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 ∼ 440km/X ∼ 4km.

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

References

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.