JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2025MS005340

Responses to Humidity and Temperature Perturbations in High-Resolution Simulations of Convection

Timothy H. Raupach[1,2,3,4], Chimene L. Daleu[5], Robert S. Plant[5], Steven C. Sherwood[2,3,4], and Yi-Ling Hwong[6]

[1]Institute for Climate Risk and Response, University of New South Wales, Sydney, NSW, Australia, [2]Climate Change Research Centre, University of New South Wales, Sydney, NSW, Australia, [3]ARC Centre of Excellence for Climate Extremes, Sydney, NSW, Australia, [4]ARC Centre of Excellence for 21st Century Weather, Sydney, NSW, Australia, [5]Department of Meteorology, University of Reading, Reading, UK, [6]International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

Abstract Convection parameterization is a leading source of uncertainty in global and regional climate models, and a lack of ground truth complicates the assessment of convection scheme performance. Here we test a linear framework for quantifying convective responses, using two models run at convection-permitting resolution, to examine model responses to temperature and humidity perturbations. For the models examined, a grid spacing finer than ∼1 km was required for consistent (hence potentially accurate) responses, implying good representation of convectively-coupled dynamics. The results from the tests at 100 and 250 m grid spacing could reasonably be used in idealized tests of convective schemes, as their spread is mostly small compared to the spread of previously reported scheme behavior.

Plain Language Summary In the atmosphere, convection is heat-driven movement of air. Convection forces changes in water state, causing clouds and rain to form, and thus further moving heat around. Weather and climate models must account for these effects. A problem arises because models are often run with spacings between modeled (grid) points that are larger than the scale of typical convective cells. In this case, models resort to simplified approximations, called convection schemes, to represent convection effects and correct for not explicitly simulating convective cells. It is difficult to assess the performance of these schemes, because there is no reliable ground truth that says exactly how the schemes should perform. Here we study how schemes "should" perform by examining how two models react to small changes in temperature and humidity when they are run at fine-enough resolution that no convective scheme is required. We show that models require a grid spacing of under ∼1 km to properly capture the effects of convection, and propose that the responses at the finest resolutions could reasonably be used to test convection schemes.

Key Points:

- Time-averaged responses to local humidity and temperature perturbations in idealized models at cloud-resolving resolutions were studied
- For the two models examined, accurate simulation of convectively-coupled dynamics required a grid spacing lower than ∼1 km
- The linear responses of large eddy simulation models in these idealized tests may facilitate Global Climate Model scheme testing

Correspondence to:
T. H. Raupach,
timothy.h.raupach@gmail.com

Citation:
Raupach, T. H., Daleu, C. L., Plant, R. S., Sherwood, S. C., & Hwong, Y.-L. (2026). Responses to humidity and temperature perturbations in high-resolution simulations of convection. *Journal of Advances in Modeling Earth Systems*, *18*, e2025MS005340. https://doi.org/10.1029/2025MS005340

Received 7 JUL 2025
Accepted 24 MAR 2026

Author Contributions:
Conceptualization: Timothy H. Raupach, Chimene L. Daleu, Robert S. Plant, Steven C. Sherwood, Yi-Ling Hwong
Formal analysis: Timothy H. Raupach, Chimene L. Daleu, Robert S. Plant, Steven C. Sherwood, Yi-Ling Hwong
Methodology: Timothy H. Raupach, Chimene L. Daleu, Robert S. Plant, Steven C. Sherwood, Yi-Ling Hwong
Software: Timothy H. Raupach, Chimene L. Daleu
Writing – original draft: Timothy H. Raupach, Chimene L. Daleu, Robert S. Plant, Steven C. Sherwood, Yi-Ling Hwong

1. Introduction

Global and regional climate models (respectively GCMs and RCMs), when run at grid spacings larger than the horizontal scale of convection, cannot resolve individual convective cells (e.g., Bryan et al., 2003) and use convection schemes to parameterize convection effects (e.g., Arakawa, 2004; Lin et al., 2022; Rio et al., 2019). Models are highly sensitive to convective cloud physics (e.g., Emanuel & Živković Rothman, 1999), and convection parameterization is a leading source of model uncertainty (e.g., Hwong et al., 2021, hereafter H21). It is difficult to compare convection schemes, since traditional comparisons of simulations with observations (e.g., Grell & Freitas, 2014; Kwon & Hong, 2017; Zhang & Wang, 2017; Zhang et al., 2011) are limited by observation selection and accuracy (H21).

Kuang (2010), hereafter K10, suggested that although there are many non-linear processes in convection, the statistics of a cumulus ensemble can be represented by smooth linear functions that describe its reactions to small changes in the large-scale environment. K10's mathematical framework uses a linear tangent model to approximate the feedback of small-scale processes on the large-scale environment, effectively playing the role of a (linear) convective scheme trained on a large cloud-resolving model (CRM) data set, such that

$$\frac{\partial \mathbf{X}}{\partial t} = \mathbf{M}\mathbf{X}, \tag{1}$$

Writing – review & editing: Timothy H. Raupach, Chimene L. Daleu, Robert S. Plant, Steven C. Sherwood, Yi-Ling Hwong

where the partial time derivative represents the feedback on the column state variable **X**, and **M** is a matrix obtained from perturbed runs of a cloud-resolving model. In this method, the responses of the model (and thus the convection scheme) are examined in idealized simulations under small perturbations of the humidity and temperature tendencies. K10 showed that linear responses sufficiently captured the dynamics of convectively-coupled waves, which evolve slowly compared to typical convective response times and have relatively small amplitudes. Even if a linear approximation cannot fully describe convection, it can be a useful first-order characterization relevant to scheme evaluation, since current GCMs often fail to correctly represent coupled waves (Ahn et al., 2020; Khouider et al., 2011).

Herman and Kuang (2013) applied the linear response framework to single column models (SCMs) to test two convection schemes against a CRM reference. H21 used the framework to evaluate convection, planetary boundary layer (PBL), and microphysics schemes' responses to perturbations in SCMs, and established that the response is dominated by the choice of convection scheme. Herman and Kuang (2013) had identified a wide range of responses among current schemes, showing the potential of the method, but did not examine which setup was the most accurate because that would require a "truth" response which is not available from observations. K10 evaluated his linear responses using rather coarse resolution (2 km) runs of the System for Atmospheric Modeling (SAM, e.g. Khairoutdinov & Randall, 2003), and in a relatively small domain (128 km square). This resolution would not be considered large-eddy resolving by today's standards, and the experiment has not been repeated in any other cloud-resolving model. A study of high-resolution linear responses of a model that uses the least possible parameterization is therefore required, along with an evaluation of robustness.

In this study, we ran perturbation experiments with two different models across a range of CRM and large eddy simulation (LES) resolutions, seeking to produce benchmark test results. This enabled us to test the conclusion of H21 that results are not sensitive to the treatment of the PBL for configurations in which convection is resolved explicitly. The key research question (RQ) is whether the linear responses are robust, or sensitive to decisions made in high-resolution modeling such as the grid length, model dynamical core, or treatment of microphysics or eddy mixing. In particular, can we establish a reference "truth" that is less uncertain than the variations between convection schemes themselves (RQ1)? A second question addressed here, which is becoming more relevant as we enter an age of global CRMs, is what grid resolution they would require to properly represent convective-scale feedbacks on large-scale flows—arguably the key benefit of global CRMs—and how sensitive the feedback might be to parameterizations remaining in those models or to the numerics or other details of the CRM (RQ2). In summary, here we use CRM and LES models to address how we should judge the differences between model schemes reported in H21.

2. Methods

We used two numerical models in this work: the Advanced Research version of the Weather Research and Forecasting (WRF) model (Skamarock et al., 2021), version 4.1.4, and the Met Office/NERC Cloud Model (MONC, Brown et al., 2020) version 0.7. We assess results by comparing the model responses with each other and with the original results of K10.

2.1. Model Settings

Model settings are summarized in Table 1. We generally followed the model settings of H21 and MONC settings were based on Daleu et al. (2023). Both models were used to simulate convection at high resolution over a water surface with periodic boundary conditions and constant sea surface temperature (SST), horizontal surface wind, and surface exchange coefficient. Radiation was idealized (Herman & Kuang, 2013) with a radiative cooling tendency of $\partial Q_{rad}/\partial t = Q_t/\Pi$ K day^{-1}, where Π is the Exner function and Q_t is a temperature tendency. Level-mean horizontal winds were relaxed to constant values. Following Chua et al. (2019), evaporation was idealized. Surface heat flux (SH) and latent heat flux (LH) were calculated as

$$\mathrm{SH} = 0.001\rho_a W_s c_p (T_s - T_a), \tag{2}$$

$$\mathrm{LH} = 0.001\rho_a W_s L\left(q_{sat} - q_a\right), \tag{3}$$

Table 1
Model Settings Per Model

	WRF		MONC
Grid spacings	4 km, 1 km	100 m	1 km, 500 m, 250 m
Vertical levels	74	370	85
PBL scheme	Enabled	Disabled	NA
Model time step (s)	6	1	1–2 (variable)
Output frequency (hr)	1	2	1
Initial profile source	RCEMIP profile Wing et al. (2018)	Mean from 1 km WRF runs	RCE profile from Daleu et al. (2015)
Value of `km_opt`	4	2	NA
Domain height (m)	33,250	32,809	20,000
Boundary-layer scheme	YSU (Hong et al., 2006)	None	NA
SST (K)	301.15		301.15
Surface wind speed (m s^{-1})	5		5
Surface exchange coefficient	0.0005		0.0005
Domain size (km^2)	20×20		64×64
Q_t (K day^{-1})	−1.5 below 200 hPa 0 above 100 hPa linear from 100 to 200 hPa		−1.5 below 12 km height 0 above 16 km height linear from 12 to 16 km height
Level-mean u target (m s^{-1})	0		0
Level-mean v target (m s^{-1})	5		5
u, v relaxation time (hr)	3		3
Microphysics scheme	Thompson 2-moment (Thompson et al., 2008)		CASIM (Field et al., 2023)
Surface-layer scheme	Revised MM5 (Jiménez et al., 2012)		Monin-Obukhov similarity (Bull & Derbyshire, 1990)
Sub-grid eddy flux scheme	Prognostic TKE at 100 m (Skamarock et al., 2021)		3D Smagorinsky (Lilly, 1967; Smagorinsky, 1963)
Convection scheme	None		None

Note. PBL stands for Planetary Boundary Layer. RCEMIP refers to the Radiative-Convective Equilibrium Model Intercomparison Project. SST stands for sea surface temperature. u and v are orthogonal horizontal wind components. See Skamarock et al. (2021) for `km_opt` diffusion options details.

where ρ_a (kg m^{-3}) is the near surface air density (we used the model-produced value of around 1.16 kg m^{-3} rather than the value of 1 kg m^{-3} assumed by Chua et al. (2019)), W_s (m s^{-1}) is surface wind speed, c_p (J kg^{-1} K^{-1}) is the specific heat capacity of dry air at constant pressure, T_s and T_a (both K) are the surface temperature (SST) and near- surface air temperature, respectively, L (J kg^{-1}) is the latent heat of vapourization of water, and q_{sat} and q_a (both kg kg^{-1}) are the mixing ratios of, respectively, saturated water vapor at T_s (in WRF we used "ground saturated mixing ratio") and water vapor at the lowest model level. To calculate near- surface air temperature we used

$$T_a = T_l \left(\frac{p_s}{p_l}\right)^{\frac{R_d}{c_p}}, \tag{4}$$

where T_l (K) is the temperature at the first model level above the surface, p_s and p_l (both hPa) are surface pressure and pressure at the first model level, respectively, and R_d (J kg^{-1} K^{-1}) is the gas constant for dry air. Temperature and water vapor mixing ratio in the stratosphere (above 160 hPa) were relaxed to initial profile values to avoid model drift (Herman & Kuang, 2013).

Because the WRF pressure levels were not exactly at the points around which we wanted to perturb, we adapted the form of perturbations used by Herman and Kuang (2013) so that only the Gaussian portion of the perturbation

Table 2
Perturbation Simulations Conducted

Model			WRF		MONC		
Grid spacing							
Variable	Perturbation amplitude	Level (hPa)	4 and 1 km	100 m	1 km	500 m	250 m
T	+0.5 K day^{-1}	415	•	•	•		•
		500	•		•	•	•
		600	•		•		•
		730	•		•		•
		850	•	•	•	•	•
	−0.5 K day^{-1}	415	•	•	•		•
		500	•		•		•
		600	•		•		•
		730	•		•		•
		850	•	•	•		•
q	+0.0002 kg kg^{-1} day^{-1}	415	•	•	•		
		500	•		•	•	•
		600	•		•		
		730	•		•		
		850	•	•	•	•	•
	−0.0002 kg kg^{-1} day^{-1}	415	•	•	•		
		500	•		•		
		600	•		•		
		730	•		•		
		850	•	•	•		

Note. A • symbol indicates that the given model was run at the given grid spacing with the given perturbation level. T is temperature (K) and q is water vapor mixing ratio (kg kg^{-1}). "415 hPa" refers to 412 hPa in the WRF runs.

function was used, and instead of selecting a level k, a pressure p_p (hPa) around which to perturb was selected. The perturbation function used for the ith model level in WRF was

$$\frac{d}{dt}\chi = A \exp\left[-\left(\frac{p_p - p_i}{75\,\text{hPa}}\right)^2\right], \tag{5}$$

where p_i (hPa) is the pressure at level i, χ is the variable subject to the extra forcing, and A is the amplitude of forcing. MONC perturbations were made using a combination of the delta and Gaussian functions (Herman & Kuang, 2013).

2.2. Perturbations to Temperature and Humidity

Following K10, for each resolution, the models were run until they reached radiative convective equilibrium (RCE). Small perturbations in the imposed temperature or moisture forcings were then introduced, while a control was continued with no perturbations. The simulations were then run to RCE once more.

Perturbations were constant in time and applied to potential temperature or water vapor tendencies at given levels. Given computational constraints, not all perturbations were applied for all of the resolutions (Table 2). Based on the WRF results of H21, we determined that the two perturbation levels that provided the most information about perturbations across a range of other levels were 850 and 412 hPa. In MONC, it was convenient to displace the highest perturbation level to 415 hPa to match with the model grid, and for simplicity we refer to the highest perturbation level in both models as 415 hPa.

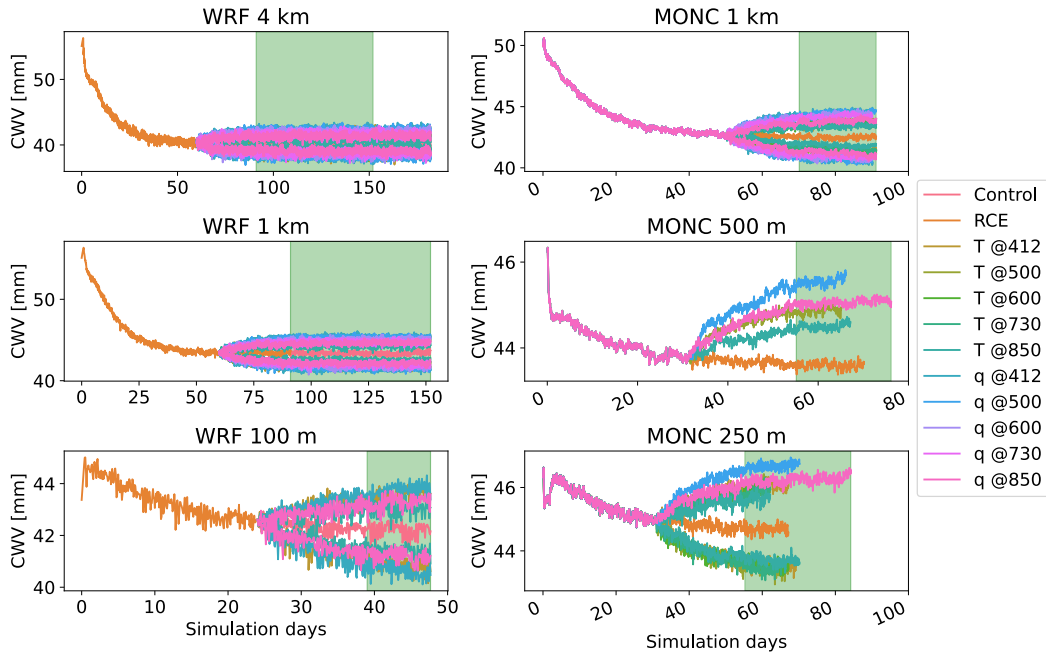

Figure 1. Time series of spatial mean column water vapor by grid spacing in Weather Research and Forecasting (WRF) and Met Office/NERC Cloud (MONC) models. "RCE periods" over which mean profiles were calculated for comparison are highlighted in green. In MONC runs, radiative convective equilibrium periods were the last 20 (10) days in each simulation for 1 km (500 m, 250 m) runs. Positive and negative perturbations are shown in the same color.

3. Results

3.1. Radiative-Convective Equilibrium Mean State

We used time series of spatially-averaged columnar water vapor (Figure 1) to determine when simulations had reached RCE. Mean profiles from control runs in RCE (Figure 2) show that mean thermodynamic states were similar for the two models, especially compared to the spread of RCE states in single-column models with parameterized physics (e.g., H21, Wing et al., 2018), with smoother wind profiles for the lower resolution runs. Hydrometeor profiles were sensitive to resolution and model, with WRF producing more cloud liquid water and snow, and less graupel, than MONC. Finer resolution runs generally produced more cloud water and snow, and differences between resolutions were more pronounced for WRF than MONC. It is uncertain whether these microphysical changes are a consequence of mean moisture changes or reflect cloud turbulence differences.

3.2. Convective Organization

We examined the spatial variance of precipitable water scaled by its spatial mean for both models (not shown), and for MONC we examined the "subsidence fraction," the fractional area of the domain where the vertically-integrated mass-weighted vertical wind is negative (not shown). There was no evidence of meaningful organization, as is expected when radiation is fixed everywhere and there are no contrasting radiative cooling profiles in moist and dry regions (Muller & Bony, 2015).

3.3. Responses to Perturbations

For perturbations at 415 and 850 hPa, respectively, the responses to heating are shown in Figures 3 and 4 and responses to moistening are shown in Figures 5 and 6. Results for perturbations at other levels (500, 600, and 730 hPa) are shown in the appendix (Figures A1–A3 for heating, Figures A4–A6 for moistening). All perturbation responses are much smaller than the spread in mean profiles across different resolution resolutions (Figure 2), indicating that their linear responses are less sensitive to model resolution than their mean states. This implies that the signal from the linear perturbations is cleaner than the background mean state differences. WRF and MONC produced responses of similar amplitude, except for cloud liquid water and ice where the changes in MONC were

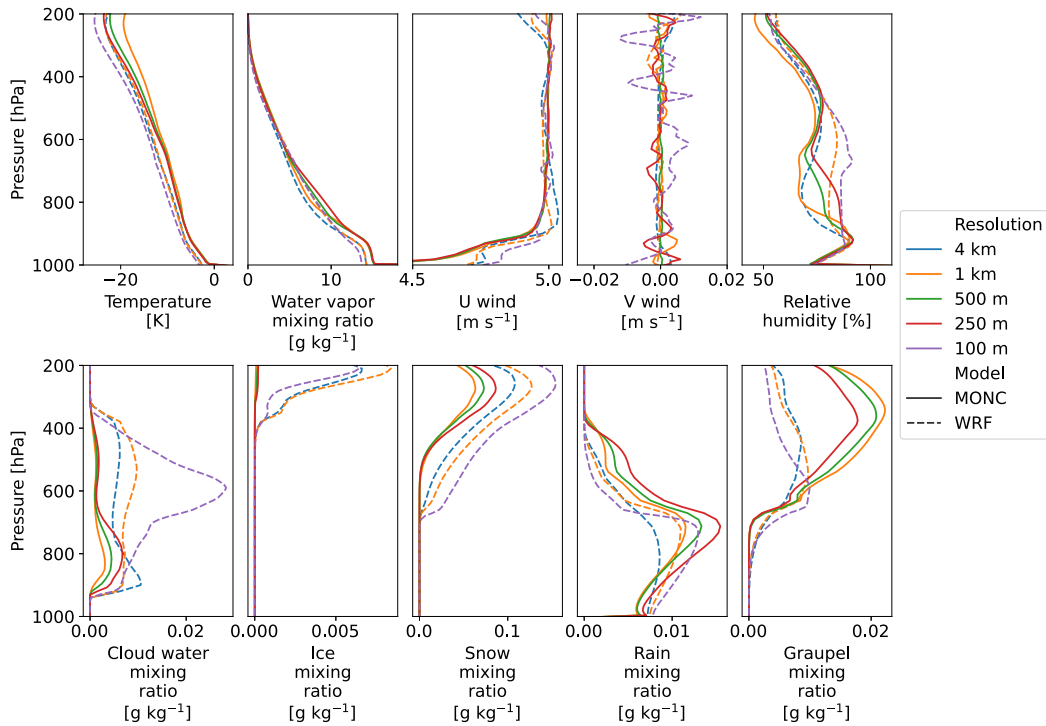

Figure 2. Mean profiles for selected variables in the control runs' radiative convective equilibrium periods, by grid spacing. All plots stop at 200 hPa (a thin cirrus layer that formed at the tropopause in the Met Office/NERC Cloud simulations is not shown). Temperatures are shown as anomalies from a reference moist pseudo-adiabat calculated using a surface temperature of 300 K.

smaller. Uncertainty on instantaneous domain-average responses was large, especially for hydrometeors (Figures A7–A10), but uncertainty reduced with longer averaging times over the RCE period (not shown). Further, most of the experiments show approximately linear behavior, with similar responses per unit forcing regardless of forcing sign, which indicates a certain stability in the responses, although some strong nonlinearities are found in the hydrometeor and relative humidity (RH) responses for finer resolutions, which is probably attributable to noise given the shorter duration and high variability of responses in the 100 m experiment.

3.3.1. Responses to Heating

Heating caused local increases in temperature and moisture around the perturbation level. From the perturbation level, temperature changes decreased down to the surface, and increased above by transmission of heat through a warming pseudoadiabat (K10, H21). There was a “dog leg” in the temperature response around the perturbed level that is consistent with local stabilization of the atmosphere inhibiting upward heat transport. WRF generally warmed more than MONC for upper-level perturbations, but less than MONC in the upper troposphere for the lowest perturbation level of 850 hPa.

Heating induced an increase in water vapor mixing ratio at the perturbation level, which decreased above and increased roughly linearly down to the surface. Lower resolution MONC runs showed less linearity in changes than the other runs. Heating of coarser resolution runs with perturbations at higher levels produced a minimum in the moisture increases near the top of the boundary layer which we hypothesize may be due to under-resolved shallow convection. Relative humidity (RH) around the perturbation level was generally increased by heating in the mid-upper troposphere and decreased when the heating was at or below 600 hPa, with decreases just above the perturbation level in coarser-resolution runs attributable to heating-induced stabilization. For finer resolution runs, low-level minima in moisture responses to upper-level heating were produced by both models and are not evident in K10. Moisture responses also showed a steeper gradient around 900–950 hPa than was shown by K10. These differences suggest that while the relatively coarse K10 simulations were surprisingly accurate overall they

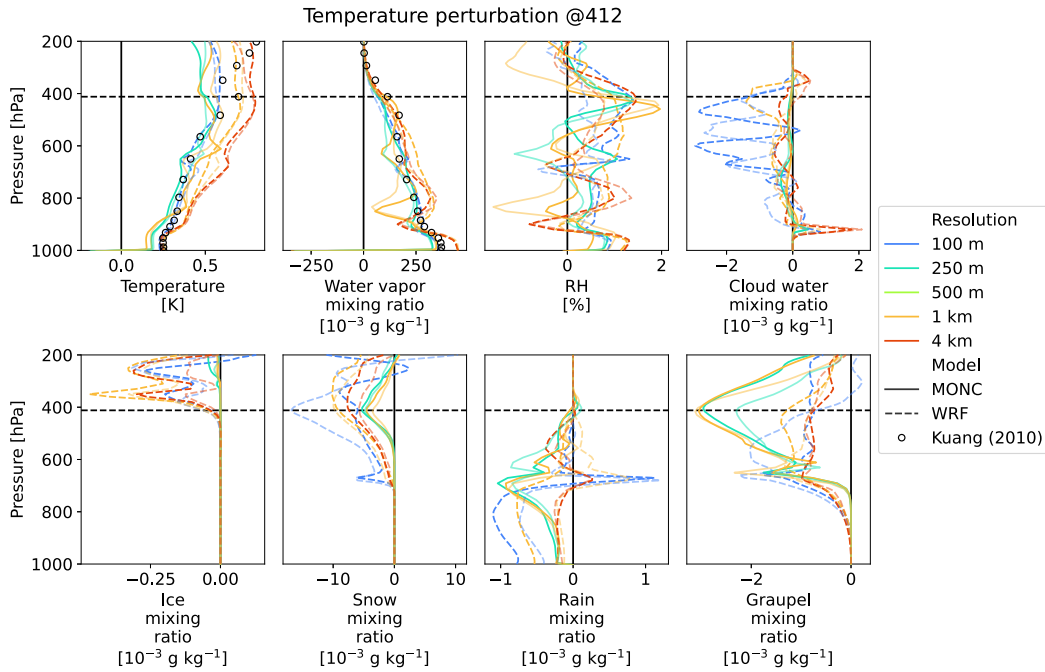

Figure 3. Responses by pressure for a temperature perturbation forcing of ±0.5 K at 415 hPa. The vertical black line shows zero difference, the dashed line shows the approximate level of maximum perturbation amplitude. Response lines are opaque (transparent) for positive (negative) perturbations. Negative perturbation responses are multiplied by −1 so they overlay positive perturbation responses if the two are symmetrical (indicating perfect linearity). Black circles show K10's responses to the same perturbation.

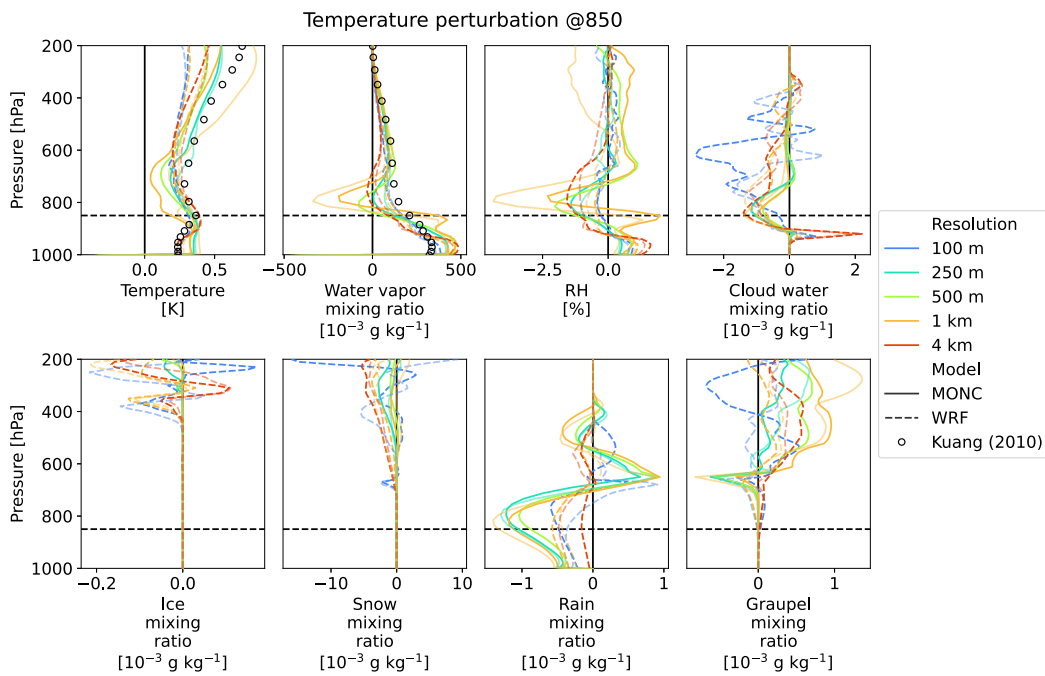

Figure 4. As in Figure 3, but for a temperature perturbation at 850 hPa.

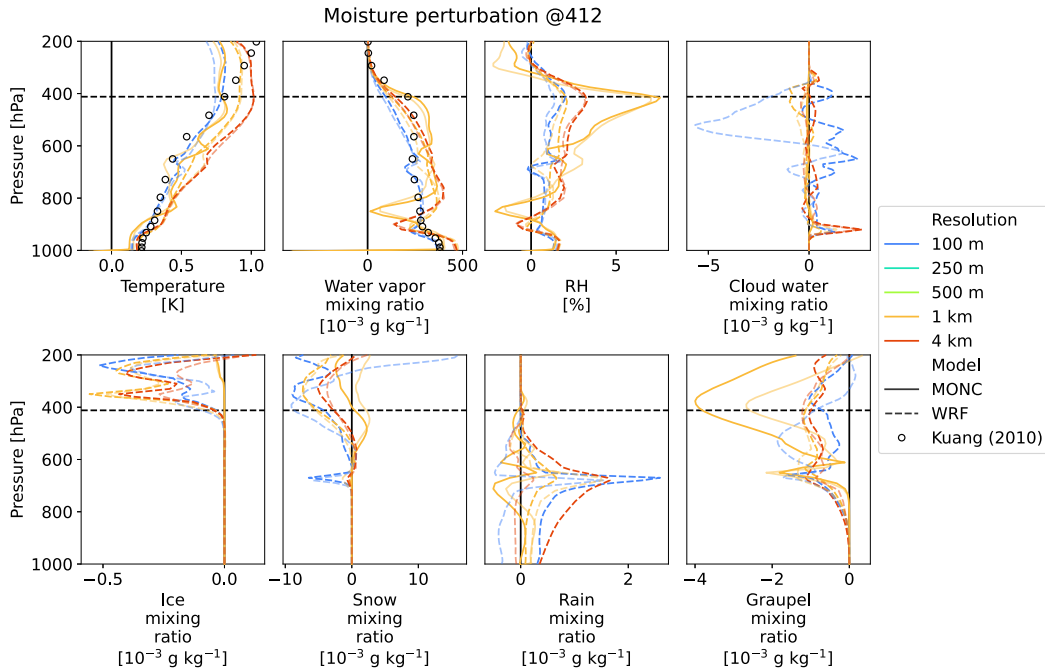

Figure 5. As for Figure 3, but for differences in model values by pressure, after a perturbation forcing of amplitude 0.2 g kg^{-1} was introduced in the water vapor mixing ratio tendency field at 415 hPa.

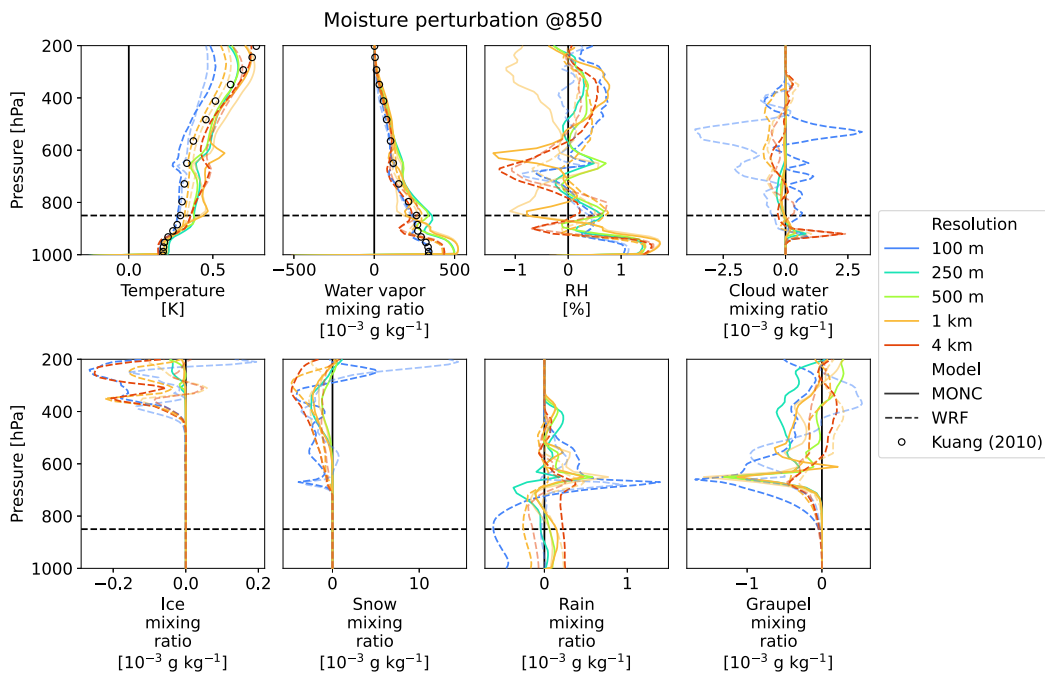

Figure 6. As in Figure 5, but for a water vapor mixing ratio perturbation at 850 hPa.

were unable to resolve structural changes to boundary layer moisture, due perhaps to humid convective downdrafts or changes to boundary-layer mixing. Responses in temperature and water vapor mixing ratio in the low-mid troposphere tended to be closer to those of K10 when the resolution was finer.

Changes in RH throughout the column varied but were more consistent between models for heating at lower levels. Hydrometeor mixing ratio responses were noisy and showed large variability between runs (Figures A7 and A8). The broad character of hydrometeor changes, however, was similar between the two tested models, with differences presumably attributable to differences in microphysics schemes. In general, heating reduced cloud liquid water (except for an increase in the boundary layer), decreased rain water content below the perturbation level and increased rain water above, and decreased graupel content except when the perturbation was at the lowest level. There were reductions in snow and increases in graupel at upper levels but to a highly variable degree.

3.3.2. Responses to Moistening

Moistening caused similar changes in temperature and moisture content as heating, but naturally with more local moisture increase around the perturbation level and without the pronounced "dog leg" in temperature response that heating provoked. With moistening perturbations, temperature increases in the boundary layer were consistent between resolutions in WRF, while those in the free troposphere were more consistent from moistening than from heating perturbations but did increase with increasing grid size. This suggests a convergence of model responses as spatial resolution increases (RQ2). MONC responses to moistening showed only weak resolution dependence. Boundary-layer temperature increases were more consistent between WRF and MONC for moistening than for heating, while free-tropospheric increases were stronger (weaker) in MONC than WRF for a low-level (upper-level) moisture perturbation. It is notable that, for instance, the temperature response above 300 hPa to a moisture perturbation at 850 hPa shows greater variability than the corresponding spread observed in the SCMs in H21, indicating a lower degree of uncertainty in the responses of high-resolution, convection-resolving models (RQ1). Perhaps more than for heating, low-mid tropospheric temperature and water vapor mixing ratio responses to moisture perturbations were closer to those of K10 when the model resolution was finer. Moistening increased RH at the perturbation level whereas heating reduced RH when applied in the low-mid troposphere. Similar to the responses to heating, hydrometeor responses to moistening had large variability (Figures A9 and A10). Graupel decreased almost universally with moistening at any level, contrasting with heating at low-levels which increased graupel content. Moistening at mid-upper levels tended to increase rain water content, in contrast with heating effects.

4. Conclusions

We examined time-averaged responses to temperature and moisture forcing perturbations in simulated RCE states, using two models at a range of grid spacings including LES runs without PBL schemes. The aim was to test the sensitivity of responses to model design, parameterization, and resolution. The study builds on Kuang (2010), here K10, who first calculated such responses, and other studies (in particular Hwong et al. (2021), here H21) who used them to test convective schemes used in global models.

In RQ1, we sought a "truth" response or benchmark results against which other models could be tested. The linear responses in our results were sensitive to model and model resolution, but showed a tendency to produce smoother vertical structures and indications of convergence as grid spacing is reduced, suggesting that model disagreement and artifacts from parameterizations or numerics were reduced at finer resolutions. In the low-mid troposphere, the temperature and moisture responses tended to be closer to the results of K10 as grid spacing became finer, indicating the accuracy of K10's 2-km grid calculations. However, above ∼500 hPa we found a smaller mean response than in K10, while in the boundary layer we found gradients in responses in the lowest kilometer of the atmosphere that were not found in the earlier study. The question is, then, what resolution is required to consider a response a benchmark result?

In RQ2, we asked what grid resolution is required to represent convective-scale feedbacks on large-scale flows. We can narrow down the grid spacing required by examining the sensitivity of the feedbacks to model settings. In our results, model-based differences between responses to temperature perturbations did not start to disappear

until the grid spacing was below ∼1 km, and likely lower for responses in the upper troposphere. We note however that variability across perturbations was present and computational limitations precluded us from performing rigorous convergence tests. Nonetheless, our results suggest that for most models, to simulate convectively-coupled dynamics correctly will likely require better resolution than is currently used in, for example, DYAMOND (∼5 km) (Stevens et al., 2019). In the lower troposphere, tangent linear responses were similar across the two LES models tested here, implying they could make reasonable benchmark estimates of model responses. In the upper troposphere the results diverged more and therefore the error bars on responses in this region are larger. Responses for grid spacings larger than ∼1 km diverged significantly from the LES responses.

The variations in responses across different convective parameterizations in H21's SCM runs were generally more pronounced than the discrepancies within our higher-resolution CRM and LES responses as well as K10. This suggests that high-resolution model outputs can serve as a valuable benchmark for identifying substantial deficiencies in convective parameterizations, particularly when the latter deviate significantly from the responses observed in these simulations. For instance, the pronounced kinks frequently observed in the SCM responses—especially near critical atmospheric levels such as the freezing level and cloud base—appeared less prominent or even disappeared at finer resolutions in our results, underscoring the problematic threshold-like behavior of convective parameterizations. However, this conclusion does not hold uniformly throughout the entire atmospheric column: the temperature responses showed increasing spread at higher altitudes, at times exceeding those of the SCMs in H21.

A novel aspect of this study is our comparisons of hydrometeors and their responses. Concentrations of snow and graupel in control runs, and responses of these (and of rain mixing ratio) in experiments, were highly diverse and appear likely to be responsible for some of the surprisingly large spread, for example, in upper-troposphere temperature responses. The treatment (often implicit) of microphysics within convection schemes may lack the diversity or sensitivity exhibited here in the LES and CRM models, explaining the narrower range of temperature responses in H21; this would imply that the true uncertainty of some responses exceeds the spreads reported in H21.

In summary, the key conclusions of our study are as follows:

1. For the two models tested here, correct simulation of convectively-coupled dynamics appears to require a grid spacing finer than ∼1 km.
2. The linear responses of the two LES models tested here could reasonably be used as ground truth for GCM scheme testing within the mid and lower troposphere.

We note that boundary-layer responses may still not be accurate without even finer grids than used here, and that better representation of microphysics in deep convection (for graupel, e.g.) will be critical to further constrain the results, especially in the upper troposphere. While we used finer grid spacing than previous work (e.g., K10) we did not address the issue of the use of small domains. Future work should investigate still finer resolutions, and could attempt to determine whether the response is dominated by the modeled convection itself, or is sensitive to the specification of surface properties in an idealized modeling setting.

Appendix A: Supplementary Figures

Supplementary figures are presented as Figures A1–A10.

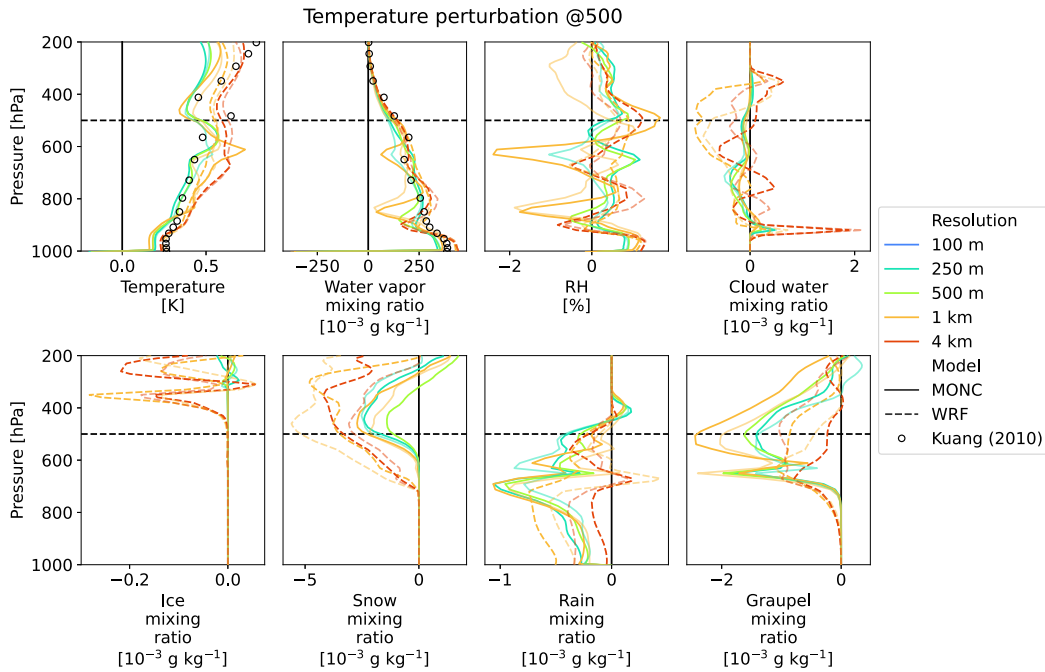

Figure A1. As in Figure 3, but for a temperature perturbation at 500 hPa, and with black circles showing the responses of K10 to a temperature perturbation at 483 hPa.

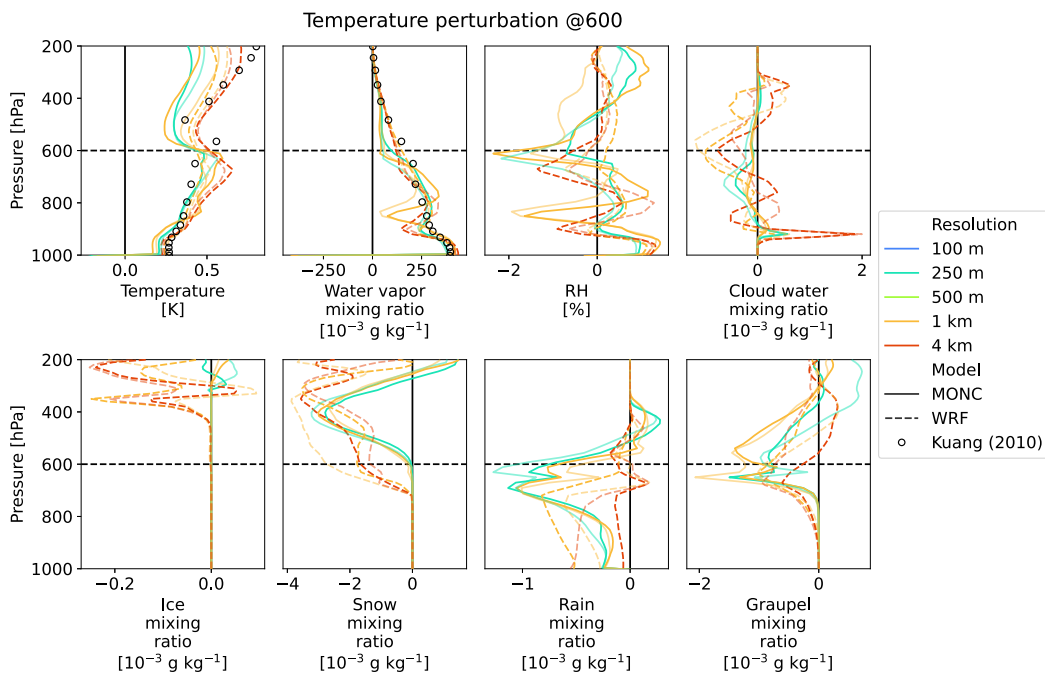

Figure A2. As in Figure 3, but for a temperature perturbation at 600 hPa, and with black circles showing the responses of K10 to a temperature perturbation at 565 hPa.

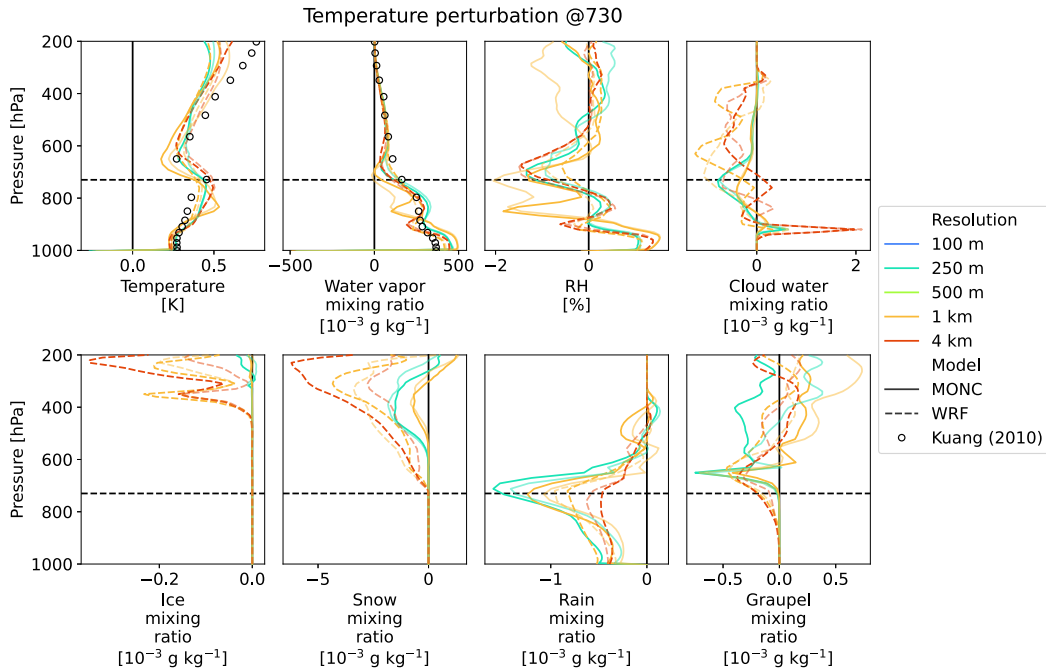

Figure A3. As in Figure 3, but for a temperature perturbation at 730 hPa, and with black circles showing the responses of K10 to a temperature perturbation at 729 hPa.

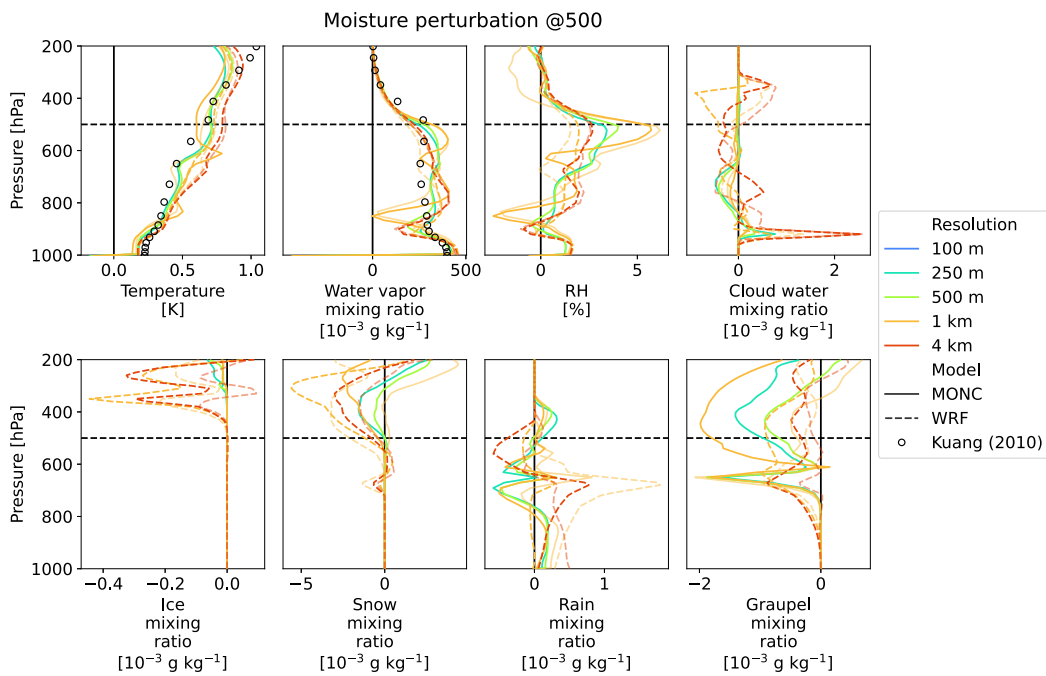

Figure A4. As in Figure 5, but for a water vapor mixing ratio perturbation at 500 hPa, and with black circles showing the responses of K10 to a specific humidity perturbation at 483 hPa.

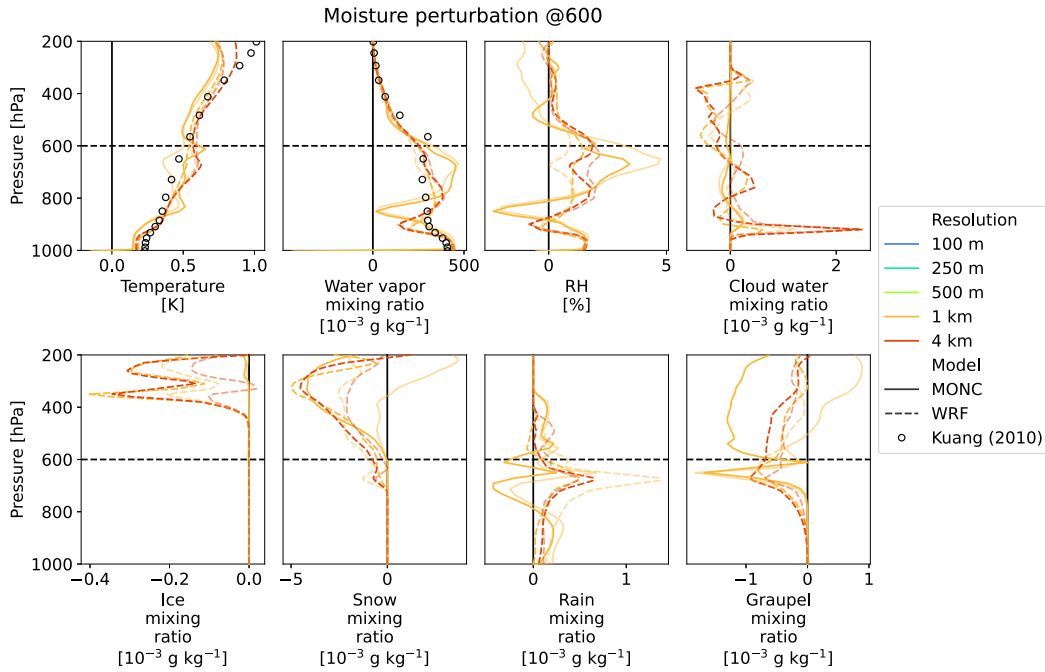

Figure A5. As in Figure 5, but for a perturbation at 600 hPa, and with black circles showing the responses of K10 to a specific humidity perturbation at 565 hPa.

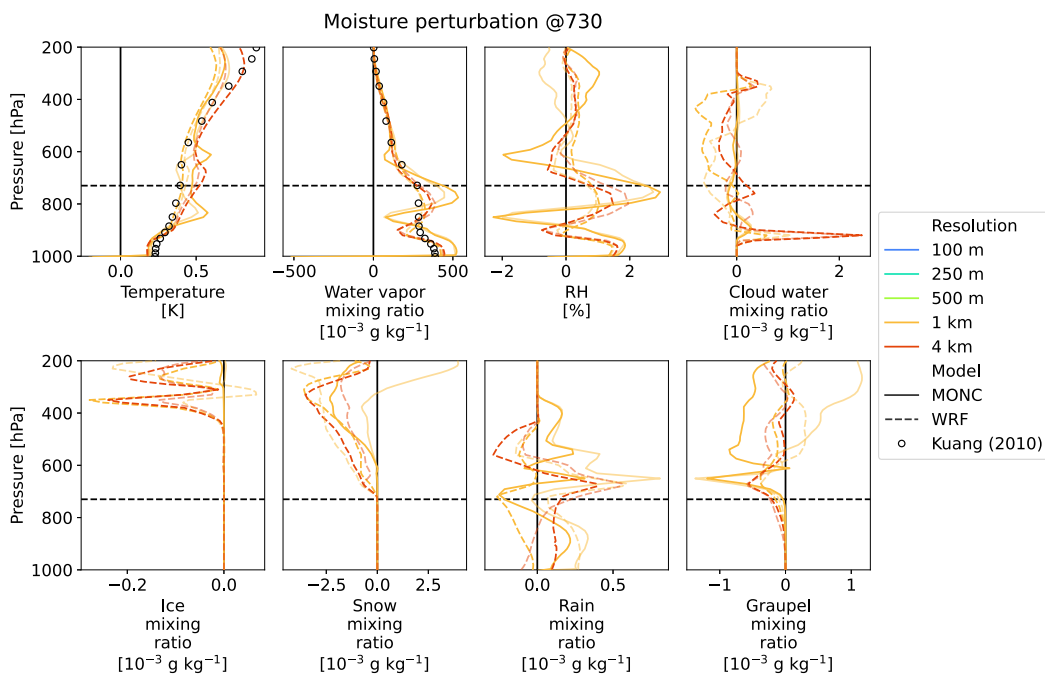

Figure A6. As in Figure 5, but for a perturbation at 730 hPa, and with black circles showing the responses of K10 to a specific humidity perturbation at 729 hPa.

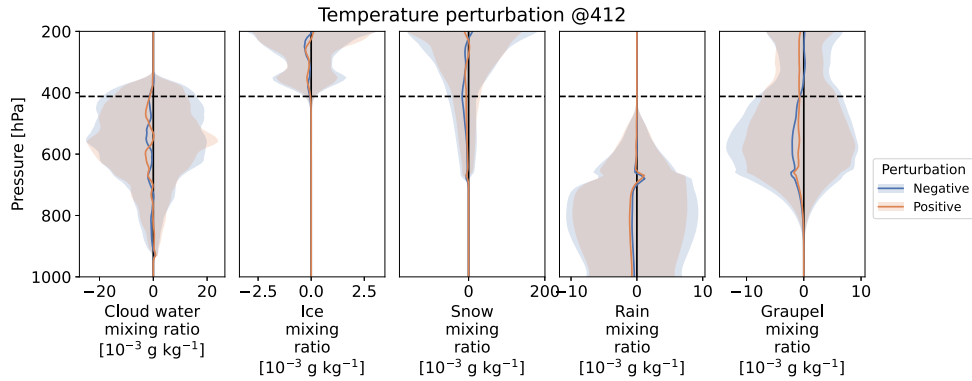

Figure A7. Mean responses (solid line) and the ±1 standard deviation range (shaded) for a temperature perturbation of ±0.5 K at 412 hPa in the Weather Research and Forecasting runs at 100 m grid spacing. The solid vertical line shows zero response, the dashed horizontal line shows the level of maximum perturbation.

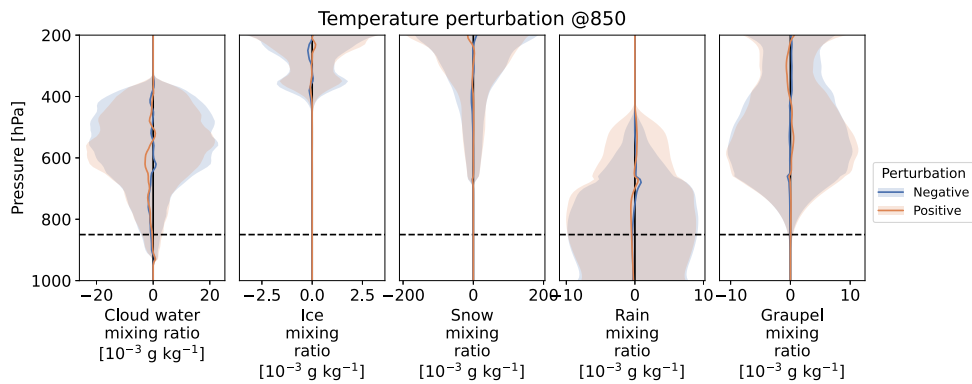

Figure A8. As in Figure A7 but for a temperature perturbation at 850 hPa.

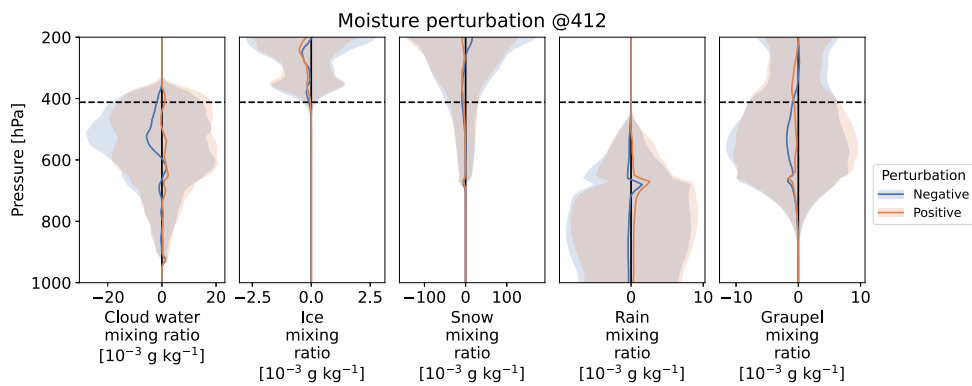

Figure A9. As in Figure A7 but for a water vapor mixing ratio perturbation of 0.2 g kg^{-1} at 412 hPa.

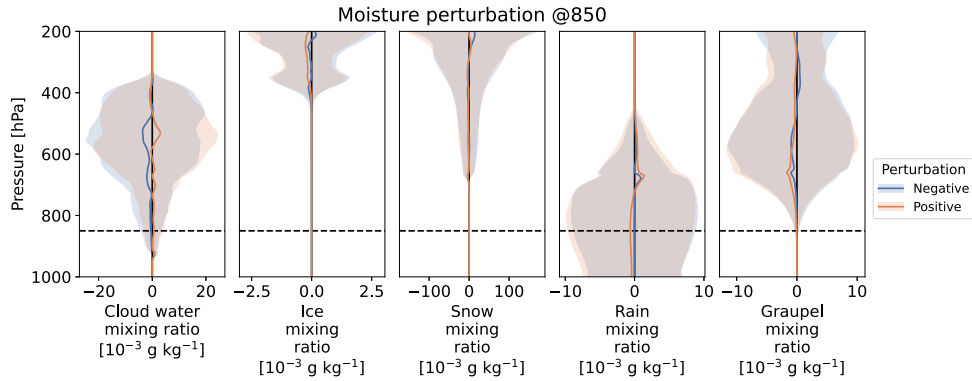

Figure A10. As in Figure A9 but for a water vapor mixing ratio perturbation of 0.2 g kg^{-1} at 850 hPa.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

Analysis code for this work, including modified WRF implementation code, example WRF and MONC namelists, and all responses reported here in machine readable form, are provided at Raupach (2026). The version used for this manuscript was v1.0.7.

Acknowledgments
THR's position at UNSW Sydney is supported by QBE Insurance. This research used resources from the National Computational Infrastructure (NCI Australia), an NCRIS-enabled capability supported by the Australian Government, with some compute time provided by the UNSW HPC Scheme (https://doi.org/10.26190/PMN5-7J50). CLD was funded by NERC Grant NE/N013743/1. SCS and THR received funding from the Australian Research Council Grant FL150100035. We thank two anonymous referees for reviewing our work. Open access publishing facilitated by University of New South Wales, as part of the Wiley - University of New South Wales agreement via the Council of Australasian University Librarians.

References

Ahn, M.-S., Kim, D., Kang, D., Lee, J., Sperber, K. R., Gleckler, P. J., et al. (2020). MJO propagation across the Maritime Continent: Are CMIP6 models better than CMIP5 models? *Geophysical Research Letters*, *47*(11), e2020GL087250. https://doi.org/10.1029/2020GL087250

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<2493:ratcpp>2.0.co;2

Brown, N., Weiland, M., Hill, A., Shipway, B., Maynard, C., Allen, T., & Rezny, M. (2020). A highly scalable Met Office NERC cloud model. https://doi.org/10.48550/arXiv.2009.12849

Bryan, G. H., Wyngaard, J. C., & Fritsch, J. M. (2003). Resolution requirements for the simulation of deep moist convection. *Monthly Weather Review*, *131*(10), 2394–2416. https://doi.org/10.1175/1520-0493(2003)131⟨2394:RRFTSO⟩2.0.CO;2

Bull, J. M., & Derbyshire, S. H. (1990). Numerical solution of the surface layer equations (Technical Report). In *Met office: Turbulence and diffusion technical note 197*.

Chua, X. R., Ming, Y., & Jeevanjee, N. (2019). Investigating the fast response of precipitation intensity and boundary layer temperature to atmospheric heating using a cloud-resolving model. *Geophysical Research Letters*, *46*(15), 9183–9192. https://doi.org/10.1029/2019GL082408

Daleu, C. L., Plant, R. S., Stirling, A. J., & Whitall, M. (2023). Evaluating the CoMorph-A parametrization using idealized simulations of the two-way coupling between convection and large-scale dynamics. *The Quarterly Journal of the Royal Meteorological Society*, *149*(757), 3087–3109. https://doi.org/10.1002/qj.4547

Daleu, C. L., Plant, R. S., Woolnough, S. J., Sessions, S., Herman, M. J., Sobel, A., et al. (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

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⟨1766:DAEOAC⟩2.0.CO;2

Field, P. R., Hill, A., Shipway, B., Furtado, K., Wilkinson, J., Miltenberger, A., et al. (2023). Implementation of a double moment cloud microphysics scheme in the UK Met Office regional numerical weather prediction model. *The Quarterly Journal of the Royal Meteorological Society*, *149*(752), 703–739. https://doi.org/10.1002/qj.4414

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

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

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

Hwong, Y. L., Song, S., Sherwood, S. C., Stirling, A. J., Rio, C., Roehrig, R., et al. (2021). Characterizing convection schemes using their responses to imposed tendency perturbations. *Journal of Advances in Modeling Earth Systems*, *13*(5), e2021MS002461. https://doi.org/10.1029/2021MS002461

Jiménez, P. A., Dudhia, J., González-Rouco, J. F., Navarro, J., Montávez, J. P., & García-Bustamante, E. (2012). A revised scheme for the WRF surface layer formulation. *Monthly Weather Review*, *140*(3), 898–918. https://doi.org/10.1175/MWR-D-11-00056.1

Khairoutdinov, M. F., & Randall, D. A. (2003). Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities. *Journal of the Atmospheric Sciences*, *60*(4), 607–625. https://doi.org/10.1175/1520-0469(2003)060⟨0607:CRMOTA⟩2.0.CO;2

Khouider, B., St-Cyr, A., Majda, A. J., & Tribbia, J. (2011). The MJO and convectively coupled waves in a coarse-resolution GCM with a simple multicloud parameterization. *Journal of the Atmospheric Sciences*, *68*(2), 240–264. https://doi.org/10.1175/2010JAS3443.1

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

Lilly, D. K. (1967). The representation of small-scale turbulence in numerical simulation experiments. In *Proceedings of IBM scientific computing symposium on environmental science* (pp. 195–210).

Lin, J., Qian, T., Bechtold, P., Grell, G., Zhang, G. J., Zhu, P., et al. (2022). Atmospheric convection. *Atmosphere-Ocean*, *60*(3–4), 422–476. https://doi.org/10.1080/07055900.2022.2082915

Muller, C., & Bony, S. (2015). What favors convective aggregation and why? *Geophysical Research Letters*, *42*(13), 5626–5634. https://doi.org/10.1002/2015GL064260

Raupach, T. (2026). Traupach/wrf_lrf_les: V1.0.7 [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.15826348

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

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., et al. (2021). A description of the advanced research WRF model version 4.1 (Technical Report Nos. NCAR/TN-556+STR). *NCAR*. https://doi.org/10.5065/1dfh-6p97

Smagorinsky, J. (1963). General circulation experiments with the primitive equations. *Monthly Weather Review*, *91*(3), 99–164. https://doi.org/10.1175/1520-0493(1963)091⟨0099:GCEWTP⟩2.3.CO;2

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., et al. (2019). DYAMOND: The DYnamics of the Atmospheric general circulation modeled on Non-hydrostatic Domains. *Progress in Earth and Planetary Science*, *6*(1), 61. https://doi.org/10.1186/s40645-019-0304-z

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Monthly Weather Review*, *136*(12), 5095–5115. https://doi.org/10.1175/2008MWR2387.1

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

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