Response to Reviewer #1

The authors present initial conditions and results of an idealized dynamical core test.

The contribution is very welcome and the attention to detail appreciated. Too often are test case descriptions incomplete, which can result in difficulties if not the failure when trying to implement and or reproduce the results.

The Scientific reproducibility was rated as good. It would be excellent if the source datasets (NetCDF files) for the figures were uploaded to, for example, figshare, and cited with their DOI. We would strongly suggest the authors do this.

Thank you for your comprehensive review.

In response to data being made available with a DOI, we should note that we (erroneously) omitted the formal public references for both the initialization code via Zenodo (DOI:10.5281/zenodo.1298671) and the model source code (DOI:10.5194/gmd-10-4477-2017) in the original submission. These have both been added to the body of the revised manuscript, which should allow for full reproducibility.

We hope we have addressed all of the comments satisfactorily below.

On page 2, lines 15 and 16: It is not clear that it is necessary to add the (1°) and (1km). Rather it would be clearer to write, for example, "... the resolution is 4°/X $\sim$ 440km/X $\sim$ 4km.

Per this suggestion, changed to 'Therefore, for a 1° mesh, the grid spacing of the reduced radius sphere is approximately 1°/X $\sim$ 111km/X $\sim$ 111km/120 $\sim$ 1km near the equator.'

On page 2, line 23: maybe rewrite the current "... were required with time between outputs required to be ..." as " were stored for post-processing with a frequency of every 15 min or higher."

Changed to '... to the microphysical routines were stored for post-processing with a frequency of every 15 min or finer.'

Page 3, line 7, 8, and 9: I would suggest to rewrite equation 3 without the cos ϕ term and hence replace u -(ϕ, z) with ueq(z) and add to line 4: u(z, ϕ) = ueq(z) cos ϕ. Remove any overbar on U and H. It is not subsequently used, and in this particular case, it is the mean of a constant, and therefore potentially misleading as one might think that there is a zonal variation.

We have implemented a 'hybrid' modification, which is a mix of the original formulation and the reviewer's suggestion. Specifically, we have chosen to remove the overbars in that our intent was to emphasize that these are mean background state values (that the warm bubble perturbation is placed upon), but the reviewer is correct to note that the notation may actually introduce more confusion than clarification.

However, we have chosen to leave the definition of wind velocity as is. This is done primarily for continuity with Klemp et al. [2015], as well as the published DCMIP2016 initialization code (DOI:10.5281/zenodo.1298671).

Page 5, line 27: Please add a summary of the results presented to prepare the reader and fill the void in between 3. and 3.1

We have added a paragraph which broadly summarizes the results to be presented in the following sections. We also chose to use this space to emphasize that attributing model spread to particular model design choices is beyond the scope of this manuscript; this paper's purpose is to catalog and define the set of model results from DCMIP2016 efforts to be used for future reference.

Page 6, line 18: Maybe "... a notable difference exists through the end of the runs ..." would better read "... a notable difference exists towards the end of the runs ..."? "through" to me would indicate from beginning to end. "This spread in model solution..." potentially has an unclear precedent. Maybe "..., the models start to diverge towards the end of the run. This divergence ..." without the emphasis, of course, would be more explicit?

We agree this passage was somewhat sloppy. This has been modified to read '... notable differences exist, particularly towards the end of the runs. This divergence can be seen as early as 30min in some cases but is most notable at the test conclusion.'

Page 15, Figure 5: either replace the "increasing darker" with the actual color names or refer to the legend. I find it difficult to say which color is darker or lighter (except the black and light pink, naturally).

This has been changed to only note that black is the finest resolution applied in this test. The reader is referred to the legend to match the rest of the lines to their associated grid spacing.

On page 2, line 22 (and elsewhere in the document): There is no need to state that 120min=7200s. Just say that it is 120min.

Corrected.

Page 7, line 3: As noted above, no need to add the 7200 sec

Corrected.

Apart from these minor and mostly technical issues, we would strongly suggest to accept and congratulate the authors for the clear and detailed presentation.

Thank you.

Response to Reviewer #2

The authors present an idealized test case (supercell storm) for global atmospheric models and briefly discuss the results from a large set of models that participated in a model comparison project (DCMIP2016). The premise of the study is good and useful. Demonstrating variability in the simulation of atmospheric phenomena due to differences in the dynamic core is an important topic and one that is often overlooked by the science community. The methods, analysis and even the figures are almost identical to a paper by Klemp et al. (2015), so the paper adds nothing new in this respect. The usefulness of the paper comes from the presentation of results from a large and diverse set of models. Despite this, the paper needs significant revisions to improve the presentation and additional analysis/simulations are needed to make the results clearer.

Thank you for your thorough review. We agree that the main thrust of this manuscript is to document a set of well-vetted solutions from many world-class dynamical cores. This not only provides a snapshot of model status at the time of DCMIP2016, but also acts as an inventory for which future modeling endeavours can compare their high-resolution solutions against; both as a sanity check as well as hypothesizing dynamical core impacts on non-hydrostatic phenomena.

The introduction was disappointing. No literature review is done on the impacts of dynamic cores on atmospheric problems such as extreme weather events. There are recently published papers on this topic that should be cited. In addition, more discussion is needed on the motivation for the study. Why do we need to compare dynamic cores? Do we expect significant differences and if so, do these differences explain some of the uncertainty observed in forecasts of extreme weather?

We have added some additional citations regarding recent work tying dynamical cores to extreme weather. That said, we would be happy to include discussion of additional literature that we may have overlooked.

We have also added some discussion of DCMIP's motivation and why it matters at 'the end of the pipeline' with respect to weather and climate simulations.

The results of the resolution sensitivity study don't appear to be converged at 0.5 km for several models. For example, there are significant differences in one or more fields going from 1 km to 0.5 km resolution for the following models: CSU, NICAM, ICON, GEM, TEMPEST. The authors should show a 0.25 km resolution simulation for one or two of these models to demonstrate better convergence properties. In addition, shouldn't we expect that as the grid spacing becomes very small, the model solutions and bulk statistics should be fairly similar across models? This is not the case with the current results so they are likely far from converged. Some discussion and additional example tests (as described above) are needed.

We would expect intramodel model solutions and bulk statistics to begin the converge at the higher resolutions in this study. That said, one of the obvious findings of DCMIP2016 is that non-hydrostatic phenomena at the resolutions tested here remain sensitive to numerical scheme, diffusion, and physics-dynamics coupling in a way that large-scale features such as baroclinic instabilities are not [Jablonowski et al., 2016].

While additional convergence tests may be enlightening, it is also computationally (and logistically) burdensome to complete new simulations at 250m resolution. That said, we agree that the reviewer's concerns are valid and have made care to note that these are targets for additional simulations either at the individual modeling center level or for future DCMIP projects.

Page 1, line 8; What is meant by "physics-dynamics coupling"?

This is the technical coupling between the dynamical core and representations of subgrid processes; commonly referred to as 'physics.' [Gross et al., 2016, 2018].

Gross et al. [2018] defines *physics-dynamics coupling* (PDC) as '... bringing together all the various discretized components to create a coherent model will be referred to here as physics–dynamics coupling. The term physics–dynamics coupling has evolved from the fact that the resolved fluid dynamics components are commonly known as the dynamical cores or simply "dynamics," and the physical parameterizations that represent the unresolved and underresolved processes and the nonfluid dynamical processes are collectively referred to as "physics."'

Page 1, line 11; What is the difference between convective-permitting and convective-allowing? These mean the same thing to me.

There is a bit of a 'gray' area in the definition of this by various modelers. Colloquially, a model 'permits' the simulation of a phenomenon if it can be discerned, regardless of whether it is under-resolved or not. A model 'allows' a phenomenon if it can be discerned and is credibly resolved. That said, we understand now where this can be confusing and isn't critical within the abstract so we have chosen to just include the term 'convective-allowing' especially since regimes pertaining to resolved convection in the atmosphere are a continuum and not cut discretely.

Page 2, sentence starting with "It is based on the work of. . ."; Sentence doesn't read well and needs a re-write.

As requested, this passage has been rewritten as follows 'This test is based on the work of Klemp and Wilhelmson [1978] and Klemp et al. [2015] and assesses the performance of global models at extremely high spatial resolution. It has recently been used in the development of next-generation numerical weather prediction systems [Ji and Toepfer, 2016].'

Page 2, line 15; Need few sentences on what a "reduced-radius sphere" is and what it intends to represent. Is the use of a reduced radius sphere a good approximation to true radius, global dynamics? Why do you need to simulate a supercell to study the performance of a global model? Wouldn't a global phenomenon (with non-hydrostatic processes) be better suited for studying the performance of a global model? How many horizontal grid points are used in the reduced radius simulations?

Using a reduced radius sphere allows for computationally-efficient simulations of O(1km) grid spacings in global models without modifying the numerical framework. Simulating supercells are important for non-hydrostatic development because A) the storms strongly stress non-hydrostatic numerics, B) they represent key atmospheric phenomena with high societal relevance, making them of importance to both weather and climate modelers, and C) they are currently unresolved in most global numerical modeling frameworks but that is projected to change over the coming decade or two.

We have added 'Reducing the model's planetary radius allows for fine grid spacing to be achieved without the added computational expense associated with adding grid cells a standard global mesh in order to achieve non-hydrostatic resolutions [Kuang et al., 2005]. Wedi and Smolarkiewicz [2009] provide a detailed overview of the reduced-radius framework for testing global models.' to the text to shed light on this approach.

Section 2.1; The notation in this section is confusing. What is Ueq a function of , z? If Ueq is the zonal velocity at the equator, then isn't u = Ueq? I also don't understand the transition to defining ubar after u. This whole section needs to be described better.

This notation has been modified per the suggestions of both Reviewers #1 and #2.

Page 3, line 1; This can't be gradient wind balance because no Coriolis force is shown in the equation.

Our apologies for this confusion. Technically, this is more akin to cyclostrophic balance. The text has been clarified to address this. The equation here can be derived from the gradient wind equation by setting the Coriolis term equal to zero and allowing a strong local pressure gradient force to be balanced by centrifugal force at an arbitrary latitude ϕ.

Page 4, line 7; what is machine epsilon? Are you referring to machine precision?

They are quite similar but also technically distinct.

Machine precision is effectively the accuracy of the basic arithmetic operations.

Machine epsilon is the discrete distance between (for example) 1 and the next 'resolved' floating point number.

In this case, we use machine epsilon because we define convergence as the time when the 'distance' between iteration n and $n+1$ is less than can be 'resolved' by the minimum gap between two floating point numbers.

Page 5, line 3; Why is the warm bubble hydrostatically balanced? The introduction and abstract highlight the importance of non-hydrostatic processes for testing global models, so this doesn't make sense. The vertical accelerations should be fairly small for this warm bubble relative to the supercell that is simulated with very large vertical velocities. The balance adjustment is a physical process that should be tested in the models.

The bubble is hydrostatically balanced in order to result in a smoothly-evolving solution at test case onset when the flow has not developed strong non-hydrostatic characteristics. Technically, there is no requirement that the bubble be balanced; however a less carefully-designed perturbation will result in gravity waves associated with flow adjustment in the first few timesteps and/or a less realistic supercell evolution. However, the long-term behavior of the solution is largely insensitive to whether or not the field is rebalanced.

Page 5, line 13; Should say "second-order diffusion operator with a constant sub-grid scale viscosity (value) is applied...". Are you using the same values in the horizontal and vertical dimensions? Is that appropriate for the grid cell aspect ratio?

Changed to 'as resolution is increased for a given model, a second-order diffusion operator with a constant viscosity (value) is applied to all momentum equations.'

The same value is used in the horizontal and vertical directions (unless specified in the relevant passage in the text). Since this diffusion is added to mimic turbulent dissipation within supercells, the choice of the same value for all three dimensions should be reasonable and is consistent with Klemp et al. [2015].

Page 5, sentence starting with "ICON applied. . ."; This doesn't sound like a departure from the default described above. Also, is diffusion really applied to the mass continuity equation (rho) in ICON? This is not typically done.

ICON departed from the prescribed test since it did not apply any diffusion in the vertical. The formal definition of the test case includes diffusion on the three-dimensional momentum equations (vertical diffusion is only applies to the background state perturbation).

The developers of ICON confirmed that the inclusion of ρ in the list of variables where diffusion is applied was erroneous. This has been corrected and the developers greatly appreciate the reviewer noticing this oversight.

Section 3.2; This whole section is very subjective with no analysis to back up any of the claims. In order to comment on the reasons for the differences in the model runs, some analysis is needed.

We have added additional text to help buttress some of the hypotheses here. However, we should emphasize that the primary goal of this manuscript is not to do a formal deep dive into all model differences but rather define the test case and scope of solutions from modeling groups that participated in DCMIP2016. We have chosen to leave formal attribution studies to individual modeling centers (or groups of modeling centers) as model design choices are implemented when accounting for a host of considerations versus the outcome of a particular test case.

Page 7, line 11; How do you know that this is noise? This statement is subjective and no quantitative analysis is done to back up this claim.

This was poor verbiage on our part, we apologize. We have replaced 'noise' with 'small-scale structure' since we are not trying to argue these solutions necessarily contain numerical or physical 'noise;' spurious or otherwise.

Page 7, line 15; This is a very weak statement and should be removed. What is meant by "coupling between the dynamical core and subgrid parameterization" and why would this lead to these differences?

We have added text to clarify that this is a hypothesis and a target for future work. There has been previous work indicating that the coupling mechanisms between the dynamical core and subgrid physics parameterizations can drive sensitivity in solution output [Gross et al., 2018]. DCMIP2016 did not formally control for this, largely because it is very difficult to tightly specify a coupling framework that satisfies the multitude of different numerical schemes and software infrastructures used by various modeling centers. Rather, the majority of work investigating physics-dynamics coupling has been done within individual modeling frameworks, and is something a subset of participating DCMIP models will likely look into over the next few years.

Page 7, line 21; This is not a bulk, integrated measure of supercell intensity. How about showing the domain integrated kinetic energy or total energy?

For this test, integrated kinetic energy is a difficult value to evaluate, as the atmosphere is not at rest at initialization (and the storm dynamics' contribution to the kinetic energy budget is actually quite 'minimal' relative to the KE of the background environment – on the order of 1% or so). Most of the energetic deviation from the initial sheared state is due to the strong vertical motion associated with the intensifying supercell.

We chose maximum updraft velocity over the entire domain as a metric for two reasons. One, it is a tangible quantity that is commonly reported in both observational and modeling studies of supercells. Two, we apply the assumption that maximum updraft velocity is a first-order proxy for 'storm intensity' as measured by a more 'dynamical' quantity than rainfall.

However, we note that the language here was not precise and understand why it may cause confusion, so we have modified to read '... more storm-wide measures...'

References

Markus Gross, Sylvie Malardel, Christiane Jablonowski, and Nigel Wood. Bridging the (knowledge) gap between physics and dynamics. *Bulletin of the American Meteorological Society*, 97(1):137–142, Jan 2016. doi: 10.1175/bams-d-15-00103.1. URL `http://dx.doi.org/10.1175/BAMS-D-15-00103.1`.

Markus Gross, Hui Wan, Philip J. Rasch, Peter M. Caldwell, David L. Williamson, Daniel Klocke, Christiane Jablonowski, Diana R. Thatcher, Nigel Wood, Mike Cullen, and et al. Physics–dynamics coupling in weather, climate, and earth system models: Challenges and recent progress. *Monthly Weather Review*, 146(11):3505–3544, Nov 2018. doi: 10.1175/mwr-d-17-0345.1. URL `http://dx.doi.org/10.1175/MWR-D-17-0345.1`.

C. Jablonowski, C. M. Zarzycki, K. A. Reed, P. A. Ullrich, J. Kent, P. H. Lauritzen, and R. D. Nair. The Dynamical Core Model Intercomparison Project (DCMIP-2016): Results of the Moist Baroclinic Wave Test Case. *AGU Fall Meeting Abstracts*, December 2016.

Ming Ji and Frederick Toepfer. Dynamical core evaluation test report for NOAA's Next Generation Global Prediction System (NGGPS). Technical report, National Oceanic and Atmospheric Administration, 2016. URL `https://repository.library.noaa.gov/view/noaa/18653`.

J. B. Klemp, W. C. Skamarock, and S.-H. Park. Idealized global nonhydrostatic atmospheric test cases on a reduced-radius sphere. *Journal of Advances in Modeling Earth Systems*, 7(3):1155–1177, 2015.

Joseph B Klemp and Robert B Wilhelmson. The simulation of three-dimensional convective storm dynamics. *Journal of the Atmospheric Sciences*, 35(6):1070–1096, 1978.

Zhiming Kuang, Peter N. Blossey, and Christopher S. Bretherton. A new approach for 3D cloud-resolving simulations of large-scale atmospheric circulation. *Geophysical Research Letters*, 32(2), 2005.

Nils P. Wedi and Piotr K. Smolarkiewicz. A framework for testing global non-hydrostatic models. *Quarterly Journal of the Royal Meteorological Society*, 135(639):469–484, 2009.

DCMIP2016: The Splitting Supercell Test Case

Colin M. Zarzycki[1], Christiane Jablonowski[2], James Kent[3], Peter H. Lauritzen[1], Ramachandran Nair[1], Kevin A. Reed[4], Paul A. Ullrich[5], David M. Hall[6], Don Dazlich[7], Ross Heikes[7], Celal Konor[7], David Randall[7], Xi Chen[8], Lucas Harris[8], Marco Giorgetta[9], Daniel Reinert[10], Christian Kühnlein[11], Robert Walko[12], Vivian Lee[13], Abdessamad Qaddouri[13], Monique Tanguay[13], Hiroaki Miura[14], Tomoki Ohno[15], Ryuji Yoshida[16], Sang-Hun Park[17], Joseph Klemp[1], and William Skamarock[1]

[1]National Center for Atmospheric Research
[2]University of Michigan
[3]University of South Wales
[4]Stony Brook University
[5]University of California, Davis
[6]University of Colorado, Boulder
[7]Colorado State University
[8]Geophysical Fluid Dynamics Laboratory (GFDL)
[9]Max Planck Institute for Meteorology
[10]Deutscher Wetterdienst (DWD)
[11]European Centre for Medium-Range Weather Forecasts (ECMWF)
[12]University of Miami
[13]Environment and Climate Change Canada (ECCC)
[14]University of Tokyo
[15]Japan Agency for Marine-Earth Science and Technology
[16]RIKEN
[17]Yonsei University

Correspondence to: Colin Zarzycki (zarzycki@ucar.edu)

Abstract. This paper describes the splitting supercell idealized test case used in the 2016 Dynamical Core Model Intercomparison Project (DCMIP2016). These storms are useful testbeds for global atmospheric models because the horizontal scale of convective plumes is O(1km), emphasizing non-hydrostatic dynamics. The test case simulates a supercell on a reduced radius sphere with nominal resolutions ranging from 4km to 0.5km and is based on the work of Klemp et al. (2015). Models are initialized with an atmospheric environment conducive to supercell formation and forced with a small thermal perturbation. A simplified Kessler microphysics scheme is coupled to the dynamical core to represent moist processes. Reference solutions for DCMIP2016 models are presented. Storm evolution is broadly similar between models, although differences in final solution exist. These differences are hypothesized to result from different numerical discretizations, physics-dynamics coupling, and numerical diffusion. Intramodel solutions generally converge as models approach 0.5km resolution. These results can be used as a reference for future dynamical core evaluation, particularly with the development of non-hydrostatic global models intended to be used in convective-permitting ~~and convective-allowing~~ regimes.

1 Introduction

Supercells are strong, long-lived convective cells containing deep, persistent rotating updrafts that operate on spatial scales O(10km). They can persist for many hours and frequently produce large hail, tornados, damaging straight line winds, cloud-to-ground lightning, and heavy rain (Browning, 1964; Lemon and Doswell, 1979; Doswell and Burgess, 1993). Therefore, accurate simulation of these features is of great societal interest and critical for atmospheric models.

The supercell test applied in DCMIP2016 (Ullrich et al., 2017) permits the study of a non-hydrostatic moist flow field with strong vertical velocities and the associated precipitation. ~~It~~ This test is based on the work of Klemp and Wilhelmson (1978) and Klemp et al. (2015) ~~, assesses the behavior of global modeling systems~~ and assesses the performance of global models at extremely high spatial resolution~~, and is~~ . It has recently been used in the development of next-generation numerical weather prediction ~~capabilities.~~ systems (Ji and Toepfer, 2016).

While some recent work has targeted the impact of dynamical cores on extreme weather – primarily tropical cyclones (e.g., Zhao et al. (2012); Reed et al. (2015)) – these studies have almost exclusively employed hydrostatic dynamical cores at grid spacings approximately 0.25° and coarser. Conversely, the supercell test here emphasizes resolved, non-hydrostatic dynamics. In this regime the effective grid spacing is very similar to the horizontal scale of convective plumes~~, emphasizing resolved, non-hydrostatic dynamics~~. Further, the addition of simplified moist physics injects energy ~~at, or near ,~~ near the grid-scale in a conditionally-unstable atmosphere, which imposes significant stress on model numerics. The supercell test case therefore sheds light on the interplay of the dynamical core and subgrid parameterizations and highlights the impact of both implicit and explicit numerical diffusion on model solutions. It also demonstrates credibility of a global modeling framework to simulate extreme phenomena, essential for future weather and climate simulations.

2 Description of test

The test case is defined as follows. All test cases are run on a non-rotating, reduced radius sphere with scaling factor X = 120. ~~Therefore, for a 4(~~Reducing the model's planetary radius allows for fine grid spacing to be achieved without the increased computational expense associated with adding grid cells to a standard global mesh in order to achieve non-hydrostatic resolutions (Kuang et al., 2005). Wedi and Smolarkiewicz (2009) provide a detailed overview of the reduced-radius framework for testing global models. For a 1° ~~)~~ mesh, the grid spacing of the reduced radius sphere is approximately ~~4km (~~$1^\circ/X \sim$ 111km/X $\sim$ 111km/120 $\sim$ 1km ~~)~~ near the equator. Klemp et al. (2015) demonstrated excellent agreement between simulations using this value of X and those completed on a flat plane with equivalent resolution. The model top is placed at 20km with uniform vertical grid spacing (Δz) equal to 500m, resulting in 40 full vertical levels. No surface drag is imposed at the lower boundary (free slip condition). Water vapor (q_v), cloud water (q_c) and rain water (q_r) are handled by a simple Kessler microphysics routine (Kessler, 1969). In particular, the ~~configuration of~~ Kessler microphysics used here is outlined in detail in Appendix C of Klemp et al. (2015) ~~.~~ and code for reproducing this configuration is available via the DCMIP2016 repository (http://dx.doi.org/10.5281/zenodo.1298671).

All simulations are integrated for 120min~~(7200s)~~. Outputs of the full three-dimensional prognostic fields as well as all variables pertaining to the microphysical routines were ~~required with time between outputs required to be no longer than~~ stored for post-processing with a frequency of every 15min or finer. Four different horizontal resolutions were specified; 4°, 2°, 1°, and 0.5°. When considering the radius reduction, this results in approximate grid spacings of 4km, 2km, 1km, and 0.5km, respectively. Note that here we use '(nominal) resolution' and 'grid spacing' interchangeably to refer to the length of a single grid cell or distance between gridpoints. All relevant constants mentioned here and in the following section are defined in Table 1.

2.1 Mean atmospheric background

The mean atmospheric state is designed such that it consists of large instability (convective available potential energy (CAPE) of approximately 2200 $\mathrm{m^2\ s^{-1}}$) and strong low-level wind shear, both of which are strong precursors of supercell formation (Weisman and Klemp, 1982).

The definition of this test case relies on hydrostatic and ~~gradient~~ cyclostrophic wind balance, written in terms of Exner pressure π and virtual potential temperature θ_v as

$$\frac{\partial \pi}{\partial z} = -\frac{g}{c_p \theta_v}, \quad \text{and} \quad u^2 \tan\varphi = -c_p \theta_v \frac{\partial \pi}{\partial \varphi}. \tag{1}$$

Defining $u = u_{eq}\cos\varphi$ to maintain solid body rotation, where u_{eq} is the equatorial wind velocity, these equations can be combined to eliminate π, leading to

$$\frac{\partial \theta_v}{\partial \varphi} = \frac{\sin(2\varphi)}{2g}\left(u_{eq}^2 \frac{\partial \theta_v}{\partial z} - \theta_v \frac{\partial u_{eq}^2}{\partial z}\right). \tag{2}$$

The wind velocity is analytically defined throughout the domain. Meridional and vertical wind is initially set to zero. The zonal wind is obtained from

$$u(\varphi, z) = \begin{cases} \left(U_s \frac{z}{z_s} - U_c\right)\cos(\varphi) & \text{for } z < z_s - \Delta z_u, \\ \left[\left(-\frac{4}{5} + 3\frac{z}{z_s} - \frac{5}{4}\frac{z^2}{z_s^2}\right)U_s - U_c\right]\cos(\varphi) & \text{for } |z - z_s| \le \Delta z_u \\ (U_s - U_c)\cos(\varphi) & \text{for } z > z_s + \Delta z_u \end{cases} \tag{3}$$

The equatorial profile is determined through numerical iteration. Potential temperature at the equator is specified via

$$\theta_{\text{eq}}(z) = \begin{cases} \theta_0 + (\theta_{tr} - \theta_0)\left(\frac{z}{z_{tr}}\right)^{\frac{5}{4}} & \text{for } 0 \le z \le z_{tr}, \\ \theta_{tr}\exp\left(\frac{g(z - z_{tr})}{c_p T_{tr}}\right) & \text{for } z_{tr} \le z \end{cases} \tag{4}$$

And relative humidity is given by

$$H(z) = \begin{cases} 1 - \frac{3}{4}\left(\frac{z}{z_{tr}}\right)^{5/4} & \text{for } 0 \le z \le z_{tr}, \\ \frac{1}{4} & \text{for } z_{tr} \le z. \end{cases} \tag{5}$$

It is assumed that the saturation mixing ratio is given by

$$q_{vs}(p,T) = \left(\frac{380.0}{p}\right)\exp\left(17.27\times\frac{T-273.0}{T-36.0}\right) \tag{6}$$

Pressure and temperature at the equator are obtained by iterating on hydrostatic balance with initial state

$$\theta^{(0)}_{v,\text{eq}}(z) = \theta_{\text{eq}}(z), \tag{7}$$

and iteration procedure

$$\pi^{(i)}_{\text{eq}} = 1 - \int_0^z \frac{g}{c_p\theta^{(i)}_{v,\text{eq}}}dz \tag{8}$$

$$p^{(i)}_{\text{eq}} = p_0(\pi^{(i)}_{\text{eq}})^{c_p/R_d} \tag{9}$$

$$T^{(i)}_{\text{eq}} = \theta_{\text{eq}}(z)\pi^{(i)}_{\text{eq}} \tag{10}$$

$$q^{(i)}_{\text{eq}} = H(z)q_{vs}(p^{(i)}_{\text{eq}},T^{(i)}_{\text{eq}}) \tag{11}$$

$$\theta^{(i+1)}_{v,\text{eq}} = \theta_{\text{eq}}(z)(1+M_v q^{(i)}_{\text{eq}}) \tag{12}$$

This iteration procedure generally converge to machine epsilon after approximately 10 iterations. The equatorial moisture profile is then extended through the entire domain,

Once the equatorial profile has been constructed, the virtual potential temperature through the remainder of the domain can be computed by iterating on (2),

$$\theta^{(i+1)}_v(z,\varphi) = \theta_{v,\text{eq}}(z) + \int_0^\varphi \frac{\sin(2\phi)}{2g}\left(u^2_{eq}\frac{\partial\theta^{(i)}_v}{\partial z} - \theta^{(i)}_v\frac{\partial\overline{u}^2_{eq}}{\partial z}\frac{\partial u^2_{eq}}{\partial z}\right)d\varphi. \tag{14}$$

Again, approximately 10 iterations are needed for convergence to machine epsilon. Once virtual potential temperature has been computed throughout the domain, Exner pressure throughout the domain can be obtained from (1),

and so

$$p(z,\varphi) = p_0\pi(z,\varphi)^{c_p/R_d}, \tag{16}$$

$$T_v(z,\varphi) = \theta_v(z,\varphi)(p/p_0)^{R_d/c_p}. \tag{17}$$

Note that, for (**??**-14), Smolarkiewicz et al. (2017) also derive an analytic solution for the meridional variation of the initial background state for shallow atmospheres.

2.2 Potential temperature perturbation

To initiate convection, a thermal perturbation is introduced into the initial potential temperature field:

$$\theta'(\lambda,\phi,z) = \begin{cases} \Delta\theta \cos^2\left(\frac{\pi}{2} R_\theta(\lambda,\varphi,z)\right) & \text{for } R_\theta(\lambda,\varphi,z) < 1, \\ 0 & \text{for } R_\theta(\lambda,\varphi,z) \geq 1, \end{cases} \tag{18}$$

where

$$R_\theta(\lambda,\varphi,z) = \left[\left(\frac{R_c(\lambda,\varphi;\lambda_p,\varphi_p)}{r_p}\right)^2 + \left(\frac{z-z_c}{z_p}\right)^2\right]^{1/2}. \tag{19}$$

An additional iterative step is then required to bring the potential temperature perturbation into hydrostatic balance. Without this additional iteration, large vertical velocities will be generated as the model flow rapidly adjusts to hydrostatic balance since the test does not possess strong non-hydrostatic characteristics at initialization. Plots showing the initial state of the supercell are shown in Figs. 1 and 2 for reference. Code used by modeling centers during DCMIP2016 for initialization of the supercell test case is archived via Zenodo (http://dx.doi.org/10.5281/zenodo.1298671).

The test case is designed such that the thermal perturbation will induce a convective updraft immediately after initialization. As rain water is generated by the microphysics, reduced buoyancy and a subsequent downdraft at the equator in combination with favorable vertical pressure gradients near the peripheral flanks of the storm will cause it to split into two counterrotating cells that propagate transversely away from the equator until the end of the test (Rotunno and Klemp, 1982, 1985; Rotunno, 1993; Klemp et al., 2015).

2.3 Physical and Numerical Diffusion

As noted in Klemp et al. (2015), dissipation is an important process near the grid-scale, particularly in simulations investigating convection in unstable environments such as this. To represent this process and ensure facilitate solution convergence as resolution is increased for a given model, a constant second-order diffusion operator with a constant viscosity (value) is applied to all momentum equations (ν = 500 m^2 s^{-1}) and scalar equations (ν = 1500 m^2 s^{-1}). In the vertical, this diffusion is applied to the perturbation from the background state only in order to prevent the initial perturbation from mixing out.

Models that contributed supercell test results at DCMIP2016 are listed in Table 2. They are formally described in Ullrich et al. (2017) and the references therein. Further, specific versions of the code used in DCMIP2016 and access instructions are also listed in Ullrich et al. (2017). Note that not all DCMIP2016 participating groups submitted results for this particular test.

Due to the multitude of differing implicit and explicit diffusion in the participating models, some groups chose to apply variations in how either horizontal or vertical diffusion were treated in this test case. Deviations from the above specified diffusion are as follows. CSU applied uniform three-dimensional second order diffusion with coefficients of $\nu = 1500\,\mathrm{m^2 s^{-1}}$ for q_v and θ_v, $\nu = 1000\,\mathrm{m^2 s^{-1}}$ for q_c and q_r, and $\nu = 500\,\mathrm{m^2 s^{-1}}$ for divergence and relative vorticity. FV^3 applied divergence and vorticity damping separately to the velocity fields along the floating Lagrangian surface. A Smagorinsky diffusion is also applied to the horizontal wind. ICON applied constant horizontal second-order diffusion to the horizontal and vertical velocity components ($\nu = 500\,\mathrm{m^2s^{-1}}$) as well as the scalar variables ~~ρ,~~ θ_v and $q_{v,c,r}$ ($\nu = 1500\,\mathrm{m^2s^{-1}}$). No explicit diffusion was applied in the vertical. NICAM applied a dynamically-defined fourth-order diffusion to all variables in the horizontal with vertical dissipation being implicitly handled by the model's vertical discretization.

3 Results

The following section describes the results of the supercell test case at DCMIP2016, both from a intermodel time evolution perspective and intramodel sensitivity to model resolution and ensuing convergence. Note that there is no analytic solution for the test case, but features specific to supercells should be observed and are subsequently discussed. It is not the intent of this manuscript to formally explore the precise mechanisms for model spread or define particular solutions as superior, but rather, to publish an overview set of results from a diverse group of global, non-hydrostatic models to be used for future development endeavours. Future work employing this test case in a more narrow sense can isolate some of the model design choices that impact supercell simulations.

3.1 Time evolution of supercell at control resolution

Fig. 3 shows the temporal evolution (every 30min, out to 120min test termination) of the supercell for contributing models at the control resolution of 1km. The top four panels for each model highlight a cross-section at 5km elevation through vertical velocity (w) while the bottom four show a cross-section (at the same elevation) through the rain water (q_r) field produced by the Kessler microphysics. For w, red contours represent rising motion while blue contours denote sinking air. Note that the longitudes plotted vary slightly in each of the four time panes to account for zonal movement. This analysis framework closely follows that originally outlined in Klemp et al. (2015).

All model solutions show bulk similarities. With respect to vertical velocity, a single, horseshoe-shaped updraft is noted at 30min in all models, although the degree to which the maximum updraft velocities are centered on the equator vary. A corresponding downdraft is located immediately to the east of the region of maximum positive vertical velocity. This downdraft is single-lobed (e.g., ACME-A) or double-lobed (e.g., GEM) in all simulations. Separation of the initial updraft occurs by 60min across all models, although variance begins to develop in the meridional deviation from the equator of the splitting supercell. Models such as NICAM, FV^3, OLAM, and ICON all have larger and more distinct north-south spatial separation, while FVM, GEM, ACME-A, and TEMPEST show only a ~~couple degrees~~ few degrees of latitude between updraft cores.

Structural differences also begin to emerge at 60min. For example, FVM, GEM, ACME-A, and TEMPEST all exhibit three local maxima in vertical velocity~~at 60min~~; two large updrafts mirrored about the equator with one small maximum still located over equator centered near the initial perturbation. Similar behavior is noted in the q_r fields. This is in contrast with other models which lack a third updraft on the equatorial plane. Generally speaking, q_r maxima are collocated with the locations of maximum updraft velocities, and thereby conversion from q_v and q_c to q_r in the Kessler microphysics.

While the aggregate response of a single updraft eventually splitting into poleward-propagating symmetric storms about the equator is well-matched between the configurations, notable differences exist~~through~~ , particularly towards the end of the runs.

~~This spread in model solutions continues to increase through 90min and is most notable at the test conclusion (~~At 120min~~).~~, FVM, GEM, ACME-A, OLAM, and MPAS all show two discrete supercells approximately 30° from the equator. FV^3 and TEMPEST both produce longitudinally-transverse storms that stretch towards the equator in addition to the two main cells. Each of the splitting supercells split a second time in ICON, forming, in conjunction with a local maximum at the equator, five maxima of vertical velocity (and correspondingly rainwater). NICAM produces two core supercells (as more clearly evident in the q_r field at 120min), but has noticeable alternating weak updrafts and downdrafts in the north-south space between the two storm cores.

The relative smoothness of the storms as measured by the vertical velocity and rain water fields also varies between models, particularly at later times. ACME-A, FVM, GEM, OLAM, and MPAS produce updrafts that are relatively free of additional, small-scale local extrema in the vicinity of the core of the splitting supercell. Conversely, CSU, FV^3, ICON, NICAM, and TEMPEST all exhibit solutions with additional convective structures, with multiple updraft maxima versus two coherent cells. This spread is somewhat minimized when looking at rain water, implying that the overall dynamical character of the cells as noted by precipitation generation is more similar, with all models other than ICON showing cohesive rain water maxima O(10 g/kg).

3.2 Resolution sensitivity of supercell

Fig. 4 shows the same cross-section variables as Fig. 3 except across the four specified test resolutions (nominally 4km, 2km, 1km, 0.5km, from left to right) at test termination of 120min~~(+7200sec)~~. Therefore, the third panel from the left for each model (1km) should match the fourth panel from the left for each model in Fig. 3.

As resolution increases (left to right) models show increasing horizontal structure in both the vertical velocity and rain water fields. Updraft velocity generally increases with resolution, particularly going from 4km to 2km, implying that the supercell is underresolved at 4km resolution. This is supported by previous mesoscale simulations investigating supercells in other frameworks (Potvin and Flora, 2015; Schwartz et al., 2017), although it should be emphasized that this response is also subject to each numerical scheme's effective resolution (Skamarock, 2004) and that the resolvability of real-world supercells can depend on the size of individual storms.

At the highest resolutions, there is a distinct group of models that exhibit more ~~noise~~small-scale structure, particularly in vertical velocity, at +120min at higher resolutions. CSU, GEM, and NICAM appear to have the largest vertical velocity variability at 0.5km, while ACME-A, FVM, MPAS, and TEMPEST appear to produce the smoothest solutions. This result is likely due to the differences in explicit diffusion treatment as noted before, as well as differences in the numerical schemes' implicit diffusion~~.~~, particularly given the large impact of dissipation on kinetic energy near the grid scale (Skamarock, 2004; Jablonowski and Williamson, 2011). Additional focused sensitivity runs varying explicit diffusion operators and magnitude may be insightful for developers to explore. It is also ~~assumed~~ hypothesized that differences in the coupling between the dynamical core and subgrid parameterizations may lead to some of these behaviors ~~(Thatcher and Jablonowski, 2016) .~~(e.g., Staniforth et al. (2002); Malardel (2010); Thatcher and Jablonowski (2016); Gross et al. (2018)) although more constrained simulations isolating physics-dynamics coupling in particular modeling frameworks is a target for future work. As before, rain water cross-sections tend to be less ~~noisy~~ spatially variable at 0.5km than vertical velocity, although CSU and NICAM both show some additional local maxima in the field associated with some of the aforementioned w maxima.

3.3 Convergence of global supercell quantities with resolution

While Fig. 4 highlights the structural convergence with resolution ~~, more bulk, integrated~~ more storm-wide measures of supercell intensity are also of interest. Fig. 5 shows the maximum resolved updraft velocity over the global domain as a function of time for each dynamical core and each resolution (~~increasing~~ finer model resolution is denoted by progressively darker lines). Maximum updraft velocity is chosen as a metric of interest due to its common use in both observational and modeling studies of supercells. All models show increasing updraft velocity as a function of resolution, further confirming that, at 4km, the supercell is underresolved dynamically. For the majority of models and integration times, the gap between 4km and 2km grid spacing is the largest in magnitude, with subsequent increases in updraft velocity being smaller as models further decrease horizontal grid spacing. At 0.5km, the majority of models ~~appear~~ are relatively converged, with FV^3, ICON, and MPAS showing curves nearly on top of one another at these resolutions. Other models show larger differences between 0.5km and 1km curves, implying that these configuration may not yet be converged in this bulk sense~~, and further~~ . Further grid refinement or modifications to the dissipation schemes are necessary to achieve convergence; this is left to the invidual modeling groups to verify.

The maximum updraft velocity as a function of resolution for particular model configurations varies quite widely. NICAM produces the weakest supercell, with velocities around 30 m s^{-1} at 0.5km, while ACME-A, TEMPEST, GEM, and CSU all produce supercells that surpass 55 m s^{-1} at some point during the supercell evolution. Models that have weaker supercells at 0.5km tend to also have weaker supercells at 4km (e.g., NICAM) while the same is true for stronger supercells (e.g., TEMPEST), likely due to configuration sensitivity. This agrees with the already discussed structural plots (Fig. 4) which demonstrated model solutions were generally converging with resolution on an intramodel basis but not necessarily across models.

Fig. 6 shows the same analysis except for area-integrated precipitation rate for each model and each resolution. Similar results are noted as above – with most models showing large spread at ~~4km~~ the coarsest resolutions, but general convergence in precipitation by 0.5km. All models produce the most precipitation at 120min with the 4km simulation. This is consistent with Klemp et al. (2015), who postulated this behavior is due to increased spatial extent of available q_r to fall out of the column at these grid spacings, even though updraft velocities are weaker at coarser resolutions. Unlike maximum vertical velocity, the integrated precipitation rate does not monotonically increase with resolution for most models. At 120min, integrated rates at 0.5km range by approximately a factor of two, from a low of 70-75 x10^5 kg s^{-1} (ACME-A, FVM, OLAM) to a high of 150-160 x10^5 kg s^{-1} (CSU, FV^3), highlighting the sensitivity of final results that have been already been discussed.

4 Conclusions

Non-hydrostatic dynamics are required for accurate representation of supercells. The results from this test case show that clear differences and uncertainties exist in storm evolution when comparing identically initialized dynamical cores at similar nominal grid resolutions. Intramodel convergence in bulk, integrated quantities appears to generally occur at approximately 0.5km grid spacing. However, intermodel differences are quite large even at these resolutions. For example, maximum updraft velocity within a storm between two models may vary by almost a factor of two even at the highest resolutions assessed at DCMIP2016.

Structural convergence is weaker than bulk integrated metrics. Two-dimensional horizontal cross-sections through the supercells at various times show that some models are well-converged between 1km and 0.5km, while results from other models imply that finer resolutions are needed to assess whether convergence will occur with a particular test case formulation and model configuration. Interestingly, in some cases maximum, bulk quantities converge faster than snapshots of cross-sections.

We postulate that these differences and uncertainties likely stem from not only the numerical discretization and grid differences outlined in Ullrich et al. (2017), but also from the form and implementation of filtering mechanisms (either implicit or explicit) specific to each modeling center. The simulation of supercells at these resolutions are particularly sensitive to numerical diffusion since damping of prognostic variables in global models is occurring at or near the scales required for resolvability of the storm. This is different from other DCMIP2016 tests (baroclinic wave and tropical cyclone), which produced dynamics that were less non-hydrostatic in nature and required resolvable scales well coarser than the grid cell level. Further, since DCMIP2016 did not formally specify a particular physics-dynamics coupling strategy, it would not be surprising for particular design choices regarding how the dynamical core is coupled to subgrid parameterizations to also impact results.

Given the lack of an analytic solution, we emphasize that the goal of this paper is not to define particular supercells as optimal answers. Rather, the main intention of this test at DCMIP2016 was to produce a verifiable database for models to use as an initial comparison point when evaluating non-hydrostatic numerics in dynamical cores. Pushing grid spacings to 0.25km and beyond to formalize convergence would be a useful endeavour in future application of this test, either at the modeling center level or as part of future iterations of DCMIP. Variable-resolution or regionally-refined dynamical cores may reduce the burden of such simulations, making them more palatable for researchers with limited computing resources.

We acknowledge that as groups ~~continued~~ continue to develop non-hydrostatic modeling techniques that small changes in the treatment of diffusion in the dynamical core will likely lead to changes in their results from DCMIP2016. We recommend ~~that~~ modeling centers developing or optimizing non-hydrostatic dynamical cores perform this test and compare their solutions to the baselines contained in this manuscript as a check of sanity relative to a large and diverse group of next-generation dynamical cores ~~currently under development in~~ actively being developed within the atmospheric modeling community.

Code availability

Information on the availability of source code for the models featured in this paper can be found in Ullrich et al. (2017). For this particular test, the initialization routine, microphysics code, and sample plotting scripts are available at http://dx.doi.org/10.5281/zenodo.1298671.

Author contributions. CMZ prepared the text and corresponding figures in this manuscript. Model data and notations about model-specific configurations were provided by the individual modeling groups. PAU assisted with formatting of the test case description in Section 2.

Acknowledgements. DCMIP2016 is sponsored by the National Center for Atmospheric Research Computational Information Systems Laboratory, the Department of Energy Office of Science (award no. DE-SC0016015), the National Science Foundation (award no. 1629819), the National Aeronautics and Space Administration (award no. NNX16AK51G), the National Oceanic and Atmospheric Administration Great Lakes Environmental Research Laboratory (award no. NA12OAR4320071), the Office of Naval Research and CU Boulder Research Computing. This work was made possible with support from our student and postdoctoral participants: Sabina Abba Omar, Scott Bachman, Amanda Back, Tobias Bauer, Vinicius Capistrano, Spencer Clark, Ross Dixon, Christopher Eldred, Robert Fajber, Jared Ferguson, Emily Foshee, Ariane Frassoni, Alexander Goldstein, Jorge Guerra, Chasity Henson, Adam Herrington, Tsung-Lin Hsieh, Dave Lee, Theodore Letcher, Weiwei Li, Laura Mazzaro, Maximo Menchaca, Jonathan Meyer, Farshid Nazari, John O'Brien, Bjarke Tobias Olsen, Hossein Parishani, Charles Pelletier, Thomas Rackow, Kabir Rasouli, Cameron Rencurrel, Koichi Sakaguchi, Gökhan Sever, James Shaw, Konrad Simon, Abhishekh Srivastava, Nicholas Szapiro, Kazushi Takemura, Pushp Raj Tiwari, Chii-Yun Tsai, Richard Urata, Karin van der Wiel, Lei Wang, Eric Wolf, Zheng Wu, Haiyang Yu, Sungduk Yu and Jiawei Zhuang. We would also like to thank Rich Loft, Cecilia Banner, Kathryn Peczkowicz and Rory Kelly (NCAR), Carmen Ho, Perla Dinger, and Gina Skyberg (UC Davis) and Kristi Hansen (University of Michigan) for administrative support during the workshop and summer school.

References

Browning, K. A.: Airflow and Precipitation Trajectories Within Severe Local Storms Which Travel to the Right of the Winds, Journal of the Atmospheric Sciences, 21, 634–639, doi:10.1175/1520-0469(1964)021<0634:AAPTWS>2.0.CO;2, 1964.

Doswell, C. A. and Burgess, D. W.: Tornadoes and toraadic storms: A review of conceptual models, The Tornado: Its Structure, Dynamics, Prediction, and Hazards, pp. 161–172, 1993.

Gross, M., Wan, H., Rasch, P. J., Caldwell, P. M., Williamson, D. L., Klocke, D., Jablonowski, C., Thatcher, D. R., Wood, N., Cullen, M., and et al.: Physics–Dynamics Coupling in Weather, Climate, and Earth System Models: Challenges and Recent Progress, Monthly Weather Review, 146, 3505–3544, doi:10.1175/mwr-d-17-0345.1, http://dx.doi.org/10.1175/MWR-D-17-0345.1, 2018.

Jablonowski, C. and Williamson, D. L.: The pros and cons of diffusion, filters and fixers in atmospheric general circulation models, in: Numerical Techniques for Global Atmospheric Models, pp. 381–493, Springer, 2011.

Ji, M. and Toepfer, F.: Dynamical Core Evaluation Test Report for NOAA's Next Generation Global Prediction System (NGGPS), Tech. rep., National Oceanic and Atmospheric Administration, https://repository.library.noaa.gov/view/noaa/18653, 2016.

Kessler, E.: On the Distribution and Continuity of Water Substance in Atmospheric Circulations, pp. 1–84, American Meteorological Society, Boston, MA, doi:10.1007/978-1-935704-36-2_1, 1969.

Klemp, J. B. and Wilhelmson, R. B.: The simulation of three-dimensional convective storm dynamics, Journal of the Atmospheric Sciences, 35, 1070–1096, 1978.

Klemp, J. B., Skamarock, W. C., and Park, S.-H.: Idealized global nonhydrostatic atmospheric test cases on a reduced-radius sphere, Journal of Advances in Modeling Earth Systems, 7, 1155–1177, doi:10.1002/2015MS000435, 2015.

Kuang, Z., Blossey, P. N., and Bretherton, C. S.: A new approach for 3D cloud-resolving simulations of large-scale atmospheric circulation, Geophysical Research Letters, 32, doi:10.1029/2004gl021024, 2005.

Lemon, L. R. and Doswell, C. A.: Severe Thunderstorm Evolution and Mesocyclone Structure as Related to Tornadogenesis, Monthly Weather Review, 107, 1184–1197, doi:10.1175/1520-0493(1979)107<1184:STEAMS>2.0.CO;2, 1979.

Malardel, S.: Physics–dynamics coupling, in: Proceedings of ECMWF Workshop on Non-hydrostatic Modelling, vol. 6777, 2010.

Potvin, C. K. and Flora, M. L.: Sensitivity of Idealized Supercell Simulations to Horizontal Grid Spacing: Implications for Warn-on-Forecast, Monthly Weather Review, 143, 2998–3024, doi:10.1175/MWR-D-14-00416.1, 2015.

Reed, K. A., Bacmeister, J. T., Rosenbloom, N. A., Wehner, M. F., Bates, S. C., Lauritzen, P. H., Truesdale, J. E., and Hannay, C.: Impact of the dynamical core on the direct simulation of tropical cyclones in a high-resolution global model, Geophysical Research Letters, 42, 3603–3608, doi:10.1002/2015GL063974, 2015.

Rotunno, R.: Supercell thunderstorm modeling and theory, The tornado: Its structure, dynamics, prediction, and hazards, pp. 57–73, 1993.

Rotunno, R. and Klemp, J.: On the Rotation and Propagation of Simulated Supercell Thunderstorms, Journal of the Atmospheric Sciences, 42, 271–292, doi:10.1175/1520-0469(1985)042<0271:OTRAPO>2.0.CO;2, 1985.

Rotunno, R. and Klemp, J. B.: The Influence of the Shear-Induced Pressure Gradient on Thunderstorm Motion, Monthly Weather Review, 110, 136–151, doi:10.1175/1520-0493(1982)110<0136:TIOTSI>2.0.CO;2, 1982.

Schwartz, C. S., Romine, G. S., Fossell, K. R., Sobash, R. A., and Weisman, M. L.: Toward 1-km Ensemble Forecasts over Large Domains, Monthly Weather Review, 145, 2943–2969, doi:10.1175/MWR-D-16-0410.1, 2017.

Skamarock, W. C.: Evaluating Mesoscale NWP Models Using Kinetic Energy Spectra, Monthly Weather Review, 132, 3019–3032, doi:10.1175/MWR2830.1, 2004.

Smolarkiewicz, P. K., Kühnlein, C., and Grabowski, W. W.: A finite-volume module for cloud-resolving simulations of global atmospheric flows, Journal of Computational Physics, 341, 208–229, doi:10.1016/j.jcp.2017.04.008, 2017.

Staniforth, A., Wood, N., and Côté, J.: Analysis of the numerics of physics–dynamics coupling, Quarterly Journal of the Royal Meteorological Society, 128, 2779–2799, doi:10.1256/qj.02.25, 2002.

Thatcher, D. R. and Jablonowski, C.: A moist aquaplanet variant of the Held–Suarez test for atmospheric model dynamical cores, Geoscientific Model Development, 9, 1263–1292, doi:10.5194/gmd-9-1263-2016, https://www.geosci-model-dev.net/9/1263/2016/, 2016.

Ullrich, P. A., Jablonowski, C., Kent, J., Lauritzen, P. H., Nair, R., Reed, K. A., Zarzycki, C. M., Hall, D. M., Dazlich, D., Heikes, R., Konor, C., Randall, D., Dubos, T., Meurdesoif, Y., Chen, X., Harris, L., Kühnlein, C., Lee, V., Qaddouri, A., Girard, C., Giorgetta, M., Reinert, D., Klemp, J., Park, S.-H., Skamarock, W., Miura, H., Ohno, T., Yoshida, R., Walko, R., Reinecke, A., and Viner, K.: DCMIP2016: A Review of Non-hydrostatic Dynamical Core Design and Intercomparison of Participating Models, Geoscientific Model Development Discussions, pp. 1–49, doi:10.5194/gmd-2017-108, 2017.

Wedi, N. P. and Smolarkiewicz, P. K.: A framework for testing global non-hydrostatic models, Quarterly Journal of the Royal Meteorological Society, 135, 469–484, doi:10.1002/qj.377, 2009.

Weisman, M. L. and Klemp, J. B.: The Dependence of Numerically Simulated Convective Storms on Vertical Wind Shear and Buoyancy, Monthly Weather Review, 110, 504–520, doi:10.1175/1520-0493(1982)110<0504:TDONSC>2.0.CO;2, 1982.

Zhao, M., Held, I. M., and Lin, S.-J.: Some Counterintuitive Dependencies of Tropical Cyclone Frequency on Parameters in a GCM, Journal of the Atmospheric Sciences, 69, 2272–2283, doi:10.1175/JAS-D-11-0238.1, 2012.

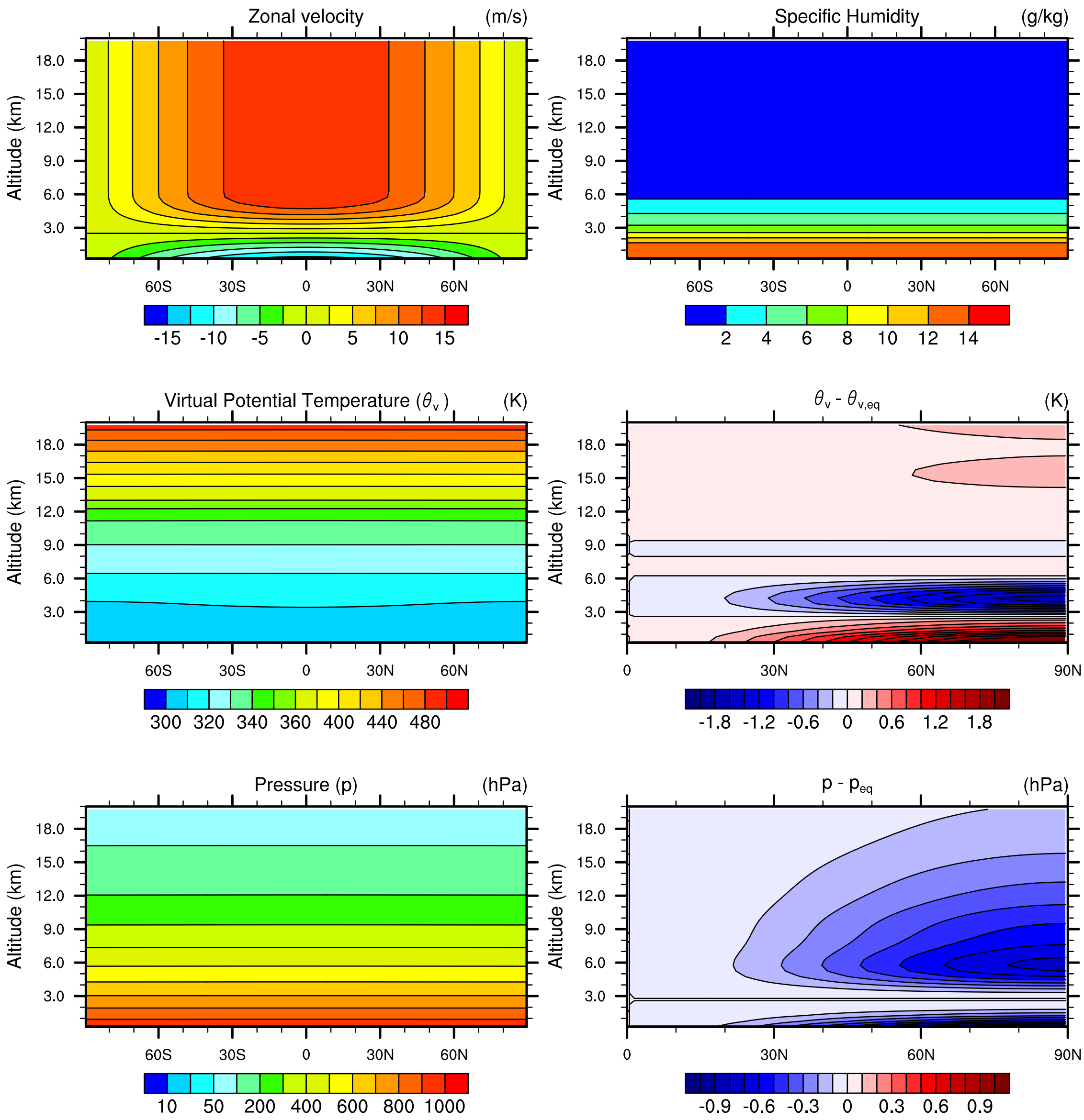

Figure 1. Initial state for the supercell test. All plots are latitude-height slices at 0° longitude. Deviations from equatorial values are shown for virtual potential temperature and pressure.

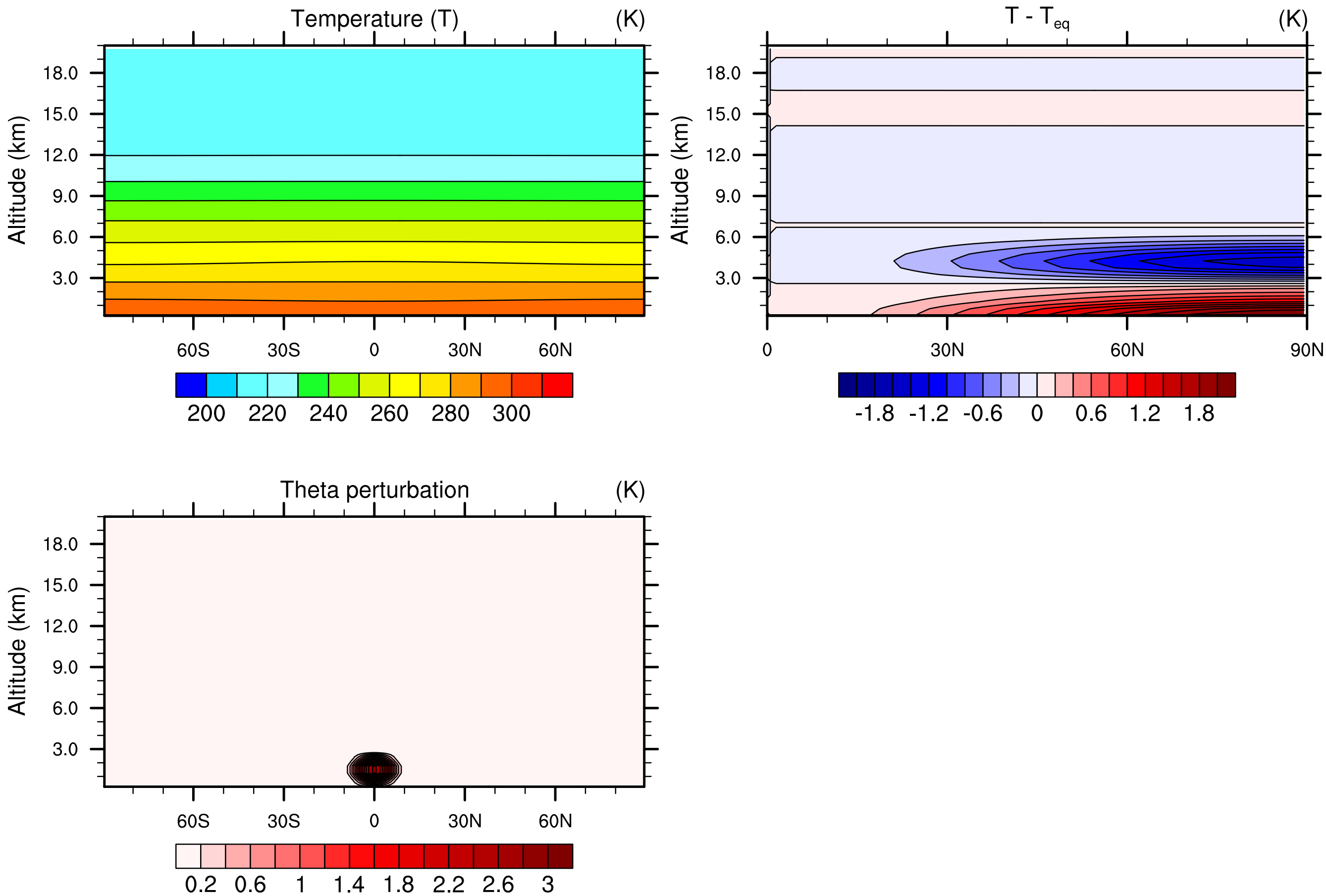

Figure 2. Same as Fig. 1 for temperature and potential temperature.

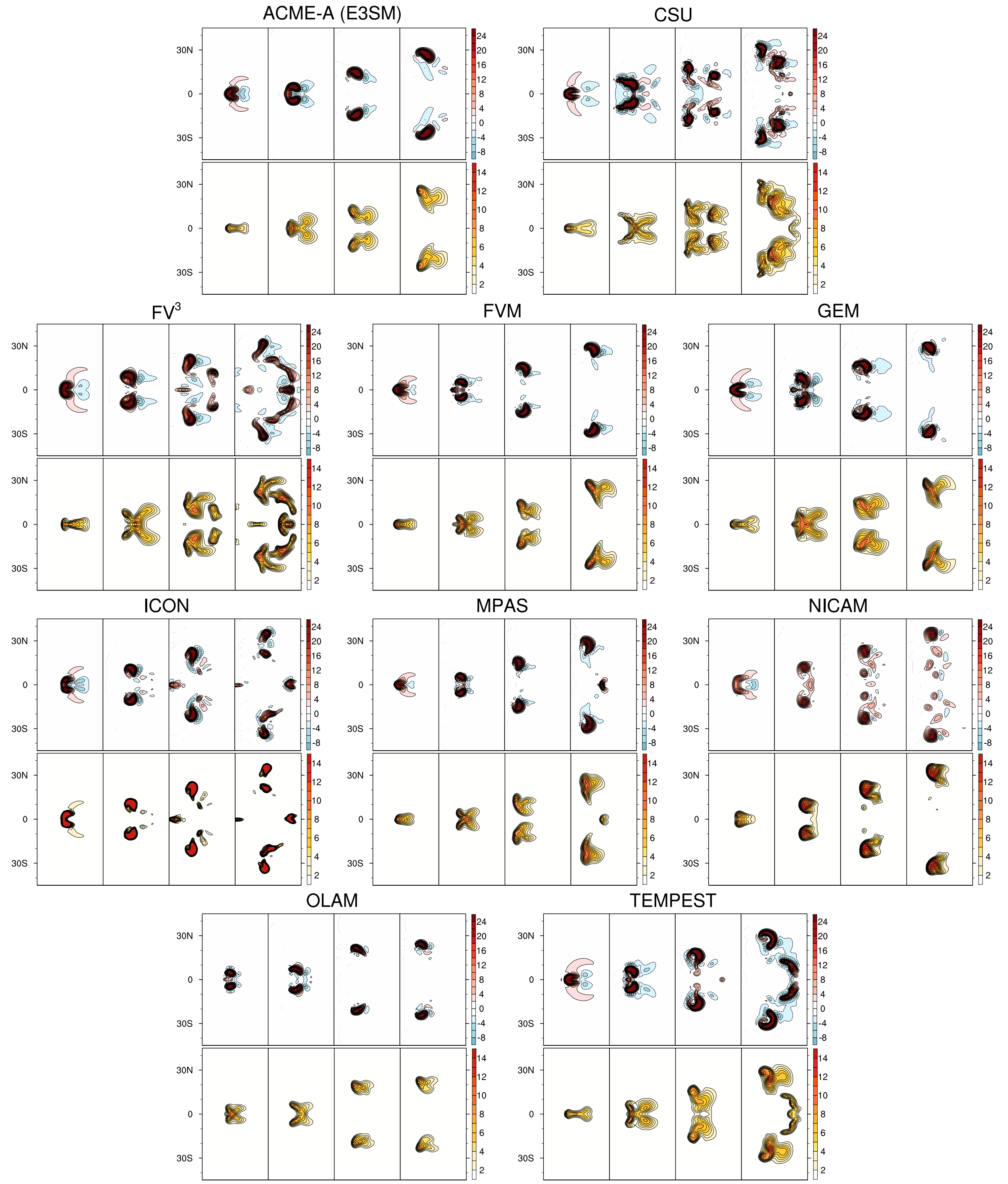

Figure 3. Time evolution of cross-sections of 5km vertical velocity (top) and 5km rain water (bottom) for each model with the r100 configuration of the test case. From left to right, fields are plotted at 30min~~(1800s)~~, 60min~~(3600s)~~, 90min~~(5400s)~~, and 120min~~(7200s)~~.

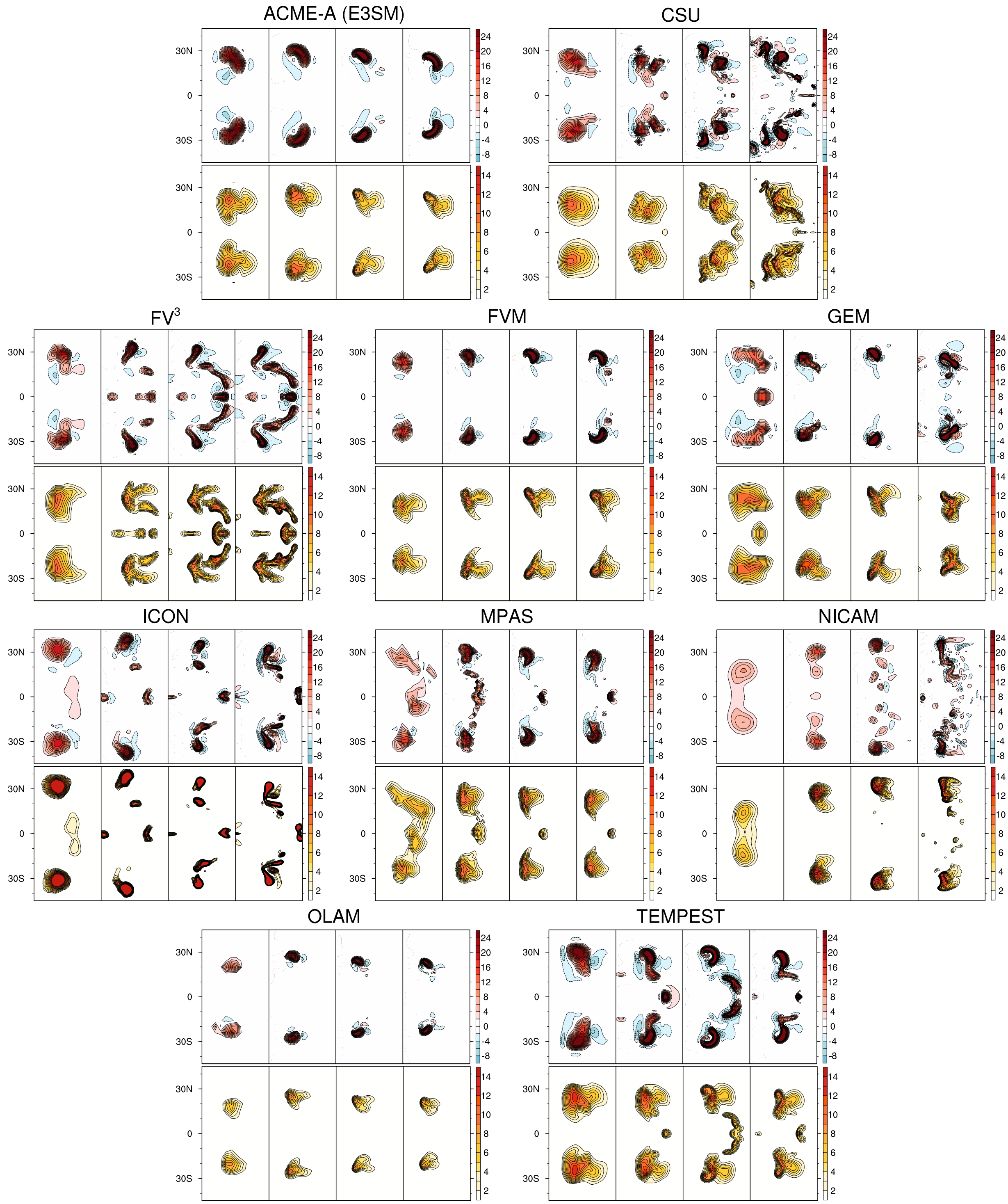

Figure 4. Resolution sensitivity of cross-sections of 5km vertical velocity (top) and 5km rain water (bottom) plotted at 120min ~~(7200s)~~ for each model. From left to right, nominal model resolutions are 4km, 2km, 1km, and 0.5km.

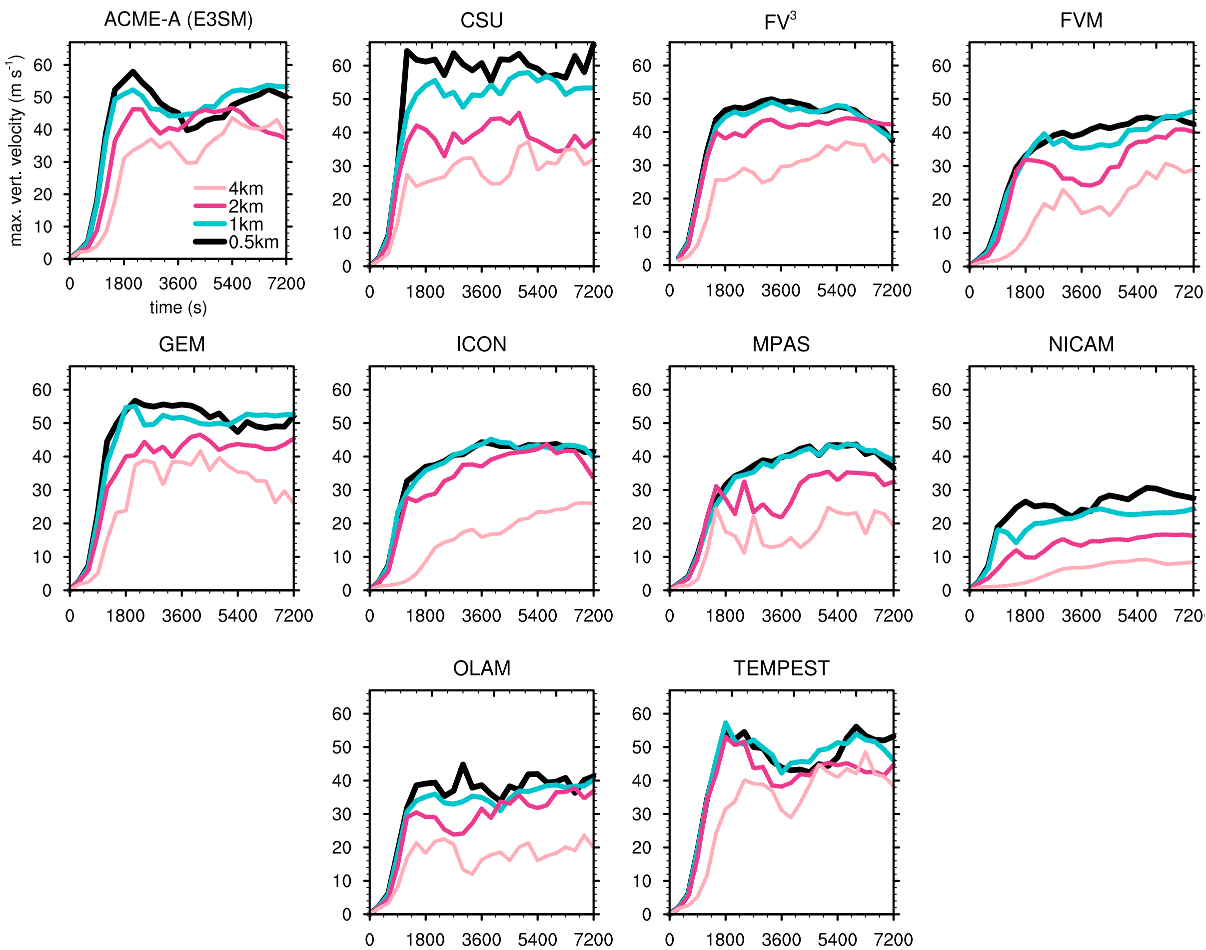

Figure 5. Maximum domain updraft velocity (m s^{-1}) as a function of time for each model at each of the four specified resolutions. ~~Increasing darkness of each~~ Note that the black line ~~denotes increased resolution~~ is the finest grid spacing (~~from 4km to~~ 0.5km) in this test.

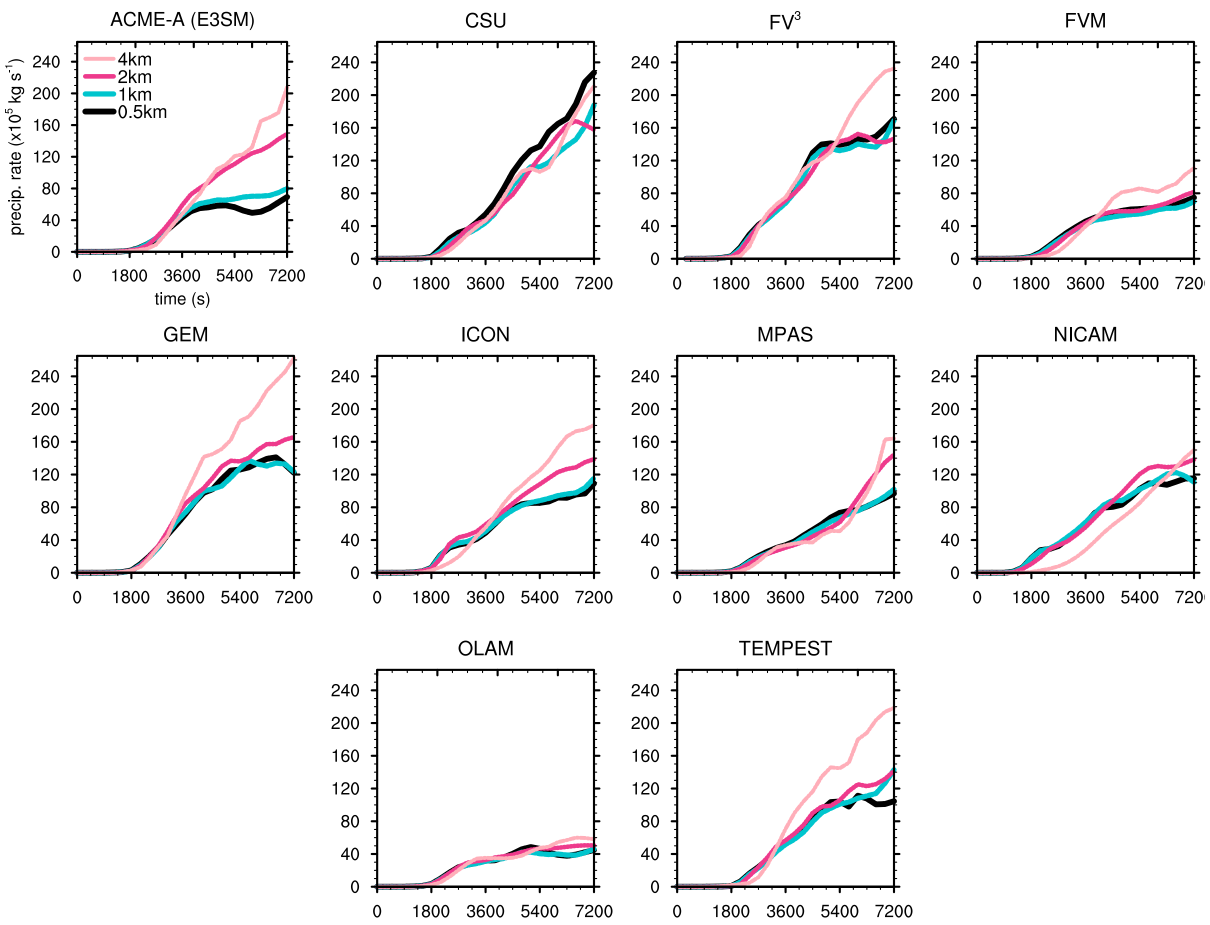

Figure 6. Same as Fig. 5 except showing area-integrated instantaneous precipitation rate (x10^5 kg s^{-1}).

Table 1. List of constants used for the Supercell test.

Constant	Value	Description
X	120	Small-planet scaling factor (reduced Earth)
θ_{tr}	343 K	Temperature at the tropopause
θ_0	300 K	Temperature at the equatorial surface
z_{tr}	12000 m	Altitude of the tropopause
T_{tr}	213 K	Temperature at the tropopause
U_s	30 m/s	Wind shear velocity
U_c	15 m/s	Coordinate reference velocity
z_s	5000 m	Height of shear layer top
Δz_u	1000 m	Transition distance of velocity
$\Delta\theta$	3 K	Thermal perturbation magnitude
λ_p	0	Thermal perturbation longitude
φ_p	0	Thermal perturbation latitude
r_p	$X \times 10000$ m	Perturbation horizontal half-width
z_c	1500 m	Perturbation center altitude
z_p	1500 m	Perturbation vertical half-width

Table 2. Participating modeling centers and associated dynamical cores that submitted results for the splitting supercell test.

Short Name	Long Name	Modeling Center or Group
ACME–A (E3SM)	Energy Exascale Earth System Model	Sandia National Laboratories and University of Colorado, Boulder, USA
CSU	Colorado State University Model	Colorado State University, USA
FV^3	GFDL Finite-Volume Cubed-Sphere Dynamical Core	Geophysical Fluid Dynamics Laboratory, USA
FVM	Finite Volume Module of the Integrated Forecasting System	European Centre for Medium-Range Weather Forecasts
GEM	Global Environmental Multiscale model	Environment and Climate Change Canada
ICON	ICOsahedral Non-hydrostatic model	Max-Planck-Institut für Meteorologie / DWD, Germany
MPAS	Model for Prediction Across Scales	National Center for Atmospheric Research, USA
NICAM	Non-hydrostatic Icosahedral Atmospheric Model	AORI / JAMSTEC / AICS, Japan
OLAM	Ocean Land Atmosphere Model	Duke University / University of Miami, USA
TEMPEST	Tempest Non-hydrostatic Atmospheric Model	University of California, Davis, USA