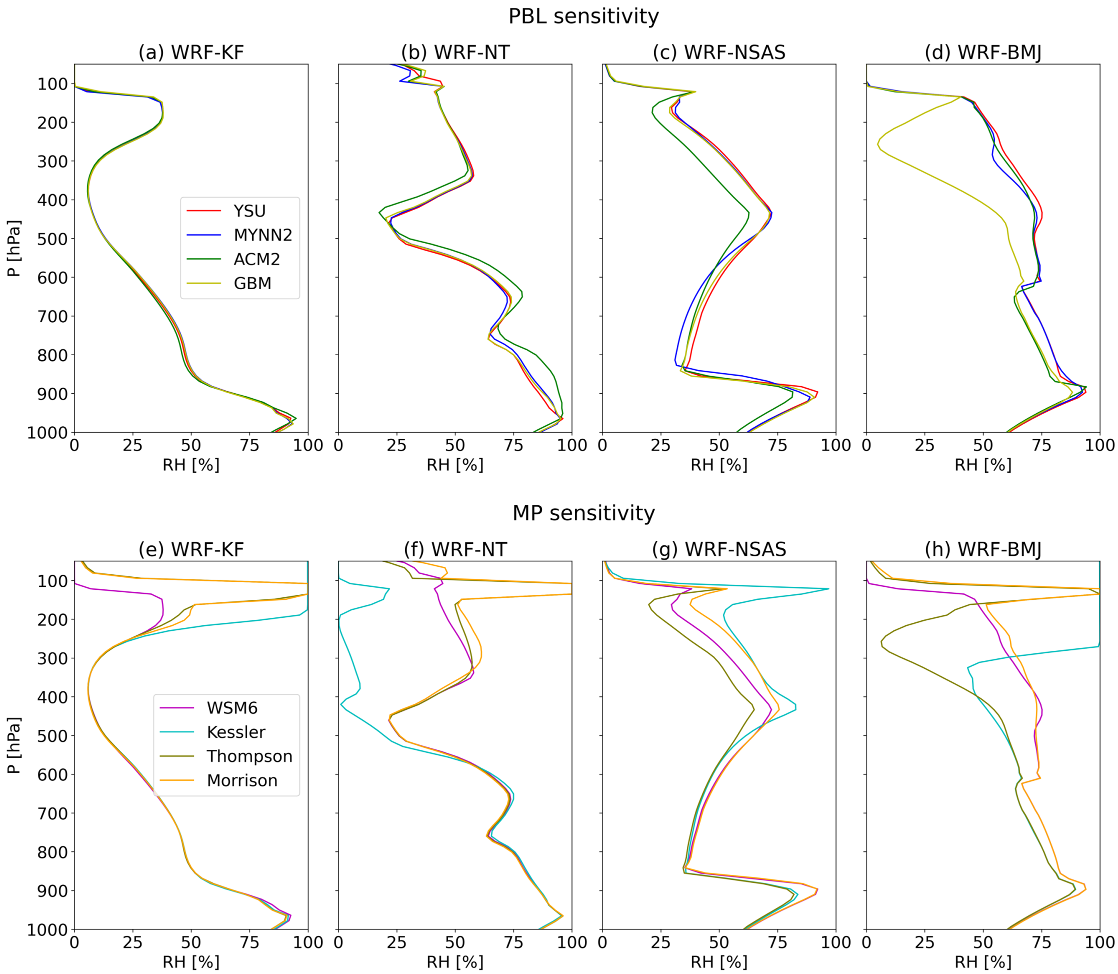

Figure 10. As in Figure 9 but for RCE relative humidity

Next, we present the sensitivity of T and q responses to the choice of PBL and MP schemes (Figures 11 and 12). To explore this we only perturbed the two levels shown in Figures 2 and 3 (850 and 650 hPa). As perturbing both levels return similar results, only results from the 850 hPa perturbation case are shown. We also combine the results for temperature and moisture tendency perturbations and show only the average as their sensitivities are very similar. Overall, the responses are not sensitive to MP schemes (Figure 12), and slightly more sensitive to PBL schemes (Figure 11). WRF-KF is not sensitive to changes in either PBL or MP schemes. For WRF-NT, WRF-NSAS, and WRF-BMJ, the responses to temperature and moisture tendency perturbations when combined with different PBL schemes retain their general shape, except for the case of ACM2 PBL scheme, which shows outlying q response when combined with WRF-NT.

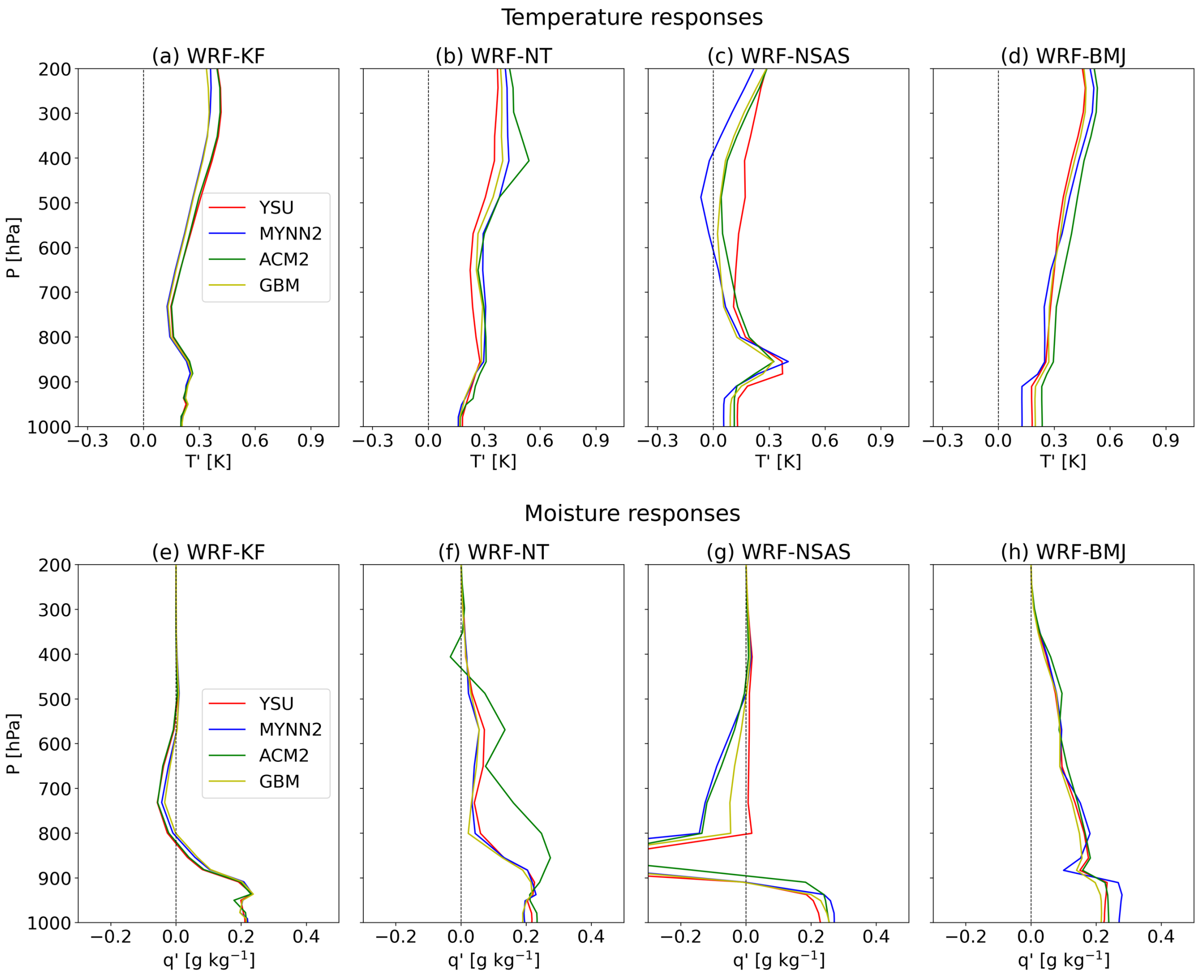

Figure 11. Sensitivity to PBL schemes of *T* (top) and *q* (bottom) responses to perturbations at 850 hPa (averaged between temperature and moisture tendency perturbations) for WRF-KF (a, e), WRF-NT (b, f), WRF-NSAS (c, g), and WRF-BMJ (d, h).

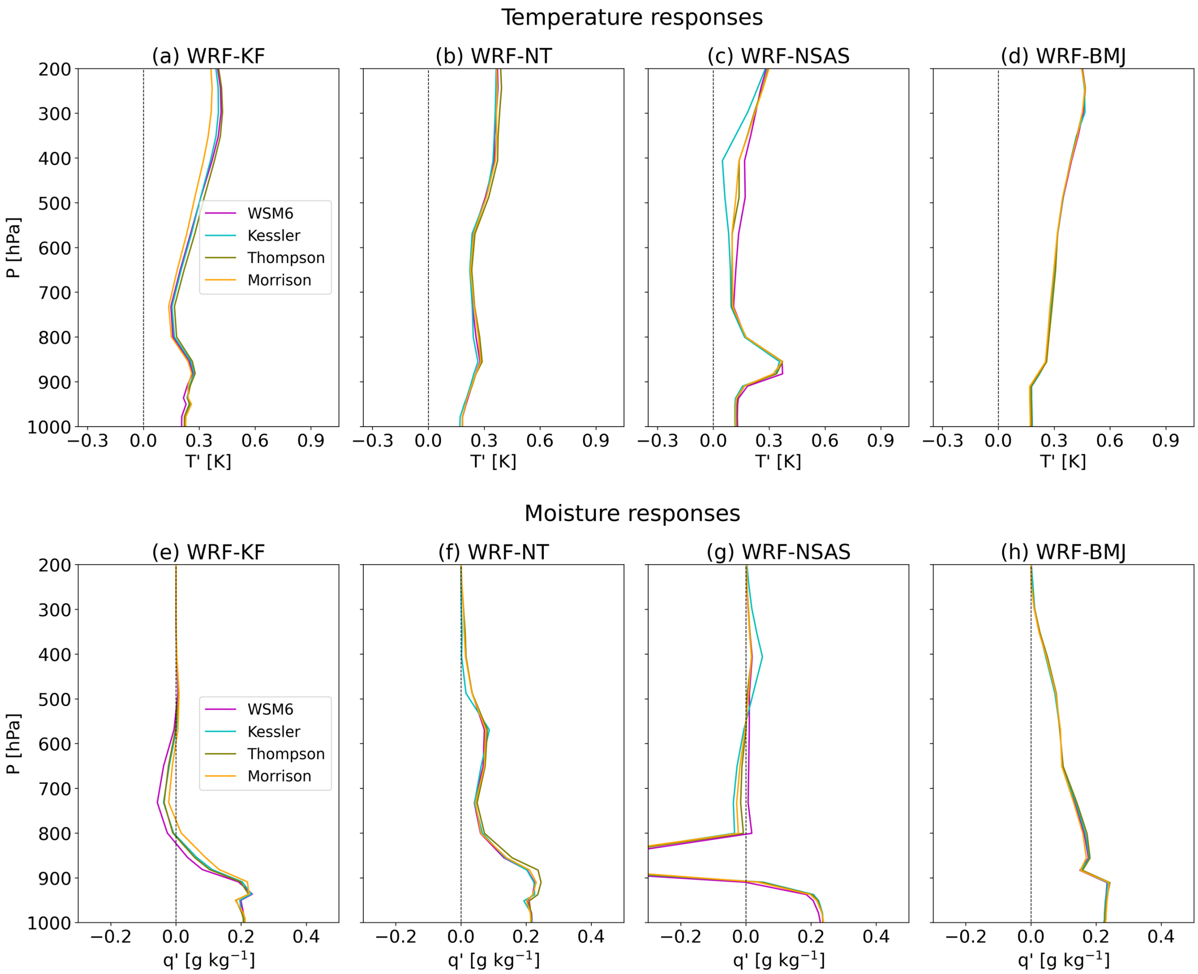

Figure 12. As in Figure 11 but for sensitivity to MP schemes

In summary, we find that the T and q responses are much more sensitive to the convection schemes than the PBL or MP schemes, indicating that our perturbation experiments can isolate the impact of convection schemes. This is also true for the RCE RH profile, but only at low and mid levels, above which it is affected by microphysics. However, there are important caveats to these findings. These experiments have only been conducted in the WRF model, which has a modular design and relatively independent physics schemes. The same insensitivity might not hold in other models that employ a more integrated approach in the design of its model physics, where there is a tighter coupling between the schemes. See, for example, the differences between the response matrices of LMDZ6A and LMDZ6Ab, where only the large-scale cloud scheme has been modified. Note, however, the large-scale cloud scheme in LMDZ also handles shallow clouds (in the WRF cases shallow clouds are handled by the convection scheme), hence it is still reasonable to postulate that convective parameterization (including convective MP) dominates the linear responses at least in the lower- and mid-troposphere. Also, the weak sensitivity to PBL and MP schemes is most likely exaggerated by our experimental setup. Specifically, the use of RCE with an idealized radiative cooling profile, and a constrained surface flux computation. If these sensitivity tests are repeated with interactive radiation, surface wind and exchange coefficients, sensitivity to PBL and MP schemes becomes more significant (see Appendix A).

7 Conclusions

The overall goal of this paper is to advance our understanding of what can be learned about model physics from single-column models (SCMs) run in radiative-convective equilibrium (RCE) configurations. The objectives are threefold: first, to compare the RCE mean states of a few SCMs containing state-of-the-art physics currently used in atmospheric modeling; second, to compare and examine the behavior of the SCMs by observing their steady temperature and moisture responses to small temperature and moisture tendency perturbations (M^{-1} matrices) using the linear response framework (Kuang 2010; Herman & Kuang 2013); and third, to determine which physical schemes control the RCE mean state and/or linear responses.

In terms of the first objective, similar to other recent intercomparison studies (e.g., Wing et al., 2020) we found substantial differences between the SCMs in their RCE temperature and relative humidity (RH) profiles, with ~ 5 K differences in absolute temperature in the near-surface levels and ~ 8 K in the free troposphere (with the exception of one outlying SCM) and free-tropospheric RH spanning nearly the entire possible range (0 – 100%). Even between the SCMs that use similar convection schemes, the difference in their RCE profiles is nontrivial: the two Zhang-McFarlane cases (WRF-ZM and SCAM) show similar shapes in their RH profiles but WRF-ZM is consistently somewhat drier than SCAM and the temperatures vary by several K at some levels, while the RH profiles of the two Betts-Miller cases (Betts-Miller-Janjic in WRF and Simplified Betts-Miller in UM) differ in both shape and magnitude.

In addressing the second and third objectives, we arrive at the following main conclusions:

1. The idealized SCM testing framework appears capable of isolating the behavior of convection schemes, thus enabling direct evaluation of these schemes against CRM or LES reference calculations.
2. This framework identifies areas of agreement, but also substantial differences in behavior among the models, which in some cases can be related to scheme design.
3. Some linear responses correlate with the RCE mean profiles (RH in particular), while others do not and hence constitute independent information. While the RCE RH profile is strongly influenced by the convection scheme, it is more sensitive to other physics schemes than are the linear responses. The RCE temperature profile is however insensitive to schemes other than the convection scheme, in this setup.
4. Almost all SCMs show irregularities or discontinuities in behavior that are likely related to switches or thresholds built into the convection scheme(s), and which do not appear in the SAM CRM.

These conclusions will now be briefly discussed in turn.

First, our experiments manage to largely isolate the behavior of the convection schemes in the SCMs. We found multiple lines of evidence for this. In the WRF model, the temperature and moisture responses to applied heating and moistening vary greatly among the convection schemes but do not deviate much when different microphysics (MP) or planetary boundary layer (PBL) schemes are used. This shows that—although in some cases the PBL scheme exerts some influence—the T and q responses are predominantly decided by the convection scheme. Also, the linear responses of the same or comparable convection schemes (the two Zhang-McFarlane and Betts-Miller cases) are considerably more alike than their RCE profiles are, supporting this finding.

Second, our framework highlights the areas of agreement and disagreement between the SCMs, and between them and the CRM, which can potentially be linked to the convection scheme design of the SCMs. The SCMs in our experiment generally reproduce the broad behavior of the CRM, albeit to different degrees. Their responses are often not as smooth and contain more splotchy and irregular patterns. Nevertheless, many SCMs exhibit behavior that is closer to the CRM than the SCMs in HK13. In general, heating perturbations lead to more diverse responses among the SCMs than do moistening ones. These disparities in response point to the different characteristics of the convection schemes and provide clues as to where to focus further investigations. Overall, two main groups emerge from inspecting their responses: the first group exhibits smooth responses akin to that of the CRM and the second displays more jumpy responses. The former group includes two variations of an adjustment-type convection scheme (Betts-Miller) and a buoyancy-based mass-flux convection scheme (CNRM), while the latter contains only mass-flux based convection schemes with CAPE closures. A scheme's responsiveness in the vertical might hold the key to the smoothness of its response. The CRM-like responses of the Betts-Miller cases point to the efficiency of adjustment-type schemes to counteract the applied localized perturbations, while the dependency of the mass-flux based schemes on vertically integrated quantities perhaps hindered their responsiveness and contributed to their bumpier responses. Our experiments also highlight important discrepancies between the three versions of the LMDZ model that employ different physical packages, uncovering shortcomings in LMDZ6A that previous studies using traditional methods have not discovered. Notably, LMDZ6A and LMDZ6Ab display almost identical RCE mean states, but very different linear responses, with LMDZ6A exhibiting abnormally strong q responses within the shallow convective cloud layer. Following an update in the way evaporation of precipitation is represented in the model (LMDZ6Ab), a marked improvement in the model's moisture responses in the shallow cloud layer was observed, demonstrating the usefulness of our framework in parameterization development.

Third, some aspects of the linear responses correspond to features of the RCE mean profiles, while others do not and can be regarded as independent diagnostics. As mentioned above our experimental setup can isolate the behavior of the convection scheme. This is also true for the RCE temperature profiles although they provide less information about the differences between convection schemes. It is partially true for the RH profiles, where the convection scheme has the strongest influence, but only at low and mid-level altitudes, above which the MP scheme plays a significant role. In other words, multiple physics schemes could potentially exert control on a model's RCE mean state, whereas its *T* and *q* responses depend mainly on the convection scheme. It is unclear how to physically interpret links between the RCE mean profile and linear responses, since either could affect the other. The extent to which the models' diverse RCE mean states directly influence their responses is hard to estimate. Like HK13, we did not attempt to tune the parameters of the SCMs to bring their mean states closer to each other[1]. Nonetheless, in our experiments we found evidence that the two measures are correlated to some extent, particularly the RCE RH profiles and the perturbation responses. The responses correspond to the model's RH profile in two ways. First, the shape of the RH profile is related to the shape of the responses in the sense that kinks in the RH profiles often locally coincide with kinks in the responses (both *T* and *q* responses). The models that display more uniform responses also produce smoother RH profiles in RCE (the SAM CRM, Betts-Miller schemes, and CNRM). Second, the magnitude of RH is positively correlated with the magnitude of *q* (but not *T*) responses locally, as well as higher above, suggesting that a wetter environment corresponds with convective activity that introduces moisture locally, and hence when we apply perturbation the models with bigger RH react more vigorously in their moisture response, possibly caused by detrainment. It is noteworthy that the shape of RH corresponds to the shape of both *T* and *q* responses, while the magnitude of RH is linked only to the magnitude of *q* responses. This implies that the two moisture-related variables (RH and *q* responses) tend to behave in a consistent manner, while *T* responses can be regarded as a complementary diagnostic.

Fourth, all SCMs in our study show discontinuities in their behavior that are likely associated with switches or thresholds embedded in the convection scheme design, and which are not observed in the CRM. Although the responses of our SCMs are linear to a large extent, the locations (heights) of the bigger non-linearities often coincide with discontinuities in their responses, suggesting a common cause. Since switches are inherently non-linear, it is reasonable to suggest that they are a possible explanation for both non-linearity and response discontinuities. These discontinuities manifest themselves as horizontal stripes in the M^{-1} matrices, which often divide the responses into regions with distinctive behaviors. For example, a discontinuity is observed around the model-predicted cloud base level in all the SCMs. In a few SCMs, discontinuity is also observed around the freezing level, indicating an inability of the scheme to respond smoothly to phase transition. Admittedly, the vertical transport of heat and moisture through transitioning levels is challenging to parameterize (Neggers et al., 2017). To simplify matters, convection schemes often use switch-like mechanisms in their design. Thresholds are also a common feature used for the triggering of deep convection. For example the Arakawa-Schubert scheme uses the threshold value for a concept called cloud work function to trigger convection (Arakawa & Cheng, 1993). Suhas and Zhang (2014) analyze the triggering systems of a few widely-used convection schemes and found that some of their performance can be improved by optimising the threshold values used. Ultimately, these threshold values are often

[1] The vertical resolution of a model likely also has an impact on its responses, which we also did not standardize between the SCMs.

subjective and sometimes arbitrary. They are at best *ad-hoc* limitations placed in a scheme to represent processes that we do not yet fully understand, and our experiment captures this flaw.

By expanding on the experiments of HK13 to a few widely-used models and convection schemes, we demonstrate that the idealized framework based on a model's responses to small heating and moistening perturbations is a useful approach to study the behavior of models and their parameterizations. In this study we compare our results to the CRM (2 km resolution) results of K10 as it is the most viable option available to us. However, we caution that these CRM results cannot be regarded as the "truth", as past studies have shown that CRMs can potentially return different results depending on model resolution (Fan et al., 2017; Lebo & Morrison, 2015; Varble et al., 2014) and other parameterized physics such as the microphysics schemes (Khain et al., 2015; Kim et al., 2014; Liu & Moncrieff, 2007). There is a need for more studies to be done—large-eddy simulations (LES), for example—to verify K10's results. Nevertheless, the T and q responses presented here are a simple and helpful way to characterize and evaluate convection schemes. Clues for deficiencies in a scheme can be diagnosed from the irregularities in the M^{-1} matrices and the location of these irregularities could provide guidance in examining the causes of errors in model physics. Further investigations into potential physical explanations for the behaviors identified here form part of our ongoing work.

Acknowledgments

We thank Zhiming Kuang for sharing his CRM results; Xinrong Chua and Shuguang Wang for the source code of the surface flux computation in WRF; I-Kuan Hu, Wayne Angevine, and Hugh Morrison for useful discussions. YLH, SS, SCS, and DF were supported by the Australian Research Council (grant FL150100035). CLD and RSP were supported by Natural Environment Research Council (NERC) funding (grant NE/T003871/1) under the joint NERC/Met Office ParaCon programme led by AJS from the Met Office. CR, RR acknowledge support from the DEPHY research group, funded by CNRS/INSU and Météo-France. PM is funded by the NERC and Met Office as part of the ParaCon project (Grant NE/N013123/1 and NE/T003863/1).

The data and scripts required to reproduce the results described in this paper are available in a Zenodo repository: https://zenodo.org/record/4433226 (DOI: 10.5281/zenodo.4433226)

Appendix A

We report here the impact of our idealized experimental procedure on the sensitivity of the responses to PBL and MP schemes. Specifically, we show the effects of the idealized radiative profile and surface flux computation. In the first set of simulations we enabled interactive radiation and kept the idealized surface flux computation (following Equations 3 and 4); in the second set of simulations we kept the idealized radiative profile and enabled fully interactive surface flux computation. These simulations were carried out with two WRF cases: WRF-KF (a mass-flux scheme) and WRF-BMJ (an adjustment-type scheme).

Results are shown in Figures A1 and A2. For both SCMs, PBL and MP schemes affect results when either the radiation or surface wind and exchange coefficients are made interactive, but to varying degrees. For WRF-KF, the responses are sensitive to the choice of MP scheme when the radiation is interactive, likely due to the impact of cloud changes on radiation (which are negated in the idealized setup), while fully interactive surface fluxes only slightly decrease the sensitivity. For WRF-BMJ, the responses are significantly more sensitive to the choice of PBL scheme and slightly more sensitive to the choice of MP scheme when either of the idealized settings is disabled. In any case, applying both idealized settings decreases the dependence of the responses on PBL and MP schemes.

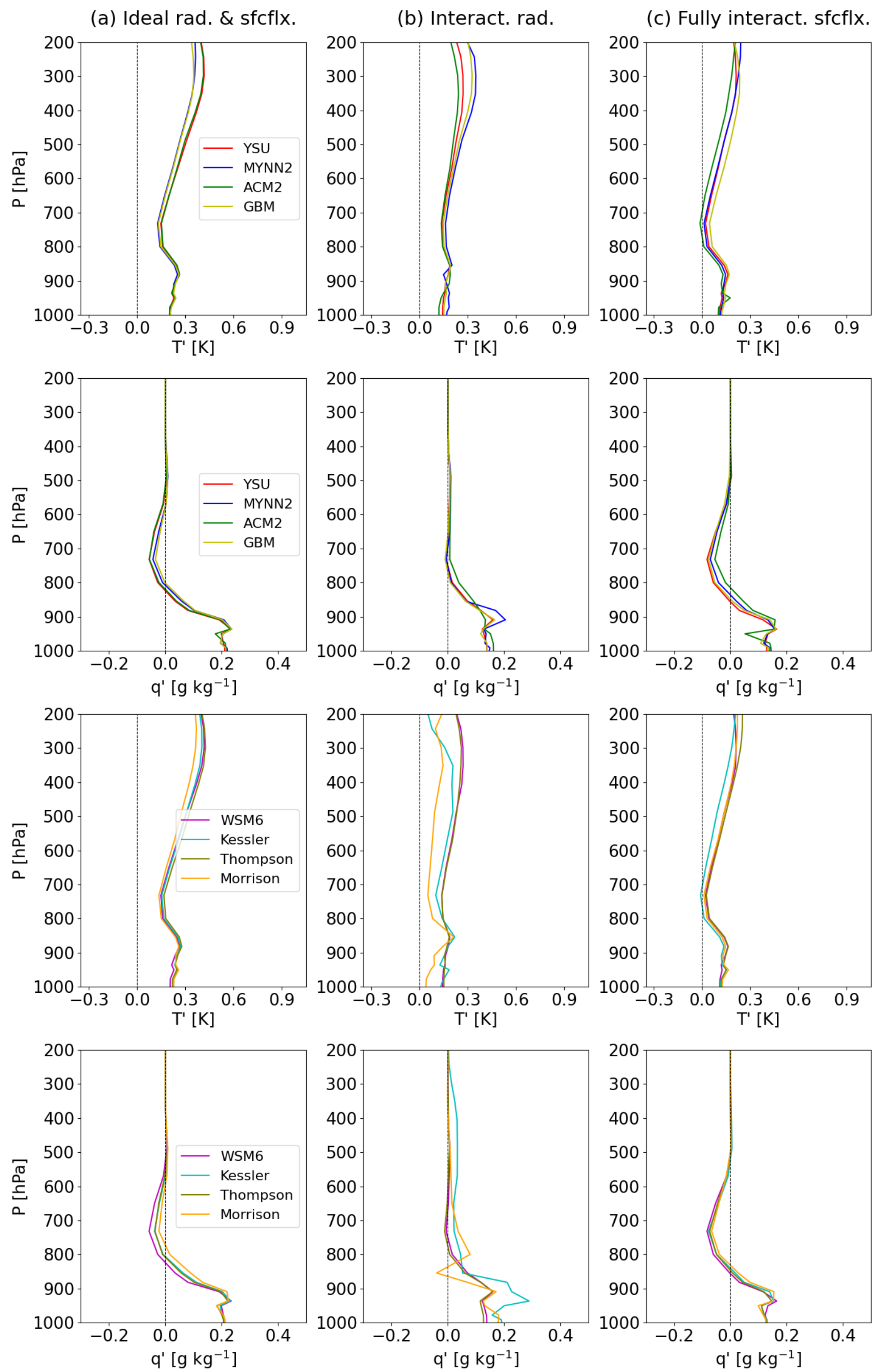

Figure A1. WRF-KF sensitivities comparison for (a) ideal radiation and surface fluxes, (b) interactive radiation and ideal surface fluxes, and (c) fully interactive surface fluxes and ideal radiation. As in Section 6, responses to *dT/dt* and *dq/dt* perturbations are averaged. PBL sensitivities are shown in first and second rows, and MP sensitivities in third and fourth rows.

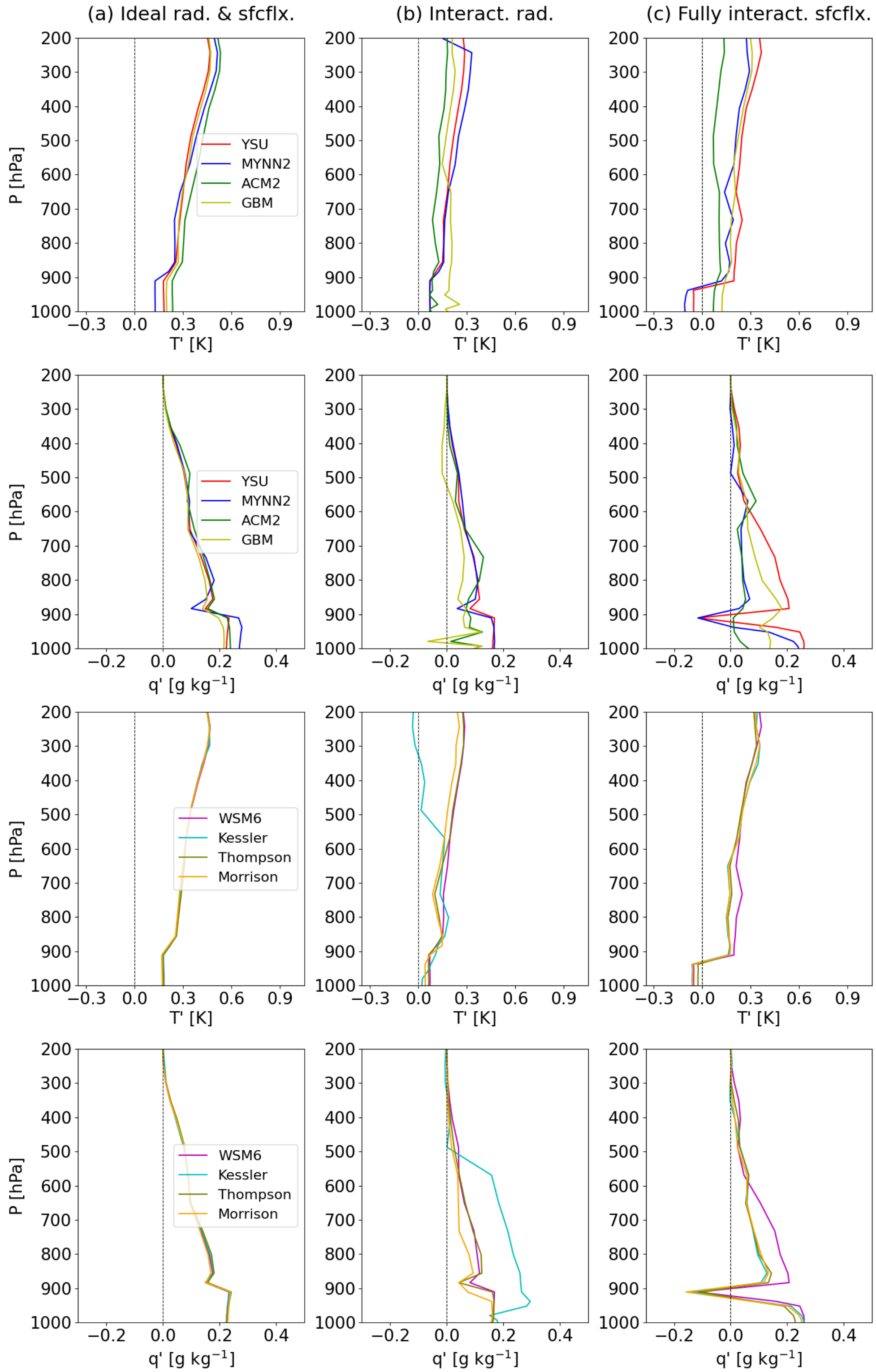

Figure A2. As in Figure A1 but for WRF-BMJ.

References

Arakawa, A. (2004). The cumulus parameterization problem: Past, present, and future. Journal of Climate, 17(13), 2493-2525. https://doi.org/10.1175/1520-0442(2004)017%3C2493:RATCPP%3E2.0.CO;2

Arakawa, A., & Cheng, M. D. (1993). The Arakawa-Schubert cumulus parameterization. In The representation of cumulus convection in numerical models (pp. 123-136). American Meteorological Society, Boston, MA. http://doi.org/10.1007/978-1-935704-13-3_10

Arakawa, A., & Schubert, W. H. (1974). Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. Journal of the Atmospheric Sciences, 31(3), 674-701. https://doi.org/10.1175/1520-0469(1974)031%3C0674:IOACCE%3E2.0.CO;2

Bechtold, P., Bazile, E., Guichard, F., Mascart, P., & Richard, E. (2001). A mass-flux convection scheme for regional and global models. Quarterly Journal of the Royal Meteorological Society, 127(573), 869-886. https://doi.org/10.1002/qj.49712757309

Bechtold, P., Semane, N., Lopez, P., Chaboureau, J. P., Beljaars, A., & Bormann, N. (2014). Representing equilibrium and nonequilibrium convection in large-scale models. Journal of the Atmospheric Sciences, 71(2), 734-753. https://doi.org/10.1175/JAS-D-13-0163.1

Berner, J., Achatz, U., Batte, L., Bengtsson, L., Cámara, A. D. L., Christensen, H. M., ... & Franzke, C. L. (2017). Stochastic parameterization: Toward a new view of weather and climate models. Bulletin of the American Meteorological Society, 98(3), 565-588. https://doi.org/10.1175/BAMS-D-15-00268.1

Betts, A. K. (1986). A new convective adjustment scheme. Part I: Observational and theoretical basis. Quarterly Journal of the Royal Meteorological Society, 112(473), 677-691. https://doi.org/10.1002/qj.49711247307

Betts, A. K., & Miller, M. J. (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. https://doi.org/10.1002/qj.49711247308

Bony, S., & Dufresne, J. L. (2005). Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models. Geophysical Research Letters, 32(20). https://doi.org/10.1029/2005GL023851

Bony, S., & Emanuel, K. A. (2001). A parameterization of the cloudiness associated with cumulus convection; evaluation using TOGA COARE data. Journal of the atmospheric sciences, 58(21), 3158-3183. https://doi.org/10.1175/1520-0469(2001)058%3C3158:APOTCA%3E2.0.CO;2

Bony, S., Stevens, B., Frierson, D. M., Jakob, C., Kageyama, M., Pincus, R., ... & Watanabe, M. (2015). Clouds, circulation and climate sensitivity. Nature Geoscience, 8(4), 261-268. https://doi.org/10.1038/ngeo2398

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., ... & Rasch, P. (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.

Bougeault, P. (1981). Modeling the trade-wind cumulus boundary layer. Part I: Testing the ensemble cloud relations against numerical data. Journal of the Atmospheric Sciences, 38(11), 2414-2428. https://doi.org/10.1175/1520-0469(1981)038%3C2414:MTTWCB%3E2.0.CO;2

Bretherton, C. S., McCaa, J. R., & Grenier, H. (2004). A new parameterization for shallow cumulus convection and its application to marine subtropical cloud-topped boundary layers. Part I: Description and 1D results. Monthly weather review, 132(4), 864-882. https://doi.org/10.1175/1520-0493(2004)132%3C0864:ANPFSC%3E2.0.CO;2

Brown, R. G., & Zhang, C. (1997). Variability of midtropospheric moisture and its effect on cloud-top height distribution during TOGA COARE. Journal of the atmospheric sciences, 54(23), 2760-2774. https://doi.org/10.1175/1520-0469(1997)054%3C2760:VOMMAI%3E2.0.CO;2

Chepfer, H., Bony, S., Winker, D., Chiriaco, M., Dufresne, J. L., & Sèze, G. (2008). Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model. Geophysical Research Letters, 35(15). https://doi.org/10.1029/2008GL034207

Cuxart, J., Bougeault, P., & Redelsperger, J. L. (2000). A turbulence scheme allowing for mesoscale and large-eddy simulations. Quarterly Journal of the Royal Meteorological Society, 126(562), 1-30. https://doi.org/10.1002/qj.49712656202

Daleu, C. L., Plant, R. S., Woolnough, S. J., Sessions, S., Herman, M. J., Sobel, A., ... & Peyrille, P. (2015). Intercomparison of methods of coupling between convection and large-scale circulation: 1. Comparison over uniform surface conditions. Journal of Advances in Modeling Earth Systems, 7(4), 1576-1601. https://doi.org/10.1002/2015MS000468

Davies, L., Plant, R. S., & Derbyshire, S. H. (2013). Departures from convective equilibrium with a rapidly varying surface forcing. Quarterly Journal of the Royal Meteorological Society, 139(676), 1731-1746. https://doi.org/10.1002/qj.2065

Deardorff, J. W. (1972). Theoretical expression for the countergradient vertical heat flux. Journal of Geophysical Research, 77(30), 5900-5904. https://doi.org/10.1029/JC077i030p05900

Del Genio, A. D., Wu, J., Wolf, A. B., Chen, Y., Yao, M. S., & Kim, D. (2015). Constraints on cumulus parameterization from simulations of observed MJO events. *Journal of Climate*, *28*(16), 6419-6442. https://doi.org/10.1175/JCLI-D-14-00832.1

DeMott, C. A., Randall, D. A., & Khairoutdinov, M. (2007). Convective precipitation variability as a tool for general circulation model analysis. Journal of Climate, 20(1), 91-112. https://doi.org/10.1175/JCLI3991.1

Derbyshire, S. H., Beau, I., Bechtold, P., Grandpeix, J. Y., Piriou, J. M., Redelsperger, J. L., & Soares, P. M. M. (2004). Sensitivity of moist convection to environmental humidity. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 130(604), 3055-3079. https://doi.org/10.1256/qj.03.130

Derbyshire, S. H., Maidens, A. V., Milton, S. F., Stratton, R. A., & Willett, M. R. (2011). Adaptive detrainment in a convective parametrization. Quarterly Journal of the Royal Meteorological Society, 137(660), 1856-1871. https://doi.org/10.1002/qj.875

Emanuel, K. A. (1993). A cumulus representation based on the episodic mixing model: the importance of mixing and microphysics in predicting humidity. In The representation of cumulus convection in numerical models (pp. 185-192). American Meteorological Society, Boston, MA. https://doi.org/10.1007/978-1-935704-13-3_19

Emanuel, K. A., & Živković-Rothman, M. (1999). Development and evaluation of a convection scheme for use in climate models. Journal of the Atmospheric Sciences, 56(11), 1766-1782. https://doi.org/10.1175/1520-0469(1999)056%3C1766:DAEOAC%3E2.0.CO;2

Emanuel, K. A., David Neelin, J., & Bretherton, C. S. (1994). On large-scale circulations in convecting atmospheres. Quarterly Journal of the Royal Meteorological Society, 120(519), 1111-1143. https://doi.org/10.1002/qj.49712051902

Fan, J., Han, B., Varble, A., Morrison, H., North, K., Kollias, P., ... & Lin, Y. (2017). Cloud-resolving model intercomparison of an MC3E squall line case: Part I—Convective updrafts. Journal of Geophysical Research: Atmospheres, 122(17), 9351-9378. https://doi.org/10.1002/2017JD026622

Frierson, D. M. (2007). The dynamics of idealized convection schemes and their effect on the zonally averaged tropical circulation. Journal of the atmospheric sciences, 64(6), 1959-1976. https://doi.org/10.1175/JAS3935.1

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G., & Yacalis, G. (2018). Could machine learning break the convection parameterization deadlock?. Geophysical Research Letters, 45(11), 5742-5751. https://doi.org/10.1029/2018GL078202

Grabowski, W. W. (2003). MJO-like coherent structures: Sensitivity simulations using the cloud-resolving convection parameterization (CRCP). Journal of the Atmospheric Sciences, 60(6), 847-864. https://doi.org/10.1175/1520-0469(2003)060%3C0847:MLCSSS%3E2.0.CO;2

Grandpeix, J. Y., & Lafore, J. P. (2010). A density current parameterization coupled with Emanuel's convection scheme. Part I: The models. Journal of the Atmospheric Sciences, 67(4), 881-897. https://doi.org/10.1175/2009JAS3044.1

Grandpeix, J. Y., Phillips, V., & Tailleux, R. (2004). Improved mixing representation in Emanuel's convection scheme. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *130*(604), 3207-3222. https://doi.org/10.1256/qj.03.144

Gregory, D. (1997). The Mass Flux Approach to the Parametrization of Deep Convection. In The Physics and Parameterization of Moist Atmospheric Convection (pp. 297-319). Springer, Dordrecht. https://doi.org/10.1007/978-94-015-8828-7_12

Gregory, D., & Miller, M. J. (1989). A numerical study of the parametrization of deep tropical convection. Quarterly Journal of the Royal Meteorological Society, 115(490), 1209-1241. https://doi.org/10.1002/qj.49711549003

Gregory, D., & Rowntree, P. R. (1990). A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. Monthly Weather Review, 118(7), 1483-1506. https://doi.org/10.1175/1520-0493(1990)118%3C1483:AMFCSW%3E2.0.CO;2

Grell, G. A., & Freitas, S. R. (2014). A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling. Atmos. Chem. Phys, 14(10), 5233-5250. https://doi.org/10.5194/acp-14-5233-2014

Guérémy, J. F. (2011). A continuous buoyancy based convection scheme: one-and three-dimensional validation. Tellus A: Dynamic Meteorology and Oceanography, 63(4), 687-706. https://doi.org/10.1111/j.1600-0870.2011.00521.x

Hacker, J. P., & Angevine, W. M. (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. https://doi.org/10.1175/MWR-D-12-00280.1

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. https://doi.org/10.1002/jame.20037

Heymsfield, A. J., & Donner, L. J. (1990). A scheme for parameterizing ice-cloud water content in general circulation models. Journal of the Atmospheric Sciences, 47(15), 1865-1877. https://doi.org/10.1175/1520-0469(1990)047%3C1865:ASFPIC%3E2.0.CO;2

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. https://doi.org/10.1175/MWR3199.1

Jakob, C., & Klein, S. A. (2000). A parametrization of the effects of cloud and precipitation overlap for use in general-circulation models. Quarterly Journal of the Royal Meteorological Society, 126(568), 2525-2544. https://doi.org/10.1002/qj.49712656809

Jam, A., Hourdin, F., Rio, C., & Couvreux, F. (2013). Resolved versus parametrized boundary-layer plumes. Part III: Derivation of a statistical scheme for cumulus clouds. Boundary-layer meteorology, 147(3), 421-441. https://doi.org/10.1007/s10546-012-9789-3

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. https://doi.org/10.1175/1520-0493(1994)122%3C0927:TSMECM%3E2.0.CO;2

Janjić, Z. I. (2000). Comments on “Development and evaluation of a convection scheme for use in climate models”. Journal of the Atmospheric Sciences, 57(21), 3686-3686. https://doi.org/10.1175/1520-0469(2000)057%3C3686:CODAEO%3E2.0.CO;2

Kain, J. S. (2004). The Kain–Fritsch convective parameterization: an update. Journal of applied meteorology, 43(1), 170-181. https://doi.org/10.1175/1520-0450(2004)043%3C0170:TKCPAU%3E2.0.CO;2

Kain, J. S., & Fritsch, J. M. (1990). A one-dimensional entraining/detraining plume model and its application in convective parameterization. Journal of the Atmospheric Sciences, 47(23), 2784-2802. https://doi.org/10.1175/1520-0469(1990)047%3C2784:AODEPM%3E2.0.CO;2

Khain, A. P., Beheng, K. D., Heymsfield, A., Korolev, A., Krichak, S. O., Levin, Z., ... & Van Den Heever, S. C. (2015). Representation of microphysical processes in cloud-resolving models: Spectral (bin) microphysics versus bulk parameterization. Reviews of Geophysics, 53(2), 247-322. https://doi.org/10.1002/2014RG000468

Kim, J. H., Shin, D. B., & Kummerow, C. (2014). Effect of microphysics scheme in cloud resolving models in passive microwave remote sensing of precipitation over ocean. EGUGA, 9934.

Kokoska, S., & Zwillinger, D. (2000). CRC standard probability and statistics tables and formulae. Crc Press.

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. https://doi.org/10.1175/2009JAS3260.1

Kwon, Y. C., & Hong, S. Y. (2017). A mass-flux cumulus parameterization scheme across gray-zone resolutions. Monthly Weather Review, 145(2), 583-598. https://doi.org/10.1175/MWR-D-16-0034.1

Lambert, F. H., Challenor, P. G., Lewis, N. T., McNeall, D. J., Owen, N., Boutle, I. A., ... & Webb, M. J. (2020). Continuous structural parameterization: A proposed method for representing different model parameterizations within one structure demonstrated for atmospheric convection. Journal of Advances in Modeling Earth Systems, 12(8), e2020MS002085. https://doi.org/10.1029/2020MS002085

Langhans, W., Schmidli, J., Fuhrer, O., Bieri, S., & Schär, C. (2013). Long-term simulations of thermally driven flows and orographic convection at convection-parameterizing and cloud-resolving resolutions. Journal of Applied Meteorology and Climatology, 52(6), 1490-1510. https://doi.org/10.1175/JAMC-D-12-0167.1

Laval, K., Sadourny, R., & Serafini, Y. (1981). Land surface processes in a simplified general circulation model. Geophysical & Astrophysical Fluid Dynamics, 17(1), 129-150. https://doi.org/10.1080/03091928108243677

Lebo, Z. J., & Morrison, H. (2015). Effects of horizontal and vertical grid spacing on mixing in simulated squall lines and implications for convective strength and structure. Monthly Weather Review, 143(11), 4355-4375. https://doi.org/10.1175/MWR-D-15-0154.1

Liu, C., & Moncrieff, M. W. (2007). Sensitivity of cloud-resolving simulations of warm-season convection to cloud microphysics parameterizations. Monthly weather review, 135(8), 2854-2868. https://doi.org/10.1175/MWR3437.1

Lock, A. P., Brown, A. R., Bush, M. R., Martin, G. M., & Smith, R. N. B. (2000). A new boundary layer mixing scheme. Part I: Scheme description and single-column model tests. Monthly weather review, 128(9), 3187-3199. https://doi.org/10.1175/1520-0493(2000)128%3C3187:ANBLMS%3E2.0.CO;2

Lopez, P. (2002). Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and data-assimilation purposes. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 128(579), 229-257. https://doi.org/10.1256/00359000260498879

Manabe, S., & Strickler, R. F. (1964). Thermal equilibrium of the atmosphere with a convective adjustment. Journal of the Atmospheric Sciences, 21(4), 361-385. https://doi.org/10.1175/1520-0469(1964)021%3C0361:TEOTAW%3E2.0.CO;2

Mapes, B. E. (1997). Equilibrium vs. activation control of large-scale variations of tropical deep convection. In The physics and parameterization of moist atmospheric convection (pp. 321-358). Springer, Dordrecht. https://doi.org/10.1007/978-94-015-8828-7_13

Mauritsen, T., Stevens, B., Roeckner, E., Crueger, T., Esch, M., Giorgetta, M., ... & Mikolajewicz, U. (2012). Tuning the climate of a global model. Journal of advances in modeling Earth systems, 4(3). https://doi.org/10.1029/2012MS000154

Morrison, H., & Gettelman, A. (2008). A new two-moment bulk stratiform cloud microphysics scheme in the Community Atmosphere Model, version 3 (CAM3). Part I: Description and numerical tests. Journal of Climate, 21(15), 3642-3659. https://doi.org/10.1175/2008JCLI2105.1

Neggers, R. A. J., Ackerman, A. S., Angevine, W. M., Bazile, E., Beau, I., Blossey, P. N., ... & Fletcher, J. (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. https://doi.org/10.1002/2017MS001064

O'Gorman, P. A., & Dwyer, J. G. (2018). Using machine learning to parameterize moist convection: Potential for modeling of climate, climate change, and extreme events. Journal of Advances in Modeling Earth Systems, 10(10), 2548-2563. https://doi.org/10.1029/2018MS001351

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. https://doi.org/10.1175/2008JCLI2557.1

Park, S., Bretherton, C. S., & Rasch, P. J. (2014). Integrating cloud processes in the Community Atmosphere Model, version 5. Journal of Climate, 27(18), 6821-6856. https://doi.org/10.1175/JCLI-D-14-00087.1

Parsons, D. B., Redelsperger, J. L., & Yoneyama, K. (2000). The evolution of the tropical western Pacific atmosphere-ocean system following the arrival of a dry intrusion. Quarterly Journal of the Royal Meteorological Society, 126(563), 517-548. https://doi.org/10.1002/qj.49712656307

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. https://doi.org/10.1002/2016GL071285

Piriou, J. M., Redelsperger, J. L., Geleyn, J. F., Lafore, J. P., & Guichard, F. (2007). An approach for convective parameterization with memory: Separating microphysics and transport in grid-scale equations. Journal of the Atmospheric Sciences, 64(11), 4127-4139. https://doi.org/10.1175/2007JAS2144.1

Raymond, D. J. (1995). Regulation of moist convection over the west Pacific warm pool. Journal of the atmospheric sciences, 52(22), 3945-3959. https://doi.org/10.1175/1520-0469(1995)052%3C3945:ROMCOT%3E2.0.CO;2

Redelsperger, J. L., Parsons, D. B., & Guichard, F. (2002). Recovery processes and factors limiting cloud-top height following the arrival of a dry intrusion observed during TOGA COARE. Journal of the atmospheric sciences, 59(16), 2438-2457. https://doi.org/10.1175/1520-0469(2002)059%3C2438:RPAFLC%3E2.0.CO;2

Rennó, N. O., Emanuel, K. A., & Stone, P. H. (1994). Radiative-convective model with an explicit hydrologic cycle: 1. Formulation and sensitivity to model parameters. Journal of Geophysical Research: Atmospheres, 99(D7), 14429-14441. Ricard, J. L., & Royer, J. F. (1993). A statistical cloud scheme for use in an AGCM. In Annales Geophysicae (Vol. 11, pp. 1095-1115). https://doi.org/10.1029/94JD00020

Ricard, J. L., & Royer, J. F. (1993). A statistical cloud scheme for use in an AGCM. AnGeo, 11, 1095-1115.

Rio, C., Del Genio, A. D., & Hourdin, F. (2019). Ongoing Breakthroughs in Convective Parameterization. Current Climate Change Reports, 5(2), 95-111. https://doi.org/10.1007/s40641-019-00127-w

Rio, C., Grandpeix, J. Y., Hourdin, F., Guichard, F., Couvreux, F., Lafore, J. P., ... & Lefebvre, M. P. (2013). Control of deep convection by sub-cloud lifting processes: the ALP closure in the LMDZ5B general circulation model. Climate dynamics, 40(9-10), 2271-2292. https://doi.org/10.1007/s00382-012-1506-x

Rio, C., Hourdin, F., Grandpeix, J. Y., & Lafore, J. P. (2009). Shifting the diurnal cycle of parameterized deep convection over land. Geophysical Research Letters, 36(7). https://doi.org/10.1029/2008GL036779

Rio, C., Hourdin, F., Couvreux, F., & Jam, A. (2010). Resolved versus parametrized boundary-layer plumes. Part II: Continuous formulations of mixing rates for mass-flux schemes. Boundary-layer meteorology, 135(3), 469-483. https://doi.org/10.1007/s10546-010-9478-z

Roehrig, R., Beau, I., Saint-Martin, D., Alias, A., Decharme, B., Guérémy, J. F., ... & Blein, S. (2020). The CNRM global atmosphere model ARPEGE-Climat 6.3: Description and evaluation. Journal of Advances in Modeling Earth Systems, 12(7), e2020MS002075. https://doi.org/10.1029/2020MS002075

Sherwood, S. C., Minnis, P., & McGill, M. (2004). Deep convective cloud-top heights and their thermodynamic control during CRYSTAL-FACE. Journal of Geophysical Research: Atmospheres, 109(D20). https://doi.org/10.1029/2004JD004811

Singh, M. S., & O'Gorman, P. A. (2015). Increases in moist-convective updraught velocities with warming in radiative-convective equilibrium. Quarterly Journal of the Royal Meteorological Society, 141(692), 2828-2838. https://doi.org/10.1002/qj.2567

Sobel, A. H., & Bretherton, C. S. (2000). Modeling tropical precipitation in a single column. Journal of climate, 13(24), 4378-4392. https://doi.org/10.1175/1520-0442(2000)013%3C4378:MTPIAS%3E2.0.CO;2

Suhas, E., & Zhang, G. J. (2014). Evaluation of trigger functions for convective parameterization schemes using observations. Journal of Climate, 27(20), 7647-7666. https://doi.org/10.1175/JCLI-D-13-00718.1

Sundqvist, H. (1978). A parameterization scheme for non-convective condensation including prediction of cloud water content. Quarterly Journal of the Royal Meteorological Society, 104(441), 677-690. https://doi.org/10.1002/qj.49710444110

Takemi, T., Hirayama, O., & Liu, C. (2004). Factors responsible for the vertical development of tropical oceanic cumulus convection. Geophysical research letters, 31(11). https://doi.org/10.1029/2004GL020225

Tawfik, A. B., & Dirmeyer, P. A. (2014). A process-based framework for quantifying the atmospheric preconditioning of surface-triggered convection. Geophysical Research Letters, 41(1), 173-178. https://doi.org/10.1002/2013GL057984

Tompkins, A. M. (2001). Organization of tropical convection in low vertical wind shears: The role of water vapor. Journal of the atmospheric sciences, 58(6), 529-545. https://doi.org/10.1175/1520-0469(2001)058%3C0529:OOTCIL%3E2.0.CO;2

Tulich, S. N., & Mapes, B. E. (2010). Transient environmental sensitivities of explicitly simulated tropical convection. Journal of the atmospheric sciences, 67(4), 923-940. https://doi.org/10.1175/2009JAS3277.1

Varble, A., Zipser, E. J., Fridlind, A. M., Zhu, P., Ackerman, A. S., Chaboureau, J. P., ... & Shipway, B. (2014). Evaluation of cloud-resolving and limited area model intercomparison simulations using TWP-ICE observations: 1. Deep convective updraft properties. *Journal of Geophysical Research: Atmospheres*, *119*(24), 13-891. https://doi.org/10.1002/2013JD021371

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. https://doi.org/10.1007/s00382-013-1725-9

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., ... & Morcrette, C. (2019). The Met Office Unified Model global atmosphere 7.0/7.1 and JULES global land 7.0 configurations. Geoscientific Model Development, 12(5), 1909-1963. https://doi.org/10.5194/gmd-12-1909-2019

Wang, X., & Zhang, M. (2013). An analysis of parameterization interactions and sensitivity of single-column model simulations to convection schemes in CAM4 and CAM5. Journal of

Geophysical Research: Atmospheres, 118(16), 8869-8880. https://doi.org/10.1002/jgrd.50690

Wilson, D. R., & Ballard, S. P. (1999). A microphysically based precipitation scheme for the UK Meteorological Office Unified Model. Quarterly Journal of the Royal Meteorological Society, 125(557), 1607-1636. https://doi.org/10.1002/qj.49712555707

Wilson, D. R., Bushell, A. C., Kerr-Munslow, A. M., Price, J. D., & Morcrette, C. J. (2008). PC2: A prognostic cloud fraction and condensation scheme. I: Scheme description. Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 134(637), 2093-2107. https://doi.org/10.1002/qj.333

Wing, A. A., Stauffer, C. L., Becker, T., Reed, K. A., Ahn, M. S., Arnold, N. P., ... & de Roode, S. R. (2020). Clouds and Convective Self-Aggregation in a Multi-Model Ensemble of Radiative-Convective Equilibrium Simulations. Journal of Advances in Modeling Earth Systems, e2020MS002138. https://doi.org/10.1029/2020MS002138

Wolding, B., Dias, J., Kiladis, G., Ahmed, F., Powell, S. W., Maloney, E., & Branson, M. (2020). Interactions between moisture and tropical convection. Part I: The coevolution of moisture and convection. Journal of the Atmospheric Sciences, 77(5), 1783-1799. https://doi.org/10.1175/JAS-D-19-0225.1

Yamada, T. (1983). Simulations of nocturnal drainage flows by a q^2 / turbulence closure model. Journal of the Atmospheric Sciences, 40(1), 91-106. https://doi.org/10.1175/1520-0469(1983)040%3C0091:SONDFB%3E2.0.CO;2

Yanai, M., Esbensen, S., & Chu, J. H. (1973). Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. Journal of the Atmospheric Sciences, 30(4), 611-627. https://doi.org/10.1175/1520-0469(1973)030%3C0611:DOBPOT%3E2.0.CO;2

Yano, J. I., & Plant, R. S. (2012). Convective quasi-equilibrium. Reviews of Geophysics, 50(4). https://doi.org/10.1029/2011RG000378

Zender, C. S., & Kiehl, J. T. (1997). Sensitivity of climate simulations to radiative effects of tropical anvil structure. *Journal of Geophysical Research: Atmospheres*, *102*(D20), 23793-23803. https://doi.org/10.1029/97JD02009

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. https://doi.org/10.1175/JCLI-D-16-0597.1

Zhang, C., Wang, Y., & Hamilton, K. (2011). Improved representation of boundary layer clouds over the southeast Pacific in ARW-WRF using a modified Tiedtke cumulus parameterization scheme. Monthly Weather Review, 139(11), 3489-3513. https://doi.org/10.1175/MWR-D-10-05091.1

Zhang, G. J., & 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. https://doi.org/10.1080/07055900.1995.9649539

Zhang, G. J., & Mu, M. (2005). Effects of modifications to the Zhang-McFarlane convection parameterization on the simulation of the tropical precipitation in the National Center for Atmospheric Research Community Climate Model, version 3. Journal of Geophysical Research: Atmospheres, 110(D9). https://doi.org/10.1029/2004JD005617

Zhang, Y., Klein, S. A., Boyle, J., & Mace, G. G. (2010). Evaluation of tropical cloud and precipitation statistics of Community Atmosphere Model version 3 using CloudSat and CALIPSO data. Journal of Geophysical Research: Atmospheres, 115(D12). https://doi.org/10.1029/2009JD012006