Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-255-RC2, 2020

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "Impact of scale-aware deep convection on the cloud liquid and ice water paths and precipitation using the Model for Prediction Across Scales (MPAS-v5.2)" *by* Laura D. Fowler et al.

Anonymous Referee #2

Received and published: 27 January 2020

In this manuscript, the authors utilize satellite data (CERES and TRMM), primarily to evaluate cloud fraction (CF), ice water path (IWP), liquid water path (LWP), and precipitation over the tropical Pacific Ocean using the Model for Prediction Across Scales (MPAS). They evaluate 30-day simulations with four configurations: two horizontal grids (a uniform 30km mesh and a variable-resolution mesh with 6km refinement) and two convective parameterizations (Grell-Freitas (GF) and Multi-Scale Kain-Fritsch (MSKF)). Variable-resolution simulations show increased mass flux, stronger updrafts, and a higher fraction of precipitation from large-scale condensation, in accordance with past

Printer-friendly version

Discussion paper

work. They find some systematic biases across all configurations (e.g., IWP uniformly too high, which leads to TOA biases), as well as differences between GF and MSKF simulations (e.g., shallow convective treatment leading to LWP being overestimated with GF). They also note that resolution alone does not necessarily correct solution errors and propose using this framework moving forward to better evaluate moist physical processes, particularly in model configurations used across a variety of grid spacings.

In general, the manuscript is grammatically sound, albeit with a few awkward phrasings and typographical errors occasionally cropping up. The figures are clear and color maps easy to visualize (at least for a non color-blind individual). The notion of using satellite data to evaluate moist processes in weather/climate models is not overly novel, but one that is still used less frequently than it perhaps should be. Therefore, I appreciate the author effort in this endeavor and think this work has the potential to be a useful reference for the numerical modeling community, particularly as weather and climate models push into the gray zone where the line between resolved and parameterized precipitation and clouds is quite blurry. The potential impact of this work is also heightened given the growing interest over the past few years in using variable-resolution grid configurations for regional weather and climate applications.

That said, I have two primary concerns with the manuscript regarding assessing climatology from single-simulation, 30-day averages and the impact of differing physics timesteps and numerical diffusion between the two grid configurations. While my experience (and other previously published work) implies that the latter concern may necessitate additional simulations, I am not fully confident in this. Therefore, my recommendation would be for major revisions, with publication possibly warranted assuming both points are addressed either A) by providing evidence that these concerns are unfounded (or minor relative to the conclusions drawn by the manuscript) or B) with additional simulations to constrain said free variables and/or evaluate their impact. I have tried to highlight specific questions that could be answered.

One other note, the manuscript is quite lengthy and therefore an intensive read. This isn't specific to the authors, as this (longer papers) has been a common trend lately in the model evaluation field. That said, for the sake of readability, a suggestion of mine would be to undertake a fairly deep assessment of what figures/passages are critical to the overall results and eliminate superfluous text where possible. In some cases the authors expend a great deal of time describing a figure panel-by-panel even if the description is not critical to their findings. This both "muddies the water" with judging what the authors deem to be important and what is not and makes for a longer read, therefore, leaving those non-critical observations to the reader may be warranted. Another example of reduction could be model configuration; the authors overwhelmingly detail MPAS, but aspects such as the dynamical core, land surface, gravity wave parameterizations, etc. are not explicitly discussed again in the manuscript. Perhaps these can be left for an in-line citation and readers can be referred to a different overview paper if they desire such information. Please note that these are merely suggestions and I will leave final judgement to the editor(s) and authors.

Major concerns:

1.) Applicability of 30-day runs for evaluating "average" model state, systematic biases

The authors run a single, deterministic 30-day simulation for each of the four model configurations (after 2 day model spinup, initialized from ERA-Interim) and then assess the average solution results against the satellite data. Anecdotally, I have seen model evaluation of this type follow two general time horizons. Numerical weather prediction is assessed on deterministic scales (out to 10-14 days or so) and addresses whether individual features or phenomena are well-simulated. Climate models, conversely, are typically evaluated over multiple seasons (if not years/decades) to minimize the impact of internal variability projecting onto the mean state.

30 days is a duration that doesn't fit neatly into one of these bins. Some initial condition persistence constrains the solutions early in the simulation (probably the first week or so), with solutions likely diverging to more "chaotic" solutions after that. Therefore,

differences in aspects such as the ITCZ location (noted in Lines 545-546, for example) may not be *systematic* for that grid/configuration but may just be "accidental" due to the internal variability of the system as it develops its own climatology.

Primary question: Since the model is not run for long enough times to "average" this internal variability (at least, given commonly accepted temporal scales), how confident can the authors be that the solutions they are seeing are robust? If they ran ten different Decembers, would the same general results arise each time? In other words, are we sure that large-scale biases in the difference plots are truly differences due to physics/resolution and not due to a "poorer" weather forecast over the month in question?

One potential way to at least think about this would be to evaluate on shorter scales (perhaps the first week of each simulation) and see if the same behaviors are apparent. Some work by Steve Klein and Hsi-Yen Ma (see Cloud-Associated Parameterizations Testbed (CAPT)) in the past decade or so has shown that errors in moist climate parameterizations (particularly clouds and precipitation) are fast-evolving and obvious after just a few days in an initialized hindcast. Seeing if the four simulations here show similar errors in days 1-7 when compared to the full average may be insightful. I understand that this type of analysis reduces the satellite coverage and may introduce other averaging errors, but the only way to truly evaluate this is to run multiple years (if not longer) of MPAS in "climate mode" which seems like a more difficult path, at least given the simulations already completed.

2.) Potential sensitivity of moist results to A) physics timestep and B) numerical diffusion

Perhaps my more substantive concern is the differences in the coarse domain ("far-field") of the variable-resolution mesh (V) relative to the uniform run (U). For both GF and MSKF, large mean differences are sometimes seen even in this far-field (where resolution is identical = 30km). Some examples are Fig. 6e-f, Fig. 7e-f, Fig. 11, Fig 14e, as well as the results in Tables 2 and 3. In many cases these 30km-30km differences are of equivalent magnitude to the differences between the 30km-6km grid spacings.

The authors hypothesize this is due to upscale effects (I interpret this as the high-resolution patch impacting larger circulation patterns), although given the dominant flow patterns in this region, it seems implausible that it is a singular factor given that this behavior occurs on both the upwind and downwind sides of the mesh. Three papers come to mind that haven't shown this response with longer simulations, two of which used MPAS (Rauscher et al., 2013, Zarzycki et al., 2014, Sakaguchi et al., 2015).

Questions: Do the authors have any evidence that supports the large "fingerprinting" of the high-resolution patch on some of these moist quantities even far away from the patch itself? What does the simulation look like on the other side of the globe between the U and V runs (where the model should be performing essentially identically in the absence of global upscaling)?

Admittedly, this is a bit of a leading question; in Table 1, it appears that both the physics timestep and numerical diffusion are different between the U and V simulations. Therefore, while the grid is different, there are also differences in model viscosity and timestepping that complicate the apples-to-apples comparison of the two sets of simulations (i.e., there are actually three key run-time differences, not just the model grid).

First, can the authors confirm that the timestep reduction between U and V is for *both* the dynamics and physics? Or is the dynamics subcycled 5x in the V run (such that the physics timestep is still 150s)? Second, is the diffusion (length scale of 6km) uniform across *all* cells or is it scale-aware in the V simulations (i.e., 6km in the high resolution patch, 30km elsewhere)? If the latter, then the far-field should see 30km in both U and V.

One potentially relevant finding that should be addressed (and is not discussed in this manuscript) is that of Williamson (2013) who found that merely varying the physics timestep (holding all other aspects of the model identical) produced differing responses in the moist physics. This is due to the fact that large-scale precipitation "operates" on the physics timestep, while convective parameterizations are dominated by relaxation timescales that do *not* necessarily scale with the physics timestep. In the aforementioned study, both parameterized and resolved precipitation did not adjust "in lock step," therefore differences in simulated precipitation (and precipitation rates) arose. This may be particularly relevant for MSKF, which applies an adjustment time scale for deep and shallow convection that is explicitly dependent on local grid spacing alone (not timestep) (see Eq. 2).

Using this example, a 30km grid cell in MPAS could (in theory) be seeing exactly the same adjustment scale between V and U, but the large-scale precipitation sees two different timesteps (30s or 150s). It would be insightful to evaluate how much (if at all) this configuration mismatch influences the moist variables evaluated here. An obvious way to verify this is to run one of the U simulations with a 30s timestep and the same diffusion length scale as the V simulation (call this sensitivity run UV). While the timestep may be overly restrictive computationally, it would provide a true apples-to-apples comparison in the far-field when compared to V. If the results (difference between UV and V) are the same as contained in this manuscript, it would lend strong support for hypothesized upscale effects and pure resolution dependence between GFU and GFV (MSKFU and MSKFV). If the differences between UV and V in the far-field become much weaker, however, this would imply that the differing numerical settings are playing a role in mean-state differences the authors are seeing (above and beyond the "pure" impact of grid spacing).

What are the authors' thoughts on the above? Perhaps this has been evaluated previously and can be referenced (either in the supporting information or by pointing to a different manuscript)?

I will also note that this resolution sensitivity and the partitioning of large-scale and convective precipitation has been addressed by a few manuscripts in the Earth system modeling community recently. Some references of interest are listed below. This is certainly not exhaustive, but it may be useful for the authors to see what (if any) of the below studies would be applicable to their results.

- O'Brien, T. A., Collins, W. D., Kashinath, K., Rübel, O., Byna, S., Gu, J., et al. (2016). Resolution dependence of precipitation statistical fidelity in hindcast simulations. Journal of Advances in Modeling Earth Systems, 8(2), 976–990. https://doi.org/10.1002/2016ms000671

- Herrington, A. R., & Reed, K. A. (2017). An Explanation for the Sensitivity of the Mean State of the Community Atmosphere Model to Horizontal Resolution on Aquaplanets. Journal of Climate, 30(13), 4781–4797. https://doi.org/10.1175/jcli-d-16-0069.1

- Rauscher, S. A., Ringler, T. D., Skamarock, W. C., & Mirin, A. A. (2013). Exploring a Global Multiresolution Modeling Approach Using Aquaplanet Simulations. Journal of Climate, 26(8), 2432–2452. https://doi.org/10.1175/jcli-d-12-00154.1

- Williamson, D. L. (2012). The effect of time steps and time-scales on parametrization suites. Quarterly Journal of the Royal Meteorological Society, 139(671), 548–560. https://doi.org/10.1002/qj.1992

- Zarzycki, C. M., Levy, M. N., Jablonowski, C., Overfelt, J. R., Taylor, M. A., & Ullrich, P. A. (2014). Aquaplanet Experiments Using CAM's Variable-Resolution Dynamical Core. Journal of Climate, 27(14), 5481–5503. https://doi.org/10.1175/jcli-d-14-00004.1

- Sakaguchi, K., Leung, L. R., Zhao, C., Yang, Q., Lu, J., Hagos, S., et al. (2015). Exploring a Multiresolution Approach Using AMIP Simulations. Journal of Climate, 28(14), 5549–5574. https://doi.org/10.1175/jcli-d-14-00729.1

Minor comments:

Line 148: Typo Line 175-176: What is the reasoning behind choosing one closure

Printer-friendly version

Discussion paper

versus the ensemble approach noted above? Line 194: Can an extra sentence be added that summarizes why 0.7 can be exceeded since this seems relevant to the manuscript? Line 210-211: This sentence is a bit confusing – is the mixed layer only ∼50mb deep then? Line 316: "cloud model to being described..." awkward, would rephrase Line 412: What is the reasoning behind not weighting the MPAS data as a function of height? Or maybe I am misinterpreting this sentence? Lines 513-515: "MSKF reduces convective precipitation more efficiently than GF" is this true in a fractional sense? It seems like GF has a larger baseline such that the percentage decrease is more similar that it looks at first blush in Fig. 8. Line 532: "evaluated" should probably be "compared" Lines 537-538: Should be Figs. 10b-e and Figs. 11a-d (range, not comma). Fig 12: Would add "averaged between 5S and 5N" to the caption. Lines 622-624: Which would be closest to the true way the satellite observes these variables?

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-255, 2019.