ADVANCING EARTH AND SPACE SCIENCE

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2020MS002280

Key Points:

- We present an Low Mach number Approximate Riemann Solver-based unstaggered shallow-water model on arbitrary gnomonic cubed-sphere grids
- The solver demonstrates a broad range of diffusion control without any explicit filters
- A newly introduced splash on the sphere test verifies the solver's desirable dispersion properties

Correspondence to:

X. Chen,
xic@princeton.edu

Citation:

Chen, X. (2021). The LMARS based shallow-water dynamical core on generic gnomonic cubed-sphere geometry. *Journal of Advances in Modeling Earth Systems*, *13*, e2020MS002280. https://doi.org/10.1029/2020MS002280

Received 11 AUG 2020
Accepted 4 DEC 2020

The LMARS Based Shallow-Water Dynamical Core on Generic Gnomonic Cubed-Sphere Geometry

Xi Chen[1,2]

[1]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, NJ, USA, [2]Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, NJ, USA

Abstract The rapidly increasing computing powers allow global atmospheric simulations with aggressively high resolutions, challenging traditional model design principles. This study presents a Low Mach number Approximate Riemann Solver (LMARS) based unstaggered finite-volume model for solving the shallow-water equations on arbitrary gnomonic cubed-sphere grids. Using a novel reference line-based grid-generation process, it unifies the representation of arbitrary gnomonic cubed-sphere grid projections and permits high-efficiency 1D reconstruction in the halo regions. The numerical discretization also extends a widely used pressure gradient algorithm with the LMARS viscous term, thus improves the model's stability for various numerical applications. The solver demonstrates a broad range of organic diffusion control without any explicit filters, validated by a comprehensive set of test cases. Lastly, a newly introduced splash on the sphere test verifies the solver's desirable dispersion properties and consistent performance among different grid types. This study paves a solid foundation for a new generation of global circulation models with kilometer horizontal scales.

Plain Language Summary Computing powers and architectures historically influence the numerical algorithm designs of global atmospheric simulations at the fundamental levels. The next generation of the global circulation models can push its resolution to kilometer horizontal scales, which requires vital capabilities in a balanced representation of all motion modes and handling sharp discontinuities such as topography. Here, we demonstrate a new framework of a dynamical core that inherits advantages in both traditional geophysical fluid dynamics (GFD) modelings and versatile general computational fluid dynamics (CFD) techniques. This new development introduces several innovations, including the unified grid description, numerical optimization, stability enhancement, and a newly designed test to illustrate specific numerical properties cleanly. The desirable results of this work increase our confidence in creating a unique global circulation model to leverage next-generation high-performance computing and to improve our fundamental understanding of the atmospheric processes.

1. Introduction

Building a dynamical core for the atmospheric model is an art of balancing the requirements between accuracy and efficiency. Although a model's accuracy can be more definitive to measure in benchmark experiments—at least for smooth solutions, computational efficiency is a relative concept that strongly associates with the computing platform characteristics. In a way, the available computational power and machine architectures have been dictating the scope of the research topics and the design of the numerical simulation. For example, coming out of the wide adoption of the massively parallel computing, many modern models prefer quasi-uniform computational grids, such as the icosahedral (Du et al., 2003; Ringler et al., 2000; Rípodas et al., 2009; Tomita et al., 2001) and cubed-sphere grids (Adcroft et al., 2004; Marras et al., 2015; Putman & Lin, 2007), over the traditional latitude-longitude grids to evenly distribute the computation tasks over a large number of processors. However, the concepts of both icosahedral and cubed-sphere grids were explored very early in Sadourny et al. (1968), Sadourny (1972), and Williamson (1968), but gained very little attention. During this time, single-processor-based latitude-longitude or Gaussian spectral methods demonstrated excellent efficiency in delivering accurate results (Bourke, 1972, 1974)—a perfect example of the evolving standard of the numerical algorithm efficiency.

The essential motivation of implementing the quasi-uniform grids is to transform the severe two parallel-unfriendly singular points from the latitude-longitude grid (the north and south poles) to less severe eight singularities in the cubed-sphere grid (corners of the six cubed-sphere tiles) or 12 singularities in the icosahedral grid (pentagon cells in a Voronoi perspective). Although the quasi-uniform grids avoid polar filters and other numerical damping schemes due to the polar singularities, they have both pros and cons. In some cases, the numerical artifacts and errors caused by the 12 icosahedral singularities are less severe than the ones caused by cubed-sphere corners. The cubed-sphere grid, however, has several desirable properties. First, many attractive high-order multi-dimensional algorithms achieved optimal computational efficiency by taking advantage of logically Cartesian grids (i.e., rectangles) on each cubed-sphere tile (C. Chen & Xiao, 2008; Putman & Lin, 2007; Rossmanith, 2006; Taylor & Fournier, 2010; Ullrich et al., 2010). Second, the logical 3D cubic geometry on each cube-tile creates one additional dimension to the computational data structure, which offers more flexibility in the parallel computational optimization designs in both horizontal and vertical directions. Lastly, each tile of the cubed-sphere geometry has a significant similarity with computational grids in regional models. In fact, with separately defined metric terms, the global cubed-sphere dynamical core requires almost no code change to convert into a regional solver in orthogonal grids. This property provides exceptional convenience and flexibility in unifying real-world applications and experimenting new algorithms with idealized validation tests.

The global circulation models (GCMs) cover a vast range of scales and operate on the most powerful high-performance computers available. Therefore, geophysical fluid models usually discretize the governing equations with various grid staggering choices (Arakawa & Lamb, 1977, namely A-, B-, C-, D-Grid) to gain numerical advantages in resolving the smallest waves. In particular, the C-Grid results in natural pressure gradient in the momentum equations and straightforward divergence representation (Adcroft et al., 2019); the D-Grid forms a perfect vorticity mode (Harris & Lin, 2013; Lin, 2004), which dominants large-scale to mesoscale atmospheric motions. With the recent rapid increase of computational power, the GCM community is pushing the model resolution aggressively to about a globally kilometer horizontal scale (Satoh et al., 2019; Stevens et al., 2019). At this scale, both divergence and vorticity play equally important roles. Therefore, the unstaggered discretization could be a balanced choice in global kilometer-scale model development.

Although most GCMs use staggered discretization, a few influential models utilize the A-Grid, including NICAM (Tomita et al., 2001) with a stencil-based scheme, and the E3SM (Taylor et al., 2020) and NEPTUNE/NUMA (Müller et al., 2019) with a compact Spectral-Element-based algorithm. Colocating the velocity components and scalars with a stencil-based scheme has several unique and attractive advantages. It allows direct coupling between the explicitly simulated dynamical process and parameterized physics processes. No interpolation of the velocity components is required, therefore, eliminating errors associated with such practices. Furthermore, it can yield energy conservation in nonhydrostatic models. With colocated prognostic variables, it is possible to formulate a flux-form prognostic total-energy governing equation, thus automatically provide energy conservation to the models. Energy conservation is beyond the scope of the shallow-water implementation and will be extensively discussed in future work. Preliminary work has demonstrated promising results in Li and Chen (2019). Last but not least, unstaggered schemes are widely implemented in traditional computational fluid dynamics (CFD) applications. Many well-tested techniques can inspire the creation of algorithms in the geophysical fluid simulations.

Despite many efforts, historically, the unstaggered discretization was overwhelmed by staggered methods due to, at least, two primary considerations. One challenge is that compared with the staggered schemes, the unstaggered discretization tends to produce severe errors in resolving phase speed for waves with very short wavelengths. Additionally, it is more challenging to design numerically robust algorithms with an unstaggered grid. For example, an unstaggered central-differencing of gradient term could produce a so-called "grid decoupling" problem and lead to checkerboard-pattern noise. Indeed, the long-term success of an unstaggered model depends on the proper treatments to accommodate the above difficulties.

To address the first concern, the previous installment of this work thoroughly investigated the dispersive and dissipative relations between different grid staggering choices and orders of accuracy (Chen et al., 2018). One of the principal findings is that with high-order algorithms, for example, using 3- or 5-points stencil polynomial schemes, waves with problematic phase speeds will be largely pushed to four-grid-spacing and below. Increasing recent linear analysis on compact schemes, such as finite element, spectral-element methods, also found similar behavior (Ainsworth, 2014; Le Roux et al., 2020; Ullrich et al., 2018). On the other hand, in practice, such smallest resolvable wavelengths are also contaminated by many sources of errors, such as numerical diffusion, strong gradients in the solutions, and, accordingly, heavily damped or removed by numerical techniques. Therefore, using high-order numerical schemes can prevent the unstaggered model from problematic phase speeds. Additionally, since it is difficult to connect and generalize the linear dispersion analysis results in real-world applications, Chen et al. (2018) also introduced simple-to-setup tests to validate the dispersion and dissipation properties of any sophisticated solvers.

Unstaggered algorithms for geophysical flows can inherit various robust approaches from other fields. Unlike the geophysical fluid modeling field, the general CFD community widely adopts unstaggered algorithms via the implementation of the approximate Riemann Solvers. There is a small but growing literature that takes advantage of the Riemann Solvers in the A-Grid GCMs and achieved stability, high order accuracy, and are free of explicit diffusion (C. Chen & Xiao, 2008; Giraldo et al., 2002; Ullrich et al., 2010; Yang et al., 2010). However, compared with state-of-science GCMs, the traditional Riemann solvers can be less efficient or more diffusive. Moreover, the Riemann solvers with carefully designed mathematical expressions are most effective in simulating generic flow types, including the vacuum, shock, sharp discontinuities. Therefore, it is not easy to understand the Riemann solvers' inherent numerical properties and make the comparison to traditional GCM implementations. Lastly, traditional Riemann solvers require careful adjustments to make them accurate on the unique geometry on the Earth's surface with the gravity and the Coriolis forces.

The Low Mach number Approximate Riemann Solver (LMARS) introduced in Chen et al. (2013) is a highly efficient tool in the finite-volume method based GCMs. The design of LMARS bases on the fact that the geophysical flows do not create any vacuum or sharp discontinuities and the subsonic flows Mach number is always less than 1. The resulting simple mathematical expression of LMARS requires only one approximation, which assumes the sound wave speed (or gravity wave speed in a shallow-water model) is continuous at the finite volume interfaces. The LMARS discretization in a full 3D atmosphere takes the gravity into account and results in better accuracy than traditional Riemann solvers (Li & Chen, 2019). Previous work has implemented LMARS in many atmospheric applications, including hydrostatic and nonhydrostatic flows with either vertical Lagrangian coordinates or Eulerian coordinates (Chen et al., 2013). It is also implemented in the multi-gas planet atmosphere environment with an intrinsic energy-conserving framework (Li & Chen, 2019). As a promising candidate for a new unstaggered GCM, LMARS is not yet tested with a cubed-sphere geometry.

This study aims to develop an LMARS-based unstaggered finite-volume shallow-water solver on the gnomonic cubed-sphere grids. The shallow-water equations offer a standard testbed to validate the horizontal advection algorithms. Whereas previous studies implemented various cubed-sphere grids in different solvers, this work is the first to unify arbitrary gnomonic cubed-sphere grid generation processes. Although the numerical discretization of LMARS on the cubed-sphere grid does not contain any explicit filter, this solver can still exhibit a broad range of diffusion properties by controlling the polynomial reconstructing methods and the strength of the forward-backward techniques (Mesinger, 1977). Lastly, various model numerical properties are illustrated and gauged by a comprehensive set of idealized tests, including the traditional famous Williamson et al. (1992) shallow water test suite, and a recently introduced colliding modons test (Lin et al., 2017). Given the traditional discussions about unstaggered grid dispersion uncertainties, this study designs a new “splash on the sphere” test to illustrate and warrant satisfactory dispersion and dissipation properties on specific wavelengths.

The remainder of the study is organized as follows. Section 2 provides a brief primer of the gnomonic cubed-sphere grids, in which a unified grid generation process is introduced. Section 3 describes both spatial and temporal numeric discretization. The model is validated in Section 4. In this section, a novel splash on the sphere test is introduced to gauge the dispersion and dissipation properties. Finally, the main findings and discussions are concluded in Section 5. Appendix A collects all major symbols to make the notations consistent and clear. B provides a mathematical description of the grid generating process. The governing equations in the cubed-sphere grids are provided in C. In this section, some optimizations are discussed to yield more efficient mathematical expressions.

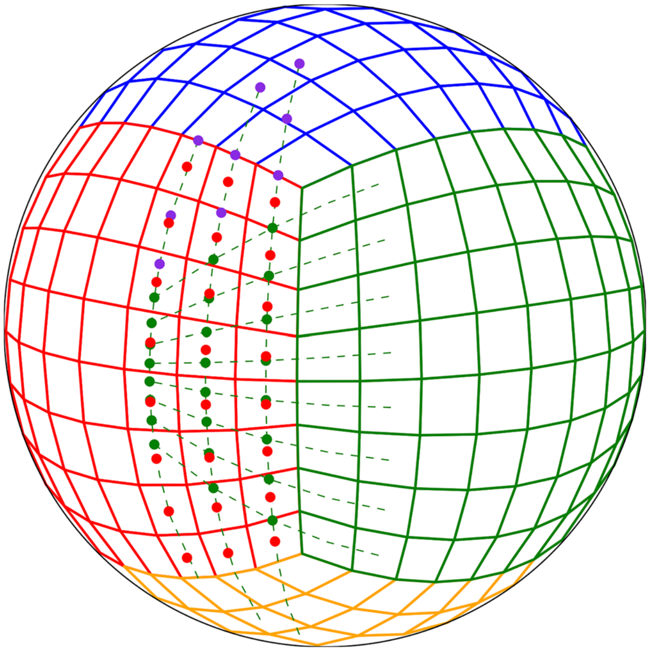

Figure 1. C8 (8 × 8 × 6) cubed-sphere grid with three layers of ghost cells. When aligned on the same great circle, a simple 1-D polynomial interpolation from the red dots can provide the green dots' values with optimum accuracy and speed.

2. A Brief Primer on Gnomonic Cubed-Sphere Grids and the Duo-Grid System

The FV3 (Finite-Volume Cubed-Sphere Dynamical Core) (Harris & Lin, 2013; Lin, 2004; Putman & Lin, 2007) has strongly influenced the development of this study. Although being different in staggering choices, the new unstaggered dynamical core is designed to be a seamless evolution within the existing FV3 framework. Therefore, this work follows a significant amount of naming conventions from the existing FV3 code base, including grid type names and dynamical parameter definitions and symbol names.

The cubed-sphere grid is obtained from projecting a gridded cube onto the surface of the sphere, which avoids the polar singularities due to the convergence of the meridians from the traditional latitude-longitude coordinate system. The cubed-sphere grid shifts the severe polar convergence of meridians to eight weaker singularities at the corners where cube tiles intersect. Figure 1 illustrates the mapping from the cube to the sphere in a C8 resolution. The naming convention C [N] denotes that each tile of the cube-sphere has N by N cell distribution, which results in $N \times N \times 6$ total cells on the sphere.

2.1. Interlock Patterns

There are two conventional logical arrangements to interlock the six cubed-sphere tiles. Figure 2a illustrates the first pattern (Guba et al., 2014; Nair et al., 2005; Ronchi et al., 1996; Rossmanith, 2006; Yang et al., 2010) with four tropical tiles and two polar tiles, denoted by the "tropic-belt" logical arrangement. Unfortunately, the interlock between the tiles is not symmetric and require separate distinction for each of the tile. The second pattern (Adcroft et al., 2004) is illustrated in Figure 2b, and denoted by the "staircase" logical arrangement. The "staircase" logical arrangement has better symmetry when exchanging information between tiles. For example, the calculation of fluxes between two adjacent tiles needs to determine the tile numbers and the tile-local coordinates rotations. The "staircase" arrangement simplifies the neighbor tile-number and rotation patterns into two odd-index-tile (tiles 1, 3, 5) and even-index-tile (tiles 2, 4, 6) scenarios, illustrated by Figures 2c and 2d. Both logical arrangements are valid for cubed-sphere applications. This work implements the "staircase" arrangement to gain some programming simplicities.

2.2. Unified Gnomonic Projections

The grid lines connecting the grid points are continuous great circles on the sphere. Therefore, the locations of vertices can uniquely determine the full grid system on each cube-sphere tile. On each tile, for example, the tile centered at $\left(\lambda, \phi\right) = (0,0)$, the vertices can be generated by either the gnomonic (Ronchi et al., 1996; Sadourny, 1972) or the conformal (McGregor, 1996; Rančić et al., 1996) projections. The gnomonic projection projects a Cartesian grid from six straight-lines-meshed cube surfaces to the sphere surface. The conformal projection maximizes the orthogonality of the coordinates. Putman and Lin (2007) examined the most popular approaches: the equidistant projection (Sadourny, 1972), the equiangular projection (Ronchi et al., 1996), the more orthogonal conformal mapping (Rančić et al., 1996), the numerical modification to analytical mappings by an elliptic solver (Khamayseh & Mastin, 1996) or the spring dynamics generator (Tomita et al., 2001). Considering the eight singularities at cube corners remain nonorthogonal in the conformal grid, and the cell size distribution in conformal grids is usually less uniform than the gnomonic choices, this work implements the gnomonic grid.

Figure 3 illustrates a typical gnomonic projection between a grid point on the cube-tile and the sphere-tile. Let the grid points on a cube-tile be indexed by $[i, j]$, with local coordinate $[X, Y]$. In a gnomonic projection,

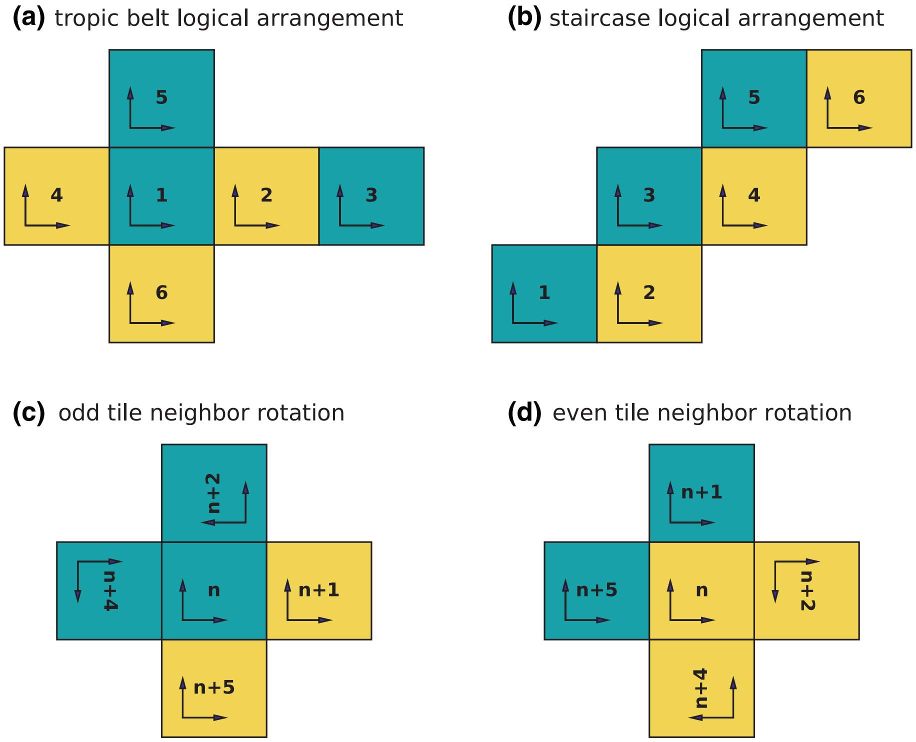
(a) tropic belt logical arrangement
5
4
1
2
3
6
(b) staircase logical arrangement
5
6
3
4
1
2
(c) odd tile neighbor rotation
n+2
n+4
n
n+1
n+5
(d) even tile neighbor rotation
n+1
n+5
n
n+2
n+4

Figure 2. The logic arrangement of the six cubed-sphere tiles. (a) the “tropic belt” arrangement has four tropic tiles and two polar tiles, the local coordinate directions are as illustrated at the lower-left corners. (b) the “staircase” arrangement and the local coordinates illustration. The “staircase” arrangement simplifies the determination of the neighbor tile indices and orientations into (c) odd-index-tile (blue) and (d) even-index-tile (yellow) cases. Note if the number $n + l > 6$, the corresponding tile number is $n + l - 6$.

each row of grid points [*i*,:] shares the same *Y* value, and each column of grid points [:, *j*] shares the same *X* value. The projection of the rows and columns on each cube face forms great circles on the sphere. Furthermore, the coordinates on the sphere converge to a pair of local north-/south-poles, and “west-/east-poles” for each sphere tile (Rossmanith, 2006; Figure 1).

This study offers a novel approach to unite generic gnomonic grid descriptions. Although previous work demonstrated different ways to create various gnomonic grids, the gnomonic grids share distinct properties. Given the orthogonality and symmetry of the gnomonic projected grid points on each cube-tile, the locations of grid points in a single row or column can fully determine the entire gnomonic cubed sphere. Therefore, Figure 3 can determine three widely used gnomonic projections in the literature by three pairs of reference lines with grid points equally distributed. The red reference lines stand for the most traditional equidistance projection (Sadourny, 1972). The equiangular grid (Ronchi et al., 1996) can be obtained by projecting the green lines with grid points equally distributed back to the cube tile and populating them to the entire [*X*, *Y*] space. Lastly, although not formally documented, FV3 (Harris et al., 2016; Harris & Lin, 2013; Putman & Lin, 2007) introduced a gnomonic projection with more uniformly distributed cells on the tile interfaces, and is now adopted in the Next Generation Global Prediction System (NGGPS) project (Zhou et al., 2019). The FV3 grid is denoted by the “equi-edge” grid and can be obtained similarly by populating the grid-points-equally distanced blue lines on the sphere. The following procedures describe the process to populate the reference line grid points to the full gnomonic projections:

1. Project the equally distributed grid points from the reference lines onto the gray-color cube-tile (this procedure is redundant for the red line).
2. Mesh the gray cube-tile with the projected 1D distribution of the grid points from the cube surface.
3. Project the fully meshed grid points from the gray cube tile to the green sphere tile.

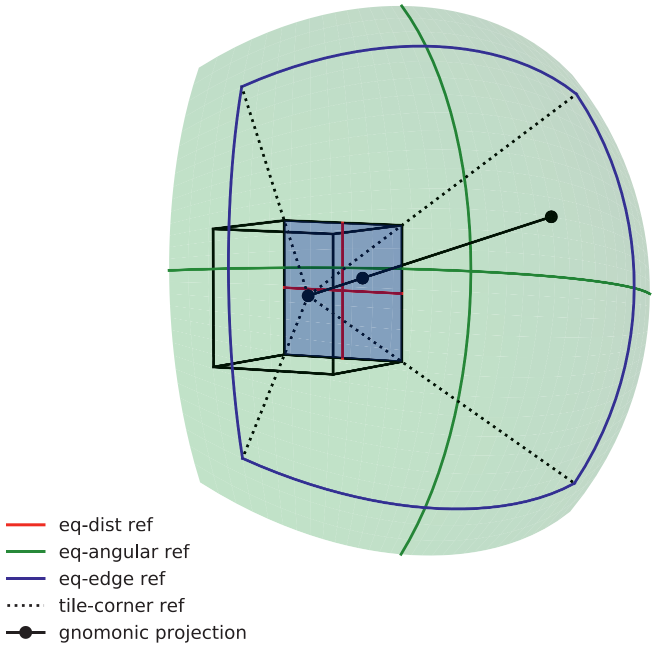

Figure 3. The gnomonic projection. The blue shaded surface is a ***cube-tile***, and the green shaded surface is a ***sphere-tile***. The solid black line from the sphere center to the sphere surface illustrates a ***gnomonic projection*** between a grid point on the cube-tile and a grid point on the sphere-tile. The dotted lines define the corners of the bounds to one of the six surfaces of the cubed-sphere grid. The red lines are the ***gnomonic projection reference lines*** of the equidistance projection, and the green and blue lines are of the equiangular and ***equi-edge*** projections. The extension of the green surface indicates the ghost cell region for a finite-volume application.

The equidistance grid produces less uniformly distributed volumes on the sphere. Therefore, most modern cubed-sphere models do not implement the equidistance grid. Although the equiangular projection yield slightly more uniform grid point distribution on the sphere, the equi-edge grid distributes grid points evenly near the sphere tile connections, which are the primary sources that cause grid imprinting with a cubed sphere. Additionally, in a stretched grid (Harris et al., 2016), the equi-edge grid produces the focused tile with more evenly distributed cells. This work only discusses the equi-edge and the equiangular grids, namely grid_type 0 and 2, following FV3 naming convention.

2.3. The Duo-Grid System to Handle Tile-Edge Connections

This work implements a maximum 5-point-stencil polynomial reconstruction for the finite volume scheme. Therefore each tile requires three more layers of ghost cells by the wave propagation method (Chen et al., 2013, 2018; Li & Chen, 2019). Unfortunately, a main numerical challenge with the gnomonic cubed-sphere grid is that the grid coordinates are not continuous across tile interfaces. Figure 1 illustrates the ghost cells on the local west-side of the green tile, which creates a west halo region. The native neighbor cell centers in the halo region are the red dots, which forms “kinks” in coordinates between the green tile and its west halo. Therefore, the name “kinked grid” denotes this halo type constructed by directly copying values from the neighboring tiles. A natural extension of the green tile coordinates (green dashed lines) into the halo region results in green dots, denoted by the “extended grid.”

There are two reasonable choices for the ghost cell numerical algorithms. FV3 powered models, for example, directly employ the red dots in the kinked-grid to form one-sided flux calculations to the tile boundaries. This method has several advantages. The operations in the halo region (e.g., west halo of the green tile) can be identical to the native calculations on the neighboring tile (red tile) from different calculating processors (Harris & Lin, 2013; Putman & Lin, 2007). Additionally, calculations associated with halo create almost no computational overhead, a valuable property to massively parallel computing. This study implements the other approach, which is to remap the red dots from the kinked-grid to the green dots on the extended grid (Katta et al., 2015; Rossmanith, 2006; Ullrich et al., 2010; Yang et al., 2010). Therefore, the halo operations are a natural extension from the interior calculations, and no extra adjustment is required for the ghost cells. The second approach can effectively reduce the cubed-sphere grid imprinting at the tile edges, but creates undesired overhead in the parallel calculations. The extra calculations are mainly: (1) Tile-interface flux synchronization between different tiles; (2) Remapping of the prognostic variable values from the kinked grid to the extended grid.

The first synchronization overhead is due to inconsistent coordinate directions between tiles. The fluxes at the same interfaces by different tiles need to be sent to the adjacent tile and get averaged for conservation properties. This operation is only performed at the end of a full cycle of integration to minimize the message passing.

The interpolating algorithm must be at minimum complexity to alleviate the duo-grid remapping overhead. Most equiangular projection applications take advantage of that the green dots and red dots are aligned on the same great circle. Thus, 1-D interpolation is sufficient for the remapping procedure. This property is denoted by duo-grid 1D alignment. Unfortunately, this 1D alignment only applies to equiangular projection. A direct halo generation with equidistance or equi-edge projections (i.e., by extending three more layers into the halo during the interior grid creation) does not produce the duo-grid 1D alignment. A few extra steps by the following procedures can solve this problem and warrant the duo-grid 1D alignment for arbitrary gnomonic projections:

1. Populate the nonhalo gnomonic projection grid points based on the reference lines (red for equidistance, blue for equi-edge, and green for equiangular).
2. Identify the grid points on the resulting sphere tile center lines (i.e., the green lines in Figure 3).
3. Mirror the outermost three points from the green lines to the ghost region.
4. Populate ghost cell grid points in the halo region base on the mirrored points on the green lines.

To minimize the remapping overhead, in this study, the minimum 1D piecewise linear interpolation is used to remap values from the red to the green points. No discernible degradation is observed compared with high-order 1D polynomial remapping algorithms. Detailed mathematic procedures in the gnomonic cubed-sphere generation are described in (B).

3. The Numerical Discretization

3.1. The Governing Equations Discretization

A standard mathematical derivation of the governing equations on the cubed-sphere is included in C. Before rearranging and discretizing the governing equations, denote η an arbitrary variable, the following finite-volume (FV) operators are defined:

$$\delta_x\left[\eta\right]_{i,j} = \eta_{i+0.5,j} - \eta_{i-0.5,j}, \tag{1}$$

$$\delta_y\left[\eta\right]_{i,j} = \eta_{i,j+0.5} - \eta_{i,j-0.5}, \tag{2}$$

where $\eta_{i\pm0.5,j}$ are the values of η evaluated at cell interfaces in the x-direction, and $\eta_{i,j\pm0.5}$ are the values of η evaluated at cell interfaces in the y-direction. Unless specifically specified, the subscripts (i,j) are omitted for cleaner mathematical expressions, and the variables without subscripts are evaluated at the cell centers at location (i,j).

The continuity equation can be rearranged in the flux form:

$$\frac{\partial h}{\partial t} = -\frac{1}{\Delta A}\left(\delta_x\left(hu_\perp^x\Delta y\right) + \delta_y\left(hu_\perp^y\Delta x\right)\right), \tag{3}$$

Define the flux coefficient:

$$f_{i+0.5,j}^x\left(u_\perp^x, \Delta t\right) = u_\perp^x\Delta t\Delta y_{i+0.5,j}, \tag{4}$$

$$f_{i,j+0.5}^y\left(u_\perp^y, \Delta t\right) = u_\perp^y\Delta t\Delta x_{i,j+0.5}. \tag{5}$$

Therefore, the final discretization of the continuity equation is:

$$h^{n+1} = h^n - \frac{1}{\Delta A}\left(\delta_x\left(f^x h\right) + \delta_y\left(f^y h\right)\right), \tag{6}$$

and a straightforward discretization to update the cell center velocity components is:

$$u^{n+1} = u^n + \left(\sqrt{g}\hat{v}\left(\zeta + f\right) - \frac{1}{\Delta x}\delta_x\left(\Pi + \Phi_s + K\right)\right)\Delta t, \tag{7}$$

$$v^{n+1} = v^n + \left(-\sqrt{g}\hat{u}\left(\zeta + f\right) - \frac{1}{\Delta y}\delta_y\left(\Pi + \Phi_s + K\right)\right)\Delta t, \tag{8}$$

The relative vorticity ζ at cell centers can be numerically evaluated by:

$$\zeta = \frac{1}{\Delta A}\left(\delta_x\left(u_{\parallel}^{x}\Delta y\right) - \delta_y\left(u_{\parallel}^{y}\Delta x\right)\right), \tag{9}$$

and the kinetic energy K can be numerically evaluated at cell interfaces by:

$$K = \frac{1}{2}\left(u_{\perp}u_{\perp} + u_{\parallel}u_{\parallel}\right). \tag{10}$$

Note that K or $u_{\perp}$, $u_{\parallel}$ are all defined at the cell interfaces, and the labels in x and y directions are omitted.

All cell-center or cell-averaged variable values are straightforward to calculate. The next step is to determine the values at the cell interfaces using an efficient approximate Riemann solver and evaluate the terms in the δ_x and δ_y operators.

3.2. Calculation of Cell-Interface Values by LMARS

The remaining variables to be determined at the cell interfaces are: $u_{\perp}$, h, Π, Φ_s, and K (i.e., $u_{\perp}$, $u_{\parallel}$).

Following the naming conventions in Chen et al. (2018), $\eta_{i,j}^{W}$ and $\eta_{i.j}^{E}$ denote the "west" and "east" volume boundary values of arbitrary variable η calculated by polynomial reconstruction within volume $\left(i, j\right)$. Therefore, the mismatching $\eta_{i,j}^{E}$ and $\eta_{i+1,j}^{W}$ are evaluated at two sides of the interface between cell $\left(i, j\right)$ and $\left(i+1, j\right)$. The first step of LMARS is to evaluate cell interface values of Π and $u_{\perp}$. In the x-direction:

$$\left(u_{\perp}^{x}\right)_{i+0.5,j} = \frac{1}{2}\left(\left(u_{\perp}^{x}\right)_{i,j}^{E} + \left(u_{\perp}^{x}\right)_{i+1,j}^{W}\right) + \frac{1}{2a_{i+0.5,j}}\left(\Pi_{i,j}^{E} - \Pi_{i+1,j}^{W}\right), \tag{11}$$

$$\Pi_{i+0.5,j} = \frac{1}{2}\left(\Pi_{i,j}^{E} + \Pi_{i+1,j}^{W}\right) + \frac{a_{i+0.5,j}}{2}\left(\left(u_{\perp}^{x}\right)_{i,j}^{E} - \left(u_{\perp}^{x}\right)_{i+1,j}^{W}\right), \tag{12}$$

where a is the gravity wave speed (group velocity) estimated at cell interface:

$$a_{i+0.5,j} = \sqrt{\frac{1}{2}\left(\Pi_{i,j}^{E} + \Pi_{i+1,j}^{W}\right)}. \tag{13}$$

Once the velocity normal to the cell interface determined, the values of the variables to be transported are chosen using upwind values:

$$h_{i+0.5,j} = \begin{cases} h_{i,j}^{E} & \text{if} \left(u_{\perp}^{x}\right)_{i+0.5,j} > 0 \\ h_{i+1,j}^{W} & \text{else} \end{cases}, \tag{14}$$

$$\left(u_{\parallel}^{x}\right)_{i+0.5,j} = \begin{cases} \left(u_{\parallel}^{x}\right)_{i,j}^{E} & \text{if} \left(u_{\perp}^{x}\right)_{i+0.5,j} > 0 \\ \left(u_{\parallel}^{x}\right)_{i+1,j}^{W} & \text{else} \end{cases}, \tag{15}$$

and the operations on the y-direction is symmetric and analogous.

Although algebraically h and Π are interchangeable with the relation $\Pi = Gh$, they play different roles in the governing equation. Π stands for pressure forcing term, and h is associated with material transportation. Therefore, their treatments are different in the LMARS solver. In a fully compressible model, their corresponding terms are the density and pressure (Chen et al., 2013; Li & Chen, 2019).

Table 1
The Coefficient to Calculate "East" Side Midpoint Value on a Stencil Number of $N_c = 2N_g - 1$ From Point-Value Variables

N_g	$E_{pt}^{(-2)}$	$E_{pt}^{(-1)}$	$E_{pt}^{(0)}$	$E_{pt}^{(1)}$	$E_{pt}^{(2)}$
1	-	-	1	-	-
2	-	−1/8	3/4	3/8	-
3	3/128	−5/32	45/64	15/32	−5/128

The last piece for a single sub-cycle update is to reconstruct the prognostic variables' values at four horizontal cell edges. Following Chen et al. (2018) conventions, assuming an N_c-point stencil polynomial reconstruction for an arbitrary variable η:

$$\eta_{i,j}^{W} = \sum_{N_g-1}^{l=1-N_g} W^{(l)}\eta_{i+l,j}, \tag{16}$$

$$\eta_{i,j}^{E} = \sum_{N_g-1}^{l=1-N_g} E^{(l)}\eta_{i+l,j}, \tag{17}$$

$$\eta_{i,j}^{S} = \sum_{N_g-1}^{l=1-N_g} W^{(l)}\eta_{i,j+l}, \tag{18}$$

$$\eta_{i,j}^{N} = \sum_{N_g-1}^{l=1-N_g} E^{(l)}\eta_{i,j+l}, \tag{19}$$

where $W^{(l)} = E^{(-l)}$, $N_c = 2N_g - 1$, and N_g is the layers of ghost cells required at the tile edges. The lookup coefficient tables for point-value reconstruction and finite-volume reconstruction from Chen et al. (2018) is repeated in Tables 1 and 2 for reference. Although this work does not implement any monotonic filters, most filtering schemes can be considered equivalent to altering the reconstruction coefficient values locally according to the shape of the solutions.

3.3. Extension of the Pressure Gradient Algorithm in L97 with LMARS Viscous Terms and the Forward-Backward Algorithm

Up to this part, the numerical discretization closely follows the original LMARS approach described by Chen et al. (2013; 2018), which is sufficient for this shallow water solver. Alternatively, Lin (1997) (L97) introduces an innovative spatial discretization to the pressure gradient term. Although mathematically equivalent to this work, the L97 pressure gradient algorithm considerably simplifies the discretization in a full 3D compressible model and makes the numerical expression backward compatible with a shallow-water solver. One drawback of L97 is that the expression is a geometric realization of the volume interface pressure force integration. Therefore, it lacks the implicit diffusion calculated by a Riemann solver, and extra filters are required to stabilize the model. Fortunately, LMARS simple expression allows the separation of the geometric derivative terms and viscous terms. It is straightforward to extend the L97 pressure gradient algorithm with the LMARS viscous contribution and result in a fast and stable pressure gradient scheme.

As discussed previously, the volume interface pressure force obtained by LMARS described in Equation 12 can be represented by an averaging term and a viscous term:

$$\Pi_{i+0.5,j} = \overline{\Pi}_{i+0.5,j} + \Pi_{i+0.5,j}^{\text{vis}}, \tag{20}$$

where

$$\overline{\Pi}_{i+0.5,j} = \frac{1}{2}\left(\Pi_{i,j}^{E} + \Pi_{i+1,j}^{W}\right), \tag{21}$$

$$\Pi_{i+0.5,j}^{\text{vis}} = \frac{a_{i+0.5,j}}{2}\left(\left(u_{\perp}^{x}\right)_{i,j}^{E} - \left(u_{\perp}^{x}\right)_{i+1,j}^{W}\right). \tag{22}$$

Therefore, the pressure gradient with the topographic contribution, for example, in the x-direction is discretized:

$$\frac{\partial}{\partial x}\left(\Pi + \Phi_s\right) = \frac{1}{\Delta x}\delta_x\left(\overline{\Pi} + \overline{\Phi}_s\right) + \frac{1}{\Delta x}\delta_x\Pi^{\text{vis}}. \tag{23}$$

Table 2
The Coefficient to Calculate "East" Side Midpoint Value on a Stencil Number of $N_c = 2N_g - 1$ From Volume-Mean Variables

N_g	$E_{vm}^{(-2)}$	$E_{vm}^{(-1)}$	$E_{vm}^{(0)}$	$E_{vm}^{(1)}$	$E_{vm}^{(2)}$
1	-	-	1	-	-
2	-	−1/6	5/6	1/3	-
3	1/30	−13/60	47/60	9/20	−1/20

Note that X. Chen et al. (2018) explained that the viscous term Π^{vis} in 1D is equivalent to a $2N_g$-order diffusion term, which automatically matches the stencil size of the reconstruction schemes. The same function evaluates the surface geopotential at cell interface to warrant numerical consistency:

$$\left(\overline{\Phi}_s\right)_{i+0.5,j} = \frac{1}{2}\left(\left(\overline{\Phi}_s\right)_{i,j}^{E} + \left(\overline{\Phi}_s\right)_{i+1,j}^{W}\right). \tag{24}$$

The following equivalent expressions are defined to simplify the expression:

$$\eta_{(\pm,\pm)}^{x} \equiv \overline{\eta}_{i\pm0.5,j,k\pm0.5}, \tag{25}$$

and the first term in pressure gradient can be rearranged in L97 format:

$$\delta_x\left(\overline{\Pi} + \overline{\Phi}_s\right) = -\frac{\left(\Phi_{(-,+)}^{x} - \Phi_{(+,-)}^{x}\right)\left(p_{(+,+)}^{x} - p_{(-,-)}^{x}\right) + \left(\Phi_{(-,-)}^{x} - \Phi_{(+,+)}^{x}\right)\left(p_{(-,+)}^{x} - p_{(+,-)}^{x}\right)}{p_{(-,+)}^{x} + p_{(+,+)}^{x} - p_{(-,-)}^{x} - p_{(+,-)}^{x}}. \tag{26}$$

The expression in the y-direction is analogous. With this expression, the SWE and 3D compressible pressure gradient numerical discretization can share the identical code.

Another expansion of the pressure gradient calculation is to implement the forward-backward (FB) technique (Mesinger, 1977) to enhance the stability of the model. Chen et al. (2018) conducted several numerical analysis on the FB technique. One of the key findings in the 1D linearized environment is that the FB operation is equivalent to a second-order diffusion term on the velocity components, which is vital for the stability of a single-step scheme. In this model with multi-step schemes, the FB scheme is not a necessary component but can be useful in the future fully coupled 3D model as a diffusion moderator. Considering the FB is almost computationally equivalent to an explicit scheme, a relaxed-FB scheme is implemented by setting a parameter $\beta \in \left[0,1\right]$ to the geometric pressure gradient term following Harris and Lin (2013):

$$\delta_x\left(\overline{\Pi} + \overline{\Phi}_s\right) = \left(1-\beta\right)\delta_x\left(\overline{\Pi} + \overline{\Phi}_s\right)^{n+1} + \beta\delta_x\left(\overline{\Pi} + \overline{\Phi}_s\right)^{n}. \tag{27}$$

Therefore, the parameter β in the FB scheme is an implicit control coefficient of a 2nd-order diffusion term. The response to various beta values will be validated in the test section.

3.4. Temporal Discretization

The linearized two-step and three-step Runge-Kutta schemes (rk2 and rk3) are implemented for time integration. In each substep, denote prognostic variable array U and right-hand-side term RHS(U, N_g, β), where the dynamical parameters N_g controls the stencil size of the reconstruction scheme and β controls the strength of the FB algorithm in the pressure gradient evaluation. Therefore, the numerical properties of each subcycle update can be tailored to various purposes. Unless explicitly declared, the dynamical parameters in rk2 scheme is:

$$U^{*} = U^{n} + \frac{1}{2}\Delta t\mathrm{RHS}\left(U^{n}, N_g = 1, \beta = 0\right), \tag{28}$$

$$U^{n+1} = U^{n} + \Delta t\mathrm{RHS}\left(U^{*}, N_g = 3, \beta = 0\right). \tag{29}$$

The default rk2 configuration represents a dynamical core with the least operation counts in each time step and most implicit diffusion.

The default rk3 scheme is:

$$U^{(1)} = U^{n} + \frac{1}{3}\Delta t\mathrm{RHS}\left(U^{n}, N_g = 3, \beta = 1\right), \tag{30}$$

$$U^{(2)} = U^n + \frac{1}{2}\Delta t\mathrm{RHS}\left(U^{(1)}, N_g = 3, \beta = 1\right), \tag{31}$$

$$U^{n+1} = U^n + \Delta t\mathrm{RHS}\left(U^{(2)}, N_g = 3, \beta = 1\right). \tag{32}$$

The default rk3 configuration represents dynamical core with minimized implicit diffusion.

Although the rk2 scheme cost fewer operations in each full cycle, the rk3 scheme allows larger time steps. In practice, a fully optimized rk2 should still maintain better computational efficiency over rk3, majorly due to the first-order substep is significantly cheaper than the standard high-order substep. The linearized Runge-Kutta schemes can highly optimize the computational memory storage cost since the values of RHS terms in each substep are not retained, and the storage of updated prognostic variables can be reused. Lastly, linearized rk3 only improves the CFL conditions and is 2nd-order accurate in time. Standard high-order accurate Runge-Kutta schemes can be found in X. Chen et al. (2013); Ullrich et al. (2010); C. Chen and Xiao (2008).

4. Tests and Results

4.1. Testing Plan

Table 3 introduces the naming convention to label the model configurations for a shallow water test simulation. For example, a run label C48. g2. rk3. b1.00 stands for a simulation of a 48 × 48 × 6 cubed-sphere grid with equiangular grid point distribution, and the discretization configuration is the rk3 scheme described in the previous subsection. The value of the forward-backward parameter is 1, which is fully explicit. As described above, the default rk2 configuration not only uses fully implicit forward-backward pressure gradient evaluation but also minimizes the first substep complexity by using a 1st-order accurate reconstruction scheme. The rk3 configuration implements a fully explicit pressure gradient algorithm and uses 5-point-stencil reconstruction for all three substeps. Therefore, the rk2 and rk3 configurations represent the two extremes of the numerical diffusion and performance properties spectrums, with rk2 at the more computationally efficient, more diffusive end, and vice versa. The diffusivities of other configurations lie between the two ends.

Most conventional tests require the simulation data on a traditional latitude-longitude grid for analysis. Therefore, all cubed-sphere simulation results are remapped by a 1st-order conservative algorithm onto a 144 × 72 spherical grid, which is about 2.5-degree grid spacing to minimize regridding introduced artifacts.

The validation of this model contains three stages. The first stage employs the widely adopted Williamson et al. (1992) shallow-water test suite (W92) to match the convergence of the results to the literature. The smooth solutions in W92 also allow the validations of the order of accuracy and model's overall response to diffusion. The sensitivities of the model-specific parameters, such as grid choices, discretization configurations, beta values, are tested. In GCM development, the rotational mode plays a dominant role in large to mesoscale motions. Therefore in stage 2, the recently proposed modon test (Lin et al., 2017) is implemented to check the model's quality in representing the vorticity. Last but not least, a new "splash test" inspired by Chen et al. (2018) is proposed in stage 3 to validate the model's dispersion and dissipation properties. Minimizing the phase speed errors at short wavelengths and maintaining scale-selective diffusion properties are vital for unstaggered algorithms.

The time step in each run configuration is set at the maximum allowed values that the simulations are stable in all resolutions. The courant number for LMARS based solver is estimated by:

$$c_{\mathrm{CFL}} = \frac{\left(a_{\max} + a_{\min}\right)\Delta t}{\left(\Delta A / \Delta x\right)_{\min}}, \tag{33}$$

where $a_{\max}$ and $a_{\min}$ stand for the highest and lowest gravity wave speed evaluated at the initial condition, and $\left(\Delta A / \Delta x\right)_{\min}$ is the shortest distance across a grid cell. Granting that the maximum flow speed

Table 3
Naming Convention to Label the Model Configurations of a Shallow Water Test Simulation

Notation	Description
$C\left[N\right]$	Cubed-sphere grid with N cells along each side of the cubed sphere tile. $N \times N \times 6$ total grid cells. $N \in \left[48, 96, 192, 384, 768\right]$.
g0	Equi-edge grid
g2	Equiangular grid
rk2	2-substep time scheme with 1-point- and 5-point-stencil polynomial reconstruction in two substeps.
rk3	3-substep time scheme with 5-point-stencil polynomial reconstruction in all substeps.
$b\left[\beta\right]$	The forward-backward pressure gradient algorithm control parameter $\beta \in [0, 1]$. (only appear when β value is nondefault)

Note. For example, a run label C48.g2.rk3.b1.00 stands for a simulation with a 48 × 48 × 6 cubed-sphere grid with equal-angular tile grid point distribution, and the time advancing is a 3-step Runge-Kutta scheme, the value of the forward-backward parameter is 1, which is fully explicit.

in LMARS should be less than the lowest gravity wave speed, the estimated courant number is an upper bound in the simulations. This definition can be extended to the compressible 3D model by replacing the gravity wave speed with acoustic wave speed. In all simulations, the dynamic time step Δt is calculated via: $\Delta t = \mathrm{dt\_atmos} / \left(\mathrm{k\_split} \times \mathrm{n\_split}\right)$, where dt_atmos = 3,600 s is a base time step, the cycling parameter n_split is unchanged in all resolutions, and k_split scales with resolutions with values [1, 2, 4, 8, 16] corresponding to [C48, C96, C192, C384, C768]. Among all parameters, the choice of the time-marching scheme is the most dominant factor in determining the maximum allowed time steps. The equiangular grid allows slightly larger time steps than the equi-edge grid. Table 4 lists the time step parameter n_spit and the corresponding courant numbers for all test cases.

Error norm measures follow the W92 in the height field h:

$$l_1\left(h\right) = \frac{I\left(\left|h - h_T\right|\right)}{I\left(\left|h_T\right|\right)}, \tag{34}$$

$$l_2\left(h\right) = \sqrt{\frac{I\left(\left(h - h_T\right)^2\right)}{I\left(h_T^2\right)}}, \tag{35}$$

$$l_\infty\left(h\right) = \frac{\max\left|h - h_T\right|}{\max\left|h_T\right|}, \tag{36}$$

with global mean operator:

$$I\left(\eta\right) = \frac{\sum \eta \Delta A}{\sum \Delta A}, \tag{37}$$

where ΔA is the discretized area of the cell that variable η occupies, h_T is the true solution of fluid depth, which is estimated by the most accurate C768. g2. rk3. b1.00 results if no analytical solution is available.

4.2. W92 Shallow-Water Tests

In W92, test cases 1, 2, 5, 6 (see test names in Table 4) are widely used in global shallow-water solver development. These four tests are conducted with various configurations and resolutions, as described in the testing plan. In particular, case 1 and 2 usually tests different rotating angles. Since the solutions do not show significant dependency on the rotation directions, only 45-degree rotation results are presented for

Table 4
Time Step Parameter n_split and the Corresponding Courant Number in Each Test Case

Test label	Test description	n_split (rk2)	cfl (g0.rk2)	n_split (rk3)	cfl (g0.rk3)
Case 1	Advection of a cosine bell	8	-	8	-
Case 2	Steady-state geostrophic balanced flow	10	0.7	7	1
Case 5	Zonal flow over an isolated conical mountain	13	0.85	9	1.22
Case 6	Rossby-Haurwitz wave	19	0.81	13	1.18
Modon	Colliding modons	11	1.02	7	1.61
Splash	Splash on the sphere	2	-	2	-

Note. The courant numbers are associated with equi-edge grid configuration.

brevity. The 45-degree advection of the cosine bell and the solid body rotation tests are labels as case1a45 and case2a45.

Figure 4 is the l_1, l_2, and l_{inf} error norms plotted against the simulation time at C48 resolution. A general finding is that the errors of most nonstationary runs depend mostly on rk2 and rk3 differences rather than the grid type choices, indicating diffusion properties are the dominant factor in error growth. Case 1 uses constant flows, which do not evaluate the pressure gradient. Therefore, the results represent the errors solely due to advection schemes. Additionally, no sudden error spike is observed when the cosine bell travels across cube-sphere tile boundaries in case 1. Thus, the singularities in a cubed-sphere geometry are sufficiently handled, and no grid imprinting is observed. In the solid-body rotation tests with stationary solu-

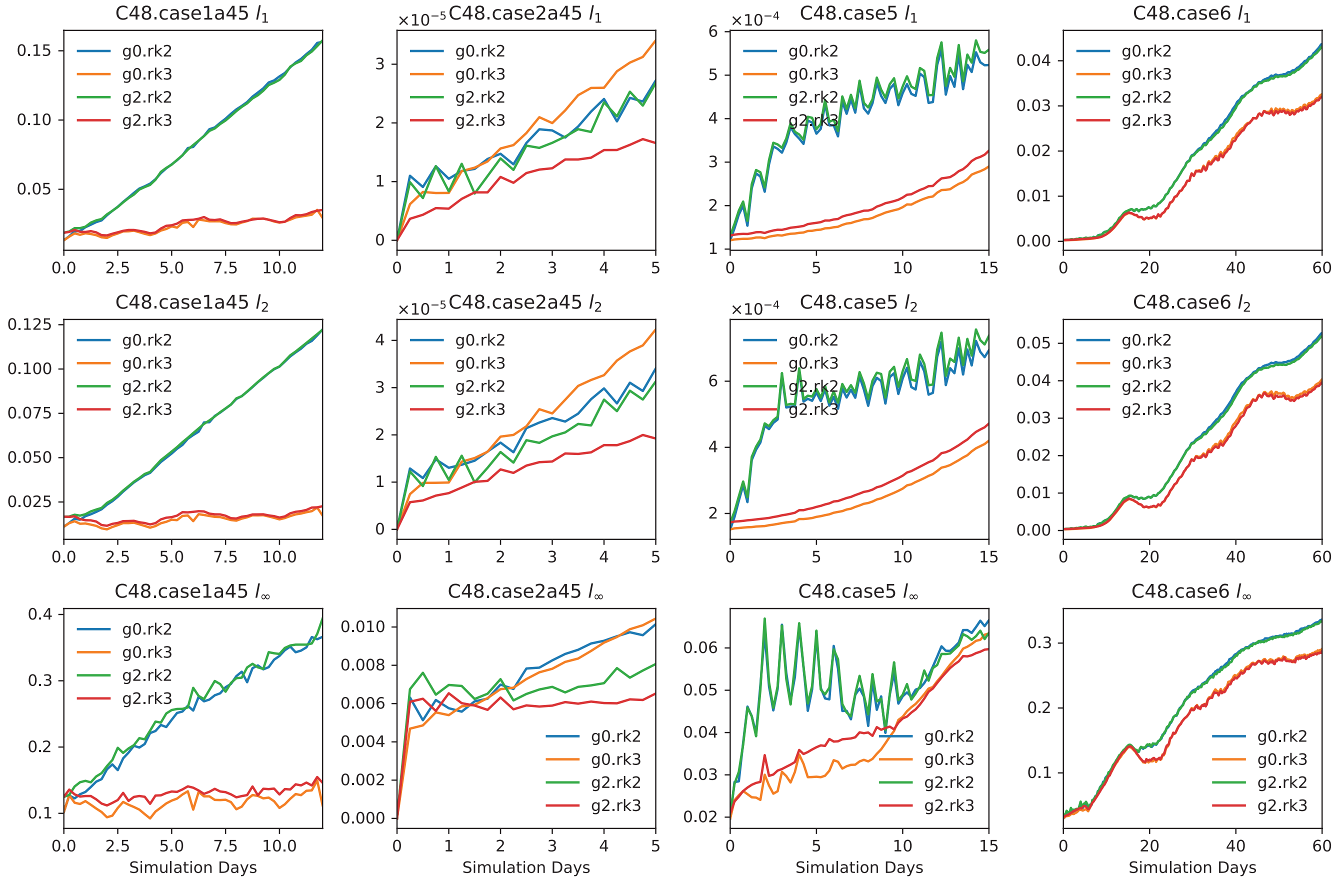

Figure 4. l_1, l_2, and l_{inf} error norms growth with the simulation time of W92 test cases 1, 2, 5, and 6 at the C48 resolution.

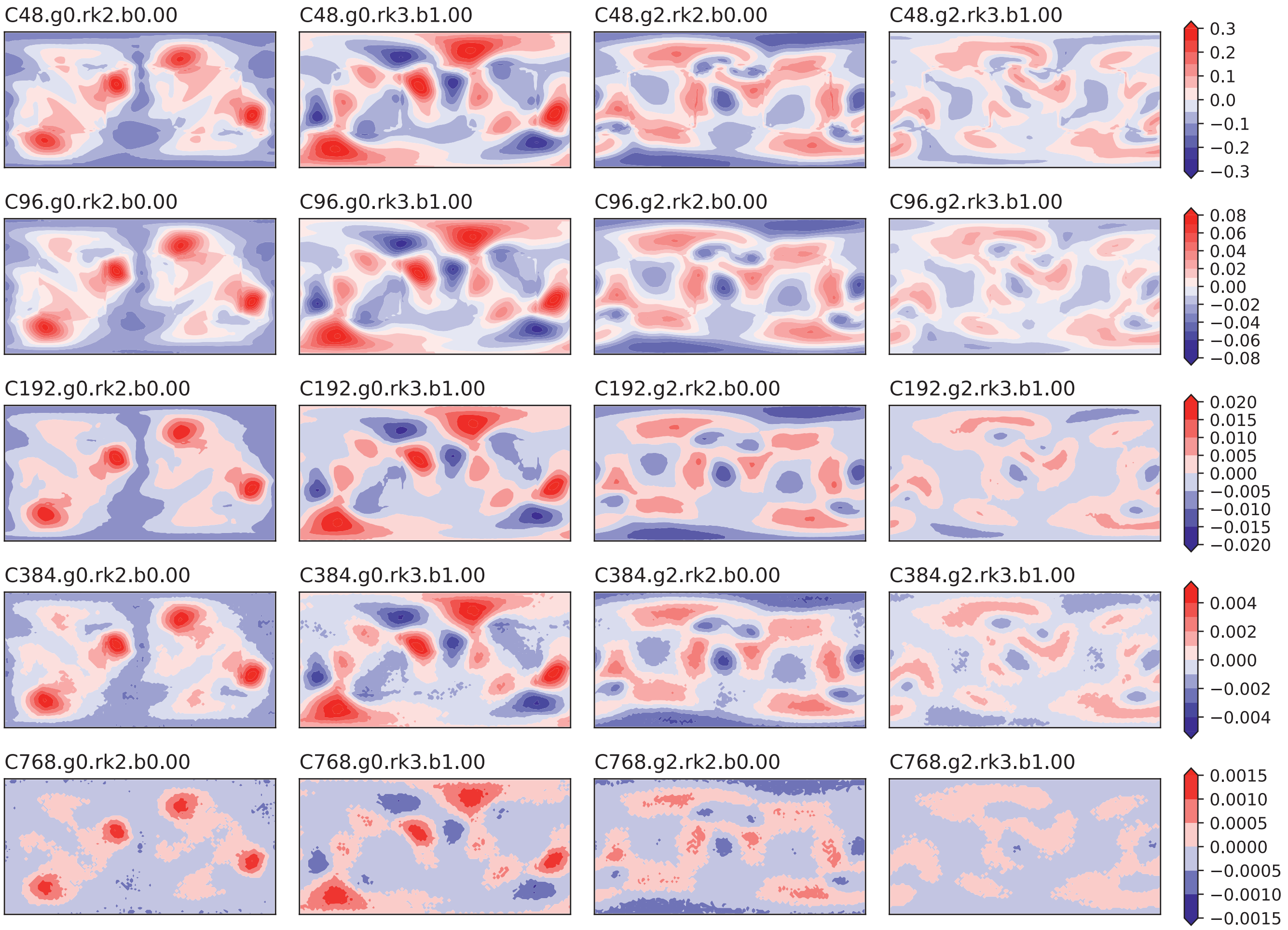

Figure 5. l_1, l_2, and l_{inf} error norms growth with the simulation time of W92 test cases 1, 2, 5, and 6 at the C48 resolution.

tions, which are different from other tests, all the runs produce similar magnitudes of errors. In particular, the equi-edge grid (g0) produces slightly larger errors than the equiangular grid (g2). Lastly, in case 6, the error growth curves do not saturate to a steady number, indicating that the solutions remain in-phase (X. Chen et al., 2018) even after a significant period of simulation time.

The distributions of the errors of case 2 offer a measure to the cubed-sphere grid imprinting. Figure 5 plots the bias of the solid-body rotation tests with a 45-degree configuration. The bias is the difference between the solutions at the final time and the initial time. All simulations are regridded to a 144 × 72 resolution. The results are consistent with Figure 5, C48. case2a45 subplots, especially the l_{inf} plot, where the equi-edge results in slightly larger errors than the equiangular grid. All solutions employ a simple 1D remap at tile-halo connections and therefore improve markedly than the more direct nonremapped approaches (e.g., Harris & Lin, 2013). Although small cube-edges can still be observed at coarser resolutions (C48, C96), high resolution runs almost eliminate all direct grid-imprinting errors. Lastly, it is worth repeating that although case 2 is a good measure for cubed-sphere grid imprinting, it DOES NOT provide sufficient test coverage of error analysis for generic nonlinear shallow water flows, as discussed in Figure 4.

To determine the general order of accuracy in different configurations, Figure 6 plots the l_2 error against different resolutions at specific dates. The results indicate that rk3 runs consistently maintain second-order overall accuracy, which is expected since no high-order multi-dimension scheme is employed in this study. The previous section explains that the forward-backward scheme (FB) is numerically equivalent to adding a 2nd-order stabilizing term to the advection equation, which will degrade the overall 2nd-order accuracy.

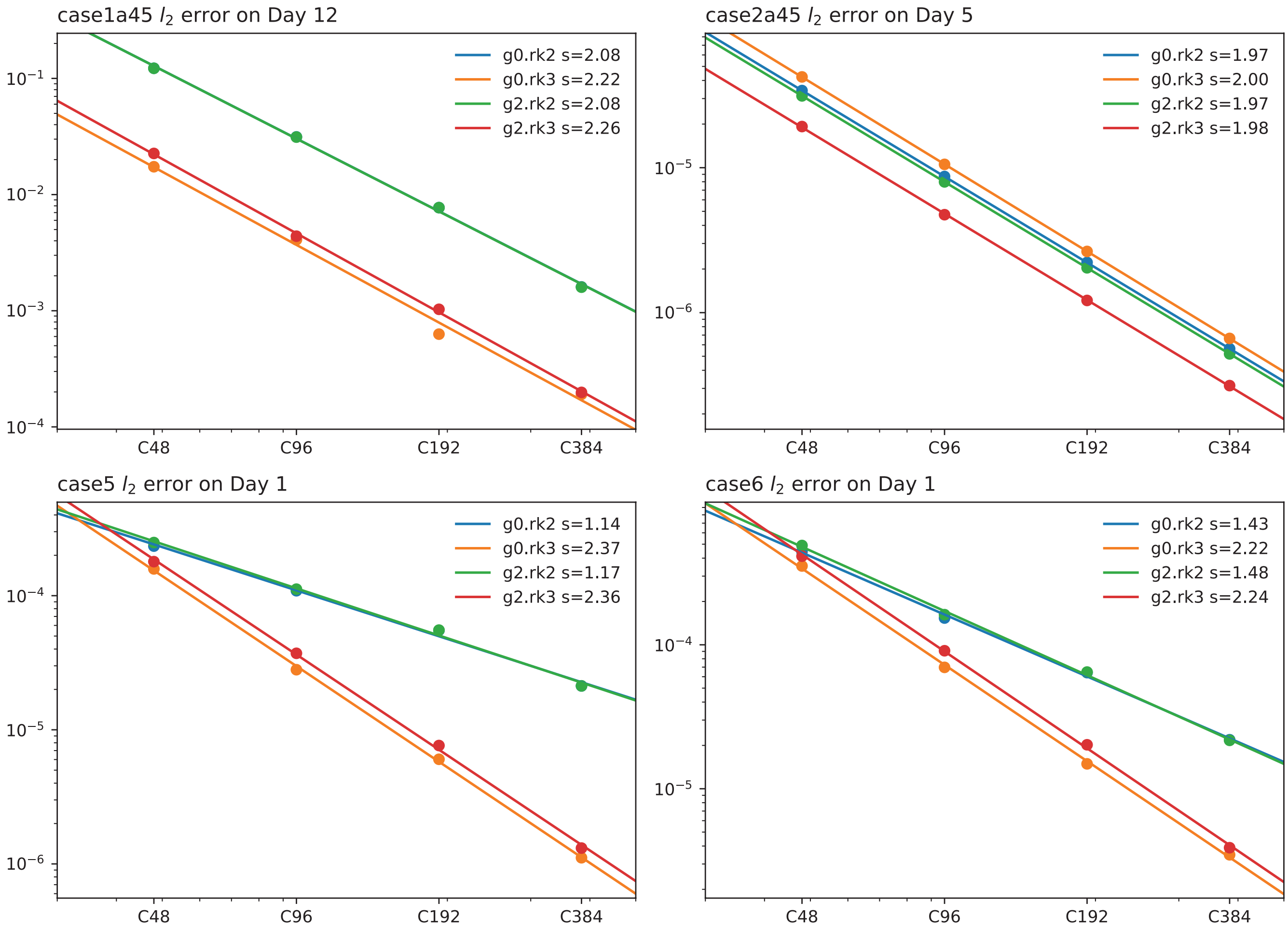

Figure 6. l_2 error against different resolutions at specific dates with different resolutions. The value of s in the legends are slopes of each line that indicates the overall order of accuracy.

Therefore, in cases 5 and 6, rk2 runs, which employ a fully forward-backward pressure gradient algorithm, demonstrate lower than 2nd-order overall accuracy. Case 1 does not involve the pressure gradient. Thus, it is not affected by FB settings. It appears that the balanced stationary case 2 is also free of FB accuracy degradation.

A more in-depth analysis of FB impact is demonstrated in Figures 7 and 8. Interestingly, the overall order of accuracy gradually increases with higher beta values in Figure 7. In Figure 8, although all results have good convergence toward C768 high-resolution solutions, FB parameter β has a significant influence on the solver's diffusion properties. The fully explicit C48. rk3. b1.00 run produces even higher maximum values (the small circle at the eastern equator) than the C768. rk3. b0.00 run with fully implicit pressure gradient evaluation.

The Rossby-Haurwitz test (case6) with a wavenumber four is adopted in various works. This test can validate the solvers' robustness in preventing the instability due to truncation error in the initial conditions. Figure 9 shows the height field at day 14, 40, 80 at resolutions C48 and C768 with rk2 and rk3 configurations. Only the equal-angular grid is displayed because no strong dependency from the grid type choices is found. The solver achieves optimum symmetry for an extended simulation time of

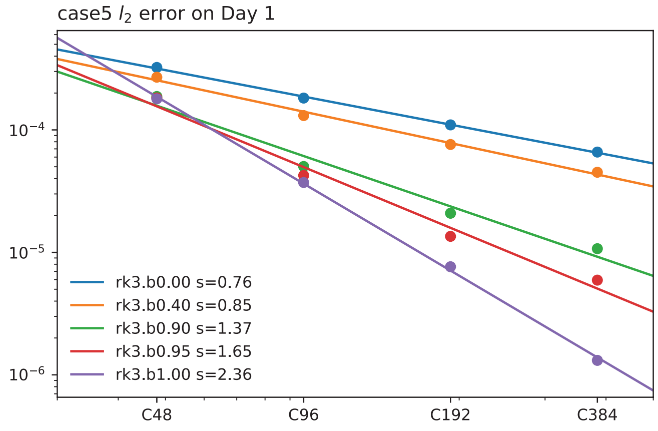

Figure 7. Same as Figure 6, but different beta values in case 5 are compared.

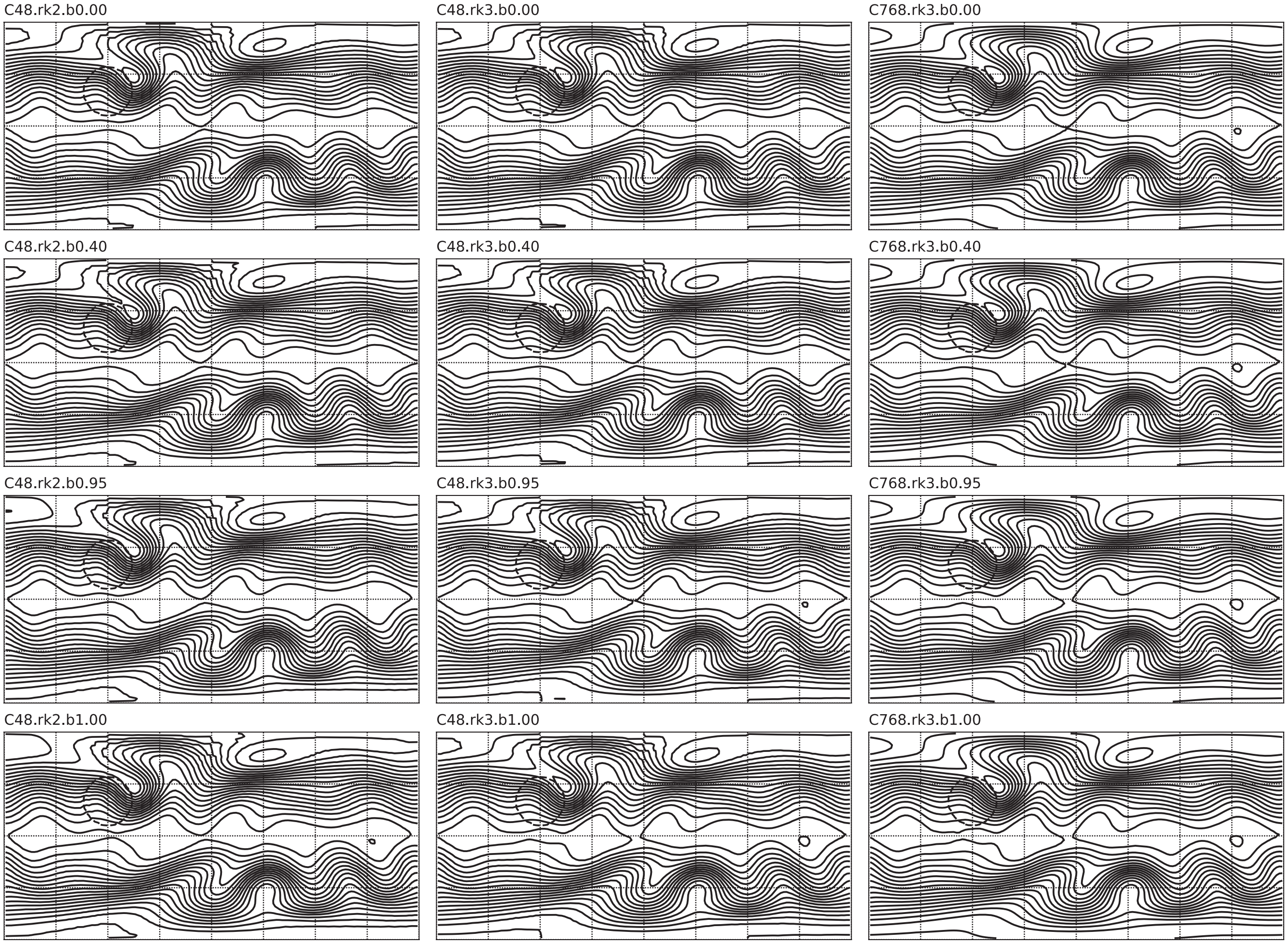

Figure 8. The final day flow height field of case 5 with C48 and C768 resolutions. Different configurations of rk2, rk3 and different strength of FB parameter β values are also compared. The dashed circle is the location of the conical mountain. The contour interval is 50 m.

over 80 days. Although both being symmetric, the high-resolution C768 reference solutions maintain much better initial modal structures and intensity than the low-resolution solutions.

4.3. Colliding Modons

The recently proposed colliding modons test (Lin et al., 2017; Wang et al., 2019; Zhang et al., 2019) evaluates the solver's ability to simulate vorticity dynamics, which plays a dominant role in large to mesoscale motions. The colliding modon pairs can fully reach steady shapes after one day and travel back to approximately their original locations at around day 100. A successful cycle can produce four symmetrical tracks in four quadrants, namely Q_{sw}, Q_{se}, Q_{nw}, Q_{ne}. The quality of the results can be evaluated by comparing the travel distances and amplitudes of the vortex.

Figure 10 demonstrates the colliding modons simulations with g2. rk3 configuration at various resolutions. The results show that the symmetry is well maintained in all runs. The simulations with C192 and better resolutions show good convergence. Figure 11 compares the travel distance and amplitude of the modons in the northwest quadrant at day 100. Note that the travel distance is measured starting at the end of day 1 to ignore initial adjustments. Although the grid choices, discretization configurations have impacts on both values, the resolution is the deterministic factor to resolve the modons movement correctly, and C192 and better resolutions result in converged values. Compared to the original work that introduced the modon

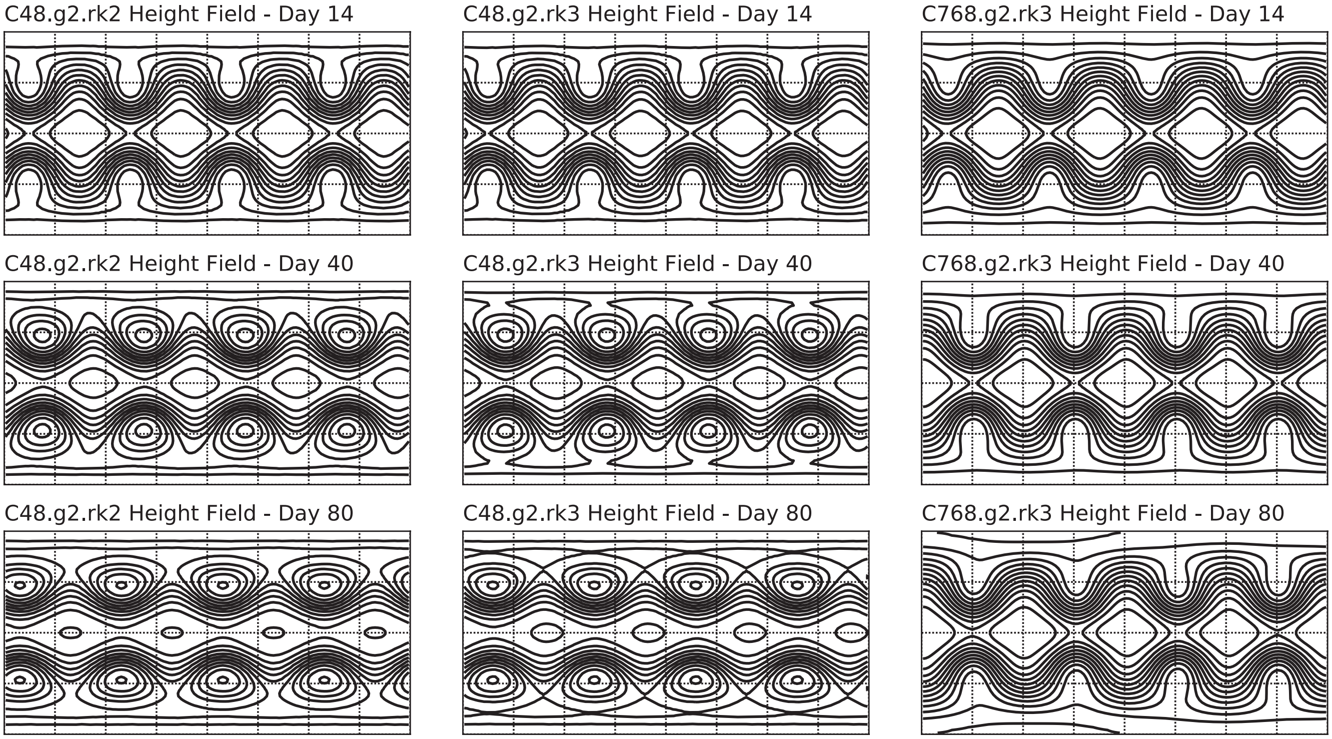

Figure 9. Height field at day 14, 40, 80 at resolutions C48 and C768 with rk2 and rk3 configurations. Only the equal-angular grid (g2) is displayed because no strong dependency from the grid type choices is found. The contour interval is 200 m. The dashed lines are the gridded latitudes and longitudes with intervals of 45 degrees.

tests, the converged final locations of the modons are consistent with both FV3 C192/C384 results and the GFDL-Spectral model T213/T511 results (see Figure 4 in Lin et al., 2017). The bar plots in Figure 11 provide a quantitative baseline for future development.

The colliding modons test provides an interesting perspective in atmospheric modeling applications. The spatial scale of each modon is at around 1,000 km. The experiments show that although a 200 km resolution (C48) model can capture the motion of the modon, the speed is not correctly simulated, and the strength of the vortex is rapidly dissipated by model internal diffusion. Considering that many rotational processes that are of interest in global circulation models are of horizontal scales less than 1,000 km, increasing the model resolution may be the most effective way to improve the accuracy in simulating many rotational events such as cyclones and storms. Recent studies also suggest significant improvements in simulating atmospheric vortices by enhancing the model resolution. For example, Murakami et al. (2015) found that by improving resolution from 50 to 25 km, the model can better capture categories 4 and 5 hurricanes. Gao et al. (2019) showed better hurricane structures by locally increasing the model resolution from 25 km to 8 km.

4.4. Splash on the Sphere

Previous subsections cover a comprehensive range of numerical model properties such as the correctness, orders of accuracy, faithful vortex representation. This section validates the dispersion and dissipation properties at various wavelengths by introducing a “splash on the sphere” test. Historically, unstaggered solvers are considered inferior in representing the short wave propagation due to poor dispersive properties. However, such conclusions usually are based on the low-order linearized analysis, which does not reflect high-order complex real-world applications. By discretizing the high-order linearized analysis in Courant numbers and numerical phase, Chen et al. (2018) found the grid staggering choices have negligible phase speed influence in high-order algorithms. Furthermore, Chen et al. (2018) also introduced a simple 1D test to visualize the dispersion behavior at short waves and the diffusion control of grid-scale noise. Inspired by

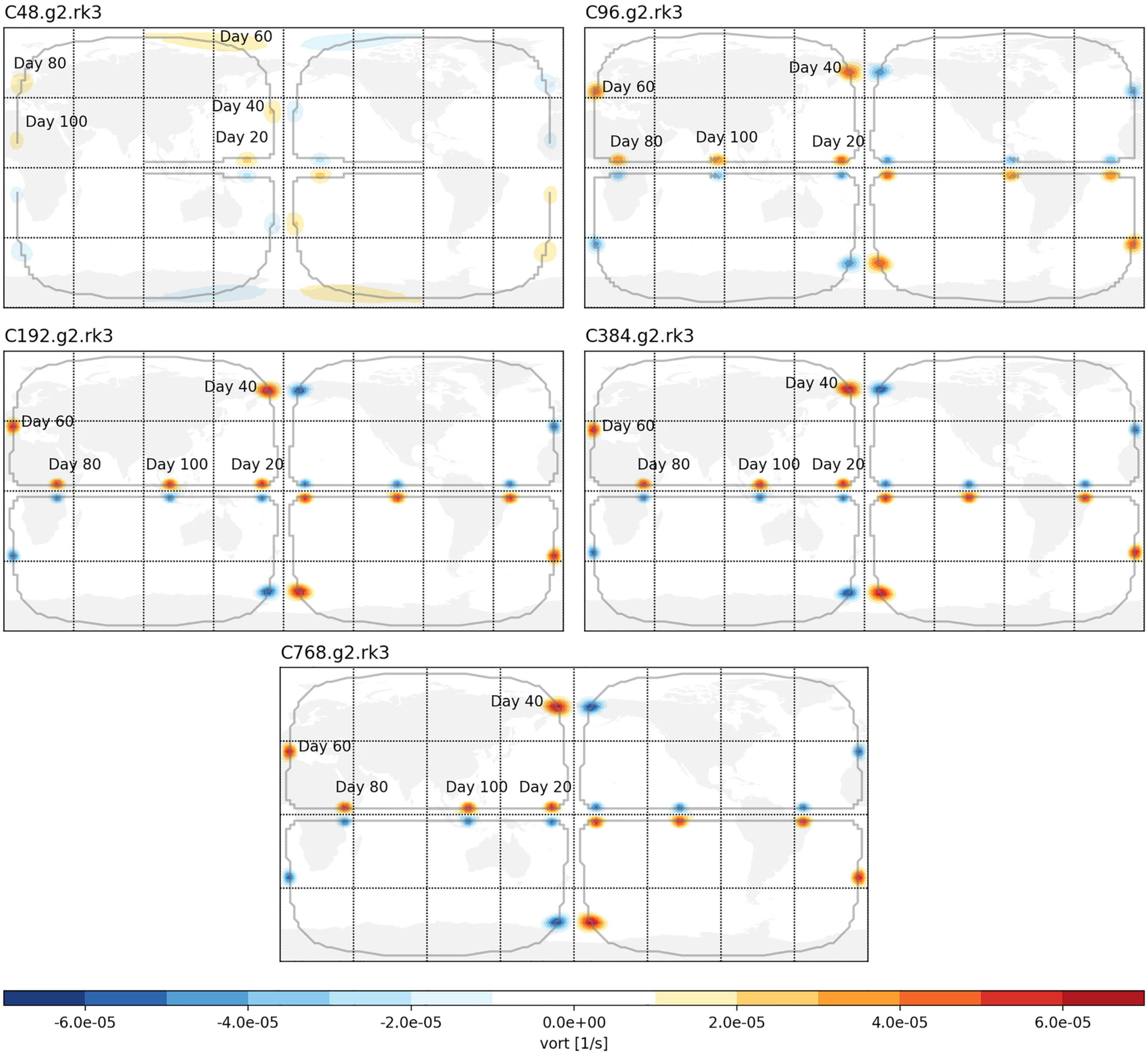

Figure 10. Colliding modons simulations with g2. rk3 configuration at various resolutions.

this work, this section presents an extended 2D version of a "splash on the sphere" test to evaluate solver capabilities in faithfully represent short wave propagations.

The test has a simple physical process to splash a sinusoidal droplet at the North Pole and freely propagate on the nonrotational sphere. The choice of the sinusoidal signal is to limit the waves in a single modal mode and test the sharp gradient between the perturbation and the background when a monotonic algorithm is implemented in the future. The Coriolis coefficient *f* and the topography are both zero, and the initial velocity field is stationary. The height field with the initial splashing perturbation is defined:

$$\Phi = \begin{cases} \Phi_0 + \Phi' \cos\left(\dfrac{\pi}{2}\dfrac{r}{R}\right) & \text{if } r < R, \\ \Phi_0 & \text{else,} \end{cases} \tag{38}$$

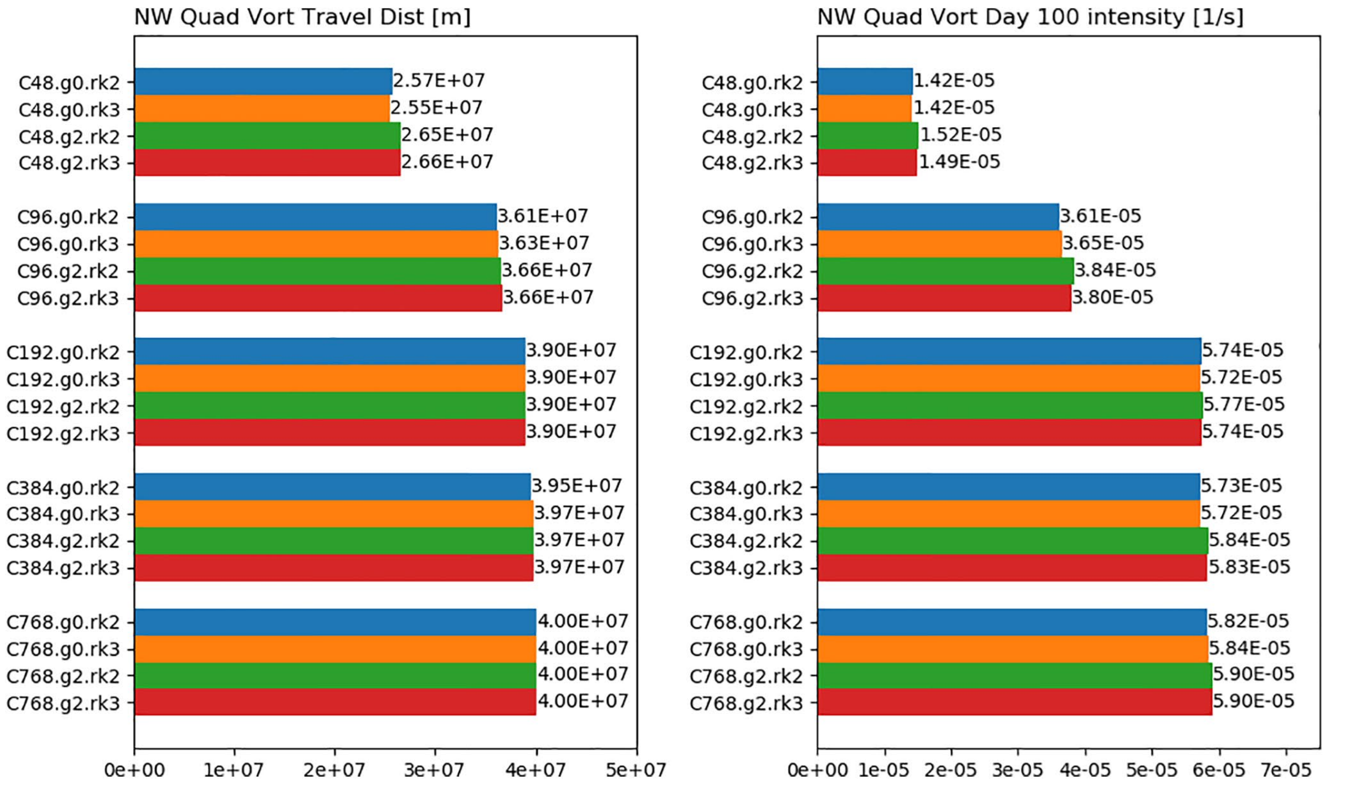

Figure 11. The travel distance and amplitude of the modons in the northwest quadrant at day 100. Note that the travel distance is measured starting at the end of day 1 to ignore initial adjustments.

where r is the distance from the north pole, and $\Phi_0 = 50\,\text{m} \times G \approx \left(\pi R_e / 10\,\text{Day}\right)^2$, $\Phi' = 1\ \text{m} \times G$, and $R = 500$ km. The design of a 50 m background flow height allows the propagating waves to reach the South Pole at approximately 10 days, which is an arbitrary choice. Figure 12 illustrates the exaggerated height field at Days 0, 2, 5, and 7.

The analysis of this work focuses on day-5 snapshots when the splash waves propagate to the equatorial region. The results in all resolutions are regridded to 144 × 72 latitude-longitude resolution. Figure 13 shows the day-5 snapshots in the meridional direction of g2. rk3 runs at various resolutions. All simulations propagate the splash at the same location, indicating no phase error is observed. The C48 run shows an obvious spreading of the wave due to the excessive diffusion at this wavelength. Since this work does not implement any monotonic algorithm, overshoots and undershoots are also observed in the results.

Figure 14 quantitatively compares day-5 snapshots of distances to the north pole (NP), peak flow height, and maximum zonal wind absolute values with different discretization settings, gird choices, and resolutions. The identical distances to NP indicate the solver's excellent dispersive properties, which is highly desired with the unstaggered algorithm. In contrast to the modon tests, the peak flow height comparison shows an intriguing observation that the model's intrinsic diffusion properties have an even more significant influence on the results than the resolution differences. In the last column, a perfect numerical solver should not produce any zonal wind in this test. The nonzero values are due to the imperfect representation of the curvilinear cubed-sphere geometry. Therefore, it is an excellent test to check the grid discretization performance. It is reassuring to find that the different grid choices do not produce excessive discrepancies in geometric errors. The resolutions and discretization play more critical roles than grid choices.

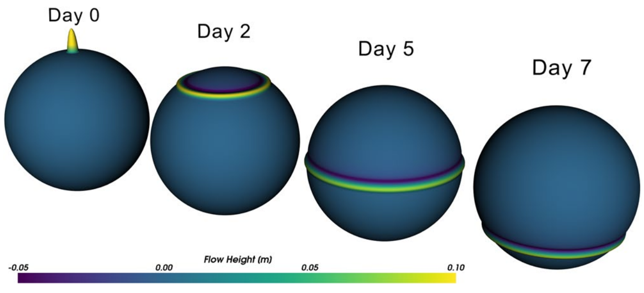

Figure 12. Exaggerated height field of the splash on the sphere test at Days 0, 2, 5, and 7.

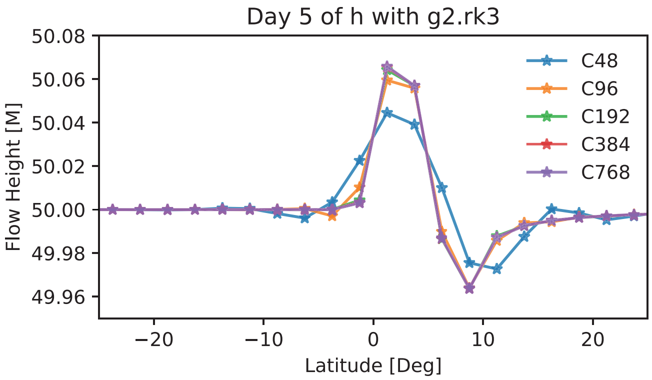

Figure 13. Day-5 snapshots in the meridional direction of g2. rk3 runs at various resolutions.

This work only tests one splash configuration for brevity. In Chen et al. (2018), a square wave is also an informative setup to investigate sharp gradient handlings. In future work, a "square splash" can be an appealing option once more sophisticated local monotonic algorithms are included. The tracking of maximum zonal wind absolute values can be a useful gauge of high-order multi-dimension algorithms.

4.5. Discussion

The model demonstrates competitive performance in all numerical tests. Even without any explicit filter, the solver exhibits a broad range of organic implicit diffusion properties, mainly controlled by reconstruction stencil size and strength of forward-backward pressure gradient settings. Considering a full-fledged GCM is an interplay between numerical methods and highly uncertain physics parameterization, flexibility in diffusion control is helpful.

The results provide some perspectives in faithfully simulating various processes. The rotational motions in the colliding modon tests and the propagation of characteristics in the splash tests are both about a 1,000-km spatial scale. Although the C48 simulations can capture both events, the qualities are significantly less reliable than those of higher resolution counterparts. C96 is the recommended minimum resolution to accurately represent these processes, which is about 10 times grid-spacing. This guideline also indicates the unstaggered solver has balanced capabilities in resolving the rotational mode, and the divergence mode.

With the duo-grid system and the tile-edge remapping, the solver is free of any grid imprinting issues. The grid-type choices show negligible differences in the solutions compared to discretization differences and resolutions, which indicates the desired consistency in grid choices for various purposes.

Lastly, most numerical decisions are balanced assessments with both accuracy and performance in consideration. For example, the theoretically best accuracy provided by the fully nonlinear 4-step Runge-Kutta

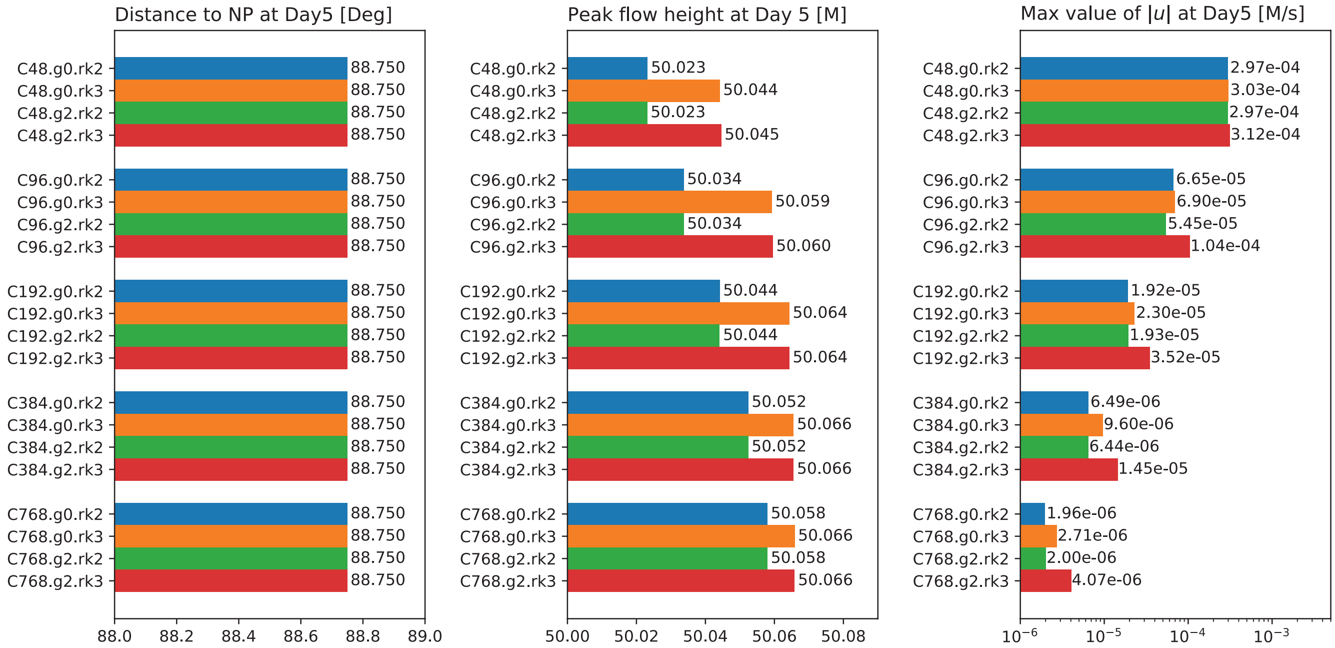

Figure 14. Day-5 snapshots of distances to the north pole (NP), peak flow height, and maximum zonal wind absolute values with different discretization settings, different gird choices, and resolutions.

schemes, or multi-dimension cubed-sphere halo remapping schemes cannot justify the considerable overhead in both computing memory requirements and computational costs or the workload imbalance to the modern paralleled computing platforms. On the other hand, a single-step time-marching scheme provides the best speed but cannot avoid undesired low-accuracy due to excessive diffusion for the stability requirement, such as the fully forward-backward pressure gradient scheme (see discussions in Chen et al., 2018). The more deterministic numerical decisions can form upon creating the full 3D model, and in-depth software-level optimization, such as Müller et al. (2019), is crucial for the final production applications.

5. Conclusions and Future Work

In this study, we have successfully demonstrated the LMARS-based shallow-water solver on two gnomonic cubed-sphere grids with controllable variations of computational efficiency and diffusion properties. We use the W92 test suite to warrant the model to produce results that converge to the literature. The model illustrates a broad range of diffusion and controls the delicate balance between the order of accuracy and the robustness by various dynamical core parameters. In particular cases, numerical tests demonstrate that many resolvable processes are not necessarily reliable. We also designed a test to isolate numerical properties by visualizing the dissipation and dispersion. Lastly, Griffies (2004, Chapter 17) discussed that explicit frictional operators might alter pressure gradient and angular momentum due to imbalanced design. With no explicit damping, the control of implicit diffusion in this model is organic and does not cause artifacts in pressure gradient calculations.

There are several new ideas in this study. We introduced the duo-grid system that unifies arbitrary gnomonic cubed-sphere grid representations. It also provides the 1D alignment for arbitrary gnomonic grids at halo regions. This work also extends the L97 pressure-gradient integration technique with LMARS diffusion contribution, resulting in a fast and stable discretization. In the numerical test section, the extension of more quantitative metrics to the modon tests allows straightforward comparison between different models and configurations. The splash on the sphere test offers an economical way to get 1D dispersion and dissipation properties of the model without extracting the schemes into a 1D testbed.

Balancing the computational performance and numerical accuracy is one of the top motivations in this development. Therefore, this work implements several optimizations. Some improvement does not sacrifice numerical accuracy. For example, the choice of unit-length basis vectors significantly simplifies the mathematical expressions of the curvilinear system, thus avoiding the computational storage of redundant vector transformation matrices. On the other hand, although halo region 1D remap can be made more sophisticated with higher-order schemes, we choose the simple piecewise linear reconstruction to maintain a minimum unbalanced workload in parallel computing. Lastly, a shallow-water testbed cannot determine the most optimal solver diffusion properties in a fully coupled GCM. Therefore, this work provides a guideline for a controllable range of organic numerical properties once integrated into a full model.

The next step is the extension of this work to a 3D compressible atmosphere with the cubed-sphere grid. Previous work has implemented LMARS in a vertical 2D compressible atmosphere with both Eulerian and Lagrangian vertical coordinates (Chen et al., 2013). Li and Chen (2019) demonstrated an energy-conserving model with simple microphysics under the LMARS Eulerian framework. Therefore, the LMARS 3D extension on the cubed-sphere grid has several core components validated in various environments. Another future topic is to include more sophisticated advection components, such as monotonic constraints (van Leer, 1979), high-order multi-dimension transport schemes (Lin & Rood, 1996) in the horizontal direction. With the vertical Lagrangian coordinates, this development shares substantial similarities with the FV3 framework, which can guide future research.

Appendix A: List of Constants, Symbols and Their Values and Units

To make the notations consistent and clear, Table A1 summarizes all important constants, symbols, their values and units in this study for reference.

Table A1
List of Parameters and Symbols Used Throughout This Document

Symbol	Description	Values	Units
G	Gravitational constant	9.80665	ms^{-2}
Ω	Rotational speed of the Earth	7.292e-5	s^{-1}
R	Radius of the Earth	6.3712e6	m
$\hat{e}_\lambda$	Unit vector in zonal direction	1	NA
$\hat{e}_\phi$	Unit vector in meridional direction	1	NA
$\hat{e}_1$	Unit vector in x-direction (local to cubed-sphere tile)	1	NA
$\hat{e}_2$	Unit vector in y-direction (local to cubed-sphere tile)	1	NA
$\hat{k}$	Unit vector in sphere radius direction	1	NA
λ	Longitude	-	rad
Φ	Latitude	-	rad
i, j	Indices to label grid points in x- and y-directions	-	NA
α	angle between two curvilinear coordinates	-	rad
g_{ij}	Covariant 2D metric tensor of the curvilinear grid system, note the *ij* in this context is the dimension iteration in Einstein Notation	-	NA
g^{ij}	Contra-variant 2D metric tensor of the curvilinear grid system	-	NA
h	Thickness of the fluid	-	m
Π	Thickness of the fluid in term of geopotential	-	m^2s^{-2}
Φ	Geopotential	-	m^2s^{-2}
Φ_s	Surface geopotential	-	m^2s^{-2}
$\vec{u}$	2D velocity vector	-	ms^{-1}
f	Coriolis parameter	-	s^{-1}
ζ	Vertical component of the relative vorticity	-	s^{-1}
u	Covariant wind component in x-direction	-	ms^{-1}
v	Covariant wind component in y-direction	-	ms^{-1}
$\tilde{u}$	Contra-variant wind component in x-direction	-	ms^{-1}
$\tilde{v}$	Contra-variant wind component in y-direction	-	ms^{-1}
g	Metric coefficient of the curvilinear geometry	-	NA
u_λ	Wind component in zonal direction	-	ms^{-1}
u_ϕ	Wind component in meridional direction	-	ms^{-1}
J^{c2l}	Matrix to convert $\left(u, v\right)$ to $\left(u_\lambda, u_\phi\right)$	-	NA
J^{l2c}	Matrix to convert $\left(u_\lambda, u_\phi\right)$ to $\left(u, v\right)$	-	NA
$u_\perp^x$	Wind projection perpendicular to cell interface in x-direction	-	ms^{-1}
$u_{\|}^x$	Wind projection parallel to cell interface in x-direction	-	ms^{-1}
$u_\perp^y$	Wind projection perpendicular to cell interface in y-direction	-	ms^{-1}
$u_{\|}^y$	Wind projection parallel to cell interface in y-direction	-	ms^{-1}
a	Gravity wave speed (group speed of the fluid)	-	ms^{-1}

Appendix B: Gnomonic Projection Initialization Algorithms

This section provides a detailed mathematical reference-line based gnomonic cubed-sphere generation process for reference.

B1 Basic geometric algorithms

Each gnomonic projection can be described by a 3D Cartesian vector $\vec{p}(x, y, z)$, which is uniquely determined by a point on the sphere or a 3D vector on the spherical coordinate with unit length $\vec{p}(\lambda, \phi)$. The valid information is the direction of the vector, and the length of the vector is trivial. The conversion can be written:

$$x = \cos\phi \cos\lambda, \tag{B1}$$

$$y = \cos\phi \sin\lambda, \tag{B2}$$

$$z = \sin\phi, \tag{B3}$$

and

$$\lambda = \arctan\frac{y}{x}, \tag{B4}$$

$$\phi = \arcsin\frac{z}{\sqrt{x^2 + y^2 + z^2}}, \tag{B5}$$

where the Fortran intrinsic functions atan2 () is used to place λ in the range of $-\pi < \lambda \leq \pi$.

To calculate the great circle distance between two points:

$$\mathrm{dist}\left(\vec{p}_1(\lambda_1, \phi_1), \vec{p}_2(\lambda_2, \phi_2)\right) = R \arccos\left(\frac{\vec{p}_1 \cdot \vec{p}_2}{|\vec{p}_1||\vec{p}_2|}\right), \tag{B6}$$

or

$$\mathrm{dist}\left(\vec{p}_1(\lambda_1, \phi_1), \vec{p}_2(\lambda_2, \phi_2)\right) = R \arccos\left(\cos\phi_1 \cos\phi_2 \cos(\lambda_1 - \lambda_2) + \sin\phi_1 \sin\phi_2\right). \tag{B7}$$

To calculate the area of a grid cell:

$$\mathrm{area}\left(\vec{p}_1, \vec{p}_2, \vec{p}_3, \vec{p}_4\right) = R^2\left(\alpha_{412} + \alpha_{123} + \alpha_{234} + \alpha_{341} - 2\pi\right), \tag{B8}$$

where α_{ijk} is the angle between three points $\left(\vec{p}_i, \vec{p}_j, \vec{p}_k\right)$ with “right-hand rule” from $\vec{p}_j$, and can be calculated by:

$$\alpha_{ijk} = \arccos\left(\frac{\left(\vec{p}_i \times \vec{p}_j\right) \cdot \left(\vec{p}_j \times \vec{p}_k\right)}{\left|\vec{p}_i \times \vec{p}_j\right| \cdot \left|\vec{p}_j \times \vec{p}_k\right|}\right). \tag{B9}$$

The mid-point of two vectors $\vec{p}_1$ and $\vec{p}_2$ is:

$$\mathrm{mid}\left(\vec{p}_1, \vec{p}_2\right) = \frac{\vec{p}_1 + \vec{p}_2}{2}. \tag{B10}$$

B2 Initializing gnomonic projection with reference lines on tile one

As illustrated in Figure 3, the eight corners of the tiles in 3D Cartesian coordinates are $(x, y, z) = (\pm1, \pm1, \pm1)$. The resolution of the cubed-sphere grid is denoted by $C[N]$, which indicates $N \times N \times 6$ total grid cells. Let the first tile be centered at $(\lambda, \phi) = (0, 0)$ or the cube-tile with vertices $(x, y, z) = (1, \pm1, \pm1)$ and denote (X, Y) the local coordinates on this cube-tile, the 3D Cartesian coordinates can be written by:

$$(x, y, z) = (1, X, Y), \tag{B11}$$

and the projection vector can be also written $\vec{p}(X, Y)$. Additionally, let (i, j) be the indices of the grid points on tile one, and the indices start counting from 1. Therefore, the coordinate sets (x, y, z), (λ, ϕ), (X, Y) and (i, j) are interchangeable and representing the same gnomonic projection vector. The vertices on the tile can be uniquely determined by a column of Y values on the cube tile surface.

B21 Equidistance projection

To get the Y values of the red reference column:

$$Y(i_{\mathrm{ref}}, j) = -1 + (j - 1)\delta h, \tag{B12}$$

where $\delta h = 2/N$, $i_{\mathrm{ref}} = N/2 + 1$. Then the red reference row is:

$$X(i, j_{\mathrm{ref}}) = Y(i_{\mathrm{ref}}, i), \tag{B13}$$

where $i_{\mathrm{ref}} = N/2 + 1$, and the grid mesh on the cube tile can be populated from the two reference lines accordingly:

$$X(i, j) = X(i, j_{ref}) = Y(i_{\mathrm{ref}}, i), \tag{B14}$$

$$Y(i, j) = Y(i_{\mathrm{ref}}, j). \tag{B15}$$

Therefore, $\vec{p}(X, Y)$ can be easily converted to $\vec{p}(x, y, z)$ and $\vec{p}(\lambda, \phi)$, and the grid points on the cubed-sphere tile is generated. One may observe that the values of i_{ref} and j_{ref} are trivial

B22 Equiangular projection

Comparing to equidistance projection, the equiangular grid generation requires an extra first step to project grid points from the green reference column to the cube tile and get the Y values. The angular step of the grid points on the green row is:

$$\delta\zeta = \frac{2\alpha_{\mathrm{ref}}}{N}, \tag{B16}$$

where $\alpha_{\mathrm{ref}} = \pi/4$. Then the Y values are:

$$Y(i_{\mathrm{ref}}, j) = R_{\mathrm{ref}} \tan\left((j - 1)\delta\zeta - \alpha_{\mathrm{ref}}\right), \tag{B17}$$

where $R_{\mathrm{ref}} = 1$. Then the entire set of grid points on the tile can be populated following the same procedure as the equidistance projection.

B23 Equi-edge projection

The procedures to generate the equi-edge projection are identical to the equiangular projection except different parameters: $\alpha_{\mathrm{ref}} = \arcsin(\sqrt{3})$, $R_{\mathrm{ref}} = \sqrt{2}$.

B3 Populating grid points to six tiles

Once the grid points on tile one are generated, the points on the rest of the tiles can be populated using the "staircase" tile-interlock pattern. The basic rotations to any 3D vector can be decomposed by the combination of the rotations about each axes of the Cartesian coordinate system. The rotation about each axes by an angle β following right-hand rule:

$$R_x(\beta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\beta & -\sin\beta \\ 0 & \sin\beta & \cos\beta \end{pmatrix}, \tag{B18}$$

$$R_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{pmatrix}, \tag{B19}$$

$$R_z(\beta) = \begin{pmatrix} \cos\beta & -\sin\beta & 0 \\ \sin\beta & \cos\beta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{B20}$$

In particular,

$$R_x(90^\circ) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{B21}$$

$$R_y(90^\circ) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \tag{B22}$$

$$R_z(90^\circ) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{B23}$$

Therefore, the rotations for each tile:

$$\vec{p}(i,j;\text{tile}=2) = R_z(90^\circ)\,\vec{p}(i,j;\text{tile}=1), \tag{B24}$$

$$\vec{p}(i,j;\text{tile}=3) = R_x(90^\circ)R_z(90^\circ)\,\vec{p}(i,j;\text{tile}=1), \tag{B25}$$

$$\vec{p}(i,j;\text{tile}=4) = R_x(90^\circ)R_z(180)\,\vec{p}(i,j;\text{tile}=1), \tag{B26}$$

$$\vec{p}(i,j;\text{tile}=5) = R_y(90^\circ)R_z(270)\,\vec{p}(i,j;\text{tile}=1), \tag{B27}$$

$$\vec{p}(i,j;\text{tile}=6) = R_y(90^\circ)\vec{p}(i,j;\text{tile}=1), \tag{B28}$$

or:

$$\vec{p}(i,j;\text{tile}=1)=[1,X,Y], \tag{B29}$$

$$\vec{p}(i,j;\text{tile}=2)=[-X,1,Y], \tag{B30}$$

$$\vec{p}(i,j;\text{tile}=3)=[-X,-Y,1], \tag{B31}$$

$$\vec{p}(i,j;\text{tile}=4)=[-1,-Y,-X], \tag{B32}$$

$$\vec{p}(i,j;\text{tile}=5)=[Y,-1,-X], \tag{B33}$$

$$\vec{p}(i,j;\text{tile}=6)=[Y,X,-1]. \tag{B34}$$

Appendix C: The Governing Equations

This shallow water model employs the vector-invariant form governing equations:

$$\frac{\partial h}{\partial t}=-\nabla\cdot\left(h\vec{u}\right), \tag{C1}$$

$$\frac{\partial\vec{u}}{\partial t}=-(\zeta+f)\vec{k}\times\vec{u}-\nabla\left(\left(\Pi+\Phi_s\right)+\frac{1}{2}\vec{u}\cdot\vec{u}\right), \tag{C2}$$

where

$$\Pi=Gh, \tag{C3}$$

$$\zeta=\hat{k}\cdot(\nabla\times\vec{u}), \tag{C4}$$

$$f=2\Omega\sin\phi. \tag{C5}$$

The definition of the variables, constants, and their values are summarized in Table A1.

C1 Generic form in a curvilinear grid system

This solver uses covariant wind components as the prognostic variable. The covariant wind components are the projections of the wind vector on the curvilinear coordinates:

$$u=\vec{u}\cdot\hat{e}_1, \tag{C6}$$

$$v=\vec{u}\cdot\hat{e}_2, \tag{C7}$$

and the corresponding contra-variant wind components satisfy:

$$\vec{u}=\tilde{u}\hat{e}_1+\tilde{v}\hat{e}_2. \tag{C8}$$

A covariant 2D metric tensor g_{ij} of the curvilinear grid system on a tile of the cubed-sphere is defined by Einstein Notation:

$$g_{ij}=\hat{e}_i\cdot\hat{e}_j, \tag{C9}$$

with the metric coefficient:

$$g = \det(g_{ij}). \tag{C10}$$

Note $\hat{e}_i$ and $\hat{e}_j$ in the Einstein Notation are iterations along dimensions $\hat{e}_1$ and $\hat{e}_2$. The contra-variant 2D metric tensor g^{ij} is the inverse of g_{ij}, and the conversions between co-variant and contra-variant vector component are:

$$\begin{bmatrix} u \\ v \end{bmatrix} = g_{ij} \begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix}, \tag{C11}$$

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} = g^{ij} \begin{bmatrix} u \\ v \end{bmatrix}. \tag{C12}$$

C2 The choice of the basis vectors and the optimizations

The choice of the definition of $\hat{e}_1$ and $\hat{e}_2$ are flexible. Modelers made various definitions of these terms in the literature. To yield concise mathematical expressions in numerical discretization, the covariant unit vectors are set to unit length in space, or $|\hat{e}_1| = |\hat{e}_2| = 1$. Thus, the values at the each grid point (i, j) on a cubed-sphere tile can be calculated in a discretized form:

$$(\hat{e}_1)_{i,j} = \frac{\left(\hat{k}_{i,j} \times \hat{k}_{i+1,j}\right) \times \hat{k}_{i,j}}{\left|\left(\hat{k}_{i,j} \times \hat{k}_{i+1,j}\right) \times \hat{k}_{i,j}\right|}, \tag{C13}$$

$$(\hat{e}_2)_{i,j} = \frac{\left(\hat{k}_{i,j} \times \hat{k}_{i,j+1}\right) \times \hat{k}_{i,j}}{\left|\left(\hat{k}_{i,j} \times \hat{k}_{i,j+1}\right) \times \hat{k}_{i,j}\right|}, \tag{C14}$$

with:

$$\hat{k}(\lambda, \phi) = \begin{bmatrix} \cos\phi \cos\lambda \\ \cos\phi \sin\lambda \\ \sin\phi \end{bmatrix}. \tag{C15}$$

This algorithm to determine $\hat{e}_1$ and $\hat{e}_2$ offers better flexibility when constructing various implementations of the cubed-sphere grids.

Define α the angle between two local curvilinear coordinates, or the unit vectors $\hat{e}_1$ and $\hat{e}_2$ on a cubed-sphere tile, it yields:

$$\sin\alpha = \hat{k} \cdot (\hat{e}_1 \times \hat{e}_2), \tag{C16}$$

$$\cos\alpha = \hat{e}_1 \cdot \hat{e}_2. \tag{C17}$$

Note that discrete grid subscripts i, j are omitted for cleaner expression, since all variables are co-located.

This set of basis produces concise mathematical forms of the metric tensors:

$$g_{ij} = \begin{bmatrix} 1 & \cos\alpha \\ \cos\alpha & 1 \end{bmatrix}, \tag{C18}$$

$$g^{ij} = \frac{1}{\sin^2\alpha}\begin{bmatrix} 1 & -\cos\alpha \\ -\cos\alpha & 1 \end{bmatrix}, \tag{C19}$$

$$\sqrt{g} = \sin\alpha. \tag{C20}$$

C3 Conversions of vector components to the regular latitude-longitude 924 coordinates

The latitude-longitude basis vectors are:

$$\hat{e}_\lambda(\lambda,\phi) = \frac{1}{\cos\phi}\frac{\partial\hat{k}}{\partial\lambda} = \begin{bmatrix} -\sin\lambda \\ \cos\lambda \\ 0 \end{bmatrix}, \tag{C21}$$

$$\hat{e}_\phi(\lambda,\phi) = \frac{\partial\hat{k}}{\partial\phi} = \begin{bmatrix} -\sin\phi\cos\lambda \\ -\sin\phi\sin\lambda \\ \cos\phi \end{bmatrix}, \tag{C22}$$

and define the regular zonal and meridional wind components u_λ and u_θ. The conversion matrices are:

$$J^{c2l} = \begin{bmatrix} \hat{e}_1\cdot\hat{e}_\lambda & \hat{e}_2\cdot\hat{e}_\lambda \\ \hat{e}_1\cdot\hat{e}_\phi & \hat{e}_2\cdot\hat{e}_\phi \end{bmatrix} g^{ij}, \tag{C23}$$

$$J^{l2c} = \frac{g_{ij}}{\det\left(J^{c2l}\right)}\begin{bmatrix} J_{11}^{c2l} & -J_{12}^{c2l} \\ -J_{21}^{c2l} & J_{22}^{c2l} \end{bmatrix}, \tag{C24}$$

with the conversion relations:

$$\begin{bmatrix} u_\lambda \\ v_\phi \end{bmatrix} = J^{c2l}\begin{bmatrix} u \\ v \end{bmatrix}, \tag{C25}$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = J^{l2c}\begin{bmatrix} u_\lambda \\ v_\phi \end{bmatrix}. \tag{C26}$$

C4 Conversions of vector components to the local orthogonal coordi-939 nates at cell interfaces

The preparation of the Riemann solver requires the wind vectors to be projected to the local orthogonal coordinates at the cell interfaces. Therefore, at cell interfaces in x-direction:

$$u_\perp^x = \tilde{u}\sin\alpha, \tag{C27}$$

$$u_\parallel^x = \tilde{u}\cos\alpha + \tilde{v} = v, \tag{C28}$$

or:

$$\begin{bmatrix} u_\perp^x \\ u_\parallel^x \end{bmatrix} = \begin{bmatrix} \sin\alpha & 0 \\ \cos\alpha & 1 \end{bmatrix}\begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix}, \tag{C29}$$

with inversion:

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{\sin\alpha} & 0 \\ -\dfrac{\cos\alpha}{\sin\alpha} & 1 \end{bmatrix} \begin{bmatrix} u_{\perp}^{x} \\ u_{\parallel}^{x} \end{bmatrix} \tag{C30}$$

At cell interfaces in y-direction:

$$u_{\perp}^{y} = \tilde{v}\sin\alpha, \tag{C31}$$

$$u_{\parallel}^{x} = \tilde{v}\cos\alpha + \tilde{u} = u, \tag{C32}$$

or:

$$\begin{bmatrix} u_{\perp}^{y} \\ u_{\parallel}^{y} \end{bmatrix} = \begin{bmatrix} 0 & \sin\alpha \\ 1 & \cos\alpha \end{bmatrix} \begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix}, \tag{C33}$$

with inversion:

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} = \begin{bmatrix} -\dfrac{\cos\alpha}{\sin\alpha} & 1 \\ \dfrac{1}{\sin\alpha} & 0 \end{bmatrix} \begin{bmatrix} u_{\perp}^{y} \\ u_{\parallel}^{y} \end{bmatrix} \tag{C34}$$

C5 Projecting the governing equations to the curvilinear coordinates

A small optimization can be observed from the above subsections. With basis vectors $\hat{e}_1$ and $\hat{e}_2$ being restricted to unit length, many vector conversions can be carried out with simple operations using pre-calculated values of $\sin\alpha$ and $\cos\alpha$. Therefore, the model saves considerable memory storage and computational cost by dropping many metric matrices and reducing operation counts. Moreover, the metric terms generation are more flexible in arbitrary gnomonic projections and simpler than most of the literature.

Projecting the continuity equation and the vector-invariance form momentum equation to the curvilinear coordinates yields:

$$\frac{\partial h}{\partial t} = -\frac{1}{\sqrt{g}}\left(\frac{\partial}{\partial x}\left(\sqrt{g}h\tilde{u}\right) + \frac{\partial}{\partial y}\left(\sqrt{g}h\tilde{v}\right)\right), \tag{C35}$$

$$\frac{\partial u}{\partial t} = \sqrt{g}\hat{v}\left(\zeta + f\right) - \frac{\partial}{\partial x}\left(\Pi + \Phi_s + K\right), \tag{C36}$$

$$\frac{\partial v}{\partial t} = -\sqrt{g}\hat{u}\left(\zeta + f\right) - \frac{\partial}{\partial y}\left(\Pi + \Phi_s + K\right), \tag{C37}$$

with the relative vorticity and the kinetic energy defined by:

$$\zeta = \frac{1}{\sqrt{g}}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right), \tag{C38}$$

$$K = \frac{1}{2}\left(u\tilde{u} + v\tilde{v}\right). \tag{C39}$$

These two terms can be further rearranged and optimized with better finite-volume representations in the numerical discretization process.

Data Availability Statement

The source code is hosted at https://doi.org/10.5281/zenodo.4287609. The resulting data is at https://doi.org/10.5281/zenodo.4276378. This report was prepared by Xi Chen under award NA18OAR4320123 from the National Oceanic and Atmospheric Administration, U.S. Department of Commerce. The statements, findings, conclusions, and recommendations are those of the author(s) and do not necessarily reflect the views of the National Oceanic and Atmospheric Administration, or the U.S. Department of Commerce.

Acknowledgments
The author wants to thank Lucas Harris and Stephen Griffies for the constructive GFDL internal reviews. The author would also like to thank the editors Dr. Lemariê, Dr. Pincus, and two anonymous reviewers for their valuable comments and comprehensive assessment.

References

Adcroft, A., Anderson, W., Balaji, V., Blanton, C., Bushuk, M., Dufour, C. O., & Zhang, R. (2019). The GFDL global ocean and sea ice model OM4.0: Model description and simulation features. *Journal of Advances in Modeling Earth Systems*, *11*(10), 3167–3211. https://doi.org/10.1029/2019MS001726

Adcroft, A., Campin, J.-M., Hill, C., & Marshall, J. (2004). Implementation of an atmosphere-ocean general circulation model on the expanded spherical cube. *Monthly Weather Review*, *132*(12), 2845–2863. https://doi.org/10.1175/MWR2823.1

Ainsworth, M. (2014). Dispersive behavior of high order finite element schemes for the one-way wave equation. *Journal of Computational Physics*, *259*, 1–10. https://doi.org/10.1016/j.jcp.2013.11.003

Arakawa, A., & Lamb, V. R. (1977). Computational design of the basic dynamical processes of the UCLA general circulation model. *Methods in Computational Physics*, *17*, 173–265.

Bourke, W. (1972). An efficient, one-level, primitive-equation spectral model. *Monthly Weather Review*, *100*(9), 683–689. https://doi.org/10.1175/1520-0493(1972)100⟨0683:AEOPSM⟩2.3.CO;2

Bourke, W. (1974). A multi-level spectral model. I. Formulation and hemispheric integrations. *Monthly Weather Review*, *102*(10), 687–701. https://doi.org/10.1175/1520-0493(1974)102⟨0687:AMLSMI⟩2.0.CO;2

Chen, X., Andronova, N., Van Leer, B., Penner, J. E., Boyd, J. P., Jablonowski, C., & Lin, S.-J. (2013, 7). A control-volume model of the compressible Euler equations with a vertical Lagrangian coordinate. *Monthly Weather Review*, *141*(7), 2526–2544. https://doi.org/10.1175/MWR-D-12-00129.1

Chen, X., Lin, S.-J., & Harris, L. M. (2018). Toward an unstaggered finite-volume dynamical core with a fast Riemann solver: 1-D linearized analysis of dissipation, dispersion, and noise control. *Journal of Advances in Modeling Earth Systems*, *10*(9), 2333–2356. https://doi.org/10.1029/2018MS001361

Chen, C., & Xiao, F. (2008). Shallow water model on cubed-sphere by multi-moment finite volume method. *Journal of Computational Physics*, *227*(10), 5019–5044. https://doi.org/10.1016/j.jcp.2008.01.033

Du, Q., Gunzburger, M. D., & Ju, L. (2003). Voronoi-based finite volume methods, optimal Voronoi meshes, and PDEs on the sphere. *Computer Methods in Applied Mechanics and Engineering*, *192*(35–36), 3933–3957. https://doi.org/10.1016/S0045-7825(03)00394-3

Gao, K., Harris, L., Chen, J.-H., Lin, S.-J., & Hazelton, A. (2019). Improving AGCM hurricane structure with two-way nesting. *Journal of Advances in Modeling Earth Systems*, *11*(1), 278–292. https://doi.org/10.1029/2018MS001359

Giraldo, F., Hesthaven, J., & Warburton, T. (2002). Nodal high-order discontinuous Galerkin methods for the spherical shallow water equations. *Journal of Computational Physics*, *181*(2), 499–525. https://doi.org/10.1006/jcph.2002.7139

Griffies, S. (2004). *Fundamentals of ocean climate models*. Princeton, NJ: Princeton University Press. ISBN: 978-0-691-11892-5.

Guba, O., Taylor, M. A., Ullrich, P. A., Overfelt, J. R., & Levy, M. N. (2014). The spectral element method (SEM) on variable-resolution grids: Evaluating grid sensitivity and resolution-aware numerical viscosity. *Geoscientific Model Development*, *7*(6), 2803–2816. https://doi.org/10.5194/gmd-7-2803-2014

Harris, L. M., & Lin, S.-J. (2013). A two-way nested global-regional dynamical core on the cubed-sphere grid. *Monthly Weather Review*, *141*(1), 283–306. https://doi.org/10.1175/MWR-D-11-00201.1

Harris, L. M., Lin, S.-J., & Tu, C. (2016). High-resolution climate simulations using GFDL HiRAM with a stretched global grid. *Journal of Climate*, *29*(11), 4293–4314. https://doi.org/10.1175/JCLI-D-15-0389.1

Katta, K. K., Nair, R. D., & Kumar, V. (2015). High-order finite-volume transport on the cubed sphere: Comparison between 1D and 2D reconstruction schemes. *Monthly Weather Review*, *143*(7), 2937–2954. https://doi.org/10.1175/MWR-D-13-00176.1

Khamayseh, A., & Mastin, C. (1996). Computational conformal mapping for surface grid generation. *Journal of Computational Physics*, *123*(2), 394–401. http://dx.doi.org/10.1006/jcph.1996.0032

Le Roux, D. Y., Eldred, C., & Taylor, M. A. (2020). Fourier analyses of high-order continuous and discontinuous Galerkin methods. *SIAM Journal on Numerical Analysis*, *58*(3), 1845–1866. https://doi.org/10.1137/19M1289595

Li, C., & Chen, X. (2019). Simulating Nonhydrostatic Atmospheres on Planets (SNAP): Formulation, validation, and application to the Jovian atmosphere. *The Astrophysical Journal: Supplement Series*, *240*(2), 37. https://doi.org/10.3847/1538-4365/aafdaa

Lin, S.-J. (1997). A finite-volume integration method for computing pressure gradient force in general vertical coordinates. *Quarterly Journal of the Royal Meteorological Society*, *123*(542), 1749–1762. https://doi.org/10.1002/qj.49712354214

Lin, S.-J. (2004). A "vertically Lagrangian" finite-volume dynamical core for global models. *Monthly Weather Review*, *132*(10), 2293–2307. https://doi.org/10.1175/1520-0493(2004)132⟨2293:AVLFDC⟩2.0.CO;2

Lin, S.-J., Harris, L., Chen, X., Yao, W., & Chai, J. (2017). Colliding modons: A nonlinear test for the evaluation of global dynamical cores. *Journal of Advances in Modeling Earth Systems*, *9*(7), 2483–2492. https://doi.org/10.1002/2017MS000965

Lin, S.-J., & Rood, R. B. (1996). Multidimensional flux-form semi-Lagrangian transport schemes. *Monthly Weather Review*, *124*(9), 2046–2070. https://doi.org/10.1175/1520-0493(1996)124⟨2046:MFFSLT⟩2.0.CO;2

Marras, S., Kopera, M. A., & Giraldo, F. X. (2015). Simulation of shallow-water jets with a unified element-based continuous/discontinuous Galerkin model with grid flexibility on the sphere. *Quarterly Journal of the Royal Meteorological Society*, *141*(690), 1727–1739. https://doi.org/10.1002/qj.2474

McGregor, J. L. (1996). Semi-Lagrangian advection on conformal-cubic grids. *Monthly Weather Review*, *124*(6), 1311–1322. https://doi.org/10.1175/1520-0493(1996)124⟨1311:SLAOCC⟩2.0.CO;2

Mesinger, F. (1977). Forward-backward scheme, and its use in a limited area model. *Contributions to Atmospheric Physics*, *50*, 200–210.

Müller, A., Kopera, M. A., Marras, S., Wilcox, L. C., Isaac, T., & Giraldo, F. X. (2019). Strong scaling for numerical weather prediction at petascale with the atmospheric model NUMA. *The International Journal of High Performance Computing Applications*, *33*(2), 411–426. https://doi.org/10.1177/1094342018763966

Murakami, H., Vecchi, G. A., Underwood, S., Delworth, T. L., Wittenberg, A. T., Anderson, W. G., & Zeng, F. (2015). Simulation and prediction of category 4 and 5 hurricanes in the high-resolution GFDL HiFLOR coupled climate model. *Journal of Climate*, *28*(23), 9058–9079. https://doi.org/10.1175/JCLI-D-15-0216.1

Nair, R. D., Thomas, S. J., & Loft, R. D. (2005). A discontinuous Galerkin transport scheme on the cubed sphere. *Monthly Weather Review*, *133*(4), 814–828. https://doi.org/10.1175/MWR2890.1

Putman, W. M., & Lin, S.-J. (2007). Finite-volume transport on various cubed-sphere grids. *Journal of Computational Physics*, *227*(1), 55–78. https://doi.org/10.1016/j.jcp.2007.07.022

Rančić, M., Purser, R. J., & Mesinger, F. (1996). A global shallow-water model using an expanded spherical cube: Gnomonic versus conformal coordinates. *Quarterly Journal of the Royal Meteorological Society*, *122*(532), 959–982. https://doi.org/10.1002/qj.49712253209

Ringler, T. D., Heikes, R. P., & Randall, D. A. (2000). Modeling the atmospheric general circulation using a spherical geodesic grid: A new class of dynamical cores. *Monthly Weather Review*, *128*(7), 2471–2490. https://doi.org/10.1175/1520-0493(2000)128⟨2471:MTAGCU⟩2.0.CO;2

Rípodas, P., Gassmann, A., Förstner, J., Majewski, D., Giorgetta, M., Korn, P., & Heinze, T. (2009). Icosahedral shallow water model (ICOSWM): Results of shallow water test cases and sensitivity to model parameters. *Geoscientific Model Development*, *2*(2), 231–251. https://doi.org/10.5194/gmd-2-231-2009

Ronchi, C., Iacono, R., & Paolucci, P. (1996). The "cubed sphere": A new method for the solution of partial differential equations in spherical geometry. *Journal of Computational Physics*, *124*(1), 93–114. https://doi.org/10.1006/jcph.1996.0047

Rossmanith, J. A. (2006). A wave propagation method for hyperbolic systems on the sphere. *Journal of Computational Physics*, *213*(2), 629–658.doi http://dx.doi.org/10.1016/j.jcp.2005.08.027

Sadourny, R. (1972). Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids. *Monthly Weather Review*, *100*, 2136–2144. https://doi.org/10.1175/1520-0493(1972)100⟨0136:CFAOTP⟩2.3.CO;2

Sadourny, R., Arakawa, A., & Mintz, Y. (1968). Integration of the nondivergent barotropic vorticity equation with an icosahedral-hexagonal grid for the sphere. *Monthly Weather Review*, *96*(6), 351–356. https://doi.org/10.1175/1520-0493(1968)096⟨0351:IOTNBV⟩2.0.CO;2

Satoh, M., Stevens, B., Judt, F., Khairoutdinov, M., Lin, S.-J., Putman, W. M., & Düben, P. (2019). Global cloud-resolving models. *Current Climate Change Reports*, *5*(3), 172–184. https://doi.org/10.1007/s40641-019-00131-0

Stevens, B., Satoh, M., Auger, L., Biercamp, J., Bretherton, C. S., Chen, X., & Zhou, L. (2019). DYAMOND: The DYnamics of the Atmospheric general circulation Modeled On Non-hydrostatic Domains. *Progress in Earth and Planetary Science*, *6*(1), 61. https://doi.org/10.1186/s40645-019-0304-z

Taylor, M. A., & Fournier, A. (2010. A compatible and conservative spectral element method on unstructured grids. *Journal of Computational Physics*, *229*(17), 5879–5895. https://doi.org/10.1016/j.jcp.2010.04.008

Taylor, M. A., Guba, O., Steyer, A., Ullrich, P. A., Hall, D. M., & Eldrid, C. (2020). An energy consistent discretization of the nonhydrostatic equations in primitive variables. *Journal of Advances in Modeling Earth Systems*, *12*(1), 1–20. https://doi.org/10.1029/2019MS001783

Tomita, H., Tsugawa, M., Satoh, M., & Goto, K. (2001). Shallow water model on a modified icosahedral geodesic grid by using spring dynamics. *Journal of Computational Physics*, *174*(2), 579–613. https://doi.org/10.1006/jcph.2001.6897

Ullrich, P. A., Jablonowski, C., & van Leer, B. (2010). High-order finite-volume methods for the shallow-water equations on the sphere. *Journal of Computational Physics*, *229*(17), 6104–6134. https://doi.org/10.1016/j.jcp.2010.04.044

Ullrich, P. A., Reynolds, D. R., Guerra, J. E., & Taylor, M. A. (2018). Impact and importance of hyperdiffusion on the spectral element method: A linear dispersion analysis. *Journal of Computational Physics*, *375*, 427–446. https://doi.org/10.1016/j.jcp.2018.06.035

van Leer, B. (1979). Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *Journal of Computational Physics*, *32*(1), 101–136. https://doi.org/10.1016/0021-9991(79)90145-1

Wang, L., Zhang, Y., Li, J., Liu, Z., & Zhou, Y. (2019). Understanding the performance of an unstructured-mesh global shallow water model on kinetic energy spectra and nonlinear vorticity dynamics. *Journal of Meteorological Research*, *33*(6), 1075–1097. https://doi.org/10.1007/s13351-019-9004-2

Williamson, D. L. (1968). Integration of the barotropic vorticity equation on a spherical geodesic grid. *Tellus*, *20*(4), 642–653. https://doi.org/10.1111/j.2153-3490.1968.tb00406.x

Williamson, D. L., Drake, J. B., Hack, J. J., Jakob, R., & Swarztrauber, P. N. (1992). A standard test set for numerical approximations to the shallow water equations in spherical geometry. *Journal of Computational Physics*, *102*(1), 211–224. https://doi.org/10.1016/S0021-9991(05)80016-6

Yang, C., Cao, J., & Cai, X.-C. (2010). A fully implicit domain decomposition algorithm for shallow water equations on the cubed-sphere. *SIAM Journal on Scientific Computing*, *32*(1), 418–438. https://doi.org/10.1137/080727348

Zhang, Y., Li, J., Yu, R., Zhang, S., Liu, Z., Huang, J., & Zhou, Y. (2019). A layer-averaged nonhydrostatic dynamical framework on an unstructured mesh for global and regional atmospheric modeling: Model description, baseline evaluation, and sensitivity exploration. *Journal of Advances in Modeling Earth Systems*, *11*(6), 1685–1714. https://doi.org/10.1029/2018MS001539

Zhou, L., Lin, S.-J., Chen, J.-H., Harris, L. M., Chen, X., & Rees, S. L. (2019). Toward convective-scale prediction within the next generation global prediction system. *Bulletin of the American Meteorological Society*, *100*(7), 1225–1243. https://doi.org/10.1175/BAMS-D-17-0246.1