Numerical Integration of the Shallow-Water Equations on a Twisted Icosahedral Grid. 
Part I: Basic Design and Results of Tests

ROSS HEIKES AND DAVID A. RANDALL

Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

(Manuscript received 23 February 1994, in final form 14 November 1994)

ABSTRACT

The streamfunction–velocity potential form of shallow-water equations, implemented on a spherical geodesic grid, offers an attractive solution to many of the problems associated with fluid-flow simulations in a spherical geometry. Here construction of a new type of spherical geodesic grid is outlined, and discretization of the equations is explained. The model is subjected to the NCAR suite of seven test cases for shallow-water models.

1. Introduction

A number of problems are associated with trying to numerically solve differential equations in spherical geometry. These difficulties, collectively known as the pole problem, can arise from the use of a particular coordinate system and from trying to discretize the surface of the sphere.

Before any numerical considerations are addressed, we run into difficulties in simply specifying a coordinate system. In the following discussion it is useful to distinguish between a true scalar-valued and vector-valued functions. The value of a scalar function, such as temperature, is independent of coordinate system. On the other hand, the individual components of a vector-valued function, such as velocity, obviously differ depending on the coordinate system used. In the spherical coordinate system, the lines of constant longitude converge at the poles, so longitude is multivalued at the poles. This means that the components of the wind are discontinuous at the poles, although the velocity itself is continuous there. True scalars are not subject to such discontinuities. The streamfunction–velocity potential form of the shallow-water equations has the advantage that momentum is expressed in terms of true scalars.

An early alternative to spherical coordinates was to project the equations from the sphere to a plane. The surface of a sphere and that of a plane are not topologically equivalent, however. In other words, there does not exist a one-to-one mapping from every point on the sphere to every point on a finite plane. We can, however, partition the sphere into pieces and project the pieces separately. The equations for each projected piece then get their boundary conditions from the solutions of the other pieces. One example of this approach is the study of Phillips (1957). In a more recent study, Browning et al. (1989) partitions the sphere into hemispheres and maps each to a plane with polar stereographic projections.

One obvious way to discretize the fluid-flow equations expressed in the spherical coordinate system is to use a regular latitude–longitude grid in which the grid intervals (Δθ, Δλ) are constant. A discretization scheme is then straightforward, except for the row of grid points next to the pole, where special considerations are necessary. While the latitudinal distance between grid points can be constant in spherical coordinates, the longitudinal distance decreases rapidly toward the poles. This is unfortunate because the scale of meteorological phenomena does not vary dramatically as a function of latitude. Ideally, then, the average distance between neighboring grid points should not depend on latitude, as it does in a latitude–longitude grid, nor should the distance between grid points in the zonal direction be substantially different from the distance in the meridional direction. Apart from these aesthetic considerations, latitude–longitude grids suffer from computation complications associated with the convergence of the meridians near the poles. Typically, selected variables are filtered near the poles to allow a practical time step without numerical instability (e.g., Arakawa and Lamb 1977). In addition, Lindzen and Fox-Rabinovitz (1989), among others, emphasize the problems that can arise if the horizontal and vertical resolution of a model are not properly matched. In striving to achieve such a match in a global model, it is clearly desirable to maintain uniform horizontal resolution over the sphere. This is achieved by triangularly truncated spectral models (e.g., Baer 1972), but conventional finite-difference models with latitude–lon-
titude grids give horizontal resolution that varies systematically and anisotropically with latitude, even when polar filtering is used. Geodesic grids are attractive in part because they hold the promise of quasi-homogeneous and quasi-isotropic horizontal resolution over the entire sphere.

Researchers have attempted to use grids with increased grid spacing at higher latitudes. Kurihara (1969) proposed a grid in which the number of grid points along latitude circles decreases toward the poles, thus covering the sphere more homogeneously. Kurihara built a model based on the shallow-water equations and tested it with a simulation of a Rossby–Haurwitz wave. Although the Kurihara grid was used at Geophysical Fluid Dynamics Laboratory for a number of years, it has been abandoned now because of numerical problems.

Williamson (1968) and Sadourny et al. (1968) simultaneously introduced a new approach to more isotropically and homogeneously discretize the sphere. Their grids are constructed from spherical triangles that are nearly equal in area and that are equatorially flat. The grid was inspired by Buckminster Fuller’s geodesic dome and is called a spherical geodesic grid. While Williamson and Sadourny worked with finite-difference models, Cullen (1974) developed finite-element models based on spherical geodesic grids. An example of a simple spherical geodesic grid is shown in section 3. Since the grid points are not regularly spaced and do not lie in orthogonal rows and columns, specially designed finite-difference schemes are needed (e.g., Masuda 1969). Williamson (1968) and Sadourny (1968) chose the nondivergent shallow-water equations to test the new grids, while Sadourny and Morel (1969) developed a model for the shallow-water equations on a plane using a triangular mesh. Masuda and Ohnishi (1987) created an elegant geodesic grid model for the shallow-water equations on the sphere. The design of his model is discussed in detail in section 4.

This paper discusses some of the problems associated with fluid flow on a rotating sphere and shows that through the use of a spherical geodesic grid and the streamfunction–velocity potential form of the shallow-water equations, many of these difficulties can be overcome. A companion paper, Heikes and Randall (1995; hereafter referred to as Part II) describe the grid in more detail.

2. The streamfunction–velocity potential form of the shallow-water equations

The shallow-water equations are the simplest form of the equations of motion capable of describing the response of a fluid to both gravitational and rotational accelerations. Their solutions represent many of the types of motion found in the real atmosphere, including slow-moving Rossby waves and fast-moving gravity waves.

The motions of the atmosphere are nearly geostrophic. For this reason, it is useful to partition the wind into rotational and divergent parts; this can be done using Helmholtz’s theorem, which states that any vector field \( \mathbf{V} = (u, v) \) can be written \( \mathbf{V} = \nabla \psi + \nabla \chi \), where \( \nabla \cdot \mathbf{V}_\psi = 0 \) and \( \nabla \times \mathbf{V}_\chi = 0 \). Let the vector field be the horizontal wind. We can define a horizontal streamfunction \( \psi \) to express the rotational part \( \mathbf{k} \times \nabla \psi \) and a velocity potential \( \chi \) to express the divergent part \( \nabla \chi \), that is

\[
v = k \times \nabla \psi + \nabla \chi. \tag{1}
\]

Using (1), we can state the streamfunction–velocity potential (SFVP) form of the shallow-water equations. The time rates of change of absolute vorticity \( \eta \) and divergence \( \delta \) are given by

\[
\frac{\partial \eta}{\partial t} + \nabla \cdot (\eta \nabla \chi) - J(\eta, \psi) = 0 \tag{2}
\]

and

\[
\frac{\partial \delta}{\partial t} - \nabla \cdot (\eta \nabla \psi) - J(\eta, \chi) + \nabla^2 (K + gh) = 0, \tag{3}
\]

respectively, where the Jacobian operator is defined by \( J(A, B) = \mathbf{k} \cdot (\nabla A \times \nabla B) \). In the divergence equation, the kinetic energy \( K \) is given by

\[
K = \frac{1}{2} [\nabla \cdot (\psi \nabla \psi) - \psi \nabla^2 \psi + \nabla \cdot (\chi \nabla \chi) - \chi \nabla^2 \chi] + J(\psi, \chi). \tag{4}
\]

The streamfunction and velocity potential satisfy a pair of elliptic equations,

\[
\nabla^2 \psi = \zeta = \eta - f \quad \text{and} \quad \nabla^2 \chi = \delta, \tag{5}
\]

where \( f \) is the Coriolis parameter.

The third prognostic equation describes the evolution of the height field. Let \( h \) be the height of the free surface and \( h_0 \) be the height of the underlying surface topography. Then \( h^* = h - h_0 \) is the depth of the fluid. With these definitions, the continuity equation can be written as

\[
\frac{\partial}{\partial t} h^* + \nabla \cdot (h^* \nabla \chi) - J(h^*, \psi) = 0. \tag{6}
\]

The SFVP form of the shallow-water equations, given by (2), (3), and (6), has the advantage that all the fields are expressed in terms of true scalars. This avoids the difficulties associated with vector-valued quantities near the poles. The SFVP equations can be written using only three differential operators, namely the flux divergence, the Jacobian, and the Laplacian. Furthermore, Randall (1994) shows that geostrophic adjustment is well simulated using the linearized form of the vorticity and divergence equations on an unstaggered grid.
3. A simple spherical geodesic grid

Next we consider a simple algorithm to construct a spherical geodesic grid on which to approximate the solution. Such a grid can be constructed by starting with an ordinary icosahedron inscribed inside a unit sphere, as shown in Fig. 1a. First each face of the icosahedron is subdivided into four new faces by bisecting the edges. The result of this process is shown in Fig. 1b. Next, the new vertices are projected onto the unit sphere, creating the polyhedron shown in Fig. 1c. Masuda and Ohnishi (1987) continue this process, further partitioning the spherical triangles and projecting onto the sphere. The resulting spherical geodesic grid is not symmetric across the equator, however. In Masuda’s numerical results, initial conditions that are symmetric across the equator slowly evolve to a state that is asymmetric across the equator, presumably as a result of the flow interacting with the grid. This suggests that it is desirable to construct a spherical geodesic grid that is symmetric across the equator. By simply rotating all the faces of Fig. 1c in the Southern Hemisphere through $\pi/5$ rad, we obtain a polyhedron that is symmetric, as shown in Fig. 1d. This polyhedron is called the twisted icosahedron. Grids based on this polyhedron will be called twisted icosahedral grids, or twigs. By recursively subdividing each triangular face and projecting the new vertices onto the unit sphere, we can generate polyhedra that progressively approximate a sphere. Examples are shown in Fig. 2.

We define the vertices of these polyhedra to be the grid points, and the two terms are synonymous in the discussion to follow. All three prognostic variables as well as $\psi$ and $\chi$ are defined at the grid points of this unstaggered grid. We associate a Voronoi cell, defined below, with each grid point. Augenbaum and Peskin (1985) discuss the construction of a Voronoi grid on a sphere. Given a set of $N$ grid points $\{P_1, P_2, \ldots, P_N\}$ on the unit sphere $S$, the Voronoi cell $k$ associated with $P_k$ is defined by

$$\text{cell}_k = \{ |p - P_k| \leq |p - P_l| \quad \forall p \in S \quad \text{and} \quad l \in \{1, 2, \ldots, k-1, k+1, \ldots, N\} \}. \quad (7)$$

where $|x_1 - x_2|$ is the distance between points $x_1, x_2 \in S$ measured along the surface of the sphere. In other words, the Voronoi cell of a grid point consists of all the points on the sphere equidistant or closer to that grid point than any other grid point. Let $P_0$ and $P_1$ be neighboring grid points, as shown in Fig. 3. From the above definition, it follows that a Voronoi cell has the important property that the cell wall shared by $P_0$ and $P_1$ is the perpendicular bisector of the circular arc between the two grid points. The segment connecting $P_0$ and $P_1$ does not bisect the cell wall, however. This point is of some importance and will be expanded upon in Part II. A cell wall is a segment of a great circle. The end points of a wall are the Voronoi corners, the points where three cell walls intersect. The position of a corner is computed using three grid points. Again, consider $P_0$, $P_1$, and $P_2$ in Fig. 3. Consider these points in three-dimensional Cartesian coordinates, that is, $P_0 = (x_0, y_0, z_0)$, $P_1 = (x_1, y_1, z_1)$, and $P_2 = (x_2, y_2, z_2)$. From the definition of a Voronoi cell, the Voronoi corner $C$ is the point on the sphere equidistant from these three grid points:

$$C = \frac{(P_2 - P_0) \times (P_1 - P_0)}{|(P_2 - P_0) \times (P_1 - P_0)|}. \quad (8)$$

Note that the points are ordered in a clockwise fashion to ensure that $C$ lies in the same hemisphere as $P_0, P_1$, and $P_2$. 

![Fig. 3. Properties of Voronoi cells.](image-url)
and $P_3$. Ordering in a counterclockwise fashion gives the point on the opposite side of the sphere from $C$.

Voronoi grids constructed from the polyhedra, like those in Fig. 2, consist of hexagons and 12 pentagons. An example is shown in Fig. 4. Each pentagon corresponds to a vertex of the original icosahedron. It can be shown that the number of cells $N_c$ is related to the number of recursive subdivisions $q$ by

$$N_c = 5(2^{2q+3}) + 2. \tag{9}$$

The grid described above is similar to but not exactly the same as the grid used to obtain the numerical results presented in section 6. In Part II we show that the finite-difference operators discussed below are inconsistent when used in conjunction with a grid constructed using this simple algorithm. For this reason, the cells of the grid used to obtain the numerical results have been moved slightly to improve the accuracy of the finite-difference operators. The construction of this alternative, the tweaked grid is described in detail in Part II. Some basic properties of the untweaked grid are shown in Table 1.

### 4. Finite-difference form of the equations

Our derivation of the finite-difference operators follows exactly the methods of Masuda and Ohnishi (1987). Consider arbitrary scalars $\alpha$ and $\beta$. Using Gauss' theorem, we can write the three operators we need, namely the Jacobian, the flux divergence, and the Laplacian, as follows:

$$\int_A J(\alpha, \beta) dA = \oint_C \alpha \frac{\partial \beta}{\partial s} ds, \tag{10}$$

$$\int_A \nabla \cdot (\alpha \nabla \beta) dA = \oint_C \alpha \frac{\partial \beta}{\partial n} ds, \tag{11}$$

$$\int_A \nabla^2 \alpha dA = \oint_C \frac{\partial^2 \alpha}{\partial n^2} ds. \tag{12}$$

The finite-difference approach we use approximates the line integrals on the right-hand sides of (10), (11), and (12). Referring to Fig. 3, let $\alpha_i$ and $\beta_i$ be the values of functions $\alpha$ and $\beta$ evaluated at the cell centers $P_i$. In the finite-difference form of the Jacobian operator, we need to interpolate to the corners of the cells. Let

$$b_i = \frac{1}{3} (\beta_0 + \beta_{i-1} + \beta_i) \tag{13}$$

be an approximation for $\beta$ at the cell corner $C_i$, where $C_i$ is the Voronoi corner common to the three cells with centers $P_0$, $P_{i-1}$, and $P_i$. It can be shown that this is the only choice of weights that conserves energy and potential enstrophy in the case of nondivergent flow. Let $A_c$ be the area of a cell. If the cell is small, $J(\alpha, \beta)$ can be considered constant across the cell, and we can approximate the left-hand side of (10) by

$$\int_A J(\alpha, \beta) dA \approx J(\alpha, \beta)|_{P_0} \int_A dA$$

$$= J(\alpha, \beta)|_{P_0} A_c, \tag{14}$$

where $J(\alpha, \beta)|_{P_0}$ is the Jacobian evaluated at $P_0$. The line integral on the right-hand side of (10) is replaced with a summation, so we approximate the Jacobian by

$$J(\alpha, \beta)|_{P_0} \approx \frac{1}{A_c} \sum_{i=1}^{N} \left( \frac{\alpha_0 + \alpha_i}{2} \right) \left( \frac{b_{i+1} - b_i}{l_i} \right) l_i, \tag{15}$$

where $N = 5$ for pentagons and $N = 6$ for hexagons. In (15), the expression

$$\frac{1}{2} (\alpha_0 + \alpha_i) \tag{16}$$

<table>
<thead>
<tr>
<th>$q$</th>
<th>$N_c$</th>
<th>$N_c$ along equator</th>
<th>Average cell area (km$^2$)</th>
<th>Ratio of smallest cell to largest cell</th>
<th>Average distance between cell centers (km)</th>
<th>Ratio of smallest distance to largest distance between cell centers</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>42</td>
<td>10</td>
<td>$1.214 \times 10^4$</td>
<td>0.885</td>
<td>3755.5</td>
<td>0.881</td>
</tr>
<tr>
<td>1</td>
<td>162</td>
<td>20</td>
<td>$3.149 \times 10^4$</td>
<td>0.774</td>
<td>1912.8</td>
<td>0.848</td>
</tr>
<tr>
<td>2</td>
<td>642</td>
<td>40</td>
<td>$7.946 \times 10^4$</td>
<td>0.763</td>
<td>961.0</td>
<td>0.840</td>
</tr>
<tr>
<td>3</td>
<td>2562</td>
<td>80</td>
<td>$1.991 \times 10^4$</td>
<td>0.742</td>
<td>481.1</td>
<td>0.838</td>
</tr>
<tr>
<td>4</td>
<td>10 242</td>
<td>160</td>
<td>$4.980 \times 10^4$</td>
<td>0.736</td>
<td>240.6</td>
<td>0.837</td>
</tr>
<tr>
<td>5</td>
<td>40 962</td>
<td>320</td>
<td>$1.245 \times 10^4$</td>
<td>0.733</td>
<td>120.3</td>
<td>0.837</td>
</tr>
</tbody>
</table>
approximates $\alpha$ evaluated along the cell wall. This is reasonable since each wall bisects the arc between cell centers. Let $l_i$ be the length of the arc between corners $C_i$ and $C_{i+1}$. The expression

$$\frac{1}{l_i} (b_{i+1} - b_i) \quad (17)$$

approximates the tangential derivative along the cell wall. If the grid consists of perfect hexagons, then the arc between cell centers is also the perpendicular bisector of the cell wall. In this case, the approximation of the tangential derivative is second-order accurate at the same point where $\alpha$ is approximated along the cell wall. The final $l_i$ approximates the infinitesimal element of curve $ds$. Equation (15) can be rewritten as

$$J(\alpha, \beta)|_{P_0} \approx \frac{1}{6A_c} \sum_{i=1}^{N} (\alpha_i + \alpha_i)(\beta_{i+1} - \beta_{i-1}). \quad (18)$$

Similarly, we can approximate flux-divergence and Laplacian operators. Let $L_i$ be the length of the arc between cell centers $P_0$ and $P_i$. Using (11), the flux-divergence can be approximated by

$$\nabla \cdot (\alpha \nabla \beta)|_{P_0} \approx \frac{1}{2A_c} \sum_{i=1}^{N} L_i (\alpha_i + \alpha_i)(\beta_i - \beta_0). \quad (19)$$

where the expression

$$\frac{1}{L_i} (\beta_i - \beta_0) \quad (20)$$

approximates the derivative normal to the curve. Finally, using (12), the Laplacian can be approximated by

$$\nabla^2 \alpha|_{P_0} \approx \frac{1}{A_c} \sum_{i=1}^{N} L_i (\alpha_i - \alpha_0). \quad (21)$$

5. Time stepping

The model uses the third-order Adams–Bashforth scheme discussed by Durran (1991). This is an explicit scheme requiring only one function evaluation per time step. It has the advantages (over the more common leapfrog time-stepping scheme) of a reduced phase-speed error and a damped computational mode. Its major disadvantage is an increased storage requirement. This is not a serious limitation with current computers, however.

Consider an ordinary differential equation of the form

$$\frac{d\xi}{dt} = F(\xi). \quad (22)$$

The third-order Adams–Bashforth scheme approximates (22) with

$$\frac{\xi^{n+1} - \xi^n}{\Delta t} = \frac{1}{12} [23F(\xi^n) - 16F(\xi^{n-1}) + 5F(\xi^{n-2})], \quad (23)$$

where $\xi^n$ is the numerical approximation to $\xi(n\Delta t)$. We see from (23) that $\xi^n$ (the initial condition), $\xi^1$, and $\xi^2$ are required to compute $\xi^3$. In the model $\xi^1$ and $\xi^2$ are computed using two successive Matsuno time steps.

6. Multigrid elliptic equation solver

One computational cycle of the model can be divided into two steps. First, the model time steps the prognostic (2), (3), and (6) for new values of absolute vorticity, divergence, and geopotential. Next, the new absolute vorticity and divergence are used to solve (7) for a new streamfunction and velocity potential, and then the cycle repeats. Efficient numeric implementation of this cycle hinges on being able to solve the elliptic equations quickly. The recursive numeric nature of the grid lends itself well to multigrid methods.

Multigrid algorithms are discussed by Fulton et al. (1986) and Brandt (1977). Because the twigs are not composed of rectangular elements as is the Cartesian grid, interpolating from a coarse to a fine grid is a little complicated. Figure 5 shows the relation between hexagonal cells for two different resolutions. The residual correction at cells common to both grids can be transferred simply from the coarse grid to the fine grid. The remaining grid points of the fine grid all lie on a cell wall of the coarse grid. Figure 5b shows six cells on the coarse grid, their walls shown with dashed lines, and one cell on the fine grid, its walls shown with thin, solid lines. This stencil is used to associate every grid point on the fine grid with six grid points on the coarse grid. For the cells $P_0$ that have a pentagonal neighbor, the stencil is the same except that the point labeled $P_1$ in Fig. 5b is associated with a pentagonal cell. We never have to interpolate to the pentagons themselves since they belong to the coarsest grid and therefore belong to all grid resolutions. Each point has spherical coordinates $P_i = (\theta_i, \lambda_i)$ for $0 \leq i \leq 6$. Suppose we

![Image](image-url)
assign arbitrary values \( f_1, f_2, \ldots, f_6 \) at grid points \( P_1, P_2, \ldots, P_6 \). Then there exists a quadratic surface determined by constants \( c_1, c_2, \ldots, c_6 \) such that

\[
\begin{align*}
&c_1 \theta_1^2 + c_2 \theta_1 \lambda_1 + c_3 \lambda_1^2 + c_4 \theta_1 + c_5 \lambda_1 + c_6 = f_1 \\
&c_1 \theta_2^2 + c_2 \theta_2 \lambda_2 + c_3 \lambda_2^2 + c_4 \theta_2 + c_5 \lambda_2 + c_6 = f_2 \\
&\quad \vdots \\
&c_1 \theta_6^2 + c_2 \theta_6 \lambda_6 + c_3 \lambda_6^2 + c_4 \theta_6 + c_5 \lambda_6 + c_6 = f_6.
\end{align*}
\]

(24)

The corresponding matrix is known as the Vandermonde matrix; see Press et al. (1988). Solving the linear system, the constants have the form \( c_i = c_i(\theta_1, \ldots, \theta_6, \lambda_1, \ldots, \lambda_6, f_1, \ldots, f_6) \) for \( 1 \leq i \leq 6 \). At the fine grid point \( P_0 \), we can write

\[
c_i \theta_0^2 + c_2 \theta_0 \lambda_0 + c_3 \lambda_0^2 + c_4 \theta_0 + c_5 \lambda_0 + c_6 = f_0.
\]

(25)

Substituting \( c_1, c_2, \ldots, c_6 \) and rearranging gives

\[
w_1 f_1 + w_2 f_2 + w_3 f_3 + w_4 f_4 + w_5 f_5 + w_6 f_6 = f_0,
\]

(26)

where \( w_i = w_i(\theta_1, \ldots, \theta_6, \lambda_1, \ldots, \lambda_6) \).

We would like the weights to depend only on the relative distances and angles between each set of seven points, but because of the use of spherical coordinates, the weights, if computed as shown above, differ depending on the position of the points on the sphere. Therefore, before solving the system in (24), each group of seven points is translated, preserving their relative positions, so that \( P_0 = (0, 0) \). Then \( P_1, P_2, \ldots, P_6 \) are drawn closer to \( P_0 \) by replacing each point \( P_i, i = 1, \ldots, 6 \), with the midpoint of the arc between \( P_0 \) and \( P_i \). This process repeated until all the points are nearly coplanar. A Mathematika program (Wolfram 1988) computes the weights \( w_1, w_2, \ldots, w_6 \) in advance and stores them in a file that is read by the model at the beginning of execution. We can use the arbitrary high numerical precision available in Mathematika to overcome the ill conditioning associated with the Vandermone matrix.

Next we need to perform relaxation on the grid. Using (21) and referring to Fig. 3, we can transform the continuous problem \( \nabla^2 U = f \) to 3, we can transform the continuous problem \( \nabla^2 U = f \) to

\[
\frac{1}{A_i} \sum_{i=1}^{L_i} \left( u_i - u_0 \right) = f_0 \quad \forall \text{cells},
\]

(27)

where \( f_0 \) is evaluated at \( P_0 \). Solving (27) for \( u_0 \), we get

\[
u_{0,\text{new}} = \left( \sum_{i=1}^{N} \frac{L_i}{L} \right)^{-1} \left( \sum_{i=1}^{N} \frac{l_i u_{i,\text{old}}}{L_i} - A_i f_0 \right).
\]

(28)

We use the \( \psi \) and \( \chi \) from the previous time step as very good initial guesses for the new \( \psi \) and \( \chi \). In practice only one multigrid V-cycle (see Fulton et al. 1986) is necessary to compute a new solution for Eq. (5) on each time step.

7. Numerical results

a. Introduction

The model described in sections 4 to 6 has been subjected to the suite of seven test cases proposed by Williamson et al. (1992). For purposes of comparing height field errors, some of the test cases were also run using the National Center for Atmospheric Research (NCAR) spectral transform model at T42 resolution (see Hack and Jakob 1992) and the Arakawa and Lamb (1981; hereafter AL) shallow-water model with 4° latitude by 5° longitude resolution. Results from the NCAR spectral transform model are further described in Jakob et al. (1993).

The physical parameters of the earth are the same as those used by Williamson et al. (1992) [their Eqs. (72)–(74)]. Also, the norms used to measure differences between solutions are the same as those used by Williamson et al. (1992) [their Eqs. (81)–(84)].

In the following test results, twig02562 refers to tests performed on a grid with 2562 cells, corresponding to \( q = 3 \) in (9). This is the grid shown in Fig. 3b. Similarly, twig10242 refers to test results obtained using the grid with 10 242 cells, corresponding to \( n = 4 \).

For test cases 3 and 5 we measure the model’s performance in part by recording how well it conserves certain global invariants. Following Williamson et al. (1992), the quantities we check include total energy \( \xi = 0.5[h^\ast v \cdot v + g(h^\ast - h^2)] - E_r \) m² s⁻², where \( E_r \) is the potential energy of the “reference state” in which the surface of the water is flat and potential enstrophy \( \zeta = \eta^2(2h^\ast)^{-1} \) m⁻² s⁻². For each time step, we compute and record the normalized integral

\[
I[\xi_i(t)] = \frac{I[\xi_i(\theta, \lambda, t)] - I[\xi_i(\theta, \lambda, 0)]}{I[\xi_i(\theta, \lambda, 0)]}
\]

(29)

for \( i = 1, 2 \), where \( I[\xi] \) is the global integral of \( \xi \). It should be noted that the model conserves only mass in case of divergent flow and energy and enstrophy in the case of nondivergent flow.

b. Test case 1: Advection of a cosine-shaped bell

This test case differs from the remaining six in that the winds are prescribed rather than predicted. That is, instead of time stepping the wind equations, the initial wind field remains constant with time. Since the initial wind field is nondivergent, the height equation reduces to the advection equation for \( h^\ast \) with constant coefficients, and the initial height field is simply advected along by the wind. In other words, \( \nabla \cdot v = 0 \), and \( v \) is constant in time, so the mass conservation equation

\[
\frac{\partial}{\partial t} h^\ast + \nabla \cdot (v h^\ast) = 0
\]

(30)

reduces to the linear advection equation

\[
\frac{\partial}{\partial t} h^\ast + v \cdot \nabla h^\ast = 0.
\]

(31)
In the absence of computational dispersion, the height field should maintain its initial shape for all time. The advecting wind is a jet along a great circle, and the initial height field is a cosine-shaped bump of fluid. We define the parameter $\alpha$ to be the angle between the axis of solid-body rotation and the polar axis of the numerical grid. In testing the models, we used $\alpha = 0.0, 0.05, \pi/2 - 0.05,$ and $\pi/2$. These choices allow comparison of advecting the bell along the equator and advecting it over the pole. Only results using $\alpha = 0.0$ and $\alpha = \pi/2$ are shown here. The initial conditions are given in Eqs. (75)–(80) in Williamson et al. (1992).

From Table 1, the grid with 2562 cells has 80 cells along the equator. The average distance between grid points is approximately

$$\frac{2\pi a}{80} = \Delta x \approx 500.4 \text{ km.} \quad (32)$$

We set the Courant number to $\mu = 0.5$. The speed of the advecting current is about $c \approx 38.6 \text{ m s}^{-1}$, so the CFL condition requires

$$\Delta t \leq \frac{\mu \Delta x}{c} \approx 6480 \text{ s.} \quad (33)$$

This is the $\Delta t$ used in twig02562 tests, and $\Delta t = 3240 \text{ s}$ is used in the twig10242 tests.

Because of computational dispersion, as the bell is advected, it tends to flatten toward the upwind side and to steepen on the downwind side, and regions of negative height form behind the bell. The phase speed of the bell is underestimated. The computational dispersion can be clearly seen in Fig. 6. Because the spectral transform model suffers no dispersion error for linear advection, its results, although not shown here, are essentially perfect. This test is somewhat misleading, however, because spectral models do produce dispersion errors for advection by more realistic nonuniform currents.

All four angles were run for both the AL and twig models, and the height fields, to the eye, are nearly indistinguishable for all four angles. The height field errors are computed using Eqs. (81)–(84) of Williamson et al. (1992) and are shown in Fig. 7. The AL model and the twig02562 model look pretty comparable. With $\alpha = 0.0$ the bell is advected along the equator. The twig02562 grid has 80 grid points along the equator, while the AL model at this resolution has 72 grid points along the equator. These two models have comparable resolution for this value of $\alpha$. We can see a slight glitch in the infinity norm in Fig. 7d as the bell passes over the poles.

c. Test case 2: Global steady-state nonlinear geostrophic flow

This test case represents a steady-state solution of the shallow-water equations. To test for possible corruption of the numerical solution resulting from interaction with the grid, the Coriolis force is rotated, making it a function of latitude, that is,

$$f = 2\Omega(-\cos \lambda \cos \theta \sin \alpha + \sin \theta \cos \alpha). \quad (34)$$

Clearly, (34) reduces to the usual Coriolis parameter when $\alpha = 0$. If the computational grid is considered fixed in space, then by changing the Coriolis parameter we are, in effect, changing the direction of the axis of the planet’s rotation so that it is no longer coincident with the axis
through the poles of the grid. In this way, flow can be directed over the poles of the grid. We will direct the flow to be parallel and nearly parallel to the equator of the grid and nearly directly over and directly over the pole. That is, as in test case 1, the parameter \( \alpha \) in (34) is set \( \alpha = 0.0, \alpha = 0.05, \alpha = \pi/2 - 0.05, \) and \( \alpha = \pi/2 \). As the models run, the initial fields should remain constant. The initial conditions are given in Eqs. (90)–(96) in Williamson et al. (1992). For this test case we set the maximum depth of the water \( h_0 = 2998 \) m, and the maximum speed of the current is \( u_0 = 38.6 \) m s\(^{-1}\).

Even though the solution is nondivergent, errors tend to accumulate in the numerical solution of \( \nabla^2 \chi = \delta \). The spurious gravity waves generated in this way limit the length of a time step and create noise in the solution. Somewhat arbitrarily we pick \( \Delta t = 600 \) s for the AL model and the twig02562 model and \( \Delta t = 300 \) s for the twig10242 model. The Arakawa–Lamb model filters the gradient of total energy and the mass fluxes near the poles.

Height errors for \( \alpha = 0 \) in the AL, twig02562, and twig10242 models and \( \alpha = \pi/2 \) in the twig10242 model are shown in Fig. 8. There are no noticeable differences between the results for \( \alpha = 0.0 \) and \( \alpha = \pi/2 \), indicating that the flow is hardly affected by the poles of the grid. A triaxially truncated spectral model would produce no error at all for this case, except for that due to round-off. Although not shown, the NCAR model’s error is indeed several orders of magnitude smaller than that of either finite-difference model.

Figure 9 shows the true height field minus the simulated height field after 5 days for the case with \( \alpha = 0.05 \). Clearly, there is a weak wavenumber 5 periodicity to the error, indicating that the flow is slightly distorted by the grid.

d. Test case 3: Steady-state nonlinear zonal geostrophic flow with compact support

This test case was designed by Browning et al. (1989) to mimic the real atmosphere. The wind is purely zonal, as in test case 2, but it is nonzero only in a range of latitudes. This choice is modeled after the observed zonal component of the wind at about 10 km. Again, as the integration progresses, the solution
should remain constant. The initial conditions are given in Eqs. (101)–(115) in Williamson et al. (1992).

Although nondivergent, as in test case 2, the numerical computation of $\nabla^2 \chi = \delta$ limits the time step. Again, we pick $\Delta t = 600$ s for the AL model and the twig02562 model, and $\Delta t = 300$ s for the twig10242 model.

Height error plots for $\alpha = 0$ in the AL and twig models are shown in Fig. 10. Again, a spectral transform model should produce no error, and although the results are not shown, the spectral model is far superior. The twig models perform a little better than the AL model. The twig10242 results are quite a bit better than those of twig02562. There appears to be no difference when the flow is directed over the pole.

The normalized conservation plots for total energy and potential enstrophy are shown in Fig. 11. These values are computed using (29). These quantities are better conserved with higher resolution. It can be shown this finite-difference scheme will exactly conserve energy and enstrophy in the case of purely nondivergent flow, and they are conserved pretty well here.

\[ e. \] **Test case 4: Forced nonlinear system with a translating low**

There are few examples of analytic solutions to the unsteady nonlinear equations. However, by including forcing terms in the shallow-water equations, we can test the model in an unsteady nonlinear simulation for which we know an analytic solution. In this case, introduced by Browning et al. (1989), the forcing terms are designed to simulate a short-wave trough embedded in a midlevel tropospheric jet. In our simulation the jet has a maximum speed of $u_0 = 20$ m s$^{-1}$.

The shallow-water equations written in advective form in spherical coordinates with forcing terms $F_u$, $F_v$, and $F_h$ are given by

\[
\frac{du}{dt} - \frac{uv \tan \theta}{a} - \frac{g}{a \cos \theta} \frac{\partial h}{\partial \lambda} - f v = F_u, \tag{35}
\]

\[
\frac{dv}{dt} + \frac{uu \tan \theta}{a} + \frac{g}{a \partial \theta} \frac{\partial h}{\partial \theta} + f u = F_v, \tag{36}
\]

\[
\frac{dh}{dt} + \frac{h}{a \cos \theta} \left[ \frac{\partial u}{\partial \lambda} + \frac{\partial}{\partial \theta} (v \cos \theta) \right] = F_h. \tag{37}
\]
Fig. 9. Test case 2. True height minus twig10242 height after 5 days with $\alpha = 0.05$; 0.4-m contour interval.

Fig. 10. Test case 3. Height field errors. (a) Arakawa–Lamb with $72 \times 44$ resolution, (b) twig02562, (c) twig10242 model, $\alpha = 0.0$, (d) twig10242, $\alpha = \pi/2$. 
where we have assumed \( h = h^* \). The forcing functions are given in Eqs. (120)–(129) in Williamson et al. (1992). A Mathematica program was used to symbolically perform the calculus in Eqs. (120)–(129) in Williamson and to produce Fortran code for \( F_h, F_c \), and \( F_s \).

The initial height is shown in Fig. 12a. The maximum zonal velocity \( u_0 = 20 \) m s\(^{-1}\) advects the region of low pressure
\[
2\pi \frac{u_0}{2\pi a \cos \theta_0} \approx 22^\circ \text{day}^{-1},
\] (38)
where \( \theta_0 = \pi/4 \) is the latitude of maximum zonal velocity. It should be noted that we modified Eq. (129) in the Williamson paper to
\[
C = \sin \theta_0 \sin \theta + \cos \theta_0 \cos \theta \cos \left( \lambda - \frac{u_0}{a \cos \theta_0} t - \lambda_0 \right).
\] (39)
As the simulation evolves, the region of low pressure moves toward the east, maintaining its initial shape.

Figure 12b shows the true solution minus the numerical solution after 5 simulated days. It appears most of the difference is due to phase error between the two solutions. The model does not move the region of low pressure fast enough.

f. Test case 5: Zonal flow over an isolated mountain

This is the only test case in the suite with topography. Zonal flow is directed over a mountain, a 2000-m-tall hemispherical bump centered at 30°N, 90°W, which on earth would be near New Orleans. Takacs (1988) first used this set of initial conditions to study an a posteriori method for conservation of integral invariants. The initial conditions are given in Eqs. (90)–(95) and Eq. (134) in Williamson et al. (1992). However, now the maximum depth of the water is \( h_0 = 5960 \) m and the maximum speed of the current is \( u_0 = 20 \) m s\(^{-1}\).

We used \( \Delta t = 300 \) s for the twig02562 and AL models, and \( \Delta t = 150 \) s for the twig10242 model. The spectral transform model was run with a time step of \( \Delta t = 600 \) s. The semi-implicit time stepping used in the spectral model allows for a larger time step.
Fig. 12. Test case 4. (a) Initial height field: $u_0 = 20$ m s$^{-1}$; 25-m contour interval. (b) True height field minus twig10242 height field after 5 days; 2-m contour interval.

Fig. 13. Test case 5. Height field errors. High-resolution spectral transform model minus (a) Arakawa–Lamb model with $72 \times 44$ resolution, (b) spectral transform T42, (c) twig02562, (d) twig10242.
This test case, unlike those discussed above, does not have an analytic solution with which to make comparisons. Instead, the numerical results are compared to those of the NCAR spectral shallow-water model run with T213 resolution. The results of that model, truncated to T106 resolution and stored in a netCDF file called REF0114.cdf, are available from NCAR.

The height field errors for AL, the spectral transform model with T42 resolution, twig02562, and twig10242 are shown in Fig. 13. The initial infinity norm in each of these plots is not zero. This can be attributed to the spectral model's inability to resolve the sharp edges near the base of the bell. This makes the infinity-norm suspect for both finite-difference models. The AL and twig02562 models look similar, and the spectral transform model and twig10242 look similar. The height field after 15 simulated days for twig10242 resolution is shown in Fig. 14. The difference plots for the AL, twig02562, and twig10242 are shown in Fig. 15. There appears to be phase error in the wave train in the Southern Hemisphere.

The normalized conservation plots for total energy and potential enstrophy are shown in Fig. 11. These values are computed using (29). Clearly the total energy is not conserved well. The finite-difference operators do not exactly conserve energy in the exchange between kinetic and potential energy. Note the improvement in energy conservation with increasing resolution in test case 3, which is more or less nongradient. We are currently working to solve this problem following methods similar to those of Arakawa and Lamb (1981).

g. Test case 6: Rossby–Haurwitz wave

Phillips (1959) devised this set of initial conditions, and it has been used by many investigators since. It was used by Masuda and Ohnishi (1987), although with different parameters, so at least qualitatively we can compare our results with Masuda's results. The initial conditions for this test case are given in Eqs. (141)–(149) in Williamson et al. (1992). In particular the zonal pattern has wavenumber 4.

Again, since there is no analytic solution to the divergent shallow-water equations for these initial conditions, we will compare our results with those of the NCAR spectral model run at T213. In this case the daily spectral coefficients truncated to T106 are available from NCAR in a netCDF file called REF0092.cdf.

The same time steps used in test case 5 are used here. The spectral model was run without diffusion. The height field errors for this test case are shown in Fig. 16. The Arakawa–Lamb model appears to lie between the twig02562 and twig10242 models. Interestingly, the error of the spectral transform T42 model periodically improves unlike the other models in which the errors monotonically worsen. Figure 17 shows the spectral transform T213 height minus the twig10242 height field after 14 days. There is a large error with zonal wavenumber 4 since the pattern moves slower in the twig10242 model than in the spectral transform model. Also, we can see that the initial pattern that starts with a zonal wavenumber 4 structure has changed to a wavenumber 1 pattern at higher latitudes. The same thing occurs in the results of Masuda and Ohnishi (1987) results. However, while Masuda's fields develop wavenumber 1 patterns that are antisymmetric across the equator, the twig height field remains symmetric across the equator because of the symmetry of the grid. Lorenz (1972) and Hoskins and Hollingsworth (1973) discuss the conditions necessary for the instability of a Rossby–Haurwitz wave for the case of nongradient shallow water on a β plane.

![Fig. 14. Test case 5. Height field for the twig10242 model after 15 simulated days; 50-m contour interval.](image)
Fig. 15. Test case 5. Height field differences after 15 simulated days. High-resolution spectral transform model minus (a) Arakawa–Lamb model 72 × 44, 10-m contour interval; (b) twig02562 model, 10-m contour interval; (c) twig10242 model, 5-m contour interval.
Fig. 16. Test case 6. Height field errors. High-resolution spectral transform model minus (a) Arakawa–Lamb model with 72 × 44 resolution, (b) spectral transform T42, (c) twig02562, (d) twig10242.

Fig. 17. Test case 6. Height field difference after 14 days. Spectral transform minus twig10242; 20-m contour interval.
h. Test case 7: Analyzed 500-mb height and wind field initial conditions

The last test case initializes the model with observed 500-mb height and wind fields from 0000 UTC 21 December 1978. This set of initial conditions produces strong flow over the North Pole and is therefore a good test of the model's performance at the poles. Ritchie (1988) used this case to test a semi-Lagrangian model. The initial conditions are contained in a netCDF file called REFO077.cdf. These data are truncated at T106 and have been processed with nonlinear normal-mode initialization to prevent spurious gravity waves from contaminating the results. The mean height field has been set to 10 km.

The same time steps used in test case 5 are used here, and again the spectral model was run without diffusion. The height field error plots are shown in Fig. 18. The results for the twig02562 grid for this test case were quite dissatisfying. The poor performance can be attributed to the fact that this resolution has only 176 grid points north of 60°N. This seems to be insufficient to resolve the structures in the region of interest. In contrast, the Arakawa–Lamb model with $72 \times 44$ resolution has 576 grid points north of 60°N. Here we show the results from the twig10242 and twig40962. Figure 19 can be used to compare the spectral transform model and the twig10242 model. The region of high pressure located near the pole along the prime meridian is moving toward the northeast. In the twig10242 model, the high has not moved as far to the east as it has in the spectral transform model. Also, the region of low pressure along the 90°E meridian is much deeper in the twig10242 than it is in the spectral transform model. These factors combine to produce jets in different directions in the two models. The jet in the twig model is directed roughly parallel to the 45°W meridian, while the jet in the spectral transform model is parallel to the 90°W meridian. This discrepancy causes large errors near the pole as shown in Fig. 20. Figure 20 also shows that the simulation is improved by increasing the resolution.

8. Performance results

To test the computational performance of the model, test case 5 was run for 3000 time steps, with grid resolutions of 642, 2562, 10 242, and 40 962 cells. Efforts

---

**Fig. 18.** Test case 7. Height field errors. High-resolution spectral transform model minus (a) Arakawa–Lamb model with $72 \times 44$ resolution, (b) spectral transform T42, (c) twig10242, (d) twig40962.
were taken to minimize I/O and to perform only those calculations necessary to time step the equations. The model was run on one processor of a CRAY Y-MP C90, which has a peak speed of 750 Mflops, and the performance results were obtained using the CRAY utility perfview. Figure 21 shows the time per cell per time step for the twig model and the Arakawa–Lamb model. The twig model was run at the four resolutions listed above, and the AL model was run with three grid resolutions $64 \times 32$, $128 \times 64$, and $256 \times 128$. The time required per cell per time step decreases as the vector length increases with increasing resolution. The AL model is about three times faster than the twig model. In a GCM with full physics this would correspond to a factor of 2 slow down. We believe that a parallel version of the twig model will be considerably more competitive, relative to a parallel version of the AL model.

9. Conclusions

We have explained the basic design of a new shallow-water model for spherical geometry, based on a
spherical geodesic grid derived from a twisted icosahedron. The grid is symmetrical across the equator. The grid actually used is slightly modified for reasons discussed in Part II. Following the approach of Masuda and Ohnishi (1987), the model uses the streamfunction–velocity potential form of the shallow-water equations.

The model has been tested using the suite of seven test cases devised by Williamson (1992). The results are compared with exact solutions where these are known and also with the NCAR spectral transform shallow-water model and with a shallow-water model based on the potential enstrophy–conserving scheme of Arakawa and Lamb (1981). The results are encouraging. The comparisons between the AL and twig models in Fig. 8 and Fig. 10 show the twig model to be more accurate. In Fig. 13 and Fig. 16 we see the accuracy of the AL model and the T42 spectral model lies between the twig02562 model and the twig10242.

We are undertaking further development of this model. Planned improvements include modifying the finite-difference schemes so as to obtain formal guarantees for conservation (in appropriate limits) of potential vorticity, potential enstrophy, and total energy; addition of a tracer equation; and enhancement of the computational performance. We also plan to port the model to a massively parallel architecture; in fact, the code has been written with this in mind. Finally, we plan to construct a multilayer version and test it using the benchmarks devised by Held and Suarez (1994).

Acknowledgments. Support for this research was provided by the U.S. Department of Energy’s CHAMMP Program under Grant DE-FG02-91ER61218 to Colorado State University. Computing support was provided by the National Energy Research Supercomputer Center at Lawrence Livermore National Laboratory.

REFERENCES


