gmd-2023-87 : DCMIP2016: the tropical cyclone test case

Justin L. Willson et al.

1 General comments

This paper summarizes some aspects of the results obtained with the academic tropical cyclone test case that was part of the intercomparison project DCMIP2016. The purpose of the publication is to document the results of a set of 9 GCMs which have been selected from among the project participants in order to create a bench-mark for the RJ TC test case that is frequently used by the GCM community to validate model developments.

It is clearly said in the introduction that it is not the purpose of this article to explain the reasons for the differences between the models. In its current state, the article is very descriptive indeed, so descriptive that the scientific level of the paper is not very high for a scientific publication (it looks more like a student report, something like that).

I understand that the purpose of the paper is not a deep analysis of the reasons behind these differences. But a few more basic diagnostics may help to make the article more interesting. For example, the authors suggest the effective resolution as one of the reason for the large differences. I would then suggest to compute global spectra that would help to compare the effective resolution of the different GCMs.

I am also wondering why there are no information about the trajectory of the TC in the participating models. Even if the trajectories are so close that a figure is not necessary, it would be interesting to know about the TC trajectory in the different GCMs.

The difference between the models in this intercomparison is very large (about 50 hPa for MSP, 30 m/s for MWS in Fig.1...). The discrepancy between the results may call into question the utility of these results as a bench mark. Anyway, I would suggest to discuss more openly the implications of such large differences for the TC intensities in the conclusion.

My opinion is that in its current state this article is too "light" to be published as a scientific paper. At the very least, I'd advise adding a few more advanced diagnostics and further interpretation and analysis of the implications of such large differences between the models for a moist test case where "only" the dyncores are different (same initial condition, same horizontal and vertical resolution, same physics package). Also, the style shows a clear lack of experience in writing scientific papers, so it should be carefully checked in a potential revised version.

2 Specific comments

- introduction: it could be useful to say somewhere in the introduction that there is no "truth" for this test case. And then discuss a bit more how such a case can be used by model developers. For example, does the intercomparison gives an ensemble which could be used as an ensemble forecast, i.e. it describes some kind of PDF of the possible TC forecasts.
- l.6 and l.88: "1 km azimuthally average wind speed" : I guess it is 1 km above the surface, is it? (I don't think it is a standard diagnostic, so it should be stated clearly). Also, why

not use the usual 10m wind as in the IBTRACKS data base?

- l.10 : it is not said in the abstract that all GCMs use the same physics package, so it seems strange to conclude about the dynamical cores only. I think it would be important to already mention in the abstract that all models use the same physics parametrizations.
- I suggest to move tables 1, 2 and 3 in an appendix.
- l.99 : it is important to say that it is **the same** simplified physical parametrization package. TC simulations are even more sensitive to the choice of parametrizations than they are to dynamical cores so, in order for the comparison to conclude about the sensitivity with respect to dynamical cores, the physics package must be the same.
- l.175 versus FIG. 1 : there are 9 models, but 10 curves on Fig. 1. What is the difference between CSU-CP and CSU-LZ?
- l.185 : "Simulations were performed ... on an interpolated lat-lon grid" ? I don't think the models use a lat-lon grid for their computation, and anyway, it is in contradiction with the next sentence and table 3. I would also remove the "identical" in the next sentence as they are clearly not identical (maybe the protocol is, but the runs are not).
- l. 196-208 : move the reference to the physics package to the beginning of the paragraph and then give the main characteristics of this simple parametrization package. Also, comment about the lack of deep convection scheme in this package as, for models using grid space of 50-25 km, the convection scheme is essential to get realistic moist processes, especially in the tropics.
- section 2.3 : I suggest to move the commands used for the software "TempestExtremes" in appendix. A more detailed description of how the tracking algorithm works would however be useful.
- l. 273 :"rapid intensification" : there is an "official" definition of rapid intensification (more than +30 kt in 24h). Is it what you mean?
- l.312 : it may be exactly the definition of the center (it depends of the detail of the tracking algorithm, so please give more details in section 2.3) so not a very useful comment if is it actually the case...
- l.319 : Should NICAM really stay in the intercomparison? It should be possible to verify if there is, or not, a problem in the initialization/initial conditions of this model. If there is, NICAM should be removed from the intercomparison. But maybe there isn't and there is an extreme numerical diffusion in this model which is smoothing a lot the TC intensification and change the total mass of the model (that would also explain the pressure increase far from the TC centre). A simple plot of the fields at t=0 and a KE spectra later in the forecast should help decide what's going on with NICAM (and if it is worth keeping it in the intercomparison).
- l.316 : what is exactly the "likely reason", and why only likely?
- l.319 : what else would you expect for an azymutally average mslp or low level wind profile around a TC center? Seems to be a very basic characteristics of a TC. Not sure it needs a reference.

- l. 334-335 : I don't see the point of this sentence here. Even real TCs "look" like that. Maybe this sentence should be moved at the beginning of Fig. 4 description to say that the general shape of the mean vertical wind structure in all models is consistent with what is expected in a TC, and then explain the specific differences between the models.
- l. 384 : If you could "prove" that this is the main reason for these very large differences, it would add a lot of value to the paper. But there may also be other reasons for the differences, such as, for example, the way the physics and dynamics are coupled (where and how the physics is called in the time step, etc). These other reasons could also be discuss in the conclusion.
- conclusion : It would be interesting to discuss the magnitude of the differences with regard to, for example, the mean intensity errors in the current NWP models and the TC intensity trends predicted in the climate change scenarios.
- conclusion : I would also suggest to discuss the fact that this test case doesn't use a convection scheme unlike what is usually done in GCM at this resolution. This may "amplify" the impact of the differences between the dynamical cores in these simulations compared to what it would be in a context where the deep convection is treated by a (common) convection scheme. Not sure, but it could be interesting to discuss this aspect in the conclusion.

3 More technical/form/writing corrections

My mother tongue is not English, but I have nevertheless the impression that the level of language of the text is not very good. The authors should check it carefully (words missing, bad choice of words etc).

- l 8. : "results are generally similar between all models" : not sure it is what is shown by Fig. 1. There is quite a difference between a TC with a MWS of 25 m/s or 55 m/s. You need to be more accurate with what is actually similar.
- l. 65-65 : a linear response to what? Not clear.
- l.100 : what other characteristics? This sentence (starting l. 98) is not very informative.
- l.123 : not sure "equation" is needed here.
- l. 129 : "perturbation pressure" $\rightarrow$ pressure perturbation
- l.133-135 : this paragraph needs to be reformulated. I don't think "There are multiple pressure differences" means something here. Maybe something like "There are several contributions to the pressure perturbation definition". Then, r_p is not a pressure change, it is a distance, same for z_p. These parameters are characteristic scales used to define the shape of the pressure perturbation change in the horizontal and the vertical, they are not the pressure change/decay themselves.
- l. 165 : move the reference to equation 18 just after velocity or at the end of the sentence.
- l. 207 : give a reference for "time-splitting method" and/or give a short explanation. What about the physics-dynamics coupling? Is there any information about that for each model?

- l. 243 : "vary randomly around a fixed mean" : I don't think there are any random processes is these models, reformulate.
- l.244-245 : I think something is missing in the sentence (maybe "to document"?)
- l.264-265: Long and complicated but not very clear sentence : do you mean "decreases linearly with a constant trend".
- l.273-274 : I don't understand what is quantified in section 2.3? What do you mean?
- l.274 : suggestion : the MWS in NICAM....
- suggestion for section 3.1 : move the last sentence to the beginning of the section.
- l.292 : what characteristics? Does it refer to the next sentence or the one before?
- l. 294 : areas **of** high intensity
- l.299 : most linear what?
- l.299 : What does variability means here? Is it fluctuation around the fitting curve?
- l.379: "different dynamical core" then add **but the same physics package**
- l.380 : what do you mean by "physical environment"? Is it the environment of the TC? Or the physics parametrization? Not clear...
- l.382 : be more precise about what you mean by "relatively consistent results" and be more specific about what's different and what's similar (NICAM should probably be treated as an outsider and discuss separately, after checking that there is no error in its setup).