Review of *Configuration and Evaluation of a Global Unstructured Mesh Model based on the Variable-Resolution Approach*

Zhou et al.

General Impressions

This study evaluates the performance of the variable-resolution configuration of a newer global model GRIST, and seeks to understand the various strengths and weaknesses of different refinement meshes. The authors provide results from both dry and moist idealized experiments that illustrate that the solution in the refined regions resemble the uniform high-resolution solutions. While this take home message is clear, I would like to see further analysis/discussion on why the errors tend to be larger in VR compared with the uniform resolution runs, examples that I point out specifically in the comments section, and also how the Smagorinsky operators are implemented in VR. After addressing these minor revisions, I think this manuscript is acceptable for publication in GMD.

Comments

L64: CAM has multiple dycores, each with distinct numerical properties, and so this statement can be misleading. I think the authors should consider mentioning that the Zarzycki study cited used the spectral-element dycore.

L88: This statmentt "[a] series of numerical tests was performed to examine the model reliability under more challenging conditions," reads like there are more chellenging tests than the TC test-case, but the TC test-case is the most complex case used in this study.

L108: If I recall correctly, the Smagorinsky coefficients scale with grid spacing. Is the density function used to determine the Smagorinsky coefficients?

Model and configurations: Can the authors include the number of vertical levels used in the simulations?

L160: The authors argue that the densification ratio should be no larger than 1:4, and point to a citation that I can't seem to get access to. I'm having trouble interpreting this statement. Do the authors mean *no less* than 1:4? Would this then mean the refined grid spacing should be no less than a 1/4 of the coarser region grid spacing? If so, I can think of many spectral-element VR studies that use a much smaller ratio without having reported any serious errors. I could be misunderstanding entirely here, but I think this densification ratio and implications of some lower limit should be spelled out more clearly for the general reader.

L197: The authors keep referring to grid imprinting in this paragraph. Am I to infer that they are only talking about the spurious waves being generated in the southern

hemisphere, in the coarse region of the grid? These features seem to become less noisy when the coarse region increases its resolution, as one would expect. I think it should be stated that the coarser region of G5B3X4 is higher resolution than the coarse region of G6X4.

L218: This assertion seems to be mostly true. But I am struck by the oscillations in northern Alaska that are absent in the uniform resolution runs, and which coincide with the mesh transition zone. I think these are real errors. Similar errors are discussed in the context of the SURX4 grid in the following paragraph, but there is no mention of these oscillations in the other VR grids (albeit, they are less noisy than SURX4).

L270: Similarly, it looks to me that the vorticity field in 8a is rather oscillatory, especially in the tails of the vortices. I think the authors should investigate whether these are real errors, an artifact of the vorticity calculation, or something else. It would also be interesting to understand the sensitivity of these spurious structures (if they are indeed spurious) to the Smagorinsky coefficient.

L288: Can the authors provide the rationale for using different physics-dynamics-tracer coupling methods for hydrostatic vs. non-hydrostatic runs?

L307: "During its movement from the 2nd-refinement into the 1st-refinement region, the change in the grid size leads to little distortion on the tropical cyclone in each experiment." This sentence would be more substantiated if the authors provided a look at how the tropical cyclone fares as it crosses the transition, not just the final structure after it already passed the transition (e.g., Figure 3 in your Zarzycki et al 2013 citation).

L310: The minor disturbance described near where the cyclone was initiated is a common feature of dycores in DCMIP2016. Might be worth looking into whether this result has been published before.

L321: It's unclear to me what the first sentence of this paragraph referencing Ringler has to do with the rest of the paragraph. Could the authors clarify?

L324: "clone" should say "cyclone."

L348: More important than what? I'd suggest removing the "more" from the last sentence.

Conclusions: I would think that the larger errors found using the SUR generator is a notable conclusion of this paper.

Figure 4: In the caption "the quasi-uniform G7 and G8 cases" should probably say "G6 and G7 cases," since the l2 norms are defined w.r.t to G8, no?