Modeling the Atmospheric General Circulation Using a Spherical Geodesic Grid: A New Class of Dynamical Cores

Todd D. Ringler, Ross P. Heikes, and David A. Randall

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

ABSTRACT

This paper documents the development and testing of a new type of atmospheric dynamical core. The model solves the vorticity and divergence equations in place of the momentum equation. The model is discretized in the horizontal using a geodesic grid that is nearly uniform over the entire globe. The geodesic grid is formed by recursively bisecting the triangular faces of a regular icosahedron and projecting those new vertices onto the surface of the sphere. All of the analytic horizontal operators are reduced to line integrals, which are numerically evaluated with second-order accuracy. In the vertical direction the model can use a variety of coordinate systems, including a generalized sigma coordinate that is attached to the top of the boundary layer. Terms related to gravity wave propagation are isolated and an efficient semi-implicit time-stepping scheme is implemented. Since this model combines many of the positive attributes of both spectral models and conventional finite-difference models into a single dynamical core, it represents a distinctively new approach to modeling the atmosphere's general circulation.

The model is tested using the idealized forcing proposed by Held and Suarez. Results are presented for simulations using 2562 polygons (approximately 4.5° × 4.5°) and using 10 242 polygons (approximately 2.25° × 2.25°). The results are compared to those obtained with spectral model simulations truncated at T30 and T63. In terms of first and second moments of state variables such as the zonal wind, meridional wind, and temperature, the geodesic grid model results using 2562 polygons are comparable to those of a spectral model truncated at slightly less than T30, while a simulation with 10 242 polygons is comparable to a spectral model simulation truncated at slightly less than T63.

In order to further demonstrate the viability of this modeling approach, preliminary results obtained from a full-physics general circulation model that uses this dynamical core are presented. The dominant features of the DJF climate are captured in the full-physics simulation.

In terms of computational efficiency, the geodesic grid model is somewhat slower than the spectral model used for comparison. Model timings completed on an SGI Origin 2000 indicate that the geodesic grid model with 10 242 polygons is 20% slower than the spectral model truncated at T63. The geodesic grid model is more competitive at higher resolution than at lower resolution, so further optimization and future trends toward higher resolution should benefit the geodesic grid model.

1. Introduction

All atmospheric general circulation models (AGCMs) numerically solve a set of equations that describe the evolution of the general circulation and thermal state. The component of an AGCM that describes the general circulation is now commonly called the dynamical core. In the context of climate modeling, the atmospheric dynamical core has a strong influence on physical processes within the atmosphere, such as cloud processes and radiation. In addition, the dynamical core strongly influences other climate subsystems, such as ocean, sea ice, and land surface. This emphasizes the importance of developing a dynamical core that can accurately represent the dynamical processes of the general circulation.

Several families of dynamical cores have been developed. Since virtually all dynamical cores use a similar set of governing equations, they are generally distinguished by the numerical methods used to solve the equations. The two largest families of dynamical cores are finite-difference methods and spectral methods. The birth of finite-difference models of the atmosphere's circulation came with Richardson's (1922) pioneering attempt to predict the weather. In the following decades, numerical methods were developed and simple numerical models of the atmosphere were successful in capturing some gross features of the general circulation (e.g., Charney et al. 1950). A paradigm shift in the modeling of the atmospheric circulation came with the discovery of the fast Fourier transform (FFT) by Cooley

Corresponding author address: Dr. Todd D. Ringler, Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523-1371.
E-mail: todd@placer.atmos.colostate.edu

and Tukey (1965) and the transform technique to evaluate vector-coupled sums (Orszag 1970; Eliasen et al. 1970). The efficiency of the FFT made spectral methods an attractive alternative to finite-difference methods. These two approaches to discretization, finite-difference methods and spectral methods, have been competing ever since, with each method having enough positive attributes to maintain its own viability.

Spectral models often solve the vorticity–divergence form of the primitive equations, as opposed to solving for the vector components of the momentum equation (e.g., Bourke 1974). Unlike components of the momentum vector, vorticity and divergence are true scalars whose values are invariant with respect to the choice of coordinate system. Furthermore, conservation principles for potential vorticity and potential enstrophy are easily formulated within the context of the vorticity and divergence equations. Solving the vorticity and divergence equations involves extra computational effort since at each time step elliptic equations must be inverted to determine the vector velocity. This overhead has generally inhibited finite-difference modelers from using the vorticity and divergence equations. In contrast to finite-difference methods, spectral decomposition reduces the elliptic operators to separable algebraic equations that are trivial to solve. By solving for vorticity and divergence, spectral methods isolate the terms related to gravity wave propagation and allow efficient semi-implicit schemes to be implemented. Furthermore, spectral methods do an exceptional job of modeling linear dynamics. For example, spectral methods give the exact phase speeds for the individual spherical harmonic components of linearly propagating waves. Also, the linear growth rates of instabilities, such as baroclinic instability, are very accurately simulated by spectral models.

When the state variables and boundary conditions that are expanded in the appropriate spectral space are smooth (i.e., infinitely differentiable), spectral methods converge extremely rapidly as the number of retained modes increases. In fact, for such a case spectral methods converge exponentially fast, which is faster than finite-difference methods of any accuracy (Jarrud and Simmons 1983). On the other hand, if state variables or boundary conditions are not infinitely differentiable and are "lumpy," as is the case with cloud liquid water and surface topography, the rate of convergence is dramatically reduced (from exponential to algebraic) and spectral ringing develops in regions of rapid change. One effect of spectral ringing is to generate negative values of fields that should be positive definite. Generally, an ad hoc method is required to overcome these deficiencies (Royer 1986). When the variable itself is discontinuous or nearly so (e.g., cloud liquid water), the spectral expansion is only first-order accurate (Tennekes and Lumley 1972), which would be the same accuracy as finite-difference methods in this case. Furthermore, in this instance the spectral expansion converges only in

the L^2 sense, not in the L^∞ sense. Therefore large local errors remain, regardless of the number of spectral modes retained (Greenberg 1978). Increasing computer power translates into retaining more wavenumbers, but may not translate into increased fidelity of simulations based on spectral methods when key variables are nearly discontinuous in space.

As the vertical coordinate systems of dynamical cores move away from the conventional σ coordinate to a more physical coordinate, finite-difference methods may be the only viable approach. An example would be the use of potential temperature as the vertical coordinate, in which case coordinate surfaces intersect the surface boundary and layers become "massless" (e.g., Hsu and Arakawa 1990). Such a mass distribution can be handled in a straightforward manner by finite-difference methods based on flux-corrected transport algorithms (Zalesak 1979). As stated above, a spectral expansion will converge very slowly with large local errors regardless of the spectral resolution.

Finite-difference methods do, of course, have disadvantages. In those instances where the dependent variables are smoothly varying in space, finite-difference methods will generally be less accurate than spectral methods. Also, finite-difference methods usually introduce anomalous dispersion in the simulation of advection and wave propagation (Haltiner and William 1980) and anomalous diffusion in the solution of the Poisson equation (Jaluria and Torrance 1986). Conventional finite-difference models integrate the components of the momentum equation, as opposed to the vorticity and divergence equations. Gravity wave propagation, which is associated with the divergent portion of the vector wind field, is contained in both components of the momentum equation. As a result, the implementation of semi-implicit schemes within finite-difference models is in general slightly more complicated than in spectral models.

Some of the commonly cited disadvantages of finite-difference models are not inherent in the finite-difference methods, but rather in the manner in which the methods have been implemented. For instance, the grids used to discretize the surface of the globe have generally been regular latitude–longitude grids, which have strong singularities at the North and South Poles. Spectral methods do not suffer from this problem. Using latitude–longitude grids results in extremely small grid cells areas near the poles relative to grid cell areas near the equator. Substantial computational effort and ad hoc methods are needed to suppress the numerical problems arising from having singularities at the grid poles (Williamson 1976). Given the relative smallness of grid cells in the polar regions, a disproportionately large amount of computational work is used to simulate that region. Furthermore, for those finite-difference models that required the inversion of elliptic operators, the disparate spatial scales of regular latitude–longitude grids reduces the efficiency of iterative elliptic solvers.

The purpose of this work is to develop a new three-dimensional dynamical core that combines many of the advantages of both finite-difference methods and spectral methods, while eliminating key deficiencies of each. This new dynamical core is radically different from conventional finite-difference models in that it uses a geodesic grid. The surface of the globe is discretized using an assembly of hexagons and 12 pentagons, as opposed to the quasi-rectangular grid cells of latitude–longitude grids. The geodesic grid gives nearly uniform grid cell areas across the entire globe, and does not suffer from the strong polar singularities present in regular latitude–longitude grids. Multigrid methods make it practical to solve the vorticity and divergence equations. Furthermore, this model uses a single unstaggered grid, the Z grid (Randall 1994), on which all of the prognostic variables are defined. While the geodesic grid allows a more elegant discretization of the primitive equations, it does lose some of the simplicity in its numerical implementation. The development of this geodesic dynamical core (GDC) grew naturally out of the work by Heikes and Randall (1995a,b), who solved the shallow water equations on a geodesic grid.

Section 2 describes the governing equations. Section 3 provides an introduction to geodesic grids and describes the numerical methods used to discretize the continuous equations. Results obtained using the idealized forcing proposed by Held and Suarez (1994) are shown in section 4. In addition, section 4 contains preliminary results from a full-physics AGCM. Section 4 also includes a comparison of computational efficiency between the geodesic grid model and a spectral model. A discussion with some conclusions is given in section 5.

2. Governing equations

The continuous governing equations used here are similar to those used in most present-day AGCMs. A set of nonlinear partial differential equations is used to describe the temporal and spatial evolution of mass, momentum, potential temperature, and an arbitrary number of passive tracers. While the final form of the governing equations used in this work does not make reference to any global horizontal coordinate system, an explicit vertical coordinate is used. The model has been tested with a variety of vertical coordinate systems, including those described by Phillips (1957), Suarez et al. (1983), and Hsu and Arakawa (1990). In this work we focus mainly on results obtained using the Phillips (1957) sigma coordinate. The coordinate is defined by

$$\sigma = \frac{p - p_T}{p_S}, \qquad (1)$$

where p is the pressure, p_S is the surface pressure, and p_T is a constant defining the pressure at the top of the model. Phillips (1957) used $p_T = 0$.

In the equations written below, we define ∇_p to operate along surfaces of constant pressure, ∇ to operate

along surfaces of constant σ, and $\pi \equiv p_S - p_T$. The prognostic equations describe the evolution of velocity, mass-weighted potential temperature, mass-weighted passive tracers, and surface pressure, respectively:

$$\frac{\partial}{\partial t}\mathbf{V} + \left(\frac{\varsigma + f}{\pi}\right)\mathbf{k} \times \pi\mathbf{V} + \nabla K + \dot\sigma\frac{\partial}{\partial\sigma}\mathbf{V} = -\nabla_p\Phi + \mathbf{G},$$
$$(2)$$

$$\frac{\partial}{\partial t}(\pi\theta) + \nabla \cdot (\pi\theta\mathbf{V}) + \frac{\partial}{\partial\sigma}(\dot\sigma\pi\theta) = \frac{\pi\theta Q}{c_p T}, \quad (3)$$

$$\frac{\partial}{\partial t}(\pi q) + \nabla \cdot (\pi q\mathbf{V}) + \frac{\partial}{\partial\sigma}(\dot\sigma\pi q) = S, \qquad (4)$$

$$\frac{\partial p_S}{\partial t} = \int_{\sigma=0}^{\sigma=1} -\nabla \cdot (\pi\mathbf{V})\, d\sigma. \qquad (5)$$

Nonconservative sources on the right-hand sides of (2), (3), and (4) are denoted by $\mathbf{G}$, Q, and S, respectively. The symbols ς and f denote the relative and planetary vorticities, while K and Φ represent the kinetic energy and geopotential height, respectively. Using the relation $\nabla_p = \nabla + (\nabla_p\sigma)(\partial/\partial\sigma)$, the pressure gradient term of (2) can be rewritten as $\nabla_p\Phi = \nabla\Phi + \sigma\alpha\nabla\pi$, where α is the specific volume.

As is well known, any vector field can be separated into purely rotational and divergent vector fields by using Helmhotz's decomposition (e.g., Panton 1984). The rotational and divergent components of the velocity fields can be expressed in terms of scalar potential fields. This relationship is expressed as

$$\mathbf{V} = \mathbf{k} \times \nabla\psi + \nabla\chi, \qquad (6)$$

where $\mathbf{k}$ is the unit vector normal to the coordinate surface, ψ is the streamfunction describing the rotational part of the vector field, and χ is the velocity potential describing the divergent part of the vector field. If we substitute (6) into (2), and take the curl and divergence of the result, we obtain the vorticity and divergence equations, respectively. These equations have the forms

$$\frac{\partial\eta}{\partial t} - J(\eta, \psi) + F(\eta, \chi) + F\left(\dot\sigma, \frac{\partial\psi}{\partial\sigma}\right) + J\left(\dot\sigma, \frac{\partial\chi}{\partial\sigma}\right)$$
$$+ J(\sigma\alpha, \pi) = \mathbf{k} \cdot \nabla \times \mathbf{G} \quad \text{and} \qquad (7)$$

$$\frac{\partial\delta}{\partial t} - J(\eta, \chi) - F(\eta, \psi) + F\left(\dot\sigma, \frac{\partial\chi}{\partial\sigma}\right) - J\left(\dot\sigma, \frac{\partial\psi}{\partial\sigma}\right)$$
$$+ L(K + \Phi) + F(\sigma\alpha, \pi) = \nabla \cdot \mathbf{G}, \qquad (8)$$

where $J(A, B) = \mathbf{k} \cdot (\nabla A \times \nabla B)$, $F(A, B) = \nabla \cdot (A\nabla B)$, $L(A) = \nabla^2 A$, η is the absolute vorticity, and δ is the divergence. Vorticity and divergence are related to the streamfunction and velocity potential through the diagnostic relations

$$\eta - f = \nabla^2\psi \quad \text{and} \qquad (9)$$

$$\delta = \nabla^2\chi. \qquad (10)$$

Note that (9) and (10) are used to determine ψ and χ, respectively. This involves inverting the elliptic operator ∇^2.

In order to close the system, we need equations for the geopotential height, Φ, and the vertical velocity, $\dot{\sigma}$. The geopotential height is obtained by integrating the hydrostatic equation from the model lower boundary upward. In differential form, the equation for Φ is written as

$$\partial\Phi = -\alpha\partial p = -\alpha\pi\partial\sigma. \tag{11}$$

The vertical velocity is obtained from mass continuity as

$$(\pi\dot{\sigma})|_{\sigma=\sigma'} = -\sigma'\frac{\partial\pi}{\partial t} - \int_{\sigma=0}^{\sigma=\sigma'} \nabla\cdot(\pi\mathbf{V})\,\partial\sigma, \tag{12}$$

with the boundary conditions $\dot{\sigma}|_{\sigma=0} = 0$ and $\dot{\sigma}|_{\sigma=1} = 0$. We can eliminate the vector velocity in (12) by substituting (6) and expanding. This manipulation gives the vertical velocity in terms of the streamfunction and velocity potential:

$$(\pi\dot{\sigma})|_{\sigma=\sigma'} = -\sigma'\frac{\partial\pi}{\partial t}$$
$$+ \int_{\sigma=0}^{\sigma=\sigma'} [J(\pi, \psi) - \nabla\cdot(\pi\nabla\chi)]\,\partial\sigma. \tag{13}$$

Similar manipulations can be performed on (3), (4), and (5) to express those equations in terms of streamfunction and velocity potential. Equations (3)–(5) and (7)–(12) constitute the "dynamical core" of the general circulation model. Note that no horizontal coordinate system is explicitly referenced; the equations are written in invariant form.

3. Numerical methods

Given the nonlinear nature of the governing equations, we must use discrete numerics to obtain anything more than trivial solutions. Since large-scale atmospheric motion is predominantly two-dimensional, we can think of the atmosphere as being an assembly of vertically stacked layers. The methods to discretize the three-dimensional fluid are then, to a large extent, separable into the horizontal discretization of each layer and the vertical discretization. Since the problem is somewhat separable, we can draw on previous work with the shallow water equations to aid in the horizontal discretization of this three-dimensional fluid.

a. Horizontal discretization on the geodesic grid

The idea of using a geodesic grid to discretize the spherical globe was put forth by Williamson (1968) and Sadourny et al. (1968). Both solve for the velocity components and height in the nondivergent shallow water equations. Masuda and Ohnishi (1986) made progress using geodesic grids by solving the vorticity-divergence form of the shallow water equations. They show that their approach conserves mass, energy, and potential enstrophy. As shown above, when the vorticity-divergence form of the shallow water equations is used, elliptic operators relating the vorticity and divergence to their respective potential fields must be inverted each time step. While this is trivial to do within spectral models, it is a major obstacle to overcome when using finite-difference methods. Heikes (1993) and Heikes and Randall (1995a,b) further develop the numerical methods related to geodesic grids by implementing a multigrid method to efficiently invert the elliptic operators. Thuburn (1997) used this multigrid method in the development of a potential vorticity-based shallow water model discretized on a spherical geodesic grid. Thuburn (1997) also developed a semi-implicit time-stepping scheme to further increase the computational efficiency of this modeling framework.

A geodesic grid covers the surface of a sphere with an assembly of polygons. In general, the generation of geodesic grids involves the repeated use of a set of simple rules. As shown in Fig. 1 each face of the initial icosahedron (Fig. 1a) is bisected to form four faces (Fig. 1b). Each vertex of this new grid is then projected to the surface of the sphere (Fig. 1c). Applying the rules of bisection and projection to subsequent grids allows finer meshes to be generated (as in Figs. 1d–f). The initial icosahedron has 12 vertices and the first subdivision has 42 vertices. Each vertex is associated with a grid point. The formula relating the number of grid cells, N_c, to the level of recursion, R, is $N_c = 5 \cdot 2^{2R+3} + 2$, where $R = -1$ corresponds to the initial icosahedron. The dynamical core is often integrated with $R = 3$ or $R = 4$, which correspond to 2562 and 10 242 grid cells, respectively. The average distance between grid cell centers with $R = 3$ and $R = 4$ is 481.6 and 240.9 km, respectively.

In the context of the shallow water equations, two modeling frameworks have evolved from Fig. 1. One method, which is extended to three-dimensions in this paper, defines the vertices of the spherical geodesic grid to be grid points. As shown in Fig. 2 and discussed below, this results in set of grid cells that are hexagonal or pentagonal in shape. This is the approach used by Masuda and Ohnishi (1986), Heikes and Randall (1995a,b), and Thuburn (1997). Another approach is to define the vertices as corners of the triangular elements shown in Fig. 1. This method is followed by Baumgardner and Frederickson (1985), Swarztrauber et al. (1997), and Stuhne and Peltier (1996, 1999).

In Fig. 2, the area associated with grid point P_0 is that area that lies closer to P_0 than to any other grid point. As discussed by Heikes and Randall (1995a), this grid has many attractive properties. The areas of the grid cells are nearly uniform across the entire sphere. This eliminates the common problems encountered near the grid poles when using conventional latitude–longitude grids. The recursive nature of the grid lends itself
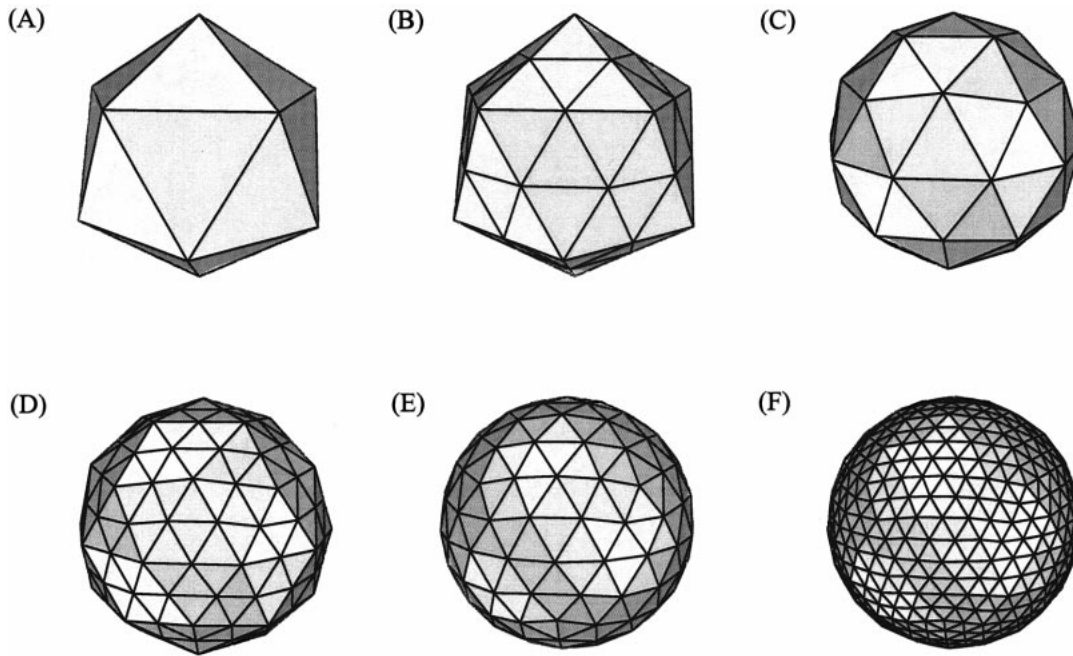
FIG. 1. Generating geodesic grids by recursive bisection and projection. (a) Starting with an icosahedron, (b) each face is bisected to form four new faces, (c) each new vertex is projected onto the unit sphere. (d), (e), and (f) The results of continued bisection, projection, and bisection. Adopted from Heikes and Randall (1995a).

nicely to the use of multigrid methods that allow the efficient inversion of elliptic operators. Another important property of this grid, which will be explained in more detail immediately below, is that it allows a simple

FIG. 2. The area associated with each vertex, such as P_0, is the set of all points closer to P_0 than any other vertex. Each vertex is referred to as a grid point. Both α and β are arbitrary scalar fields evaluated at each grid point; b_i is a sequential listing of grid cell corners.

second-order accurate discretization of the analytic operators J, F, and L.

The J, F, and L operators must be numerically approximated at a grid point such as P_0 (Fig. 2). If we assume that the grid cell area is sufficiently small, we can approximate the values of the operators at P_0 as the respective mean values integrated over the cell's area. Using the Jacobian operator as an example, this is expressed mathematically as

$$J(\alpha, \beta)|_{P_0} \approx \frac{1}{A_c} \iint_{A_c} J(\alpha, \beta) \, dA, \qquad (14)$$

where α and β are arbitrary scalar fields and A_c is the area of the cell. With this approximation, we can use Green's theorem to reduce the area integral to a line integral as

$$J(\alpha, \beta)|_{P_0} \approx \frac{1}{A_c} \iint_{A_c} J(\alpha, \beta) \, dA = \frac{1}{A_c} \oint_C \alpha \frac{\partial \beta}{\partial s} \, ds, \quad (15)$$

where C is the contour enclosing A_c and ds is an infinitesimal segment along C. Referring to Fig. 2, the contour integration is carried out by summing over the cell walls associated with P_0 as

$$J(\alpha, \beta)|_{P_0} \approx \frac{1}{A_c} \sum_{i=1}^{N} \left(\frac{\alpha_0 + \alpha_i}{2} \right) \left(\frac{b_i - b_{i+1}}{l_i} \right) l_i, \quad (16)$$

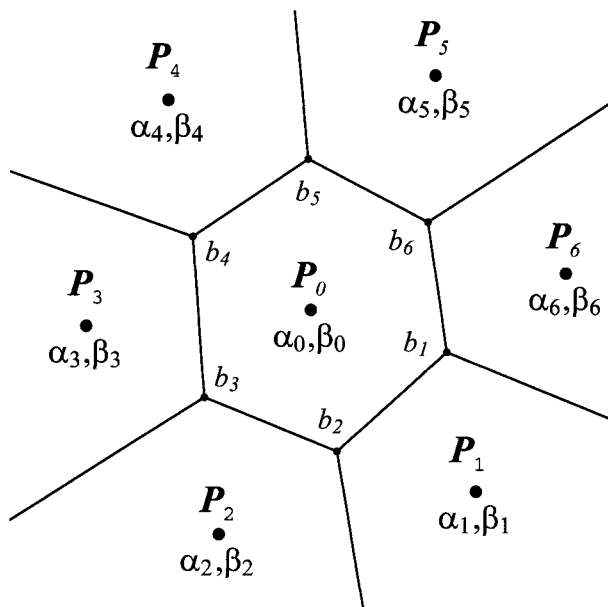where l_i is the length of the arc of the cell wall and

$$b_i = \frac{\beta_0 + \beta_i + \beta_{i-1}}{3}. \tag{17}$$

So α is approximated at the cell wall as the average of the two nearest gridpoint values, and β is approximated at the cell corners as the average of the three closest gridpoint values. Interpolation to the cell-wall center can be interpreted as fitting a line through the closest two grid points, while interpolation to the grid cell corners can be interpreted as fitting a plane through the closest three grid points. Formally, both of these approximations lead to a finite-difference stencil that is second-order accurate (Heikes and Randall 1995b). Equation (16) can be simplified to yield

$$J(\alpha, \beta)|_{P_0} \approx \frac{1}{6A_c} \sum_{i=1}^{N} (\alpha_0 + \alpha_i)(\beta_{i-1} - \beta_{i+1}). \tag{18}$$

Operators F and L can be reduced to line integrals and discretized in a similar fashion (see Heikes and Randall 1995a). The analytic operators satisfy

$$\iint_A J(\alpha, \beta)\, dA = 0, \qquad \iint_A F(\alpha, \beta)\, dA = 0, \quad \text{and}$$

$$\iint_A L(\alpha)\, dA = 0, \tag{19}$$

where the integrations are carried out over the entire sphere. Using the fact that the discrete versions of these operators are evaluated by integrating along cell walls and each cell wall is shared by two grid cells, we can show that the numerical operators satisfy

$$\sum_{c=1}^{N_c} A_c \cdot J(\alpha, \beta)|_c = 0, \qquad \sum_{c=1}^{N_c} A_c \cdot F(\alpha, \beta)|_c = 0, \quad \text{and}$$

$$\sum_{c=1}^{N_c} A_c \cdot L(\alpha)|_c = 0, \tag{20}$$

where the summations are over the total number of grid points, N_c. Furthermore, it can be shown that this discretization conserves kinetic energy and enstrophy under advection in purely rotational flows.

Along with vorticity and divergence, mass is defined at the grid cell center P_0. This unstaggered grid system in which vorticity, divergence, and mass are all defined at the same location is referred to as the Z grid (Randall 1994). Within the context of the shallow water equations, Randall (1994) showed that the Z grid does a satisfactory job in simulating geostrophic adjustment and performs better than the more conventional B and C grids. Furthermore, the Z grid is the natural choice for the numerical integration of the vorticity-divergence equations.

b. Vertical discretization

Within the context of the Phillips sigma coordinate, most of the numerical issues relating to vertical discret-

ization were worked out by Arakawa and Lamb (1977). They formulated the numerics of the discrete primitive equations so as to conserve a host of quantities, including mass, entropy, and total energy. While the conservation relations were originally formulated with the momentum form of the primitive equations, each of these relations has been implemented in the vorticity-divergence form of the primitive equations used in this model [see Ringler et al. (1998) for more details]. A goal that Arakawa and Lamb (1977) discuss is to conserve kinetic energy under vertical advection. By beginning with the discrete form of the vertical advection of momentum that conserves kinetic energy, the terms in (7) and (8) that are related to vertical advection can be formulated such that the discrete form of the vorticity-divergence equations also possess this conservation property. This derivation is shown in appendix A.

As in Arakawa and Lamb (1977), the vertical staggering of variables uses the Lorenz vertical grid, which places θ at the layer centers. This is in contrast to the Charney–Phillips vertical grid on which θ is placed at the layer edges. A disadvantage of the Lorenz grid is that there is an extra degree of freedom for θ in the vertical direction. This unconstrained degree of freedom often manifests itself as vertical grid-scale noise in θ and related fields, such as geopotential. Our experience has been that such vertical grid-scale noise is not strongly excited within our model.

c. Temporal discretization

Truncation errors due to horizontal discretization are generally much larger than those due to temporal discretization (Haltiner and Williams 1980). As a result, increasing the length of the time step generally does not degrade the simulation, as long as the numerical scheme is stable. In general, the stability criterion limiting the maximum allowable time step with an explicit scheme is the phase speed of the gravity waves. The idea leading to semi-implicit schemes is to integrate all terms related to gravity wave propagation in an implicit, or unconditionally stable, manner, while integrating all other terms explicitly (Hoskins and Simmons 1975). Since linear gravity waves are irrotational, no terms in the vorticity equation are essential for gravity wave propagation and so the vorticity equation can be integrated explicitly. All terms that are integrated explicitly use the third-order Adams–Bashforth method (Durran 1991). The details of the semi-implicit scheme are given in appendix B.

The semi-implicit scheme involves first evaluating all terms that are integrated explicitly, then making a first guess at the values of the prognostic variables that include terms that are integrated implicitly. One iteration through the semi-implicit scheme involves evaluating the implicit terms on the right-hand side of the prognostic equations, updating the prognostic variables, and then updating the diagnostic equations. Following an

iteration cycle, convergence is assessed by differencing the values of the prognostic variables at the current iteration and their values at the previous iteration. If this difference is smaller (in an absolute value) than a specified value for each prognostic variable at every grid point, then the solution is deemed to have converged and the integration continues to the next time level. This approach allows time steps of 30 and 20 min with horizontal resolutions of 2562 and 10 242, respectively. The lengths of these time steps are similar to those we have been able to use in a spectral model with comparable resolutions.

4. Results

Until recently, the evaluation and intercomparison of dynamical cores have received little attention. Even though a wide variety of numerical techniques, such as semi-Lagrangian methods, spectral methods, finite-element methods, and finite-difference methods, are being used in dynamical cores, no systematic intercomparison has been completed. One reason for this lack of intercomparison is related to the lack of exact solutions to the primitive equations with realistic forcing. Without the aid of analytic solutions, identifying and quantifying errors in the numerical models is difficult.

Several benchmark calculations have been proposed as a means of evaluating primitive equation dynamical cores. One such benchmark that has met with considerable acceptance is the calculation proposed by Held and Suarez (1994). In the Held–Suarez Test Case (HSTC), the thermal structure of the model atmosphere is relaxed to a prescribed "radiative–convective equilibrium" on a timescale of approximately 40 days. The HSTC emphasizes simplicity by specifying the forcing and boundary conditions to be zonally symmetric and keeping parameterizations to an absolute minimum. The simplicity of the formulation makes it readily "portable" to a wide variety of models and, thereby, facilitates intercomparisons. A weakness of the HSTC is that the real climate is composed of much more than the zonally symmetric statistics. Regardless, the HSTC has proven to be useful as a first step in evaluating dynamical cores, and we will use it here.

Several important points should be considered when analyzing the HSTC simulations. First, since analytical solutions are not known, separating right from wrong can be difficult or impossible. Enough simulations have been completed using the HSTC forcing that a consensus could be formed, but consensus is a poor substitute for truth. As with almost all numerical modeling, there is an underlying belief that as the resolution is increased, the model simulation will converge to the (statistically) correct solution. Held and Suarez (1994) point out that even at a spectral truncation of T63 the simulations have still not converged. As the model resolution increases, the variance of basic model statistics such as $\overline{u'u'}$, $\overline{v'v'}$, and $\overline{u'v'}$ increases. As a result, the fidelity of model simulations is often judged based on the amount of variance produced at a given resolution.

### a. Experimental design of the Held–Suarez test case

The initial condition for all the experiments is an isothermal atmosphere at rest. Random noise is added to the surface pressure field (± 0.5 Pa) and to the potential temperature field (± 0.5 K) in order to break the symmetry of the initial conditions. The lower-resolution (2562 and T30) integrations are carried out for 1200 days with averages and statistics computed from the last 1000 days. The higher-resolution integrations (10 242 and T63) are carried out for 600 days with averages and statistics computed from the last 450 days. Since the HSTC forcing is zonally symmetric, additional statistical significance can be obtained by analyzing zonal-mean statistics instead of sections at a given longitude. If we assume that each hemisphere is a nearly independent realization, then the statistical significance of the results can be evaluated by comparing the results between hemispheres.

In order to compare the geodesic grid model to an "independent" model, we also present results from a spectral model that is identical to that used by Held and Suarez (1994). Simulations from the GDC at resolutions of 2562 and 10 242 are compared to spectral model simulations completed at T30 and T63, respectively. All results shown use the Phillips sigma coordinate (Phillips 1957) with 17 levels spaced evenly in pressure. The time-stepping scheme is semi-implicit in all cases. In addition, both models incorporate a ∇^4 diffusion on relative vorticity, divergence, and potential temperature. While both models have biharmonic diffusion, the diffusion is implemented differently in each model. In the GDC the coefficient for the biharmonic diffusion is uniformly set to 4.0×10^{16} and 7.5×10^{15} m^4 s^{-1} at resolutions of 2562 and 10 242, respectively. In the spectral model, the coefficient is a function of wavenumber with the smallest resolved scale dissipated with an e-folding period of 0.1 days. At the smallest resolved scales the GDC diffusion coefficients are approximately equivalent to e-folding periods of 0.15 days and 0.06 days at resolutions of 2562 and 10 242, respectively.

b. Instantaneous fields in a 10 242 simulation

A qualitative means of evaluating the quality of the simulation is to look at instantaneous "snapshots" of the model state. In particular, it is useful to analyze the structure of mature extratropical baroclinic eddies. Figure 3a shows the lowest-layer potential temperature and the surface pressure on day 225 in a simulation using 10 242 grid cells. The figure shows only a part of the global domain in order to isolate two mature baroclinic eddies, which appear as two nearly occluded low pressure systems with surface pressure minima of 965 and 960 hPa. Regions of warm advection to the east of the

(A) Potential Temperature (K) and Surface Pressure (Pa)
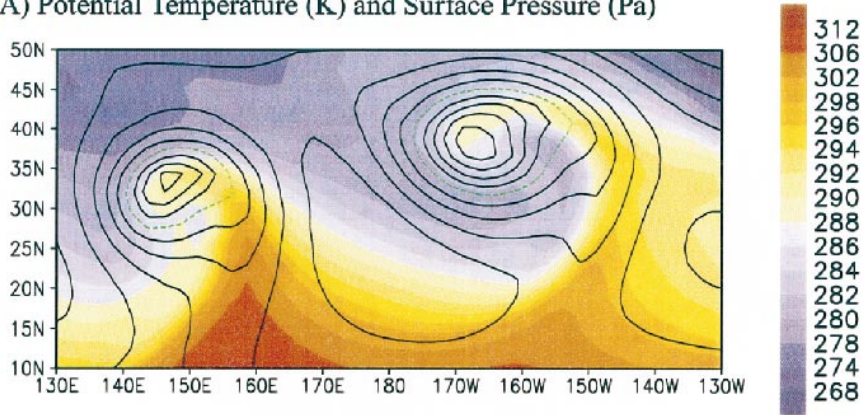
(B) Relative Vorticity (s^{-1}*1e5) and Wind Vectors (m s^{-1})
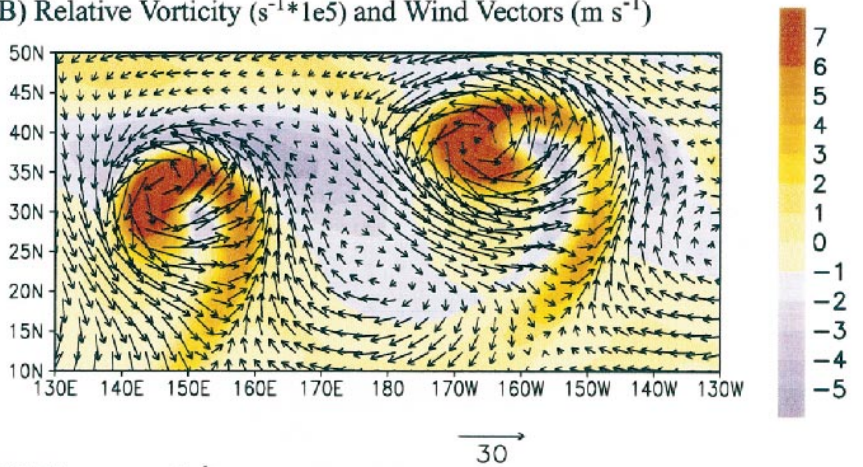
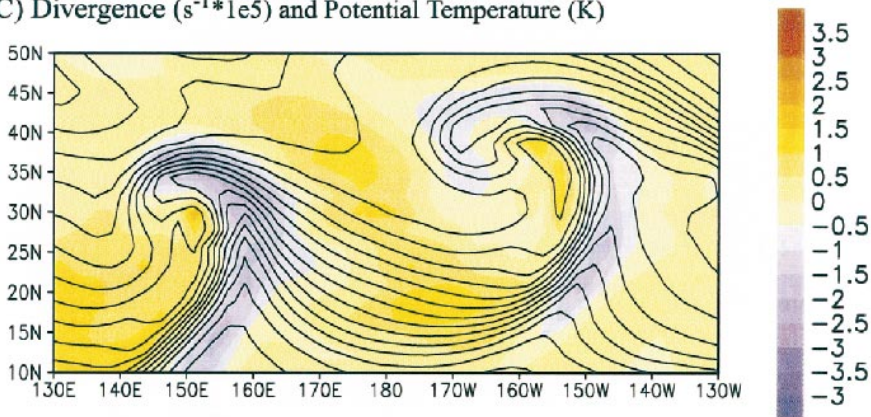## (C) Divergence (s^{-1}*1e5) and Potential Temperature (K)

FIG. 3. Snapshot of model state at day 225 in a simulation using the geodesic grid model with 10 242 cells. (a) Lowest-layer potential temperature (K) and surface pressure contours with the 980-hPa contour (dashed) and a contour interval of 5 hPa. (b) Lowest-layer relative vorticity (1×10^5 s^{-1}) with lowest-layer surface winds. (c) Lowest-layer divergence (1×10^5 s^{-1}) with potential temperature contours.

low pressure centers and cold advection to the west are evident. The eastern low pressure system has wrapped warm air almost completely around the low pressure center. Figure 3b shows the lowest-layer relative vor-

ticity along with the near-surface winds. (Note that the model does not prognostically solve for the wind vector field. Wind vectors are computed using bilinear interpolation to map the streamfunction and velocity poten-
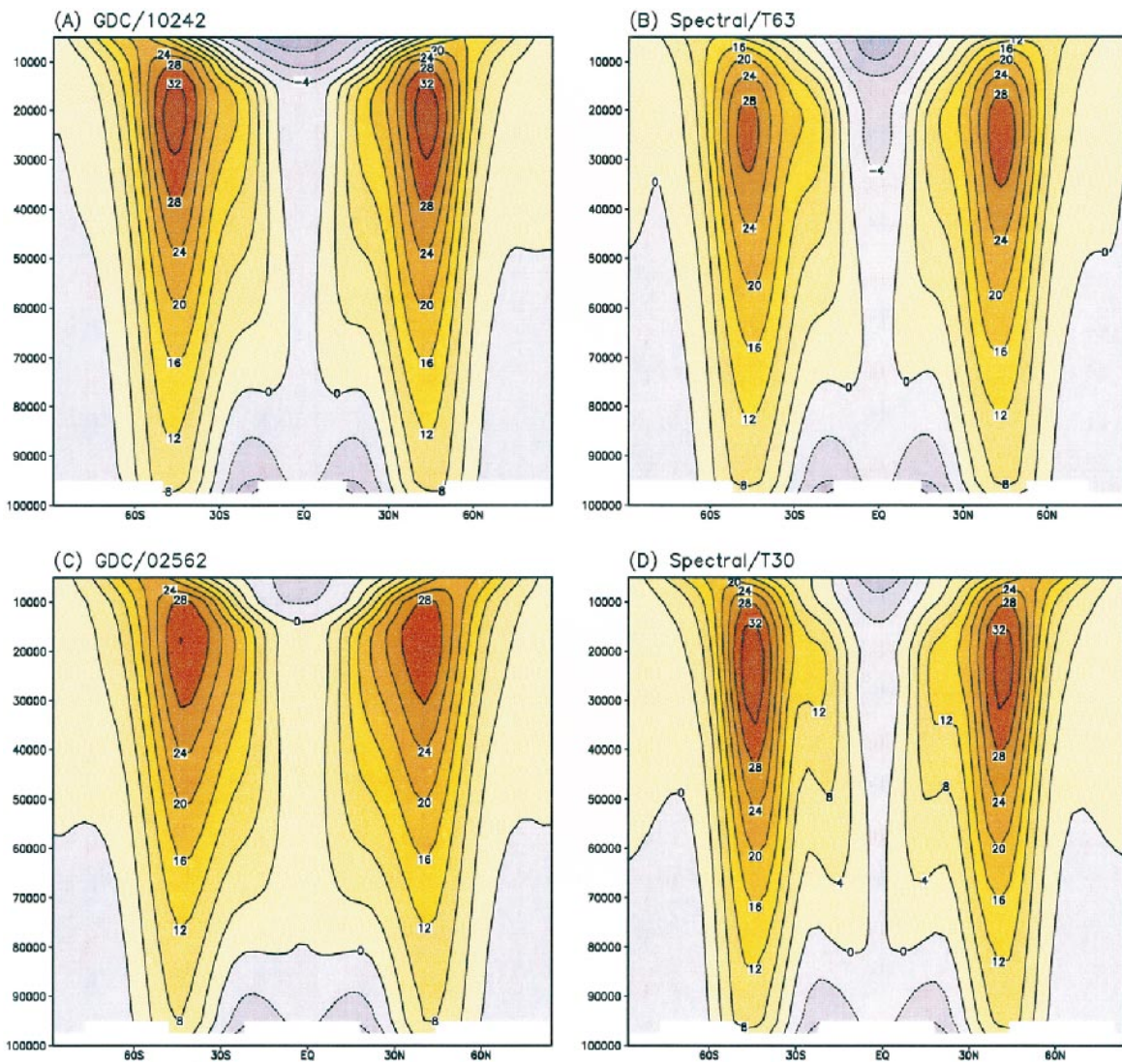
FIG. 4. Zonal-mean zonal wind (m s⁻¹) for four simulations. (a) The geodesic grid model with 10 242 cells, (b) the spectral model truncated at T63, (c) the geodesic grid model with 2562 cells, and (d) the spectral model truncated at T30.

tial from the geodesic grid to a regular latitude–longitude grid and then constructing the velocity field on that grid.) The low pressure centers are correlated with regions of strong cyclonic rotation. The relative vorticity extends away from the low pressure centers along the temperature front. Small regions of anticyclonic vorticity are developing immediately east of the low pressure centers due to low-level cold air advection. These regions are being wrapped in a spiral around the center of the cyclonic circulation. Figure 3c shows the lowest-layer surface divergence along with contours of near-surface potential temperature. The regions of convergence are correlated to regions of large temperature gradient (i.e., temperature fronts). Figures 3b and 3c clearly show that the system is attempting to further strengthen the front. This is consistent with analytical theories of frontogenis (Orlanski 1985).

c. Comparison of geodesic grid model simulations to spectral model simulations

In this section we compare simulations completed with 2562 grid cells to simulations completed with 10 242 grid cells. In addition, we compare the GDC results to results obtained with a spectral model truncated at T30 and T63. Statistics were computed on 39 evenly spaced pressure levels. Pressure surfaces that exist less than 20% of the time are masked out. Unless otherwise stated, each figure shown in this section has four panels where (a), (b), (c), and (d) show results from the 10 242, T63, 2562, and T30 simulations, respectively.

Figure 4 shows the zonal-mean zonal wind for the four experiments. The simulations show many common features. The midlatitude jet is located at approximately
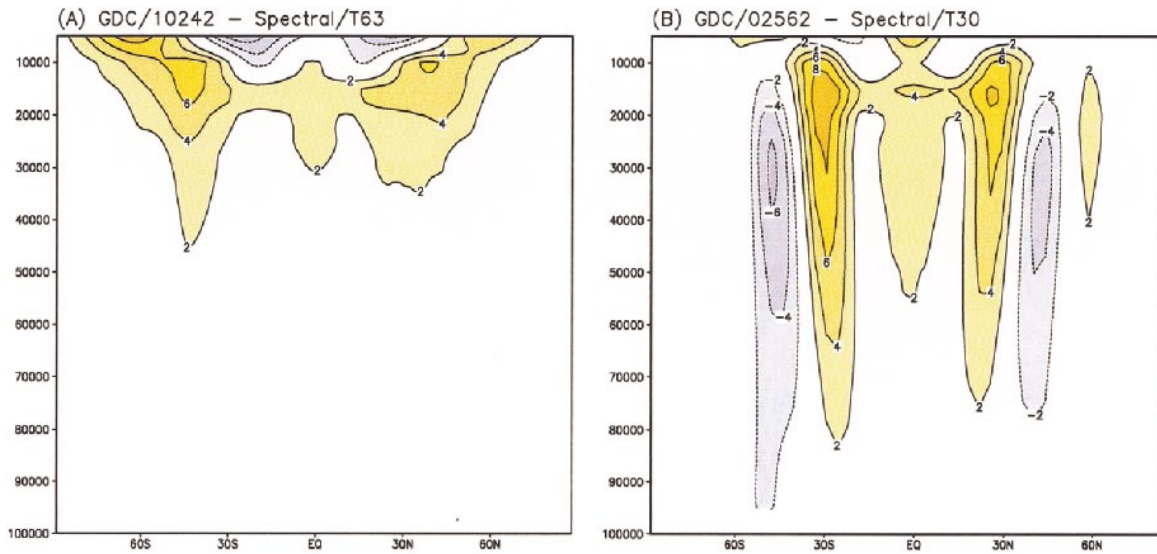
FIG. 5. Difference in zonal-mean zonal wind: (a) GDC/10 242–spectral/T63 and (b) GDC/2562–spectral/T30. Contour interval is 2 m s^{-1} with the zero contour omitted.

45° latitude and has a strength of about 30 m s^{-1} in all simulations. In addition, the surface easterlies are similar with maximum zonal winds of approximately 8 m s^{-1}, and all the simulations produce tropical stratospheric easterlies. Tropical easterlies extend through the entire depth of the atmosphere in all cases except the 2562 simulation where weak westerlies (less than 1 m s^{-1}) exist in the midtroposphere. The differences in the zonal-mean winds (10 242–T63 and 2562–T30) are shown in Fig. 5. At the higher resolution the differences in the zonal-mean zonal wind are small and are confined primarily above 200 hPa. At the lower resolution the difference is mainly in the barotropic component of the wind, with the 2562 simulation producing a jet that is equatorward of that obtained in the T30 simulation.

Overall measures of the baroclinic wave activity can be obtained by computing variances and covariances of fields such as zonal wind, meridional wind, and temperature. The zonal-mean temperature variance shown in Fig. 6 indicates that both models at both resolutions contain, to greater and lesser extents, baroclinic wave activity. As expected, the maximum in temperature variance occurs in the lower troposphere where the eddies are transporting heat poleward. Increasing the resolution from 2562 to 10 242 and from T30 to T63 results in approximately a 50% increase in variance. The higher-resolution integrations show a well-defined maximum at 800 hPa.

While the zonal-mean statistics, such as temperature variance, provide good measures of the overall amplitude of the variance and the vertical distribution of variance, they do not provide any information on the spatial scales in which this variance resides. One way of determining the scales that are producing the variance is to compute the vertically averaged zonal spectra of qua-

dratic quantities such as $\overline{u'u'}$, $\overline{v'v'}$, $\overline{u'v'}$, and $\overline{v'T'}$. Figure 7 shows the zonal spectrum of the vertically averaged zonal wind variance plotted against latitude. The simulations tend to produce three maxima in each hemisphere with one maximum at low wavenumber near 45° latitude and two maxima near wavenumber 5 straddling 45° latitude. The two maxima straddling 45° latitude are separated by approximately one baroclinic eddy length scale. Surface low pressure systems that move along the storm track, which in this case is located at approximately 45° latitude, will show maxima in zonal wind variance along the flanks of the storm track due to geostrophic balance. Increasing the resolution of the geodesic grid model results in an overall increase in variance of more than a factor of 2. Within both models the most pronounced increase in variance occurs at low wavenumber near 45° latitude and near the poles.

The vertically averaged divergence variance is plotted in Fig. 8 and shows a dramatic increase in variance with increasing resolution. In the lower-resolution simulations, maxima in the divergence variance related to baroclinic eddies are found in each hemisphere near zonal wavenumber 7, but relatively little variance occurs along the equator. The higher-resolution simulations show a marked increase in divergence variance along the equator with a maximum occurring near wavenumber 1. A Fourier decomposition in time of the divergence variance along the equator for wavenumber 1 (not shown) shows a typical red spectrum. Analyzing the vertical structure of the equatorial variance shows that the variance is contained predominately in the higher baroclinic modes. While the mechanism responsible for this dramatic increase in divergence variance along the equator has not been identified, it is possibly being
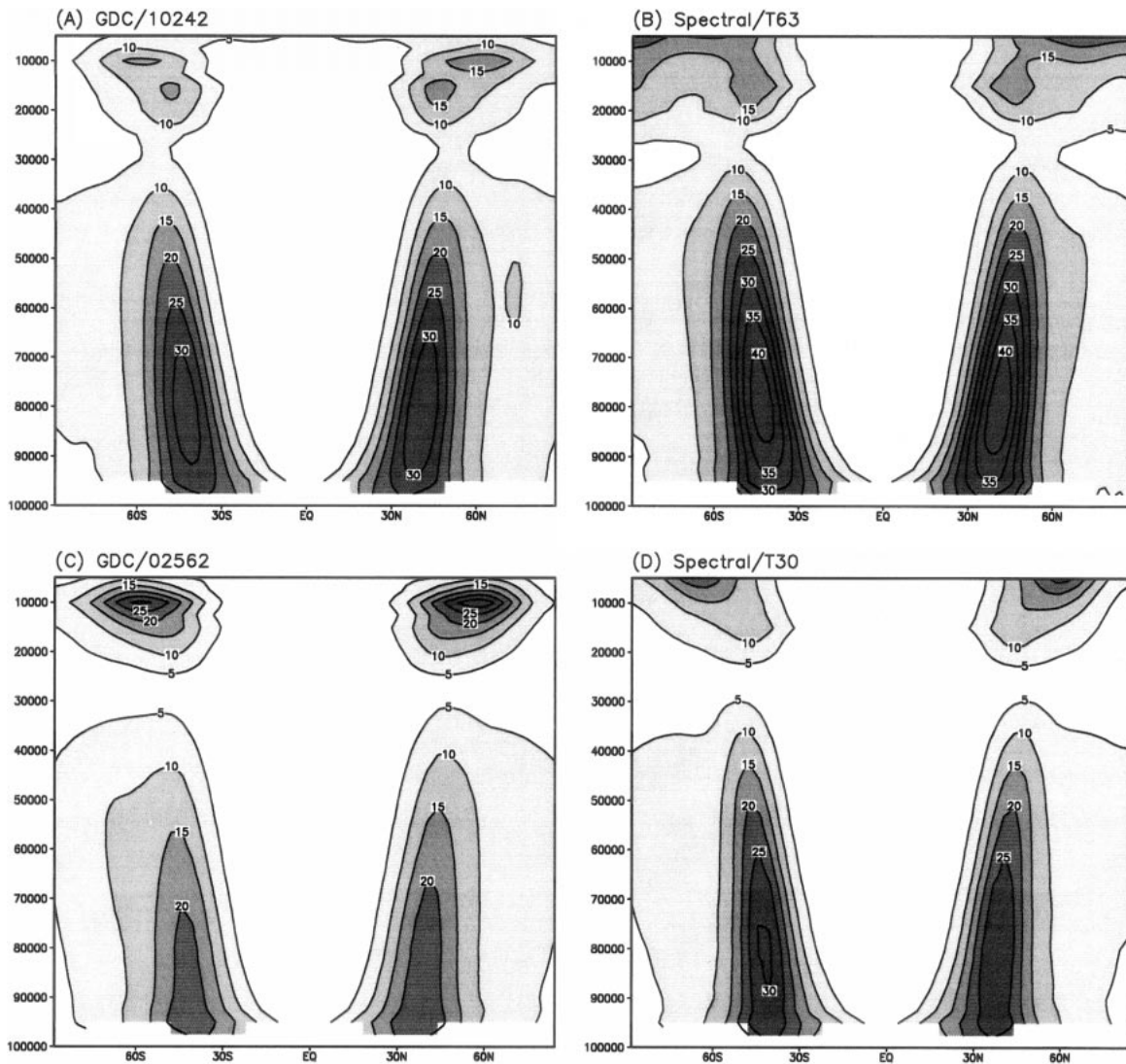
FIG. 6. Zonal-mean temperature variance for four simulations. (a) The geodesic grid at 10 242, (b) the spectral model at T63, (c) the geodesic grid at 2562, and (d) the spectral model at T30. Contour interval is 5 K^2 with the zero contour omitted.

forced by the increased extratropical wave activity. Figure 7 is consistent with this hypothesis.

d. Full-physics GCM results

The GDC has been merged with a state-of-the-art physics package to produce a full AGCM (GDC–AGCM). The physics package is exactly the same as is used in the latitude–longitude version of the Colorado State University (CSU) AGCM. The radiation scheme is based on the work of Harshvardhan et al. (1989) while the cloud microphysics was formulated by Fowler et al. (1996). The cumulus mass flux parameterization that allows for multiple cloud bases was developed by Ding and Randall (1998). The AGCM also includes version 2 of the Simple Biosphere Model land surface parameterization (Sellers et al. 1996). The gravity wave

scheme is that proposed by Palmer et al. (1986). All of these physical parameterizations are discretized on the same geodesic grid used in the dynamical core. The vertical coordinate used in the AGCM simulations is the generalized sigma coordinate (Suarez et al. 1983), which includes a prognostic treatment of the PBL depth. In addition to the biharmonic diffusion on vorticity, divergence, and potential temperature, which was used in the HSTC, a biharmonic diffusion is applied to water vapor in this full-physics simulation.

Beginning from a dry isothermal atmosphere at rest, the AGCM was spun up for two simulated years using 2562 grid cells in the horizontal and 17 layers (2562/17). At the end of this integration, the atmospheric state was interpolated to 10 242/17 and the integration carried forward for another five simulated years. Surface orography was averaged from a 1° × 1° dataset to the geo-
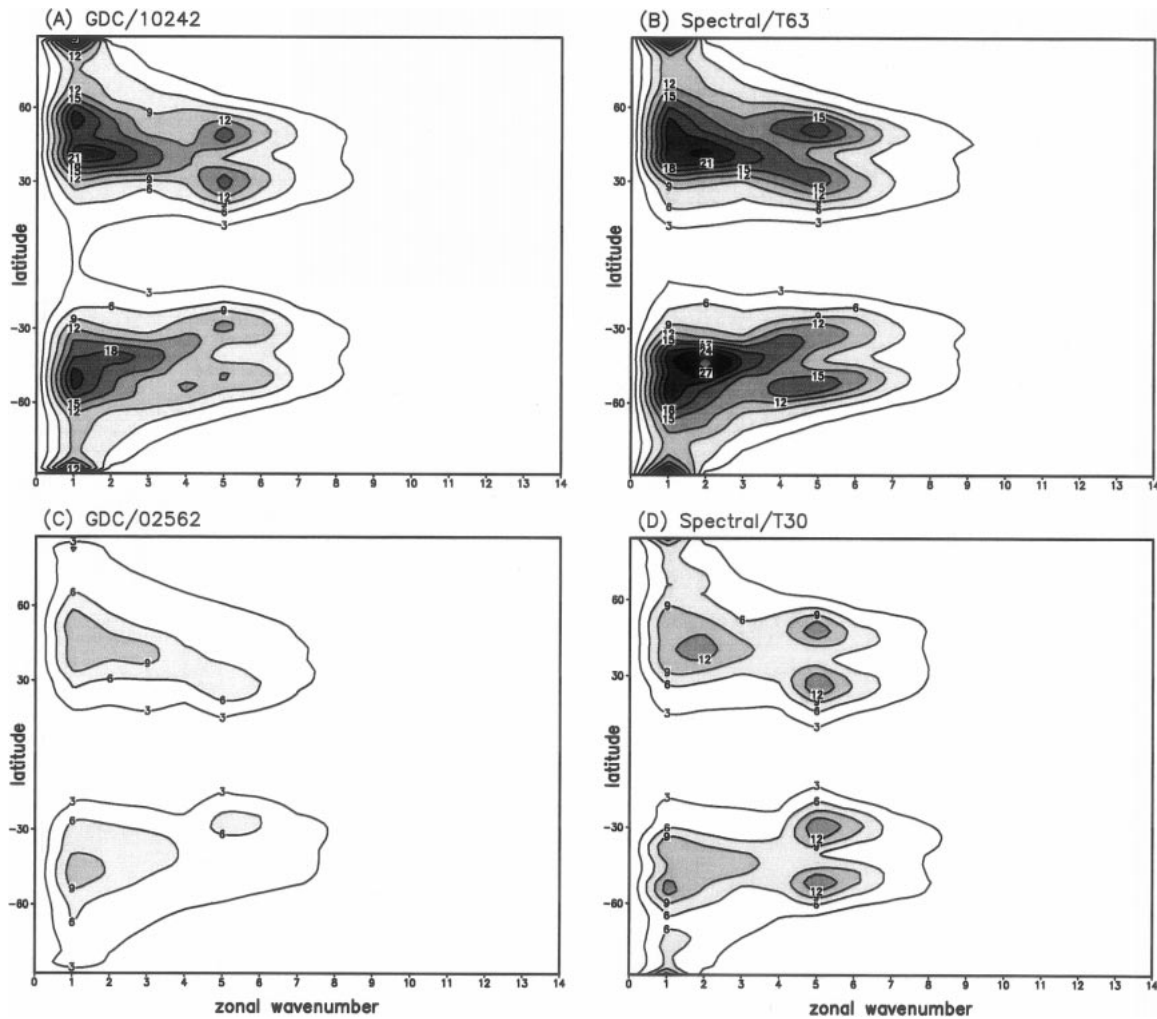
FIG. 7. Zonal spectrum of vertically averaged zonal wind variance for four simulations. (a) The geodesic grid at 10 242, (b) the spectral model at T63, (c) the geodesic grid at 2562, and (d) the spectral model at T30. Contour interval is 3.0 m^2 s^{-2} with the zero contour omitted.

desic grids (Jones 1999). Sea surface temperature (SST) was specified using the Atmospheric Modeling Intercomparison Project Phase 2 (AMIP2) monthly varying climatological SST dataset (Phillips 1996).

Figures 9a and 9b show the zonal-mean zonal wind for December–January–February (DJF) from the GDC–AGCM and from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis project (Kalnay et al. 1996). The model surface winds are of the correct amplitude and structure, with the exception of the Southern Hemisphere surface westerlies, which are shifted equatorward relative to the reanalysis data. The tropical easterlies in the model do not extend through the entire depth of the troposphere as in the observational data. The model produces stratospheric easterlies across most of the Southern Hemisphere, which is consistent with the NCEP–NCAR data. The Northern Hemisphere jet has the correct position and correct amplitude, expect

within the core region where the jet is approximately 10 m s^{-1} too strong. Similar to the surface winds, the Southern Hemisphere jet is shifted equatorward relative to the reanalysis data.

The zonal-mean temperature, which is shown in Figs. 9c and 9d, indicates that the model is developing a tropopause that is too strong relative to the observed data. Throughout the rest of the upper troposphere and stratosphere the model and reanalysis data are in qualitative agreement. The largest discrepancy occurs near the surface in the Southern Hemisphere. The model exhibits a uniform near-surface temperature profile poleward of 60° latitude, whereas the NCEP–NCAR data show a uniform gradient in near-surface temperature that extends to the southern pole. This discrepancy is strongly linked to the equatorward shift of the surface westerlies relative to the NCEP–NCAR data.

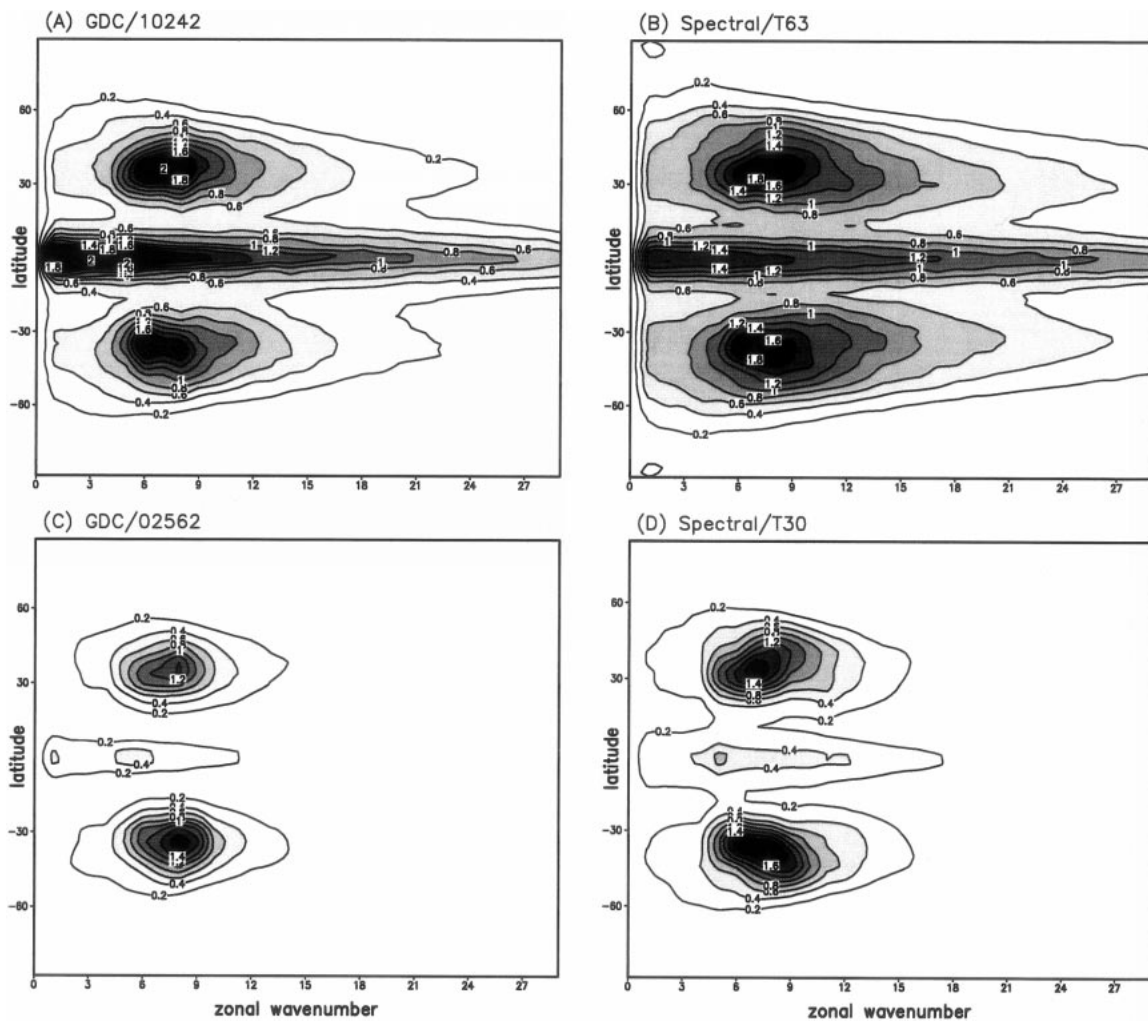Figure 10 shows DJF sea level pressure and 850-hPa streamlines from the GDC–AGCM simulation and from

FIG. 8. Zonal spectrum of vertically averaged divergence variance for four simulations. (a) The geodesic grid at 10 242, (b) the spectral model at T63, (c) the geodesic grid at 2562, and (d) the spectral model at T30. Contour interval is 0.2×10^{-12} s^{-2} with the zero contour omitted.

the NCEP–NCAR reanalysis data. Both the Aleutian and Icelandic low pressure systems are present in the model simulation, but are slightly weaker than found in the NCEP–NCAR data. The continental high pressure systems over Asia and North America are also present in the model simulation. In the Pacific basin, the subtropical anticyclones are weaker in the model data than in the NCEP–NCAR data. Within the tropical Pacific the reanalysis data show easterlies extending completely across the basin, whereas the model shows a region of convergence near the date line. In the polar regions, the area-averaged sea level pressure in the model simulation is approximately 10 hPa too high in the Arctic and 10 hPa too low in the Antarctic.

The model-simulated DJF precipitation is compared to the Global Precipitation Climatology Centre (GPCC) dataset (Xie and Arkin 1996) in Fig. 11. Regions of localized precipitation related to the Northern Hemisphere storm tracks in the North Pacific and North At-

lantic are well simulated in terms of spatial structure and amplitude. Within the tropical Pacific, the model simulates a coherent ITCZ across the entire basin. The western Pacific precipitation is shifted eastward in the model relative to the GPCC dataset. The model generates a South Pacific convergence zone and a South Atlantic convergence zone, which are both in good agreement with the GPCC dataset. The precipitation simulated by the model over the Amazon Basin and across the tropical Atlantic are also consistent with the GPCC dataset.

e. Computational efficiency

The initial development of this model has concentrated on the fidelity of the simulations, as opposed to the computational efficiency of the numerical algorithm. Given the satisfactory quality of simulation results presented above, we are now in a position to optimize the
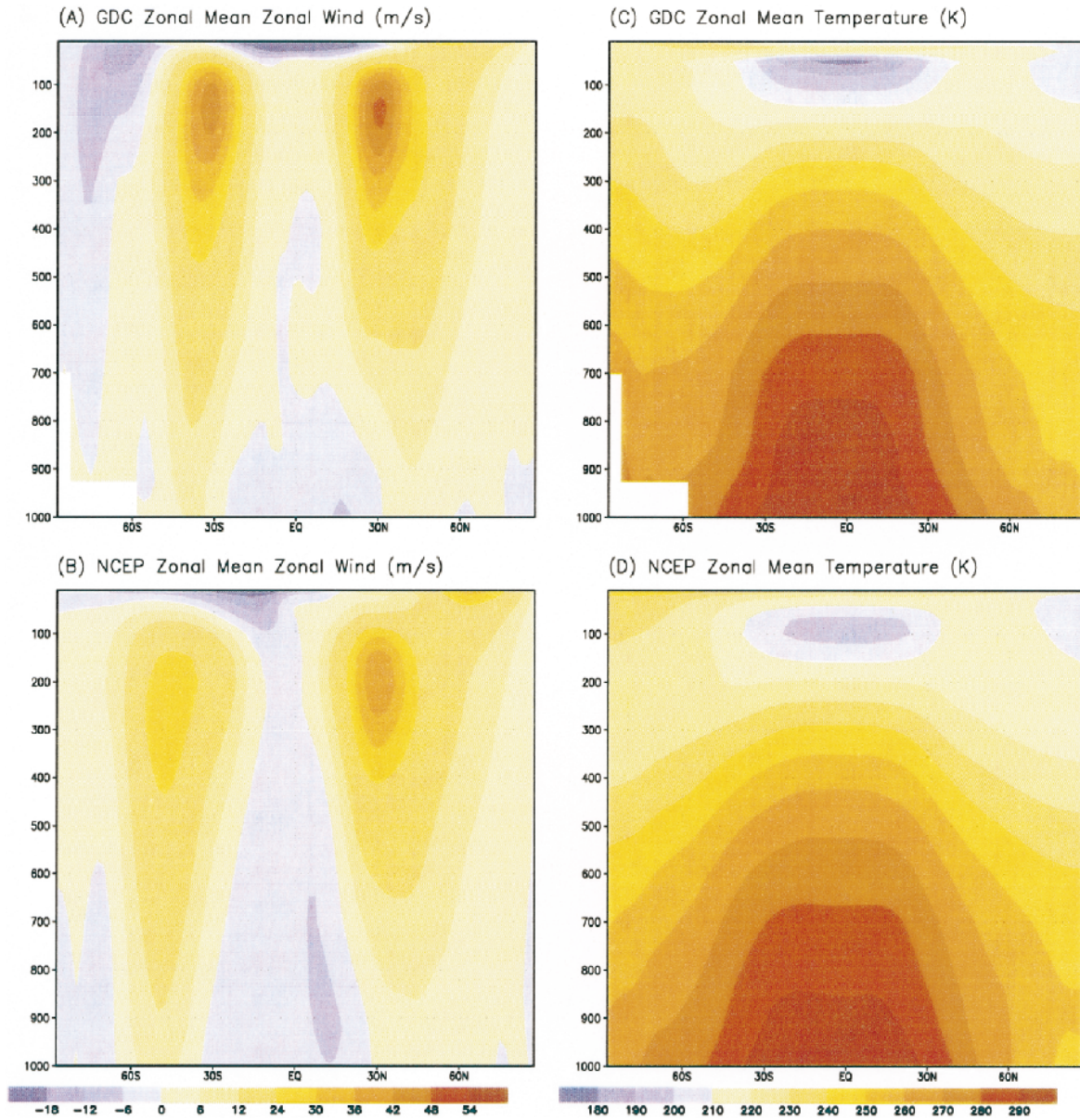
FIG. 9. DJF zonally averaged zonal wind (m s^{-1}) for the (a) GDC–AGCM and (b) NCEP–NCAR reanalysis datasets. DJF zonally averaged temperature (K) for the (c) GDC–AGCM and (d) NCEP–NCAR reanalysis datasets.

GDC such that it will be competitive not only in terms of simulation quality, but also in terms of computational efficiency. This section provides a baseline comparison between the GDC and the spectral dynamical core.

Table 1 provides a comparison of the models' computational speeds in terms of model resolution, rate of floating point operations (MFlops), and CPU time required per simulated day (computational efficiency). This comparison was completed on a SGI Origin 2000 (O2K) using a single processing element. The O2K is a shared/distributed memory machine with relatively slow access to main memory. The two resolutions for each model are the same as shown in the results above:

2562 and 10 242 for the GDC versus T30 and T63 for the spectral model. As seen in Table 1, the geodesic grid dynamics with 10 242 grid cells is 20% slower than the spectral model truncated at T63 (the T63 transform grid has 18 432 grid points). With 2562 grid cells the geodesic model is approximately 2.5 times slower than the spectral model truncated at T30. In terms of MFlop rate, the geodesic grid model at 10 242 is approximately 50% faster than the spectral model truncated at T63. The geodesic grid model is, nevertheless, slower than the spectral model because it requires more to complete a simulated day.

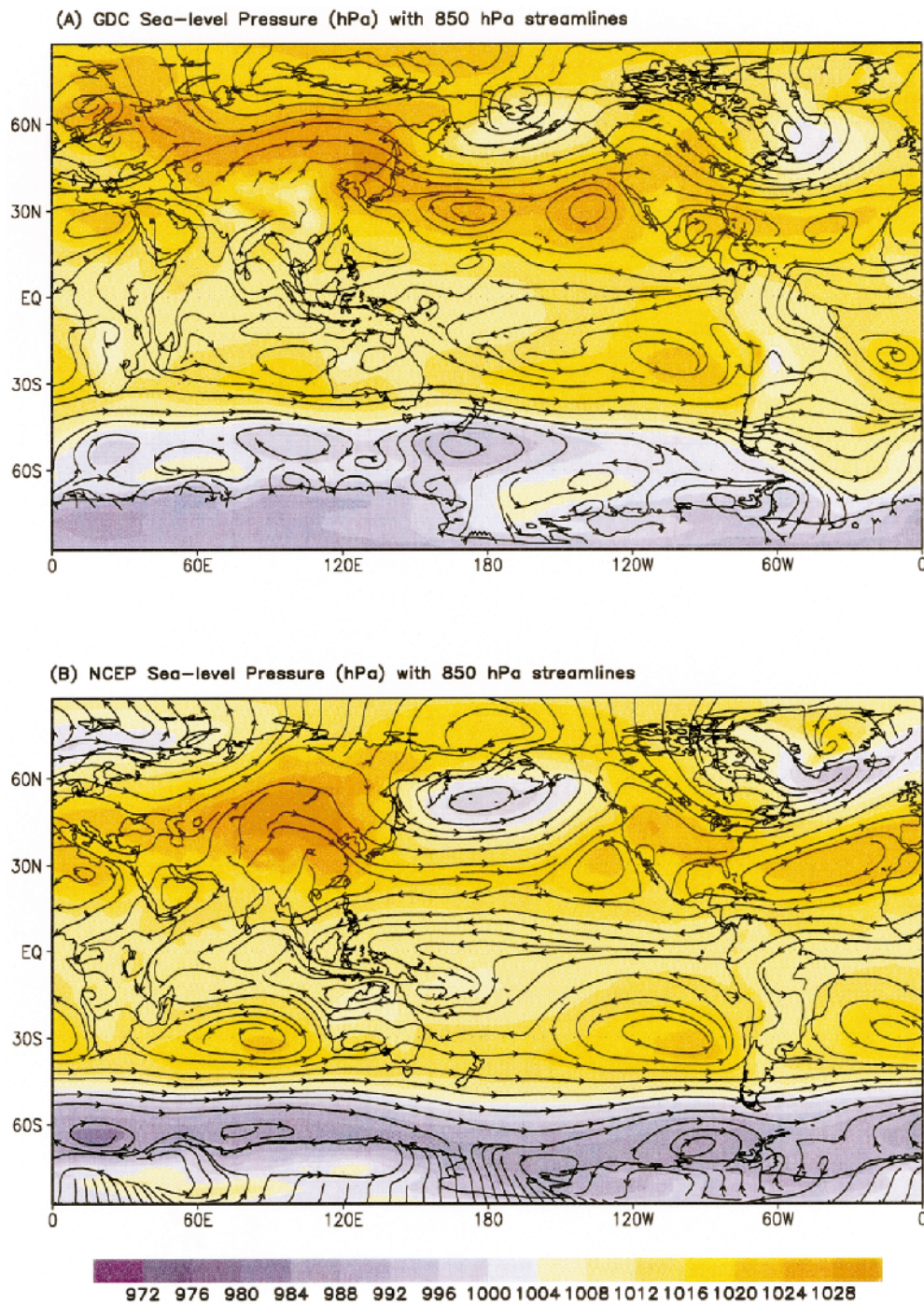One reason that more floating point operations are

FIG. 10. Sea level pressure and 850-hPa velocity streamlines for DJF in the (a) GDC–AGCM and (b) NCEP–NCAR reanalysis datasets.

required is that the line integrals used to evaluate the Jacobian, flux-divergence, and Laplacian operators [see, e.g., (7) and (8)] are computed at all cell walls for every grid cell. Referring to Fig. 2, this means that the line integral from b_2 to b_1 is evaluated twice, once for cell P_0 and once for cell P_1. The numerical operators were

originally developed for vector architecture and it was determined that this method was more efficient than storing and reusing results. We are exploring the alternative method of saving the computations at the cell walls and using them at the adjoining cells instead of recomputing them. While this alternative will signifi-
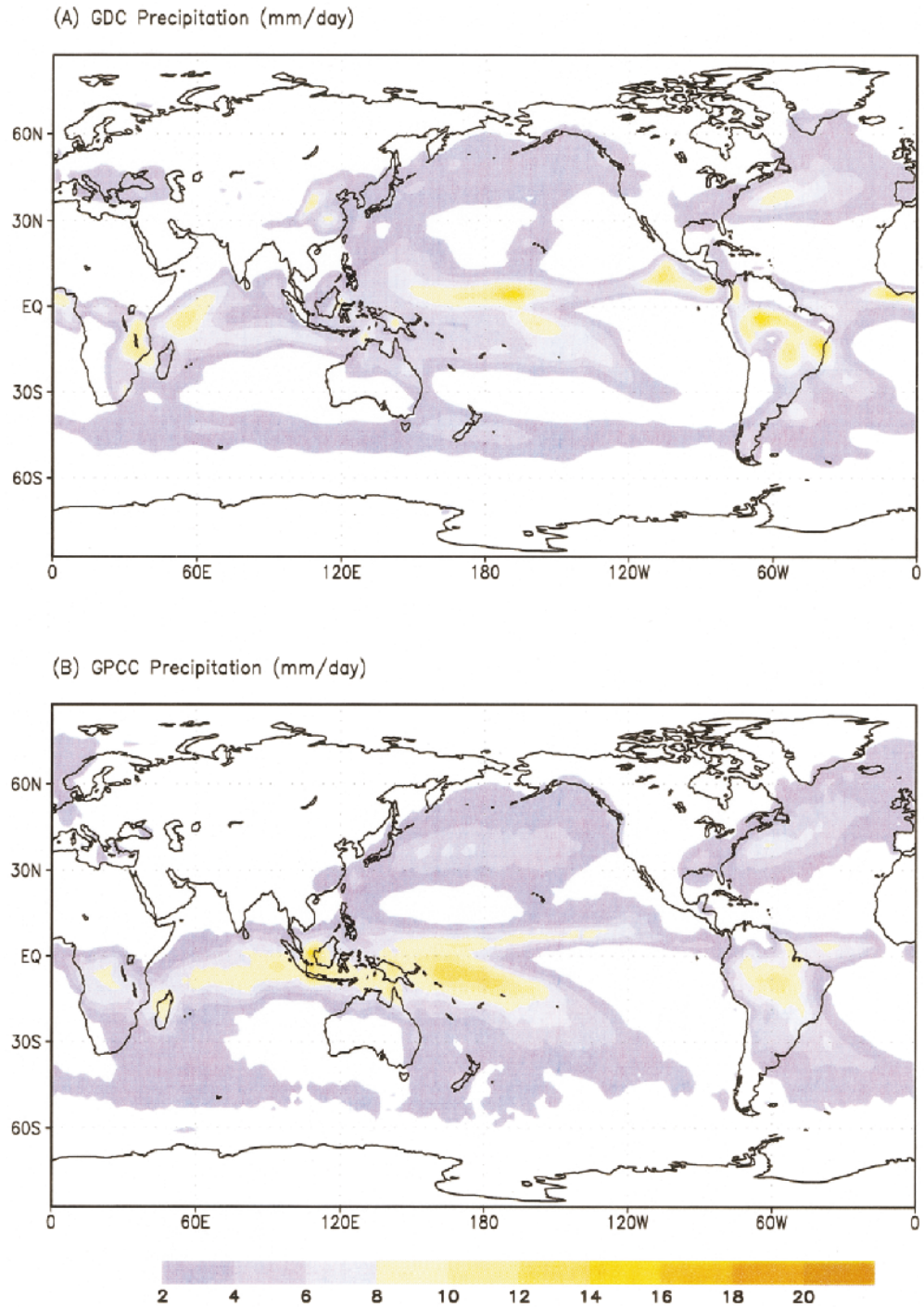
(A) GDC Precipitation (mm/day)

(B) GPCC Precipitation (mm/day)

FIG. 11. Total precipitation (mm day^{-1}) for DJF in the (a) GDC–AGCM and (b) GPCC precipitation datasets.

cantly reduce the number of floating point operations per time step, it may also reduce the MFlop rate due to increased communication.

Several caveats should be kept in mind when evaluating these results. First, the GDC is a relatively new model that, unlike the spectral model, has not had the benefit of years of optimization. Improvements in the semi-implicit time-stepping scheme and the reduction of redundant floating point operations could easily result in the geodesic grid model being as efficient as the spectral model, particularly at higher resolution. Second, these tests were conducted on a single processing element. The current trend in computer platforms is toward massively parallel systems where "turnaround time" is

TABLE 1. Comparison of computational efficiency between the geodesic grid model and a standard spectral model. All timings were completed on a single processing element on a SGI Origin 2000.

Model	Resolution	Time step (min)	Mflop rate	CPU time (s) per simulated day
Geodesic	10 242	20	106.0	490.3
Geodesic	2562	30	91.2	65.7
Spectral	T63	20	67.9	411.9
Spectral	T30	30	94.3	25.7

a strong function of how efficiently the problem can be spread across many processing elements. Models based on finite-difference schemes, such as the Los Alamos National Laboratory Parallel Ocean Program, have demonstrated computational efficiency on as many as 512 processing elements. In contrast, spectral models generally utilize significantly fewer elements. For example the spectral dynamical core used in NCAR's Community Climate Model version 3.2 is currently limited to 64 processing elements. Our intention is to implement a massively parallel version of the GDC in the future.

5. Conclusions

Overall, the results indicate that the development of the geodesic dynamical core has been very successful. The dynamical core incorporates many of the positive features of spectral models and finite-difference models into a single framework. As with many spectral models, this dynamical core solves the vorticity-divergence form of the primitive equations. This allows a straightforward implementation of semi-implicit time differencing. Furthermore, by choosing the vorticity-divergence form of the governing equations instead of the momentum equations, the model is written in terms of true scalars and is, therefore, invariant to coordinate system transformations. All of the scalars are defined on an unstaggered grid, the Z grid (Randall 1994), which has been shown to simulate geostrophic adjustment better than the more conventional B and C grids. The equations are discretized on a spherical geodesic grid that is nearly uniform over the entire globe. This grid eliminates the strong singularities that are present at the poles of regular latitude–longitude grids. Similar to other finite-difference models, the finite-difference stencils are local, which makes the model amenable to massively parallel systems.

In the idealized test case of Held and Suarez (1994), the overall level of meteorological variability simulated by this model using 2562 grid cells is comparable to that produced by a spectral dynamical core truncated at slightly less than T30 (Figs. 6 and 7). With 10 242 grid cells, the variability of the GDC is comparable to that of a spectral model truncated at slightly less than T63. Comparing resolutions of finite-difference models and spectral models can be misleading, but the transform grids of T30 and T63 are comparable in resolution to the geodesic model using 2562 and 10 242 grid cells, respectively.

Comparison of the zonal-mean statistics and zonal spectra of the variance fields indicates that the GDC and the spectral model are producing very similar general circulations. Spectral models are at their best within the framework of the dry primitive equations with no surface topography. In this case the numerical solutions are relatively smooth and the spectral expansion converges rapidly with increasing wavenumber. Making a comparison between the geodesic grid model and the spectral model within the framework of a full AGCM is relatively difficult given the strong influence that physical parameterizations have on the general circulation. While the Held–Suarez Test Case is a useful first step in evaluating dynamical cores it lacks orography and moisture, both of which have a strong impact on the climate. We are currently developing an intermediate test case that uses real topography and incorporates moisture in a simple way. In some sense, this would be a more useful test case since it includes some aspects of the asymmetries present in the real atmosphere and since moisture is an integral part of the general circulation and accurately accounting for its transport and phase changes is a critical task for dynamical cores.

Results from an AGCM that uses the GDC and the CSU physics package were presented to demonstrate the viability of this modeling approach. Both the dynamical core and the physical parameterizations were discretized on the same geodesic grid using 10 242 grids cells in the horizontal and 17 vertical layers. While further development is required, the resulting large-scale atmospheric structure (Fig. 9), surface winds (Fig. 10), and precipitation (Fig. 11) compare favorably to the observational data.

Comparing AGCM results with the GDC to previous results is a critical test since the results will contain the feedbacks between the atmospheric dynamics and the other climate subsystems. Standard AMIP2 simulations will be conducted as a part of the evaluation process. Coupled ocean–atmosphere simulations using the GDC are also planned.

Within the geodesic grid model, further improvement in the order of accuracy of the finite-difference operators is possible and we are currently working on this. We expect a substantial improvement in the simulations, particularly at lower resolution, when the accuracy of the finite-difference operators is increased.

Further optimization, in terms of computational efficiency, is also possible in the GDC (Table 1). While the model demonstrates a relatively high Mflop rate, the number of floating point operations required to complete a simulated day is higher than that required by the spectral model. Our intention is to work on improving the semi-implicit time-stepping scheme and reducing the number of redundant floating point operations in order to improve the overall computational efficiency.

While we are pleased with the results shown in this paper, we are pursuing the possibility of using vertical coordinate systems different from the conventional σ

coordinate. One version of the model that is currently being tested uses potential temperature as the vertical coordinate, following Hsu and Arakawa (1990). Another version incorporates the generalized vertical coordinate of Konor and Arakawa (1997).

APPENDIX A

Conserving Kinetic Energy under Vertical Advection

Beginning with Arakawa and Lamb [1977, Eq. (169)], the material derivative in the momentum equation can be expressed as

$$\left(\pi_k \frac{d\mathbf{V}_k}{dt}\right) = \pi_k\left(\frac{\partial}{\partial t} + \mathbf{V}_k \cdot \mathbf{\nabla}\right)\mathbf{V}_k$$

$$+ \frac{1}{\Delta\sigma_k}[(\pi\dot\sigma)_{k+1/2}(\hat{\mathbf{V}}_{k+1/2} - \mathbf{V}_k)$$

$$+ (\pi\dot\sigma)_{k-1/2}(\mathbf{V}_k - \hat{\mathbf{V}}_{k-1/2})], \quad \text{(A1)}$$

where k is the index representing layer center and $k \pm$ ½ represents layer edges. The index increases with increasing σ. Here, $\mathbf{V}$ is defined at layer centers, so $\hat{\mathbf{V}}$ represents an averaging of velocity from layer centers to layer edges. Arakawa and Lamb (1977) show that the choice of averaging that conserves kinetic energy under vertical advection is

$$\hat{\mathbf{V}}_{k+1/2} = \frac{1}{2}(\mathbf{V}_k + \mathbf{V}_{k+1}) \quad \text{and}$$

$$\hat{\mathbf{V}}_{k-1/2} = \frac{1}{2}(\mathbf{V}_k + \mathbf{V}_{k-1}). \quad \text{(A2)}$$

We obtain the discrete form of the vertical advection in the vorticity equation by 1) dividing (A1) by π_k, 2) substituting (6) into (A1), and 3) taking the curl of the resulting equation. This results in

$$\mathbf{\nabla} \times \left\{\frac{1}{\pi_k\Delta\sigma_k}\left[(\pi\dot\sigma)_{k+1/2}(\hat{\mathbf{V}}_{k+1/2} - \mathbf{V}_k) + (\pi\dot\sigma)_{k-1/2}(\mathbf{V}_k - \hat{\mathbf{V}}_{k-1/2})\right]\right\}$$

$$= \left\{F\left[\frac{(\pi\dot\sigma)_{k+1/2}}{\pi_k\Delta\sigma_k}, (\hat\psi_{k+1/2} - \psi_k)\right] + F\left[\frac{(\pi\dot\sigma)_{k-1/2}}{\pi_k\Delta\sigma_k}, (\psi_k - \hat\psi_{k-1/2})\right]\right\}$$

$$+ \left\{J\left[\frac{(\pi\dot\sigma)_{k+1/2}}{\pi_k\Delta\sigma_k}, (\hat\chi_{k+1/2} - \chi_k)\right] + J\left[\frac{(\pi\dot\sigma)_{k-1/2}}{\pi_k\Delta\sigma_k}, (\chi_k - \hat\chi_{k-1/2})\right]\right\}. \quad \text{(A3)}$$

The continuous counterparts of the discrete form of the vertical advection can be found in (7). The discrete form of the vertical advection in the divergence equation is found in similar way, except we take the divergence instead of the curl. This results in

$$\mathbf{\nabla} \cdot \left\{\frac{1}{\pi_k\Delta\sigma_k}\left[(\pi\dot\sigma)_{k+1/2}(\hat{\mathbf{V}}_{k+1/2} - \mathbf{V}_k) + (\pi\dot\sigma)_{k-1/2}(\mathbf{V}_k - \hat{\mathbf{V}}_{k-1/2})\right]\right\}$$

$$= \left\{F\left[\frac{(\pi\dot\sigma)_{k+1/2}}{\pi_k\Delta\sigma_k}, (\hat\chi_{k+1/2} - \chi_k)\right] + F\left[\frac{(\pi\dot\sigma)_{k-1/2}}{\pi_k\Delta\sigma_k}, (\chi_k - \hat\chi_{k-1/2})\right]\right\}$$

$$- \left\{J\left[\frac{(\pi\dot\sigma)_{k+1/2}}{\pi_k\Delta\sigma_k}, (\hat\psi_{k+1/2} - \psi_k)\right] + J\left[\frac{(\pi\dot\sigma)_{k-1/2}}{\pi_k\Delta\sigma_k}, (\psi_k - \hat\psi_{k-1/2})\right]\right\}, \quad \text{(A4)}$$

which can be compared to (8). The discrete forms shown on the right-hand sides of (A3) and (A4) conserve kinetic energy under vertical advection.

APPENDIX B

Semi-Implicit Time Differencing

Equations (B1)–(B6) are included in the semi-implicit time-stepping scheme. Those terms that are underlined are integrated implicitly, while all other terms are integrated explicitly:

$$\frac{\partial \delta}{\partial t} - J(\eta, \chi) - \nabla \cdot (\eta \nabla \psi) + \nabla \cdot \left(\dot{\sigma} \frac{\partial \chi}{\partial \sigma} \right) - J\left(\dot{\sigma}, \frac{\partial \psi}{\partial \sigma} \right)$$
$$+ \nabla^2 K + \underline{\nabla^2 \Phi} + \underline{\nabla \cdot (\sigma \alpha \nabla \pi)} = \nabla \cdot \mathbf{G}, \qquad (B1)$$

$$\frac{\partial p_S}{\partial t} = \int_{\sigma=0}^{\sigma=1} [J(\pi, \psi) - \underline{\nabla \cdot (\pi \nabla \chi)}] \, d\sigma', \quad (B2)$$

$$\frac{\partial}{\partial t}(\pi \theta) - J(\pi \theta, \psi) + \underline{\nabla \cdot (\pi \theta \nabla \chi)} + \underline{\frac{\partial}{\partial \sigma}(\dot{\sigma} \pi \theta)} = \frac{\pi \theta Q}{c_p T},$$
$$(B3)$$

$$\partial \Phi = -\alpha \pi \partial \sigma, \qquad (B4)$$

$$(\pi \dot{\sigma})|_{\sigma = \sigma'} = -\sigma' \frac{\partial \pi}{\partial t} + \int_{\sigma=0}^{\sigma=\sigma'} [J(\pi, \psi) - \nabla \cdot (\pi \nabla \chi)] \, \partial \sigma,$$
$$(B5)$$

and

$$\chi = \nabla^{-2} \delta. \qquad (B6)$$

The semi-implicit scheme is iterative in nature. First guesses are made for the values of the prognostic variables, and all terms that are evaluated explicitly are computed and saved. The iteration cycle consists of 1) evaluating the implicit terms, 2) updating the prognostic variables, 3) evaluating the diagnostic variables, 4) evaluating convergence, and 5) returning to step 1 if convergence is not met. One iteration through the semi-implicit scheme involves evaluating the implicit terms on the right-hand side of (B1)–(B3), updating the prognostics in (B1)–(B3), and then updating the diagnostic equations (B4)–(B6). Multigrid methods are only used to solve (B6). Following an iteration cycle, convergence is assessed by differencing the values of the prognostic variables at the current iteration and their values at the last iteration. If this difference is smaller than the specified criteria for each prognostic variable at every grid point, then the solution is deemed to have converged and the integration continues to the next time level.

The criteria we use to assess convergence for divergence, temperature, and surface pressure are 1.0×10^{-6} s^{-1}, 0.25 K, and 15 Pa, respectively. All three of these conditions must be met before the solution is deemed convergent. Typically, the scheme requires about four iterations to converge. In order to ensure stability of the iteration process, we have also found it advantageous to underrelax divergence during the iteration process. The model typically uses an underrelaxation coefficient of approximately 0.5.

REFERENCES

Arakawa, A., and V. R. Lamb, 1977: Computational design of the basic dynamical process of the UCLA general circulation model. *Methods in Computational Physics,* Vol. 17, Academic Press, 173–265.

Baumgardner, J. R., and P. O. Frederickson, 1985: Icosahedral discretization of the two-sphere. *SIAM J. Numer. Anal.,* **22,** 1107–1115.

Bourke, W., 1974: A multilevel spectral model. I. Formulation and hemispheric integrations. *Mon. Wea. Rev.,* **102,** 687–701.

Charney, J. G., R. Fjortoft, and J. von Neumann, 1950: Numerical integration of the barotropic vorticity equation. *Tellus,* **2,** 237–254.

Cooley, J. W., and J. W. Tukey, 1965: An algorithm for the machine computation of complex Fourier series. *Math. Comput.,* **19,** 297–301.

Ding, P., and D. A. Randall, 1998: A cumulus parameterization with multiple could base levels. *J. Geophys. Res.,* **103,** 11 341–11 354.

Durran, D. R., 1991: The third-order Adams–Bashforth method: An attractive alternative to leapfrog time differencing. *Mon. Wea. Rev.,* **119,** 702–720.

Eliasen, E., B. Machenhauer, and E. Rasmussen, 1970: On a numerical method for integration of the hydrodynamical equations with a spectral representation of the horizontal fields. Institut for Teoretisk Meteorologi Rep. 2, University of Copenhagen, 37 pp.

Fowler, L. D., D. A. Randall, and S. A. Rutledge, 1996: Liquid and ice cloud microphysics in the CSU general circulation model. Part I: Model description and simulated microphysical processes. *J. Climate,* **9,** 489–529.

Greenberg, M. D., 1978: *Foundations of Applied Mathematics.* Prentice-Hall, 636 pp.

Haltiner, G. J., and R. T. Williams, 1980: *Numerical Prediction and Dynamic Meteorology.* John Wiley and Sons, 477 pp.

Harshvardhan, D. A. Randall, T. G. Corsetti, and D. A. Dazlich, 1989: Earth radiation budget and cloudiness simulations with a general circulation model. *J. Atmos. Sci.,* **46,** 1922–1942.

Heikes, R., 1993: The shallow water equations on a spherical geodesic grid. Department of Atmospheric Science Tech. Rep. 524, Colorado State University, Fort Collins, CO, 175 pp.

——, and D. A. Randall, 1995a: Numerical integration of the shallow-water equations on a twisted icosahedral grid. Part I: Basic design and results of tests. *Mon. Wea. Rev.,* **123,** 1862–1880.

——, and ——, 1995b: Numerical integration of the shallow-water equations on a twisted icosahedral grid. Part II: A detailed description of the grid and an analysis of numerical accuracy. *Mon. Wea. Rev.,* **123,** 1881–1887.

Held, I. M., and M. J. Suarez, 1994: A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bull. Amer. Meteor. Soc.,* **73,** 1825–1830.

Hoskins, B. J., and A. J. Simmons, 1975: A multi-layer spectral model and the semi-implicit method. *Quart. J. Roy. Meteor. Soc.,* **101,** 636–655.

Hsu, Y.-J. G., and A. Arakawa, 1990: Numerical modeling of the atmosphere with an isentropic vertical coordinate. *Mon. Wea. Rev.,* **118,** 1933–1959.

Jaluria, Y., and K. E. Torrance, 1986: *Computational Heat Transfer.* Hemisphere Publishing, 366 pp.

Jarrud, M., and A. J. Simmon, 1983: The spectral technique. *Proc. Seminar on Numerical Methods for Weather Prediction,* Reading, United Kingdom, European Centre for Medium-Range Weather Forecasts, 1–60.

Jones, P. W., 1999: First- and second-order conservative remapping

schemes for grids in spherical coordinates. *Mon. Wea. Rev.,* **127,** 2204–2210.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Konor, C. S., and A. Arakawa, 1997: Design of an atmospheric model based on a generalized vertical coordinate. *Mon. Wea. Rev.,* **125,** 1649–1673.

Masuda, Y., and H. Ohnishi, 1986: An integration scheme of the primitive equations model with an icosahedral-hexagonal grid system and its application to the shallow water equations. *Short- and Medium-Range Numerical Weather Prediction,* T. Matsuno, Ed., Japan Meteorological Society, 317–326.

Orlanski, I., 1985: Advances in the theory of atmospheric fronts. *Advances in Geophysics,* Vol. 28, Academic Press, 223–252.

Orszag, S. A., 1970: Transform method for the calculation of vector-coupled sums: Application to the spectral form of the vorticity equation. *J. Atmos. Sci.,* **27,** 890–895.

Palmer, T. N., G. J. Shutts, and R. Swinbank, 1986: Alleviation of a systematic westerly bias in general circulation and numerical weather prediction models through an orographic gravity wave drag parameterization. *Quart. J. Roy. Meteor. Soc.,* **112,** 1001–1039.

Panton, R. L., 1984: *Incompressible Flow.* John Wiley and Sons, 780 pp.

Phillips, N. A., 1957: A coordinate system having some special advantages for numerical forecasting. *J. Meteor.,* **14,** 184–185.

Phillips, T. J., 1996: Documentation of the AMIP models on the World Wide Web. *Bull. Amer. Meteor. Soc.,* **77,** 1191–1196.

Randall, D. A., 1994: Geostrophic adjustment and the finite-difference shallow water equations. *Mon. Wea. Rev.,* **122,** 1371–1377.

Richardson, L. F., 1922: *Weather Prediction by Numerical Process.* Cambridge University Press, 236 pp.

Ringler, T. D., R. P. Heikes, and D. A. Randall, 1998: Solving the primitive equations on a spherical geodesic grid: A technical report to a new class of dynamical cores. Department of At-

mospheric Science Tech. Rep. 665, Colorado State University, Fort Collins, CO, 62 pp.

Royer, J. F., 1986: Correction of negative mixing ratios in spectral models by global horizontal borrowing. *Mon. Wea. Rev.,* **114,** 1406–1410.

Sadourny, R., A. Arakawa, and Y. Mintz, 1968: Integration of the nondivergent barotropic vorticity equation with an icosahedal-hexagonal grid for the sphere. *Mon. Wea. Rev.,* **96,** 351–356.

Sellers, P. J., and Coauthors, 1996: A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: Model formulation. *J. Climate,* **9,** 676–705.

Stuhne, G. R., and W. R. Peltier, 1996: Vortex erosion and amalgamation in a new model of large scale flow on the sphere. *J. Comput. Phys.,* **128,** 58–81.

——, and ——, 1999: New icosahedral grid-point discretizations of the shallow water equations on the sphere. *J. Comput. Phys.,* **144,** 23–58.

Suarez, M. J., A. Arakawa, and R. A. Randall, 1983: Parameterization of the planetary boundary layer in the UCLA general circulation model: Formulation and results. *Mon. Wea. Rev.,* **111,** 2224–2243.

Swarztrauber, P. N., D. L. Williamson, and J. B. Drake, 1997: The Cartesian method for solving partial differential equations in spherical geometry. *Dyn. Atmos. Oceans,* **27,** 679–706.

Tennekes, H., and J. L. Lumley, 1972: *A First Course in Turbulence.* The MIT Press, 300 pp.

Thuburn, J., 1997: A PV-based shallow-water model on a hexagonal–icosahedral grid. *Mon. Wea. Rev.,* **125,** 2328–2347.

Williamson, D. L., 1968: Integration of the barotropic vorticity equation on a spherical geodesic grid. *Tellus,* **20,** 642–653.

——, 1976: Linear stability of finite-difference approximations on a uniform latitude–longitude grid with Fourier filtering. *Mon. Wea. Rev.,* **104,** 31–41.

Xie, P. P., and P. A. Arkin, 1996: Analyses of global monthly precipitation using gauge observations, satellite estimates, and numerical model predictions. *J. Climate,* **9,** 840–858.

Zalesak, S., 1979: Fully multidimensional flux-corrected transport algorithms for fluids. *J. Comput. Phys.,* **31,** 335–362.