The manuscript describes a new numerical model for simulation of atmospheric flows. The focus is on limited-area high-resolution configurations in large-eddy-simulation-type modeling. The manuscript introduces the new model's background, philosophy, and motivation, describes the model formulation – both the continuous PDEs and discrete system, discusses some representative results, and documents parallel computing performance. Overall, the manuscript is well written, and the presentation is clear. The text is concise and provides a fair amount of detail regarding the model formulation and approximations without compromising accuracy and completeness. The main weakness of the manuscript is the presentation of the results which lack quantitative comparisons. A rigorous and quantitative test case is not presented. Overall, I believe the manuscript is suitable for publication in Geoscientific Model Development.

Major comments:

- The Numerical Experiments (Section 5) provide a nice overview of the model capabilities. Unfortunately, most of the results are presented as contour plots or as comparisons with respect to different model parameter choices (e.g., SGS closure) in the present model. This aspect is the main weakness of the manuscript, and it can be improved. For instance: The Taylor-Green vortex is an exact solution of the Navier-Stokes. For laminar flow the convergence rate can be determined. There are other solutions and/or methods that can be used to quantify the convergence rate of solution error, rather than presenting and comparing contour plots.
- Some of the numerical experiments are performed as an LES (the turbulence model is
very active) and some appear to be (almost) a DNS (no turbulence model in 5.2?). This is not clear. The Density Current appears to be a non-LES calculation. However, it appears from Figure 3 that the flow is not fully resolved because some of the solutions have the tendency to generate small scales. It seems that the numerical method is artificially stabilizing the solution by providing numerical dissipation, resulting in an implicit LES. Can the nature of the simulation be clarified? This has implications relating to the reproducibility of the benchmark.

- The results of Section 6 are somewhat misleading. The results verify the conservation property of prognostic variables. The flux formulation of the numerical method guarantees no internal “leaking” of prognostic variables. However, other types of “energy” are not conserved because the method is dissipative. Perhaps the section should be renamed as “Verification of prognostic variable conservation”. This is also stated in the abstract (line 5: “energy-conserving”). “Energy conserving” in numerical methods means that the model conserves second-order moments of prognostic variables, such as kinetic energy and scalar variance and not just the prognostic variables.

Other comments:

- The scaling results are somewhat underwhelming. A maximum of 32 ranks and 16 GPUs is used. A reader might expect that more ranks or GPUs are required by the “ClimateMachine” to simulate Earth’s climate.
- Line 68: Typically, in LES the flow variables are defined as filtered or averaged variables over the volume of the grid cell. A density weighted average is expected (similar to Appendix B). This should be corrected or clarified.
- Line 310: Is theta-v the actual buoyancy variable used in the buoyancy gradient in Ri and it is consistent with the energy equation (5)?
- Typically in LES, the characteristic length scale is modified near the surface, e.g., Mason & Callen (1986, J. Fluid Mech.) is this method applied in the model?
- Line 285: The deviatoric rate of strain tensor $S_{ij} - 1/3 \delta_{ij} \text{trace}(S_{ij})$ should be use in the Smagorinsky model, not the rate of strain which has non-zero trace for non-constant density flows.
- Appendix A3: is this how the BCs are actually applied in the DG method?
- Line 267: “flux tensor” should be “turbulent stress tensor”
- Line 160: is “divergence form” a better term in place of “compact notation”?
- Line 2: The “performance portability” of the model is not demonstrated in the manuscript.
- Line 21: There is a misrepresentation of Smagorinsky (1963) and Lilly (1962) – both here and in other places in the manuscript. These papers do not discuss LES. Smagorinsky (1963) is a pioneering paper about a GCM. Smagorinsky recognized that some form of horizontal dissipation is required to stabilize the GCM since the forward turbulence cascade tends to create smaller scales (similar to Figure 3). He used a simple eddy viscosity parameterization based on the local horizontal rate of strain. Lilly (1962) introduced the TKE parameterization correction for stratified flows which is equation (39) of the current manuscript.

The first paper that starts to resemble modern LES is:

Lilly 1966: On the Application of the Eddy Viscosity Concept in the Inertial Sub-Range of Turbulence, NCAR manuscript 123

Another useful reference is:
Smagorinsky, 1993: Some historical remarks on the use of nonlinear viscosities.